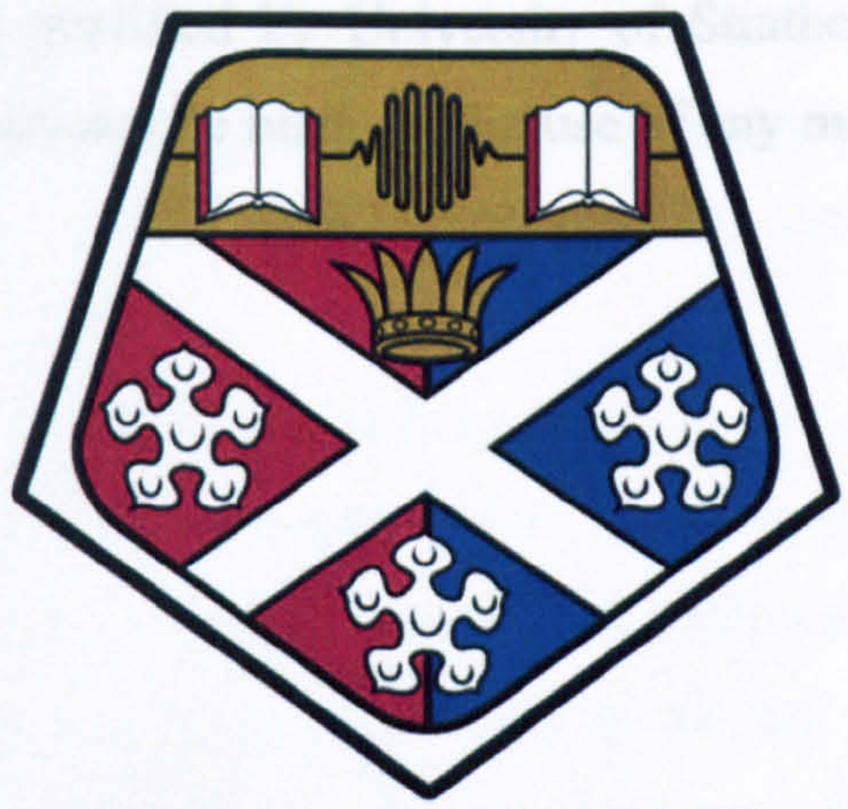


University of Strathclyde

Strathclyde Institute of Pharmacy and Bioscience

A Study of the Ecology & Diversity of Bacteriophages.

by

Emma Lisa Bell

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy

2008

Declaration

Emma Lisa Bell

Acknowledgements

I would like to express my gratitude to Paul Herron, my supervisor, for his continuous guidance, support and patience. Supporting advice from Mike Mattey and Iain Hunter was invaluable, therefore I would also like to extend my appreciation to them. Further thanks to Blaze Technologies for their funding, and to Roger Hendrix for his assistance and collaborative work.

A special thank you to Vera Gill who helped ease a new, apprehensive girl into the 5.21 lab, and who later provided laughter and "great banter". Thanks to Janice Spencer who provided me with, not only strains and technical advice, but also her memories of our shared wee town. My gratitude to all my fellow lab mates and departmental friends from over the years who provided me with a endless coffee, a caffeine addiction and good chat: Emma, Russell, Laura, Karen, Angzass, Urska, Debs, Janet, Badri, Meijing, Sarah, Gordon, Rob, Naomi, Clett and Alex. Lunchtimes shall never be the same without the daily random facts and discussions or "factoids" as we liked to call them. A very special thank you to Vinod, Chantevy, Dzeti and Liz for their friendship and support.

Finally I would like to thank my family and friends for having the patience to listen to my "cloning's not working" moans over the years - they provided invaluable support. Particular thanks must go to Misiek, my canine pal who forced me to take breathers while writing by insisting on walks and to my brother, Marek, who remained confident in my ability to succeed throughout the years. In addition, I would like to dedicate this work to and thank my Babcia, who passed away after my viva, she consistently held faith that I would do well, and was overjoyed when I passed although she "always knew I would".

Table of Contents

List of Figures

List of Tables

Abbreviations

Standard Units

°C - Degress Celsius

g – gram

k – kilo

L – litre

M – Molar

Ω - ohms

Psi – pounds per square inch

V – volt

W – Watt

DNA Bases

A - Adenine

C – Cytosine

G – Guanine

T – Thymine

U – Uracil

Textual abbreviations

ap - apramycin

bla - amplicillin

BLAST - Basic local alignment tool

BLAST N - Basic local alignment tool nuclotide

BLAST P - Basic local alignment tool protein

bp - base pairs

DNA - Deoxyribonucleic acid

DNase - Deoxyribonulease

dATP - deoxyadenisine 5'-triphosphate

dCTP - deoxycytidine 5'-triphosphate

dGTP - deoxyguanidine 5'-triphosphate

DMF – Dimethyformamide

DOC – Dissolved organic matter

dNTP - dinucleoside 5'-triphosphate

dsDNA - double stranded deoxyribonucleic acid

dTTP - deoxythymidine 5'-triphosphate

EDTA – Ethylendiaminotetraacetic acid

EM – Electron microscope

EOP – efficiency of plating

G/C – guanine/cytosine content

ICTV – International committee in taxonomy of viruses

ISP – International Streptomyces Project

Kb - kilobase

L – litre

LB - Lennox Broth

LFRFA - Low frequency restriction fragment analysis

Mc^s – methicillin susceptible

mL – millilitre

MLSA – Multi locus sequence analysis

mM – millimolar

MOI – multiplicity of infection

MRSA – Methicillin resistant *Staphylococcus aureus*

MSSA - Methicillin resistant *Staphylococcus aureus*

MS - Mannitol Soya

NB – Nutrient Broth

neo – kanomycin

NCBI - National Centre for Biotechnology Information

nt - nucleotide

OD - optical density

ORF - Open reading frame

p - promoter

PCR - Polymerase chain reaction

PEG – Polyethylenglycol

Pgl – Phage growth limitation

RASS – Reduced arganine starch salts

RBS – Ribosomal binding site

rpm - revolutions per minute

RNase - Ribonuclease

SDS - sodium deodecyl sulphate

SDW – sterile distilled water

SEM – Scanning electron microscope

SNA - Soft nutrient agar

sp - species

ssDNA - single stranded deoxyribonucleic acid

SSU – Small sub-unit

TEM – transmission electron microscope

tet – tetracycline

$\mathbf{T_m}$ – melting temperature

Tris - Trishydroxymethylaminomethane

v/v - volume to volume ratio

w/w - weight to weight ratio

w/v - weight to volume ratio

YEME - Yeast Extract-Malt Extract Medium

x-gal - 5-bromo-4chloro-3indoyl-β-D-galactoside

Abstract

Bacteriophages are ubiquitous and abundant in nature, yet our knowledge of their ecology, global impact on the ecosystem, their relationships with one another, and their hosts remains relatively primitive. Their consequence on the resident microbial population greatly depends upon their lifecycle with hosts. Their acquisition or removal of genes between the phages themselves eventually determines their global impact. Therefore by examining their genomes will contribute to our understanding of their importance and potential impact on ecology as a whole. Previous studies have shown that bacteriophages do undergo extensive genetic exchange with one another (Hendrix *et al.*, PNAS, **96**, 2192-97), however, these have examined bacteriophages from varied locations. Unlike previous studies, this work examines the ecology of phages and their hosts in a single soil source. By concentrating on single environment the dynamics of phages and hosts can be determined by examining their relationships and interactions with one another.

Bacteriophages and hosts were isolated from the same soil and placed into distinct groups according to characterisation experiments. The monitoring the phage-host system by soil microcosm experiments allowed the impact of phages upon their host populations to be followed over time. In addition, to our knowledge there are no studies examining the impact of soil phages upon their hosts whilst also examining their genomes. However there is one study examining the genetic relationships of bacteriophages from a single soil sample which found them to be highly related. As a result genomic sequence and hybridisation experiments were applied in order to quantify the degree of relatedness between phages. Genomic sequencing contributed to our understanding of the mechanisms involved in viral evolution within a single soil sample and supported the findings of previous studies.

1. Introduction

1.1. Soil composition

Soil presents a complex three-dimensional, heterogeneous environment. Its composition varies depending on location, and primarily consists of mineral particles, organic matter, water and air. The organic matter accounts for only 5%, yet despite this small amount, it is vital within the ecosystem as a whole (Pidwirny, 2006)

Figure 1). The microbial diversity within soil is dependant upon several factors. For example, a soil with pH range of 6.3 to 6.8 results in the highest abundance of bacterial species. This is indeed true for the actinomycetes, whose optimum soil pH range is 6.5 to 8.0 (Burroughs *et al.*, 2000). In addition to pH, other factors such as moisture content, temperature and nutrient availability all contribute to the levels of diversity within soil ecosystem (Egamberdlyeva and Hoflich, 2003; Williams and Khan, 1974; Williams *et al.*, 1987). Early studies on the composition of soil readily identified the importance of the microbial life within soil and the impact it has within the nutrient turnover of the ecosystem as a whole (Winogradsky, 1953).

1.2. Soil ecology and microbial diversity

Abundantly found in soil there are an estimated 10^8 prokaryotes per gram of soil, (Chibani-Chennoufi *et al.*, 2004). There are thought to be more micro-organisms in soil than in sea water (Chibani-Chennoufi *et al.*, 2004; Wommack *et al.*, 1999). Inhabitants of soil are highly adapted to their environment or are diverse organisms capable of surviving in a number of different environments. Several studies have demonstrated that the soil structure and practices, such as the repeated use of fertilizers, have dramatic effects upon microbial diversity and community structure. Indeed, it has been shown that seasonal changes do result in variations within resident microbial communities (Anderson, 2003; Torsvik and Overas, 2002).

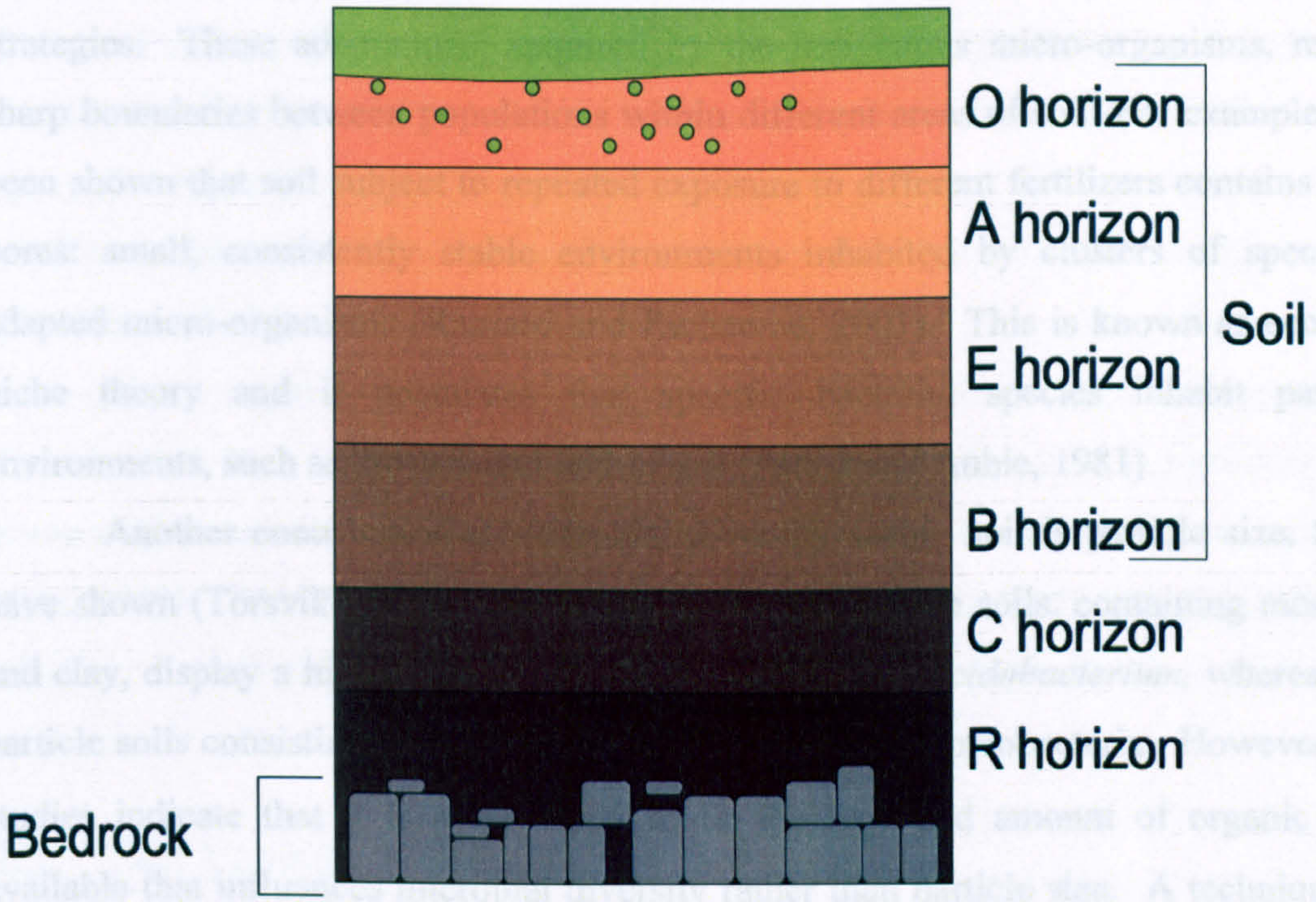


Figure 1 The horizons found within soil. Layers within soil are termed "horizons". o-horizon: organic matter, a-horizon topsoil, e-horizon: sand and silt, b-horizon: sub-soil mineral deposits, c-horizon: low in organic material, & r-horizon: solid bedrock. The top green layer represents living plant material. Deeper orange-brown layers signify the horizons found within soil. The diversity and abundance of resident bacterial communities is depth dependant, typically with greatest populations within the O and A horizons (Flerer *et al.*, 2003) (diagram adapted from Pidwirny, 2006).

Soil presents its inhabitants with an unstable, changeable environment consisting of several microhabitats, therefore forcing the organisms to have evolved coping strategies. These adaptations, acquired by the indigenous micro-organisms, result in sharp boundaries between populations within different areas of soil. For example, it has been shown that soil subject to repeated exposure to different fertilizers contains micro-pores: small, consistently stable environments inhabited by clusters of specifically adapted micro-organisms (Ranjard and Richaume, 2001). This is known as ecological niche theory and it postulates that specific bacterial species inhabit particular environments, such as those found within soil (Patten and Auble, 1981).

Another contribution to microbial diversity within soil is particle size. Studies have shown (Torsvik and Overas, 2002) that small particle soils, containing mostly silt and clay, display a higher abundance of *Holophaga* and *Acidobacterium,* whereas large particle soils consisting of sand are more abundant in α-proteobacteria. However, other studies indicate that it is more likely to be the type and amount of organic matter available that influences microbial diversity rather than particle size. A technique used to monitor nutrient supply within soil is the examination of the resident microbial populations. For example, oligotrophs are organisms found in environments with low levels of nutrients whereas copiotrophs are found in nutrient rich environments (Koch, 2001). Indeed, it is the C:N ratio which has major effects on soil microbiology as well as the NPK content (Abaye *et al.*, 2004; Sun *et al.*, 2004). However, despite the impact that these elements can have upon soil it is also evident that pH plays an important role in the composition of the resident bacterial community (Fierer and Jackson, 2006). In addition, R/k theory may be used as an indication of the soil environment as this theory states that resident organisms possess specific traits depending on the surrounding environment. R strategists multiply with high frequency and reside in unstable environments, as a result are these are typically α-proteobacteria and ϒ-proteobacteria (Andrew and Harris, 1986; Smit *et al.*, 2001). Conversely, k-strategists are found in constantly stable environments and replicate slowly. An example of k-strategists are the slow growing acidobacterium found in nutrient-poor soils (MacArthur and Wilson, 1967; McCaig *et al.*, 2001). Another study supporting this finding has shown that k-

strategists are highly dependant upon soil characteristics such as water content as well as physical and chemical characteristics (Pacsuta *et al.*, 2005). These strategies have been illustrated throughout many environments. For example, it has been found that within a lake the resident bacterial populations survive by these strategies. It was found that one strain flourished under particular conditions and perished in the absence of these conditions. In contrast, another resident strain persistently remained in smaller numbers (Weinbauer, 2004).

Finally, spatial isolation of microbial communities is indicative of the microbial diversity. A high number of isolated bacterial communities in soil represent a high amount of microbial diversity and the converse is true for low microbial diversity (Smit *et al.*, 2001; Torsvik and Overas, 2002). The diverse resident soil populations are adapted to their harsh environment and as a result have evolved several survival mechanisms such as the formation of spores or dormant cocci.

1.3. Soil Ecology of phages

The richness of the bacterial diversity and adaptation within soil is mirrored by the resident bacteriophages, which are more abundant than their hosts in soil. Indeed, it is probable that they too are subject to influences such as k-selection and r-selection (see section 1.2) (Weinbauer *et al.*, 2006). The study of phages, their hosts and the predator-prey relationships that result, has provided great insights into the dynamics of these populations within their natural environments. Such studies have greatly contributed to our understanding of the microbial world; however, there remains a distinct lack of studies examining, with any detail, the dynamics within a single ecosystem (Ashelford *et al.*, 2000). There is increasingly more interest in the natural population ecology of phages due to the realisation that phages may have a potentially significant global impact (Wiebe and Liston, 1968). Despite the large numbers of phages within soil, the majority of studies focus on the marine environment (Weinbauer, 2004) due to the ease at which they can be isolated. These studies have clearly shown that a carefully balanced equilibrium between phage and its host is maintained.

The early studies on bacterial populations in soil were carried out using the transmission electron microscope. This method of direct observation is indiscriminate in that it will account for phages of all the bacterial species within the soil. In addition, it is probable that phage heads lacking in DNA or "non-viable ghost particles" are also counted. In comparison to viable counts, TEM analysis gives approximately 40 fold less the number of phages than those counted on plates (Ashelford *et al.*, 2003). Despite this, the method is useful as it provides an overall picture of the indigenous phage populations within soil. However, it does not provide detailed descriptions of fluctuations in populations or the effect of these viruses on their hosts.

Plaque assays remain the most commonly used method of assaying phage due to their ease and efficiency of use, and are therefore used in order to determine populations in the majority of phage studies. Ashelford *et al* (2003) have demonstrated, with the first ever *in-situ* study, that within the terrestrial soil environment, phages compete with one another for the same host. It was found that if the indigenous bacterial population increased, the corresponding phage population also increased significantly. Conversely, another phage infectious to the same bacteria did not benefit from an increase in host population. Indeed, the inoculated plots appeared to have lost the presence of that phage altogether. It was therefore concluded that competition between the two phages resulted in the demise of one phage and, successful establishment of the other (Ashelford *et al.*, 2003). However, due to the complex nature of soil it is possible that some of the contending phages did survive albeit with a low population.

More recently, flow cytometry has been applied in assessing the soil ecology of phages. This allows rapid identification and enumeration of of a wide variety of phages with different morphology, genome type and size as it utilises specific, sensitive nucleic acid stains(Brussaard, 2009)

Mathematical models displaying population fluctuations may also examine the predator-prey relationships. This has been of interest in streptomycetes, as due to their complex lifecycles, they exhibit different susceptibility to phage infection at different growth stages. More specifically, phage interaction with its host is of particular interest. Liquid culture is subject to a great amount of mixing and therefore the virus has more

opportunity to make contact with the host than is likely in soil. Also, with agar, as there is no movement between phage and host, "meetings" remain confined within the colony boundary. Soil, as already stated, is a far more complex environment and therefore possesses additional factors not accounted for when utilising agar or liquid cultures (Burroughs *et al.*, 2000). Burroughs *et al* (2003) demonstrated that, despite an observed increase in phage population after spore germination, there was no apparent impact upon the bacterial population. It has been suggested that soil composition produces underlying mechanisms affecting phage-host susceptibility and accounting for the low susceptibility of hosts in soil (Burroughs *et al.*, 2000). Given that most bacterial species are not driven to extinction, it is self-evident there must be constant equilibrium between phage and host.

1.4. A general introduction to Actinobacteria

Actinomycetes are a large and diverse group of micro-organisms. Within this group there exist the disease causing *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae,* as well as, the apathogenic streptomycetes (Camus *et al.*, 2002; Cerdeno-Tarraga *et al.*, 2003). Typically considered soil dwelling organisms they also inhabit other environments such as skin and water.

The soil actinomycetes are key members of the microscopic community due to their role in degradation and replacement of organic matter within the nutrient cycle. Their ability to degrade recalcitrant compounds such as chitin and lignin are vital for humus formation (Nishimura *et al.*, 2006). Characteristically, these organisms are Gram positive with a high G/C content and have a distinctive morphology suited to their soil habitat. Indeed, early classifications of several actinomycete species considered them to be fungi (Waksman and Henrici, 1943). Their importance within the rhizosphere was first recognised due to improvements in isolation techniques in the 1950s, for example the fixing of atmospheric nitrogen by *Frankia* species. In contrast, they some act as plant pathogens, for example, *S.scabies* which cause potato scab.

1.5. Streptomycetes

Of all the actinomycetes, *Streptomyces* has been one of the most studied genera due to the ability of its members to produce antibiotics. This has resulted in the genomes of *S.coelicolor* A3(2), *S.avermitilis*, *S.griseus* and *S.scabies* being sequenced (Bentley *et al.*, 2002; Ikeda H, 2003; Ohnishi *et al.*, 2008; Omura *et al.*, 2001) (http://www.sanger.ac.uk). *Streptomyces* are abundant in all soils due to their many strategies for survival such as spore formation and antibiotic production. Indeed, they have even been found in radiation-polluted soil, which illustrates their capacity of survival in extreme environments (Mao *et al.*, 2007).

Despite the clear ability to produce many bioactive compounds such as antibiotics, this capacity has proved difficult to demonstrate in natural environments such as soil (Williams and Khan, 1974). Williams and Vickers (Williams and Vickers, 1986) proposed possible reasons why detection in soil is difficult: antibiotics are not produced in soil, they are rapidly inactivated, they are absorbed by soil colloids, soil nutrients are insufficient for growth, and detection methods are altogether insufficient. However, recently some studies have examined this (Anukool *et al.*, 2004; Turpin *et al.*, 1992) and have proved that the presence of *Streptomyces* in soil does indeed have an impact upon resident organisms, even within its own genus. For example, *Salmonella* populations are reduced in the presence of *S.bikiniensis* in soil (Turpin *et al.*, 1992). Clearly, on a large scale, antibiotic production has an effect on resident soil micriobiology.

In addition, *Streptomyces* survive in soil by utilizing carbon sources that many other organisms cannot utilize, therefore, greatly contributing to degradation of organic material, such as lignin and as a result, have an impact upon biogeochmical cycles (Braissant *et al.*, 2002). An example might be the degradation of the toxic Ca-oxalate by *Streptomyces* sp. which is partially responsible for micro-site pH regulation, energy production and detoxification of the soil (Sahin, 2004).

S.coelicolor A3(2) was sequenced due to its role as a paradigm for the streptomycetes and it is mainly in this organism that their lifecycle was deduced. One of the most striking features of these highly adaptive, sporulating bacteria is their unique

and distinctive three-dimensional morphology, which is a result of their complex lifecycles (Figure 2) (Erikson, 1947; Hopwood, 1960). The lifecycle of *Streptomyces* begins as a unigenomic spore, resistant to soil desiccation (Radajewski and Duxbury, 2001) extremes of temperature, pH as well as low nutrient availability. Indeed it has been found that the majority of *Streptomyces* in soil are in this form (Mayfield *et al.*, 1972). In addition, in this form it is dispersed by wind, rain and arthropods (Lloyd, 1969; Ruddick and Williams, 1972). The spores of streptomycetes exist in this state for extended periods until exogenous nutrients are available which act as triggers, inducing the formation of a germ tube from which the tip extends, forming a mycelium.

In 1960 Waksman confirmed that there are two distinct forms of *Streptomyces* mycelium: substrate (vegetative) and aerial (reproductive) mycelium, both grown on solid agar. Substrate mycelium grows deep into the medium whereas aerial hyphae, first thought to be branches of substrate hyphae (Hopwood and Glauret, 1961), grow upwards, away from the media due to the activation of *bld* genes. Upon activation of *bld* genes, a signalling cascade occurs, and coiled, highly hydrophobic hyphae grow outwards from the radial growth (Nodwell *et al.*, 1999). It has also been found that the *bld* genes are additionally responsible for other characteristics such as the production of antibiotics (Plaskitt and Chater, 1995). In addition, *whi* genes are responsible for the formation of mature spores. These genes result in a formation of spiral syncytium in which multigenomic compartments reside, with the eventual formation of individual septa, and lastly, a mature spore (Ryding *et al.*, 1999). This complex lifecycle contributes to the renowned difficulty in classifying these organisms.

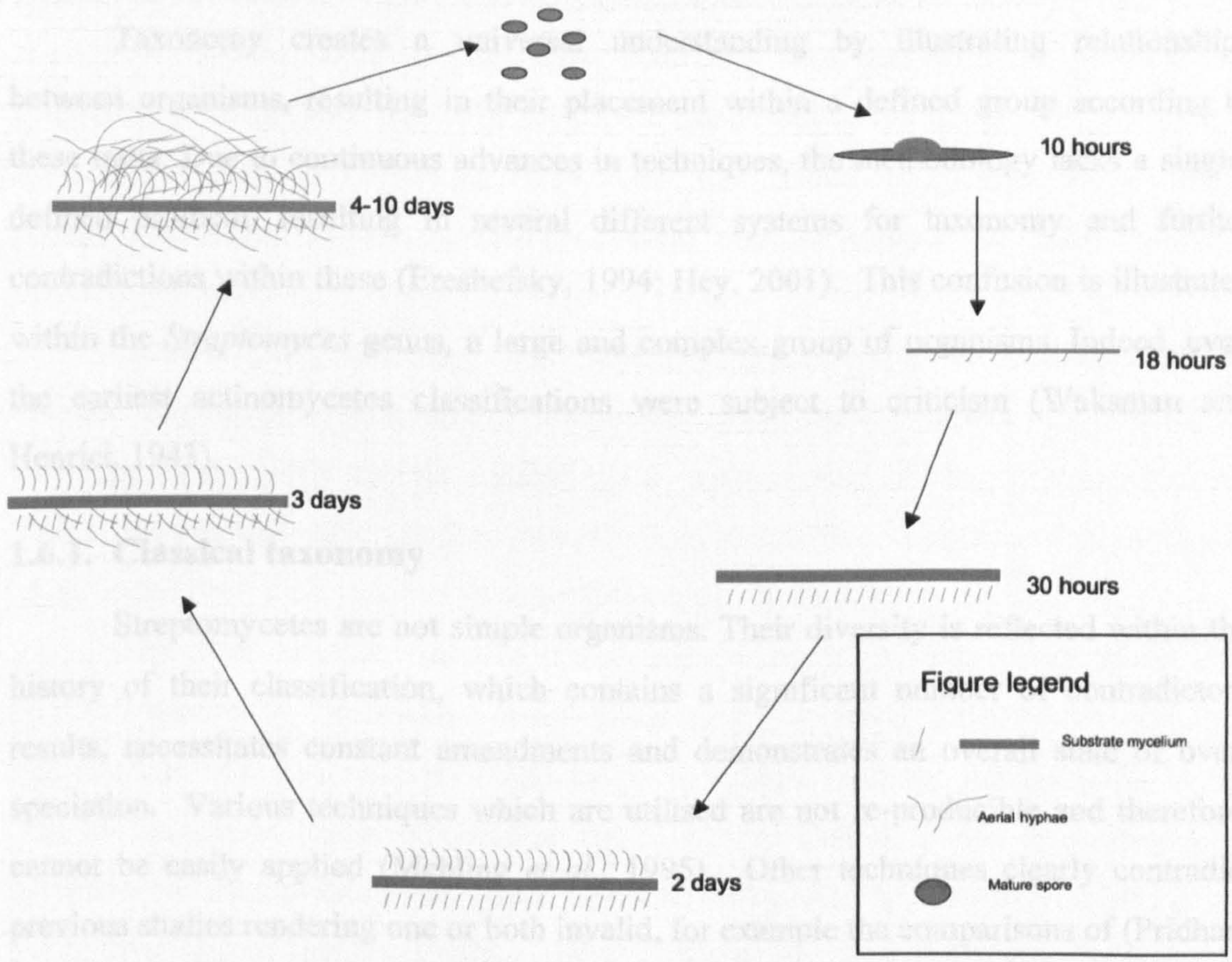


Figure 2 The *Streptomyces* lifecycle. At 10 hours under favourable conditions germ tubes emerge from a spore by tip extension and branch formation, creating substrate mycelium after 18 hours. At approximately 2 days, aerial hyphae grow up by the activation of *bld* genes. The apical compartment produce unigenomic spore compartments at 4-10 days eventually resulting in the production of mature spores (Klieneberger-Nobel, 1947).

1.6. Actinomycete taxonomy

Taxonomy creates a universal understanding by illustrating relationships between organisms, resulting in their placement within a defined group according to these traits. Due to continuous advances in techniques, the methodology lacks a single, defined protocol, resulting in several different systems for taxonomy and further contradictions within these (Ereshefsky, 1994; Hey, 2001). This confusion is illustrated within the *Streptomyces* genus, a large and complex group of organisms. Indeed, even the earliest actinomycetes classifications were subject to criticism (Waksman and Henrici, 1943).

1.6.1. Classical taxonomy

Streptomycetes are not simple organisms. Their diversity is reflected within the history of their classification, which contains a significant number of contradictory results, necessitates constant amendments and demonstrates an overall state of over-speciation. Various techniques which are utilised are not re-producible and therefore, cannot be easily applied (Mehling *et al.*, 1995). Other techniques clearly contradict previous studies rendering one or both invalid, for example the comparisons of (Pridham *et al.*, 1958) and (Waksman, 1961). These problems have been highlighted in the review by (Anderson and Wellington, 2001). The number of methods for detecting and speciating bacteria is overwhelming and causes much of the confusion associated with bacterial taxonomy.

Even the early classifications, which relied upon morphology and cell wall biochemistry, were subject to criticism. The Waksman publication of 1940 was scrutinised for the placement of the entire actinomycete family - originally a term used to describe an aerobic, pathogenic, spore-forming species. Even at this early stage of classification there existed confusion about the name used to describe these organisms (Waksman and Henrici, 1943). The morphological arrangement of the mycelia was a key feature in early classification – aerial or vegetative, chains or cocci (Waksman and

Henrici, 1943). As well as the cell wall chemistry, type I for the streptomycetes. Type I is defined by the presence of LL-diaminopimimelic acid and glycine and the absence of typical sugars found in cell walls (Lechevalier and Lechevalier, 1970). Due to the controversy surrounding streptomycete taxonomy, the International *Streptomyces* Project was established in 1964. This organisation drew up a list of specific morphologies used to describe streptomycetes: spore chain morphology, spore surface, substrate mycelium, soluble pigment production, melanin production and the utilisation of carbon sources.

However, as technology and methodology advanced, the standard criteria changed and became even more complicated. Additionally, it has been shown by (Silvestri *et al.*, 1962) that the characteristics used to describe the streptomycetes are highly variable and subject to user interpretation. All these complications are indicative of an urgent need for a simple, yet accurate system.

The most common classical technique is chemotaxonomy, a way of defining bacterial species by biochemical means (Cummins and Harris, 1956). It involves the examination of characteristics unique to the individual bacterial species under study. The methods used involve examining the fatty acid chains by gas chromatography, assessing carbon utilization and whole cell analysis by mass spectrometry (Abdel *et al.*, 1963; Shirling and Gottieb, 1968). These traditional methods continue to be improved and used alongside more modern techniques such as 16S rDNA sequencing.

Additionally, bacterial surface antigens are also used to exploit their specific receptivity to particular bacteriophages. This has resulted in phage typing, more commonly used amongst pathogenic bacteria such as *Salmonella* and *Listeria*. It has however, also been applied to streptomycetes. Despite occasional cross-reactivity (Bradley, 1961), phages are generally considered to be genus specific. Additionally, a study examined 905 Actinomycetes and concluded that overall phage typing for the actinomycetes is a useful aid in their taxonomy and has subsequently helped to re-classify some strains (Foor, 1990; Korn-Wendisch and Schneider, 1992; Kurtboke, 2005). Results provide evidence that phage typing is a useful aid for the identification of actinomycetes at the genus level but is not as successful for species definition.

Another method of classification is serological taxonomy, which involves the rising of anti-sera or highly specific mono-clonal antibodies against bacterial surface antigens. Within the *Streptomyces* species this technique proved useful as it has been shown, in some cases, to be species specific and able to determine the microbial population within a natural setting. (Wipat *et al.*, 1994) have devised mono-clonal antibodies specific to *Streptomyces lividans*. Their technique was successfully applied in order to isolate Streptomycete spores directly from soil. Other studies also support the ability of serology in order to define cluster groups, however, these are not reproducible (Kirby and Rybicki, 1986; Williams *et al.*, 1983a; Williams *et al.*, 1983b).

To date, however, the most relevant and supported, non-molecular method of classification of organisms is through numerical taxaonomy, as detailed in the renowed Bergey's manual. (Williams *et al.*, 1983b) devised a probalistic identification matrix which clarifies the intergeneric relationships within the *Streptomycetaceae* family. This method involves the clustering of species groups, within which there may be high diversity, according to their taxonomic traits: one hundred and thirty nine in total. It devised twenty-three major clusters, each with four or more strains, twenty minor clusters with two or three strains and twenty-five single member clusters. The largest cluster is cluster one, containing seventy-one strains, phenotypically described as those which produce yellow - grey pigmentation yet lack melanin production with smooth spores in straight chains, resistant to some antibiotics. This method has established the nature of several relationships, therefore resulting in the re-classification of some strains (Wayne *et al.*, 1987; Williams *et al.*, 1983b).

There is a large amount of data on bacterial taxonomy and deciphering it requires a great amount of investigation. The conventional methods described above are now making way for more modern molecular methodology, and in some cases the combination of both is used.

1.6.2. Molecular taxonomy

Great advances have been made in bacterial taxonomy and phylogeny since the advent of molecular biology. The primary procedure for initial screening in molecular

taxonomy is DNA-DNA hybridization which alongside the percentage G/C content of the bacterial genome can provide a crude measure of similarity between bacterial strains. A threshold of 70% defines a bacterial species for total chromosome DNA-DNA hybridisation with an equal to or less than 5°C melting temperature (Wayne *et al.*, 1987). The examination of DNA relatedness has allowed for the reclassification of several species once grouped in accordance to numerical taxonomy. According to the *Bergey's Manual of Systematic Bacteriology* the genus *Streptoverticillium* overlapped with the *Streptomyces* genus. However, it was proposed that by analysis of 16S rRNA that the two should be unified. (Labeda, 1996).

Regardless of the widespread acceptance of this technique, it too has been subject to criticism. It has been claimed that the 70% figure has been artificially derived and is highly dependant upon previous bacterial species definitions (Stackenbrandt *et al.*, 2002). When compared to the definition of animal species, this would result in *Homo sapiens* belonging to the same species group as almost all primates, including lemurs. Furthermore, DNA-DNA hybridisation does not take the evolutionary processes or mutations, which can affect the life-style or characteristics of the organism in question, into account. This technique is subject to high criticism as it is thought to underestimate grossly the sheer volume of bacterial diversity present within our ecosystems and to define a bacterial species in terms which are too simplistic (Govan *et al.*, 1996). Genomic phylogenetic species concept was proposed as a technique that would provide a conceptual and testable framework for bacterial taxonomy and would replace the need for DNA hybridisation. This relies on DNA/RNA sequences that are resistant to change, and make up a core set of genes which define that organism

However, there are several alternative methods utilising total chromosomal DNA for the purposes of taxonomy, for example, low frequency restriction fragment analysis which uses rare cutters. Following examination by pulsed field gel electrophoresis, a specific finger-print of that organism is derived. Under comparison, this can reveal the relatedness, and therefore the clustering, of some strains. Unfortunately, some results contradict those found in previous tests. For example, in the case of *Streptomyces*

cyaneus which when subject to phenotypic tests cluster at 90% similarity but only at 58% similarity when examined by LFRFA (Rauland *et al.*, 1995).

The most powerful tool which has transformed bacterial taxonomy is the comparison of rRNA sequences, first described by (Fox *et al.*, 1977). This rapid technique is universally supported and accepted, and is therefore applied in almost all modern taxonomic studies (Woese, 1987). This method has been applied in determining the genus, species and strains of organisms with the additional ability to clarify between intra and interspacific differences. In the use of SSU rRNA sequence taxonomy, (Stackebrandt *et al.*, 1991) the importance of the region selected for sequencing has been highlighted. The variable regions typically used are α, β and γ (Figure 3). By sequencing the variable α region out of 89 streptomycete strains, according to standard ISP, 57 types were identified and 42 possessed unique sequences. Kataoka *et al* have also undertaken the task of sequencing 485 16S rRNA γ regions of streptomycetes and placing these in GenBank., showing that clusters formed under phenotypic taxonomy. Although it can be argued that SSU rRNA is a worthwhile contribution to bacterial taxonomy, it remains limited despite the growing number of sequences in the database (Kataoka *et al.*, 1997) (Labeda, 1996). For example, when the 16S rRNA sequences of 14 marine isolates were compared, the majority did not contain a match in the database. This shortcoming reinforces the ecological niche theory of a unique highly suited population of inhabitants occupying each environment (Ranjard and Richaume, 2001). It seems that as yet the database is not complete, rendering it only useful in identifying already characterised organisms (Schmidt *et al.*, 1991). An additional problem is that, although 16S sequences can be identical between 2 'strains', DNA-DNA hybridisation displays them as the two defined species (Fox, 1992). However 16S remains a technique that readily allows for the rapid reclassification of bacteria. For example, Streptomyces *caeruleus* has been reclassified as a synonym of *Actinoalloteichus cyanofgriseus* using this technique (Tamura *et al.*, 2008).

Contradictory results have been obtained by the work on *Burkholderia* species: yet again illustrating the shortcomings of 16S rRNA sequencing. This attempt involved *B.thailandensis, B.pseudomallei* and *B.mallei,* whose 16S rRNA sequences are 99%

identical (Brett *et al.*, 1998). Whereas previous studies have shown their DNA-DNA hybridization to give contradictory results, *B.pseudomallei* and *B.mallei* display 76% DNA-DNA binding while *B.thailandensis* is clearly different from *B.pseudomallei* and displays only 47% DNA-DNA binding (Rogul *et al.*, 1970; Yabuuchi, 2000). MLSA has elucidated this conflict by examining seven housekeeping genes (Goody *et al.*, 2003). This study has clearly shown that *B.psudomallei* and *B. thailandensis* are different species and that *B.mallei* is a clone of *B.psudomallei,* and should therefore not be given different names (Gevers *et al.*, 2005).

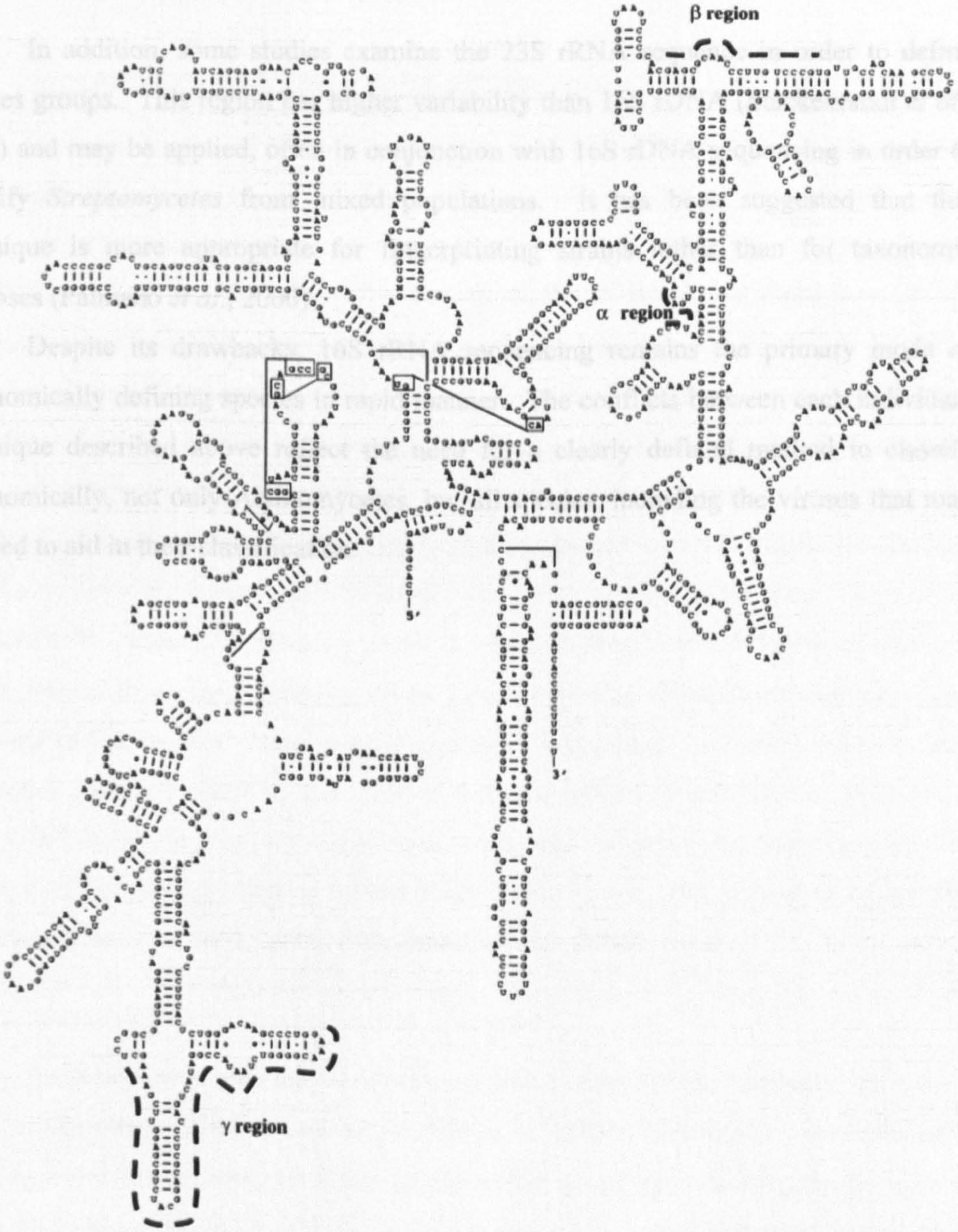


Figure 3 Secondary structure of 16S rRNA from *Streptomyces coelicolor*. Variable regions are used as diagnostics for the *Streptomyces* genus (γ, β and α). Taken from (Anderson and Wellington, 2001).

In addition, some studies examine the 23S rRNA sequence in order to define species groups. This region has higher variability than 16S rDNA (Stackebrandt *et al.*, 1991) and may be applied, often in conjunction with 16S rDNA sequencing in order to identify *Streptomycetes* from mixed populations. It has been suggested that this technique is more appropriate for fingerprinting strains rather than for taxonomic purposes (Palmano *et al.*, 2000).

Despite its drawbacks, 16S rRNA sequencing remains the primary mode of taxonomically defining species in rapid manner. The conflicts between each individual technique described above reflect the need for a clearly defined method to classify taxonomically, not only Actinomycetes, but all species, including the viruses that may be used to aid in their classification.

1.7. A general introduction to bacteriophages

Bacteriophages are the most abundant genetic entity on Earth with the total population estimated to be 10^{31} (Wommack *et al.*, 1992). This number accounts for a great mass of organic matter. Additionally, it is not an innate mass but is responsible for the stability, death and evolution of, at very least, the microbial communities it infects and, on a greater scale, the entire ecosystem it inhabits (Wihelm and Suttle, 1999). From the time of their discovery by d'Herelle and Twort in the 1910s (Twort, 1915), these ubiquitous, bacterial viruses have been detected in nearly all environments: soil, thermal vents and sulphur springs are just a few examples. Research on bacteriophages was initially carried out in the hope that they would provide a solution to hugely problematic infectious bacterial diseases (Sulakvelidze *et al.*, 2001). However, upon the discovery of antibiotics by Alexander Fleming in 1928, phage research was reduced as there no longer remained an urgent need to invest in these viruses. Despite this development, some isolated pockets of phage research remained, particularly in Eastern Europe where antibiotics were too costly to use. Today, with the advent of antibiotic resistance and molecular biology, interest in these life forms has been re-ignited and phage research has allowed for the development of these viruses as molecular tools, indicators of specific bacteria and as therapeutic agents (Sulakvelidze *et al.*, 2001).

1.7.1. Bacteriophage structure and lifecycles

Bacteriophages are structurally the simplest genetic entity, minimally comprised of a protein coat; a capsid, and ds or ssDNA or dsRNA (figure 2). Some do have additional features such as tail fibres. They utilise a very specific receptor in order to dock onto their host and begin the infection process. A good example is the well-characterised lambda phage tail protein that attaches to a maltose receptor (LamB) on the cell wall (Berkane *et al.*, 2006; Hazelbauer, 1975). Initial binding can be reversible whilst subsequent formation of covalent bonds is irreversible. Phage nucleic acid is then injected into the host by means of a tail tube. Once within the host the phage replicates

to generate new phages or, in the case of temperate phages, integrated and lays dormant within the host chromosome. This mechanism is only made possible by the possession of an integrase gene within the phage genome.

There are many studies that have successfully isolated *Streptomyces* phages from soil (Ackermann *et al.*, 1985; Hahn, 1991; Mellaert *et al.*, 1998). They were primarily conducted in order to exploit the use of phages as tools for the genetic analysis of their hosts (Stuttard, 1983). These phages typically are temperate and have a wide host range, except VWB which displays a narrow host range (Mellaert *et al.*, 1998). However, the most extensively studied actinophage is ϕC31.

1.7.2. Bacteriophage Genome Structure

When compared, the most striking feature of phage genomes is mosaic structure they exhibit in relation to one another. The first indication of this was by Simon *et al* in 1971 with the use of electron microscopy, which revealed DNA-DNA heteroduplexes in *E.coli* phages (Simon *et al*, 1971). With the advent of molecular genetics and phage sequencing this mosaicism was observed with genes found in blocks. When two phage sequences are aligned, areas of sequence similarity are clustered, followed by distinct transitions to regions of little or no sequence similarities. An illustration of this is found within the major capsid genes of phages HK97 and HK022, which share 99% similarity nucleotide sequence. However, the nucleotide sequences for their DNA replication proteins share only 33% similarity (Juhala *et al.*, 2000).

1.8. Classification and taxonomy of bacteriophages

Since their discovery, phages, like their hosts, have gone through several methods of taxonomic classification and to date there remains no universal agreement. In the 1920s and 30s, host range was thought sufficient and as a result began the advent of 'phage typing' (Lawrence *et al.*, 2002). Subsequently, with the emergence of the electron microscope, morphological differences between phages were used according to Bradley's classifications (Bradley, 1961). More recently, the post-genomic era has added yet another factor into the phage taxonomic schemes.

As there is no single feature, such as the bacterial equivalent of ribosomal DNA, phages are loosely grouped by concepts lacking in distinct boundaries, thereby resulting in quasispecies (Eigen, 1993). Consequently, this dilemma is subject to individual interpretation and creates debate amongst phage biologists. Phenotypic features have traditionally been, and perhaps still remain, the most widely used characteristics in clustering phages (see Figure 4). Some of the first studies relied solely upon host range, however, with technological advances, more features began to be considered. In 1961, soon after the development of the electron microscope, Bradley's classification system was proposed in order to group phages into one of six morphologies (Bradley, 1961). Although this system remains to be used when describing almost all newly characterised phages it is not included within the universal taxonomic system which was created by the International Congress of Microbiology. This group established the International Committee for the Taxonomy of Viruses whose objective, since 1971, is continually to update taxonomic guidelines. The ICTV are true to the original Linnaean hierarchical system which placed the tailed phages under the order of the *Caudovirales*. Within this order there are three families: the *Myoviridae*, with long, contractile tails, the *Siphoviridae* with long, non-contractile tails, and finally, the *Podoviridae* with short, stubbed tails and a striking lack of features (figure 5).

Each of these three families may be further broken down into genra by considering their host range, genome size and genome form. An example of ICTV classification is the *Podoviridae* which contains both the *Salmonella* P22 phage and the coliphage T_7 (Steinbacher *et al.*, 1996; Zhang *et al.*, 2000). They both fall under this order due to their short tails, however, it is known that P22 shares great genetic homology with the *Siphoviriade* λ phage (Susskind and Botstein, 1978). Indeed, their similarity is so great that recombination between them results in fully functional hybrids (Retallack *et al.*, 1994). There are several similar examples demonstrating the fact that taxonomic groupings according to ICTV may not reflect common ancestry and therefore that this type of classification may be a flawed, although a widely recognized method of speciation. Additionally, morphology of the viruses is essential for classification according to this system. However, in the phage genome database, a large majority have

never been examined under the microscope and consequently cannot be taxonomically grouped according to ICTV (Lawrence *et al.*, 2002).

An obvious alternative in classing viruses would be to identify their overall similarities and group them according to these, rather than using a single core trait. The phage proteomic taxonomic tree is a system which exploits phage similarity and is genome-based (Figure 5). Its construction involved all 105 sequenced phages and the comparison of their predicted protein sequences. No single common protein was found in all the genomes; the closest, with 56 matches, was the YomI, a putative transglycosylase from *Bacillus subtilis* ϕSPBc2 (Rohwer and Edwards, 2002b). Despite solving problems such as those previously discussed in the case of the P22, T7, λ classification according to ICTV, this system has been criticised. Lawrence *et al* (2002) drew up two hypothetical, nearly identical phylogenetic trees based on overall similarity. This comparison resulted in strikingly misleading information by the simple loss of a single phage from one of the trees, illustrating problems with the whole concept of overall similarity (Lawrence *et al.*, 2002).

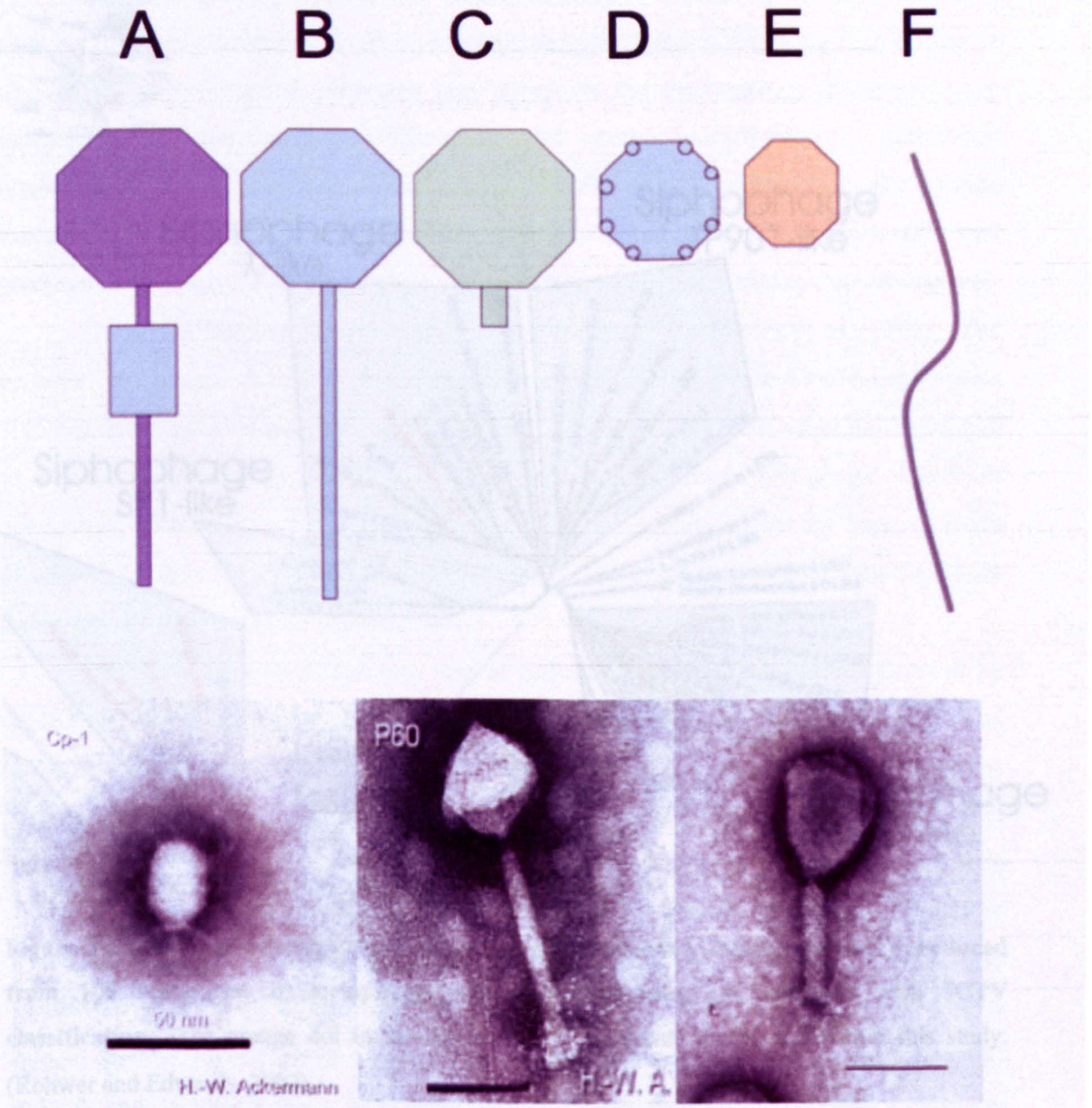


Figure 4: Classification of bacteriophages according to morphological traits. Above: The six phage morphologies according to Bradley's classification, A to F. Below: EM of three phage morphologies according to ICTV classification: *Podoviriade* (C according to Bradleys), *Siphoviriade* (B according to Bradleys),, *Myoviriade* (A according to Bradleys),, from left to right.

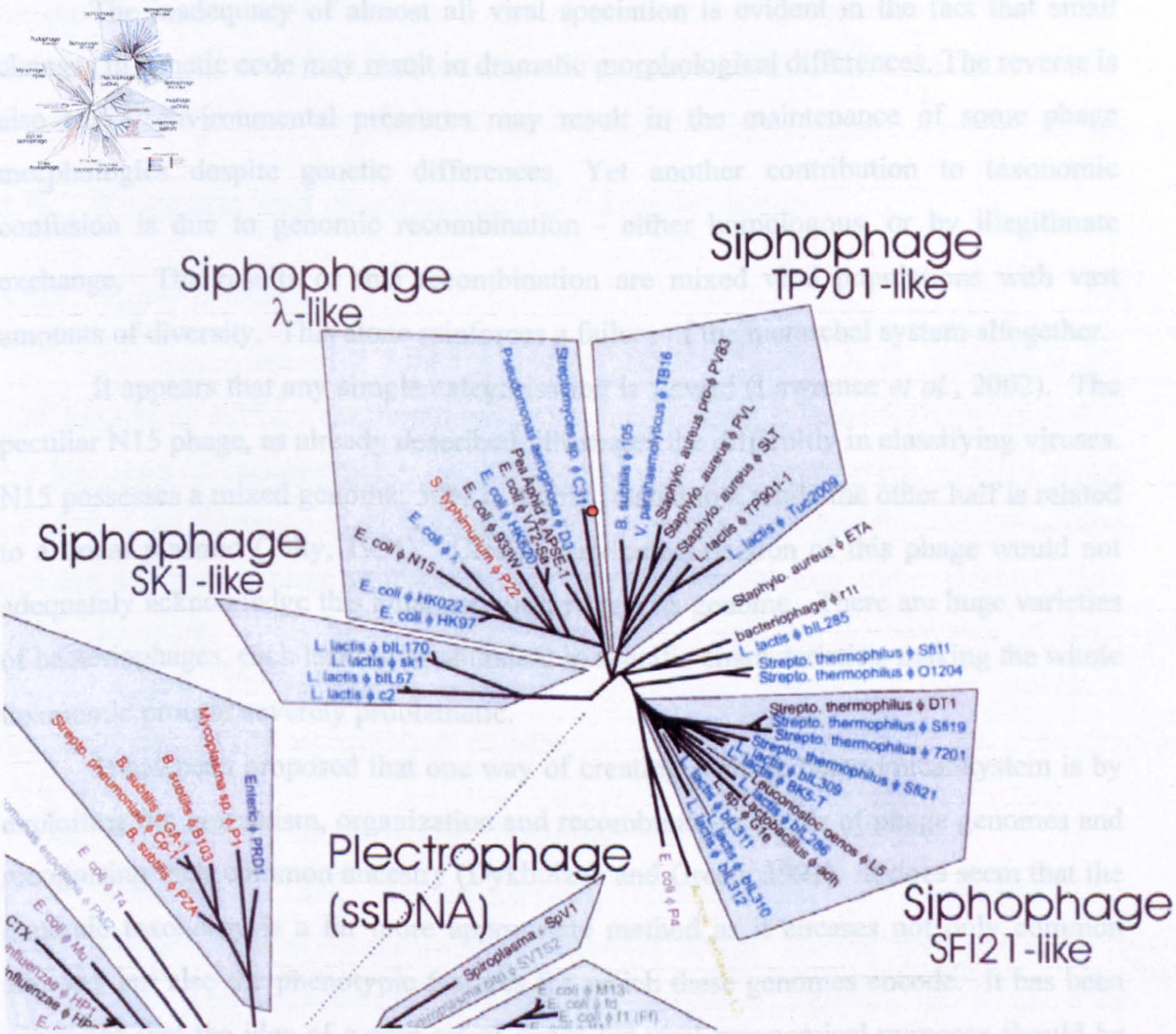


Figure 5 Siphophage branch of bacteriophage proteomic tree (full tree top left) produced from 105 sequenced bacteriophage genomes, illustrating the conflicts with ICTV classification. The orange dot in the λ-like cluster represents the phages within this study. (Rohwer and Edwards, 2002)

The inadequacy of almost all viral speciation is evident in the fact that small changes in genetic code may result in dramatic morphological differences. The reverse is also true - environmental pressures may result in the maintenance of some phage morphologies despite genetic differences. Yet another contribution to taxonomic confusion is due to genomic recombination - either homologous, or by illegitimate exchange. The results of this recombination are mixed viral populations with vast amounts of diversity. This alone reinforces a failure of the hierarchal system altogether.

It appears that any simple categorisation is flawed (Lawrence *et al.*, 2002). The peculiar N15 phage, as already described, illustrates the difficultly in classifying viruses. N15 possesses a mixed genome: 50% is highly related to λ while the other half is related to a linear plasmid (Tilly, 1991). Clearly, any categorisation of this phage would not adequately acknowledge this amalgamation within its genome. There are huge varieties of bacteriophages, each lacking or abundant in specific characteristics, making the whole taxonomic process severely problematic.

It has been proposed that one way of creating a phage taxonomical system is by exploiting the mosaicism, organization and recombination events of phage genomes and recognising their common ancestry (Dykhuizen and Green, 1991). It does seem that the genomic taxonomy is a far more appropriate method as it encases not only common ancestry but also the phenotypic features for which these genomes encode. It has been suggested that the idea of a core set of genes for viral taxonomical purposes should be abandoned in favour of relatedness by means of a shared pool. Indeed, if these phages are subject to genetic exchange amongst themselves, there are likely to be some shared aspects of their lifestyles. It appears that more recent proposals for phage taxonomy methodology recognise that a great amount of involvement is required before a virus can be clearly placed into any one group (Van Regenmortel *et al.*, 1997). Controversially, some even suggest that bacteriophages should belong to more than one taxonomic group at the same level (Lawrence *et al.*, 2002). However, in order to create the ideal taxonomical system for phages, it is first necessary to obtain the information utilising all the incomplete data. One way of doing this is by studying and classifying their hosts

before attempting phage taxonomy (see section 1.6). One of the problems posed in all prokaryotic taxonomy is the inability to culture all microorganisms from their natural environment, such as soil (Athalye *et al.*, 1985), as a result it is difficult to examine the full impact viruses have upon their surrounding environment.

1.9. Evolution and relationships of bacteriophages

1.9.1. Early Evolution

The moron accretion hypothesis postulated by Hendrix *et al* (Hendrix *et al.*, 2003) describes the very first evolutionary mechanisms that, ultimately, have resulted in the formation of bacteriophages. In this process, a rare mutation event occurred which resulted in the formation of an icosahedral protein shell. With the chance encapsulation of DNA a primitive phage would result and be preserved due to positive selection. Any future mutations, which gave a positive result, would maintain the formation of these shells. Clearly this hypothesis is merely speculative; however, it does provide a Darwinian account of bacteriophage creation.

1.9.2. Recombination events are responsible for phage evolution

Homologous and non-homologous recombination is responsible for bacteriophage evolution (Recktenwald and Schmidt, 2002). The Rec system in *E.coli* is one of the best-characterized systems in prokaryotes resulting in homologous recombination. RecA plays a major role within this system, as it is responsible for aiding the catalysation of the DNA synapsis reaction between homologous regions (Joo *et al.*, 2006; Kowalczykowski, 2000). In contrast, phage evolution may also occur by non-homologous end joining. Within this system, sequences of DNA can be brought together without the need for long sections of homologous DNA. Only short regions of homology between sequences are required to guide repair (Moore and Haber, 1996).

An example of phage evolution as a result of these events is found within the lamboid-like N15 phage. One half of this viral genome is very similar to lambda, however, the other half is from an unrelated group of phages. This illustrates the result of extensive phage evolution (Ravin *et al.*, 2000). In specialised transduction, the

incorrect excision of prophages from host chromosomes may result in the phage acquireing new bacterial genes, however, this is viewed as a limited mechanism for phage evolution. There are currently two models for recombinant bacteriophage evolution. Susskind (1978) proposed the initial modular theory was that phage evolution took place by homologous recombination at short conserved boundaries of gene modules. This accounts for the evolution of coliphages HK620 and P27, which was, at least in part, due to blocks of genes for a particular function, for example, the DNA metabolism genes, being homologous to modules with the same function in other phages (Clark *et al.*, 2001; Recktenwald and Schmidt, 2002; Susskind and Botstein, 1978). Hendrix (2003) has postulated an alternative theory which states that the majority of phage evolution is due to illegitimate exchange; this is also evident within actinophages and mycobacteriophages (Pedulla *et al.*, 2003b). This non-homologous recombination between phages is not sequence dependant, therefore, it mostly produces vast quantities of non-viable constructs. The few recombinants, which are not eliminated, maintain boundaries where the initial recombination event took place, thereby, conserving the gene junctions and therefore, the modules within the genome (Hendrix *et al.*, 2003).

Although recombination events generally take place at gene boundaries due to selective pressures, such as a sudden change in host populations, there are some exceptions. The best known is the case of coliphage tail fibre genes, which show mosaicism, not only at gene boundaries, but also within gene segments. Evidence of this has also been found in the intergrase and homologous recombination genes (Juhala *et al.*, 2000). In both cases, the event has taken place at an area of the gene corresponding to a protein domain boundary resulting in a fully functional protein. It has been proposed (Hendrix *et al.*, 2003) that the allowance for such intragenic mosaicism within the tail fibres is not purely due to the selective advantage, but due to the final conformation of the gene product. Unlike globular proteins, tail fibres are not closely packed which means that individual amino acids do not make intimate contact, therefore allowing toleration of changes in gene coding.

Conversely, recombination within or at boundaries of the head genes is rarely detected. This conservation is due to the co-evolution of capsid proteins with its

surrounding genes. As a result, each gene product requires other co-evolved proteins to interact successfully. The few examples of recombination within this area are possible due to the removal of DNA and the subsequent replacement with a similar DNA sequence. An illustration of this replacement can be found at the junction of 2 head genes in HK97 and HK022 (Juhala *et al.*, 2000).

In addition, it has recently been shown that recombination is remarkably efficient within, at least, the lamboid phages by the use of Red/Gam like systems. These systems allow for the recombination between diverged DNA (Martinsohn *et al.*, 2008).

1.9.3. Morons & more recent phage evolution

Morons are genetic elements within phage genomes with a transcription promoter and terminator. They stand out from the rest of the genome due to their difference in G/C content in comparison to the flanking regions (Hendrix *et al.*, 2000). These are present as a result of relatively recent evolution of phage genomes. It has been hypothesised that these short sequences entered into the genome by random, non-homologous recombination followed by the deletion of 'extra' DNA segments. The conservation of these morons suggests that they are advantageous. Another section of phage DNA that stands out from its surrounding sequence are groups of less tightly packed genes - intragenic mosaics derived from other phages (Hendrix *et al.*, 2003). Hypothetical codon ORF 34 within HK97 reflects the typical and unusual structure of these DNA segments (Juhala *et al.*, 2000). Following the translational start, Shine-Dalgarno sequence, a gene from phage lambda is fused to an unknown sequence which is attached to the *xis* gene from P22. This *xis* sequence is found within an incorrect reading frame, signifying that this gene does not result in a functional protein (Gottesman and Abremski, 1982; Hendrix *et al.*, 2003; Mattis *et al.*, 2008). These types of assemblies within the phage genomes do not provide any selective benefit to the phage, nor are they disadvantageous. Hendrix proposes that these genes are maintained in order for efficient packaging of DNA during phage assembly – termed "stuffer" DNA. This DNA may act as a pool of genes allowing phages to adapt according to their

surrounding environment. In addition, due to the rapid recombination events within phages, it also provides means to drive evolution within the phage community.

1.9.4. Common ancestry of phages

The use of the term 'species' is inadequate when attempting to group bacteriophages; it is due to their extensive and merged evolution over long periods of time that the renowned difficulty in their classification has resulted (Lawrence *et al.*, 2002). Sequence similarity between phages provides an indication of the relationships between recently diverged phages, however, in order to determine more distant relations other comparisons must be made. Amino acid sequence, genome organisation and architecture are capable of revealing more subtle and ancient relationships between these viruses (Rohwer and Edwards, 2002a) (Hendrix *et al.*, 2000).

It is in this way that phages of distinctly distant hosts have been identified. Most notably, the shared common ancestry between phages with the Gram negative *Escherichia, Salmonella, Haemophilus* hosts and those with the Gram positive *Mycobacterium* and *Streptomyces* hosts. A key link between these viruses is ϕflu; a cryptic prophage containing several homologous genes and gene arrangements with many phages. This prophage is accountable for the addition of a 603bp sequence within the D29 gene gp10 (Hendrix *et al.*, 1999). Yet another example of prophage ancestry is in the form of mycobacteriophages ϕRv1 and ϕRv2. These small prophages exhibit homology with ϕC31, L5, D29, TM4 and HK97 (Cole *et al.*, 1998; Duda *et al.*, 1995; Hatfull, 1993; Philipp *et al.*, 1996; Popa *et al.*, 1991). Similarly, the actinophage ϕC31 acts as a bridge between coliphages and mycobacteriophages. For example, the ORF 9a in ϕC31 shares homology with gp70 in the mycobacteriophage TM4 and also displays similar gene organisation of its head proteins to the coliphage, thus providing a link between all these phages. Phylogentically distant phages have been shown to share a common gene pool (Hendrix *et al.*, 1999).

1.10. The effect of phages on their hosts and the surrounding ecosystem

Viruses are major contributory factors in many processes involving microbial mortality to global geochemical cycles (Suttle, 2005). The influence of phages in the environment is too frequently ignored, however, there are some key studies, particularly within the marine environment, which reveal their value and influence in the ecosystems they inhabit. Microbial survival or demise is dependant on bacteriophage populations and their chosen lifecycles. Azam *et al* (1983) described the microbial loop in 1983, a system responsible for the replenishment of DOC within the marine food web (see Figure 6) (Azam *et al.*, 1983). Bacteriophages are instrumental in this process and are accountable for 30% of bacterial mortality in the cycle and as much as one quarter of DOM in the sea flows through the viral shunt. The liberation of cytoplasmic and structural matter from prokaryotes by phages is estimated to release 1μg/L of DOC per bacterial generation. Indeed, it is not only bacteriophages that are important in the viral shunt also of some importance are cyanophages - viruses which infect algal cells. The populations of these viruses have been shown to correlate with, and indeed, control algal blooms (Azam *et al.*, 1983).

Yet another example of phage populations influencing their host populations is found within cholera outbreaks. The cyclic nature of cholera outbreaks is evident in their occurrence twice a year. Faruque *et al* proposed (2005) proposed that bacteriophages might be an important intrinsic factor in these outbreaks. A study was carried out in Bangladesh, linking the virulent bacterial strain with a specific phage (JSF4). It found an inverse correlation between the presence of vibriophages and the presence of a susceptible strain. The conclusions demonstrated that the absence of phages in the water promotes cholera epidemics (Faruque *et al.*, 2005). In addition, temperate bacteriophages create bacterial diversity by causing extensive horizontal gene transfers between bacterial species. Again, using *Vibrio cholerae* as an example, the bacterium is only pathogenic if it is infected with a temperate bacteriophage carrying the toxin gene, thereby, increasing the fitness of the host (Bik *et al.*, 1995).

One of the most dramatic and obvious effect of viruses upon their hosts is that of algae, in particular the effect of viruses upon *Emiliania huxleyi* blooms. Blooms of this organism can be observed by satellite on a global scale due to reflective calcium carbonate coccoliths. Frequently these blooms are observed to disappear suddenly causing a flux in calciate and cloud forming dimethyl sulfide to the atmosphere. This global impact has been shown to be at least in part due to lysis caused by highly virulent viruses (Schroeder *et al.*, 2003).

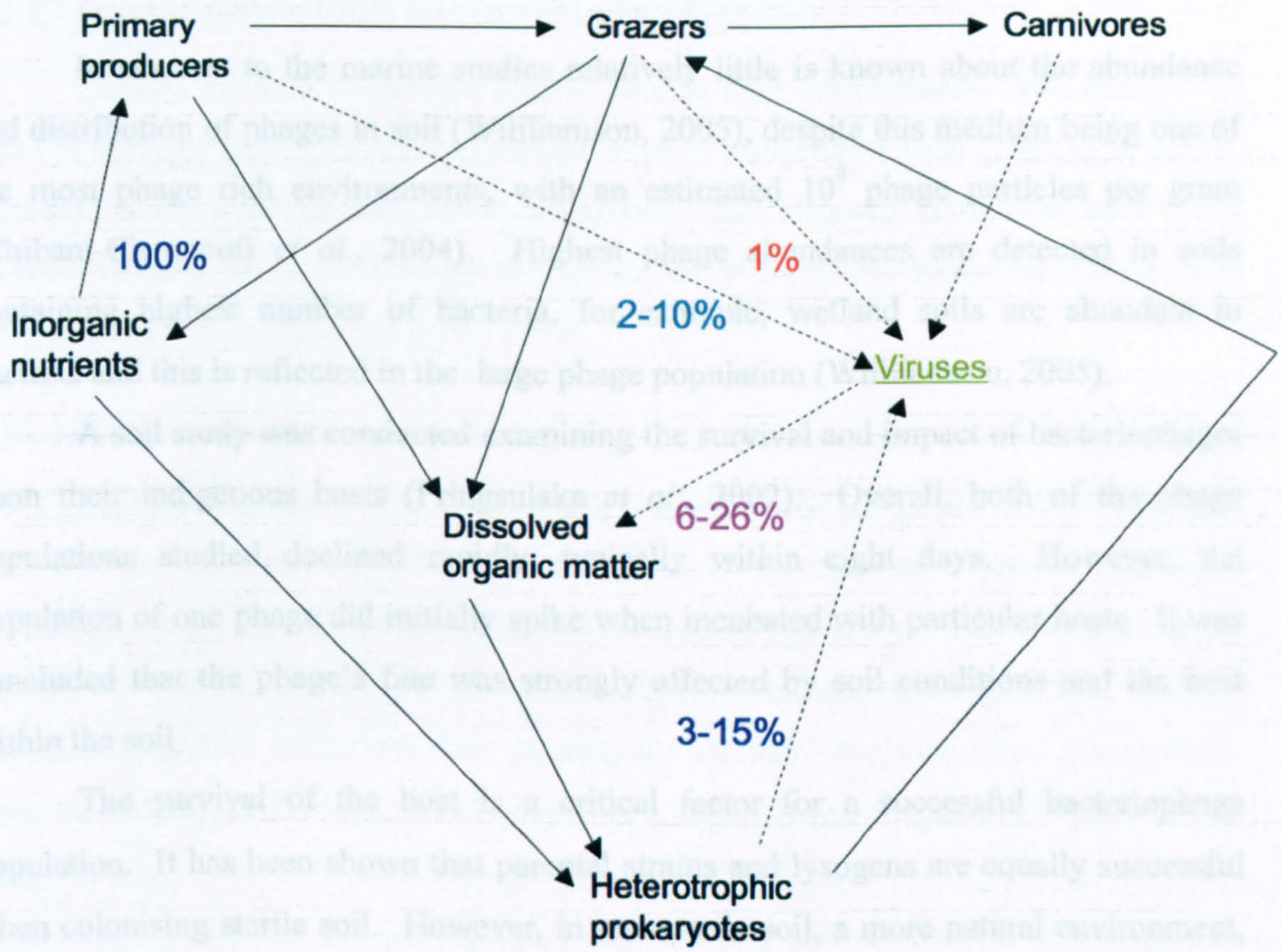


Figure 6 The grazing food chain. Dotted lines depict virus mediated pathways. Model adopted from (Fuhrman, 1999). Viruses in the marine environment account for a total of 12-52% of organic turnover in the sea. This is an example of their potential in the global turnover in the ecosystem.

In contrast to the marine studies relatively little is known about the abundance and distribution of phages in soil (Wililiamson, 2005), despite this medium being one of the most phage rich environments, with an estimated 10^9 phage particles per gram (Chibani-Chennoufi *et al.*, 2004). Highest phage abundances are detected in soils containing highest number of bacteria, for example, wetland soils are abundant in bacteria and this is reflected in the large phage population (Wililiamson, 2005).

A soil study was conducted examining the survival and impact of bacteriophages upon their indigenous hosts (Pringsulaka *et al.*, 2002). Overall, both of the phage populations studied declined rapidly, typically within eight days. However, the population of one phage did initially spike when incubated with particular hosts. It was concluded that the phage's fate was strongly affected by soil conditions and the host within the soil

The survival of the host is a critical factor for a successful bacteriophage population. It has been shown that parental strains and lysogens are equally successful when colonising sterile soil. However, in non-sterile soil, a more natural environment, the lysogens died rapidly whereas the parental strains remained viable. Therefore, the lysogenic form of the viruses may not be as advantageous to them as once thought (Herron and Wellington, 1990). Due to the impact phages can have on their surroundings, they are now being exploited in order to modify, destroy or detect their hosts.

The dramatic impact of a pseudomonas phage: ϕGP100, in soil was illustrated by Keel *et al* in 2002. The protective bacterium *Pseudomonas fluorescens* colonizes cucumber roots preventing harmful infection of other organisms. However, in the presence of ϕGP100 this protection was completely lost due to the demise of the protective bacterium. In contrast, the presence of this phage had no impact upon root colonisation of phage resistant strains of the bacterium. This is an excellent illustration of how phages may be applied as a therapeutic agent within soil (Keel *et al.*, 2002).

1.11. Host defence tools against phage infection

Clearly, bacteria do survive despite sharing their environments with phages. In order to survive, they have evolved several ways of avoiding successful phage infection. These traits are predominantly linked to plasmid DNA, although some have been found on chromosomal DNA, for example the lactococcal plasmid pNP40 mediates resistant to a phage of lactococci (Garvey *et al.*, 1996).

1.11.1.Initial infection

There are two means by which a host can prevent or halt phage entry: adherence inhibition and prevention of phage injection. One of the simplest techniques employed by bacteria is to alter or lack specific cell receptors that phages can use as docking systems, as is the case with Ox2 resistant *E.coli* (Morona and Henning, 1984). In some cases, although the bacterium does possess the cellular receptors, other components, such as exopolysaccharides, which physically mask it, are in place (Deveau, 2002). Alternatively, bacteria which possess these receptors can allow the phage to dock, rendering it immobile but prevent its DNA entering the cell, which is the defence mechanism encoded by the pNP40 (Garvey *et al.*, 1996).

1.11.2. Phage exclusion

Unsuccessful infection, also known as abortive infection, is another host defence system against phages. It permits the virus to enter the cell but prevents the phage from completing its cycle successfully, which is achieved by the host defences at several stages within the phages cycle. This abortive infection has been described by Sing and Klaenhammer in 1993 and termed Abi mechanisms (Sing and Klaenhammer, 1993). Typically plasmid encoded, these mechanisms reduce burst size, efficiency of plating, and the formation of smaller infection centres. Abi systems ensure that the host cells do not die, therefore preventing the release of new infective viral particles. Abi systems are subject to much study within dairy industry due to the importance of phage free starter cultures (Daly *et al.*, 1996; Sullivan *et al.*, 1998). To date, there are 21 Abi mechanisms found within *Lactococcus lactis*, all of which, apart from two, are plasmid encoded. Not

all of the 21 Abi traits have fully been characterised, those which have been, are briefly described in Table 1.

Abi	Phage target
A	Interfere with DNA replication
F	Interfere with DNA replication
K	Interfere with DNA replication
R	Interfere with DNA replication
B	Interfere with RNA transcription
G	Interfere with RNA transcription
C	Reduces synthesis of major capsid protein
D1	Interference of phage ORF required for phage development
Q	Prevents maturation
S	Sequesters factors needed for phage development
U	Delays phage transcription

Table 1 The mechanisms of abi systems

1.11.3. Restriction modification

Some bacteria possess type II restriction modification systems. These systems are widespread in almost all bacterial hosts and result in the prevention or retardation of phage infection. The system typically involves pairs of enzymes which work independently, but are found linked on the chromosome. One of these enzymes is DNA methyltransferase which specifically methylates A or C residues, rendering it immune to destruction by the second enzyme. This is an endodeoxyribonuclease which cleaves DNA at a specific site if it is lacking in methylation (Rocha *et al.*, 2001). Despite this being one of the most well- known host defences against phage, there are some studies which indicate that this system is extremely ineffective in contributing to overall host survival. On the whole, this system is not fully adequate in dealing with all phage invasions. It is only effective against dsDNA viruses which signifies that the host is

defenceless against ssDNA or RNA viruses (Levin, 1993). In addition, some of the dsDNA phages have developed their own defence mechanisms such as the inhibition of restriction enzymes. Kusano *et al* 1995 proposed an alternative as to why so many bacteria possess RMS. They concluded that these systems are essential for bacterial survival due to the long half-life of the nuclease. If the host were to lose its RMS, the nuclease would still be present and therefore lead to cell death. Moreover, RS avoidance is much simpler and more effective for overall survival of the bacterium. Indeed, it has been shown that RS avoidance is not only preferable for bacteria but also strongly favoured by phages themselves. It appears that RMS presents a stronger selective load on bacteria than on their viral counterparts (Kusano *et al.*, 1995). Therefore, RMS constitutes a third factor within phage-host interactions.

1.11.4. Unusual, complex defence systems

An unusual bacteriophage resistance system is encoded by *Streptomyces coelicolor* A3(2), yet not its close relative *S.lividans (Laity et al., 1993)*. The phage growth limitation system was first described by Lomovskaya *et al* in 1982. The pgl system is as follows: following a single normal phage burst in a pgl, positive host progeny are produced, which upon a second cycle of infection are severely attenuated, resulting in the inability to cause subsequent infections within that host. However, they remain virulent to other hosts of a different strain (a pgl negative strain) (Figure 7) Although it was originally thought to be due to the presence of a defective prophage (ND Lomovkskaia, 1971) this was later disproved, as no ϕC31 DNA could be detected in *S.coelicolor* (Chater, 1986). This modification is clearly not lethal to the bacteriophage as it is capable of several rounds of infection in pgl negative hosts after infection of a pgl positive host. It seems likely that this feature is advantageous to the host as it would facilitate the destruction of competitor organisms. Cox (Cox and Baltz, 1984) suggested that the pgl system works in a similar way to the restriction modification system, however, some (Chinenova *et al.*, 1982) propose that the pgl system works due to a specific ϕC31 trigger.

Bedford *et al* (Bedford *et al.*, 1995) discovered 2 adjacent genes – *pglY* and *pglZ* which were required for pgl. These genes were studied using complementation experiments and two types of pgl negative mutants were isolated: A and B mutants. A mutants were complemented by a specific plasmid and B were unable to be complemented. More recently, Sumby *et al* identified two more genes which are involved in the pgl system: *pglW* and *pglX*. These genes, like the *pglY and* pglZ, were also studied using complementation (Sumby and Smith, 2002). Phase variation from pgl negative to pgl positive, and vice versa, is common at frequencies of 10^{-3} to 10^{-4} per spore. Sumby and Smith (Sumby and Smith, 2003), demonstrated that this is caused partly by the expansion and contraction of a polyguanine tract present in the *pglX* gene. This technology has recently been exploited in order to develop a phage capable of marker deletion in *S.coelicolor* by the Cre-*lox*P system (Khodakaramian *et al.*, 2006). However, further study is required before this complex system is fully understood.

Yet another recently described host defence mechanism are the rapidly evolving short transcribed nucleotide sequences that are inserted into clustered regularly interspaced short palindromic repeats or CRISPR resulting in phage resistance. It has been found that tandem repeats, separated by non-repetitive spacer sequences similar to phage DNA occur in half of all bacterial genomes (Godde and Bickerton, 2006; Jansen *et al.*, 2002). In conjunction with Cas proteins these CRISPRs confer viral resistance by targeting and neutralizing foreign DNA. Diversity in these regions is due to the changes in selective pressures of phage predation that in turn shapes the microbial ecology(Tyson and Banfield, 2008).

1.12. Phage defences against bacteria

Bacteriophages require a plethora of mechanisms in order to be successful in a wide range of hosts. Such defences include inhibitors of restriction enzymes, phage encoded methylation, nucleotide modification and complex modifications upon superinfection, the last of which has been most recently discovered.

Some bacteriophages use proteins that modify DNA resulting in the protection of its DNA from restriction. An example of this is found in the temperate phage Mx8 as it

contains mox protein. Mox appears to modify, by methylation, its restriction sites for XhoI and PstI (Magrini *et al.*, 1997). Another anti-restriction mechanism has been described in the coliphages T_7 and T_3. This mechanism is due to the gene *ocr* (overcome classical restriction), also known as, *0.3*. This gene encodes the protein that hydrolyses intracellular methyl donors, which are required for restriction of DNA.

The entire area of phage defence mechanisms has not been studied in detail (Walkinshaw *et al.*, 2002). However, it is evident that their continued survival gives an indication that mechanisms like these are common amongst the phage populations. This arms race between phage and host results in dynamic host-phage populations inhabiting a niche within the natural environment.

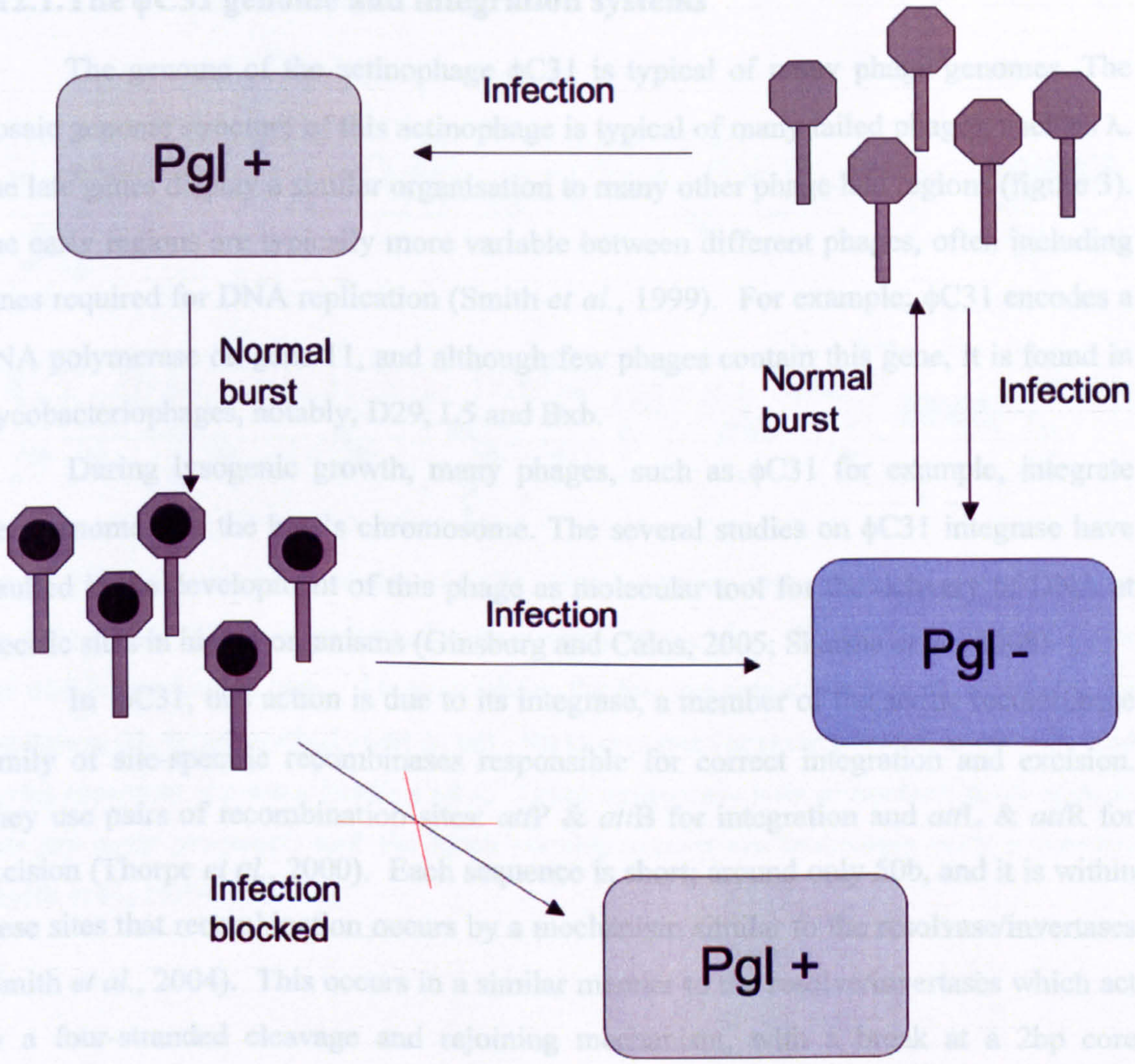


Figure 7 The Pgl system. Wild type phages (top right) are capable of infecting both pgl– and pgl+ hosts. Normal bursts result in both cases, however, the progeny from the pgl + host are modified, resulting in a loss in their ability to infect pgl + hosts. They remain infective to pgl – hosts (adapted from (Chinenova *et al.*, 1982) .

1.12.1.The ϕC31 genome and integration systems

The genome of the actinophage ϕC31 is typical of many phage genomes. The mosaic genome structure of this actinophage is typical of many tailed phages, such as λ. The late genes display a similar organisation to many other phage late regions (figure 3). The early regions are typically more variable between different phages, often including genes required for DNA replication (Smith *et al.*, 1999). For example; ϕC31 encodes a DNA polymerase on gene 11, and although few phages contain this gene, it is found in mycobacteriophages, notably, D29, L5 and Bxb.

During lysogenic growth, many phages, such as ϕC31 for example, integrate their genome into the host's chromosome. The several studies on ϕC31 integrase have resulted in the development of this phage as molecular tool for the delivery of DNA at specific sites in higher organisms (Ginsburg and Calos, 2005; Sharma *et al.*, 2008).

In ϕC31, this action is due to its integrase, a member of the serine recombinase family of site-specific recombinases responsible for correct integration and excision. They use pairs of recombination sites: *att*P & *att*B for integration and *att*L & *att*R for excision (Thorpe *et al.*, 2000). Each sequence is short; around only 50b, and it is within these sites that recombination occurs by a mechanism similar to the resolvase/invertases (Smith *et al.*, 2004). This occurs in a similar manner to the resolve/invertases which act by a four-stranded cleavage and rejoining mechanism, with a break at a 2bp core sequence. (Johnson and Bruist, 1989). It has been shown that mutations in this 2bp results in a switch of the polarity of the *att* sites resulting in incorrectly joined products (Gohosh *et al.*, 2003). These sites in *S.coelicolor* lie within the *Sco*3798 gene, a possible chromosome condensation protein. It has been shown that there is a slight detrimental effect on antibiotic biosynthesis once the phage has integrated (Gregory, 2003).

In addition, this mechanism is responsible for both phage and host evolution by the incorrect excision of the phage genome from the host chromosome. This is due to the polarity of the recombination sites *attB* and *attP*, which is dependant on the sequences of just two base pairs where the crossover occurs. A mutation within these

two nucleotides results in incorrectly joined products, resulting in the mechanism of evolution of the family of phage integrases (Smith *et al.*, 2004).

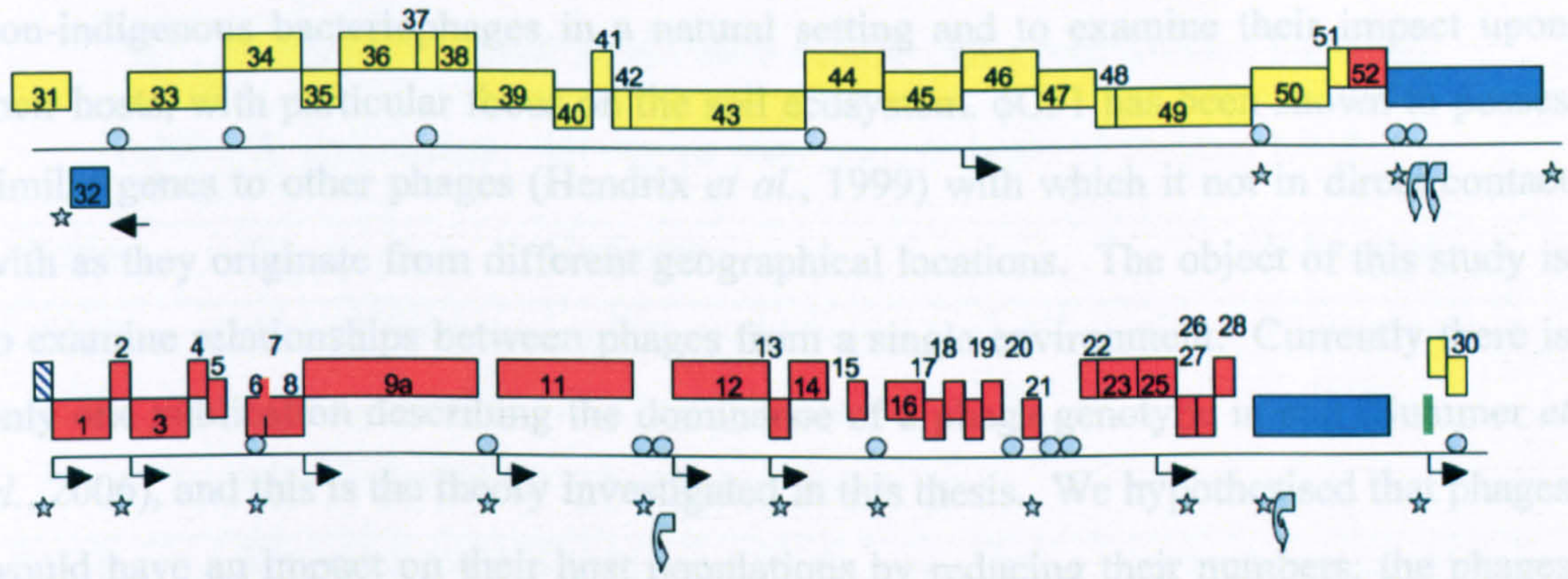


Figure 8 The genome organisation of ϕC31. The late genes are yellow and the early genes are red. Genes which are expressed only in lysogens are blue. The single tRNA is a green bar and is transcribed late. All genes above the line are transcribed left to right and the one below, 32, is transcribed right to left. Stars represent transcriptional terminators and circles represent repressor binding sites. Tailed arrows are lytic promoters, those without tails are early promoters and the block curving arrows are immediate early promoters (Smith *et al.*, 1999).

1.13. Thesis Objectives

The aim of this thesis was to decipher the relationships between indigenous and non-indigenous bacteriophages in a natural setting and to examine their impact upon their hosts, with particular focus on the soil ecosystem. ϕC31 has been shown to posses similar genes to other phages (Hendrix *et al.*, 1999) with which it not in direct contact with as they originate from different geographical locations. The object of this study is to examine relationships between phages from a single environment. Currently there is only one publication describing the dominance of a phage genotype in soil (Summer *et al.*, 2006), and this is the theory investigated in this thesis. We hypothesised that phages would have an impact on their host populations by reducing their numbers; the phages would share extensive homology and be unique to their location. In order to test this we examined hosts and phages from the same ecological context and compared their fitness in soil to introduced strains. The use of this phage throughout the study allowed for a clear comparison of indigenous soil hosts and phages with adaptable, yet not indigenous strains.

2. Materials and Methods

2.1. Bacterial Strains

A variety of *E.coli* strain were used to clone recombinant plasmids.

Table 2 *E. coli* strains. All strains were used as hosts of plasmid DNA.

Strain	Genotype	Reference
JM109 K12	endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB$^+$ Δ(lac-proAB) e14- [F' traD36 proAB$^+$ lacI^q lacZΔM15] hsdR17($r_K^- m_K^+$)	(Yanisci-Perron *et al.*, 1985)
DH5α	F$^-$ *end*A1 *gln*V44 *thi*-1 *rec*A1 *rel*A1 *gyr*A96 *deo*R *nup*G Φ80d*lacZ*ΔM15 Δ(*lac*ZYA- *arg*F)U169, *hsd*R17(r_K^- m_K^+), λ–	(Taylor *et al.*, 1993)
MC1061	*hsd*R2 *hsd*M+ *hsd*S+ *ara*D139 Δ(ara-leu)7697 Δ(lac)X74 *gal*E15 *gal*K16 rpsL (*Str*R) *mcr*A *mcr*B1	(Casadaban and Cohen, 1980)
TOP10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ* ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(ara-leu)7697 *gal*U *gal*K *rps*L (str^R) *end*A1 *nup*G	Invitrogen

Table 3 *S. aureus* strains. These were used in order to accept plasmid DNA or as bacteriophage hosts.

Strain	Genotype	Reference
RN4220	Mc^s; restriction-negative derivative of 8325-4	(Kreiswirth *et al.*, 1983)
8588	Mc^s	Janice Spencer (Uni of Strathclyde)

Table 4 Streptomycete strains. All used as bacteriophage hosts.

Strain	Genotype	Reference
S.lividans 1326	$SLP2^{+}$ $SLP3^{+}$, *pgl*$^{-}$	(Hopwood *et al.*, 1983)
S.lividans TK24	pro-2 str-6 $SLP2^{-}$ $SLP3^{-}$,*pgl*$^{-}$	(Hopwood *et al.*, 1983)
S.coelicolor A3(2) M145	SCP1-, SCP2-, *pgl*$^{+}$	(Hopwood *et al.*, 1969)
S.avermitilis	pgl-	(Ikeda *et al.*, 1987)
SP01	-	This study
SP02	-	This study
SP03	-	This study
SP04	-	This study
SP05	-	This study
SP06	-	This study
SP07	-	This study
SP08	-	This study
SP09	-	This study

2.2. Bacteriophages

Throughout this work several bacteriophages were isolated and comparisons draw between them and previously characterised bacteriophages (Table 5).

Table 5 Bacteriophages. Typed and isolated strains.

Bacteriophage	Reference
ϕC31	(Lomovskaya *et al.*, 1972)
ϕR4	(Chater and Carter, 1979)
ϕHau3AI	(Zhou *et al.*, 1994)
ϕELB1	This study
ϕELB2	This study
ϕELB3	This study
ϕELB4	This study
ϕELB5	This study
ϕELB6	This study
ϕELB7	This study
ϕELB8	This study
ϕELB9	This study
ϕELB17	This study
ϕELB18	This study
ϕELB19	This study
ϕELB20	This study
ϕ858 (Phage K)	(Rees and Fry, 1981)

2.3. Chemicals and Enzymes

Sigma Chemical Company Ltd supplied chemicals, unless otherwise stated. Bacterial media, including agar (Agar Technical No. 3), was purchased from Oxoid Ltd. Enzymes utilised were purchased from Promega, Invitrogen, New England Biolabs and

Invitrogen. Sequencing primers were obtained form MWG-Biotech A.G and dNTPs from Amersham Pharmacia Biotech Inc.

2.3.1. Antibiotics and reagents

All antibiotics used were stored at -20°C and thawed on ice. After use, they were immediately replaced at -20 °C. The antibiotics, indicators and inducers used are listed in Table 6.

Table 6 Reagents and antibiotics

Chemical	Stock concentration (mg/mL)	Usage concentration (μg/mL)	Dilulent
Ampicillin	100	100	Deionised H_2O
Apramycin	25	25	Deionised H_2O
Chloramphenicol	25	6-25	100% ethanol
Cyclohexamide	25	25	Deionised H_2O
Kanamycin	25	25	Deionised H_2O
Nalidixic acid	50	200	Deionised H_2O
Neomycin	1	1	Deionised H_2O
Rifampicin	2	10	100% Ethanol
Tetracycline	12.5	12.5	70% Ethanol
IPTG	20	50	Deionised H_2O
X-Gal	200	20	DMF
Lysozyme	50	1	Deionised H_2O
Lysostaphin	0.5	0.2	TE buffer

2.4. Vectors

Plasmids and vectors used are detailed in Table 7.

Table 7 Plasmids.

Vector	Structure	Reference
pUC19	*Bla, lacZ, rep* (pMB1) – 2686bp	(Norrander *et al.*, 1983)
pALTER1	Bla, lacZ, Tet - 5680bp	Promega
pBLT100	phage K ORF 36/37 Bla, lacZ, neo – 5157bp	This study
pBLT101	phage K ORF 38/39 Bla, lacZ, neo – 5772bp	This study
pBLT197	phage K ORF 36/37 tet, lac*Z* - 7673	This study
pBLT198	phage K ORF 36/37 *bla* – 4681bp	This study
pBLT200	phage K ORF 36/37 & 38/39 tet – 7081bp	This study
pBLT201	phage K ORF 36/37 & 38/39 *bla* – 6021bp	This study
pBLT203	phage K ORF 36/37 & 38/39 , *lux*AB, *tet*- 8716bp	This study
pU16AVE	Bla, lacZ, rep 3086bp	This study
pU16SO1	Bla, lacZ, rep	This study

	3086bp	
pU16SO3	Bla, lacZ, rep 3086bp	This study
pU16SO4	Bla, lacZ, rep 3086bp	This study
pU16SO5	Bla, lacZ, rep 3086bp	This study
pU16SO6	Bla, lacZ, rep 3086bp	This study
pU16SO7	Bla, lacZ, rep 3086bp	This study
pU16SO9	Bla, lacZ, rep 3086bp	This study
pVC119	*lux*AB, apra -7854bp	PR Herron unpublished
pNZ123	*Cm* – 2808bp	(de Vos, 1987)
pIJ925	*Bla*, - 2715bp	PR Herron
pSK5630	Bla, par, cm, - 5920bp	(Grkovic, 2003)
pSK5632	Bla, par, cm, lacZ, - 5912bp	(Grkovic, 2003)
pCR®4Blunt-TOPO	bla, lacZ, neo – 3957bp	Invitrogen

2.5. Media and solution preparation

Solutions and media were prepared using deionised water, unless otherwise stated. The pH of solutions and media were measured using a Thermo Orion (model 410) pH meter with the sensor stored in 3M KCl.

2.5.1. Sterilisation of media and solutions

Antibiotics and amino acid solutions were sterilised by passing through a 0.22μM filter unit (Millex®-HV). All media and buffers were autoclaved for 15 minutes at 121°C to achieve sterilisation, unless otherwise stated.

2.6. Media for the growth of streptomycete strains

2.6.1. Mannitol Soya Flour medium (Hobbs *et al.*, 1989)

20g of mannitol was made up to 1 litre of tap water. 20g/L of soya flour and 20g/L of agar technical No. 3 was added to 250mL bottles. This media was prepared using tap water and autoclaved at 121 °C for 30 minutes. Following sterilisation $MgCl_2$ was added to a concentration of 10mM (Hobbs *et al.*, 1989).

2.6.2. RASS media (Herron and Wellington, 1990).

15g/L of agar technical No. 3 was added to bottles with 12.5g/L soluble starch. 1g/L NaCl, 1g/L K_2HPO_4, 0.5g/L $MgSO_4.7H_2O$, 0.1g/L L-arginine was added. Following sterilisation 1ml (of 1g/L) of each of the following was added to molten media: $Fe(SO_4).6H_2O$, $Cu(SO_4).7H_2O$, $ZnSO_4.7H_2O$ and $MnSO_4.4H_2O$.

2.6.3. Yeast-extract malt media (Kieser *et al.*, 1985)

This media consisted of 3g/L yeast extract, 5g/L peptone, 3g/L malt extract, 10g/L glucose and 340g/L of sucrose. After autoclaving 2ml of 2.5M $MgCl_2.6H_2O$ (5mM final) was added to a litre using aseptic technique.

2.7. Media used for the growth of *E.coli*

2.7.1. Luria-Bertani Agar media (Luria and Burrous, 1957)

Constituents were: 10g/L of tryptone, 5g/L of yeast extract, 5g/L NaCl and 15g/L of agar technical No.3.

2.7.2. Lennox Broth (Lennox, 1955)

Constituents were: 10g/L of tryptone, 5g/L of yeast extract and 5g/L NaCl

2.7.3. SOC media (Hanahan, 1983)

20g/L tryptone. 5g/L yeast extract, 0.5g/L NaCl. Following sterilisation to final concentrations are 10mM of 1M $MgSO_4$, 2.5mM of 1M KCl and 20mM of 1M glucose were added.

2.8. Media used for the growth of *S.aureus*

2.8.1. B2 Broth (Lofblom *et al.*, 2006)

This media consisted of 25g/L of yeast extract, 25g/L NaCl, 10g/L casein hydrolysate. Following sterilisation there were additions of 5 g/L of glucose and 1g/L of K_2HPO_4.

2.8.2. NYE Agar (Schenk and Laddaga, 1992)

10g/L casein hydrolysate, 15g/L agar technical no. 3, 5g/L yeast extract and 5g/L NaCl.

2.8.3. Preservation of *E.coli* and *S.aureus*

For the intention of long term storage 500µL of an overnight culture was added to a cyrogenic storage tube containing an equal volume of 50% sterile glycerol and was stored at -70°C.

2.9. Media used for the propagation of bacteriophages

2.9.1. Nutrient Agar (Association, 1917)

Components were 13g/L nutrient broth and 15g/L agar technical no.3.

2.9.2. Soft Nutrient Agar

Components were 8g/L nutrient broth and 5g/L agar technical no.3.

2.10. Media used for the storage, extraction and dilution of bacteriophages

2.10.1.Nutrient Broth (Association, 1917)

This media consists of 13g/L of nutrient broth. When required, 0.1% (v/v) egg albumin was added after sterilisation.

Table 8 General use buffers, reagents and solutions (Kieser *et al.*, 1985)

Reagent	Constituents	Final volume (using deionised H_2O)
SDS mix	10% (w/v) SDS 2M Tris-HCl 0.5M Sodium EDTA (pH 7.4)	1mL
SM Buffer	1M Tris-HCl 1M MgSO4 5M NaCl 1g Gelatine	1 L
20 x SSC	175.3g NaCl 88.23g Sodium citrate (pH 7.2)	1L
¼ Strength Ringers solution	2.25g NaCl 0.105g KCl 0.12g CaCl2 0.05g NaHCO3	1L
Precipitation solution	33% PEG 3.3M NaCl	100mL
RNase solution	SM Buffer containing 40µg/mL RNase	50mL

2.11. Growth conditions and maintenance of streptomycetes

2.11.1. Growth of streptomycetes s mycelium in liquid media

Typically, a 500µl of a spore suspension was used to inoculate YEME media (2.6.3) and grown at 30°C on an orbital shaker at 200rpm for 24 hours.

2.11.2. Preparation of spores

In order to prepare spore suspensions required for plaque assays and growth curves, MS media was inoculated with 100µl of a spore suspension or streaked from a single colony to create a lawn. It was then incubated at 30°C for, typically, 4-7 days. After this, the spores were harvested by pouring 10ml sterile, deionised H_2O on the plate and, using a sterile loop, the surface growth scraped off into the water and poured into a sterile universal tube. This was then agitated using a vortex for 2 minutes in order to break spore chains. A sterile syringe, containing non-absorbent, cotton wool, was then used to filter out large clumps of spores and agar. The solution was centifuged at 10000g for 10 minutes to achieve a spore pellet which was resuspended in 20% (w/v) glycerol and stored at -20°C.

2.11.3. Spore counts

Viable spore concentrations were determined by colony counts on MS agar (2.6.1) or RASS (2.6.2) agar. Serial dilutions, typically 10^{-1} to 10^{-5} were made using SDW as the diluent. Colonies were counted after 5 days of incubation at 30°C.

2.12. Propagation, storage and assays of bacteriophages

2.12.1. Plaque assays (Dowding, 1973)

All phages used in this study were propagated on nutrient agar with a host and appropriate supplements (see 2.9). Nutrient broth or SM buffer was used as a dilulent for phages. All phages were stored at 4°C.

Table 9 Bacteriophages, their hosts and media supplements required for their propagation (AT Cater, unpublished).

Bacteriophage	**Host for propagation of high titre lystates**	**1M $MgCl_2$ (per 250mL agar)**	**0.8M $Ca(NO_3)_2$ (per 250mL agar)**	**50% Glucose (per 250mL agar)**	**Reference**
ΦC31	*S.lividans* 1326	2.5	2.5	2.5	(Lomovskaya *et al.*, 1972)
ΦR4	*S.lividans* 1326	-	1	2.5	(Chater and Carter, 1979)
ΦHau3A	*S.coelicolor* A3(2)	-	1	2.5	(Dyson and Schrempf, 1987)
ΦELB1	SPO1	-	1	2.5	This study
ΦELB2	SPO2	-	1	2.5	This study
ΦELB3	SPO3	-	1	2.5	This study
ΦELB4	SPO4	-	1	2.5	This study
ΦELB5	SPO5	-	1	2.5	This study
ΦELB6	SPO6	-	1	2.5	This study
ΦELB7	SPO7	-	1	2.5	This study
ΦELB8	SPO8	-	1	2.5	This study
ΦELB9	SPO9	-	1	2.5	This study
ΦELB17	*S.lividans* 1326	-	1	2.5	This study
ΦELB18	*S.lividans* 1326	-	1	2.5	This study
ΦELB19	*S.lividans* 1326	-	1	2.5	This study
ΦELB120	*S.lividans* 1326	-	1	2.5	This study
Φ858 (K)	*S.aureus* 8588	-	-	-	(Rees and Fry, 1981)

Nutrient agar plates were inoculated with serially diluted bacteriophages that were suspended in nutrient broth or SM buffer. Soft nutrient agar was allowed to cool to 50°C and approximately 1 x 10^8 host spores were added to each 200ml SNA. Soft nutrient agar containing host spores was poured on the nutrient agar plate (2.5ml per 9cm plate), which was previously inoculated with phages. This was then allowed to set and incubated at 30°C overnight.

2.12.2.Sterilisation and storage of phage soak outs

0.45μM filter units were used to remove host and contaminating bacteria from phage lysates. These were supplied by Millex®-HV. Phages were stored at 4°C.

2.12.3. Single Plaque Isolation of Phages

A well-isolated plaque was identified and using a sterile pipette tip a core was removed from the centre of the plaque. This was left to soak out in 1mL NB at room temperature for 2 hours. The soak out was filter sterilised and stored at 4°C.

2.12.4. Creating high titre lysates of phages and their storage

Plates with almost confluent lysis (approximately 1 x 10^4 plaques per plate) were used to make high titre lysates. Each 9cm plate was covered in 2.5ml nutrient broth. This was left to soak the phages out for 2 hours at room temperature. The soak outs were decanted and filtered as described in section 2.12.2.

2.12.5. Caesium chloride gradient isolation of phages

Six large plates (24cm^2 each) with almost confluent lysis were used for a single caesium chloride gradient. The top agar layers were scraped off using a sterile spreader and left to soak out by submerging in 180ml nutrient broth for 2 hours at room temperature. In order to remove agar, it was spun at 16000g for 15 minutes. The supernatant was decanted and centrifuged at 34000g for 75 minutes at 4°C. The resulting pellet was resuspended in a total volume of 4ml SM buffer and incubated with shaking at 37°C for 15 minutes. A further 1ml of buffer was added to give a final volume of 5ml. 4.25g Caesium chloride was added to the solution and allowed to

dissolve. This was then transferred into a Beckman L8 Ultra clear 5mL centrifuge tube and centrifuged at 85000g for 20 hours at 4°C in a Beckman Ultracentrifuge. The resulting phage band was extracted from the centrifuge tube by piercing with a sterile syringe needle. This solution was transferred to a dialysis bag. Dialysis was carried out in 200ml SM buffer which was changed every hour for 4 hours at 4°C. The phage suspension was then stored at 4°C or used immediately for DNA extraction.

2.12.6. One-step growth curves

Growth curves of phages were carried out over a time period of 6 hours with regular readings, in triplicate, every 30 minutes. A host spore suspension was used to inoculate 25mL YEME broth (see Yeast-extract malt media) in a 250mL flask, and incubated, with shaking. Bacteriophages were added to the culture at an MOI of 0.1. At 30-minute intervals 1mL of culture was removed and filter sterilised (§2.12.2). This was then assayed for plaques (seesee section2.12.1) at appropriate dilutions, in triplicate.

2.13. Soil experiments

2.13.1. Soil source

Soil was sourced from a residential area in Erskine, Renfrewshire, United Kingdom.

2.13.2. Soil preparation

The soil was stored in a cool, dark environment prior to use. After which it was passed through a sieve in order to remove large items. 1% (w/w) crab-shell chitin and 1% (w/w) starch was added to the soil. It was stored as individual 10g samples at room temperature.

2.13.3.Measuring pH of soil

Soil pH was measured with a glass electrode using a 1:2 soil-to-water ratio.

2.14. Extraction of bacteriophages from soil

2.14.1. Direct count method

9mL nutrient broth, containing 0.1% (v/v) egg albumin, was mixed with 1g soil and vortexed for 5 minutes. The solution was passed through a 0.45µM filter unit and plated in appropriate dilution factors with a chosen host (Lanning and Williams, 1982).

2.14.2. Enrichment method

As in the direct count method (see section 2.14.1), 9mL nutrient broth, containing 0.1% (v/v) egg albumin solution was mixed with 1g soil and mixed for 5 minutes. Spores of a decided host were added to this at a concentration of 1×10^5 spores per ml and left at room temperature overnight. The solution was filtered using a 0.45µM filter unit prior to assay (Lanning and Williams, 1982).

2.14.3. Extraction of actinomycetes from soil

Soil (1g) was mixed with 1/4 strength ringers solution (9mL) and agitated using a vortex for 3 minutes. This was plated out at appropriate dilutions onto RASS agar containing rifampicin (Herron and Wellington, 1990). Colony morphology and pigmentation production were used to as an indication of different strains.

2.15. Soil microcosm experiments

2.15.1.Soil sterilisation

Soil was sterilised at 121°C for 15 minutes. This was left overnight at room temperature and autoclaved again the following day.

2.15.2. Extraction of host bacteria from soil

Samples (1g) were taken from a single batch of soil, in triplicate, on each day of sampling. Quarter strength ringers solution (9mL) was added and mixed for 5 minutes by vortexing. Appropriate dilutions were made and plated onto RASS agar. This was incubated for 5 days at 30°C and colony forming units were counted.

2.16. DNA manipulation techniques

Most methods below are from Sambrook et al (1989), unless otherwise noted.

2.16.1. Commonly used solutions and buffers for molecular work

Table 10 Common molecular biology solutions and buffers

Solution	Components and concentrations
10 x TE Buffer	100mM Tris HCl (pH 8) 1mM EDTA
50 x TAE	302.4g Tris 57.1mL glacial acetic acid 100mL 0.5M EDTA (pH 8)
10 x Loading buffer	0.5% (w/v) bromophenol blue 0.05% (w/v) xylene cyanol 50% (w/v) ficoll 1% (w/v) SDS 100mM EDTA
Cell resuspension solution*	50mM glucose 25mM Tris-HCl (pH 8) 10mM EDTA
Cell lysis solution**	500µl 4M NaOH 500µl 20% SDS 9mL sterile deionised H_2O
Neutralisation solution	60mL 5M potassium acetate 11.5mL glacial acetic acid 28.5mL sterile deionised H_2O

*Stored at 4°C (Birnboim and Doly, 1979)

**Freshly prepared immediately before use (Birnboim and Doly, 1979)

All solutions were sterilized as described in see section 2.5.1.

2.17. Isolation of DNA

2.17.1. Plasmid isolation (Birnboim and Doly, 1979)

1.5mL of a 5mL overnight culture of *E.coli* was microcentrifuged at 16,000g for 1 minute to pellet cells. The supernatant was discarded and 100µl of ice-cold cell resuspension solution was added and mixed. 200µl of freshly prepared alkaline lysis solution was added and the tube inverted 4 times. This was incubated at room temperature for 3 minutes or until the lysate cleared. After the addition of 150µl of ice cold Neutralisation solution, the tube was inverted 5 times. This was centrifuged for 10 minutes at 16,000g following which the supernatant was subject to phenol:choloroform treatment as described insee section2.18.4. The aqueous phase was removed and 2.5 x volumes of ethanol added. This was incubated at room temperature for 3 minutes and then microcentrifuged at 16,000g for 10 minutes. The resulting DNA pellet was dried as insee section2.18.5. The DNA was dissolved in 50µl 1 x TE buffer containing 20µg/mL pancreatic RNase. The sample was incubated at 37°C for 20 minutes. DNA was stored at -20°C.

2.17.2. Isolation of *E.coli* plasmid DNA using a kit

Promega Wizard Plus SV Minipreps DNA Purification System was used to isolate plasmid DNA intended for cloning and sequencing. The protocol was followed as the manual instructs and the DNA stored at -20°C.

2.17.3. Isolation of plasmid DNA from *S.aureus* (Vriesema *et al.*, 1996)

10mL cultures of *S.aureus* in LB with appropriate antibiotics were grown overnight and centrifuged to obtain a bacterial pellet. This was resuspended in 376µl 50mM EDTA (pH 8). Next, 20µl of 50mg/mL lysozyme was added along with 20µL 0.5mg/mL lysostaphin. This was incubated at 37°C, with frequent mixing, for 1 hour. Plasmid DNA was then extracted according to the alkaline plasmid isolation as described in section 2.17.1.

2.18. Bacteriophage DNA Isolation

2.18.1. Small scale DNA extraction of phage DNA (DA Hopwood, 2000) (Kieser *et al.*, 1985)

Top layers of 6 plates with approximately 2 x 10^4 plaques were scrapped off and allowed to soak in 25mL NB for 2 hours at room temperature. The agar was removed by centrifugation at 2500g for 10 minutes. From the supernatant, sedimentation of the phage was achieved by centrifugation at 55,000g for 75 minutes. The resulting phage pellet was resuspended in 2mL RNase solution and incubated at 37°C, with shaking, for 20 minutes. 400µl SDS mix was added and incubated at 70°C for 30 minutes. Next, 500µl 8M potassium acetate was added and left on ice for 15 minutes. This was centrifuged at 2°C for 30 minutes at 30,000g. The supernatant was added to 7mL TE buffer. Following this, it was exposed to ethanol precipitation (see section 2.18.5) and resuspended in 100µl water.

2.18.2. Small-scale extraction of phage DNA with PEG precipitation

This method was adapted from previous work using PEG precipitation (Lis and Schleif, 1975). From three 9cm plates, with approximately 1 x 10^4 plaques on each, the soft top was scraped off into a total of 12mL NB. This was incubated for 2 hours at room temperature. As before, agar was removed by centrifugation and the supernatant collected. The supernatant was treated with RNase and DNase, both to a final concentration of 1µg/mL. This was incubated at 37°C for 30 minutes. 4mL of precipitation solution was added and incubated on ice for 1 hour. Then, supernatant was removed by centrifugation at 7500g for 10 minutes and then cleaned with phenol-choloroform treatment (see section 2.18.4) and subsequent ethanol precipitation (see 2.18.5). The final pellet was resuspended in 100µl TE buffer.

2.18.3. Large scale DNA extraction of phage DNA

Following caesium chloride gradient extraction (see section 2.12.5) the high-density phage solution was treated with 1 X Sodium chloride and Sodium citrate

solution (SSC). It was then incubated at 60°C for 10 minutes. NaCl was added to give a final concentration of 0.25M. The DNA was purified by phenol-choloform treatment (section 2.18.4) and precipitated as stated in see section 2.18.5.

2.18.4. Phenol chloroform extraction of DNA (Sambrook *et al.*, 2001)

Phenol chloroform is used in order to remove contaminating proteins from DNA. Typically, an equal volume of phenol:cholorform was added to a sample and mixed vigorously for 3 minutes. This was then centrifuged at 16,000g for 10 minutes to separate phases. The upper, aqueous, layer was carefully removed by avoiding the interface between layers.

2.18.5. Ethanol precipitation of DNA (Sambrook *et al.*, 2001)

In order to precipitate DNA, ethanol was used. This involved adding 2.5 times of the volume of ethanol to the DNA and, typically, placing at -20 °C for 2 hours or overnight. Centrifugation at 16,000g for 30 minutes resulted in a DNA pellet. This was washed with 70% ethanol and spun as before to form the DNA pellet that was dried in a speed vac for 20 minutes. This was then resuspended in either deionised H_2O or TE buffer.

2.18.6. DNA pellet drying

In order to dry a DNA pellet a Speed Vac dryer was used according to the manufactures instructions. The pellet was resuspended in an appropriate amount of TE buffer or sterile, deionised H_2O.

2.18.7. Digestion of DNA with restriction enzymes (Sambrook *et al.*, 2001)

Plasmid and phage DNA were digested with several different enzymes, all of which were Type II endonucleases. A typical reaction consisted of 150ng – 0.5µg DNA, 1 x buffer and 10-units/µL restriction enzyme. This reaction was made up to 20µl using sterile, deionised H_2O. In some cases, 2 restriction enzymes were used in one reaction, for this the total volume was increased in order to maintain the total glycerol

concentration below 5%. The reactions were incubated at 37°C for 2 hours, unless otherwise stated.

2.18.8. Ligation of DNA fragments (Sambrook *et al.*, 2001)

T4 DNA ligase was used to anneal compatible fragments together, therefore creating recombinant plasmids. Typically, 20ng linear vector was combined with insert DNA at an insert: vector ratio of 3:1 when ligating cohesive ends and at 1:1 when cloning blunt ends. The following equation was applied in order to determine appropriate concentrations.

$$ng\ insert\ DNA = \frac{ng\ vector \times kb\ size\ of\ insert}{kb\ size\ of\ vector} \times molar\ ratio\ of\ \frac{insert}{vector}$$

Ligation reactions consisted of insert and vector DNA at appropriate concentrations, 1 x T_4 DNA ligation buffer and 0.5units/μL T4 DNA ligase and made up to 15μL using sterile, deionised water. This was incubated at 16°C overnight. Controls consisted of digested, linear, vectors.

2.18.9. Dephosphorylation of vector DNA

Typically, the vector had its 5' end phosphates removed during a ligation reaction in order to prevent re-ligation without an insert. For this, an enzyme; Thermosensitive Alkaline Phosphatase, supplied by Promega, was used according to manufacturers instructions.

2.19. Competent cell propagation and transformation procedures

2.19.1. Production of electrocompetent *E.coli* cells (Sambrook *et al.*, 2001)

A 10mL culture of *E.coli* was inoculated from a single colony and grown overnight at 37°C in L-broth with shaking at 250 rpm. The following day, 5mL of this culture was added to 500mL L-broth and grown until it reached an OD_{600} of 0.6. The cells were chilled on ice for 20 minutes, after which they were centrifuged at 10000g for 15 minutes. The pellet was resuspended in 500mL ice cold 10% (w/v) glycerol and centrifuged as before. This was repeated using 250mL 10% (w/v) glycerol and then

resuspended in 20mL 10% (w/v) glycerol. Following centrifugation at 10000g for 15 minutes the cells were finally suspended in 2mL 10% (w/v) glycerol. This was then transferred in 50µl samples to microcentrifuge tubes and frozen quickly using an ethanol/dry ice bath. Cells were stored at -70°C until required.

2.19.2. Production of chemically competent *E.coli* cells (Sambrook *et al.*, 2001)

1mL of an overnight cultured was diluted in 100mL LB broth and incubated with shaking to an OD_{600} of 0.6. The culture was centrifuged and the pellet resuspended in 20mL ice-cold $MgCl_2$. After incubation on ice for 5 minutes the cells were harvested and suspended in 20mL ice-cold 100mM $CaCl_2$. This was incubated fro 30 minutes on ice. Cells were pelleted and resuspended in 2mL ice-cold 100mM $CaCl_2$ containing 20% (w/v) glycerol. These were then aliquoted in 100µl samples and quick frozen using liquid nitrogen or in an ethanol-dry ice bath. They were stored at -70°C.

2.19.3. Production of electrocompetent *S.aureus* cells (adapted from (Sambrook *et al.*, 2001))

A single colony of *S.aureus* was isolated and inoculated into 10mL of B2 broth. This was incubated, with shaking, at 37°C overnight. The culture was diluted 1/25 and grown until an OD_{600} of 0.6. This culture was pelleted by centrifugation at 10000g for 15 minutes and washed with sterile deionised water 3 times. It was finally washed with 5mL 10% (w/v) glycerol and then with 2.5mL 10% (w/v) glycerol before being resuspended in 800µl 10% (w/v) glycerol. This was divided up into eight 100µl samples and frozen using liquid nitrogen or an ethanol/dry ice bath. If not used immediately, cells were stored at -70°C for no more than 1 month.

2.19.4. Chemical transformation of *E.coli* (Sambrook *et al.*, 2001)

Typically, no more than 200ng of plasmid DNA was required for successful transformation. DNA was added to competent cells and left on ice for 5 minutes. The sample then underwent heat-shock at 42°C for 30 seconds and immediately placed on

ice for 5 minutes. 1mL SOC was added and incubated at 37°C for 90 minutes. This was then plated onto LB-agar containing appropriate antibiotics and, if necessary, substrate or x-gal.

2.19.5. Electrotransformation of *E.coli* and *S.aureus* cells

An appropriate amount of DNA (see Table 11) was added directly to electrocompentent cells, which had thawed on ice. This was incubated, on ice for 3 minutes. The cells and DNA were added to electrocuvettes (see Table 11 for sizes) and pulsed with an electric current. Immediately, 1mL broth was added to the cuvette and mixed with the cells. This was incubated for 90 minutes at 37°C with shaking. The cells were then plated onto agar containing the appropriate antibiotics and supplements and incubated overnight at 37°C.

Table 11 Transformation of *E. coli* and *S. aureus*

Feature	***E.coli*** **(Sambrook *et al.*, 2001)**	***S.aureus*** **(Vriesema *et al.*, 1996)**
DNA concentration added (ng)	100	1000
Broth	SOC	B2
Cuvette size (mm)	0.1	0.2
kV cm^{-1}	1.8	2.3
Time constant (msec)	1	2.5
Agar used	L-agar	NYE

2.19.6.Calculating transformation efficacy

The efficacy of all competent cells was calculated by transforming a known concentration of plasmid DNA (typically, 1ng pUC19 was used) according to the following calculation:

$$number\ of\ colonies \times dilution\ factor \times \frac{1}{amount\ of\ DNA\ used(\mu g)}$$

If the efficiency was less than 1 x 10^{-6} the cells were discarded.

2.19.7.Use of Blue/White screening for identifying clones (Sambrook *et al.*, 2001)

X-Gal was used with IPTG in order to identify colonies containing desired recombinant clones within multiple cloning sites of vectors carrying the *lacZ* gene. Blue colonies were formed due to the activity of ß-galactosidase. Colonies containing recombinants generally yielded white colonies.

2.20. DNA visualisation

2.20.1. Preparation, staining and running of agarose gels

DNA was visualised by size fractionation on horizontal, neutral gels at an agarose concentration, typically, of 1% (however, this varied from 0.7% to 2%). Agarose was dissolved in 1 x TAE buffer and poured into a cast (size dependent on number of samples being ran). Immediately prior to pouring into the case, ethidium bromide was added to give a final concentration of 0.5µg/mL and mixed gently. Small gels were run in 1 x TAE at 70 volts for 1.5 hours and larger gels at 100 volts for 3 hours. Where good separation was required large gels were electrophoresed overnight at 30 volts.

2.20.2. Photography of gels

A Syngene 302nm ultra violet transilluminator was used to visualise DNA in the agarose gels with the aid of GeneSnap (Syngene version 6.08).

2.20.3. DNA band sizing and quantification

DNA was quantified by signal intensity visualised on a gel by comparison to known concentrations of DNA within the ladder. Typically, 100ng/µl of λ Hind III DNA (provided by Promega) was used as a ladder on all agarose gels. The band sizes and concentrations are described in the table below:

Table 12 λ Hind III ladder

DNA Band from top	Size (base pairs)	DNA concentration (ng/5µl)
1st	23130 (cos ends)	240
2nd	9416	97
3rd	6557	68
4th	4361 (cos ends)	44
5th	2322	24
6th	2027	21
7th	564	6
8th	125	-

Other ladders used include:

Promega 1kb standard: Fragemnt sizes were 10Kb, 8Kb, 6Kb, 5Kb, 4Kb, 3Kb, 2.5Kb, 2Kb, 1.5Kb, 1Kb, 750bp, 500bp and 250bp.

Promega 100bp standard: Fragment sizes were 1.5Kb, 1Kb, 900bp, 800bp, 700bp, 600bp, 500bp, 400bp, 300bp, 200bp and 100bp.

2.20.4. DNA recovery from agarose gels

DNA fragments were isolated from a 0.07% agarose gel using a QIAquick Gel Extraction Kit® from Qiagen, following the manufacturer's instructions.

2.21. Polymerase Chain Reaction

Polymerase chain reaction was utilised in order to amplify DNA regions of interest using specifically designed primers.

2.21.1. Design of PCR primers

All primers used in this project are listed in Table 13. Primers were typically deigned to be between 15 and 25 nucleotides of length. Where possible, the initial and last nucleotide of the primer was either a guanine or cytosine in order to assure

hybridisation to the template sequence. Other than primers used for sequencing, primers used were obtained from MWG-Biotech A.G and arrived as a dry pellet. This was dissolved in nuclease free sterile deionised H_2O to a final concentration of 100pmol/µL and stored at -20°C.

2.21.2. PCR

Fermentas supplied the T_4 DNA ligase *Taq* polymerase used for reactions. PCR mixtures consisted of 25ng/µL template DNA, 1 x Polymerase buffer, 1 unit/µL *taq* polymerase, 2mM $MgCl_2$ (unless otherwise stated), 1.25mM dNTPs (1.25mM of dATP, dTTP, dGTP and dCTP), 2 pmol/µL forward and reverse primers and made up to 50µl with sterile, nuclease free deionised H_2O.

Using an Eppendorf Mastercycler gradient PCR machine the samples underwent a denaturation step at 94°C for 5 minutes, after which, 1U/µL of *Taq* polymerase enzyme was added. Amplification involved, typically, 30 cycles involving 94°C for 1 minute, followed by 30 second annealing step at the melting temperature specific for the primers being used. The last part of the cycle was the elongation step at 72°C for 1 minute per kilobase of template DNA. The final stage was the polishing step of 72°C for 10 minutes. In order to assess the PCR 10% of the final reaction was elecrophoreised on an agarose gel as insee section 2.20.1

Table 13 PCR Primers

Primer name	**Sequence (5' to 3')**	**Use**	**T_m**
EP01	AAT TTG GGA AAT ATC TAA GTA C	ORF 36 within phage K genome	50.9°C
EP02	CTA TAG TGC TAA ATA AAA TCA C	ORF 36 within phage K genome	50.9°C
EP03	TCA TCT ATT CTA TTG TAT TTA TTC	ORF 37 within phage K genome	50.8°C
EP04	TTA TAA TAT TCA TTC CTA ACA	ORF 37 within phage	50.8°C

	AAG	K genome	
M13 FOR	GTA AAA CGA CGG CCA G	Sequencing	55°C
M13 REV	CAG GAA ACA GCT ATG AC	Sequencing	55°C
F984-GC	GC-clamp ACGGGGGGAACGCGAAGAACCTTAC	16S rRNA sequencing	52°C
R1378	CGGTGTGTACAAGGCCCGGGAACG	16S rRNA sequencing	52°C

2.22. Blotting and labelling of DNA

2.22.1.Transfer of DNA from an agarose gel to Hybond-N^{TM} membrane

Southern analysis was carried out according to the method developed by (Southern, 1975), with adaptations. In order to examine DNA by southern blotting, transfer of the DNA from gel to membrane was carried out with the aid of a VacuGene™XL blotting system, according to manufacturers' guidelines.

Table 14 DNA transfer solutions (Southern, 1975)

Solution	Components
Alkali transfer buffer	0.25M NaOH 1.5M NaCl
Denaturation solution	1.5M NaCl 0.5M NaOH
Depurination solution	0.25M HCl
Neutralisation solution	1M NH_4 Acetate 0.02M NaOH
20 x SSC	3M NaCl 0.3M tri-sodium citrate

Following electrophoresis and photography, DNA fragments were transferred using alkali transfer buffer onto a Hybond-N^{TM} membrane. The gel was washed in

deionised H_2O before being immersed in depunaturation solution for 15 minutes twice. Afterwards, it was rinsed with deionised H_2O. The gel was then immersed in denaturation solution and left for 15 minutes, this was then repeated, as before, and rinsed with deionised H_2O. Finally, the gel was covered in neutralization solution for 30 minutes. 20 x SSC was used to cover the gel and left for 1 hour to allow efficient blotting. Following this, the membrane was baked at 80°C for 2 hours.

2.22.2.Random primed DIG-DNA labelling

Typically, 1µg DNA was used as a template. Following DNA denaturation, at 95°C for 10 minutes, 2µL of hexanucleotide mix, 2µL dNTP labelling mix (containing 1mmol/L dATP, 1mmol/L dCTP, 1mmol/L dGTP, 0.65mmol/L dTTP and 0.35mmol/L DIG-dUTP) was added to a total volume of 20µL using sterile, deionised H_2O. 2U of Klenow enzyme was added to this and the reaction was incubated at 37°C overnight. Immediately before use the probe was boiled at 95°C for 15 minutes.

2.23. Hybridisation of DNA

2.23.1. Solutions required for hybridisation

Table 15 Hybridisation solutions (Southern, 1975)

Solution	Components per litre
Maleic acid buffer	22.2g maleic acid 17.6g NaCl pH 7.5 (NaOH)
Blocking solution	10% (w/v) in maleic acid buffer
hybridisation solution	1.25mL 20 x SSC 0.1% (w/v) N-laurylsarcosine 0.02% (w/v) SDS 1%(v/v) Blocking solution
Stringency buffer A	2 X SSC

	0.1% SDS
Stringency buffer B	0.2% SSC 0.1% SDS
Washing buffer	Maleic acid buffer 0.3% (v/v) Tween 20
Antibody solution	Anti-DIG-AP (1 in 5000 dilution) in 1% blocking solution

A prepared membrane was wrapped in mesh and rolled. This was placed in a hybridisation tube. 20mL of hybridisation buffer was added to the tube. This was incubated, with rotation, at 65°C for 1 hour. The hybridisation solution was decanted and replaced with a fresh 20mL of the solution, containing the denatured probe (seesee section2.22.2). This was incubated overnight at 65°C with continuous rotation. The following morning, the solution was decanted and 35mL stringency buffer A was added. two washes were carried out and incubated at 65°C for 15 minutes. The solution was removed and replaced with 35mL stringency wash B and, as before, used to wash the membrane. This step was then repeated. The membrane was then ready for detection of hybridised DNA (see section 2.24).

2.24. Detection of hybridised DNA

The membrane, within a hybridisation tube, was washed with 35mL washing buffer, with rotation, at room temperature for 5 minutes. Following removal of the buffer, 35mL of blocking solution was added and incubated, with agitation, for 30 minutes. The blocking solution was removed and antibody solution added for 30 minutes. In order to remove unbound antibodies, 35mL washing buffer was added for 15 minutes. This was repeated once more before incubation in 20mL detection buffer for 2 minutes. The membrane was developed using Roche NBT tablets according to manufacturers instructions.

2.25. Bacteriophage plaque hybridisation

The detection of plaques was carried out using plaque hybridisation as developed by(Sambrook *et al.*, 2001). The protocol used adapted and taken from was derived from (Kieser *et al.*, 1985).

Table 16 Plaque hybridisation solutions (Kieser *et al.*, 1985)

Solution	Components
Alkali transfer buffer	0.5M NaOH
Denaturation solution	1M Tris HCl
Depurination solution	1M Tris HCl 1.5M NaCl

Large plates with almost confluent lysis plaques were blotted with Hybond-N™ for 15 minutes. The membrane was placed onto 3 layers of filter paper, which had been soaked in alkali transfer buffer. This was left for 10 minutes. The membrane was then transferred onto freshly soaked filter paper with 1M Tris HCl (pH 7.5) and left for 2 minutes. This step was repeated. Finally, the membrane was exposed to the third solution for 5 minutes. Afterwards the membrane was air dried for 30 minutes before being baked at 80°C for 2 hours. Following this the standard hybridisation and detection protocol was followed (2.23).

2.25.1. Genome library formation

Plasmid genome libraries were formed using pTOPO® shotgun cloning kit provided by Invitrogen. Phage genome inserts were between 1-2kb in length. Using the following calculation the total number of clones required was determined:

$$\left(\frac{kb\,genome\,size}{kb\,insert\,size}\right) x\,\mathrm{cov}erage\;\;required$$

Libraries were completed according to manufacturers' instructions and stored at -70°C.

2.25.2. DNA Sequencing

In the course of this project DNA sequencing was required to confirm the presence of inserts and for the purposes of phage genome sequencing. Mr. Roth Tate (SIBPS, University of Strathclyde) sequenced plasmid inserts on an Applied Biosystems 480. Sequence data analysis was carried out using Chromas® program. Complete phage genomes were sequenced by Prof. Roger Hendrix (University of Pittsburgh).

2.26. In silico analysis

Analysis of phage genome sequences was carried out utilising a variety of software programs.

2.26.1. BLAST Searching

Initial analysis of nucleotide and amino acid sequences was achieved by Basic local alignment tool (BLAST®). This allowed for the alignment of sequences against a publicly available database at the National Centre for Biotechnology Information Genebank. The programs BLASTP and BLASTN (Altschul *et al.*, 1997) were used for protein-protein analysis and nuclotide-nucelotide analysis, respectively.

2.26.2. Dot plots

Dot plots were created utilising the Genome Shovel computer program (Japan Science, 2008) .

2.26.3. Genome Sequence Annotation

In order to fully appreciate genome sequences of phages genome annotation was carried out. This was done with the use of GeneMark. ORF recognition allowed for more precise BLAST searches.

2.26.4. Clustal alignment

Clustal alignments were conducted using Jalview version 8.0 (Larkin *et al.*, 2007).

2.26.5. Dendrogram formation

Dendrograms were created using R software 2.7.2. Simple matching coefficients (Sokal and Michener, 1958) were used to create the matrix.

2.26.6. DNA mapping

Plasmid and phage DNA restriction maps were created using Serial Cloner (V 1.3)

2.26.7. Statistical analysis

All experimental procedures subject to statistical analysis were conducted in triplicate, unless otherwise stated. Microsoft Excel, R and MiniTab (v2.4) were applied in statistical analysis to calculate significant differences, standard error and ANOVA calculations.

3. Isolation and Characterisation of Soil Streptomycetes and their Bacteriophages

3.1. Introduction: Soil structure, its inhabitants and their isolation

In its natural state, soil is a complex group of individual habitats. It presents a heterogeneous medium, consisting of solid, liquid and gaseous phases, each section potentially varying greatly from its neighbouring habitat (Figure 9) (Sylvia *et al.*, 1999). As such the variety of bacterial life within soil has been subject of study due to the exploitable properties of the inhabiting organisms (Ashelford *et al.*, 2000; Chibani-Chennoufi *et al.*, 2004; Krugel *et al.*, 1980; Torsvik and Overas, 2002).

The actinomycetes and fungi are the most abundant soil micro-organisms (Casida, 1965; Hayakawa, 2004). Despite being present in such numerous quantities, other micro-organisms are also present in soil as dense populations, for example, pseudomonads and yeasts. Due to the great diversity and rapid growth of many soil micro-organisms, actinomycete isolation from soil is complicated by their slow growth in comparison to other soil bacteria (Hirsch and Christensen, 1983). Their isolation has been repeatedly adjusted and improved upon due to the value in their secondary metabolite production. Indeed, there are several publications, each with their own methodology and their claims of superiority over other published methods. As a result, there is currently no recognised technique for their isolation (Porter *et al.*, 1960; Van der Heijden *et al.*, 2008). It is likely that these techniques vary due to the diversity and adaptability of the soil organisms alongside their complex lifestyles. As a result it is doubtful that there will ever be a sole isolation technique for the actinomycetes.

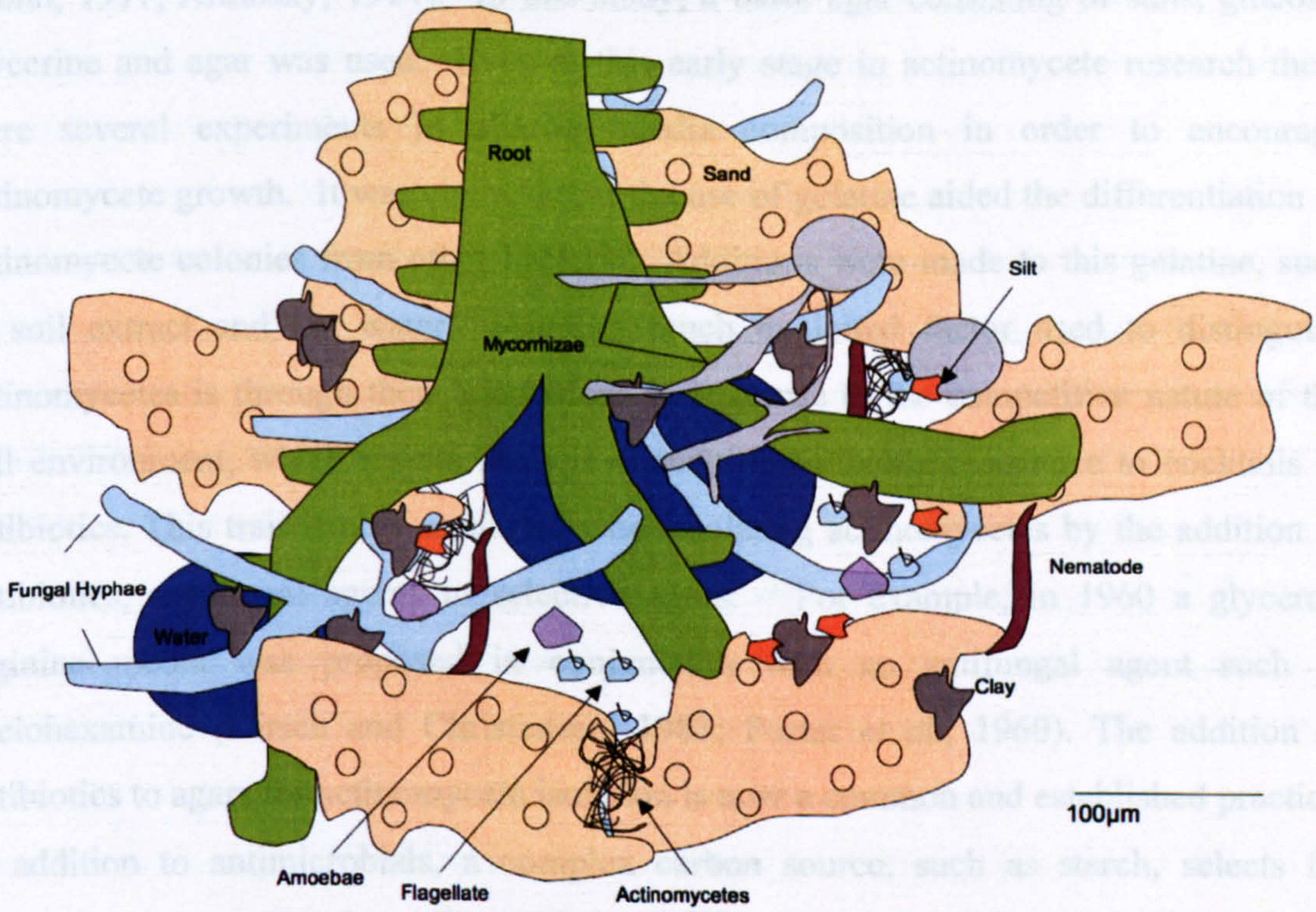


Figure 9 The soil habitat (adapted from Sylvia *et al*, 1999). The figure displays a representation of the soil microhabitat. Soil particles consist of clay, silt and water particles combined with filamentous fungi, root hairs, nematodes, amoebae, flagellates and bacteria, such as actinomycetes.

One of the very first studies in actinomycetes isolation was carried out in 1914 (Conn, 1917; Krainsky, 1914). In this study, a basic agar consisting of salts, glucose, glycerine and agar was used. Even at this early stage in actinomycete research there were several experiments in altering media composition in order to encourage actinomycete growth. It was claimed that the use of gelatine aided the differentiation of actinomycete colonies from other bacteria. Additions were made to this gelatine, such as soil extract and tap water. Another much exploited factor used to distinguish actinomycetes is through their adaptations in response to the competitive nature of the soil environment, which results in their display of antibiotic resistance to cocktails of antibiotics. This trait is now exploited when isolating actinomycetes by the addition of antibiotics, antifungal agents to selective agars. For example, in 1960 a glycerol-arginine media was proposed in conjunction with an antifungal agent such as cyclohexamide (Hirsch and Christensen, 1983; Porter *et al.*, 1960). The addition of antibiotics to agars for actinomycete isolation is now a common and established practice. In addition to antimicrobials, a complex carbon source, such as starch, selects for organisms that possess the ability of breaking down of such substances. Nevertheless, despite this research, there remains no media that exclusively selects for actinomycete growth (Athalye *et al.*, 1981; Conn, 1917; Hsu and Lockwood, 1975).

In order to identify actinomycetes, successful growth on a selective agar, alongside the correct morphological properties assigned to actinomycetes, can be used as an initial screen. However, one of the most rapid, efficient and accepted methods of identification is 16S rRNA sequencing (Fox, 1992). Dependent upon the region of the gene that is amplified, this technique can be used to identify organisms down to the species level (Stackebrandt *et al.*, 1991). The primary structures of rRNA molecules contain areas of sequences conserved to various degrees. It is these highly conserved regions that are the basis for phylogenetic analysis (Woese *et al.*, 1985). However, the variable regions have doubtful phylogenetic significance (Weisburg *et al.*, 1989). Studies applied in order to identify environmental samples typically use a combination of molecular approaches (Kremmerling *et al.*, 1989). This is required due to the great amount of diversity among prokaryotes in soil.

The abundance of *Streptomyces* is superseded by bacteriophages abundance in soil. The presence of these viruses is ten times greater than that of bacteria (Ashelford *et al.*, 2003), therefore, their approximate population is 10^9 per gram of soil. As a result, their isolation from soil may be performed with relative ease by application of the established enrichment technique or the direct plate count technique, developed and optimised by Lanning and Williams (Lanning and Williams, 1982). These techniques were optimised in order to recover the maximum number of virus particles from soil. It was found that the extraction eluent at a pH of 8 resulted in maximum recovery of 70%. In contrast, a pH of 10 resulted in a maximum phage yield of 8%. In addition, protein in the form of bovine albumin or egg albumin also resulted in an increased yield. These improvements are possible due to the fundamental nature of soil and phage particles. It has been shown that there are four characteristics which exert major influence on viral adsorption on and through soil. These are: pH, the presence of organic compounds, overall ionic strength and the nature and concentration of the ions present in the soil. It has been shown that viruses adsorb to soil at a low pH, due to the net negative charge of molecules such as silica, and the overall positive charge of viruses (McLaren *et al.*, 1996). Clearly, isolation of both host and phage is required in order to examine the dynamics of their relationship within their natural environment.

Aims of the Chapter

In order to examine the diversity of a group of phages from the same ecosystem, it was first necessary to isolate a group of host organisms within which the phage might come into contact. These hosts were then characterised initially by phenotypic traits, followed by 16S rRNA classification in order to characterise the panel of isolates. These host streptomycetes were then used to isolate phages from that environment. Bacteriophages were screened on all hosts in order to obtain infection profiles of each individual phage.

Additional experiments, such as determination of EOP within each host, were conducted in order to provide insight into their survival mechanisms and thus assess their diversity further.

3.2. Soil extraction of indigenous bacteria and their characterization

3.2.1. Soil source and treatment

Soil was sourced from a residential area in Erskine, UK (coordinates 55° 53'47. 49" N, -4° 26' 30.78" W). It had not been subject to fertilisation and therefore, the microbial populations should remain relatively stable, unlike those found in repeatedly fertilised soils. Soil was air-dried for two months before use and then sieved to remove large particles. The moisture content of soil was measured by dehydration, which was found to be 11%, pH was 5.59 but, following sterilisation by steam, it increased to 5.67. The soil type was slow permeable, seasonally wet, acid clay (British soil survey).

3.2.2. Isolation and characterisation of host bacteria from soil

In order to obtain a representative sample of free, viable bacteriophages from soil, it was first necessary to isolate streptomycete hosts from the soil that would subsequently be screened for bacteriophage lysis. Bacteria were isolated from the soil by shaking in ¼ strength ringers solution (Herron and Wellington, 1990). After seven days incubation, the colonies displaying characteristic morphology of streptomycetes

were streaked out for single colonies and incubated for a further seven days on MS agar (Hobbs *et al.*, 1989) before being examined for spore pigmentation, diffusible pigmentation and ability to degrade agar. Nine separate colony varieties were identified and each of approximately 300 colonies were assigned to one of the nine groups (Table 17, Figure 10). Morphological features of all selected colonies were consistent with their assignment to the genus *Streptomyces*.

Table 17 Characteristics of streptomycetes isolated from soil. The bacteria were initially isolated from soil using RASS agar with supplements (see 2.6.2) and separated on basis of spore colour, pigment production and the ability to digest agar. Nine distinct colony varieties were determined using these characteristics.

Host strain	Spore colour	Secondary pigment production	Agar digestion
SP01	Grey	Black	-
SP02	Grey-black	Black	+
SP03	Grey-white	Black	-
SP04	Grey	Pink & black	-
SP05	Dark grey	None	-
SP06	White	Blue & black	+
SP07	Yellow-grey	Black	+
SP08	White-yellow	Black	-
SP09	White-brown	Black	+

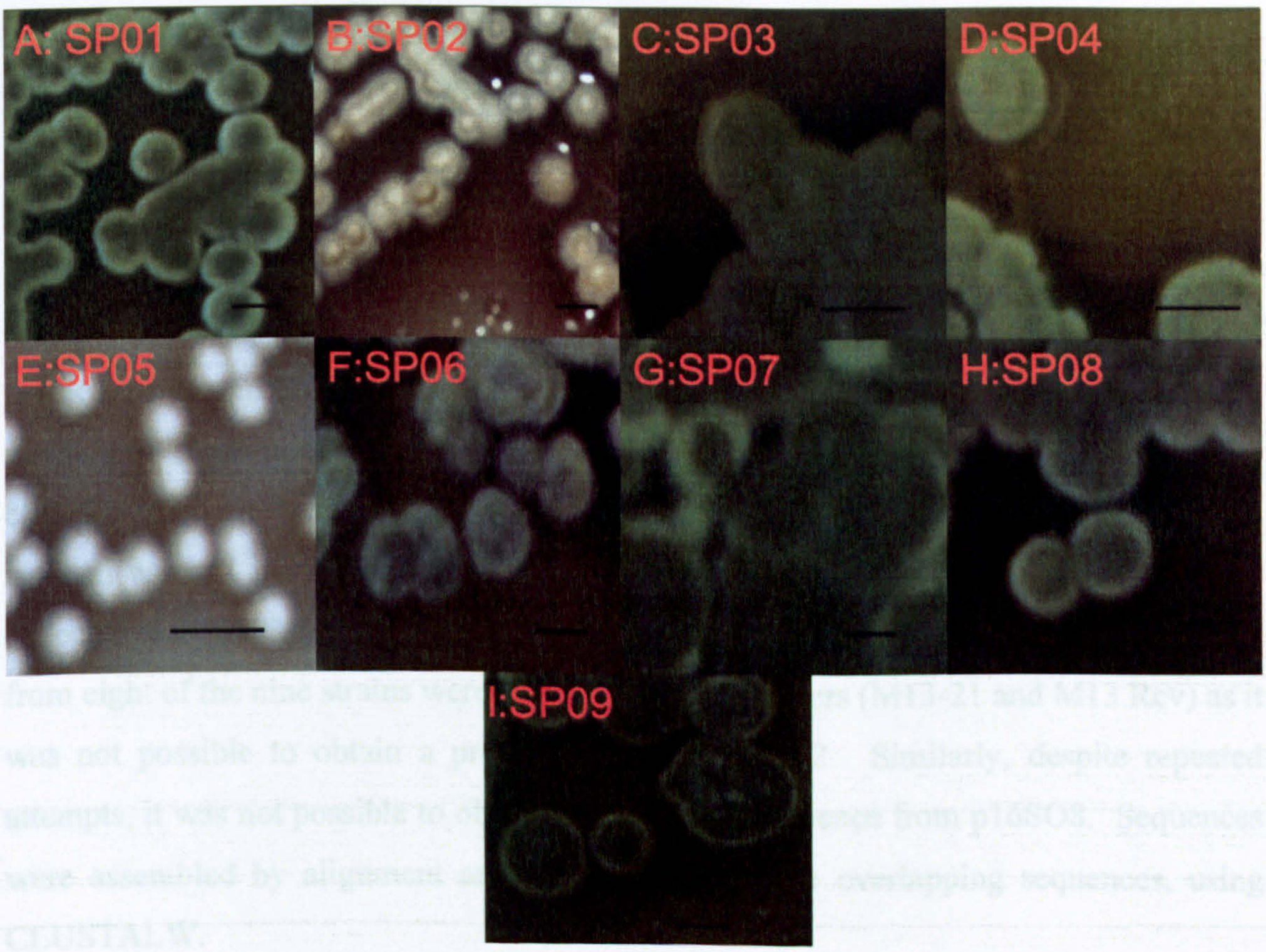


Figure 10 Soil isolates grown on MS. A to I: SP01 to SP09, respectively. All are plated on MS agar. Note the black diffused pigmentation, presumably melanin, within the agar surrounding colonies of SP01, SP02, SP03, SP04, SP06, SP08 and SP09. Black scale bars represent 5mm.

A representative strain was chosen from each of the nine morphological groups and confirmed as a streptomycetes by 16S rRNA sequencing.

16S rRNA was amplified according to Heuer *et al* (Heuer *et al.*, 1997). The primers used (Table 13) hybridised to and amplified variable regions V6 and V8; also covering the particular *Streptomyces* variable regions γ and β (Heuer *et al.*, 1997) (Figure 3). PCR products were visualised on an agarose gel (Figure 11) and the band of predicted size purified, using a Qiagen gel purification kit and cloned into pUC18, using a Novagen blunt ended cloning kit. Following transformation of *E.coli* JM109, putative clones were identified on the basis of blue/white selection (2.19.7). Plasmid DNA was isolated using a Promega Wizard kit and was confirmed as containing inserts of the expected size (400bp) following restriction digestion (Figure 12) (2.17). Clones derived from eight of the nine strains were sequenced using primers (M13-21 and M13 Rev) as it was not possible to obtain a product PCR from SPO2. Similarly, despite repeated attempts, it was not possible to obtain a meaningful sequence from p16SO8. Sequences were assembled by alignment and identification of the overlapping sequences, using CLUSTALW.

All eight clones were sequenced, however, only seven meaningful sequences were obtained, a sequence for SPO8 could not be attained, despite repeated attempts. Standard nucleotide-nucleotide BLAST search against the DDBJ/EMBL/GenBank revealed close relationships with characterised *Streptomyces* as well as uncultured *Streptomycetes*. The sequences were aligned and placed in a phylogenetic tree (Figure 13). The final dataset was then used to construct a phylogenetic tree (Figure 14).

It is clear from Figure 14 that the indigenous streptomycetes form an independent phyletic clade in the 16S rRNA gene sequence *Streptomyces* tree. This lack of a direct link to other characterised streptomycetes is consistent with previous studies involving the isolation of *Streptomyces* from soil (Kim *et al.*, 2004; Semedo *et al.*, 2004; Sun *et al.*, 2007). In addition, it exemplifies the diversity of these organisms in soil, as none of these sequences were present in the GeneBank database. Despite this, the sequence analysis shows that the indigenous cluster is loosely related to the type strain

of *Streptomyces avermitilis*. The clear distinct grouping of the indigenous bacteria indicates that within soil they share a common unique ancestor.

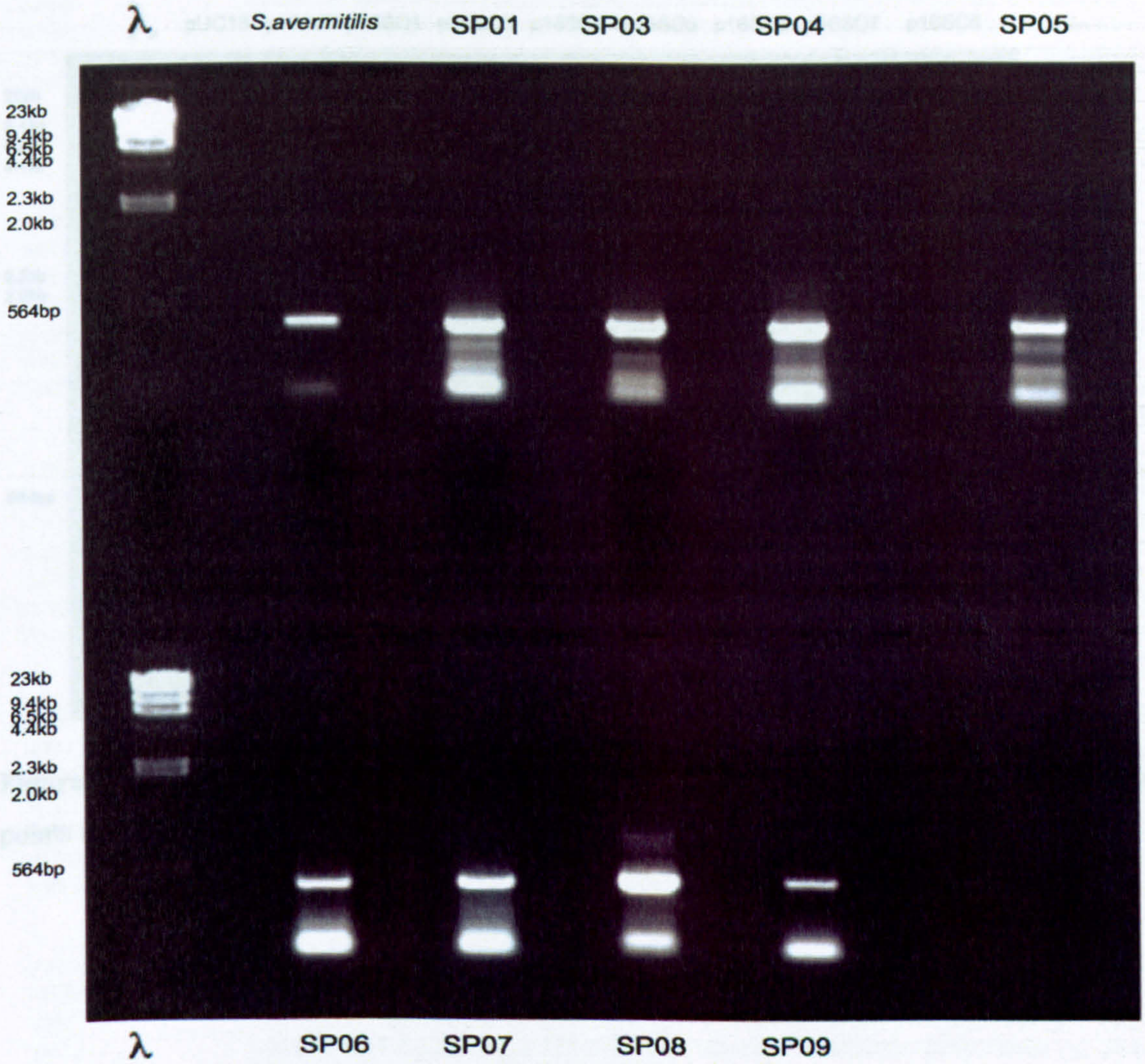


Figure 11 Agarose gel showing PCR products following amplification of the 16S rRNA sequence. Bands of ~500bp were extracted and purified using Quiagen gel extraction kit. TOP: λ HindIII, PCR products amplified from *S.avermitilis*, PCR products amplified from SPO1, PCR products amplified from SPO3, PCR products amplified from SPO4, PCR products amplified from SPO5 BOTTOM: λ HindIII, PCR products amplified from SPO6, PCR products amplified from SPO7, PCR products amplified from SPO8, PCR products amplified from SPO9. No PCR product was obtained from SPO2 despite repeated attempts.

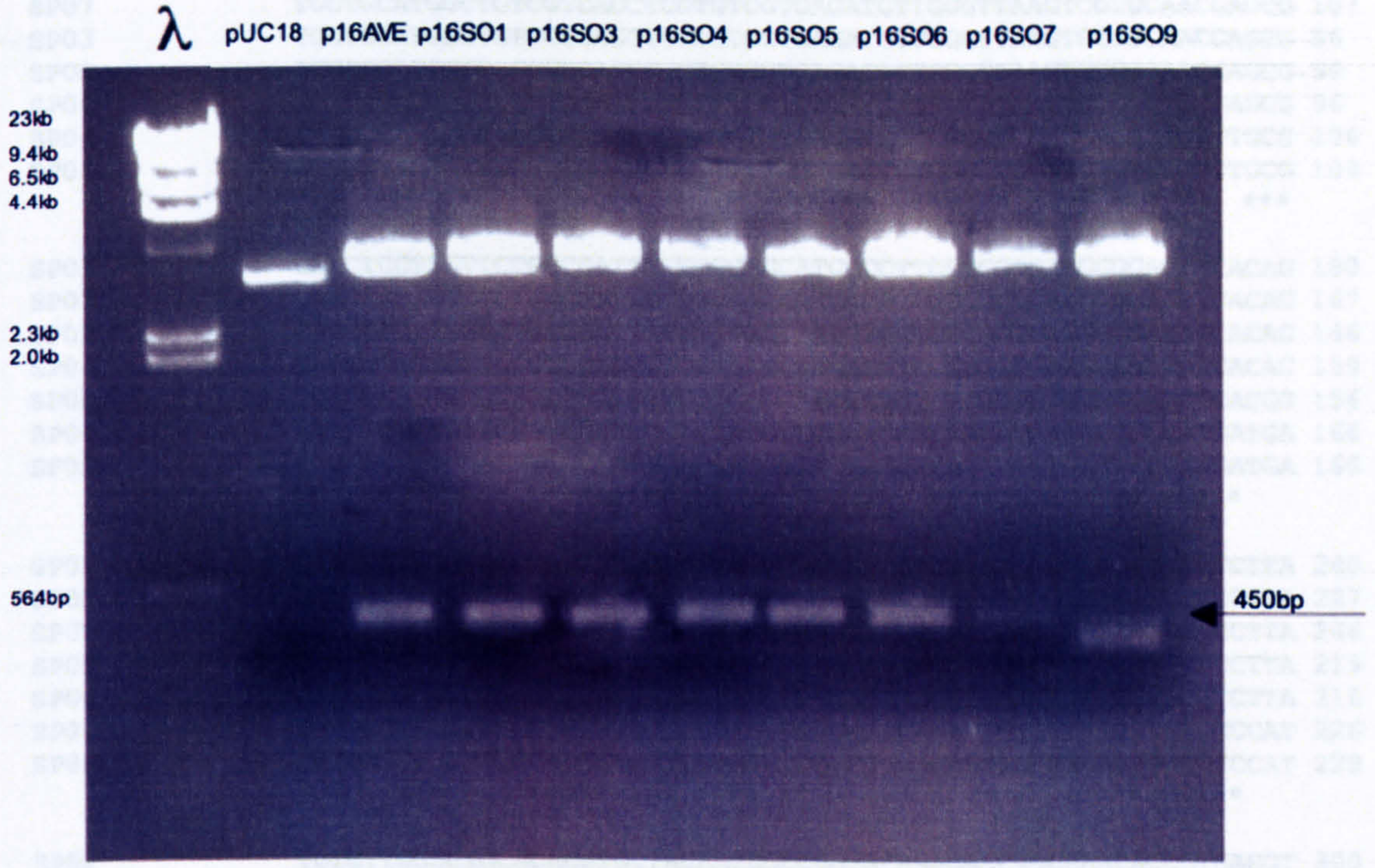


Figure 12 Agarose gel showing the screen of 16S rRNA PCR inserts using EcoRI. Arrow points to 450bp.

```
SP01    TGGTGCATGGCTGTCGTTAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 120
SP07    TGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 107
SP03    TGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 86
SP09    TGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 99
SP06    TGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG 96
SP04    CCGGCCAATCCGAACTGAGACCGGCTTTTTGAGATTCTGGGCTCCACCTCACGGTTTGCG 106
SP05    CGGCCCAATCCGAACTGAGACCGGCTTTTTGAGATTCTGGGCTCCACCTCACGGTTTGCG 109
          *  **   *   *     * * * ****** **** *    * * *     ***

SP01    CAACTCCTTGTTCTGGCCATTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACATCACAG 180
SP07    CAACTCCTTGTTCTGGCCATTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACATCACAG 167
SP03    CAACTCCTTGTTCTGGCCATTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACATCACAG 146
SP09    CAACTCCTTGTTCTGGCCATTGTTGCCAGCATGCCCTTCGGGGTGATGGGGACATCACAG 159
SP06    CAACTCCTTGTCCCGGCCATTGTTGCCAGCAAGCCCTTCGGGGTGTTGGGGACATCACGG 156
SP04    CAGCTCTTTGTACCGGCCATTGTAGCACGTGTGCAGCCCAAGACATAAGGGGCATGATGA 166
SP05    CAGCACTTTGTACCGGCCATTGTAGCACGTGTGCAGCCCAAGACATAAGGGGCATGATGA 169
        ** * * **** * ********* **  *   **    * *       *** *** *

SP01    GAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA 240
SP07    GAGACCGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA 227
SP03    GAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA 206
SP09    GAGACCGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA 219
SP06    GAGACCGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTA 216
SP04    CTTGACGTCGTCCCCACCTTCCTCCGAGTTGACCCCGGCAGTCTCCTGTGATGTCCCCAT 226
SP05    CTTGACGTCGTCCCCACCTTCCTCCGAGTTGACCCCGGCAGTCTCCTGTGATGTCCCCAT 229
           * **    ** **       ** **    ** *         * *** ***

SP01    TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAAAGAGCTGCGAACACCGTGAGGT 300
SP07    TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATAAGCTGCGATCACCGTGAGGT 287
SP03    TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAAAGAGCTGCGAACACCGTGAGGT 266
SP09    TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATCACCGTGAGGT 279
SP06    TGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATCACCGCAAGGT 276
SP04    CACCCCGAAGGGCATGCTGGCAACAATGGCCAGAACAATGGGTTGCGCTCGTTGCGGGAC 286
SP05    CACCCCGAAGGGTATGCTGGCAACAATGGCCAGAACAATGGGTTGCGCTCGTTGCGGGAC 289
          *  *    * *  *  ** ********* * ****   * ****  *   *   *

SP01    GGAGCCCAGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGCCGTCTGCAACTCGACCC 360
SP07    GGAGCCCAGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGCCGTCTGCAACTCGACCC 347
SP03    GGAGCCCAGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGCCGTCTGCAACTCGACCC 326
SP09    GGAGCCCAGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGCCGTCTGCAACTCGACCC 339
SP06    GGAGCCCAGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGCCGTCTGCAACTCGACCC 336
SP04    TTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGTACACCGACCA 346
SP05    TTAACCCAACATCTCACGACACGAGCTGACGACAGCCATGCACCACCTGAACACCGACCA 349
          * ****  ****** * * *    *   *   ** ** *** *  *****

SP01    CATGAAGTCGGACGTTGCTAGTAATCGCAGATCAGCATTAAGCTGCGGTGAATACGTTCC 420
SP07    CATGAAGTCGGACGTCGCTAGTAATCGCAGATCAGCATTAAGCTGCGGTGAATACGTTCC 407
SP03    CATGAAGTCGGACGTTGCTAGTAATCGCAGATCAGCATTAAGCTGCGGTGAATACGTTCC 386
SP09    CATGAAGTCGGACGTCGCTAGTAATCGCAGATCAGCATTAAGCTGCGGTGAATACGTTCC 399
SP06    CATGAAGTCGGACGTCGCTAGTAATCGCAGATCAGCATTAAGCTGCGGTGAATACGTTCC 396
SP04    CAAGGGGGCGACCATCTCTGGCCGTTTCCGGTGTATGTCAAGCCTTGGTAAGGTTCTTCG 406
SP05    CAAGGGGGCGACCATCTCTGGCCGTTTCCGGTGTATGTCAAGCCTTGGTAAGGTTCTTCG 409
        ** *  * **  * *  ** *   *  * * *     * ****   *** *     ***
```

Figure 13 Clustal alignment of the indigenous streptomycetes 16S rRNA sequences. This alignment was created from the PCR amplified 16S rRNA sequences from SPO1, SPO3, SPO3, SPO4, SPO5, SPO6, SPO7 and SP09 on CLUSTALW.

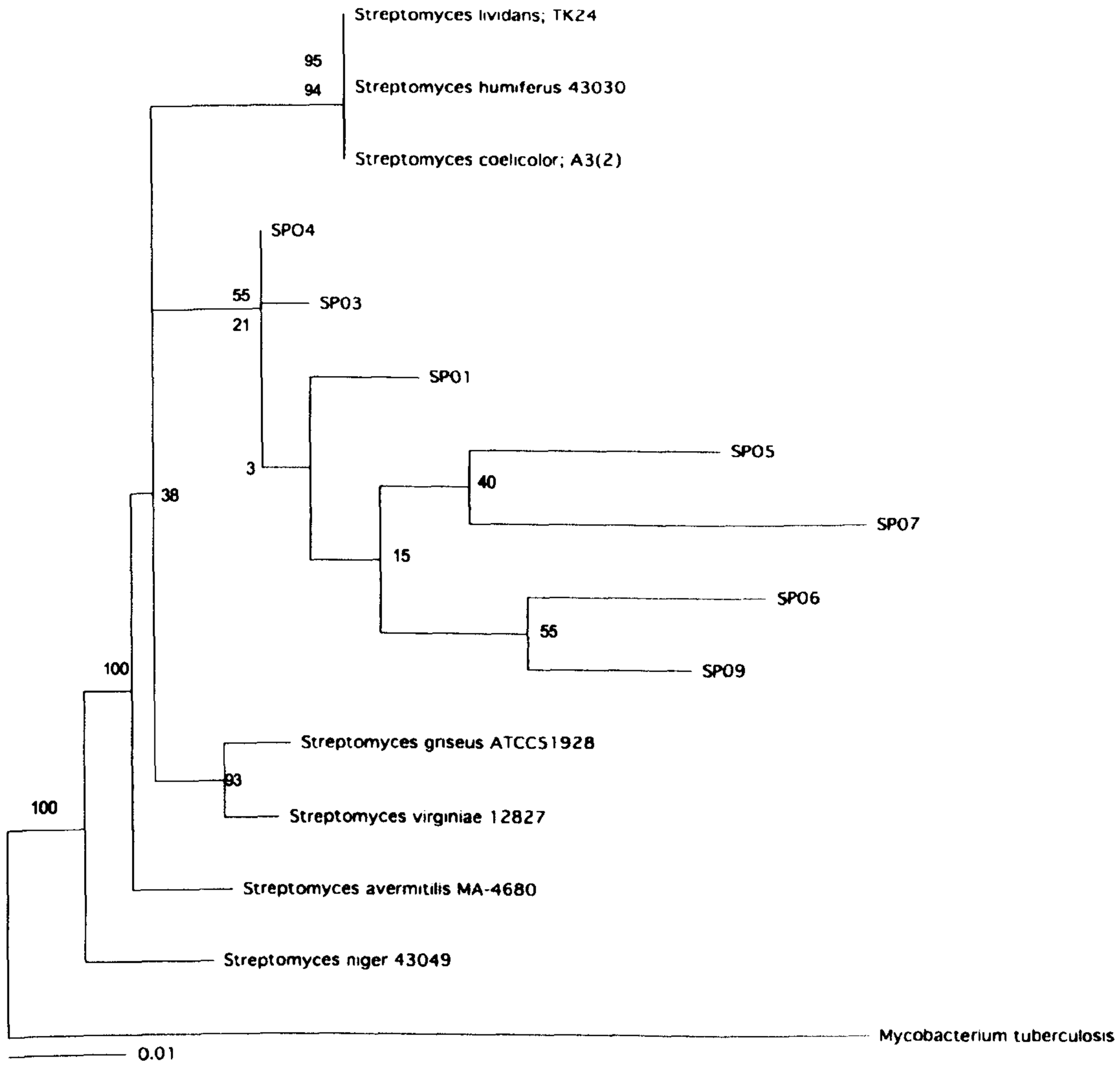


Figure 14 Unrooted neighbour joining tree based on 16S rRNA sequences showing the positions of newly isolated indigenous soil bacteria. Out-group is *Mycobacterium tuberculosis.* Bar 0.1 substitutions per nucleotide position and numbers at the nodes are bootstrap values. The indigenous strains are within their own distinct clade.

3.3. Bacteriophage isolation from soil and their characterisation.

Phage typing has shown that specific host range phages exist alongside highly polyvalent phages (Korn-Wendisch and Schneider, 1992); therefore, specific and polyvalent phages should be present within the soil used throughout this study. In order to isolate bacteriophages that infected both indigenous and non-indigenous streptomycetes, a number of strategies were employed. Firstly, *S.lividans* was employed as a host in order to optimize isolation conditions as this strain is relatively well characterised (Dyson and Schrempf, 1987; Leblond *et al.*, 1993) and is known to support infections by a range of bacteriophages (Lanning and Williams, 1982) and is pgl negative (Sumby and Smith, 2002). There are two methods by which phages can be isolated from an environmental sample, firstly, direct isolation, which yields virions present at the time of isolation. This method provides a quantitative measure of phage infecting a particular host in soil. The second means of isolating phages from an environmental sample is by enrichment. This relies on inoculation of a host organism coupled with incubation for a period of time in order to amplify indigenous phages by their infection of that host. As a result, it is not quantitative. Both these techniques were applied in this work.

In order to determine the efficacy of the direct isolation procedure of Lanning & Williams (1982) in Erskine soil, sterile soil was spiked with a known concentration of ϕC31 particles and subsequently subjected to the extraction procedure of Lanning & Williams (1982). These authors showed that eluent pH can have a dramatic impact on the efficacy of this technique, and as a result, extraction efficacy was compared in different eluent pHs (Table 18).

Table 18 Recovery of ϕC31 particles isolated from spiked sterile soil. Soil was sterilized as in section 2.13. Virus particles were added to 10g soil microcosms and incubated for 15 minutes at a concentration of 1 x 10^4 pfu/g and subject to the direct count isolation with *S.lividans* as a host.

pH of broth	**Phages recovered (pfu/mL)**	**Percentage recovery**
7	2 x 10^2	2%
8	4x 10^3	40%
9	4 x 10^1	0.4%

The maximum success of 40% re-isolation of phages was not as high as the 70% described in previous studies (Lanning and Williams, 1982). When unspiked soil was extracted using an eluent of pH8, the total indigenous phages capable of infecting *S.lividans* 1326 were 2.25 x 10^3 pfu/g. This is less than 0.002% of the total phage population estimated to be in soil. According to the literature, virus-to-bacterium ratio in soil is 0.04 (Ashelford *et al.*, 2003). Therefore, it may be estimated that there are 5.62 x 10^4 viable hosts of this phage in soil. However, it must also be noted that there are conflicting figures estimating the total number of phages in soil, therefore this figure is questionable (Danovaro and Serresi, 2000; Danovaro, 2002). Clearly, these viruses and their hosts are abundant within the soil used throughout this study.

In order to maximise the number of phages isolated, indigenous actinophages were initially extracted from soil by the enrichment process using only *S.lividans* 1326 as a host. Despite it not necessarily being indigenous to this particular soil, its role as a host was successful in isolating actinophages, which gave a variety of plaque morphologies (Figure 15). The majority of plaques were turbid, however, approximately 10% were clear. Plaque size also varied from pinprick plaques to some ~0.5cm in diameter. Initially, four actinophages were selected for further study and termed ϕELB17, ϕELB18, ϕELB19, ϕELB20; all were isolated using *S.lividans* as a host. Although, not being conclusive marker plaque morphology was used as a crude reference indicating diversity within these phages.

In order to isolate phages that infected the nine indigenous *Streptomyces* described in section 3.2.2 separate soil samples were inoculated with SPO1, SPO2, SPO3, SPO4, SPO5, SPO6, SPO7, SPO8 and SPO9, phages were extracted according to enrichment (section 2.14) and assayed using the appropriate host streptomycete. This resulted in a greater variety of plaque morphologies that infected each host. Plaque turbidity indicated that the majority of soil phages were temperate, however, it is also evident that a minority were lytic, at least while within specific hosts. Consequently, these initial experiments demonstrate that actinophages within this soil infecting the panel of hosts were abundant and could be readily isolated. A single plaque from each of the nine host strains (SPO1, SPO2, SPO3, SPO4, SPO5, SPO6, SPO7, SPO8 and SPO9) was further purified and termed ϕELB1 to ϕELB9, respectively. Although plaque morphology was more varied when isolating phages using the indigenous strains there still remained a substantial lack of variety, as the majority of plaques were small and turbid. These findings support that of previous phage soil isolation studies which state that variation in plaque traits is narrow (Sharma *et al.*, 2002).

In order to provide a measure of phage diversity, the infectivity and plaque morphology of each of the panel phage isolates, in addition to the, characterised actinophages ϕC31. ϕR4, and ϕHau3A1. Each phage was then screened for plaque formation and morphology on the panel of indigenous and typed strains. The results from the screening process indicate a large amount of diversity in patterns of strain infectivity within the soil phages (Table 19). Interestingly, no phage was isolated that was exclusively lytic in all hosts: this points to the possession of a survival mechanism.

Using non-molecular data, relationships were examined in order to provide an indication of the functional diversity of phages and hosts within soil. Dendrograms were created by comparison of susceptibility to phage infection and creation of a similarity matrix by the simple matching coefficient. This data was then processed using R. Figure 16 demonstrates the grouping of the indigenous host strains according to their susceptibility to phage infection and clearly shows the grouping of these into distinct clusters, none of which are directly related to the typed strains. Apart from the lack of a direct link to the typed strains, the 16rRNA phylogenetic tree does not support these

relationships. This illustrates the extensively documented flaws in prokaryotic taxonomy (Anderson and Wellington, 2001; Stackenbrandt *et al.*, 2002).

In a similar manner, the bacteriophages were typed according to their host range, and grouped accordingly (Figure 17). This provided an insight into community organisation of the soil phages. Unlike their non-indigenous host strains, typed, characterized non-indigenous phages appeared to have direct relationships with the indigenous phages. In some cases, this relationship was substantial, such as that of ϕELB17 and ϕR4, which shared a direct relationship. Therefore, this finding indicates that there may be one or two common ancestors of all the indigenous and non-indigenous phages, potentially reducing the extent of their predicted diversity.

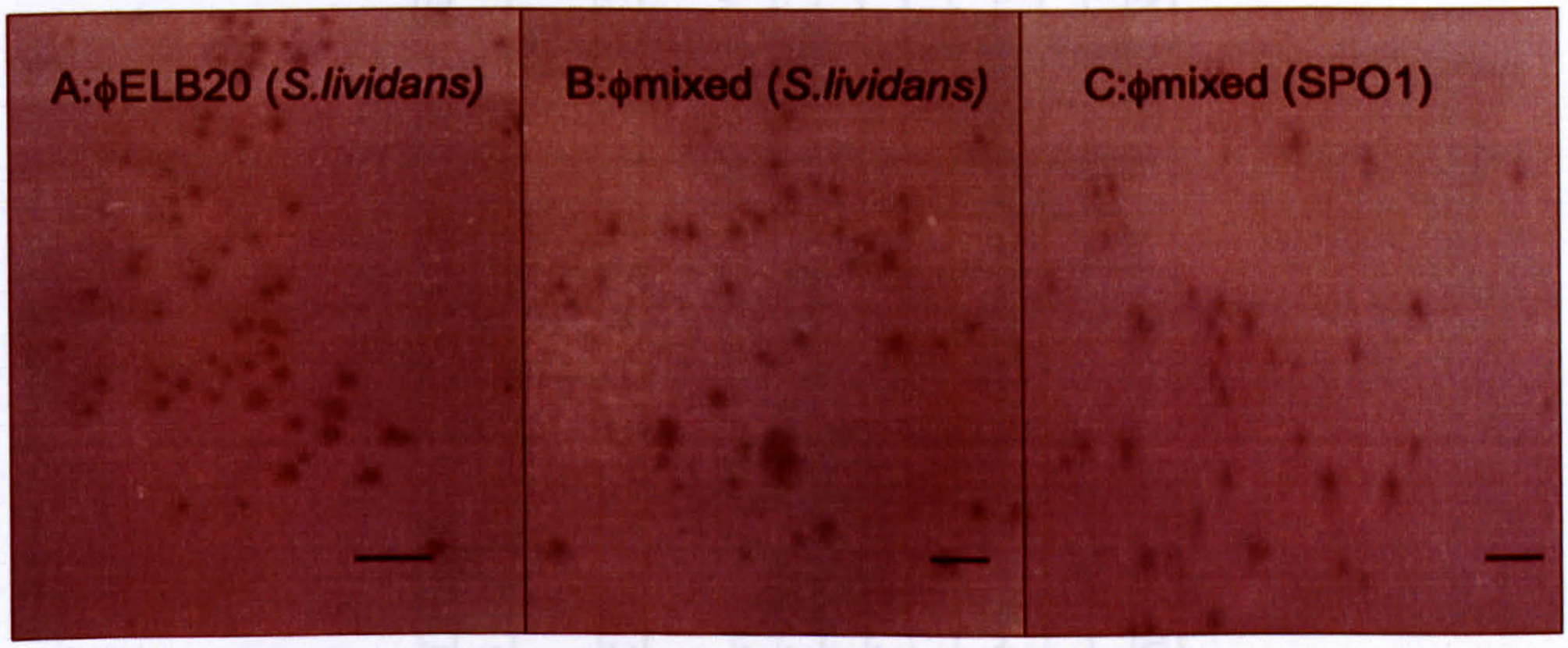


Figure 15 The variety of bacteriophage plaques. A: ϕELB20 on *S.lividans*, B: mixed phages isolated from soil on *S.lividans* 1326 using the enrichment procedure, C: mixed phages isolated from soil on SP01 using the enrichment procedure. Black bars represent 3mm.

Host Strain	φC31	φR4	φHau3 AI	φELB17	φELB18	φELB19	φELB20	φELB1	φELB2	φELB3	φELB4	φELB5	φELB6	φELB7	φELB8	φELB9
S.lividans 1326	T	T	-	T	T	T	T	-	T	-	-	C	T	T	-	-
S.avermitilis	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	T
S.coelicolor M145	-	T	T	T	T	T	T	-	-	-	-	-	-	-	T	-
SP01	C	T	C	T	T	T	T	T	C	T	T	T	-	C	-	-
SP02	T	T	C	T	T	T	T	T	T	T	-	T	-	T	C	-
SP03	T	C	C	C	C	C	C	C	C	T	-	T	-	C	T	-
SP04	C	T	T	T	T	T	T	T	-	T	T	T	-	T	T	-
SP05	T	T	T	T	T	T	T	-	T	-	-	T	-	T	C	-
SP06	T	-	T	-	-	-	-	-	-	-	-	-	T	-	T	-
SP07	T	T	T	T	T	T	T	T	-	-	-	T	T	T	C	T
SP08	T	T	C	T	T	T	T	-	-	-	-	C	-	T	T	-
SP09	-	-	T	-	-	-	-	T	-	-	-	T	T	T	-	T
φC31 lysogen	-	T	N/D	T	T	T	T	N/D	T	N/D	N/D	-	T	T	N/D	N/D

Table 19 Bacteriophage library host range and plaque morphologies. T = turbid plaques formed, C = clear plaques formed, - = does not infect. N/D: Not done.

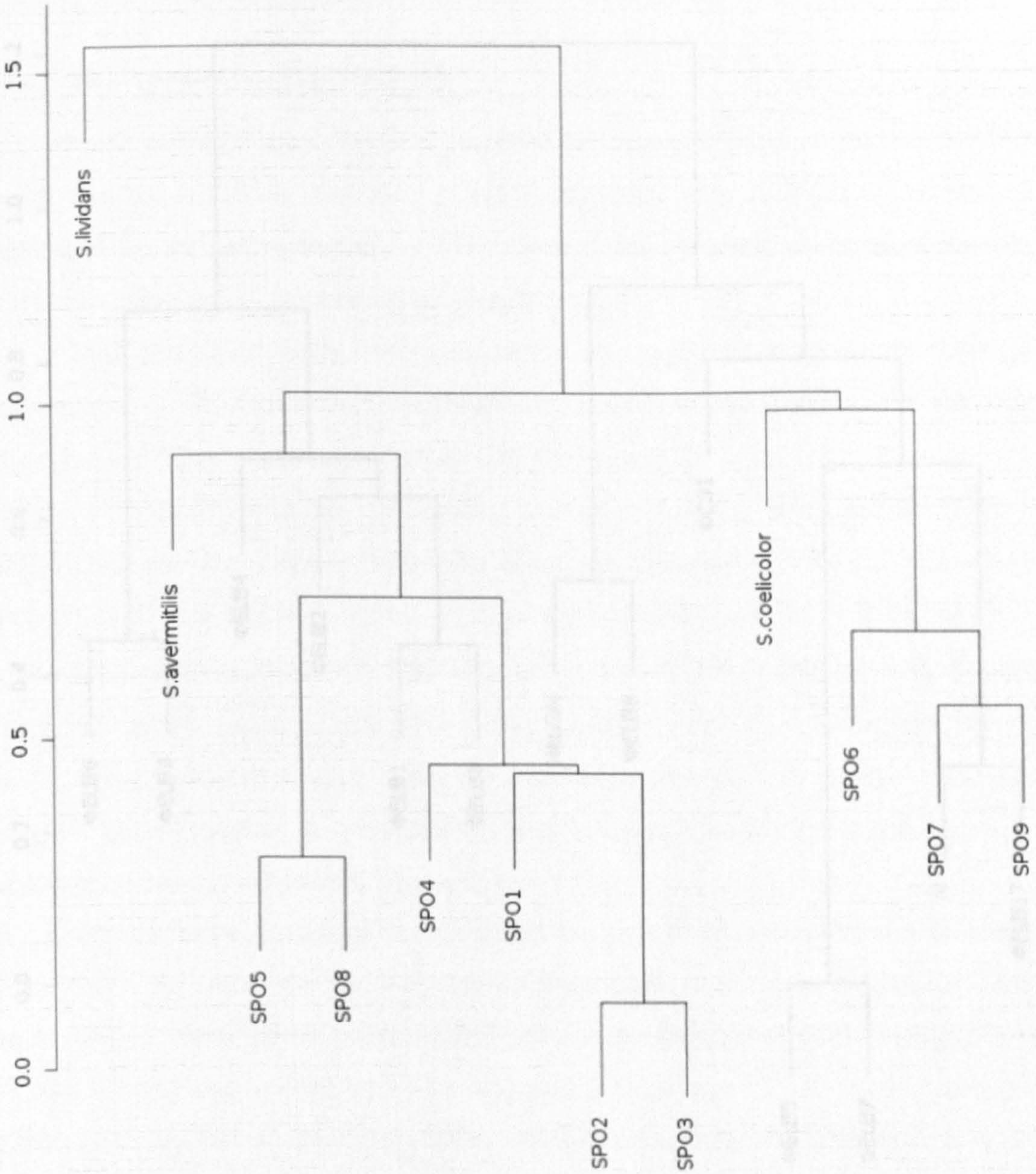


Figure 16 Relationships of indigenous and wild type strains derived by phage typing. Figure produced using R. Figure produced using R, pair group matching algorithm. Characterised, typed strains are included in order to give a comparison of their relationships. Height is the Euclidean distance.

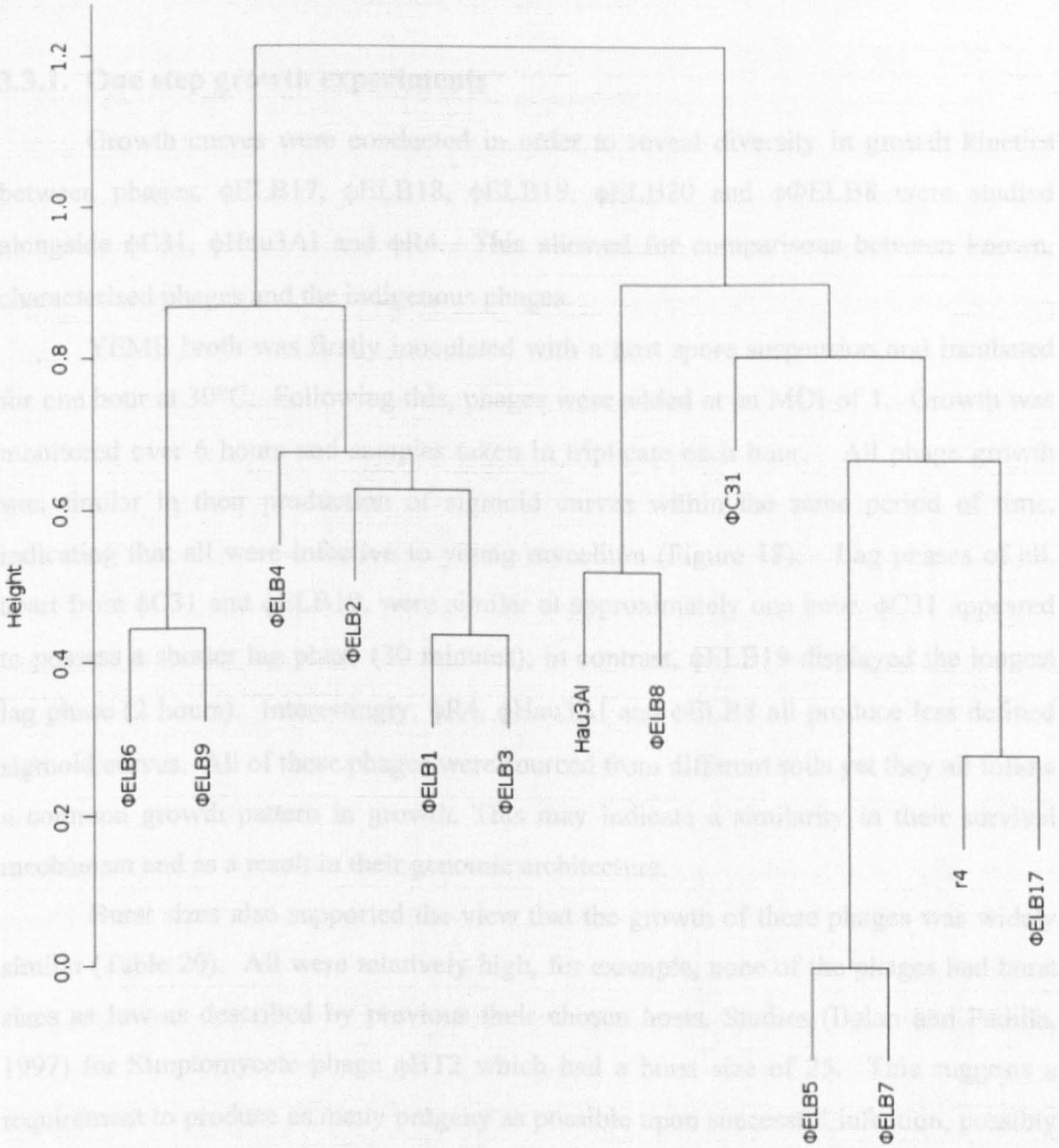


Figure 17 Relationships of indigenous and wild type bacteriophages derived from host range. Figure produced using R, pair group matching algorithm. Characterised, typed strains are included in order to give a comparison of their relationships. Height is the Euclidean distance.

3.3.1. One step growth experiments

Growth curves were conducted in order to reveal diversity in growth kinetics between phages. ϕELB17, ϕELB18, ϕELB19, ϕELB20 and ϕΦELB8 were studied alongside ϕC31, ϕHau3AI and ϕR4. This allowed for comparisons between known, characterised phages and the indigenous phages.

YEME broth was firstly inoculated with a host spore suspension and incubated for one hour at 30°C. Following this, phages were added at an MOI of 1. Growth was monitored over 6 hours and samples taken in triplicate each hour. All phage growth was similar in their production of sigmoid curves within the same period of time, indicating that all were infective to young mycelium (Figure 18). Lag phases of all, apart from ϕC31 and ϕELB19, were similar at approximately one hour. ϕC31 appeared to possess a shorter lag phase (30 minutes); in contrast, ϕELB19 displayed the longest lag phase (2 hours). Interestingly, ϕR4, ϕHau3AI and ϕELB8 all produce less defined sigmoid curves. All of these phages were sourced from different soils yet they all follow a common growth pattern in growth. This may indicate a similarity in their survival mechanism and as a result in their genomic architecture.

Burst sizes also supported the view that the growth of these phages was widely similar (Table 20). All were relatively high, for example, none of the phages had burst sizes as low as described by previous their chosen hosts. Studies (Balan and Padilla, 1997) for Streptomycete phage ϕBT2 which had a burst size of 25. This suggests a requirement to produce as many progeny as possible upon successful infection, possibly due to the small chance of a phage finding a viable host in the soil. The viruses selected for growth curve analysis were all highly polyvalent and shared many of the same hosts, which may account for their apparent similarities. These results suggest that the phages possessed similar growth and burst sizes, a characteristic which presumably optimal for their survival within soil.

Despite this diversity there are some striking similarities between ϕELB17, ϕELB18, ϕELB19, ϕELB20 and ϕR4; as their host range, and lifecycle within each host

is the same. It is possible that this is the same phage or that this profile of traits is the most adequate for their environment. As a consequence, unrelated phages may have evolved this survival strategy due to selective pressures.

This work has made evident that there is indeed a degree of diversity in both hosts and phages within the soil. The varying host ranges between phages clearly demonstrate that there are different phages within a single soil environment. As the majority are polyvalent, there is competition between viruses for the same hosts. In this study there was no bacteriophage isolated that was only capable of infecting a single host; all phages displayed polyvalencey to an extent. There are, however, less promiscuous phages, such as ϕELB9, which is limited to just a few hosts. This specificity could be detrimental to their survival as it is incapable of infecting the majority of streptomycete within its environment. However, as in the case of ϕELB9, its host, SP09 is not susceptible to infection by many other phages, unlike other indigenous strains, such as SP02. It is therefore possible that ϕELB9 has evolved to target a specific host, thereby reducing competition from other phages. This finding is supported by the EOP data (Table 21) of SPO9 as it contains one of the lowest EOPs in strains screened against ϕHau3A, indicative of a restriction modification system or similar defence. This data demonstrates that there are several different strategies employed by host organisms in order to avoid harmful phage infection.

One of these survival strategies is lysogeny of hosts. Within this study, ϕC31 lysogens were examined in order to indicate if any of the indigenous phages were similar or related to this well characterised phage. These were created by scraping from the middle of a turbid plaque of ϕC31 on *S.lividans* and streaked for a single colony. Following seven days of incubating, spores were harvested and screened as a host for all phages. All indigenous phages, aside from ϕELB5 were capable of infecting the lysogen. This inability of ϕELB5 to infect the lysogen is an example of homoimmuntity. As a result, conclusion was reached that ϕELB5 was similar to ϕC31. In contrast, remaining phages were capable of infecting the lysogen. The majority of indigenous

phages were clearly not similar to ϕC31, as the ϕC31 repressor did not prevent lysis. This indicated diversity in strategies within resident soil populations.

By examining the efficiency of plating, the degree of host strain resistance to a phage may be derived. Therefore, in order to continue phage characterisation, these were measured for ϕC31, ϕR4, ϕHau3AI, ϕELB20 and ϕELB8 within indigenous and typed hosts. Inoculating the chosen host strains with a consistent concentration of phage particles and counting the resulting plaques in each strain was carried out in order to derive the EOP. Interestingly, the organisms with high EOPs for all phages were clustered together in the phage typing dendrogram (Figure 16), yet not in the 16S phylogenetic tree (Figure 14). In contrast, those with low EOPs or resistance were also grouped together. This indicated two distinct groups of hosts within soil; those which were highly susceptible to infection and those which were not. For example, SPO6, SPO7 and SPO9 were neither susceptible to infection nor giving low EOPs. This resistance in one group may be due to the acquisition of phage defence mechanism such as restriction modification, lack of a phage receptor or Abi-like systems. Irrespective of the mechanism it is evident that they are effective against not only one phage, but also several others, particularly in the case for SPO9.

In contrast, the group susceptible to phage infection and appearing to lack defences against phage infection, for example, SP04, has clearly survived, despite its shared environment with virulent phages. Previous studies have proposed that a lack of defence mechanisms such as restriction modification systems may, in the long run, be more advantageous to the host than developing complex defences due to the continuous requirement to maintain them (Naito *et al.*, 1995).

Clearly 16S rRNA sequencing illustrates some degree of diversity, however, overall this region is highly conserved over generations, hence its application in taxonomy. In contrast, it is probable that other factors such as surface molecules on the bacteria are less conserved thereby resulting in differences in susceptibility to phage infection not illustrated by 16S rRNA sequencing. It is in this way that phage typing may provide more representative relationships between bacteria. As this study

illustrates phages are capable of infecting across different clades suggesting that unlike their hosts, phages can occupy several different niches.

Evidently, phage and host have co-existed together in soil and evolved survival mechanisms in an "arms race". This work has shown that there is indeed diversity within soil hosts and phages. Hosts can be classified into two overall groups, those that exert efforts in avoiding phage infection and those that do not. It is clear that both strategies are effective due to the ease with which organisms from each group can be isolated from the same soil.

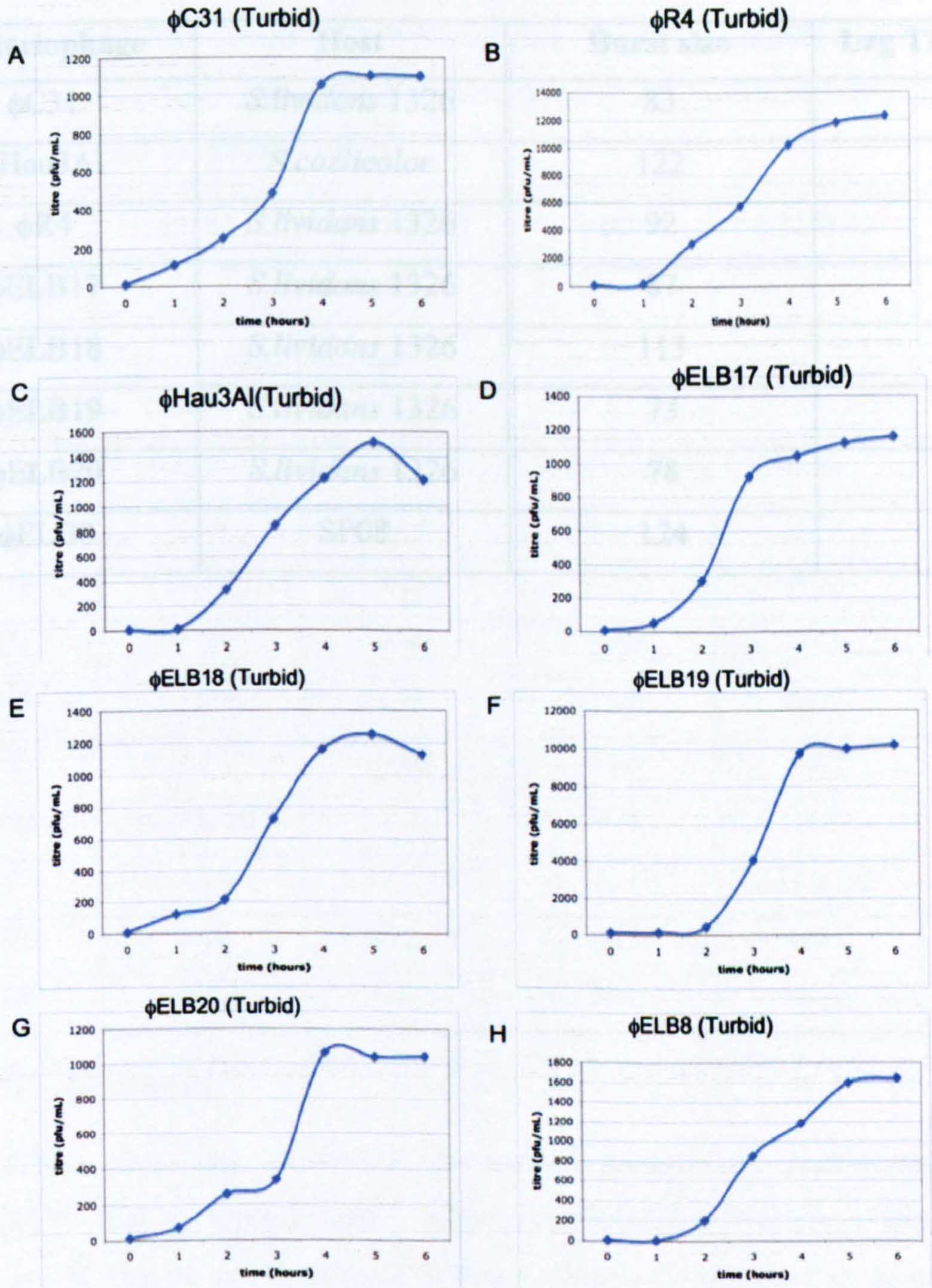


Figure 18 One step growth experiments. All were deduced from triplicate experiments. A: ϕC31 with *S.lividans* B: ϕR4 with *S.lividans* C: ϕHau3AI with *S.coelicolor* D: ϕELB17 with *S.lividans* E: ϕELB18 with *S.lividans* F: ϕELB19 with *S.lividans* G: ϕELB20 with *S.lividans* H: ϕELB8 with SPO8. Liquid cultures were inoculated with phages at MOI of 0.1. Samples were taken every hour in triplicate.

Table 20 Burst sizes and lag times of phages deduced from one-step growth curve experiments.

Bacteriophage	Host	Burst size	Lag Time (hours)
ϕC31	*S.lividans* 1326	83	1
ϕHau3AI	*S.coelicolor*	122	1
ϕR4	*S.lividans* 1326	92	1
ϕELB17	*S.lividans* 1326	87	1
ϕELB18	*S.lividans* 1326	113	1
ϕELB19	*S.lividans* 1326	73	2
ϕELB20	*S.lividans* 1326	78	1
ϕELB8	SP08	124	2

Table 21 EOP of phages. Phage lysates were produced on their original host (see Chapter 2). In triplicate, 1 x10^3 pfu/mL of phage were inoculated with the host; resulting plaques were counted, thereby allowing for EOP to be calculated. This was conducted in order to give an indication of defence mechanisms, such as restriction modification, held by the host strain.

Host	ϕC31	ϕHau3AI	ϕR4	ϕELB20	ϕELB8
S.lividans 1326	1	NA	0.14	1	NA
S.avermitilis	0.85	1	1	1	0.5
S.coelicolor	NA	1	2.07×10^{-3}	0.02	0.07
SP01	0.50	0.08	0.11	0.09	0.31
SP02	0.83	0.29	0.13	0.02	0.33
SP03	0.98	0.96	0.06	0.49	0.40
SP04	1	1	0.06	0.07	1
SP05	0.17	0.03	0.08	0.07	0.01
SP06	1.57×10^{-3}	1×10^{-3}	NA	NA	1.1×10^{-3}
SP07	6×10^{-4}	3.33×10^{-4}	4.1×10^{-4}	2.33×10^{-3}	0.01
SP08	0.67	1.42×10^{-3}	0.63	0.51	0.02
SP09	NA	1.9×10^{-3}	NA	NA	NA

3.4. Conclusions

This work has presented an overall picture of phage-host abundance, relationships and survival in soil. Initial experiments provided an indication of the abundance of phages and their hosts in that medium. Following this, host classification according to 16rRNA data illustrated that these organisms were distinct from other typed streptomycetes. Phage typing supported this distinction. It was apparent that the indigenous hosts shared a common ancestor, from which they have evolved, resulting in a distinct soil community.

Indigenous bacterial hosts were used to isolate phages from soil by the enrichment procedure, which was conducted with relative ease, an indication of their abundance in soil. It was hypothesised that the unique community of hosts would mirror the resident phage community. However, when subject to analysis it became apparent that the phage population, although diverse to an extent in their host range, employ similar strategies to aid their survival. In addition, similar traits, such as those illustrated by the one-step growth curve experiments, are not necessarily an indication of lack of diversity. These predominant traits may be due to selective pressures forcing different phages to employ the same survival mechanisms. For example, all phages are capable of infecting more than one host. This polyvalency gives each phage choice of hosts; therefore, if one were to become depleted, they have others in which they can survive. It is in this way that phages are likely to be subject to the pressures of r-K selection – just as their hosts are subjected to them in similar ways, resulting in the observed differences in their survival strategies. In addition, the distinction of highly susceptible and highly resistant indigenous hosts to the resident phages support the findings of previous studies (Holmfeldt *et al.*, 2007). This implies an extreme complexity of phage-host interactions in soil.

4. Phage-host Population Dynamics in Soil.

4.1. Introduction: the balance of phage-host dynamics in soil.

Bacteriophages have been shown to be the main mortality factor of bacteria in many ecosystems (Suttle, 1994). There are currently two broad hypotheses concerning the impact of phages upon their hosts within their natural setting. Firstly, viruses limit the productivity of the host community by lysis, for example, evidence indicates that this accounts for 10-20% of bacterial mortality in marine environments (Fuhrman, 1999). It is thought that selective infection and lysis of predominant bacterial populations control the bacterial community competition and sustain coexistence of bacterial species - "killing the winner" (Thingstad, 2000; Thingstad, 1997). Indeed, phage-host populations have been shown to follow the typical predator-prey oscillations throughout all ecological systems (Hennes, 1995; Wommack *et al.*, 1999). Secondly, viral lysis is important in maintaining genetic diversity of host populations (Waterbury and Valois, 1993). However, it is evident that an equilibrium is sustained between both populations due to the ease with which phages can be isolated from the marine environment. Despite these theories not being mutually exclusive, it has been proposed that they are too simplistic as they do not account for broad-host range phages or bacterial acquisition of resistance (Riemann and Middelboe, 2002). Indeed, some findings indicate that the extreme host specificity is a consequence of the commonly used single-host enrichment protocol (Jensen *et al.*, 1998; Wichels *et al.*, 2002).

Chemostat experiments have provided indications as to why this observed equilibrium is maintained as they have proven that bacterial development of resistance is a common defence in response to lytic phages (Fischer *et al.*, 2004; Middelboe, 2000). These studies have shown that rapid shifts occur between dominance of phage sensitive clones and that of phage resistant clones. As a consequence, phage resistant clones are

continuously being produced and they subsequently replace sensitive strains when the population is exposed to strong selective pressures from infectious phages. In the marine environment, a complex pattern of resistance and susceptibility has been observed (Moebus and Nattkemper, 1981). Therefore, the acquisition of resistance and its maintenance against phage infection is an integral part of phage-host ecology.

The complexity of actinophage-host interactions in soil ecology has not yet been studied extensively (Weinbauer, 2004; Williams *et al.*, 1987). However, it is proposed that, in the case of the actinomycetes, these systems are complicated further by the unusual lifecycles of their hosts and the soil environment (Burroughs *et al.*, 2000). Due to the decay of viable phages in soil at a rate of 0.1 day^{-1} (Cresswell *et al.*, 1992; Williams *et al.*, 1987), successful phage survival requires either continuous production or latency in the form of lysogeny. Either of these firstly requires successful infection of streptomycete mycelia. The susceptibility to phage infection decreases with age of the receptive mycelium due to higher rates of DNA synthesis or to differences in density or surface receptor molecules (Blokhina *et al.*, 1992; Gopul and Reilly, 1995; Gray *et al.*, 1990). Burroughs *et al* developed a population-dynamic mathematical model to determine the underlying mechanisms of this low susceptibility to phage attack. Their findings concluded that the spatial heterogeneity is an important factor in phage-host dynamics in soil.

Due to these interactions of phage and host to soil, a level of equilibrium between them should be maintained once the populations stabilise. This stability is particularly true of temperate phages, as their relationship with their hosts can be considered to be mutually beneficial due to homoimmunity rather than classical predator-prey relationships. In contrast, predator-prey relationships of lytic phages are expected to follow typical predator-prey oscillations. The examination of temperate phage ϕC31 with *S.lividans* in sterile soil microcosms reveals that the persistent form of phage in sterile, unamendned soil occurs as a prophage. Analysis showed that most infections occurred during the first round of filamentous growth, and any subsequent infections had no effect on lysogenised or uninfected populations. This may account for the equilibrium between free phage and prophage populations (Marsh *et al.*, 1993).

Despite equilibrium, there is also undoubtly a consequence of free phages in soil upon their hosts, either by their population depletion or by influencing their genomic composition. It has been shown that phages are, at least in part, responsible for driving bacterial evolution, and therefore attribute to their ability to adapt to new environments (Brussow *et al.*, 2004).

These adaptations result in the ability of microorganisms to be suited to their ever-changing surrounding environments. There are, however, no studies examining the potential of streptomycetes to adapt to an environment other than soil, which is most likely due to their infrequent responsibility for causing infectious diseases in animals (Dunne *et al.*, 1998). In contrast, the survival of organisms capable of causing infectious diseases has been studied extensively (Tolba *et al.*, 2008).

This survival of pathogenic bacteria has been examined extensively, for example, on inanimate objects as this has been shown to be a source of many hospital-acquired infections. The longest recorded survival for a nosocomial bacterium: *Klebsiella* species, on a hospital surface was over 30 months (Kramer *et al.*, 2006). The majority of studies devoted to examining *S.aureus,* a highly adaptable organism, are within the healthcare environment (Dancer, 2004). It has been demonstrated that without organic protection the organism fails to survive, however, even with little organic material *S.aureus* is capable of surviving for extended periods, for example, it is able to survive for over 6 months in a solution of PBS and hospital dust (Wagenvoort and Penders, 1997). On common hospital surfaces, such as plastic polymers, *S.aureus* can survive for up to 7 months (Wagenvoort *et al.*, 2000).

Such strong survival rates in very diverse environments may also apply in the non-healthcare setting (Kowalski *et al.*, 2005; Tolba *et al.*, 2008). There have been a few studies attempting to reveal the sources of community-acquired infections; these examine the survival of *S.aureus* outwith the medical setting. A recent study discussed the survival limitation of *S.aureus* in water, especially where more frequent outbreaks were occurring in healthy athletes who shared swimming pools, whirlpools and shower rooms. These were transient environmental reservoirs that permitted the colonisation of new, healthy hosts with minor skin traumas. The study found that chlorinated

(swimming pool), river and seawater MRSA was capable of surviving despite the harsh environment, and that the higher the salt concentration, the longer the organism survived (up to a maximum of 16 days (Tolba 2008)). The increasing outbreaks of antibiotic resistant infections within the community require more study on their source and survival. In addition, the effect of phages on these populations would aid in the development of a treatment of contaminated substances.

4.2. Aims of the chapter

The presence of large populations of bacteriophages has been shown to have an impact upon their hosts (Wiebe and Liston, 1968). The aim of the following Chapter is to determine the effect of phages upon a host within its natural environment. Clearly, phages do not totally eliminate their hosts as it is possible to isolate them along with their pathogenic phages with relative ease (see Chapter 3).

The hypothesis in this chapter is that indigenous bacteriophages and hosts will both survive well in soil microcosms whereas introduced non-indigenous phages and hosts will not. Additionally, the impact of lytic phages upon their hosts is hypothesised to be more dramatic than that of temperate phages. Using soil microcosms, these hypotheses were tested.

4.3. Survival of indigenous and non-indigenous bacteriophages in soil

Phage viability is a critical factor influencing their survival. Bacteriophage stability is greatly dependant upon factors of the soil, such as pH, water content, temperature, hydrostatic pressure, radiation, the ionic environment, oxygen, organic matter and host availability (Williams *et al.*, 1987). Studies have demonstrated that phage are completely inactive after five weeks of inoculation (Williams *et al.*, 1987) or at a rate of 0.1^{-1} per day (Cresswell *et al.*, 1992). Therefore, the constant requirement to find viable hosts is mandatory for phage survival.

In order to assess how phages within this study survive in soil ecosystems, microcosm experiments were established. This involved 10g of Erskine sterile soil containing 0.1% chitin and 1% starch spiked with host bacteria and phages. The

purpose of these experiments was to clarify the impact of phages upon their hosts within their natural environment.

Bacteriophages are not capable of surviving without a host. However, viability is an important factor in phage persistence. Therefore, the first experiment was planned in order to answer the question of how long they remain viable in soil by examining phage populations alone in sterile soil. ϕC31 was added to the soil and monitored using direct plate count (see section 2.12.1) over a period of a week. Due to the lack of a viable host being present, as was expected, its population dropped. This population remained stable for the first day following the inoculation, however, subsequently it dropped dramatically to barely detectable levels.

This experiment was also conducted using the indigenous soil phage ϕELB8 as it is also temperate and, as with ϕC31, its population also dropped within a week. However, it did survive slightly better than ϕC31. Unlike ϕC31, this phages' population began to decline the day after inoculation, although, its population did not drop as rapidly as that of ϕC31 (Figure 19).

The above experiments demonstrate that phages lose their viability in sterile soil within a week. Additionally, as predicted, the indigenous soil phage remains viable for longer than the introduced phage. This may be due to the indigenous phage having adapted and being more resilient to destructive factors, such as desiccation and adherence to clay particles, within the soil than the introduced strain. Alternatively, the indigenous phage may have formed its own niche within the soil thereby sustaining its own survival.

In addition, phage populations were monitored in microcosms containing a viable host (Figure 19). An increase by approximately 10% of free phage population occurred at the start of the incubation period, most likely due to initial germination of the host population. Following this, the population declined to barely detectable numbers. This rapid drop in phage stability differs from the figures proposed in previous studies (Pantastico-Caldas *et al.*, 1992). Presumably, it is factors specific to the individual soil and phage, such as pH and ionic charge that can account for this difference. However, overall there was no significant difference in survival between an

indigenous soil bacteriophage and ϕC31. This may be due to both phages being similar as they were originally isolated from soil and infect many of the same hosts (see Chapter 3).

4.4. Survival of streptomycetes with free virulent phages in soil

Due to the drop in phage populations observed within a week, subsequent soil microcosm experiments did not monitor phage populations. Only host populations were used as an indication of the phage-host dynamics in soil (Figure 19). Control experiments involved monitoring the host alone in soil over a period of four weeks.

In order to assess the survival of non-indigenous hosts in sterile soil with exposure to viable free phages, two individual soil microcosms were inoculated with *S.avermitilis* and phages, ϕC31 – a non-indigenous phage or ϕELB8, an indigenous phage (Figure 19). Both of these phages produce clear plaques on *S.avermitilis* lawns indicating that they are purely lytic within this host. Therefore, it was predicted that the presence of phages in soil would decrease the resident host population.

In the absence of phage, the *S.avermitilis* population remained relatively stable for over four weeks. However, populations behaved differently in the presence of phages. In both samples, host populations initially dropped as a result of phage infection of newly germinated spores. Bacterial populations were consistently lower than the phage free sample, as a general trend the populations recovered and remained stable one week after inoculation. This finding support previous studies, which stated that following an initial drop in population, subsequent recovery and stabilisation followed (Burroughs *et al.*, 2000; Herron and Wellington, 1990; Holmfeldt *et al.*, 2007). Interestingly, the indigenous phage had a stronger effect on the resident host population than the introduced phage. This was probably due to the adaptation of the indigenous phage to the soil; as a consequence, it had a more dramatic effect on the resident host than the introduced phage.

Following this, the survival of an indigenous host strain, SPO4, chosen due to its susceptibility to different phages, was monitored in sterile soil. In soil alone, this bacterium displayed an initial increase in population on day 2, which dropped back to

the initial inoculum concentration and was relatively stable for the duration of the experiment, with slight fluctuations. In the sample containing ϕC31 and ϕELB8, a similar trend occurred, however, the initial spike in population was not seen. Therefore, it can be concluded that this was presumably due to the lysis of newly germinated spores which reduced the host population, thus restraining the spike in population. An unexpected observation in these two experiments was the similarity in the effect of phage presence whether, it was due to a purely lytic or temperate phage lifecycle. By comparing overall survival of *S.avermitilis* (an introduced soil host) and SPO4 (a native soil host) in phage presence, it became evident that, overall, the indigenous strain was more successful at resisting a depletory effect by phage lysis. Therefore, it can be concluded that indigenous soil strains have adapted to the soil and evolved coping mechanisms in the presence of phages.

4.5. The impact of ϕELB20 and ϕR4 on non-indigenous and indigenous soil hosts in soil.

Due to the similarities revealed in experiments of Chapters 3 and 5, the impact of ϕELB20, an indigenous phage, and ϕR4 on an indigenous host and non-indigenous host were examined in soil microcosm experiments. It was hypothesised that these phages would produce the same effect on each of their hosts. SPO3 and *S.avermitilis* were chosen as hosts as in both of them phages exhibit lytic lifestyles.

In the soil containing *S.avermitilis,* both phages restricted an increase in host population observed in the virus free microcosm on day 5 Additionally, as expected, both phages exerted almost identical influence on the host population by causing a slight, overall decrease.

The effect of these two phages was also observed on indigenous soil strain SPO3. Like *S.avermitilis,* the population of this host was also affected by phage presence, however, unlike the non-indigenous strain, this population increased in the presence of lytic phages. Indeed, there was an apparent spike in host population in the first few days of inoculation, despite a drop being expected, as observed in the control phage free sample. Therefore, it can be deduced that phage presence somehow confers

an advantage to the indigenous host population. Overall, it can be concluded that, as expected, ϕELB20 and ϕR4 have similar effects on their host populations.

4.6. *S.aureus* and bacteriophage survival in the environment

The impact of phages upon pathogenic bacteria has been well documented, however, this has not been examined in a natural setting (Sulakvelidze *et al.*, 2001). *S.aureus* typically colonises the skin, however, clearly it also survives outwith this setting, for example, in such environments as hospitals (Tolba 2008). Previous studies found that *S.aureus* could survive for long periods of time; however, all of those studies focused on the hospital environments (Neely and Maley, 2000; Wagenvoort and Penders, 1997). Evidently, *S.aureus* is capable of surviving in a wide variety of environments. Therefore, soil microcosm studies were also conducted on this bacterium and its virulent bacteriophage ϕ858. As a result, it was possible to compare the survival of indigenous soil bacteria with adaptable, non-indigenous bacteria: both in the presence of bacteriophages (Burnet and Lush, 1935; Hotchin, 1951). Not being indigenous to soil, it was hypothesised that both virus and host would not survive well in this environment; surprisingly, they did survive well for the duration of the experiment.

In soil alone, the phage population, as was predicted, declined gradually in the absence of a host. In host presence, the phage population increased on days one and two, after which it sharply declined. However, when compared to the actinophages in soil, this non-native phage survived significantly better than phages originally isolated from soil.

The survival of *S.aureus* in soil was unexpectedly successful. Indeed, it was more successful in soil than indigenous soil organisms themselves. This illustrates the high adaptability of this organism. In the presence of a lytic bacteriophage, a significant impact upon the population was observed, however, despite this the host maintained a steady population. The experiment showed that *S.aureus* and a lytic phage could survive well in soil, in fact, as well as the indigenous Streptomycetes. In addition, the presence of bacteriophage, despite it being lytic did not have a dramatic effect upon its host in this environment.

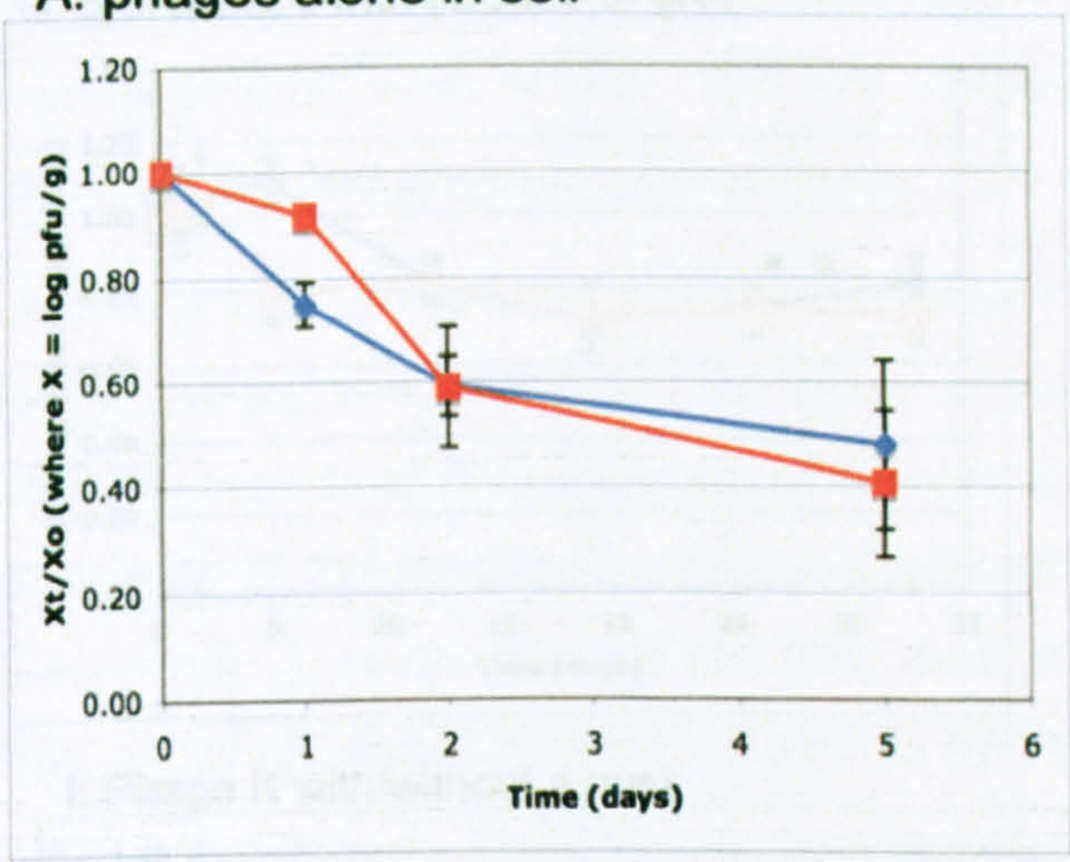
The conducted experiments supported the findings of previous studies (Ashelford *et al.*, 2000; Burroughs *et al.*, 2000; Herron and Wellington, 1990; Marsh *et al.*, 1993; Weitz and Dushoff, 2008) that phages and hosts attain equilibrium resulting in survival of both within a shared niche. It has been shown in previous studies that phages can co-exsist with hosts at lower densities (Weitz and Dushoff, 2008). This finding was also evident in these experiments. However, it was observed that early increases in host populations due to germination and sporulation of the bacteria were restrained in the presence of phages. In summary, the trend in host populations was stabilisation irrespective of phage presence or in the nature of the phage (lytic or temperate).

Within soils containing temperate phages it is probable that this observed stability was due to the formation of lysogens, which not only survived but expressed resistance by homoimmunity. However, it is also worthy of noting that previous studies have observed a compromise in survival of lysogens (Marsh *et al.*, 1993). In spite of this, the carriage of a prophage would not have completely suppressed the proliferation of host bacterium.

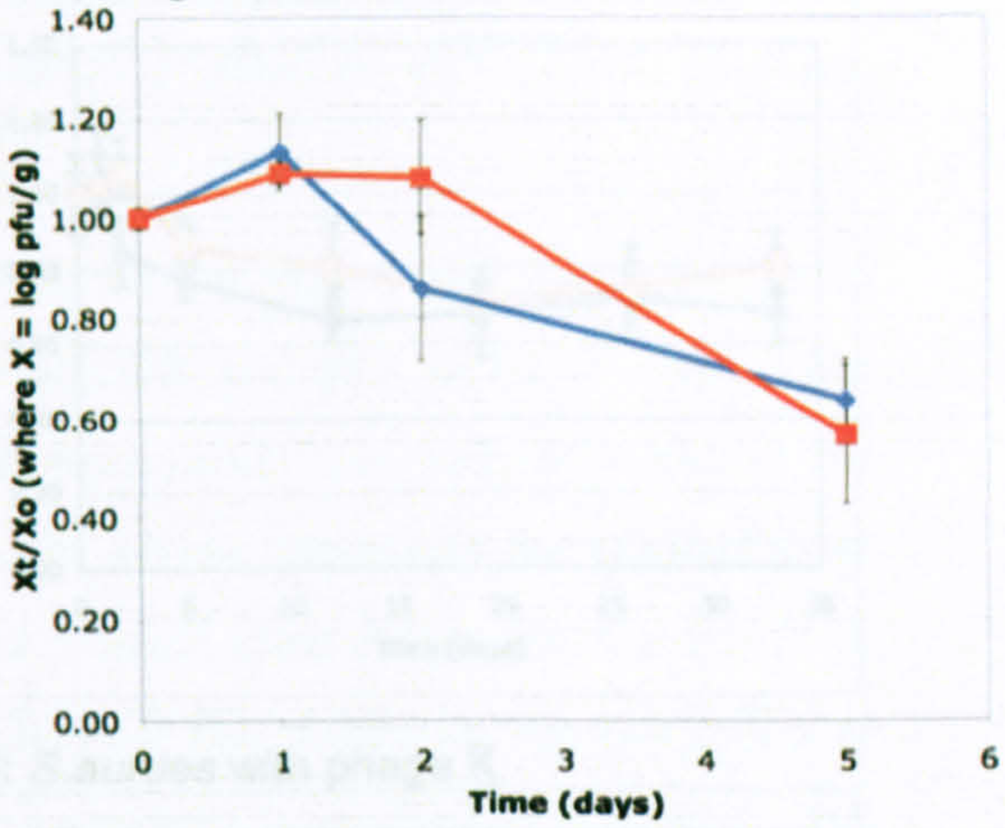
The reason for the stability observed in soils containing lytic phages is less clear, although soil characteristics, such as pH will have had an impact on phage survival and viability. The minimal impact observed in populations in presence of lytic phages may have been due to the development of resistance within the resident bacterial community, which in some cases, as in that of SPO3 and ϕELB20, actually increased the host population by triggering increased survival in the hosts.

In addition, the high adaptability of *S.aureus* and its lytic phage in soil was clearly illustrated in this Chapter. Indeed, in several instances, its survival superseded that of native soil organisms and their phages. The reasons for such good survival in soil are unclear. *S.aurues* is evidently an adaptable organism, however, it is not typically indigenous to soil. Its resilience even in this hazardous environment illustrates why this organism is such a problem in hospital acquired infections.

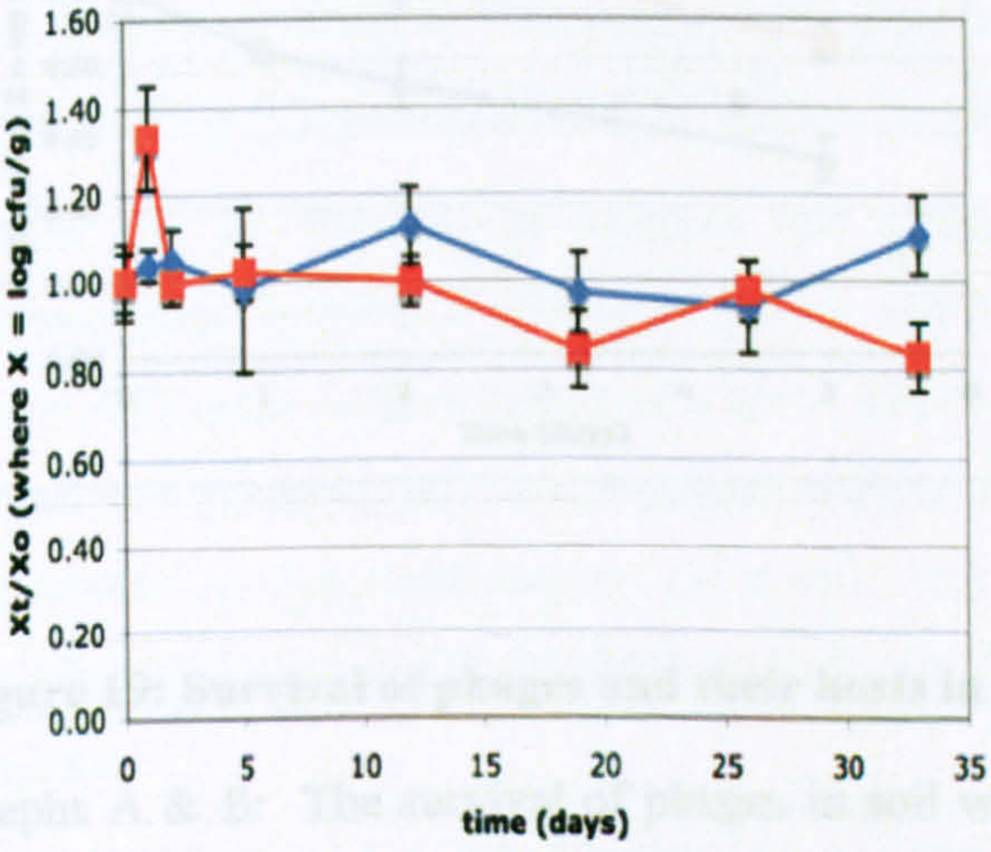
A: phages alone in soil



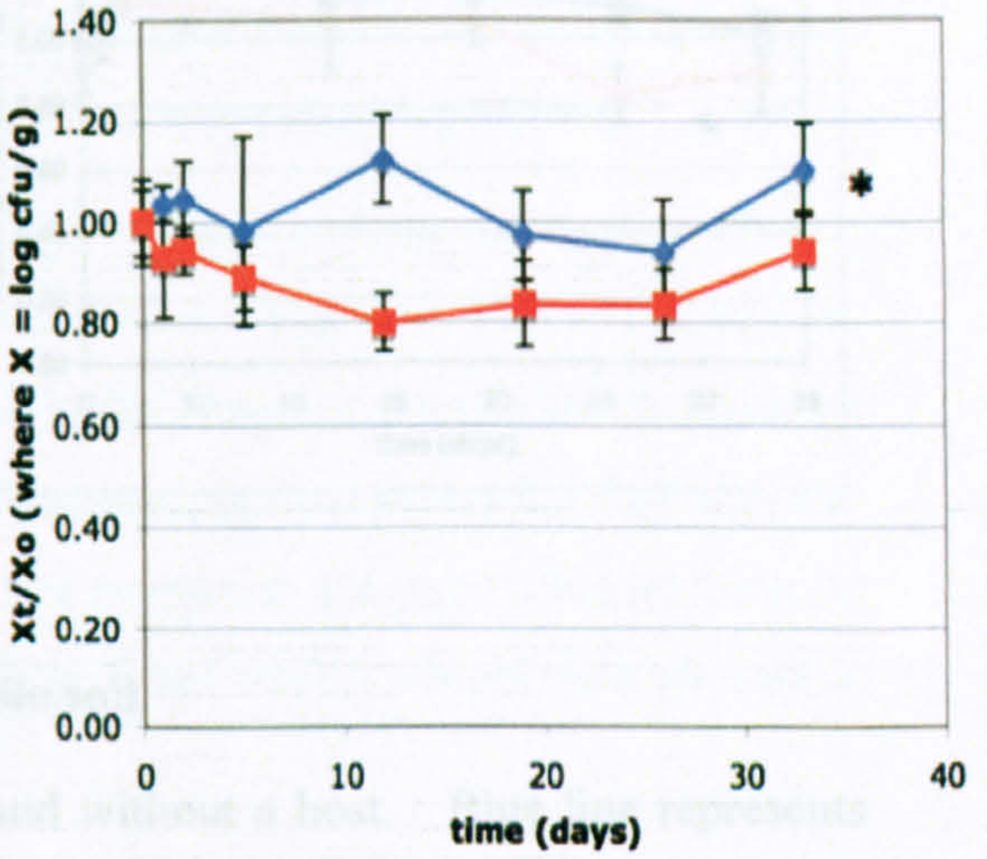
B: phages with a host in soil



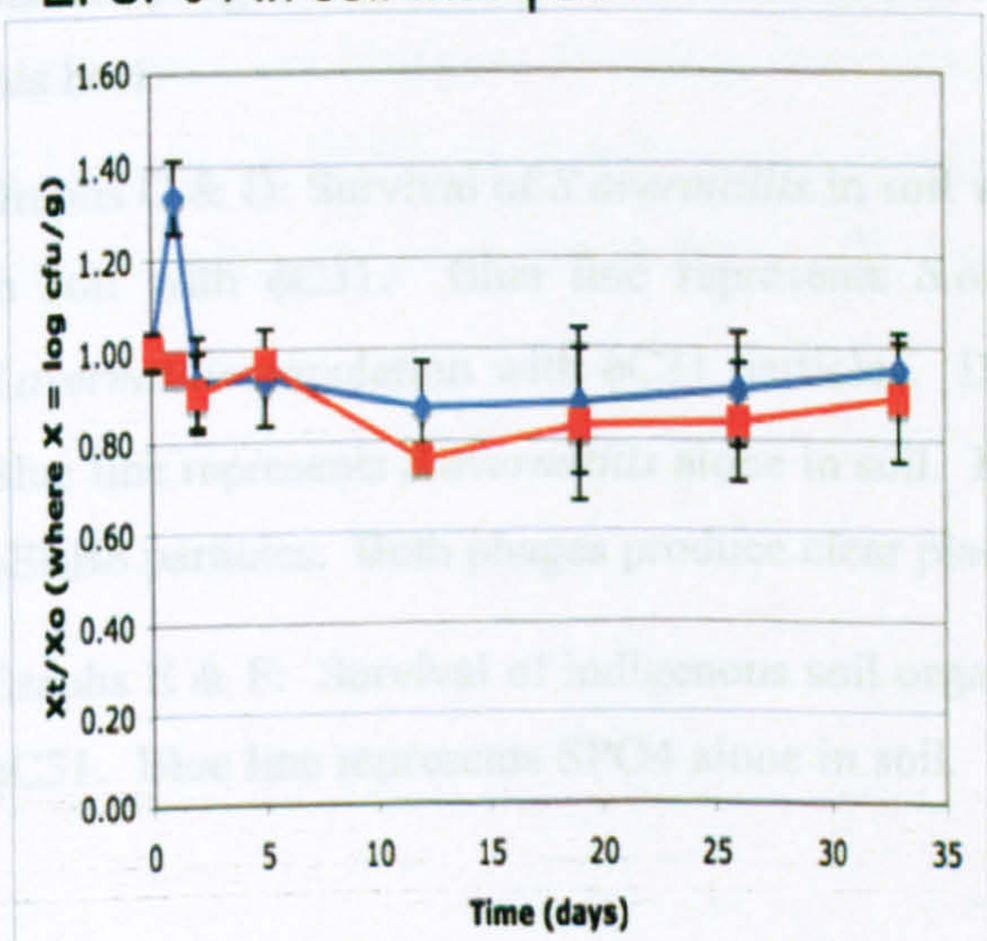
C: *S.avermitilis* with ϕC31



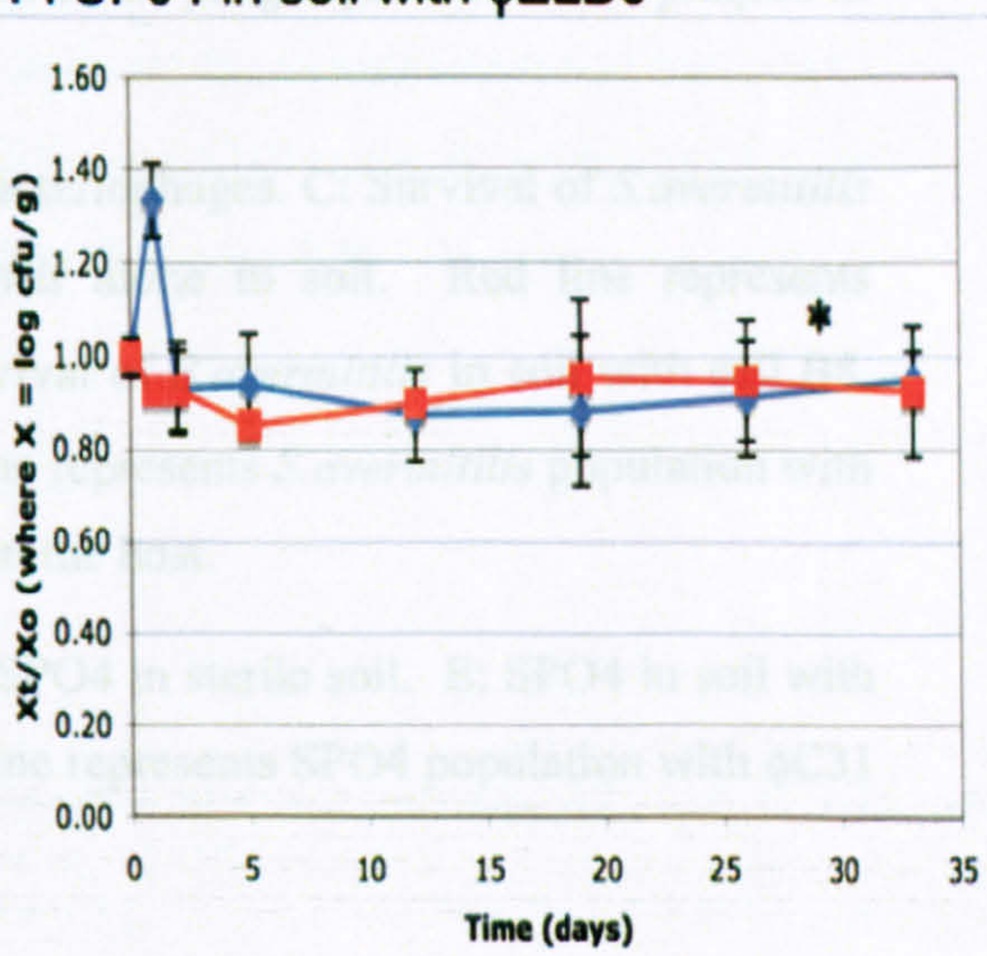
D: *S.avermitilis* with ϕELB8

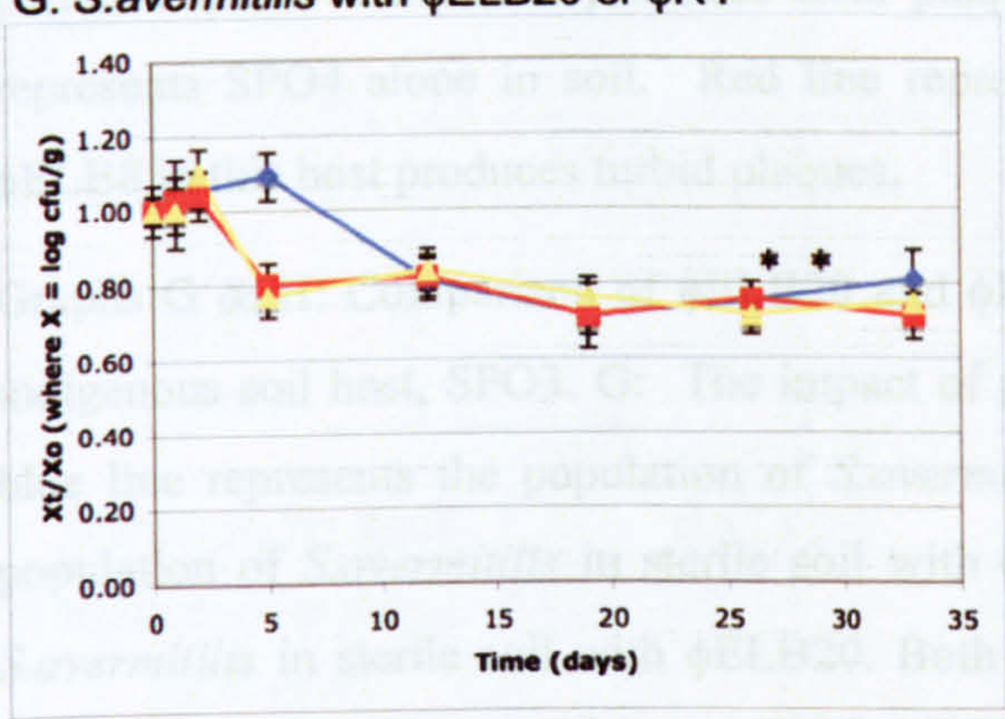


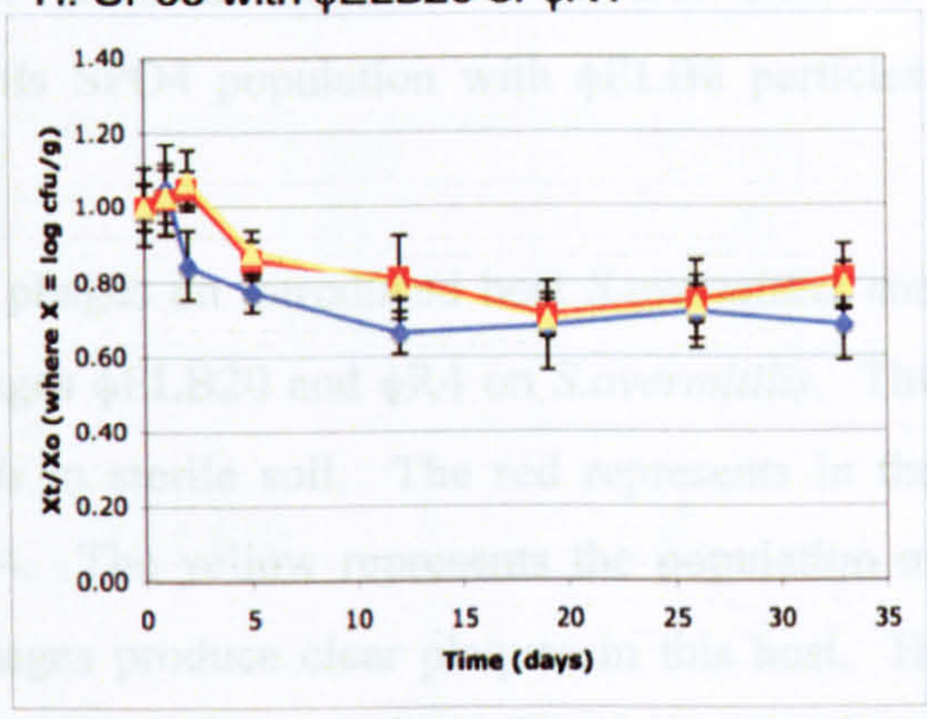
E: SP04 in soil with ϕC31

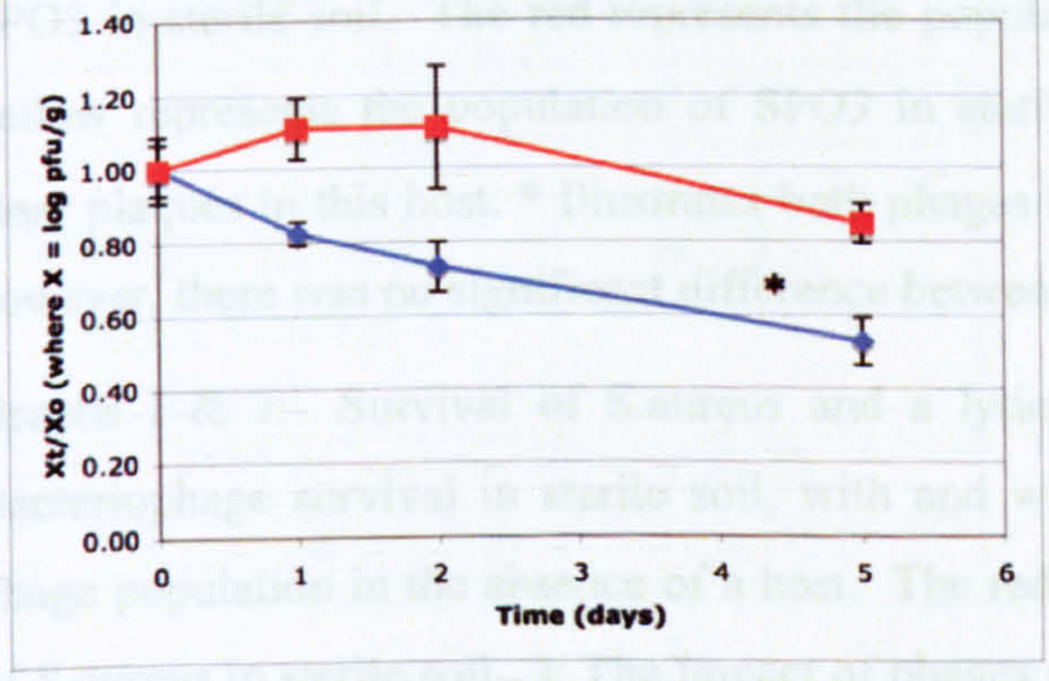


F: SP04 in soil with ϕELB8









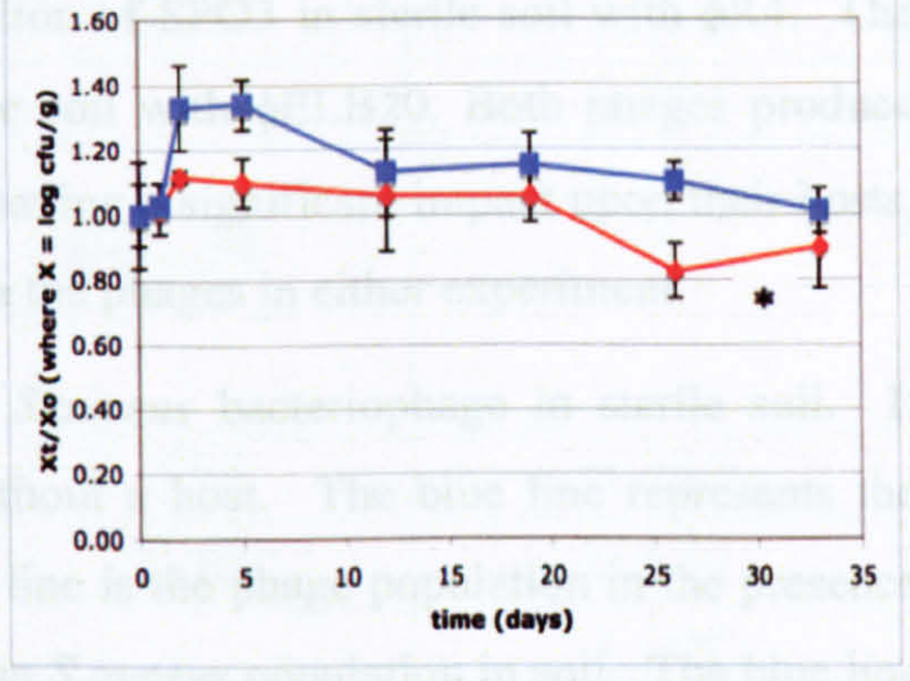


Figure 19: Survival of phages and their hosts in sterile soil.

Graphs A & B: The survival of phages in soil with and without a host. Blue line represents ϕC31. Red line represents ϕELB8. A: Bacteriophages in soil in the absence of a host. B: Bacteriophages in soil with *S.avermitilis* as a host. Both of the phages produce turbid plaques in this host.

Graphs C & D: Survival of *S.avermitilis* in soil with bacteriophages. C: Survival of *S.avermitilis* in soil with ϕC31. Blue line represents *S.avermitilis* alone in soil. Red line represents *S.avermitilis* population with ϕC31 particles. D: Survival of *S.avermitilis* in soil with ϕELB8. Blue line represents *S.avermitilis* alone in soil. Red line represents *S.avermitilis* population with ϕELB8 particles. Both phages produce clear plaques on the host.

Graphs E & F: Survival of indigenous soil organism SPO4 in sterile soil. E: SPO4 in soil with ϕC31. Blue line represents SPO4 alone in soil. Red line represents SPO4 population with ϕC31

particles. ϕC31 in this host produces clear plaques. F: SPO4 in soil with ϕELB8. Blue line represents SPO4 alone in soil. Red line represents SPO4 population with ϕELB8 particles. ϕELB8 in this host produces turbid plaques.

Graphs G & H: Comparison of ϕELB20 and ϕR4 phages on introduced host *S.avermitilis* and indigenous soil host, SPO3. G: The impact of phages ϕELB20 and ϕR4 on *S.avermitilis*. The blue line represents the population of *S.avermitilis* in sterile soil. The red represents in the population of *S.avermitilis* in sterile soil with ϕR4. The yellow represents the population of *S.avermitilis* in sterile soil with ϕELB20. Both phages produce clear plaques in this host. H: The impact of phages ϕELB20 and ϕR4 on SPO3. The blue line represents the population of SPO3 in sterile soil. The red represents the population of SPO3 in sterile soil with ϕR4. The yellow represents the population of SPO3 in sterile soil with ϕELB20. Both phages produce clear plaques in this host. * Illustrates both phages having a significant impact upon their hosts, however, there was no significant difference between the phages in either experiment.

Graphs I & J: Survival of S.aureus and a lytic *S.aureus* bacteriophage in sterile soil. I: Bacteriophage survival in sterile soil, with and without a host. The blue line represents the phage population in the absence of a host. The red line is the phage population in the presence of *S.aureus* in sterile soil. J: The impact of phages on *S.aureus* population in soil. The blue line represents *S.aureus* population in sterile soil. The red line represents the population of *S.aureus* alongside a virulent lytic bacteriophage in sterile soil. The phage used produces clear plaques in *S.aureus*.

Soil microcosms were inoculated with 1×10^4 viral particles, and, where appropriate, 1×10^4 host bacteria. Phage populations were monitored over 5 days, host populations were monitored over 33 days. Readings were taken in triplicate by extraction using method C of Lanning & Williams. * indicates a significant difference between populations ($p <$ or equal to 0.05) (see matrices in Figure 20, Figure 21 and Figure 22)

Hosts in soil without phages	*S.avermitilis*	SP03	SP04	*S.aureus*
S.avermitilis	1.000	0.000	0.000	0.0001
SP03		1.000	0.057	0.000
SP04			1.000	0.000
S.aureus				1.000

Phages in soil alone (no host present)	ϕC31	ϕELB8	K
ϕC31	1.000	0.804	0.532
ϕELB8		1.000	0.272
K			1.000

Phages in soil with viable hosts	ϕC31	ϕELB8	K
ϕC31	1.000	0.688	0.136
ϕELB8		1.000	0.131
K			1.000

Figure 20: Matrices of phage and host soil microcosm experiments displaying the p values deduced from ANOVAs (p< or equal to 0.05 is significant).

Host populations with lytic phages present	*S.avermitilis*	*S.avermitilis* plus ϕC31	*S.avermitilis* plus ϕELB8	SP04	SP04 plus ϕC31	SP04 plus ϕELB8	*S.aureus*	*S.aureus* and phage K
S.avermitilis	1.000	0.254	0.000	0.000	NA	NA	0.000	NA
S.avermitilis plus ϕC31		1.000	0.708	NA	0.784	NA	NA	0.595
S.avermitilis plus ϕELB8			1.000	NA	NA	0.000	NA	0.000
SP04				1.000	0.784	0.137	0.000	NA
SP04 plus ϕC31					1.000	1.000	NA	NA
SP04 plus ϕELB8						1.000	NA	0.172
S.aureus							1.000	0.000
S.aureus and phage K								1.000

Figure 21: Matrix displaying p values when comparing the effects of lytic phages. (p< or equal to 0.05 is significant) NA is not applicable.

Comparison of ϕELB20 vs. ϕR4	*S.avermitilis*	*S.avermitilis* plus ϕELB20	*S.avermitilis* plus ϕR4	SP03	SP03 plus ϕELB20	SP03 plus ϕR4
S.avermitilis	1.000	0.000	0.000	0.000	NA	NA
S.avermitilis plus ϕELB20		1.000	0.811	NA	0.011	0.242
S.avermitilis plus ϕR4			1.000	0.080	0.008	0.130
SP03				1.000	0.067	0.080
SP03 plus ϕELB20					1.000	1.000
SP03 plus ϕR4						1.000

Figure 22: Matrix comparing the effects of phiELB20 and R4 on different hosts in soil microcosm experiments. (p< or equal to 0.05 is significant). NA is not applicable.

4.7. Conclusions

The conducted experiments demonstrate that host populations are capable of surviving and indeed remaining stable despite sharing its environment with a pathogenic virus in soil. The early fluctuations observed in host populations are dependant upon whether the virus is lytic or lysogenic within the host. These initial differences are most probably due to, in the case of temperate phages, lysogens being formed and surviving whereas, in the case of lytic phages, the early populations of young mycelium being lysed.

Phages do have an impact upon their hosts as effects of this impact are evident in all experiments conducted in this study. There are clearly differences between populations if soils containing only bacteria and ones containing bacteria and phages. Typically, soils without phages exhibit a spike in population on days one and two, most likely due to spoulation; this is evidently restricted or prevented by the presence of phages. However, over a longer period of time the host population stabilises, even in the presence of phages. There are several reasons that can explain this observed result: a low phage population, little or no mixing of the soil, particles physically obscuring phages from their hosts, phage adherence to clay particles or host defence systems (see section 1.11). It is these factors that account for the similarity between phage-free and phage inoculated soil microcosms, however, despite this, phages do have impact upon their hosts (see 4.4). Moreover, it is probable that much of that impact of phages in soil has not been reflected in these experiments due to the unobserved genomic rearrangements and the drive of bacterial evolution: factors which cannot be examined by population studies alone. Nevertheless, it is evident by the analysis of bacterial populations in soil that a delicate balance exists between phage and host. In order for both to survive successfully, they must co-evolve mechanisms for adaptation to their surroundings. The dynamics of phage-host populations in soil is complex and subtle, resulting in diverse inhabiting populations.

The above work has demonstrated the adaptability and resilience of *S.aureus* as well as its capability to survive together with its pathogenic bacteriophage in harsh environments. This has implications for the future use of phages as therapeutics. Despite its survival the *S.aureus* phage clearly has an effect upon the bacterial population in soil, which is not as successful in the presence of this phage. Therefore, although it does not completely eliminate it, it is useful in reducing *S.aureus* populations in soil.

The experiments conducted in this chapter clearly illustrate the adaptation and resilience of bacteria and their phages in soil. Although the conclusions derived from this work are well supported in the experiments conducted, there is also a need to emphasise the enforced simplicity of using sterile, filtered, treated soil throughout this work. Within the natural environment, the complete nature of these relationships is likely to be a result of complex systems.

5. Genetic Diversity of Actinophages in Soil.

5.1. Introduction: The study of bacteriophage genetics.

Since the advent of molecular techniques our understanding of host and phage genetics has increased significantly. At present, there are over two hundred entire phage genomes which have been sequenced (Ackermann and Kropinski, 2007). As the database of phage genomes grows, their evolutionary relationships are being revealed. Additionally, the ability to apply molecular biology to bacteriophages has resulted in their exploitation as molecular tools (Karam, 2005).

Early molecular studies of phage genomes involved quantifying the G/C content of the genome primarily as a means of taxonomic classification (Wyatt and Cohen, 1953). In addition to taxonomical applications, determination of G/C content provides information about phage evolutionary relationship with their host and one another. However, there are several instances in which the phage G/C content is not reflected in the host genome, for example, the *M.tuberculosis* phage with 63.6% G/C genome whilst its host possesses a 65.6% G/C. An even more significant example is the difference in genomes between *P.aeruoginosa* (66.6%) and its phages (54.4%) (Kwan *et al.*, 2006). Lateral gene transfer can explain why these differ between host and phage. In addition, these G/C differences may be a result of the recent development of a phage's ability to infect a host. Clearly, obtaining the G/C content of the phage genome does provide useful information for taxonomy and aids clarification of phage-host interactions.

Before DNA sequencing not only was G/C content determined but also phage restriction profiling was, and continues to be applied, as it provides information about the phage genome without sequencing. It can reveal information about the genome size, homology to other phages, possession or lack of *cos* ends and phage resistance mechanisms to host defences. As restriction is a significant barrier to plaque formation,

it directly has an effect on EOP. Therefore, restriction enzymes made by the host species are often used to screen phage DNA in order to detect resistance mechanisms. For example, 38 out of 70 restriction enzymes screened for *Streptomyces* phage FP43 DNA failed to cut (Hahn, 1991). This was, in part, due to some of those being A/T rich enzymes, evidently not able to restrict a G/C rich genome. When enzymes specifically derived from *Streptomyces* spp 46% failed to digest the DNA, it was later discovered that this phage possessed a mechanism through which it managed to avoid digestion by host defences. This occurred as a consequence of a lack in target sites for restriction enzymes produced by their hosts. The same study concluded from restriction profiling alone that the genome was approximately 54kb in size. Based on this technique, primitive but useful genomic maps were created which, when combined with other experiments, such as knock out mutant formation, resulted in the mapping of phage functions.

It is the integrase of temperate phages that is often the focus of much attention while developing phages as molecular tools. As a result, the NCBI sequence database contains numerous such genome regions, without the remainder of the genome having been sequenced. One such example is that of the wide-host range actinophage ϕR4, whose *attP* site was cloned and sequenced along with the attachment sites in the host chromosome. This allowed for site-specific integration of DNA into hosts (Shirai *et al.*, 1991).

A more recent study has revealed that it is ORF469 in the phage genome that is a specific-site recombinase responsible for the integration reaction in ϕR4. The exploitation of this resolvase-DNA invertase, has resulted in the development of molecular tools, which are now applied to several fields, including mammalian cells (Matsuura *et al.*, 1996).

The first phage genome which was sequenced in its entirety was that of ϕX174, published in *Nature* in 1977 (Sanger *et al.*, 1977). This virus has a small genome of only eleven genes with the entire genome being only 5.4Kb long and its sequencing began the revolution in phage genome sequencing, now a common technique in phage biology. Of all the actinophages, the ϕC31 genome, as already described, is the best-

characterised. However, many other actinophage genomes have now been sequenced and each contributes to our understanding of phage survival, evolution and host interactions.

Additionally, until recently, all of the characterised actinophages had *cos* ends. However, the most recently published actinophage genome is that of μ 1/6 which does not have them (Figure 23), consequently, its genome is likely to replicate bidirectionally, much like that of bacteriophage T7 (Doublie, 1997). It is a high G/C content, narrow host range phage, with a 38kb genome to which 52 overlapping ORFs have been assigned. Due to the ever-expanding database of genome sequences available, it is now customary for related sequences to be identified by shared homologous regions with other phages. Similarly, the evolutionary relationships have been revealed. In the case of μ 1/6, for example, homology with *Streptomyces* phages VWB, ϕC31 and ϕBT1 has been identified. However, it is not only homology between sequences which can reveal relationships but also examination of the mosaic structure of entire genomes. Genome architecture of this phage strongly reflects that of the three other sequenced *Streptomyces* spp phages. Consequently, it can be concluded that although there was some sequence similarity, it was genome structure that was highly conserved in these phages (Farkasovska *et al.*, 2007).

These relationships between phages have been detected in viruses originating for various geographical locations, however, there are few studies of small-scale bacteriophage communities. One of which examines the virulent soil phages of *Bukholderia cepacia* (Summer *et al.*, 2006). Throughout this study, four phages were isolated from soil by the enrichment process over a period of several years. Remarkably, three of the four phages were related, much more closesy than any other phages previously examined. There were some single nucleotide polymorphisms,. However, the majority of these were synonymous, making no difference to amino acid sequence. Nevertheless, their complement of encoded proteins did not reflect their high sequence similarity. There appeared to be mosaic orthologues at protein levels yet not at the DNA sequence level. Genes had been discovered which were thought to be the result of extensive deletion or addition of related homologues. It was concluded that this

mosaicism, found in all of the phages, conferred a selective advantage to the phage survival as it allows for smooth and efficient exchange of genes under selective pressures.

Previous studies have revealed a vast amount of common ancestry between phages based on genomic architecture (Hendrix *et al.*, 1999; Pedulla *et al.*, 2003b). The many different combinations of phage gene clusters created by homologous and non-homologous recombination result in the addition of new genes, the creation of new genes, and even the formation of new bacteriophages.

In spite of the fact that the number of phage genome sequencing is growing, it should be remembered that there are a significant number of phages in nature and, consequently, advances made in the understanding of their evolution and relationships remain comparatively elementary. In order to understand the evolution of phages on a global scale, the genetic structure of smaller populations, such as those within soil, must first be derived.

5.2. Aims of the Chapter

Due to the increasing number of published genomes, more is now known about the relationships and evolutions between phages. However, there exists only one study examining the direct relationships between populations of phages from a single soil sample (Summer *et al.*, 2006). The findings in this work have demonstrated that the phages in direct contact with one another are very closely related. Due to the lack of studies focusing on a single environment, this work aimed to reveal the relationships and evolution of the panel actinophages described in Chapter 3, using genome analysis, restricting mapping, Southern blotting and genome sequencing. It is hypothesised that the bacteriophages in this study will share significant areas homology.

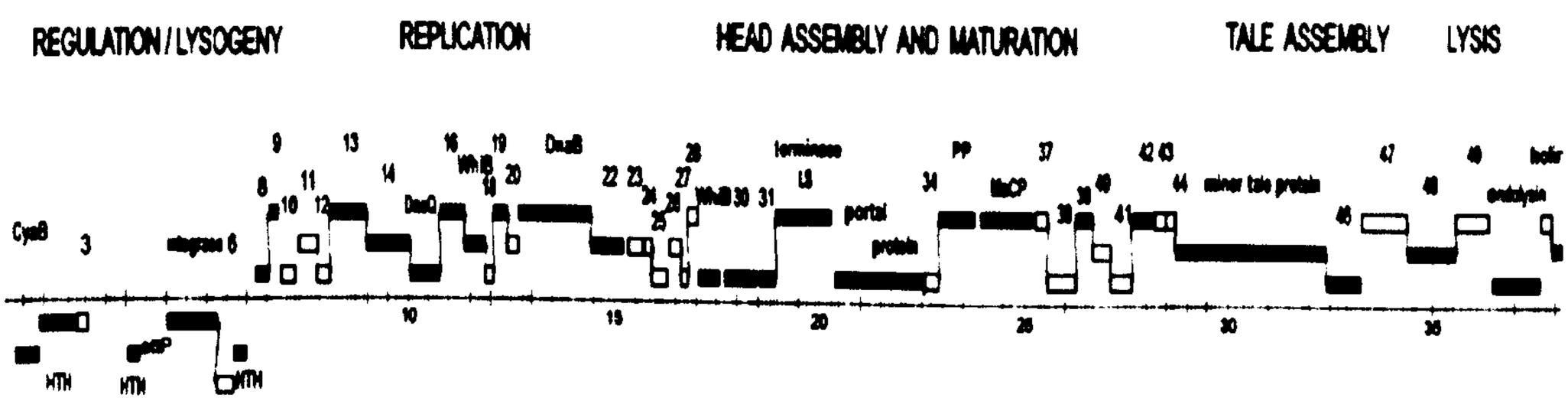


Figure 23 Typical phage genome organisation as displayed by the genome of μ 1/6 (Farkasovska *et al*, 2007). Boxed areas highlight assumed ORFs.

5.2.1. Phage DNA extraction

In order to examine the genomes of the phages in this study, firstly, their DNA had to be extracted. Currently there are many protocols for phage DNA extraction causing some confusion with regards to which is most efficient (Dale and Greenaway, 1984). Prior to embarking upon this study, several protocols for phage DNA extraction were tested in order to assess which was most suitable. Of the four techniques examined, one proved most useful. Optimal DNA isolation was achieved by purifying the phages using a Caesium chloride gradient and then subjecting the particles to high salt concentrations, followed by phenol chloroform treatment in accordance with Hopwood (Hopwood *et al.*, 2000). Despite several attempts and modifications to protocols, (such as the addition of phenol chloroform steps and the use of higher titres than the protocol required,) high quantities of clean phage DNA could not be isolated by those means. In spite of it being studied extensively at the beginning of the project. Therefore, it was concluded that there was no alternative to the large-scale extraction protocol as described by Hopwood *et al* (Figure 24).

5.2.2. Restriction mapping of phage DNA

Prior to the development of genome sequencing the primary tool in genome mapping of viruses was the mapping of phage genomes according to their restriction profiles (Hahn, 1991). Clearly, without the DNA sequences of the indigenous bacteriophages, screening for digestion by a library of restriction enzymes was required. Initial studies were solely conducted on ϕC31, ϕELB17. ϕELB18, ϕELB19 and ϕELB20. Using the restriction profile of ϕC31, and a high G/C target sequence as guide, enzymes were chosen. The majority of the restriction enzymes did restrict the DNA (Table 22). It was suspected that the indigenous phages would display some resistance high G/C content restriction enzyme, in particular *Streptomyces* spp enzymes. Overall, the majority of restriction enzymes with target sites of high G/C content were unable to digest the phage DNA. This avoidance technique is typically found in bacteriophages as

it aids their survival. It is essential that bacteriophages are able to infect their hosts successfully. Consequently, it was predicted that a resistance to G/C rich restriction enzymes such as those produced by *Streptomyces* would be exhibited. However, this did not result as many G/C rich enzymes digested the phage DNA successfully (Table 22).

In order to produce restriction maps of ϕELB17, ϕELB18, ϕELB19 and ϕELB20, restriction enzymes were screened. Of these, several enzymes did successfully digest the phage DNA. However, the majority restricted the DNA too frequently to prove useful. NcoI was deemed appropriate as its use resulted in clear restriction patterns (Figure 25). The resulting profiles, derived by NcoI digestion, did not demonstrate any similarity to ϕC31. Nevertheless, all four of the phages isolated from the soil had remarkably similar profiles (Figure 25). From these profiles, it was possible to estimate their genome sizes as ϕELB17, ϕELB18, ϕELB19 and ϕELB20 at 55.3Kb, 49.3Kb, 41.3Kb and 47Kb respectively. All had a genome size of around 50Kb, which was comparable to 41.5Kb genome of ϕC31.

The experiments carried out for this study illustrated that the genome size of each bacteriophage was approximately 50kb, which is common for many phages. Additionally, examination of their banding patterns revealed four of the phages to be related.

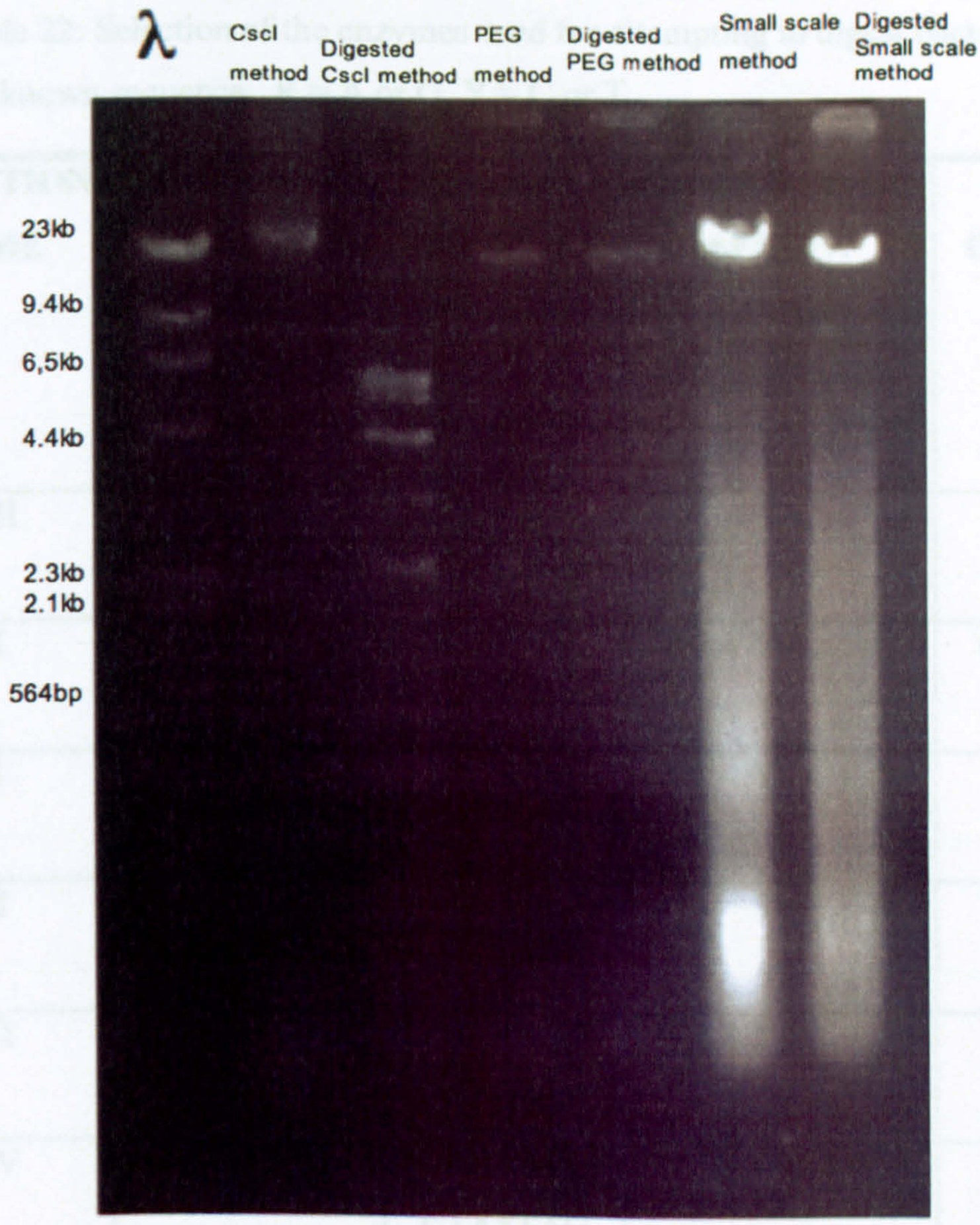
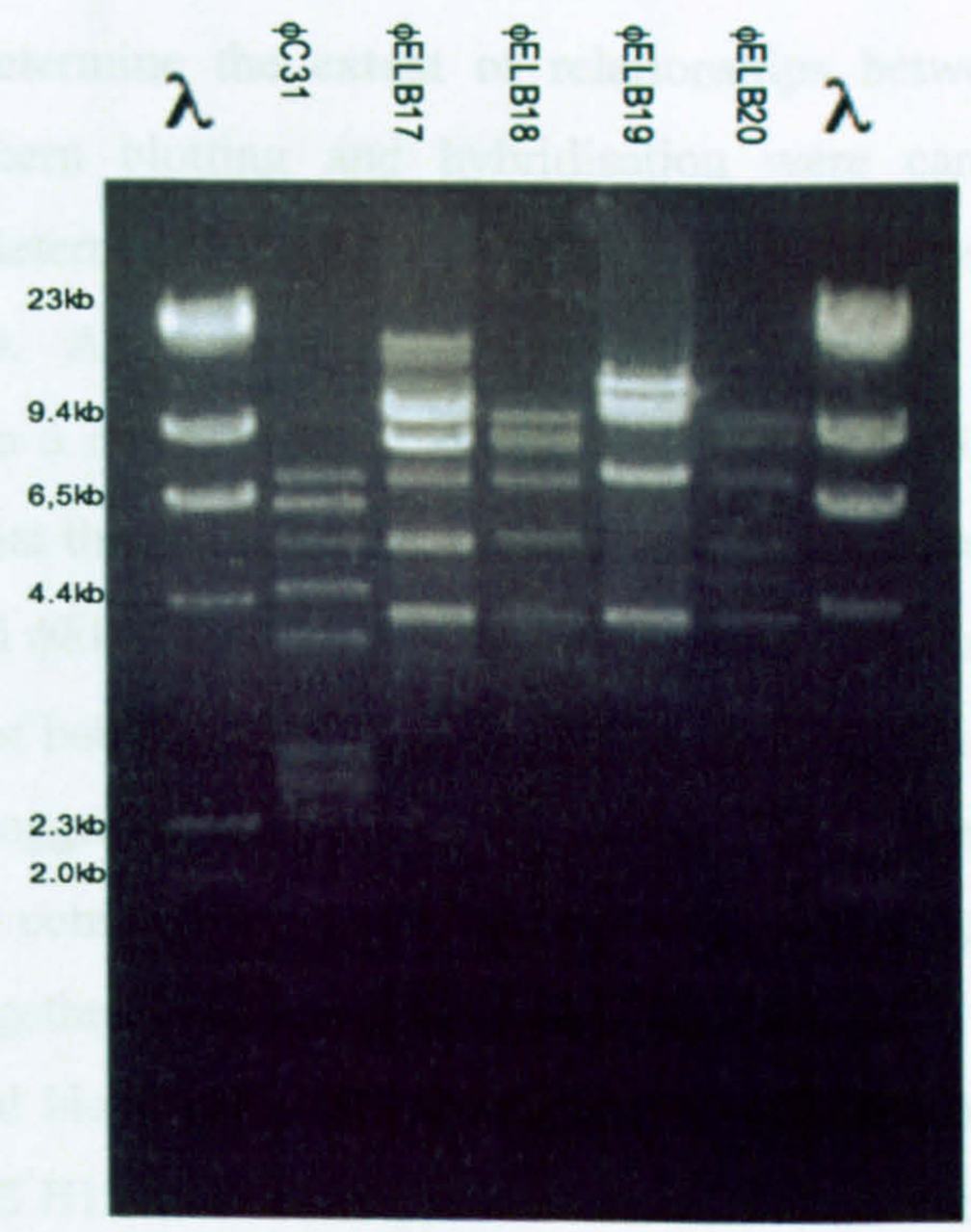


Figure 24: Comparison of phage DNA extraction protocols using only ϕC31 DNA. DNA digested was with NcoI. Refer to section 2.17 in methods for protocol.

Table 22: Selection of the enzymes used for attempting to digest bacteriophage DNA of unknown sequence. R = A or G, Y = C or T

RESTRICTION ENZYME	**SOURCE ORGANISM**	**TARGET SEQUENCE**	**DIGESTION OF ϕELB17. ϕELB18, ϕELB19 AND ϕELB20**	**GC CONTENT (%)**
BamHI	*Curtobacterium*	G*GATCC CCTAG*G	No	67
BanII	*Paenibacillus*	GRGCY*C C*YCGRG	Yes	67 – 100%
BglII	*Bacillus*	A*GATCT TCTAG*A	Yes	33
BstBI	*Geobacillus*	TT*CGAA AAGC*TT	Yes	33
EcoRI	*Escherichia*	G*AATTC CTTAA*G	Yes	33
EcoRV	*Escherichia*	GAT*ATC CTA*TAG	No	33
HindIII	*Haemophilus*	A*AGCTT TTCGA*A	No	33
KpnI	*Klebsiella*	GGTAC*C C*CATGG	Yes	67
MluI	*Micrococcus*	A*CGCGT TGCGC*A	Yes	67
NcoI	*Gordinia*	C*CATGG GGTAC*C	Yes	67
NdeI	*Nesseria*	CA*TATG	No	33

		GTAT*AC		
PciI	*Planococcus*	A*CATGT TGTAC*A	Yes	33
PvuI	*Proteus*	CGAT*CG GC*TAGC	Yes	67
PvuII	*Proteus*	CAG*CTG GTC*GAC	Yes	67
SalI	*Streptomyces*	G*TCGAC CAGCT*G	Yes	67
ScaI	*Streptomyces*	AGT*ACT TCA*TGA	Yes	33
SphI	*Streptomyces*	GCATG*C C*GTACG	No	67
XbaI	*Xanthomonas*	T*CTAGA AGATC*T	No	33



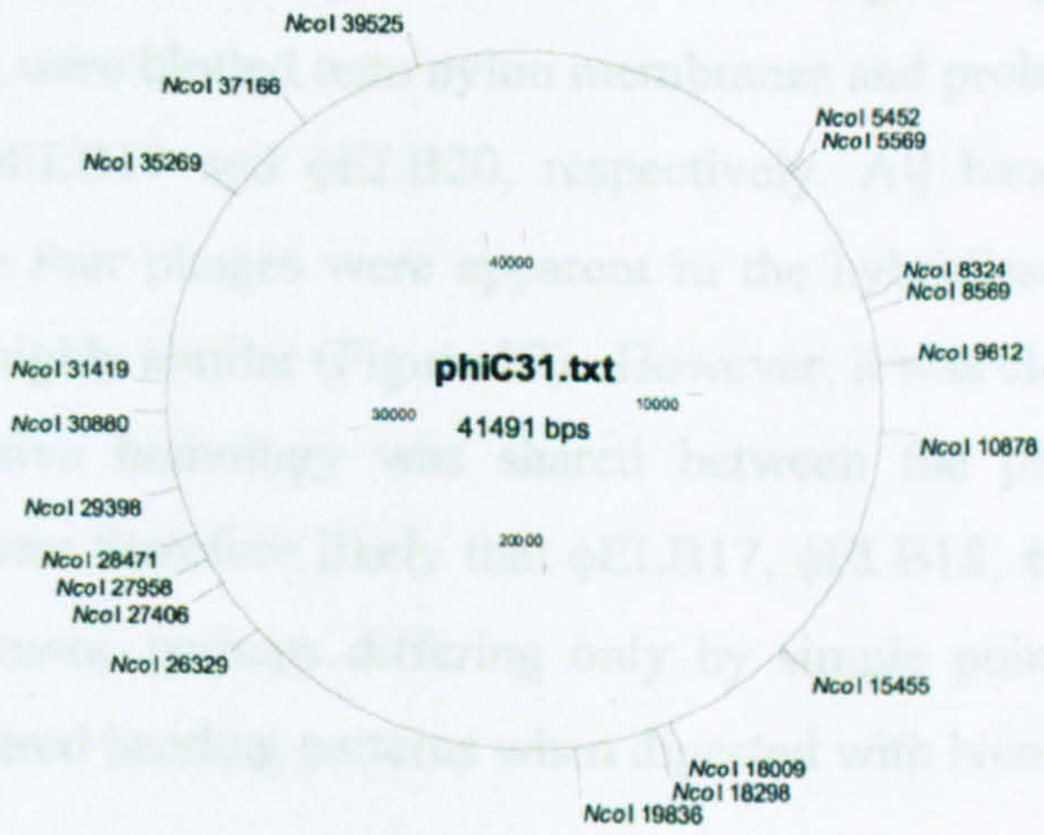


Figure 25: Phage DNA digested with NcoI. BOTTOM: ϕC31 restriction map displaying NcoI sites. ϕC31 is digested in to smaller fragments than the indigenous soil phages as illustrated on the restriction pattern above.

5.2.3. Homology between phages and determination of *cos* ends

In order to determine the extent of relationships between the *Streptomyces* bacteriophages, Southern blotting and hybridisation were carried out. The initial experiment involved determining if ϕC31 shared any homology with ϕELB17, ϕELB18, ϕELB19 and ϕELB20. An agarose gel containing DNA of the phages digested with NcoI was blotted onto a nylon membrane and probed with DIG labelled ϕC31 DNA. This blot confirmed that there was no detectable homology between ϕC31 and ϕELB17, ϕELB18, ϕELB19 and ϕELB20 (Figure 26). This lack of homology was unexpected as the wide host ranges of both sets of phages were similar (see section 3.3). Additionally, previous studies had suggested homology between phages with the same hosts. The lack of similarity may be a consequence of the fact that the source of phages varied and thus they did not evolve together.

Four additional blots were produced in order to determine the extent to which ϕELB17, ϕELB18, ϕELB19 and ϕELB20 were related. In order to attain an overview of similarities between phages entire genomes were used. Agarose gels containing NcoI restricted phage DNA were blotted onto nylon membranes and probed with DIG labelled ϕELB17, ϕELB18, ϕELB19 and ϕELB20, respectively. All bands evident from gel electrophoresis of the four phages were apparent in the hybridisation blots, indicating that the phages were highly similar (Figure 27). However, it was clear from these results that, although extensive homology was shared between the phages, their banding patterns differed. It was therefore likely that ϕELB17, ϕELB18, ϕELB19 and ϕELB20 were very similar viruses, perhaps differing only by simple point mutations, thereby producing slightly altered banding patterns when digested with NcoI.

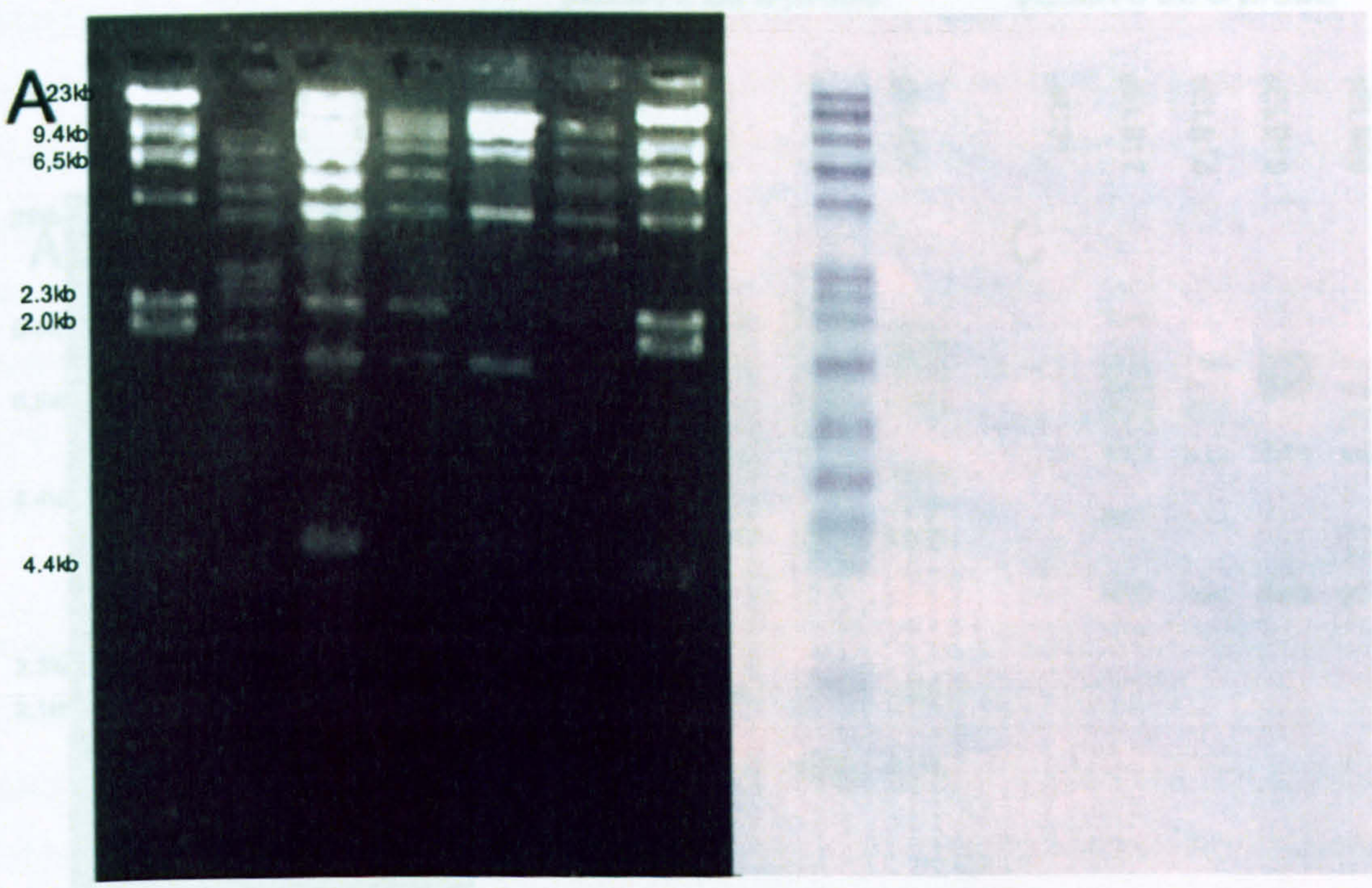


Figure 26: Restriction profiles of phages and hybridisation to phages with ϕC31 DNA. A: Restriction profiles of ϕC31, ϕELB17, ϕELB18, ϕELB19 & ϕELB20 digested with NcoI. B: Hybridisation blot of gel A using ϕC31 DNA as a probe. Gel A was blotted onto a nylon membrane and DIG labelled ϕC31 DNA was used to probe the membrane. Upon visualisation only bands in the ϕC31 lane became apparent indicating that there is no homology between this phage and ϕELB17, ϕELB18, ϕELB19 or ϕELB20.

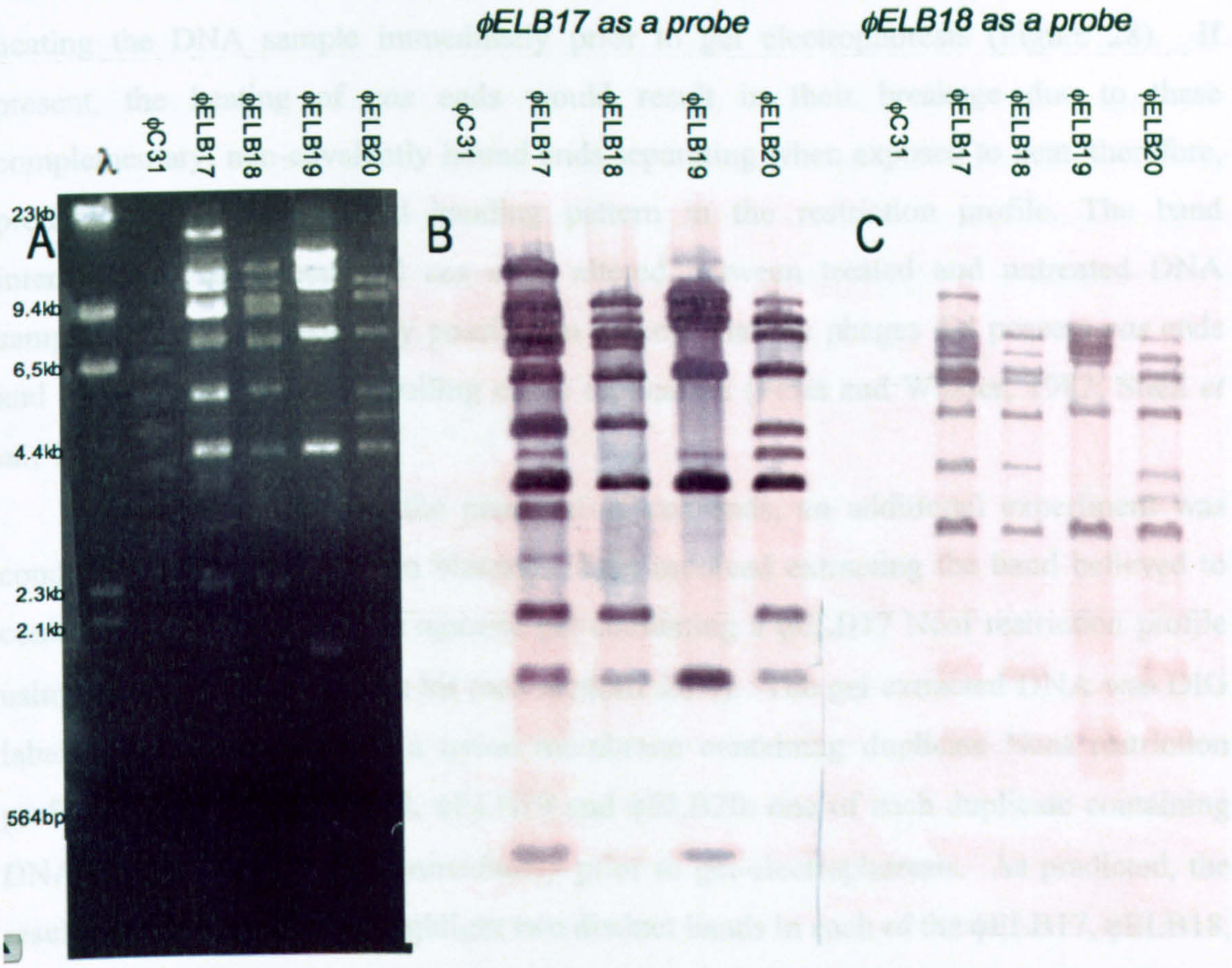


Figure 27 Homology displayed between ϕC31, ϕELB17, ϕELB18, ϕELB19 & ϕELB20.

A: Gel electrophoresis of bacteriophages digested with NcoI. B: Southern blot of ϕC31, ϕELB17, ϕELB18, ϕELB19 & ϕELB20, using ϕELB17 as a probe.

C: Southern blot of ϕC31, ϕELB17, ϕELB18, ϕELB19 & ϕELB20, using ϕELB18 as a probe. Template gels for Southern blotting were as in A; These were blotted onto nylon membranes and probed with appropriate DIG labelled phage DNA. All bands in the lanes containing ϕELB17, ϕELB18, ϕELB19 & ϕELB20 were highlighted by the probes, indicating extensive homology between the four indigenous soil phages.

The four soil bacteriophages were examined for the possession of *cos* ends by heating the DNA sample immediately prior to gel electrophoresis (Figure 28). If present, the heating of *cos* ends would result in their breakage due to these complementary, non-covalently bound ends separating when exposed to heat, therefore, producing a slightly altered banding pattern in the restriction profile. The band intensities of the speculated *cos* ends altered between treated and untreated DNA samples. It was subsequently possible to assume that the phages did possess *cos* ends and therefore replicated by rolling circle replication (Feiss and Winder, 1982; Stahl *et al.*, 1982).

In order to confirm the presence of *cos* ends, an additional experiment was conducted involving Southern blotting. This involved extracting the band believed to contain the *cos* ends from an agarose gel containing a ϕELB17 NcoI restriction profile using a Qiagen gel extraction kit (see section2.20.4). The gel extracted DNA was DIG labelled and used to probe a nylon membrane containing duplicate NcoI restriction profiles of ϕELB17, ϕELB18, ϕELB19 and ϕELB20: one of each duplicate containing DNA that had been heated immediately prior to gel electrophoresis. As predicted, the resulting Southern blot did highlight two distinct bands in each of the ϕELB17, ϕELB18, ϕELB19 and ϕELB20 profiles. Each band differed in intensity depending on whether it was heated prior to electrophoresis (Figure 29). This confirmed that ϕELB17, ϕELB18, ϕELB19 and ϕELB20 did indeed contain cohesive ends within their genomes.

The extensive homology revealed between ϕELB17, ϕELB18, ϕELB19 and ϕELB20, a fifth indigenous bacteriophage (ϕELB8), with a different host profile (see Chapter 3) was analysed. The purpose of this analysis was to determine if this phage, clearly different from ϕELB17, ϕELB18, ϕELB19 and ϕELB20, shared any homology with them or with any other well-characterised phages.

For comparison purposes, a single restriction enzyme was required for all the phages. NcoI did not restrict ϕELB8 DNA, as a result KpnI was used for all phages. It displayed a distinct restriction profile which was different from the other four phages

(see figure 10). Despite the obvious differences from the restriction profiles, hybridisation experiments illustrated that there was indeed approximately 50% homology of ϕELB8 with ϕELB17, ϕELB18, ϕELB19, ϕELB20 and ϕHau3AI (Figure 30). This result indicated that, although the phages within the same soil sample may have had different lifecycles and host range, they did share common genes. These phages may share substantial amounts of their genomes due to a common ancestor from which they have evolved. Consequently, essential genes will have been maintained whilst non-essential genes will have varied between strains.

5.2.4. Genome Sequencing of phages

For the purposes of genome sequencing, plasmid genome libraries of ϕELB17 and ϕELB18 were created as it was these phages that were first isolated from the soil and had been subject to characterisation. This was achieved using TOPO® Shotgun Subcloning kit (invitrogen) and involved shearing the phage DNA using a nebuliser, which uses pressure, (the source being nitrogen gas), to shred the DNA into appropriate sizes of approximately between 1-2Kb. In order to optimise the sizes, this was initially tested over a time course from 0 to 120 seconds and samples were taken at 15-second intervals (Figure 31). These fragments were ligated into pTOPO EcoRI sites and purified using Promega Wizard kits. Blue white selection of *E.coli* allowed for selection of the cloned plasmids. Libraries were then created with an estimated 99% coverage using TOPO Blunt ended cloning kit (Invitrogen) and stored.

Following the creation of the ϕELB17 and ϕELB18 libraries subsequent sequencing of ϕELB20, in collaboration with R.Hendrix, University of Pittsburgh was conducted, The genome of ϕELB20 was sequenced, firstly by, shearing the DNA into 1-3kb sizes and ligated into EcoRV sites of pBluescript II KS+ vector. Individual plasmids were purified using QiaPrep plasmid purification kits and these were sequenced from both ends of the inserted DNA by using Applied Biosystems BigDye v3.0 dye terminator chemistry and universal sequencing primers. Sequences were analysed using ABI Prism 3100 vDNA analyzer. Oligonucleotide primers were synthesised and used to prime sequencing reactions with whole genome templates to

provide sequence coverage of underrepresented regions and to fill gaps in sequence assembly. Approximately eight-fold-coverage was attained (Hendrix, personal communication).

Due to the complete sequencing of ϕELB20 and the similarities between the four indigenous soil phages ϕELB17, ϕELB18, ϕELB19 and ϕELB20, highlighted by hybridisation experiments, the libraries of ϕELB17 and ϕELB18 were stored, but not used for sequencing. The sequence derived from ϕELB20 was used for analysis.

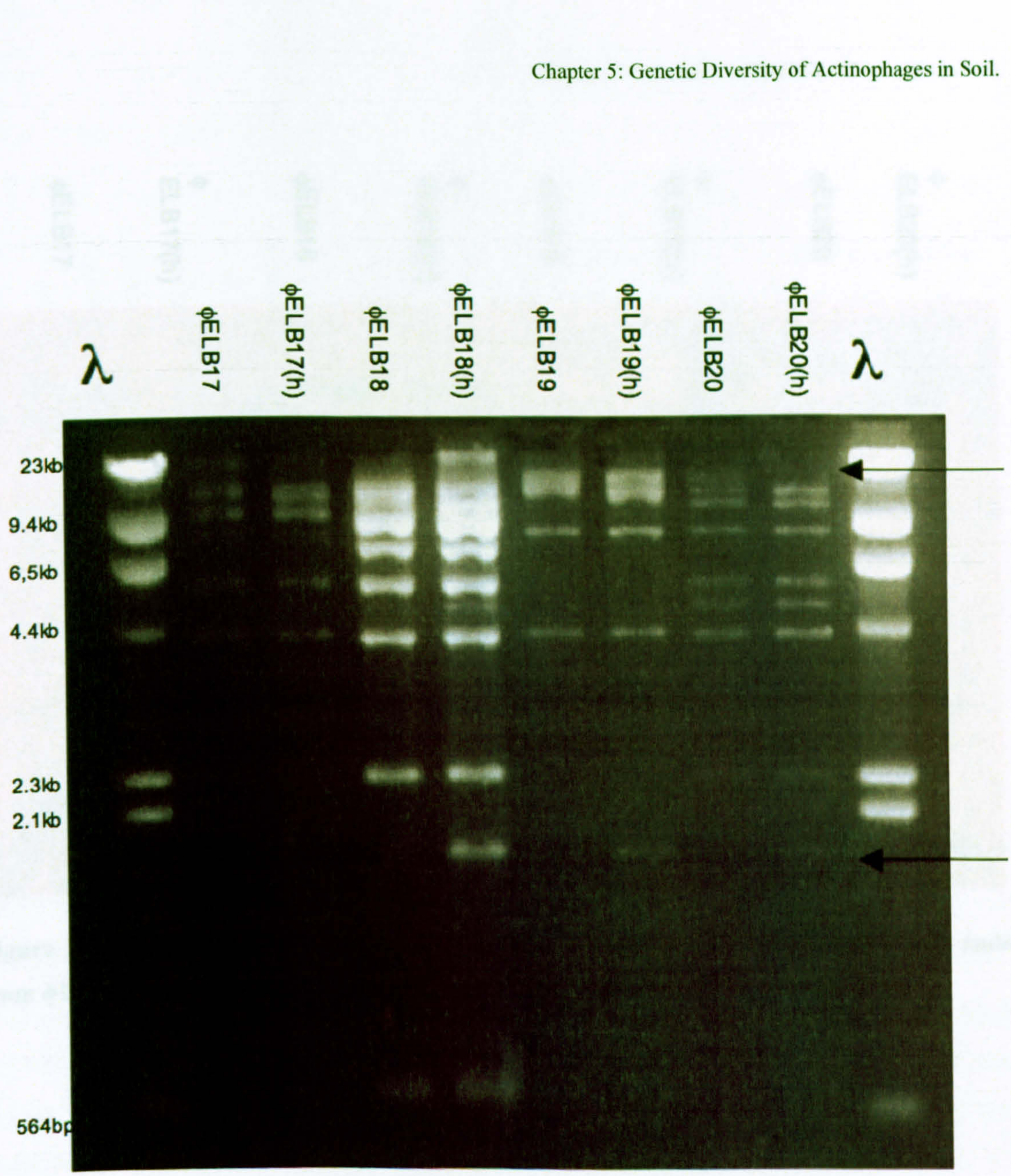


Figure 28 Restriction profiles of ϕELB17, ϕELB18, ϕELB19 & ϕELB20 digested with NcoI, displaying differences in band intensities due to the presence of *cos* ends. Black arrows indicate bands possessing *cos* ends. (h) denotes the sample was heated prior to electrophoresis. Upon heating, the *cos* ends were broken, resulting in a greater intensity of the band at around 1600bp.

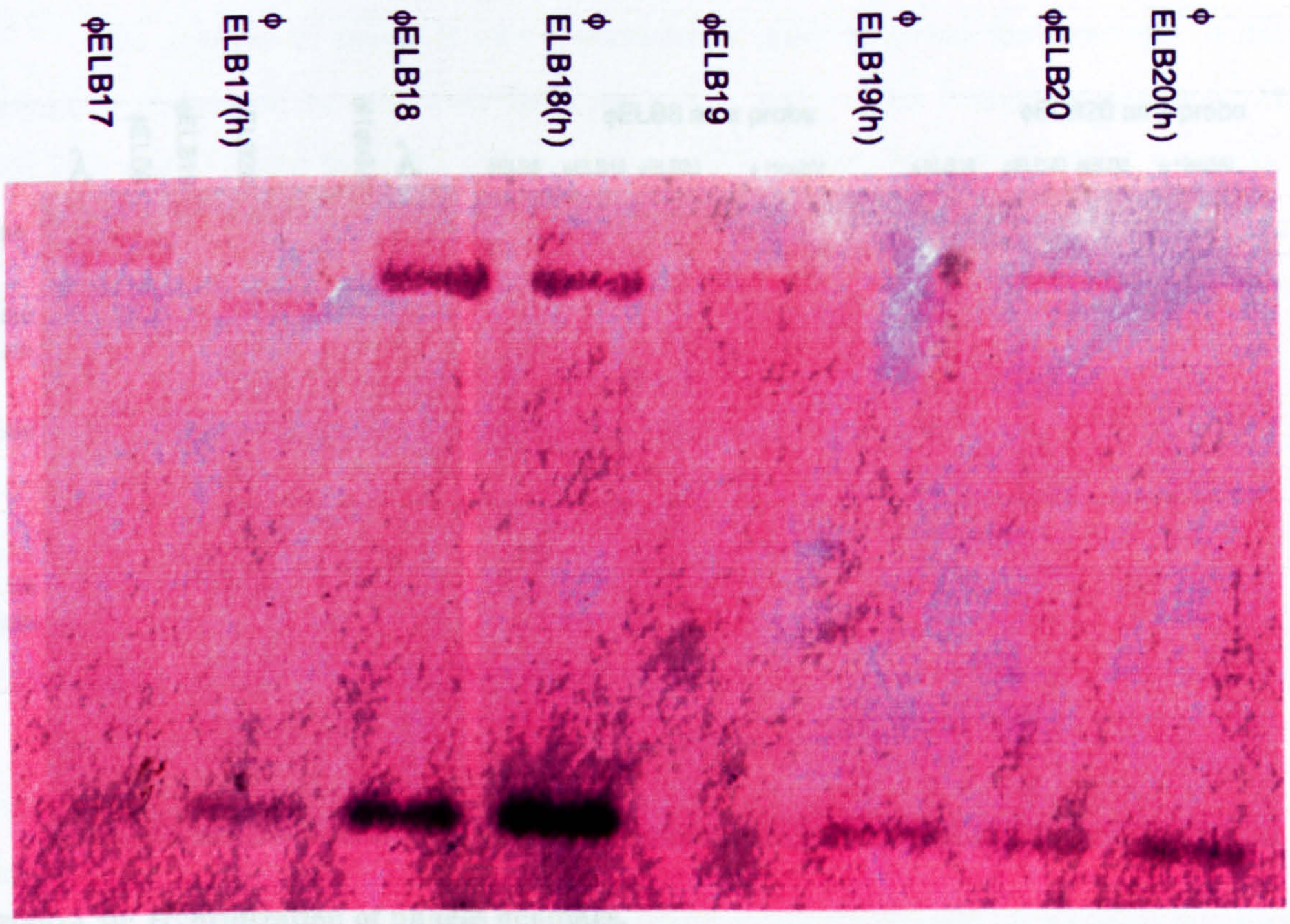


Figure 29 Hybridisation blot of phage restriction profiles probed with putative *cos* ends from ϕELB17. Approximately 1mg DNA was loaded per sample.

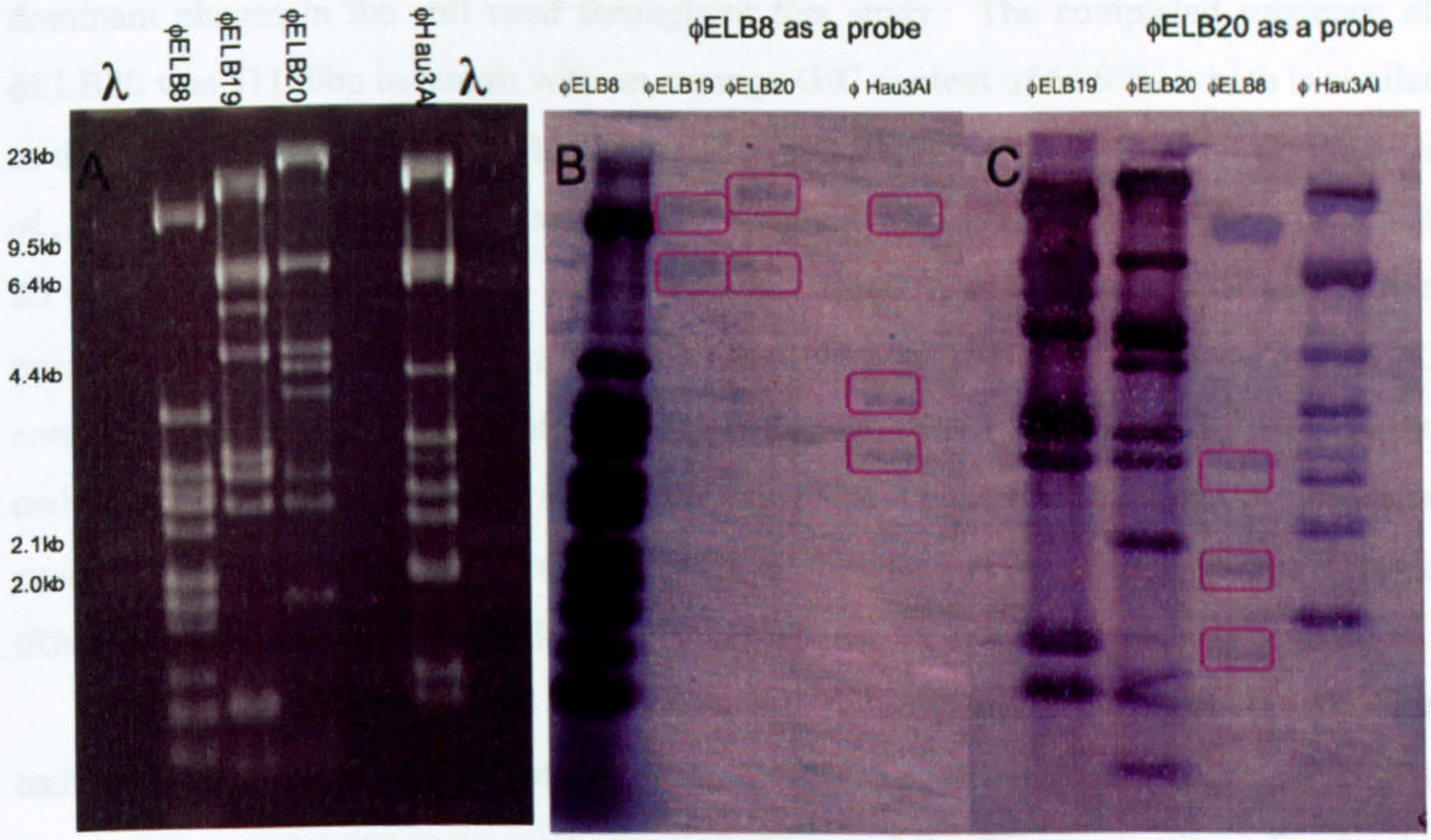


Figure 30: Hybridization of phages genomes.

A: Agarose gel showing KpnI restriction profiles of ϕELB88, ϕELB19, ϕELB20 and ϕHau3AI.

B: Southern blot of KpnI restriction profiles of ϕELB8, ϕELB19, ϕELB20, and, ϕHau3AI using ϕELB8 as a probe. The template gel (as in A) was blotted onto a nylon membrane and this probed with DIG labelled ϕELB8 DNA was used as a probe. Upon visualisation of the membrane, specific bands in all the phages were highlighted.

C: Southern blot of KpnI restriction profiles of ϕELB19, ϕELB20, ϕELB8 and ϕHau3AI, using ϕELB20 as a probe. The template gel similar to A. Containing ϕELB19, ϕELB20, ϕELB8 and ϕHau3AI respectively, was blotted onto a nylon membrane and this probed with DIG labelled ϕELB20 DNA was used as a probe. Upon visualisation of the membrane, specific bands in all the phages were evident. Boxes highlight unclear bands.

5.3. The Genome Sequence of ϕELB20 and its comparison with ϕR4

The genome of ϕELB20 was sequenced in order to reveal the genotype of the dominant phages in the soil used throughout this study. The completed sequence of ϕELB20 was 51160bp in length with an average G/C content of 66.99% which is similar to that of *S.coelicolor* (72.3% (Bentley *et al.*, 2002)) and *S.avermitilis* (71.1% (Ikeda *et al.*, 2003)). Bioinformatic analysis revealed 81 putative protein-coding regions of which 20 overlap. This bacteriophage shared 98% sequence similarity with ϕR4 and it was possible to visualise this using a similarity matrix (Figure 32). Due to the high G C content in *Streptomyces* spp codon usage in this species is biased and as a result the codons for ϕELB20 displayed a strong preference for a G or C in the third position (for example, gp47 has a G/C in the third position at 96.8%). The genome contains a single tRNA gene encoding phenylalanine.

The homology of ϕR4 was apparent despite the differences in isolation technique, location and time of isolation. ϕR4 and ϕELB20 proved to be almost identical in genome sequence and structure (Figure 34). The production of a dotplot using Genome Shovel illustrated the extensive homology and areas of insertion between these two phages (Figure 32). Upon examination of the designated ORFs, these areas of difference and insertion encoded for hypothetical proteins whose function was unknown, typically on the right arm of the genome.

The genome of ϕR4 and ϕELB20 can be divided into five main areas: structural, lysogeny, lysis,, DNA replication and unclassified or undefined. The structural gene cluster encases ORFs 2-23. Within this cluster are many genes similar to the mycobacteriophages such as Tweety, Bxz2, Che12, D29, Bethlehem as well the *Propionibacterium* PA6 in ORF 20. This cluster of genes from the mycobacteriophages in particular may have been acquired in one or more recombinational events some time ago. However, the appearance of a single gene from PA6 indicates that this was a recent event that occurred independently from the mycobacteriophages cluster. Interestingly, it has been suggested by the authors of the PA6 genome sequence that the ORF 16 (ORF 19 in ϕELB20) makes up part of the phage tail. This is as the predicted gene product is a

42.8kDa protein of which the first 80 amino acids show homology to the N terminus of a family of protease enzymes. This allows us to speculate that it may specifically bind and subsequently degrade proteins with nonpolar C termini, allowing for efficient phage assembly or alternatively, allow easier entry into the host cell. Clearly, if this were true upon attaining this gene ϕELB20 and ϕR4 would maintain this advantageous trait.

The lysogenic gene cluster is small consisting of only ORFs 52 and 53. Previous work shows this to be a site specific member of the resolvase-DNA invertase family highly similar to *Tn5021*. The lytic cluster is in the area of ORFs 27 to 30 within which a repressor is present. This confirms that ϕELB20 is truly a temperate bacteriophage capable of the lytic and lysogenic lifecycle within particular hosts.

The DNA replication cluster cannot be exactly defined however it is in the area of ORFs 30 to 64, with the exception of a few ORFs. It is within this cluster that many of the genes show homology to the phage PA6 as well as other mycobacteriophages. ORF 35 is likely to be a DNA helicase due to its match to the COG DNA helicase group.

The remaining cluster is more difficult to define as much of the ORFs have no significant hits within Genebank, as a result the boundaries of the clusters are less defined. The last cluster is particularly large within ϕELB20 and ϕR4 and is found at the right arm of the genome. This cluster consists of many, small ORFs with little or no matches in the sequence database.

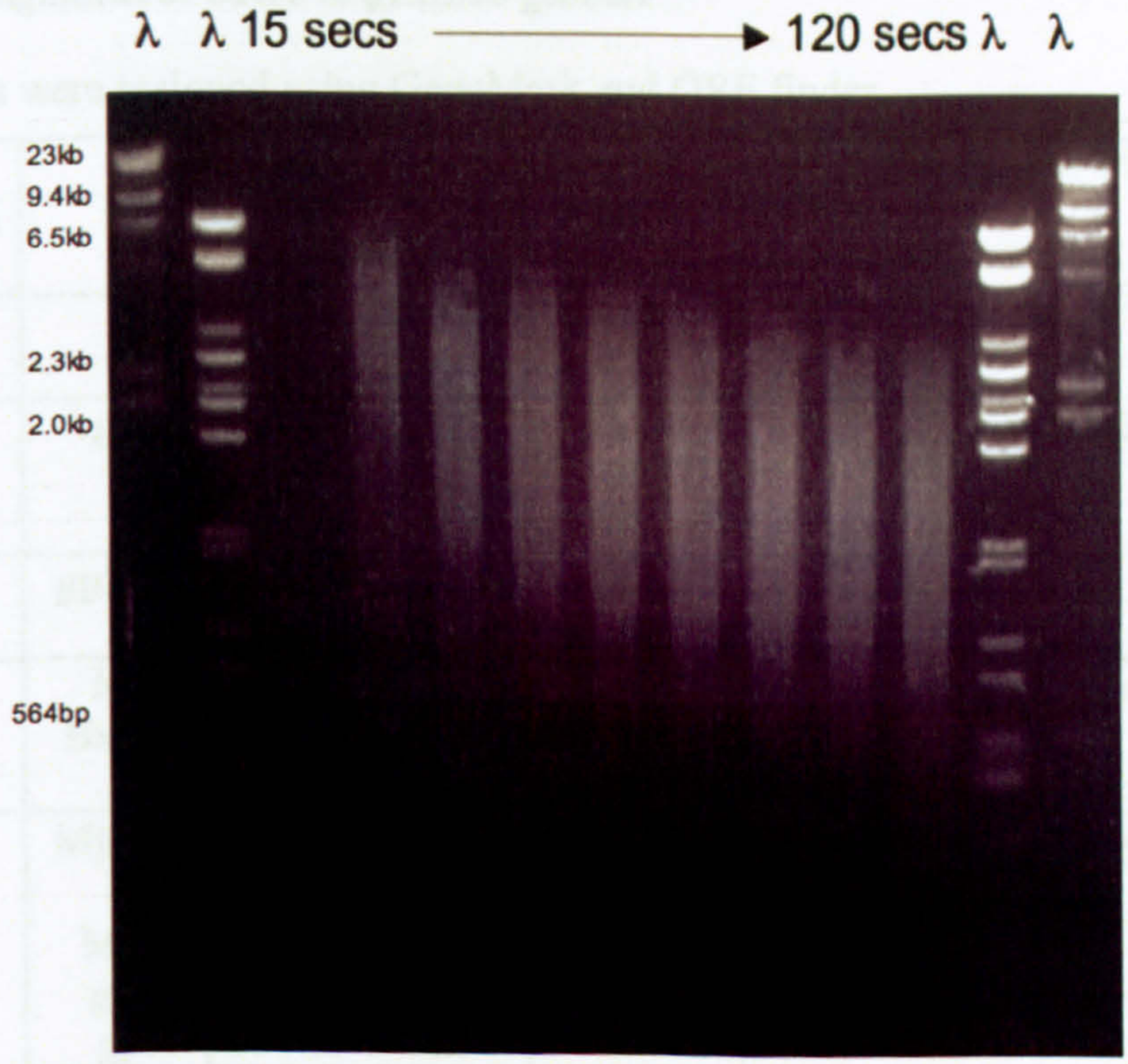


Figure 31 Shearing of ϕELB17 DNA. DNA was sheared using a nebuliser under pressure of nitrogen gas. Phage DNA was exposed for up to two minutes with samples taken every 15 seconds. Flanking lanes contain λ DNA digested with HindIII and PstI respectively. Fragment sizes of approximately 1-2 kb were desired and was obtained at 60 seconds.

5.3.1. Designated ORFs in ϕELB20

Table 23: Assignment of ORFs to ϕELB20 genome

ORFs were assigned using GeneMark and ORF finder.

Assigned ORF (bp position)	aa length	Homologues	BLASTp E-value (Score)	Putative role or Pfam search result
1 (38-220)	61			
2 (804-974)	804	gp1 Mycobacterium phage BXb1 and Pukovnik	3e-04 (47.8)	Structural protein
3 (1370-1630)	87	gp29 in ϕC31 and ϕBT1	5e-12 (73.6)	
4 (1845-2180)	112	Mycophages -gp2 in Bxz2 &Bethlehem, gp5 in D29 & Che12	4e-27 (124)	Structural protein
5 (2194-3855)	554	Mycophages gp13 Bxz2, Che12 & D29. Mycophage Puknovik gp33 & Strep phages gp33 ϕC31 & VWB	4e-110 (402)	Terminase (COG motif)
6 (3885-5357)	491	Mycophages gp4 of TM4, Ramsey, Bommer & Wildcat	1e-044 (185)	Portal protein
7 (5354-6463)	370	gp28 in mycophage Wildcat	6e-06 (55.8)	C1 domain (rich in his & cis)
8 (6466-6990)	175	Phage protein from *Streptococcus* spp.	8e-04 (46.6)	Hypothetical protein
9 (7068-8135)	356	Mycophage go30 of Wildcat	4e-29 (132)	Capsid
10 (8214-8711)	166	Gp3 in Strep phage VWB	7e-04 (43.1)	Structural protein
11 (8708-9049)	114			
12 (9049-9516)	156			
13 (9516-10013)	266			

14 (10026-10766)	247	Mycophage Bxb2 gp23	2e-16 (89.7)	Tail protein
15 (10890-11231)	114	Mycophages gp22 in Solon & gp21 in U2	5e-06 (55.1)	
16 (11291-11623)	111	Mycophages gp27 in Che12 & Puknovik, gp25 in D29	6e-04 (46.6)	
17 (11627 – 15028)	1134	Phage protein in *Salinaspora* spp.	3e-114 (417)	Tape measure
18 (15035-15892)	286	Protein in *Salinaspora* spp	2e-59 (233)	Tail protein
19 (15893-17104)	404	gp16 in PA6	7e-06 (55.8)	Tail protease?
20 (17109-17411)	101			
21 (17412-19601)	703	Strep phages gp49 in ϕC31 and ϕBT1	9e-41 (172)	Bacterial oxidase
22 (19637-20107)	157			
23 (20139-21416)	426	Mycophage go4 in Bethlehem	7e-08 (62.4)	Tail related protein
24 (21476-21796)	107	Mycophages, gp55 in Omega & gp33 in Tweety	7e-04 (46.2)	Tetracycline repressor
25 (21866-22936)	357	Strep phages ϕC31 & mu16	9e-68 (261)	Endolysin
26 (22980-23153)	58			Lysis
27 (23210-23518)	103			Lysis
28 (23634-24092)	153	Mycophages gp72 in Pukovnik & D29	5e-04 (43.5)	Repressor

29 (24348-24869)	174			Lysis
30 (24870-25844)	325	Corynebacterium phage gp40 & gp67 in mycophage Wildcat	8e-15 (85.1)	
31 (25968 – 26450)	161			
32 (26533-26970)	146	gp35 in PA6	9e-05 (49.3)	DNA replication protein
33 (27045-27530)	162	Strep phages gp20 in ϕC31 and ϕBT1	3e-31 (137)	Cytidine deaminase
34 (27587-28342)	252	gp34 in PA6	4e-47	RepA motif
35 (28408-28785)	126	Mycophage CjwI gp87 & PA6 gp33	1e-18 (95.5)	DNA replication protein
36 (28782-28919)	46			
37 (28907-29350)	148	Mycophage CjwI gp57 & PA6 gp32	3e-07 (57.8)	ZnF motif
38 (29365-30027)	221	Mycophage gp55 KBG & gp31 in PA6	1e—70 (269)	TOPRIM-primases
39 (30032-30394)	121			
40 (30391-30603)	71			
41 (30604 – 30792)	63			
42 (30794-31018)	75			
43 (31002-32195)	398	gp31 in PA6	1E-65 (254)	

44 (32192-32545)	118	Integrase/recombinase	6E0 (33.5)	Integrase/resolvease DNA binding
45 (32542-32928)	129			
46 (32925-33302)	126			
47 (33380-35494)	705	Mycophages gp50 D29, Bxb2 & Puknovik	0.00 (717)	RNR-2 motif. Ribonuclease reductase
48 (35597-35704)	36			
49 (35701-35943)	81	Strep phages gp25.3 in ϕC31	1.2E-02 (42.4)	
50 (35940-36188)	83	Hypothetical *S.avermitilis* protein	2E-10 (68.2)	
51 (36266-37357)	364	Hypothetical *Nocardia* spp protein	3E-94 (348)	DNA binding
52 (37707 – 38996)	430	Integrase (ORF 460 from bacteriophage ϕR4)	0.00 (880)	Recombinase
53 (39052-39252)	67			Lysogeny protein
54 (39249-39956)	236	*S.coelicolor* thymidylate synthase	8e-89 (330)	ThyX motif. Thymidylate synthase
55 (39953-40216)	88			
56 (40217-40708)	164			
57 (40693-40920)	76			
58 (40917-41414)	166	Hypothetical *Salinaspora* spp protein	1e-13 (78.6)	

59 (41411-41605)	65			
60 (41697-42101)	135			
61 (42098-42298)	67			
62 (42300-42887)	196	Strep phages gp52 in ϕC31 & ϕBTI	3e-26 (121)	Nucleotide kinase
63 (42884-43087)	68			
64 (43080-43394)	105			
65 (43394-43636)	81			
66 (43636-43968)	111			
67 (44061-44234)	58			
68 (44234-44653)	140			
69 (44912-45139)	76			
70 (45142-45570)	143			
71 (45692-46105)	138			
72 (46134-46277)	48			

73 (46359-46892)	178	Mycophage gp78 in BxbZ	2.5 (35.0)	
74 (47014-47862)	283	*Rhodobacter* spp dehydrogenase	1.6 (37.4)	Dehydrogenase
75 (48111-48242)	44			
76 (48315-48503)	63			
77 (48566-48964)	133			
78 (49057-49284)	76			
79 (49281-49457)	59			
80 (49568-49960)	131			
81 (49957-50160)	68			

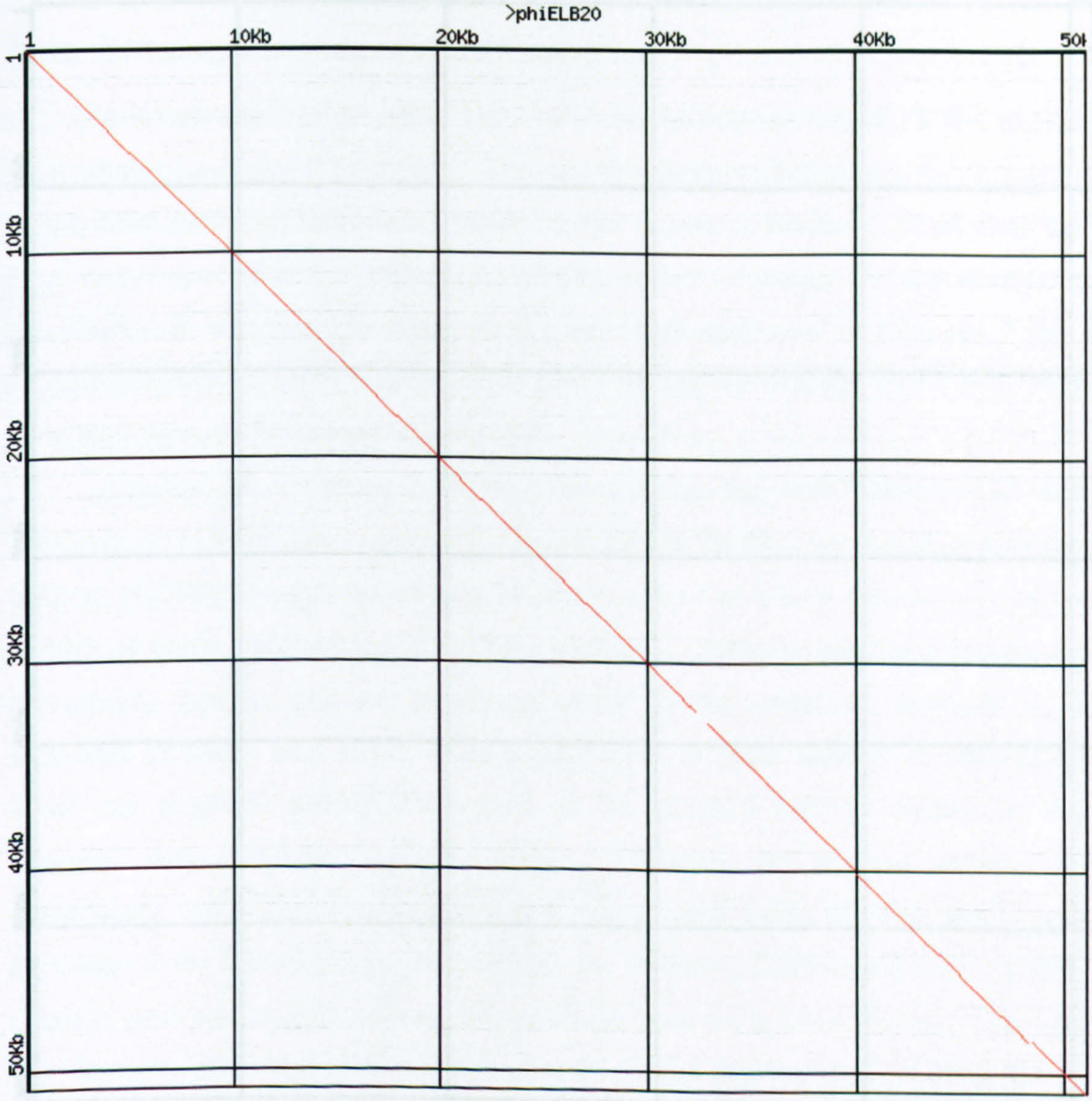


Figure 32: This plot illustrates homology in DNA sequence between ϕELB20 (horizontal) and ϕR4 (vertical), (produced using Genome shovel). The distinct diagonal line illustrates the extensive homology between the phages. A clear area of difference is highlighted by a break at around 33kb, highlighting the area of indel in ϕELB20.

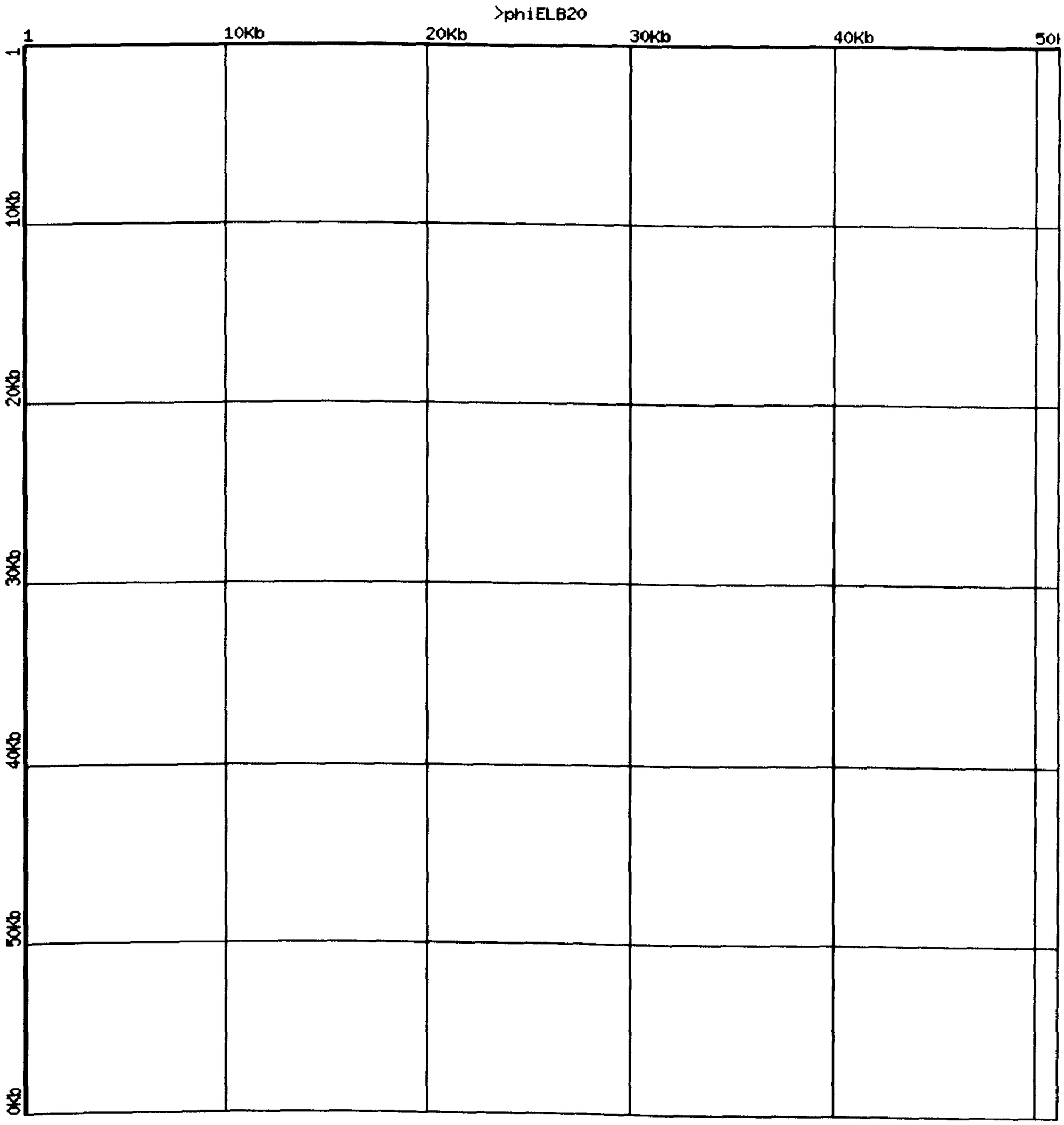


Figure 33: Dot plot of ϕELB20, and ϕHau3AI.

There is a significant amount of DNA sequence homology between ϕELB20, and ϕHau3AI. The majority of this homology of ϕHau3AI with ϕELB20 lies at the early region of the genome.

5.3.2. Differences between ϕELB20 and ϕR4

Although ϕR4 and ϕELB20 were very closely related, they did differ in some areas of their genomes (Figure 34). The majority of these differences are a result of simple point mutations that made little difference to protein function. These may have been simple spontaneous mutations, conferring neither advantage nor disadvantage to the phage, and was therefore maintained. The work conducted in Chapters 3 and 4 support that there are no significant differences between ϕR4 and ϕELB20 which impact upon their survival, host range or lifecycles.

Although the two phage genomes differed in size, they both maintained 81 ORFs This was due to additional hypothetical ORF in ϕR4 in the third frame and an additional ORF in ϕELB20 in positions 44 and 79, the presence of which ensured a consistent number of ORFs were maintained. These additional ORFs encoded for a hypothetical *S.avermitilis* protein with no significant match in the GeneBank database at the nucleotide or amino acid level. Even when subject to Pfam analysis no recognisable motif was produced making it's impact on the ϕELB20 survival difficult to fully interpret without further experimentation. However, its obvious addition and maintenance within the genome indicates it was a recent event that may confer some advantage to the bacteriophage. Interestingly the additional DNA in ϕELB20 in ORF 44 displays weak homology to an integrase/resolvase involved in DNA binding. Due to the low degree of homology it is probable that this is not functional as an integrase. Its presence illustrates the storage of "extra" DNA by bacteriophages as upon mutation a fully functional integrase may result allowing the phage to alter its life style in particular hosts.

DNA sequence comparisons revealed that generally ϕR4 and ϕELB20 shared common ancestry with several other viruses, particularly with the mycobacteriophages. However, this result might have been due to the recent publication of many mycobacteriophage genomes and the still very limited database of bacteriophage

genome sequences. In spite of this, there were clear relationships between these and other actinophages when considering genome architecture (Figure 35). The technique clearly proved that ϕR4 and ϕELB20 were related to ϕC31, ϕBT1, L5, TG1 and ϕHau3AI. By contrast, when comparing raw DNA sequence data, only the relationship with ϕHau3AI was evident (Figure 35). However, such comparisons did highlight the fact that ϕHau3AI was most closely related to ϕELB20 and ϕR4. It is possible that the subtle differences between ϕELB20 and ϕR4 were representative of the primary event that gave rise to the divergence and eventually led to ϕHau3AI. This is rather than random events resulting in ϕELB20, as small changes in sequence may confer slight advantage, therefore, reinforcing the maintenance of these changes. Figure 35 shows the differences between phages with similar genome organisation. It indicates that small, subtle differences in genomes can lead to the development of several distinct phages while genomic architecture maintains overall similarity.

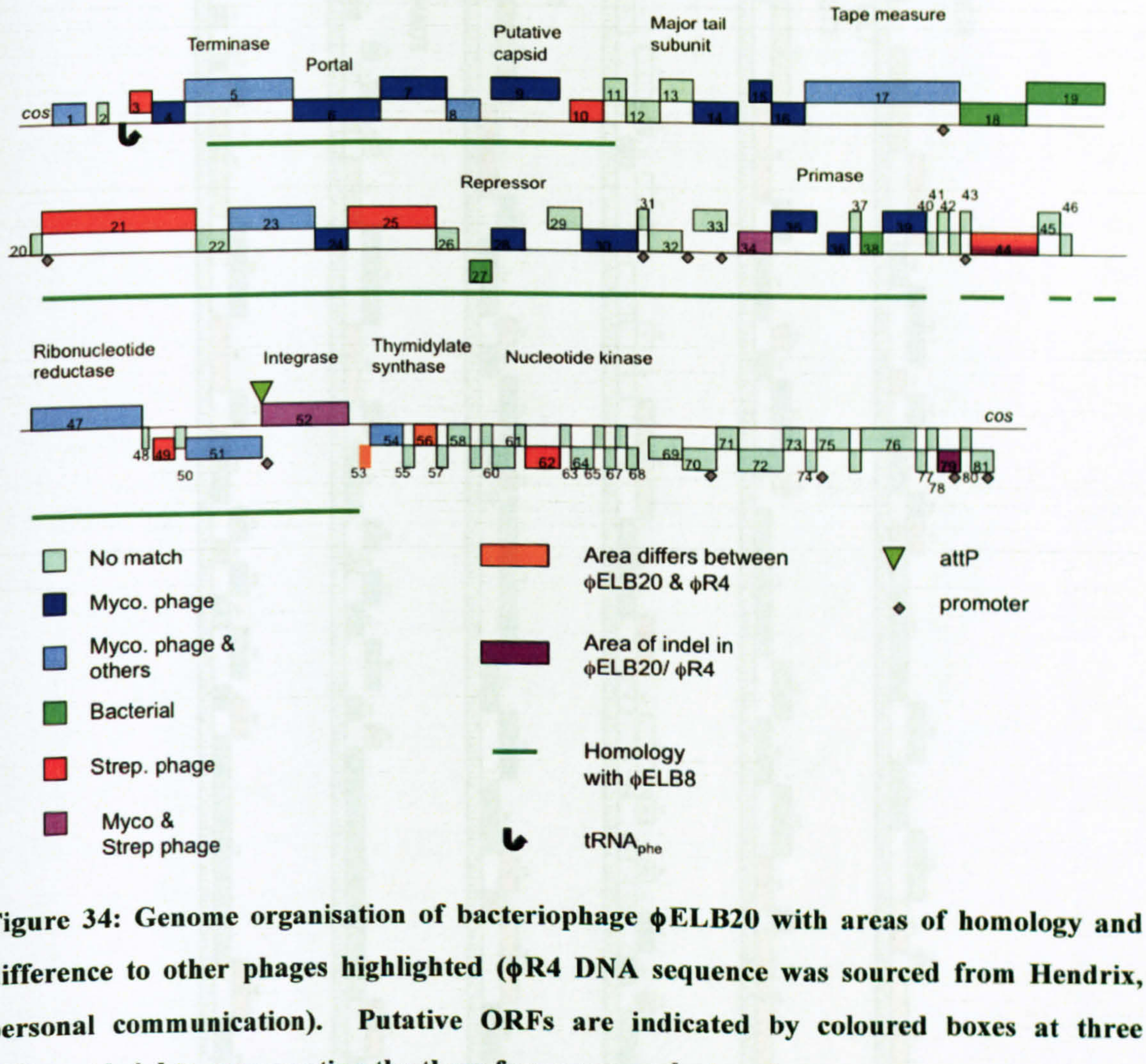
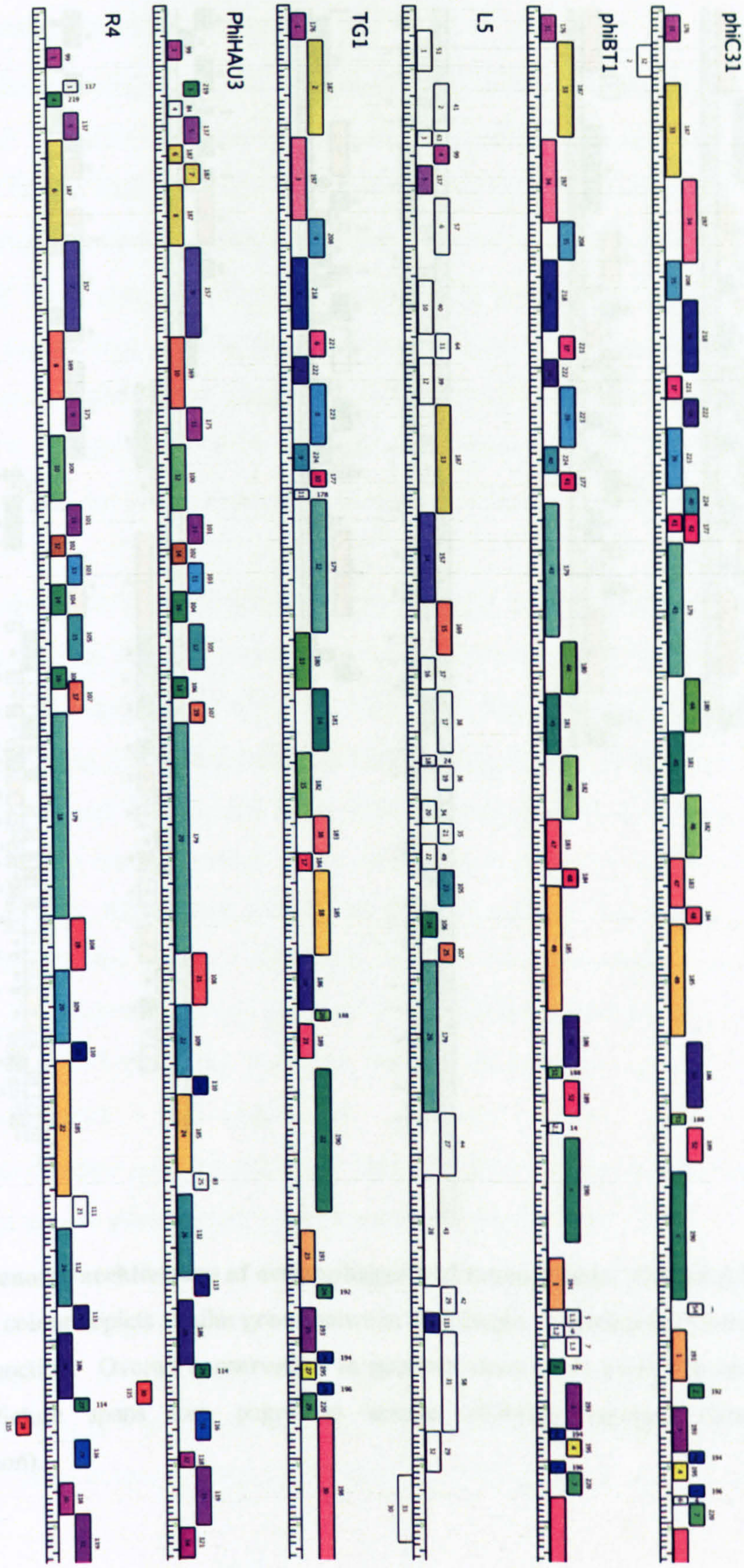


Figure 34: Genome organisation of bacteriophage ϕELB20 with areas of homology and difference to other phages highlighted (ϕR4 DNA sequence was sourced from Hendrix, personal communication). Putative ORFs are indicated by coloured boxes at three different heights, representing the three frames on each strand.

Coloured boxes display meaning in relation to homology with other phages and organisms. Those without any match in the GeneBank are light green, areas of homology with mycobacteriophages are dark blue, and light blue boxes indicate homology with mycobacteriophages and other phages. Green boxes display genes that are homologous with bacterial genes. Red is indicative of homology with *Streptomyces* phages. Purple shows genes similar to mycobacteriophage and *Streptomyces* phages. Coloured lines below sections of DNA indicate homology or differences with phages for this study. Orange boxes represent genes which are different in ϕR4. Purple boxes signify area of insertion in ϕELB20 whilst green lines indicate homology with ϕELB8.


phiC31
phiBT1
L5
TG1
PhiHAU3
R4

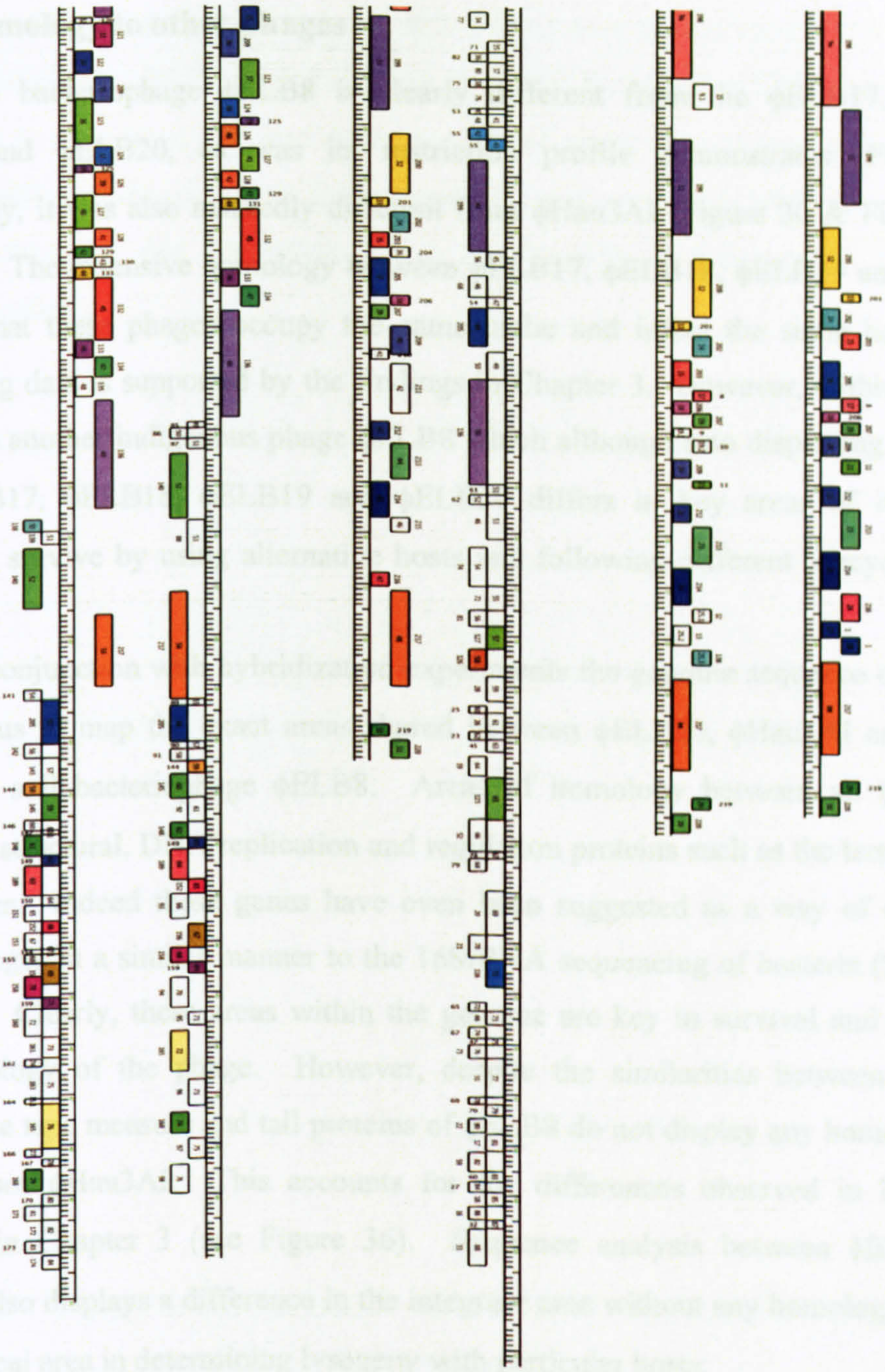


Figure 35 Genome architecture of actionphages and mycophages. Coloured boxes represent ORFs. Each colour depicts similar genes between the phages. Uncoloured boxes are ORFs with no known function. Overall conservation in genome structure is evident between these viral genomes. Picture spans two pages to ensure visibility (source: Hendrix, personal communication).

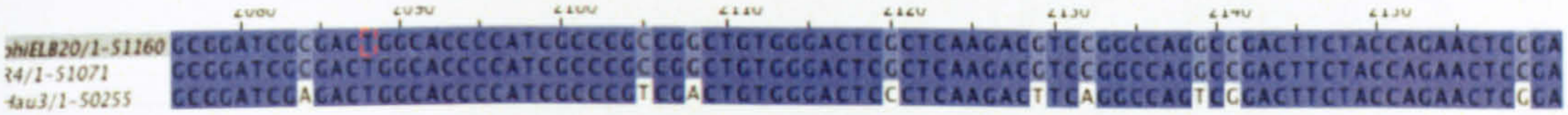
5.3.3. Homology to other phages

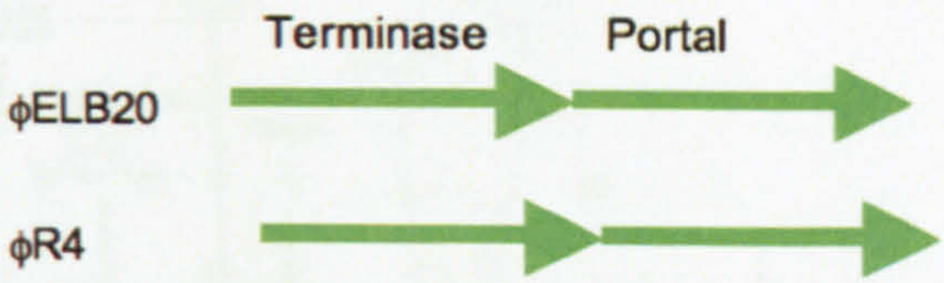
The bacteriophage ϕELB8 is clearly different from the ϕELB17, ϕELB18, ϕELB19 and ϕELB20, as was its restriction profile demonstrates (Figure 28). Additionally, it was also markedly different from ϕHau3AI (Figure 30 & Figure 33 & Figure 36). The extensive homology between ϕELB17, ϕELB18, ϕELB19 and ϕELB20 indicates that these phages occupy the same niche and infect the same hosts. This unsurprising data is supported by the findings in Chapter 3. However, within the same soil resides another indigenous phage ϕELB8 which although also displaying homology with ϕELB17, ϕELB18, ϕELB19 and ϕELB20 differs in key areas of its genome allowing it survive by using alternative hosts and following different lifecycles within these hosts.

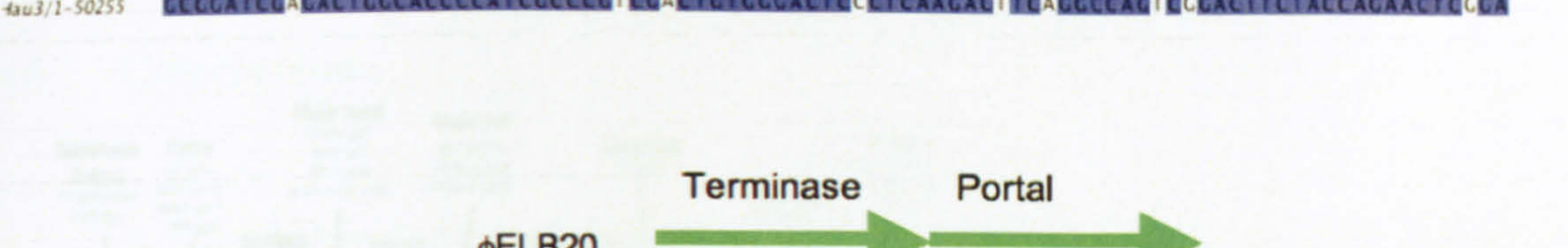
In conjunction with hybridization experiments the genome sequence of ϕELB20 can allow us to map the exact areas shared between ϕELB20, ϕHau3AI and another indigenous soil bacteriophage ϕELB8. Areas of homology between all the phages encode for structural, DNA replication and regulation proteins such as the terminase and portal genes. Indeed these genes have even been suggested as a way of classifying bacteriophages in a similar manner to the 16SrRNA sequencing of bacteria (Sullivan *et al.*, 2008). Clearly, these areas within the genome are key in survival and define the overall biology of the phage. However, despite the similarities between structural proteins, the tape measure and tail proteins of ϕELB8 do not display any homology with ϕELB20, nor ϕHau3AI. This accounts for the differences observed in host range discussed in Chapter 3 (see Figure 36). Sequence analysis between ϕELB20 and ϕHau3AI also displays a difference in the integrase area without any homology apparent in this critical area in determining lysogeny with particular hosts.

Another phage which was highlighted by genome sequencing was PA6 (Farrar *et al.*, 2007). Interestingly, some ϕR4 and ϕELB20 ORFs are homologues of a *Propionibacterium acnes* phage, PA6. Within ϕELB20, there were six ORF which showed homology to PA6: these were all clustered, suggesting their addition to be by few recombination events (Figure 37). Within the PA6 genome, these ORFs overlapped.

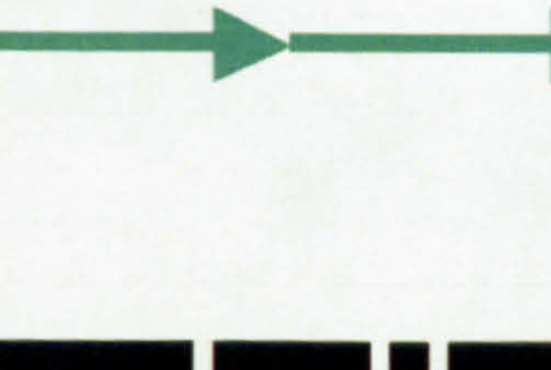
However, within ϕELB20 they did not, although there was little or no space between these and subsequent ORFs. This indicated that these homologues within ϕELB20 were not derived directly from PA6 and rather transferred on by several intermediate phages, thereby resulting in divergence of these genes. It can be postulated that the sequence of events resulting in this homology is by the transfer of these genes from other Actinomycetales







Integrase



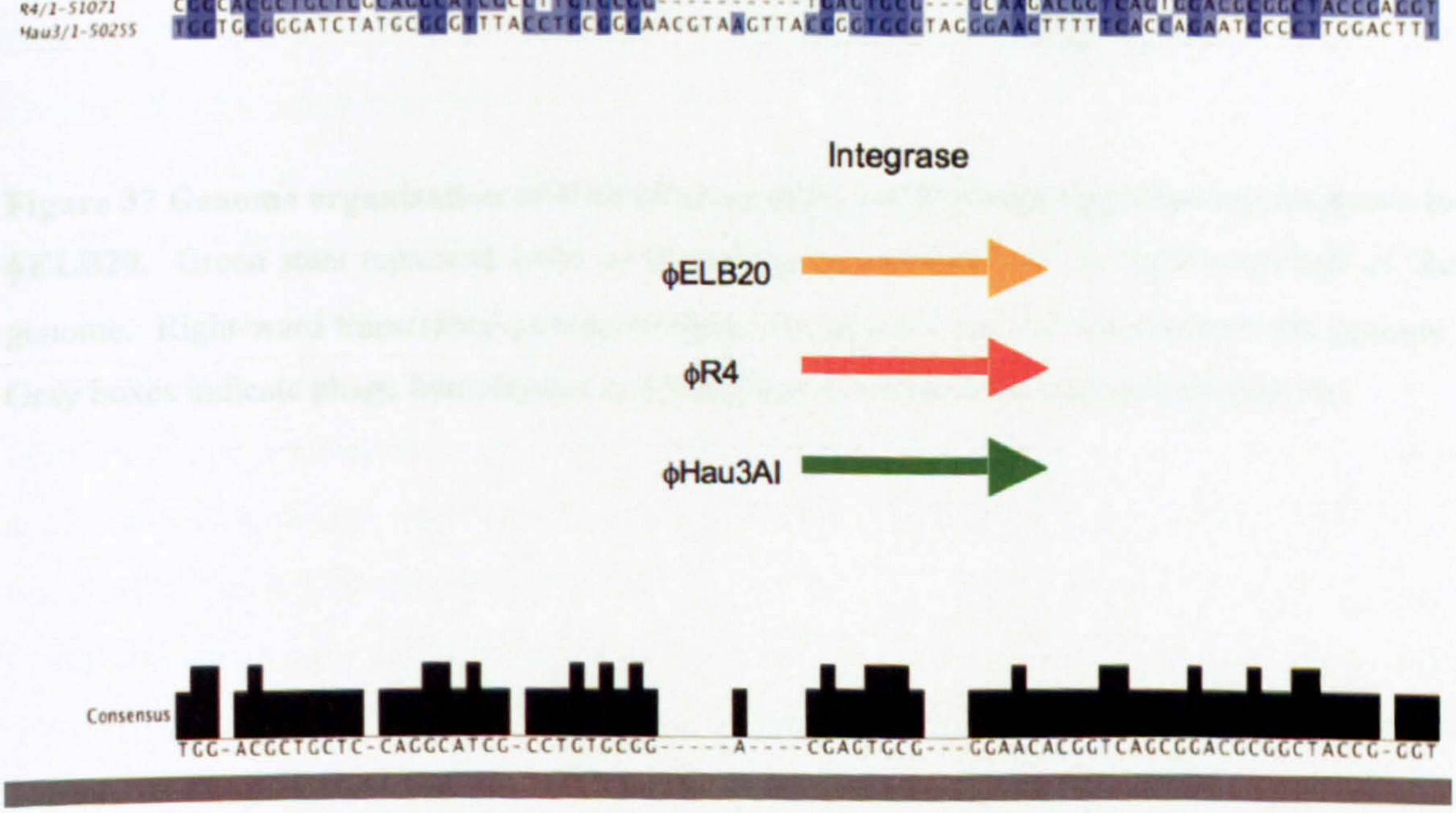


Figure 36: Areas of homology or difference between ϕELB20, ϕR4 and ϕHau3AI terminase and integrase. At the top of each illustration the deeper the purple the more identity between the three phages. (Work on Jalview 3.0)

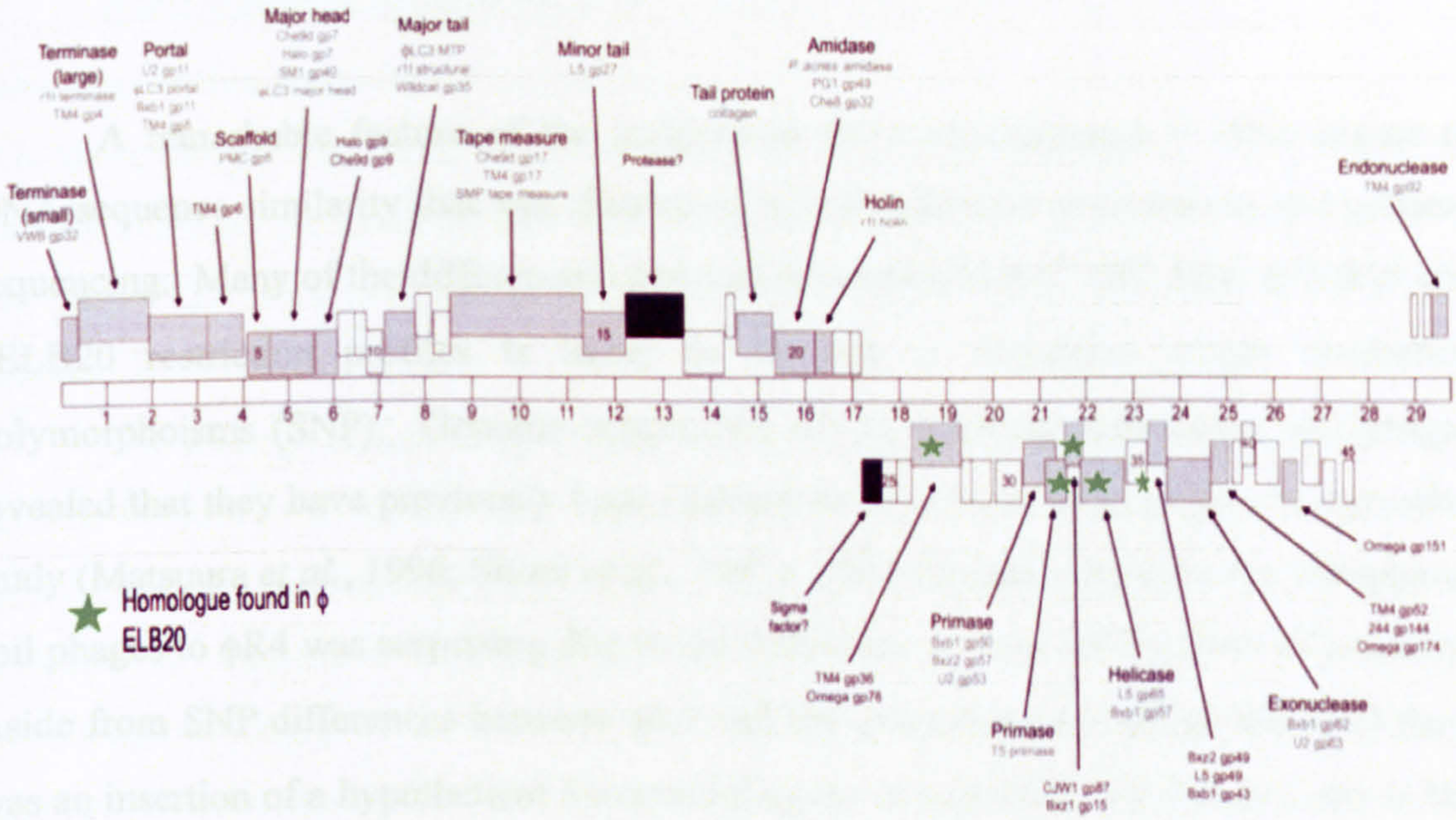


Figure 37 Genome organisation of PA6 (Farrar *et al*, 2007) displaying homologous genes to ϕELB20. Green stars represent areas of homology to ϕELB20, within the second half of the genome. Right-ward transcribed genes are above the genome and left-ward below the genome. Gray boxes indicate phage homologues and black boxes represent nonphage homologues..

5.4. Conclusions

A remarkable feature of the indigenous soil bacteriophages is their degree of DNA sequence similarity that was illustrated by hybridization experiments and genome sequencing. Many of the differences observed between ϕELB17, ϕELB18, ϕELB19 and ϕELB20 restriction profiles is likely to be due to numerous single nucleotide polymorphoisms (SNP). Genome sequencing of one of these indigenous soil phages revealed that they have previously been isolated from soil and even subject to extensive study (Matsuura *et al.*, 1996; Shirai *et al.*, 1991). This similarity between the indigenous soil phages to ϕR4 was surprising due to the difference in time and location of isolation. Aside from SNP differences between ϕR4 and the indigenous soil phage ϕELB20 there was an insertion of a hypothetical *S.avermitilis* gene in ϕELB20. This gene is sits in the cluster of small, non-essential genes with no known function towards the end of the genome, an area potentially subject to extensive genetic exchange. The second significant difference between ϕR4 and ϕELB20 is the presence of a gene with homology to an integrase. This gene sits in an important area of DNA replication in the phage genome and its integration and maintenance within the genome points to it providing an advantageous trait. The small differences observed between phages in this study illustrate phage evolution as it occurs. Eventually leading to a bacteriophage which can lead a different lifecycle much like the differences in lifecycles and genomes observed between mycobacteriophages L5 and D29 (Pedulla *et al.*, 2003a). In addition, the soil phage ϕELB8, isolated and examined due to its difference in lifecycle and host range from the other phages, also shares homology with ϕR4 ϕELB20 and ϕHau3AI. The areas of homology between all these phages can be mapped to DNA replication and structural ORFs indicating that the products of these are essential. Typically, it is the tail proteins, which determine host range and the cluster of small genes towards the right end of the genome that vary between phages. This can also be observed when examining not only DNA sequence but also the mosaicism structure of the phage

genomes, all with a similar ϕR4-like genomic architecture. The variations on ϕR4 that do succeed in soil result in such phages as ϕHau3AI and ϕELB8.

Bacteriophages of significant difference to the ϕR4-like bacteriophages are not as numerous within the soil and are most likely filling a small niche within the ecosystem. It is these phages that, with environmental pressures changing, act as a pool of genes, thereby allowing the predominant virus to evolve by the acquisition of new genes from this pool. The work in this Chapter supports the differences in host range observed in Chapter 3.

The findings of this study support previous studies such as those of Summer *et al* (2006), Pedulla *et al* (2003) and Kwan *et al* (2005). Each of these studies reveals genetic stability between viruses infecting the same host species. They also describe the presence of numerous insertions and deletions flanked by regions of significant homology, although, the synteny remains in all studies. This consistency in genomic structure is indicative of phages' requirement to maintain it. The ease with which the viruses can evolve by recombination due to their genomic architecture may be one factor selecting for their maintenance.

6. General Discussion

Upon examination of the literature in Chapter 2, it became apparent that there exists a distinct lack of studies concentrating on the dynamics of phages and hosts in a single soil sample. This overlooked area, despite its small scale, is a contributory factor of global microbial mortality (Suttle, 2005). Moreover, bacteriophages not only control bacterial populations, but also drive their evolution, and understanding of these processes within soil would allow for a more comprehensive understanding of soil ecology.

Streptomyces, the indigenous soil host organisms which was the subject of this study, were readily isolated by the established spore isolation technique (Herron and Wellington, 1992). Their placement within a phylogenetic tree found them to be clustered within their own clade, indicating that they evolved from a common ancestor, separate from other typed bacteria (see section 3.2). These findings supported those of previous studies examining indigenous soil communities explicitly displaying unique microorganism populations, regardless of their geography (Wawrik *et al.*, 2007). Indeed, it has even been found that soil communities can differ significantly despite their relative proximity to one another (within meters) in uniform soil (Sliwinski and Goodman, 2004). The mechanisms responsible for these differences in communities have not yet been resolved; however there are several hypotheses, which claim those differences to be caused by dispersal, physical isolation and genetic drift (Escobar-Paramo *et al.*, 2005; Papke *et al.*, 2003). In addition, competitive interactions and microhabitat variability may contribute (Bent *et al.*, 2003). These findings have implications for classical microbial taxonomy, which, by placing all strains into specific groups, results in a gross underestimation of their total global diversity. An example of this can be found in soil streptomycete isolation experiments which repeatedly report the isolation of *S.griseus* according to numerical taxonomy. (Arora *et al.*, 2005; Carvajal,

1946). This thesis illustrated that the use of 16S rRNA does provide an indication of the unique diversity in microbial communities within soil.

Futhermore, using indigenous and typed phages, the hosts were classified by phage typing. Overall, the resulting dendrogram did not match the phylogenetic tree produced by 16S rRNA sequencing (see section 3.3). However, it did display the two main groupings of host susceptibility to infection by the phages: highly susceptible (SPO1, SPO2, SPO3, SPO4, SPO5 & SPO8) or highly resistant (SPO6, SPO7 & SPO9) to the majority of resident phages. The majority of hosts were receptive to infection by a wide range of phages. As previous studies have stated, this may be more advantageous to the bacterial population rather than development by them of resistance mechanisms.

The above findings were supported by the EOP experiments, which confirmed that the resident soil phage and bacterial populations have evolved together, each devising their own survival tactics in response to each other's defences. The majority of the hosts appeared to have relatively high EOPs, however, a few, namely, SPO6, SPO7 and SPO9, indicated by their low susceptibility to infection, that they possessed phage defence mechanisms (see Table 21). Within the phage typing dendrogram, these organisms were grouped together, indicating relationships between them (Figure 16).

It was hypothesised that such a unique host population would also be observed in the resident phage population. The ease with which phages were isolated from soil immediately provided an indication of their abundance, which, was confirmed by the approximation of actinophage counts in soil (see 3.3). With the use of their indigenous hosts, several actinophages were selected from the soil. These varied in host range and all were polyvalent, indicating that few phages were specific to a single host. However, some of the phages possessed a narrower host range than others, indicating that there were several different mechanisms applicable for the successful propagation of phage in soil (see Table 19). There were clear similarities between four of the indigenous phages and ϕR4, a non-indigenous characterised phage. Due to those, and the similarities observed in growth and burst sizes, it was concluded that for optimal survival in this soil, a virus would possess these particular traits, which would explain the predominance of ϕR4-like phages.

As there exists no equivalent to 16S rRNA in bacteriophages, they were grouped numerically, according to host range (Figure 17). The resulting dendrogram demonstrated that the phages were related, descending initially from two ancestors. In contrast to their hosts, it became apparent that these viruses were similar to the already characterised typed phages ϕR4, ϕC31 and ϕHau3AI, irrespective of their geographical isolation. In addition, homo-immunity and molecular experiments reinforced the similarity between these phages. These findings postulated a theory that, globally, bacteriophages were not as diverse as their hosts, and reinforced the findings of past studies on relationships between viruses.

The dynamics of host and phage within soil clearly respects equilibrium, as both could be isolated from soil. In order to examine the impact of the phage upon its host, population soil microcosms proved a useful tool. However, it must be taken into account that there may be micro-sites within the soil of communities varying in phage-host ratios, therefore resulting in unrepresentative sampling of these communities. In addition, previous studies have shown that in patchy host populations, virulent phages may cause their own extinction by their destruction of hosts. In spite of this, soil micrcosm experiments are deemed suitable for examining phage-host dynamics in soil (Pantastico-Caldas *et al.*, 1992). Due to the rapid decline of phage populations, even in the presence of viable hosts, emphasis in this chapter focused on host populations.

Overall, it became evident that host populations survived well despite a presence of a parasitic phage in the soil. This was most probably due to the physical properties of the soil and the majority of streptomycetes being in the phage resistant state of spores, although host defence mechanisms would have also played a role. As described in previous studies, phage resistant hosts will continuously be created and may also account for this apparently stable host population (Fischer *et al.*, 2004; Marsh *et al.*, 1993; Middelboe, 2000). Despite the evident survival of hosts, it has been shown that phages did have a significant effect upon their host populations in soil, particularly those in which the bacteriophage was lytic. However, the unobserved impact in this study was the effect of lysogeny on their hosts, which would have undoubtedly also have taken

place in these soil microcosms. The nature of this effect would be suitable for future work, as it might provide a more complete picture of host-phage dynamics in soil.

Interestingly, it was found that *S.aureus*, not indigenous to soil, survived well even in the presence of a virulent bacteriophage. Indeed, this organism and its phages were as successful surviving in soil as the indigenous strains. This highly adaptable organism grows more rapidly than *Streptomyces* spp and as a result its re-isolation from inoculated soil yielded more organisms.

The molecular examination of phages was required not only to devise molecular and reporter tools but also to reveal relationships between them and their hosts. Examination of the actinophage genomes confirmed findings of previous studies (Hendrix 1999: Summer 2006) that they are all share genomic architecture and are closely related within a soil. The similar actinophages as indicated by characterisation studies in Chapter 3 and 4 were confirmed by homology and sequencing experiments (see section 5.2.3). This strategy confirmed the findings of the phage dendrogram and of Summer *et al*; who examined a resident population of soil phages and found them to be genetically highly similar. Overall, the majority of observed differences between soil phages are synonymous point mutations, making no impact upon the lifestyle or survival of the phage. However, even phages with different host ranges were homologous, indicating the sharing of a common ancestor. Indeed, these observed small differences might lead to the eventual creation of new phages over long periods of time and under selective pressures to alter their lifestyles.

Homology experiments by Southern blotting and hybridisation revealed that ϕR4 and ϕELB20 shared homology between ϕELB8 and ϕHau3A1 (see 5.2.4 & 5.3.3). This clearly highlighted the common ancestry of actinophages irrespective of variation in their place and time of isolation. Using this data, the areas of homology were identified on the ϕELB20 genome. In general, this identification demonstrated conservation of the essential genes for phage structure and replication but variation in tail proteins and towards the end of the genomes. These areas of variation are typically confined to small genes with no known function which may act as a pool that can be used to aid evolution

in the event of a change of surroundings, for example, availability of viable hosts or as termed by Hendrix: "stuffer DNA" (Hendrix *et al.*, 2003).

The homology observed between ϕELB20 and ϕELB8 was expected as these viruses share many of the same hosts and inhabit the same area. Within the dendrogram, they did share a common ancestor but were not directly related to one another (Figure 17). Areas of difference between these phages were the tail regions, which account for variations in host range, and at the right end of the genomes, where the small-unidentified genes reside. Homology between ϕELB20 and ϕR4 was much more extensive, indicating that these are essentially the same bacteriophage differing only in small point mutations and in a few, small unidentified genes (see 5.3). ϕHau3AI, although not indigenous to this soil, also displayed homology with ϕELB20 and ϕELB8. This occurred in areas encoding essential proteins such as the terminase and portal proteins. In contrast, the main areas of difference between ϕR4, ϕHau3AI, ϕELB20 and ϕELB8 were typically towards the right of the genomes containing small genes of unknown function, therefore, the observed differences between these phages may be due to these genes, for example, the differences in host range and EOP between ϕELB20, ϕHau3AI and ϕELB8 (Table 21).

It can also be speculated, by examining the literature, that bacteriophage Bα and SH10 are related to the phages in the soil under study due to their apparent relationship to ϕR4. (Ishihara *et al.*, 1982; Lomovkskaia, 1980). Furthermore, upon examination of genome architecture (Figure 35), relationships with many other phages, such as L5 and TG1, were evident. In conclusion, there appeared to be a dominant R4-like genomic architecture in actinophages as this formation was suited to soil. The global predominance of this phage type in soils is indicative of its ability to cross bacterial niches hence its universal abundance. Moreover, small non-essential genes varied between viruses in order to allow their evolution and adaptation in the event of an environmental change in their surroundings.

This thesis has evaluated the relationships of phages and hosts in soil and their survival mechanisms. It can be concluded that soil contains uniquely adapted bacterial

populations, alongside their bacteriophages, which, are less abundant and lack the global diversity of their hosts due to selective pressures. Resident bacteria may develop resistance mechanisms to these phages, however, the majority do not, due to the overall small impact they have on their host populations within soil. However, the impact they do have is responsible for global nutrient turnover and evolution of bacteria It is apparent that one dominant phage genomic architecture in soil is common to all actinophages due to their sharing of a common ancestor. Although the streptomycetes and their phages have adapted well to their soil environment, it became clear that highly adaptable, non-indigenous bacteria are capable of surviving as well as indigenous bacteria.

Overall the main area of future work should focus on the full characterisation of all the phages in the panel in order to obtain greater understanding of all the phages infecting streptomycetes. This would include experiments conducted throughout this study (for example, one-step growth experiments and EOP readings). Furthermore, a useful aid in their classification would be by their examination under the electron microscope. In addition, microscopy and biochemical tests of the host organisms would further classify them by application of the numerical approach. Following this phages, clearly different from ϕELB20 and ϕR4, would be examined by homology experiments in order to determine the genes responsible for any differences, thus revealing relationships between the populations. Moreover, monitoring populations of phages and hosts in non-sterile soil would provide a more realistic picture of phage-host dynamics in soil.

7. References

Abaye, D., Lawlor, K., Hirsch, P., and Brookes, P. (2004) Changes in the microbial community of an arable soil caused by long term metal contamination. *European Journal of Soil Science* **56**: 93-102.

Abdel, K., deSchmertzing, H., and Peterson, J. (1963) Classification of microorganisms by analysis of chemical composition *Journal of Bacteriology* **85**: 1039-1044.

Ackermann, H., Berthiaume, L., and Jones, L. (1985) New actinophage species. *Intervirology* **23**: 121-130.

Ackermann, H., and Kropinski, A. (2007) Curated list of prokaryote viruses with fully sequenced genomes. *Research in Microbiology* **158**: 555-566.

Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, J., and Lipman, D. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucelic Acid Reseach* **25**: 3389-3402.

Anderson, A., and Wellington, E. (2001) The taxonomy of *Streptomyces* and related genera. *International Journal of Systematic and Evolutionary Microbiology* **51**: 797-814.

Anderson, T. (2003) Microbial eco-physiological indicators to asses soil quality *Agriculture, Ecosystems and Environment* **98**: 285-293.

Andrew, J., and Harris, R. (1986) r- and K- selection and microbial ecology. *Advances in Microbial Ecology* **9**: 99-147.

Anukool, U., Gaze, W., and Wellington, E. (2004) In situ monitoring of streptothricin production by *Streptomyces rochei* F20 in soil and rhizosphere. *Applied and Environmental Microbiology* **70**: 5222-5228.

Arora, A., Nain, L., and Gupta, J. (2005) Solid-state fermentation of wood residues by *streptomyces griseus* B1, a soil isoalte, and solubilization of lignins. *World Journal of Microbiology & Biotehnology* **21**: 303-308.

Ashelford, K., Norris, S., Fry, J., Bailey, M., and Day, M. (2000) Seasonal population dynamics and interactions of competing bacteriophages and their host in the Rhizosphere. *Applied and Environmental Microbiology* **66**: 4193-4199.

Ashelford, K., Day, M., and Fry, J. (2003) Elevated abundance of bacteriophage infecting bacteria in soil. *Applied and Environmental Microbiology* **69**: 285-289.

Association, A.P.H. (1917) Standard methods of water analysis *American Public Health Association* **3rd Ed**.

Athalye, M., Lacey, J., and Goodfellow, M. (1981) Selective isolation and enumeration of actinomycetes using rifampicin. *Journal of Applied Bacteriology* **51**: 289-297.

Athalye, M., Noble, W., and Minnikin, D. (1985) Analysis of cellular fatty acids by gas chromatography as a tool in the indentification of medically important coryneform bacteria. *Journal of Applied Bacteriology* **58**: 507-512.

Azam, F., Enchel, T., Field, J., and Graf, J. (1983) The ecological role of water-column microbes in the Sea. *Marine Ecology* **10**: 257-263.

Balan, A., and Padilla, G. (1997) New thermal inducible *Streptomyces* phages isolated from tropical soils. *Brazilian Journal of Genetics* **20**.

Bedford, D., Laity, C., and Buttner, M. (1995) Two genes invovled in the phase-variable phiC31 resistance mechanism of *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **177**: 4681-4689.

Bent, S., Gucker, C., Oda, Y., and Forney, L. (2003) Spatial distribution of *Rhodopseudomonas palustris* ecotypes on a local scale. *Applied and Environmental Microbiology* **69**: 5192-5197.

Bentley, S., Chater, K., Cerdeno-Tarraga, A., Challis, G., and Thomson, N. (2002) Complete genome sequence of model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**: 141-147.

Berkane, E., Orlik, F., Stegmeier, A., Winterhalter, M., and Benz, R. (2006) Interaction of Bacteriophage Lambda with Its Cell Surface Receptor: An in Vitro Study of Binding of the Viral Tail Protein gpJ to LamB (Maltoproin). *Biochemistry* **45**: 2708-2720.

Bik, E., Bunschoten, A., Gouw, R., and Mooi, R. (1995) Genesis of the novel epidemic Vibrio cholerae 0139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis *The EMBO Journal* **14**: 209-216.

Birnboim, H., and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Research* **7**: 1513-1979.

Blokhina, T., Kostrikina, N., Sakharova, Z., and Biryunzova, V. (1992) Comparative electron microscope study of *Bacillus thuringiensis grown in chemostat*. *Microbiology* **61**: 477-481.

Bradley, D. (1961) Negative staining of bacteriophage phiR at various pH values. *Virology* **61**.

Braissant, O., Verrecchia, E., and Aragno, M. (2002) Is the contribution of bacteria to terrestrial carbon budget greatly underestimated? *Naturwissenschaften* **89**: 366-370.

Brett, P., Deshazer, D., and Woods, D. (1998) *Burkholderia thailandensis* sp. nov. a *Burkholderia pseudomallei*-like species. *International Journal of Systematic and Evolutionary Microbiology* **48**: 317-320.

Brussaard, C. (2009) Enumeration of bacteriophages using flow cytometry *Methods in Molecular Biology* **501**: 97-111.

Brussow, H., Canchaya, C., and Hardt, W. (2004) Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiology Molecular Biology Reviews* **68**: 560-602.

Burnet, F., and Lush, D. (1935) The staphylococcal bacteriophages. *Journal of Pathology & Bacteteriology* **40**: 455-469.

Burroughs, N., Marsh, P., and Wellington, E. (2000) Mathematical analysis of growth and interaction dynamics of Streptomyces and bacteriophage in soil. *Applied and Environmental Microbiology* **66**: 3868-3877.

Camus, J., Pryor, M., Medigue, C., and Cole, S. (2002) Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**: 2967-2973.

Carvajal, F. (1946) Biological strains of *Streptomyces griseus*. *Mycologia* **38**: 596-607.

Casadaban, M., and Cohen, S. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *Journal of Molecular Biology* **138**: 179-207.

Casida, L. (1965) Abundant Microorganism in Soil. *Applied Microbiology* **13**: 327-334.

Cerdeno-Tarraga, A., Efstratiou, A., Dover, L., and Holden, M. (2003) The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Research* **15**: 3516-3523.

Chater, K., and Carter, A. (1979) A new wide-host range, temperate bacteriophage (R4) of *Streptomyces* and its interaction with some restriction-modification systems. *Journal of General Microbiology* **115**: 431-442.

Chater, K. (1986) Streptomyces phages and their applications to *Streptomyces* genetics: 119-158.

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M., and Brussow, H. (2004) Phage-host interaction: an ecological perspective. *Journal of Bacteriology* **186**: 3677-3686.

Chinenova, T., Mkrtumian, N., and Lomovskaia, N. (1982) Genetic characteristics of a new phage resistance trait in *Streptomyces coelicolor* A3(2). *Genetika* **18**: 1945-1952.

Clark, A., Inwood, W., Cloutier, T., and Dhillon, T. (2001) Nucleotide sequence of coliphage HK620 and the evolution of lamboid phages. *Journal of Molecular Biology* **311**: 657-679.

Cole, S., Brosch, R., Parkhill, J., and Garnier, T. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **11**: 537-544.

Conn, H.J. (1917) A possible function of the actinomycetes in soil. *Journal of Bacteriology* **1**: 197-207.

Cox, K., and Baltz, R. (1984) Restriction of bacteriophage plaque formation in Streptomyces spp. *Journal of Bacteriology* **159**: 499-504.

Cresswell, N., Herron, P., Saunders, V., and Wellington, E. (1992) The fate of introduced streptomycetes plasmid and phage populations ina dynamic soil system. *Journal of General Microbiology* **138**: 659-666.

Cummins, C., and Harris, H. (1956) The relationship between certain members of the staphylococcus-micrococcus group as shown by their cell wall composition *International Bulletin of bacterial nomenclature and taxonomy* **6**.

DA Hopwood, M.B., MJ Bibb, T Kieser, KF Chater (2000) *Practical Streptomyces Genetics*.

Dale, J., and Greenaway, P. (1984) Preperation of Lambda DNA. *Nucelic Acids* **2**: 211-215.

Daly, C., Fitzgerald, G., and Davis, R. (1996) Biotechnology of lactic acid bacteria with special reference to bacteriophage resistance. *Antione Van Leeuwenhoek* **70**: 99-110.

Dancer, S. (2004) How do we assess hospital cleaning? A proposal for microbiological standards for surface hygiene in hospital. *Journal of Hospital Infection* **56**: 10-15.

Danovaro, R., and Serresi, M. (2000) Viral density and virus to bacterium ratio in deep-sea sediments of the Eastern Mediterranean. *Applied and Environmental Microbiology* **66**.

Danovaro, R.M., E Dell'Anno, A (2002) Hgiher abundance of bacteria than of viruses in deep Mediterranean sediments. *Applied and Environmental Microbiology* **67**: 1468-1472.

de Vos, W. (1987) Gene cloning and expression in lactic streptococci. *FEMS Microbiology Reviews* **46**: 281-295.

Deveau, H.V.C., MR Moineau, S (2002) Effect of Exopolysaccharides on Phage-Host Interactions in Lactococcus lactis. *Applied and Environmental Microbiology* **68**: 4364-4369.

Doublie, S.T., S Long, AM Richardson, CC Ellenberger, T (1997) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2A resolution. *Nature* **391**: 251-258.

Dowding, J. (1973) Characterization of a bacteriophage virulent for *Streptomyces coelicolor* A3(2). *Journal of General Microbiology* **76**: 163-176.

Duda, R., Martincic, K., Xie, Z., and Hendrix, R. (1995) Bacteriophage HK97 head assembly *FEMS Microbiology Reviews* **17**: 41-46.

Dunne, E., Burman, W., and Wilson, M. (1998) *Streptomyces pneumonia* in a patient with human immunodeficiency virus infection: case report and review of the literature on invasive streptomyces. *Clinical Infectious Diseases* **37**: 93-96.

Dykhuizen, D., and Green, L. (1991) Recombination in *Eschierichia coli* and the definition of a biological species. *Journal of Bacteriology* **173**: 7257-7268.

Dyson, P., and Schrempf, H. (1987) Genetic instability and DNA amplification in *Streptomyces lividans* 66. *Journal of Bacteriology* **169**: 4796-4803.

Egamberdlyeva, D., and Hoflich, G. (2003) Influence of growth-promoting bacteria on the growth of wheat in different soils and temperatures. *Soil Biology and Biochemistry* **35**: 973-978.

Eigen, M. (1993) Viral quasispecies. *Scientific American* **269**: 32-39.

Ereshefsky, M. (1994) Some problems with Linnaean hierarchy. *Philosophy of Science* **61**: 186-205.

Erikson, D. (1947) Differentation of the vegatitive and sporogenous phases of actinomycetes. *Journal of General Microbiology* **1**.

Escobar-Paramo, P., Ghosh, S., and DiRuggiero, J. (2005) Evidence for genetic drift in the diversification of a geographically isolated population of the hyperthermophillic archaeon *Pyrococcus*. *Molecular Biology Evolution* **22**: 2297-2303.

Farkasovska, J., Klucar, L., Vlcek, C., Kokavecc, J., and Godany, A. (2007) Complete genome sequence and analysis of the *Streptomyces aureofaciens* phage mu1/6. *Folia Microbiologica* **52**: 347-358.

Farrar, M., Howson, K., Bojar, R., West, D., and Towler, J. (2007) Genome Sequence and Analysis of a Propionibacterium acnes Bacteriophage. *Journal of Bacteriology* **189**: 4161-4167.

Faruque, S., Naser, I., Islam, M., and Faruque, A. (2005) Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *PNAS* **102**: 1702-1707.

Feiss, M., and Winder, W. (1982) Bacteriophage lambda DNA packing: scanning for the terminal cohesive end site during packing. *Proc. Natl Acad Sci* **79**: 3498-3502.

Fierer, N., and Jackson, R. (2006) The diversity and biogeography of soil bacterial communities. *PNAS* **103**: 626-631.

Fischer, C., Yoichi, M., Unno, H., and Tanji, Y. (2004) The coexsistance of *Escherichia coli* serptype O157:H7 and its specific bacteriophage in continuous culture. *FEMS Microbiology Letters* **241**: 171-177.

Flerer, N., Schimel, J., and Holden, P. (2003) Variations in microbial community composition through two soil depths *Soil Biology and Biochemistry* **35**: 167-176.

Foor, F.M., N *et al* (1990) Construction of a shuttle vector consisting of the *Escherichia coli* plasmid pACYC117 inserted into the *Streptomyces catteya* phage TG1. *Gene* **39**.

Fox, G., Magrum, L., Balch, W., Wolfe, R., and Woese, C. (1977) Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *PNAS* **74**: 4537-4541.

Fox, G. (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guaranty species identity. *International Journal of Systematic Bacteriology* **42**: 166-170.

Fuhrman, J. (1999) Marine viruses and their biogeochemical and ecological effects *Nature* **399**: 541-548.

Garvey, P., Hill, C., and Fitzgerald (1996) The Lactoccocal plasmid pNP40 encodes a third bacteriophage resistance mechanism, one which affects phage DNA penetration. *Applied and Environmental Microbiology* **62**: 676-679.

Gevers, D., Cohan, F., Lawrence, J., and Spratt, B. (2005) Re-evaluating prokaryotic species. *Nature Reviews* **3**: 733-739.

Ginsburg, D., and Calos, M. (2005) Site-specific integration with phiC31 integrase for prolonged expression of therapeutic genes. *Advances in Genetics* **54**: 179-187.

Godde, J., and Bickerton, A. (2006) The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal gene transfer among prokaryotes *Journal of Molecular Evolution* **62**: 718-729.

Gohosh, P., Kim, A., and Hatfull, G. (2003) The orientation of mycobacteriophage BxbI integration is soley dependant on the central dinucletide of *attP* and *attB*. *Molcular Cell* **12**: 1101-1111.

Goody, D., Randle, G., Simpson, A., Aanensen, D., Pitt, T., Kinoshita, R., and Spratt, B. (2003) Multilocus sequence typing and evolutionary relationships among the causative agents of melloidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. *Journal of Clinical Microbiology* **41**: 2068-2079.

Gopul, P., and Reilly, K. (1995) Molecular architecture of lactococcal cell-surface regulates important industrial properties. *International Dairy Journal* **5**: 1095-1111.

Gottesman, S., and Abremski, K. (1982) Purification of the bacteriophage xis gene product required for excisive recombination. *Journal of Biological Chemistry* **257**: 9658-9662.

Govan, J., Hughes, J., and Vandamme, P. (1996) *Burkholderia cepacia*: medical, taxonomic and ecological issues. *Jouurnal of Clinical Microbiology* **45**: 395-407.

Gray, D., Gooday, G., and Prosser, J. (1990) Apical hyphal extension in *Streptomyces coelicolor* A3(2). *Journal of General Microbiology* **136**: 1077-1084.

Gregory, M.T., R Smith, MCM (2003) Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *Journal of Bacteriology* **185**: 5320-5323.

Grkovic, S.B., MH Hardie, KM Firth, N Skurray, RA (2003) Stable low copy number *Staphylococcus aureus* shuttle vectors. *Microbiology* **149**: 785-794.

Hahn, D.M., MA Baltz, RH (1991) Properties of the Streptomycete temperate bacteriophage FP43. *Journal of Bacteriology* **173**: 3770-3775.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids *Journal of Molecular Blology* **166**: 557-580.

Hatfull, G.S., GJ (1993) DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Molcular Microbiology* **7**: 395-405.

Hayakawa, M.Y., Y Limura, Y (2004) Selective isoaltion of bioactive soil actinomycetes belonging to the *Streptomyces violaceusniger* phenotypic cluster. *Journal of Applied Microbiology* **96**: 973-981.

Hazelbauer, G. (1975) Role of the receptor for bacteriophage lambda in the functioning of maltose chemoreceptor of *Escherichia coli*. *Journal of Bacteriology* **124**.

Hendrix, R., Hatfull, G., Burns, R., Ford, M., and Smith, M. (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *PNAS* **96**: 2192-2197.

Hendrix, R., Lawrence, J., Hatfull, G., and Casjens, S. (2000) The origins and ongoing evolution of viruses *Trends in Microbiology* **8**: 504-508.

Hendrix, R., Hatfull, G., and Smith, M. (2003) Bacteriophages with tails: chasing their origins and evolution. *Research in Microbiology* **154**: 253-257.

Hennes, K.S., CA Chan, AM (1995) Fluorescently labeled virus probes show that natural virus populations can control the structure of marine microbial communities *Applied and Environmental Microbiology* **61**: 3623-3627.

Herron, P., and Wellington, E. (1990) New method for extraction of Streptomycete spores from soil and application to the study of lysogeny in sterile amended and nonsterile soil. *Applied and Environmental Microbiology* **56**: 1406-1412.

Heuer, H., Krsek, M., Baker, P., Smalla, K., and Wellington, E. (1997) Analysis of actinomycete communities by specific amplification og genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Applied and Environmental Microbiology* **63**: 3233-3241.

Hey, J. (2001) Genes, catogeries and species. *Oxford University Press, NY.*

Hirsch, C., and Christensen, D. (1983) Novel method for selective isolation of actinomycetes. *Applied and Environmental Microbiology* **46**: 925-929.

Hobbs, G., Frazer, C., Gardiner, D., Cullum, J., and Oliver, S. (1989) Dispersed growth of *Streptomyces* in liquid culture. *Applied Microbiological Biotechnology* **31**: 272-277.

Holmfeldt, K., Middleboe, M., Nybroe, O., and Riemann, L. (2007) Large variabilities in host strain sucesptibility and phage host range govern interactions between lytic marine phages and their *Flavobacterium* hosts. *Applied and Environmental Microbiology* **73**: 6730-6739.

Hopwood, D. (1960) Phase-contrast observations on *Streptomyces coelicolor Journal of General Microbiology* **22**: 295-302.

Hopwood, D., and Glauret, A. (1961) Electron microscope obsevations on the surface structures of *Streptomyces violaceoruber*. *Microbiology* **26**: 325-330.

Hopwood, D., Harold, R., Vivian, A., and Ferguson, H. (1969) A new kind of fertilitiy variant in *Streptomyces coelicolor Genetics* **62**: 461-477.

Hopwood, D., Kieser, T., Wright, H., and Bibb, M. (1983) Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66. *Journal of General Microbiology* **129**: 2257-2269.

Hopwood, D., Bibb, M., Kieser, T., Chater, K., and Buttner, M. (2000) Practical *Streptomyces* genetics. *John Innes Centre, Norwich* **Crowes, Norwich, England**.

Hotchin, J. (1951) The influence of acridines on the interaction of Staphylococcus aureus and Staphylococcus K phage. *Journal of General Microbiology* **5**: 609-618.

Hsu, S., and Lockwood, J. (1975) Powdered chitin agar as selective medium for enumeration of actinomycetes in water and soil. *Applied Microbiology* **29**: 422-426.

Ikeda, H., Kotaki, H., and Omura, S. (1987) Genetic studies of Avermectin biosynthesis in *Streptomyces avermitilis*. *Journal of Bacteriology* **169**: 5615-5621.

Ikeda, H., Ishikawa, J., Hanamoto, A., and Shinose, M. (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis Nature Biotechnology* **21**: 526-531.

Ikeda H, I.J., Hanamoto A, Shinose M *et al* (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis Nature Biotechnology* **21**: 526-531.

Ishihara, H., Nakano, M., and Ogawara, H. (1982) Restriction enzyme mapping of the DNA of *Streptomyces* bacteriophage B-alpha and its deletion derivaties. *Journal of Bacteriology* **152**: 1288-1291.

Jansen, R., van Embden, J., Gaastra, W., and Schouls, L. (2002) Identification of genes that are associated with DNA repeats in prokaryotes *Molecular Microbiology* **43**: 1565-1575.

Japan Science, T.A. (2008). *http://www-btls.jst.go.jp/cgi-bin/Tools/dotBLAST/index.cgi*.

Jensen, E., Schrader, H., Rieland, T., and Thompson, K. (1998) Prevelance of broad host-range lytic bacteriophages of *Sphaerotilus natans, Escherichia coli,* and *Pseudomans aeruginosa*. *Applied and Environmental Microbiology* **64**: 575-580.

Johnson, R., and Bruist, M. (1989) Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombinational enhancer in the strand exchange reaction. *The EMBO Journal* **8**.

Joo, C., McKinney, S., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006) Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **126**: 515-527.

Juhala, R., Ford, M., Duda, R., Youlton, A., Hatfull, G., and Hendrix, R. (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *Journal of Molecular Biology* **299**: 27-51.

Karam, J. (2005) Bacteriophages: The viruses for all seasons of molecular biology. *Virology* **2**: 1-2.

Kataoka, M., ueda, K., and Kudo, T. (1997) Application of the variable region in 16S rDNA to create an index for rapid species identification in the genus *Streptomyces*. *FEMS Microbiology Letters* **151**: 249-255.

Keel, C., Ucurum, Z., Michaux, P., Adrian, M., and Haas, D. (2002) Deleterious impact of a virulent bacteriophage on survival and biocontrol activity of *Pseudomonas fluorescens* strain CHA0 in natural soil. *Molecular Plant Microbe Interactions* **15**: 567-576.

Khodakaramian, G., Lissenden, S., Gust, B., Moir, L., Hoskisson, P., Chater, K., and Smith, M. (2006) Expression of Cre recombinase during transient phage infection permits efficient marker removal in *Streptomyces*. *Nucleic Acids Research* **34**: 1-5.

Kieser, T., Bibb, M., Buttner, M., Chater, K., and Hopwood, D. (1985) Practical *Streptomyces* Genetics. *John Innes Foundation* **0-7084-0623-8**.

Kim, S., Seong, C., Jeon, S., Bae, K., and Goodfellow, M. (2004) Taxanomic study of neutrotolerant acidophilic actinomycetes isolated from soil and description of *Streptomyces yeochonensis* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **54**: 211-214.

Kirby, R., and Rybicki, E. (1986) Enzyme-linked immunosorbent assay (ELISA) as a means of taxanomic analysis of *Streptomyces* andrelated organisms. *Journal of General Microbiology* **132**: 1891-1894.

Klieneberger-Nobel, E. (1947) The lifecycle of sporing *Actinomyces* as revealed by a study of their structure and sepation. *Journal of General Microbiology* **1**.

Koch, A. (2001) Oligotrophs versus copiotrophs. *BioEssays* **23**: 657-661.

Korn-Wendisch, F., and Schneider, J. (1992) Phage typing - a useful tool in actinomycete systematics. *Gene* **115**: 243-247.

Kowalczykowski, S. (2000) Initiation of genetic recombination and recombination-dependant replication. *Trends in Biochemistry* **25**: 156-165.

Kowalski, T., Berbari, E., and Osmon, D. (2005) Epidenology, treatment and prevention of community aquired methicillin-resistant *Staphylococcus aureus* infections. *Mayo Clinical Proceedings* **80**: 62-79.

Krainsky, A. (1914) Die Aktinomyceten und ihren Bedeutung in der Natur. *Bakteriol Parasitenkd Infektionskr Hyg Abt II* **41**: 649-688.

Kramer, A., Schwebke, I., and Kampf, G. (2006) How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infectious diseases* **6**: 1-8.

Kreiswirth, B., Lofdahl, S., Betley, M., and O'Reilly, M. (1983) The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305**: 709-712.

Kremmerling, C., Witt, C., Liesack, W., Weyland, H., and Stackenbrandt, E. (1989) Approaches for the molecular identification of streptomycetes in marine environment *Current Topics in Marine Biotechnology* **Fuji Technology Press Ltd**: 423-426.

Krugel, H., Fiedler, G., and Noack, D. (1980) Transfection of protoplasts from *Streptomyces lividans 66* with actinophage SH10 DNA. *Molecular and General Genetics* **177**: 297-300.

Kurtboke, I. (2005) Actinophages as indicators of actinomycete taxa in marine environments. *Antonie van Leeuwenhoek* **87**: 19-28.

Kusano, K., Naito, T., Handa, N., and Kobayashi, I. (1995) Restriction-modification systems as genomic parasites in competition for specific sequences. *Proc. Natl Acad Sci* **92**: 11095-11099.

Kwan, T., Liu, J., DuBow, M., Gros, P., and Pelletier, J. (2006) Comparative genomic analysis of 18 Pseudomonas aeruginosa bacteriophages. *Journal of Bacteriology* **188**: 1184-1187.

Labeda, D. (1996) DNA Relatedness among Verticil-forming *Streptomyces* species (formerlly *Streptoverticillium* species). *International Journal of Systemic Bacteriology* **46**: 699-703.

Laity, C., Chater, K., Lewis, C., and Buttner, M. (1993) Genetic analysis of the phiC31-specific phage growth limitation (pgl) system of *Streptomyces coelicolor* A3(2) *Molecular Microbiology* **7**: 329-336.

Lanning, S., and Williams, S. (1982) Methods for the direct isolation and enumeration of actinophages in soil. *Journal of General Microbiology* **128**: 2063-2071.

Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, A., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Gibson, T., Thompson, J., and Higgins, D. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Lawrence, J., Hatfull, G., and Hendrix, R. (2002) Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches *Journal of Bacteriology* **184**: 4891-4905.

Leblond, P., Redenbach, M., and Cullum, J. (1993) Physical map of the *Streptomyces lividans* 66 genome and comparison with that of the related strain *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **175**: 3422-3429.

Lechevalier, M., and Lechevalier, H. (1970) Chemical composition as a criterion in the classification of aerobic actinomycetes. *International Journal of Systematic Bacteriology* **20**: 435-443.

Lennox, E. (1955) Transduction of linked genetic charaters of the host by bacteriophage P1. *Virology* **1**: 190-206.

Levin, B. (1993) The accessory genetic elements of bacteria: existence conditions and (co)evolution. *Current Opinion in Microbiology* **3**.

Lis, J., and Schleif, R. (1975) Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucelic Acids Research* **2**: 383-390.

Lloyd, A. (1969) Behaviour of sreptomycetes in soil. *Journal of General Microbiology* **57**: 35-40.

Lofblom, J., Kronqvist, N., Uhlen, M., and Wernerus, H. (2006) Optimization of electroporation-mediated transformation: *Staphylococcus carnosus* as model organism. *Journal of Applied Microbiology* **102**: 736-747.

Lomovkskaia, N.C., KF Mkrtumian, NM (1980) Genetics and molecular biology of *Streptomyces* bacteriophages. *Microbiology Reviews* **44**: 206-229.

Lomovskaya, N., Mkrtumian, N., Gostimskaya, N., and Danilenko, V. (1972) Characterization of temperate actinophage phiC31 isolated from *Streptomyces coelicolor* A3(2). *Journal of Virology* **9**: 258-262.

Luria, S., and Burrous, J. (1957) Hybridization between *Escherichia coli* and *Shigella*. *Journal of Bacteriology* **74**: 461-476.

MacArthur, R., and Wilson, E. (1967) The theory of island biogeography. *Priceton University Press*.

Magrini, V., Salmi, D., Thomas, D., Herbert, S., Hartzell, P., and Youderian, P. (1997) Temperate *Myxococcus xanthus* phage Mx8 encodes a DNA adenine methylase, Mox. *Journal of Bacteriology* **179**: 4524-4263.

Mao, J., Tang, Q., Zhang, Z., Wang, W., and Wei, D. (2007) *Streptomyces radiopugnans* sp. nov., a raditation-resistant actinomycete isolated from radiation-polluted soil in China *International Journal of Systematic and Evolutionary Microbiology* **57**: 2578-2582.

Marsh, P., Toth, I., Meijer, M., Schilhabel, M., and Wellington, E. (1993) Survival of the temperate actinophage phiC31 and *Streptomyces lividans* in soil and the effects of competition and selection on lysogens. *FEMS Microbiology Ecology* **13**: 13-22.

Martinsohn, J., Radmam, M., and Petit, M. (2008) The λ Red Proteins Promote Efficient Recombination between Diverged Sequences: Implications for Bacteriophage Genome Mosaicism. *PLoS Genetics* **4**: 1-12.

Matsuura, M., Noguchi, T., Yamaguchi, D., and Aida, T. (1996) The *sre* gene (ORF 469) encodes a site specific recombinase responsible for integration of the R4 genome. *Journal of Bacteriology* **178**: 3374-3376.

Mattis, A., Gumport, R., and Gardner, J. (2008) Purification and characterization of bacteriophage P22 hix protein. *Journal of Bacteriology* **190**: 5781-5796.

Mayfield, C., Williams, S., Ruddick, S., and Hatfield, H. (1972) Studies on the ecology of actinomycetes in soil IV. Observations on the form and growth of of streptomycetes in soil. *Soil Biology and Biochemistry* **4**: 79-91.

McCaig, A., Glover, L., and Prosser, J. (2001) Numerical analysis of grassland bacterial community structure under different land management regimens by using 16S ribosomal DNA sequence data and denaturing gradient gel electrophoresis banding patterns. *Applied and Environmental Microbiology* **67**: 4554-4559.

McLaren, A., Bollag, J., Ladd, J., and Paul, E. (1996) Soil Biochemistry. *CRC Press*.

Mehling, A., Wehmeier, U., and Piepersberg, W. (1995) Amplification of random amplified polymorphic DNA (RAPD) assays in identifing conserved regions of actinomycete genomes. *FEMS Microbiology Letters* **128**: 119-126.

Mellaert, L., Lammertyn, L., Schacht, S., and Anne, J. (1998) Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. *Microbiology* **144**: 3351-3358.

Middelboe, M. (2000) Bacterial growth rate and marine virus-host dynamics *Microbial Ecology* **40**: 114-124.

Moebus, K., and Nattkemper, H. (1981) Bacteriophage sensitivity paterns among bacteria isoalted from marine waters. *Helgolander. Meeresunters* **36**: 375-385.

Moore, J., and Haber, J. (1996) Cell cycle and genetic requirements of two pathways of nonhomologous end-joinging repair of double strand breaks in *Saccharamoyces cerevisiae*. *Molecular and Celllular Biology* **16**: 2164-2173.

Morona, R., and Henning, U. (1984) Hoat range mutants of bacteriophage Ox2 can use two different outer membrane proteins of *Escherichia coli* K-12 as receptors. *Journal of Bacteriology* **159**: 579-582.

Naito, T., Kusano, K., and Kobayash, I. (1995) Selfish behavior of restriction-modification systems. *Science* **267**.

ND Lomovkskaia, L.E., SI Alikhanian (1971) The genetic location of prophage on the chromosome of *Streptomycyes coelicolor*. *Genetics* **68**: 341-347.

Neely, A., and Maley, M. (2000) Survival of enterococci and staphylococci on hospital fabric and plastic. *Journal of Clinical Microbiology* **38**: 724-726.

Nishimura, M., Ooi, O., and Davies, J. (2006) Isolation and characterisation of *Streptomyces* sp. NL15-2K capable of degrading lignin-related aromatic compounds. *Journal of Bioscience and Bioengineering* **102**: 124-127.

Nodwell, J., Yang, M., Kuo, D., and Losick, R. (1999) Extracellular complementation and the identification of additional genes invovled in aerial mycelium formation in *Streptomyces coelicolor*. *Genetics* **151**: 569-584.

Norrander, J., Kempe, T., and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**: 101-106.

Ohnishi, Y., J, I., Hara, H., Suzuki, H., and Ikenoya, M. (2008) The genome sequence of the Streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *Journal of Bacteriology* **1**.

Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., and Takahashi, C. (2001) Genome sequence of an industiral microorganism*Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. *Microbiology* **98**: 12215-12220.

Pacsuta, P., Kodobocz, L., and Biro, B. (2005) Drought-driven respond of fast and slow growing components of an adapted clover *Rhizobium* strain. *Acta Biologica Szegediensis* **49**: 167-169.

Palmano, S., Firrao, G., and Locci, R. (2000) Sequence analysis of domains III and IV of the 23S rRNA gene of verticillate streptomycetes. *International Journal of Systematic and Evolutionary Microbiology* **50**: 1187-1191.

Pantastico-Caldas, M., Duncan, K.I., stock, C., and Bell, J. (1992) Population dynamics of bacteriophage and *Bacillus subtilis* in soil. *Ecology* **73**: 1888-1902.

Papke, R., Ramsing, N., Bateson, M., and Ward, D. (2003) Geographical isolation in hot spring cyanobacteria. *Environmental Microbiology* **5**: 650-659.

Patten, B., and Auble, G. (1981) The theory of the ecological niche. *The American Naturalist* **117**: 893-922.

Pedulla, M., Ford, M., Houtz, J., Karthikeyan, C., Wadsworth, J., Lewis, D., Jacobs-Sera, J., Falbo, J., Gross, J., Pannunziou, N., and Brucker, V. (2003a) Origins of highly mosaic mycobacteriophage genomes *Cell* **113**: 171-182.

Pedulla, M., Ford, M., Karthikeyan, T., Houtz, J., and Hendrix, R. (2003b) Corrected Sequence of the Bacteriophage P22 Genome. *Journal of Bacteriology* **185**: 1475-1477.

Philipp, W., Poulet, S., Eiglmeier, L., and Pascopella, L. (1996) An integrated map of the genome of the tubercle bacillus *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proc. Natl Acad Sci* **93**.

Pidwirny, M. (2006) Introduction to Soils. *Fundamentals of Physical Geography* **2nd Edition**.

Plaskitt, K., and Chater, K. (1995) Influences of the developmental genes on localised glycogen deposition in colonies of a mycelial prokaryote *Streptomuces coelicolor* A3(2): a possible interface between metabolism and morphogenesis. *JSTOR* **347**: 105-121.

Popa, M., McKelvey, T., Hempel, J., and Hendrix, R. (1991) Bacteriophage HK97 structure: wholesale covalent cross-linking between the major head shell subunits. *Journal of Virology* **65**: 3227-3237.

Porter, J., Wilhelm, J., and Tresner, H. (1960) Method for the preferential isolation of actinomycetes from soils. *Applied Microbiology* **8**: 174-178.

Pridham, T., Hesseltine, C., and Benedict, R. (1958) A guide for the classification of streptomycetes according to selected groups. *Applied Microbiology* **6**: 52-79.

Pringsulaka, O., Chavanich, S., Doi, K., and Ogata, S. (2002) Changes in the population of a wide host range actinophage isoalted from Thai soil and host Streptomycetes in Thai and Japanese soil. *Actinomycetol* **16**: 21-25.

Radajewski, S., and Duxbury, T. (2001) Motility responses and desiccation survival of zoospores from the Actinomycete *Kineosporia*sp. strain SR11. *Microbial Ecology* **41**: 233-244.

Ranjard, L., and Richaume, A. (2001) Quantitative and qualitative microscale distribution of bacteria in soil. *Research Microbiology* **152**: 707-716.

Rauland, U., Glocker, I., Redenbach, M., and Cullum, J. (1995) DNA amplifications and deletions in *Streptomyces lividans 66* and the loss of one end of the linear chromosome. *Molecular General Genetics* **241**: 37-44.

Ravin, V., Ravin, N., Casjens, S., Ford, M., Hatfull, G., and Hendrix, R. (2000) Genomic sequence and analysis of the atypical temperate bacteriophage N15. *Journal of Molecular Biology* **299**: 53-73.

Recktenwald, J., and Schmidt, H. (2002) The nucleotide sequence of shiga toxin (Stx) 2e-encoding phage phiP27 is not related to other Stx genomes, but the modular genetic structure is conserved. *Infection and Immunity* **70**: 1896-1908.

Rees, P., and Fry, B. (1981) The morphology of staphylococcal bacteriophage K and DNA metabolism in infected *Staphylococcus aurues Journal of General Virology* **53**: 293-307.

Retallack, D., Johnson, L., and Friedman, D. (1994) Role for 10Sa RNA in the growth of lambda-P22 hybrid phage *Journal of Bacteriology* **176**: 2082-2089.

Riemann, L., and Middelboe, M. (2002) Viral lysis of marine bacterioplankton: implications for organic matter cycling and bacterial clonal composition. *Ophelia* **56**: 57-68.

Rocha, E., Danchin, A., and Viari, A. (2001) Evolutionary rols of restriction/modification systems as revealed by compartative genome analysis. *Genome Reseach* **11**: 946-958.

Rogul, M., Brendle, J., Haapaka, D., and Alexandar, A. (1970) Nucleic acid similarities among *Pseudomonas pseudomallei, Pseudomonas multivrans* and *Actinobacillus mallei*. *Journal of Bacteriology* **101**: 827-835.

Rohwer, F., and Edwards, R. (2002a) The Phage Proteomic Tree: a genome-based taxonomy for phage. *Journal of Bacteriology* **184**: 4529-4535.

Rohwer, F., and Edwards, R. (2002b) The phage proteomic tree: a genome-based taxonomy for phage. *Journal of Bacteriology* **184**: 4529-4535.

Ruddick, S., and Williams, S. (1972) Studies on the ecology of actinomycetes in soil. *Soil Biology and Biochemistry* **4**: 93-103.

Ryding, J., Bibb, M., Molle, V., and Fundlay, K. (1999) New sporulation loci in *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **181**: 5419-5425.

Sahin, N. (2004) Isolation and characterisation of mesophillic, oxalate-degrading *Streptomyces* from plant rhizosphere and forest soils. *Naturwissenschaften* **91**: 498-502.

Sambrook, J., Fritsch, E., and Maniatis, T. (2001) Molecular Cloning: A laboratory Manual. *3rd Ed* **CSHL Press**.

Sanger, F., Air, G., Barrell, B., Brown, N., and Coulson, A. (1977) Nucelotide sequence of bacteriophage phi X174 DNA. *Nature* **265**: 687-695.

Schenk, S., and Laddaga, R. (1992) Improved Method for Electroporation of *Staphlococcus aureus*. *FEMS Microbiology Letters* **94**: 133-138.

Schmidt, T., DeLong, E., and Pace, N. (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing *Journal of Bacteriology* **173**: 4371-4378.

Schroeder, D., Oke, J., Hall, M., Malin, G., and Wilson, W. (2003) Virus Succession Observed during an *Emiliania huxleyi* Bloom. *Applied and Environmental Microbiology* **69**: 2484-2490.

Semedo, L., Gomes, R., Linhares, A., and Duarte, G. (2004) *Streptomyces drozdowiczii* sp. nov., a novel celluolytic streptomycete from soil in Brazil. *International Journal of Systematic and Evolutionary Microbiology* **54**: 1323-1328.

Sharma, N., Moldt, B., Dalsgaard, T., Jensen, T., and Mikkelsen, J. (2008) Regulated gene insertion by steroid-induced phiC31 integrase. *Nucleic Acids Research* **36**: 1-12.

Sharma, R., Mohmmed, A., and Babu, C. (2002) Diversity among rhizobiophages from rhizospheres of legumes inhabiting three ecogeographical regions of India. *Soil Biology and Biochemistry* **34**: 965-973.

Shirai, M., Nara, H., Sato, A., Aida, T., and Takahashi, H. (1991) Site-specific integration of the actinophage R4 genome into the chromosome of *Streptomyces parvulus* upon lysogenization. *Journal of Bacteriology* **173**: 4237-4239.

Shirling, E., and Gottieb, D. (1968) Cooperative description of type cultures of *Streptomyces* II. Species descriptions from the first study. *International Journal of Systematic Bacteriology* **89**: 69-189.

Silvestri, L., Turri, M., Hill, L., and Gilardi, E. (1962) A quantative approach to the systematics of actinomycetes based on overall similarity. *Proc. Symposium of the Society for General Microbiology* **12**: 333-360.

Sing, W., and Klaenhammer, T. (1993) A Strategy for Rotation of Different Bacteriophage Defenses in a Lactococcal Single-Strain Starter Culture System. *Applied and Environmental Microbiology* **59**: 365-372.

Sliwinski, M., and Goodman, R. (2004) Spatial heterogeneity of crenarcheal assemblages within mesophilic soil ecosystems as revealed by PCR-single-stranded conformation polymorphism profiling. *Applied and Environmental Microbiology* **70**: 1811-1820.

Smit, E., Leeflang, P., Gommans, J., Saskia van Mil, B., and Wernars, K. (2001) Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Applied and Environmental Microbiology* **67**: 2284-2291.

Smith, M., Burns, R., Wilson, S., and Gregory, M. (1999) The complete genome sequence of the *Streptomyces* temperate phage φC31: evolutionary relationships to other virsues. *Nucleic Acids Research* **27**: 2145-2155.

Smith, M., Brady, R., Soultanas, P., Thorpe, H., and Smith, M. (2004) Synapsis and DNA cleavage in phiC31 integrase-mediated site-specific recombination. *Nucleic Acids Research* **32**: 2607-2617.

Sokal, R., and Michener, C. (1958) A statistical method for evaluating systematic relationships. *University of Kansas Bulletin* **38**: 1409-1438.

Southern, E. (1975) Detection of specific sequences among DNA fragments seperated by gel elecrophoresis. *Journal of Molecular Biology* **98**: 503-517.

Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R., and Liesack, W. (1991) Designation of streptomycete 16S and 23S rRNA-based target regions for oligonucleotide probes. *Environmental Microbiology* **57**: 1468-1477.

Stackenbrandt, E., Frederiksen, G., Garrity, P., and Grimont, P. (2002) Report of the ad hoc committee for the revaluation of the species doing in bacteriology *International Journal of Systematic and Evolutionary Microbiology* **52**: 1043-1047.

Stahl, F., Kobayashi, I., and Stahl, M. (1982) Distance from cohesive end site *cos* determines the replication requirement for recombination in phage lambda. *Proc. Natl Acad Sci* **79**: 6318-6321.

Steinbacher, S., Baxa, U., Miller, S., and Wientraub, A. (1996) Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc. Natl Acad Sci* **93**: 10584-10588.

Stuttard, C. (1983) Localized hydroxylamine mutagenesis and cotransduction of theronine and lysine genes in *Streptomyces venezuelae Journal of Bacteriology* **155**: 1219-1223.

Sulakvelidze, S., Alavidza, Z., and Morris, J. (2001) Bacteriophage therapy. *Antimicrobial Agents and Chemotherapy* **45**: 649-659.

Sullivan, D., Coffey, A., Fitzgerald, G., Hill, C., and Ross, R. (1998) Design of a phage-insenstive Lactococcal dairy starter via sequqential transfer of naturally occuring conjugative plasmids. *Applied and Environmental Microbiology* **64**: 4618-4622.

Sullivan, M., Coleman, M., Quinlivan, V., Rosenkrantz, J., DeFreancesco, A., Tan, G., and Fu, R. (2008) Portal protein diversity and phage ecology. *Environmental Microbiology* **10**: 2810-2823.

Sumby, P., and Smith, M. (2002) Genetics of the phage growth limitation (pgl) system of *Streptomyces coelicolor* A3(2). *Molecular Microbiology* **44**: 489-500.

Sumby, P., and Smith, M. (2003) Phase variation in the phage growth limitation system of *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **185**: 4558-4563.

Summer, E., Gonzalez, C., Bomer, M., and Carlile, T. (2006) Divergence and mosaicism among virulent soil phages of the Burkholderia cepacia complex. *Journal of Bacteriology* **188**: 255-268.

Sun, H., Deng, S., and Raun, W. (2004) Microbial community structure and diversity in a Century-Old Manure-Treated Agroecosystem. *Applied and Environmental Microbiology* **70**: 5868-5874.

Sun, W., Huang, Y., Zhang, Y., and Liu, Z. (2007) *Streptomyces emeiensis* sp. nov., a novel streptomycete from soil in China *International Journal of Systematic and Evolutionary Microbiology* **57**: 1635-1639.

Susskind, M., and Botstein, D. (1978) Molecular genetics of bacteriophage P22. *Microbiological Reviews* **42**: 385-413.

Suttle, C. (1994) The significance of viruses to mortality in aquatic microbial communities *Microbial Ecology* **28**: 237-243.

Suttle, C. (2005) Viruses in the sea. *Nature* **437**: 356-361.

Sylvia, D., Fuhrman, J., Hartel, P., and Zuberer, D. (1999) Principles and applications of soil microbiology *Prentice Hall, Upper Saddle River, USA*.

Tamura, T., Ishida, Y., Otoguro, M., Hatano, K., Labeda, D., Price, N., and Suzuki, K. (2008) Reclassification of *Streptomyces caeruleus* as a synonym of *Actinoalloteichus cyanogriseus* and reclassification of *Streptomyces spheroides* and *streptomyces laceyi* as later synonyms of *streptomyces niveus*. *International Journal of Evolutionary Microbiology* **58**: 2812-2814.

Taylor, R., Walker, D., and McInnes, R. (1993) *E.coli* host strains significantly affect the quality of small scale plasmid DNA preperations used for sequencing. *Nucelic Acids Research* **21**: 1677-1678.

Thingstad, T. (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems *Limnology and Oceanography* **45**: 1320-1328.

Thingstad, T.L., R (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand *Aquatic Microbial Ecology* **13**: 19-27.

Thorpe, H., Wilson, S., and Smith, M. (2000) Control of directionality in the site-specific recombination system of the *Streptomyces* phage phiC31. *Molecular Microbiology* **38**: 232-241.

Tilly, K. (1991) Independence of bacteriophage N15 lytic and linear plasmid replication from heat shock proteins DnaJ, DnaK, and GrpE. *Journal of Bacteriology* **173**: 6639-6642.

Tolba, O., Loughrey, A., Goldsmith, C., Millar, B., Rooney, P., and Moore, J. (2008) Survival of epidemic strains of healthcare (HA-MRSA) and community-associated (CA-MRSA) methicillin resistant *Staphylococcus aureus* (MRSA) iin river, sea and swimmin pool water. *International Journal of Hygine and Environmental Health* **211**: 398-402.

Torsvik, V., and Overas, L. (2002) Microbial diversity and functions in soil: from genes to ecosystems. *Current Opinion in Microbiology* **5**: 240-245.

Turpin, P., Dhir, V., Maycroft, K., Rowlands, C., and Wellingtion, E. (1992) The effect of *Streptomyces* species on the survival of *Salmonella* in soil. *FEMS Microbiology Letters* **101**: 271-280.

Twort, F. (1915) An investigation on the nature of ultramicroscope viruses. *Lancet* **11**.

Tyson, G., and Banfield, J. (2008) Rapidlly evolving CRISPRs implicated in aquired resistance of microorganisms to viruses. *Environmental Microbiology* **10**: 200-207.

Van der Heijden, G., Bardgett, M., and Van Straalen, R. (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11**: 296-310.

Van Regenmortel, M., Bishop, D., Fauquet, C., Mayo, M., Maniloff, J., and Calisher, C. (1997) Guidelines to the demarction of virus species. *Archives of Virology* **142**: 1289-1504.

Vriesema, A., Zaat, S., and Dankert, J. (1996) A simple procedure for isolation of cloning vectors and endogenous plasmids from viridans group of *Streptococci* and *Staphylococcus aurues*. *Applied and Environmental Microbiology* **62**: 3527-3529.

Wagenvoort, J., and Penders, R. (1997) Long-term-in-vitro survival of an epidemic MRSA phage-group III-29 strain. *Journal of Hospital Infection* **35**: 322-325.

Wagenvoort, J., Sluijsman, W., and Penders, R. (2000) Better environmental survival of outbreak vs. sporadic MRSA isolates. *Journal of Hospital Infection* **45**: 231-234.

Waksman, S., and Henrici, A. (1943) The nomenclature and classification of the Actinomycetes. *Journal of Bacteriology* **46**: 337-341.

Waksman, S. (1961) Classificaton, identification and description of of genera and species *The Actinomycetes* **2**.

Walkinshaw, M., Taylor, P., Sturrock, S., and Atanasiu, C. (2002) Structure of Ocr from Bacteriophage T7, a Protein that Mimics B-Form DNA . . *Molcular Cell* **9**: 187-194.

Waterbury, J., and Valois, F. (1993) Resistance to cocurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Applied and Environmental Microbiology* **59**: 3393-3399.

Wawrik, B., Kutliev, D., Abdivasievna, U., and Kukor, J. (2007) Biogeography of actinomycete communities and type II polyketide synthase genes in soils collected in New Jersey and Central Asia. *Applied and Environmental Microbiology* **73**: 2982-2989.

Wayne, L., Brenner, D., and Colwell, R. (1987) International committee on systematic bacteriology Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Biology* **37**: 463-464.

Weinbauer, M. (2004) Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* **28**: 127-181.

Weinbauer, M., Christen, R., and Hofle, M. (2006) The response of *Vibrio*- and *Rhodobacter*-related populations of the NW Mediterranean sea to additions of dissolved organic matter, phages, or dilution. *Microbial Ecology* **51**: 336-344.

Weisburg, W., Giovannoni, S., and Woese, C. (1989) The *Deinococcus-Thermus* phylum and the effect of rRNA composition on phylogenetic tree construction. *Systematic and applied Microbiology* **11**: 128-134.

Weitz, J., and Dushoff, J. (2008) Alternative stable states in host-phage dynamics. *Theoretical Ecology* **1**: 13-19.

Wichels, A., Gerdts, G., and Schutt, C. (2002) *Pseudoalteromonas* spp. phages, a significant group of marine bacteriophages in the North Sea *Aquatic Microbial Ecology* **27**: 233-239.

Wiebe, W., and Liston, J. (1968) Isolatin and Characterisation of marine bacteriophage. *Marine Biology* **1**: 244-249.

Wihelm, S., and Suttle, C. (1999) Viruses and nutrient cycles in the sea. *BioScience* **49**: 781-788.

Wililiamson, K.R., M Wommack, KE (2005) Abundance and diversity of viruses in six delaware soils. *Applied and Environmental Microbiology* **71**: 3119-3125.

Williams, S., and Khan, M. (1974) Antibiotics - a soil microbiologist's viewpoint *Postepy Hig Med Dosw* **28**: 395-407.

Williams, S., Goodfellow, M., Alderson, G.M., Wellington, E., and Sneath, P. (1983a) Numerical classification of *Streptomyces* and related genera. *Journal of General Microbiology* **129**: 1743-1813.

Williams, S., Goodfellow, M., Wellington, E., and Vickers, J. (1983b) A probability matrix for identificationof some streptomycetes. *Journal of General Microbiology* **129**: 1815-1830.

Williams, S., and Vickers, J. (1986) The ecology of antibiotic production. *Microbial Ecology* **12**: 43-52.

Williams, S., Mortimer, A., and Manchester, L. (1987) Ecology of soil bacteriophges. *Phage Ecology* **In SH Goyal, CP Gerba G Britton (ed)**: 157-179.

Winogradsky, H. (1953) Preliminary study of the sulfur bacteria of the Enghien-les Bains region. *Anns Inst Pasteur* **84**: 577-587.

Wipat, A., Wellington, E., and Saunders, V. (1994) Monoclonal antibodies for *Streptomyces lividans* and their use for immunological capture of spores from soil. *Microbiology* **140**: 2067-2076.

Woese, C., Stanckenbrandt, E., Macke, T., and Fox, G. (1985) A definition of the eubacterial "divsions" *Systematic and applied Microbiology* **6**: 143-151.

Woese, C. (1987) Bacterial evolution. *Microbiology Reviews* **51**: 221-271.

Wommack, K., Hill, R., Kessel, M., Russek-Cohen, E., and Colwell, R. (1992) Distribution of viruses in Chesapeake Bay. *Applied and Environmental Microbiology* **58**: 2965-2970.

Wommack, K., Ravel, J., Hill, R., and Colwell, R. (1999) Hybridisation analysis of Cheasapeake Bay virioplankton. *Applied and Environmental Microbiology* **65**: 241-250.

Wyatt, G., and Cohen, S. (1953) The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochemical Journal* **55**: 774-782.

Yabuuchi, E. (2000) *Burkholderia uboniae* sp. nov. L-arabinose-assimulating but different from *Burkholderia thailandensis* and *Burkholderia veitnamiensis*. *Micobiology and Immunology* **44**: 307-317.

Yanisci-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **1**: 81-83.

Zhang, Z., Greene, B., Thuman-Commike, P., and Jakana, J. (2000) Visualization of the maturation transition in bacteriophage P22 by electron crymoicroscopy. *Journal of Molecular Biology* **297**: 615-626.

Zhou, X., Deng, Z., Hopwood, D., and Kieser, T. (1994) Characterization of phi HAU3, a broad-host-range temperate streptomyces phage, and development of phasmids. *Journal of Bacteriology* **176**: 2096-2099.

8. Appendix

Experiment or comparison	Sum of Squares	df	p value
ϕC31 and ϕELB8 without a host			
Between Groups	0.004	1	0.804
Within Groups	1.775	28	
Total	1.779	29	
ϕC31 and phage K without a host			
Between Groups	0.023	1	0.532
Within Groups	2.692	28	
Total	2.715	29	
phage K and ϕELB8 without a host			
Between Groups	0.054	1	0.272
Within Groups	2.002	28	
Total	2.056	29	
ϕC31 and ϕELB8 with a host			
Between Groups	0.009	1	0.688
Within Groups	2.460	28	
Total	2.468	29	
ϕC31 and phage K with a host			
Between Groups	0.089	1	0.136
Within Groups	1.774	28	
Total	1.863	29	
phage K and ϕELB8 without a host			
Between Groups	0.021	1	0.131
Within Groups	0.412	28	
Total	0.433	29	
***S.avermitilis* and SP03** without phages			
Between Groups	0.285	1	0.000
Within Groups	0.483	46	
Total	0.767	47	
***S.avermitilis* and SP04** without phages			
Between Groups	0.124	1	0.000
Within Groups	1.100	46	
Total	0.224	47	

Experiment or comparison	Sum of Squares	df	p value
SP03 and SP04 without phages			
Between Groups	0.033	1	0.057
Within Groups	0.394	46	
Total	0.427	47	
***S.avermitilis* and *S.aureus* without phages**			
Between Groups	0.145	1	0.001
Within Groups	0.507	46	
Total	0.652	47	
SP03 and *S.aureus* without phages			
Between Groups	0.012	1	0.172
Within Groups	0.294	46	
Total	0.306	47	
SP04 and *S.aureus* without phages			
Between Groups	0.538	1	0.000
Within Groups	0.418	46	
Total	0.957	47	
***S.avermitilis* and *S.aureus* with lytic (ϕC31 or K) phages**			
Between Groups	0.005	1	0.595
Within Groups	0.768	46	
Total	0.773	47	
***S.avermitilis* and *S.avermitilis* with and without phages (ϕC31)**			
Between Groups	0.019	1	0.254
Within Groups	0.661	46	
Total	0.680	47	
***S.avermitilis* and *S.avermitilis* with and without phages (ϕELB8)**			
Between Groups	0.402	1	0.000
Within Groups	0.217	46	
Total	0.619	47	

Experiment or comparison	Sum of Squares	df	p value
S.avermitilis **and** ***S.avermitilis*** **with phages (ϕELB8 or ϕC31)**			
Between Groups	0.402	1	0.000
Within Groups	0.217	46	
Total	0.619	47	
S.avermitilis **and SP04 with phages (ϕC31)**			
Between Groups	0.002	1	0.784
Within Groups	1.049	46	
Total	1.051	47	
S.avermitilis **and SP04 with phages (ϕELB8)**			
Between Groups	0.206	1	0.000
Within Groups	0.158	46	
Total	0.364	47	
S.avermitilis **and** ***S.aureus*** **with phages (ϕELB8 or phage K)**			
Between Groups	0.319	1	0.000
Within Groups	0.324	46	
Total	0.643	47	
SPO4 and SPO4 with and without phages (ϕC31)			
Between Groups	0.019	1	0.784
Within Groups	0.661	46	
Total	0.680	47	
SPO4 and SPO4 with and without phages (ϕELB8)			
Between Groups	0.030	1	0.137
Within Groups	0.598	46	
Total	0.628	47	
SPO4 and SPO4 with phages (ϕELB8 or ϕC31)			
Between Groups	0.000	1	1.000
Within Groups	0.575	46	
Total	0.575	47	
SPO4 and ***S.aureus*** **with phages (ϕELB8 or phage K)**			
Between Groups	0.012	1	0.172
Within Groups	0.294	46	
Total	0.306	47	

Experiment or comparison	Sum of Squares	df	p value
S.aureus **and** ***S.aureus*** with and without phages (phage K)			
Between Groups	0.023	1	0.000
Within Groups	0.643	46	
Total	0.846	47	
S.avermitilis **and** ***S.avermitilis*** with and without phages (ϕELB20)			
Between Groups	0.319	1	0.000
Within Groups	0.586	46	
Total	0.905	47	
S.avermitilis **and** ***S.avermitilis*** with and without phages (ϕR4)			
Between Groups	0.359	1	0.000
Within Groups	0.614	46	
Total	0.973	47	
S.avermitilis **and SP03** without phages			
Between Groups	0.285	1	0.000
Within Groups	0.483	46	
Total	0.767	47	
S.avermitilis **and** ***S.avermitilis*** with phages (ϕR4 or ϕELB20)			
Between Groups	0.001	1	0.811
Within Groups	0.955	46	
Total	0.957	47	
S.avermitilis **and SP03** with phages (ϕELB20)			
Between Groups	0.143	1	0.011
Within Groups	0.928	46	
Total	1.0271	47	
S.avermitilis (ϕELB20) and SP03 (ϕR4) with phages			
Between Groups	0.030	1	0.242
Within Groups	0.981	46	
Total	1.011	47	

Experiment or comparison	Sum of Squares	df	p value
***S.avermitilis* (ϕELB20) and SP03 (ϕR4) with phages**			
Between Groups	0.161	1	0.008
Within Groups	0.955	46	
Total	1.017	47	
***S.avermitilis* and SP03 with phages (ϕR4)**			
Between Groups	0.052	1	0.130
Within Groups	1.009	46	
Total	1.061	47	
***S.avermitilis* and SP03 (ϕELB20) with phages**			
Between Groups	0.143	1	0.011
Within Groups	0.928	46	
Total	1.071	47	

Figure 38: ANOVA tables of soil microcosm experiments.

```
LOCUS       nuc_1                   51160 bp    DNA     linear   26-JAN-2009
DEFINITION  Streptomyces bacteriophage phiELB20.
ACCESSION
VERSION
KEYWORDS    .
SOURCE      Streptomyces bacteriophage phiELB20 (phiELB20)
  ORGANISM  Streptomyces bacteriophage phiELB20
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
            PhiC31-like viruses.
REFERENCE   1  (bases 1 to 51160)
  AUTHORS   Bell,E., Hendrix,R. and Herron,P.
  TITLE     Complete genome Streptomyces phage phiELB20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 51160)
  AUTHORS   Bell,E. and Hendrix,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2009) SIPBS, University of Strathclyde, 204
            George St, Glasgow G1, United Kingdom
FEATURES             Location/Qualifiers
     source          1..51160
                     /organism="Streptomyces bacteriophage phiELB20"
                     /mol_type="genomic DNA"
     gene            complement(38..220)
                     /gene="gp1"
     CDS             complement(38..220)
                     /gene="gp1"
                     /codon_start=1
                     /translation="MTSDDDLTFALQVAGAELADTPPSPDSPLGRLRLFAAANPGVEL
                     GPEHIRGAIAGTLPRLP"
     gene            804..974
                     /gene="gp2"
     CDS             804..974
                     /gene="gp2"
                     /note="Structural protein"
                     /codon_start=1
                     /product="gp1"
                     /translation="DHIVAGDDHRLSNLRSLCAWHHKQKSSREGAAAQAAKRRAIQKK
                     FQRGEQHPGLL"
     tRNA            1083..1155
                     /product="tRNA-Phe"
                     /anticodon=(pos:complement(1083..1155),aa:Phe)
     gene            1370..1630
                     /gene="gp3"
     CDS             1370..1630
                     /gene="gp3"
                     /note="Structural protein"
                     /codon_start=1
                     /product="gp3"
                     /translation="MCRMSARRPSVTGFEWAWLAWGGAFVVIEGIALKRKDKNDTLSE
                     QVWKIFHTSQGQEKTKTTQARRAVLVMFLAWLVAHFVSGGKV"
     gene            1845..2180
                     /gene="gp4"
     CDS             1845..2180
                     /gene="gp4"
                     /codon_start=1
                     /product="gp4"
                     /translation="MKAVKVPAADRDWHPIARRLWDSLKTSGQADFYQNSDWAFAYSL
                     CEDLSYYKKSGKRSGQMLQTIYSAFERLLVAEGDRRRVRIELTEPVPEQDSAAVVAIA
                     DYKKVLGVD"
     gene            2194..3855
                     /gene="gp5 - Terminase"
     CDS             2194..3855
                     /gene="gp5 - Terminase"
                     /codon_start=1
                     /translation="MAATAPLTLEEIEALEPEYLGPTWQKDAFGRWKLPEFTLGWQIA
                     GWCAEYLDGEGSTDEKRVPWKFTTEQLRFLLWWYAIDKDGEFIYRTGVLQRLKGWGKD
                     PLLAVISLVEFVGPSRFSHWDEDGNPVAKAHPQAWVQITAVNQEQTTNTMGYLPILMG
                     PKLIQTYGIKAGAELIRADRGRKKLQAVTSSYRAIEGKRTTFTLLNETHHWVAGNGGH
                     KMFETIDGNSTKMDSRYLAITNAYLPGEDSVAERTREAHNKVLEGRAIDVGTLYDSIE
                     AHPKTPLTPEALRIVLPKIRGDAVWLKIETIIKSVMNLTISASRSRRMYLNQVVAEED
                     ALYGPAEWDPLADEGKSLKPGDEIVLGFDGGKTSDSTALVALRVRDMHAVLLNIWEHP
```

```
                     EGEAAKDWSVPRHEVDSAVHEAFRLFEVRAFFADVALWESYIADWSEAYGDQLGVRSP
                     SGKDAIGWDMRSSQKTVTLAHERLMRSVFDGKLSHDGDLTLRRHVLSARRRTNNYGLS
                    FGKESKDSPRKIDAYAALMLAHEALNELRARGKKVRKRTGRGYFM"
gene                3885..5357
                     /gene="gp6 - Portal protein"
CDS                 3885..5357
                     /gene="gp6 - Portal protein"
                     /codon_start=1
                     /translation="MADTSPASLAKELLAILDRDEDRIQRIDDYIHGKHDDPYMPPHA
                     DDEYRLLAKRAVSNAMPLLIGTPAQALYVDGFRPGTTDSDLPRASSSTSPAMNHWQRS
                     RMDARQAAIYRGALGFGHSFVLTEKTKKGVISKGLSAKRTAALYEDPANDETPYAAMT
                     VTAMPRGENLGKARLFDGKREYAVTFKSRTDPKSVKVAAGKLHGASECPVTRFAAQVD
                     LEGRTIGVVEPMIPLQDRINQTIFDLLVAQTYTSHEVRYATGMAPPMQMEMLDENGDV
                     TTDPNLAVDSRPRLDDEGNPVPATINHNARRFLFAEDPDVKFGSLPAGPITPLIESVD
                     MSIRHLAAVSQTPPHHLLGQIANLSAEALLAAETALARMVAEFRSIFGESWERVFRLA
                     AEMEGDETATDDFSGEVQNRDMESRSLAQAADALGKLADMLGIPKRGLWKRVPGVTQT
                    EYEDWEQMAEEDDSVGQLANALTRATPEAAPASDPGVIAA"
gene                5354..6463
                     /gene="gp7"
CDS                 5354..6463
                     /gene="gp7"
                     /codon_start=1
                     /translation="-
                        TSPARQAEADRAAIAFQTALTQIGAGTIAEALTMWGDVPATARASTSASWLRRAITMV
                        MSRRRMSRDLARAYYRLARALHTGTTVADPYHPEPTYITLDVLRREFAELAGEPAEQP
                        QEGRSEPSDTSTPEAPSSAASSASVEADDDDSPSADQESEDELDRILVEEIEGLREAE
                        ERIEREAEQELRTVLEALGPDWLQGKVDKIDGGRTADEVDKLRDEAHQQAGGRQAAAA
                        ERVAMNGGRSTVMNHASRDRRAIGYIRLSRTGTPCGWCAMLISRGPVYRSSASAEYAD
                        GDKYHDNCHCYAETVFSREQYNSSSLYELNRRYEELWPKVTSGLSGKAAVAAWRRFIR
                     QEQKAAAQEARRQTPNVQEA"
gene                6466..6990
                     /gene="gp8"
CDS                 6466..6990
                     /gene="gp8"
                     /codon_start=1
                     /translation="-
                        PEQETPSTEETTTEETVETPPEGTAPEGGETETSTEEKPSEESVPPEVLRKKLTDANA
                        EAANYRTKLRETEAKLSSAKTVEEFESATAELKGQIAELERSILLNTVAAKYELPASL
                        AKRLTGATPEELEADAKELQKLVAPPQPESLSGGLDPEDDEDFDPVKAAQAARRNRY"
gene                7068..8135
                     /gene="gp9 - Capsid protein"
CDS                 7068..8135
                     /gene="gp9 - Capsid protein"
                     /codon_start=1
                     /translation="-
                        AHTPVKPEKIAATAAVALEQSLIVPALFQREGIDAYKGAKNDTINVKVEGVLPFRIYE
                        WRSGEPGSSTPGVRQAIEFDEYTERTVAVKFGGNIYSAVKLTDEQNEFDLPGWAKLMA
                        KQTEAIGRGLEHEAVQHLLNQEYAVTLGGDVAARGGQGDLRKTLIRAREVLNRFMVPK
                        EGRVLVVGTGWETALLSDEKLNFAGNVGEQEAVSALREASIGRRFGFDIVVSQEVPSD
                        AAFAMHRSAFIFATGAPSVPQSIKAGGTSSYEGVALRWLQDYDADHLTDRSVVNTYKG
                        FRSVEDILVGTDPEGQSFVSANEHFVRAIRLDLDATVDVLPDADGPDAKQKELFDITG
                     VGAPIV"
gene                8214..8711
                     /gene="gp10"
CDS                 8214..8711
                     /gene="gp10"
                     /codon_start=1
                        /translation="MADFATLEELKARLDWTLDADEERIATASLEDASDLARAYAGRD
                        WDPAATAPRLVRTMVLKACKRYMNNPSGYTQSRAGDETLGWSDDAGENAGTVYFTKDE
                        QEMLANLGGRRSGLVSVGVTAWNSDIRRYRNRQSWRSGNDLPAGFVPADGKDVPFFGD
                     EVEPW"
gene                8708..9049
                     /gene="gp11"
CDS                 8708..9049
                     /gene="gp11"
                     /codon_start=1
                     /translation="-
                        SSMQRRRGVTARIWKTSKQIDNRGNEVLVADADGPYEVRCALIPQRSARAEVPGQQQI
                        NITRMIVSADLENVSLWSRVEVHGAVWDVVTPPAYHHGERKTRHWSIDIRERPS"
gene                9049..9516
                     /gene="gp12"
```

```
CDS             9049..9516
                /gene="gp12"
                /codon_start=1
                   /translation="MAYIYKGLNGKKMSEVIASLDEVQGEVWERTFEIAARAEALLVQ
                   HRAEGVAQIDWAKGDIDAYVVLEDKNRTNQKHDPKKPYSNSAASIEFGRSAYDVEVVD
                   SSGQVVNEYEVGEMEGLFILTEASHLPKKRRAVRTPKTVRIKQKKRKKSGGDG"
gene            9516..10013
                /gene="gp13"
CDS             9516..10013
                /gene="gp13"
                /codon_start=1
                   /translation="MAGLPPEIKALAELSPVEDLMLAVLRQGLPGIQVRSLIAKDQKF
                   PLVLARRDPSFGNWGGDTRFLDAARVAVHTFCADPDGDEDAAILSEAVRVVMRNAWLN
                   QTVYAGLGHITRVDLISAPRRVTDWATSTGPVQYADLPTGVWRYEATYDIEIRKPRTR
PYPTP"
gene            10026..10766
                /gene="gp14 - Tail protein"
CDS             10026..10766
                /gene="gp14 - Tail protein"
                /codon_start=1
                /translation="-
                   ALNDNATLVIGSGNYLTAPVGTDIPTDLLVPVSPWEAVGHTSLEDIFSITSEGGEATT
                   IGSLQNKSLRTKYSARTETMAFTLQQFDVKGLKLYYGSNAPILPDGSVGVPTDPTPTT
                   AAFLAVFVDGENVFAFYAPKAEIYRGDDVSFGDTESLAGLPLGVKPMAHGANSWTYAI
                   TPLGTVAATGATAGTPGTFTPTGSLAPYDLVDLDLVTADPATAWTTGQYVTLGDGTNA
                YWDGADWVTGIAA"
gene            10890..11231
                /gene="gp15"
CDS             10890..11231
                /gene="gp15"
                /codon_start=1
                   /translation="MASYSLDDIRAAADAKYSSTDITVDEKTTARLLNPLRLSKEKRE
                   ELQGIQDRMDGEGVDQEALLAEAIMIVVDHPAKGKALLKAVAGDLALLAQIFNTYIEG
                AQVGEASASAA"
gene            11291..11623
                /gene="gp16"
CDS             11291..11623
                /gene="gp16"
                /codon_start=1
                /translation="-
                   DVIEGRGPSPLFVLALVRRLPDTSLTVALASGGREHFGWGIDRHLSADLYDAINQNTR
                   STGQWAKGKAPKIPAWPRPKALKKPTKSSGPKTAGKKVSVAEIYKRFTARR"
gene            11627..15028
                /gene="gp17 - Tape Measure"
CDS             11627..15028
                /gene="gp17 - Tape Measure"
                /codon_start=1
                   /translation="MPQGQVIGRVSVRVLPDTDEFRSKASKQLTKIEKQLKVEVQVMP
                   NMAGFERQMLTEISKISQRNRQSDARKVKIYTRIDTSTMNGELARAIRAYTNRAKSGQ
                   KVKLQSTLDAGAVDVEISDQSLRKMTDQLKDWRDDNSPLKIKIEPDISALSSAATSAR
                   LGILTRPRTVSIIPQLNNAAVAKVATALAALSGIRVLNNLFEKFGNILKNLDKSVPII
                   GSLASAMAGLAAMALTSASNLFALSASLAQIGALALTAPGLLGGFAVGIGITVAAFKD
                   FNKEVPQAKAALADMQNTISSNFWAEARKPISDLVDNLLPRFEKGFASSSTAIGKFFG
                   SFATDLTAALDPSLVDKMFGYLNESITTATGGTKVFSSIIAQLGEVGASYLPNLAGWF
                   VDISKQFDVWLQKKGQIGLQEEIDAGITSLKDLGGVLYETGGILAGLSRAATEAGGST
                   LGMLRDSLASLHDVVDSDGFQKGMVDVFTAAHTAMSNLANGAGPAVKNLFIELGSLLT
                   TILPQAGSILGVAMGAIADALSQTAVTDGIKSMFTALFQAVQLLAPAMAPLGQALGAL
                   MQVVAAMLPVFAGLVTAAIIPLANAFTTLAPMLIPIVELLGGALTSAFQMLTPVINQL
                   VPVVGEMLGAAFGLLAAILPPIAQLFQMIVAAVAPLVEQLIGALAPILPILGEALAQI
                   FTALQPIVQIALQILSAVLGPLIEILSTVIAAVLPPLSDAITRVVEALQPFLQALLAV
                   VDFIMPILVPIIQFLVEILAGALVAAINGVGLVLEGLKELFVGIWDYIVGYFKMIWGI
                   FEGLWSNDWSTFQEGFSQLWDGIWGIIKGIWDIILGGLEIFFNVGILGTAAKAFKAIG
                   ALFKSAWKAIVEIFTGAFAAIRGYIGVFLTGAKGLFMDGLKAIGKFFSDGWASITRGV
                   TSFFSGAVSKVTSGLASIKGFFSGAWTSIKSTAVTGFSRLVSTIAEWLGKAVAKVKEL
                   PGKAKAALGSLGSTLKDAGMKLIRGFIDGIASMFGSVKDKLGSLTDKLTSWKGPESLD
                   RVLLVDAGKLVIDGFINGLESRYDAVRKTLRGLTEDVAGTEFDALGVGQISASGSLSA
                  AVNGALAGSEGGGTTKVLNYYAAPGSSLGSEEDLFAAGNRARFGW"
gene            15035..15892
                /gene="gp18 - Tail protein"
CDS             15035..15892
                /gene="gp18 - Tail protein"
```

```
                    /codon_start=1
                    /translation="
                       PKLLLASEKDTLNLNEIMDKGLGYQAKTGVTGFGLPPVSVQWLEGAGDGAVFRRRRVL
                       TRDIDIPLEILARDRRHLQELTDRLALVLAGRCTLTLLDDDGTRWTTEVYRVGGGEYA
                       YGSDTTGEREFQTVITFRSPDPYWTSSEAQIRYIGGDTSTGSFLSSLASVPVAASQAI
                       GEIQLDNRGTAEAYPIWEITGPGTNFLAVSPAGEKLAWSGTLGANQRLIIDTRKGTVV
                       DQDGANRYAELEEAPRFWTVKPGLSTATAQLEEIGTASKITCSWRPRKWMVI"
gene                15893..17104
                    /gene="gp19 - Tail protease?"
CDS                 15893..17104
                    /gene="gp19 - Tail protease?"
                    /codon_start=1
                    /translation="
                       RLEDITVEVRDKSLTRRGLVRPEELRLELTDNFNNVGEWKLTLAAENPLAGVLRTPGA
                       GVIITGPTDILMSGPMVKAEFAATQEDPGGSITFEGVSDTVCLADALAFPDPTNPNGA
                       SQTKAHDVRQGPAEDVMHSFVMANIGPTAPIERRKPHLAASTSSGRGPAVIKSARFPV
                       LGNLLAELALLGKLGFRVVQRGSSLVFETYEITDRTALIRLDVENGSLAGQRVAISPP
                       GVTRAVVAGQGELVKRQFLQVQTTESVAAEADWGRRIEQFIDQRQTDKWEELQQAGDE
                       AMETSGFTAVNVQVVPMEDGAMRFGLDWGLGDKVSVVVEGQELTSNATGMVMKIDSDG
                       FRYGVLLGDPTGFNSGAAMAKRVTNTESRVSQLERMTDAGAAALNQLMSIMGVM"
gene                17109..17411
                    /gene="gp20"
CDS                 17109..17411
                    /gene="gp20"
                    /codon_start=1
                    /translation="
                       ATAPKNFWRGELPDLDTIIYTAPANGQAIITDIVATNVSASSALIAIKINGVPLLANV
                       GIHPNGVFTLRISQVLEPNDSISIQGNAATAYAHISGVEVA"
gene                17412..19601
                    /gene="gp21"
CDS                 17412..19601
                    /gene="gp21"
                    /codon_start=1
                    /translation="
                       PTTWYPYDNGTVGPTGPEGPQGPPGSAGTPGVVQSVNGQSVQQVVLDAADVNALPSNA
                       NAQLDATYFGINKAAGNYRVLRWLTDGASRWEAQVDDVAEGGSAAGSDFRLASRNDDG
                       SFNKTVIHAKRSDGTISFGTTVHHGAAQVTSYGAVGLRDIGADPATASGGTYVYSKSG
                       KLFVKQADGTSFQIAQVSYPVTSVNSKTGNVTLGAADVGAISSVSGGTAGGTVRVTGT
                       AGTYRTLGLSTGPTGTASDRWYFQADNAAEAGGAAGSDLVITGRSDTGTFQSHAMHAK
                       RSTGQTSFGTTRPLGEAQTTVAGAIAARDRTTDPATASGGIQLYSKAGLPYIKQGDGT
                       VFKVGSGGGTGGGAVDSVNGKTGIVTLAASDVNALPISGGDLTGAVNINPTSGNTLTA
                       YGSTDPATYFRVTADGHPYSNSLRATFYNLGVGDTGAPFGGGKFMLGIKDASVLPTSN
                       PTNGVIAYSEAGKLKVRQTDGTIVTVANAPVSSVNGLTGAVNLTIEDLEGVSVSNVGQ
                       ANGLATLDSGKRVPLAQIPLSATPGAFQPEDLGLKAWAGDPATCRSGYRDGSTGKGRM
                       AAVVVRQTTTVSKIVWHFLGYAGGMKTGSWAGIYNTSGALVRGTGDLSTAAYEPAEQH
                       GVGGGCSTSNLTSSVTLAPGTYYIAWRFNYTASPADGPLLMAFESTDVSATLYGLNAV
                       RRFGAFATSATTAPSSVSGWETDPIRFWVALA"
gene                19637..20107
                    /gene="gp22"
CDS                 19637..20107
                    /gene="gp22"
                    /codon_start=1
                       /translation="MPRGVVALTTGLDDTPYISESETMVYQLPFTAAPKRIYKVTLRV
                       GRADTNGTGDQAGDVNRYGKNSLVTRCRWASGSTVTTSGAGVGDYRVTVFDDDSTSAT
                       GVCATWFIVNPPAGQTAVGIAIYAGRTAATYGQVRYIADGGAVLVVEDVGPYSE"
gene                20139..21416
                     /gene="gp23 - Tail related protein"
CDS                 20139..21416
                     /gene="gp23 - Tail related protein"
                    /codon_start=1
                    /translation="
                       SLSSYPFDGQAVSEGQYSYLFRELASHSGIADSVNGSSFNTYGDSSGMQVKVNPGFAV
                       VRGHAVQSTATETITIAAASAAQRYDRVVLRLDPTANSITLAVVQGTSGGGLPALTQT
                       DTGIFEFPLATVLIPVGASTISAGNVTGERRFLGNTVGGWTTATRPESARIGRLGLNT
                       STNQFEFWNGTQWRDLAPTITWSTIEGAPATFPPAAHSHAWNDITSKPTSFAPSAHAH
                       EWNDITGKPSTFAPSSHSHSWSSITSKPSTFAPSSHSHSGYLTSGSTISWANGSKKPH
                       SNSASGSGTWYSVWVEGDGTFCRNTSSIKFKENVRDYAVNEDAVLSLRPVVYDRKDQR
                       NDDGSVKEGRKDEVGFIAEEVEAAGLDWMVNYLDGEVDGLRYDLLGVALVPVVQRQDK
                    QIKALEARLEALEAHLAG"
gene                21476..21796
                    /gene="gp24"
```

```
CDS             21476..21796
                /gene="gp24"
                /codon_start=1
                /translation="
                   TLIGTIVELLRRQNNAINEVRENAQEARDQVANTHSTNLRDDLDRVIDGLEAVLAGQA
                  RHDEALRQHGREIGGLRTDLAHERAERLAVAERLDGHIAVTVATNNS"
gene            21866..22936
                /gene="gp25 - Endolysin"
CDS             21866..22936
                /gene="gp25 - Endolysin"
                /codon_start=1
                   /translation="MSVSIISRSAWGARAWNGTPNYVALSQRTEFFIHYDGANHITAT
                   GNSVPQAIDRQHQAQGNAGIGYNFVVDQAGNIFEGRGWTGQGAHCPSHNVSGLGVQIA
                   IGGDQEPSAKALSAARALYDEACRKTGRTLAKKGHKDGFATACPGGKLYAWVQAGMPA
                   DGYAPSTGGSSSGGTVARYQVTINGLKYGYGAQGSHVTTVGKALVAKGFGKHYAEGPG
                   PTWSDADTLNYADFQRSLGYSGSDADGVPGEGSLKTLLGSLPGASAPAPAAKPAAKKY
                   EPFPGASFFKRAPKSAIVTAMGKRLVAEGCGVYSSGPGPQWTESDRKSYAKWQRKLGY
                TGSAADGYPGKASWDKLHVPEV"
gene            22980..23153
                /gene="gp26 - Lysis protein"
CDS             22980..23153
                /gene="gp26 - Lysis protein"
                /codon_start=1
                /translation="
                   ITQYLPTKYKSKAGLVASLVGVVLSVAVLFTTDYPQIAVAIQALTAFGFVENSDTE"
gene            23210..23518
                /gene="gp27"
CDS             complement(23210..23518)
                /codon_start=1
                /translation="
                   KLQVTACDIDKKFPAKTYTITASDGRSISKDLCEEHAEPFEEWLEEAEVEEEPEPECA
                  PEAKPEPEPAPAPKTVKRAPAKKTAARRRPKIVSLEEIEAQKN"
gene            23634..24092
                /gene="gp28 - Repressor"
CDS             23634..24092
                /gene="gp28 - Repressor"
                /codon_start=1
                   /translation="MGKRKIVDETEVIRWFEEGRTYAWMIAEYKRKYGIDTVPSMWGN
                   FRRRRGLDRRIVRDDDLIPWFVKEEHRWAYPLAMLRAEARRRAGKELTETDTSRLASW
                   LDMLKDENAVVHYDPETDEGFFYIPRQPEDDDIIHRPDEKTTPRPNADRE"
gene            24348..24869
                /gene="gp29 - Lysis protein"
CDS             24348..24869
                /gene="gp29 - Lysis protein"
                /codon_start=1
                   /translation="MSHSVGDGPSEASGGWAGEYTSSDGLIKLVVDEEEMDWHIDARP
                   GHSPACMRTVLAQARARGWEPMDADECEPELLEDGTVRIYLVFAGQSAIVAQPAIYTL
                   QEKRRTSAAKRMSLAFVLAACVAGALLTPSPLHHDFPLIEHDHHQSDTADKVVMDPGT
                TNLITVSAEREEH"
gene            24870..25844
                /gene="gp30"
CDS             24870..25844
                /gene="gp30"
                /codon_start=1
                /translation="
                   ALNLIEIPQQSKPLHPNDAVPRDGHGKPLVIPEAGGKPKALVRTTTFIDCIEDKSNLV
                   DWGKRMVLVGAQKRPSLLDAVAELDPNVSGDKRKLDALAERAMDVSGANDKREKGTYL
                   HDLSEYVDRGEQLPSHVSEQDIEDMAAYLMETAPLTVHAVEQFVVCSELGVGGTYDRT
                   YGYEGPGPHGRPVSGRFIGDLKTGSVDYGGLKMAMQLAIYSRAKKYDHTLFPAPIRDE
                   DEKGFQKWKKVEVSPEEAAKAYTVPEPVNQDWGIIVHLPSGEGVCKLYWVDLNIGWKA
                  AQLALTIREMRSMSRKAMMPFVTQVTPAGIDFG"
gene            25968..26450
                /gene="gp31"
CDS             25968..26450
                /gene="gp31"
                /codon_start=1
                /translation="
                   KITIKYGKSYEDTWAVFEGTTTELRTEIIDYFGMDPETQRGLSLSSIVVNATQVAHGK
                   GLIATSLGATVIEETNHEEPAKPAGDPWAAASAKQSPSAWSGSASVAEPKASDPNAYI
                  LGEIEKKTTVAELKKLWAEMQSFFADPAVMTAYKAKGKALSAG"
gene            26533..26970
```

```
                     /gene="gp32 - DNA rep"
CDS                  26533..26970
                     /gene="gp32 - DNA rep"
                     /codon_start=1
                     /translation="
                        ALNLIDIPVQGGGWFKPADNINSVAILEVHSFDRQRPTPNGPKDSVLADVTIFQDAA
                        SLAAGTPQVTKGQRIEQTILARDLETIVGGATIVTLAQVPPKKPGAHPAWVNRPLTDA
                      AARKAVIDYAGKRDQAAEAAMDDAPDFD"
gene                 27045..27530
                     /gene="gp33 - Deaminase"
CDS                  27045..27530
                     /gene="gp33 - Deaminase"
                     /codon_start=1
                     /translation="
                        FKRPTRDEWALGIAEAVATIADCTRAQVGAIIVAKSGHSVLGLGYNGLPRGVPGCGTG
                        GNCPRGRLSTEDCARDSDYANCAADHAERNAIRDALDVKGIHPDELRDSTLYVTRKPC
                       PACTTLITAVGIGRVVVRGEENPECSPQDEAFRFMLNQAVNSRA"
gene                 27587..28342
                     /gene="gp34"
CDS                  27587..28342
                     /gene="gp34"
                     /codon_start=1
                        /translation="MIAGRSGTQKSGFALFWTAMMNLPTLYFSADMSAFTASSRLASM
                        FTRDTTEMVEAGMAEGGKYRQAYIDALADSKITFSFGSPISMRAVDEELEAYVELMDM
                        YPEVIVMDNLMDFEGAESDYTEQMAVMQGATELARHTGATVILLHHASDKAMEAKSSP
                        MNPPSRDQIKGGLSEKPELSLSVALDPTSMAYHVACVKQRMGPCDPTAQRYATMICEP
                      EYTRFKKAEIRQVAAPVVKPAEDMSPTKVLLGS"
gene                 28408..28785
                     /gene="gp35"
CDS                  28408..28785
                     /gene="gp35"
                     /codon_start=1
                        /translation="MSNSISTRNKRNKRIGAQMETDLREGLRSEGFDVESLRLAGKED
                        EGDMVVREGDGKYLVIEAKNAKFEPGVFLGEAIVERENFAKHRGLDLENVESIVVVKR
                     RGKGMRKAFVLTTVEDYLGLDPQ"
gene                 28782..28919
                     /gene="gp36"
CDS                  28782..28919
                     /gene="gp36"
                     /codon_start=1
                     /translation="
                       NATADGIAAFHAFISDAEADILTVFRVERAYEVDMKDVGEGWLP"
gene                 28907..29350
                     /gene="gp37 - DNA replication"
CDS                  28907..29350
                     /gene="gp37 - DNA replication"
                     /codon_start=1
                     /translation="
                        AAVRFHRIDDRSGEGSDSKPLLDATMHHFDVDFNDQRNAGMAKCPLHDDNTPSMSYRL
                        DEGLWNCHSCGNGGDSFTLIEKYHDMQLNKAIDFKQAKAYAKEHGLEEGAVAKETGYT
                      SRYGGGRKAAGKKPGQKPGGGYVPAWKRSK"
gene                 29365..30027
                     /gene="gp38"
CDS                  29365..30027
                     /gene="gp38"
                     /codon_start=1
                        /translation="MAEHEPLTPLSTSQKEMLEEAVATYQGHLTPEAAGYLMARGIGR
                        DEALAFRLGIVADPAPGHEKYRGMLAIPYLGRDGQPLTVRFRCLAEHNHRDYFHGKYN
                        TIKDDIPRMFNVGAVHRAGDEIHVTEGELDAVILNKLGLPAVAIPGANMWFGRHRRNL
                        AGFSRVWTWADPDDAGAELTGKVTRALRSAKAVRLKADVTDTYMTHGAEHVLSLVETK
                     ED"
gene                 30032..30394
                     /gene="gp39"
CDS                  30032..30394
                     /gene="gp39"
                     /codon_start=1
                     /translation="
                        ADIENEKTETVETPKAKPTPRKKAPARKPDPMEALLAEMREAVSHLGPMPTKDTAPSR
                        RRGHDDRAAAWGRHYGREHTVDGFLLSLTFEALACYKQERRYALLQLAGAALHLAENP
                     VDE"
gene                 30391..30603
```

```
                     /gene-"gp40"
CDS                  30391..30603
                     /gene-"gp40"
                     /codon_start-1
                     /translation-"-
                        SAENEDFPEENEGVEVTDTAPVIDHYGAVKRAASIVGDLRKELRAEGFTKEETFELVQ
                     VYWASEMGVFD"
gene                 30604..30792
                     /gene-"gp41"
CDS                  30604..30792
                     /gene-"gp41"
                     /codon_start-1
                     /translation-"-
                        AAELKHGRLAYVVVHGTHEVVLARYTQHDPEEPGYFAFFGTEFTAWPEDTEIVELAVT
                    RSV"
gene                 30794..31018
                     /gene-"gp42"
CDS                  30794..31018
                     /gene-"gp42"
                     /codon_start-1
                     /translation-"-
                        TELPGDPGPTLIDIFAAWTLAEQLAFVPHLLGETSADWLSATLRRFDHDVSATTIRTY
                     RRALRQEGGPSERAA"
gene                 31002..32195
                     /gene-"gp43"
CDS                  31002..32195
                     /gene-"gp43"
                     /codon_start-1
                     /translation-"-
                        SELLDELLAKPVGPTVPARQTDPEKDFTRQIEVKGDAADVTVRAETFEQTETAATDVL
                        KGQGLDPAEWTVTGFRSSEWTWANGDTGVSTRFSFARAKCATVADARPPIDELLATIE
                        QYTPVEPLDEPGDHTFIVALGDMQFGKIDGDGVEGTLQRTIDCLNKAADLLYVYKQRF
                        AIDHVHIAWLGDHVEGFVSQGGANTWRTQLTLNEQIRLTRRVMLHALLLFAPRVGYLT
                        MAAVPGNHGEAIRVSGKGVTRYDDSHDTESLIAVKDAADLNPERFGHVEFFVPDTDEL
                        TVVVECSGTVVSHAHGHQFRPGKHFDWWKGQAFGRASAMHQADVLLAGHLHHEFIEAD
                       GPRTFIQVPSMESESTWFRHSKGAEGAPGLIVAVTKDGRVPVKEVVTQ"
gene                 32192..32545
                      /gene-"gp44 - integrase/recombinase"
CDS                  32192..32545
                      /gene-"gp44 - integrase/recombinase"
                     /codon_start-1
                     /translation-"-
                        NIIEITNAYETAEQASVDWTVLELKPVTDGARSAAYAFARDYEGVVEREDMEQELLTA
                        FALRPRMVREVIAEADNPAGVLNFRGYRILRDIFKTQATHRRKQVSYEANAEALGGAE"
gene                 32542..32928
                     /gene-"gp45"
CDS                  32542..32928
                     /gene-"gp45"
                     /codon_start-1
                     /translation-"-
                        SGYNRALVEHILPAVWDGEAAYGIKAEQVPDADMPRGHKDPKKGSPLFAHIADIKAAW
                        KQAPLSLVEQQAVFLRFALDEHDDEIAAFQSVSGRAVRYRVERGVGKLAAHLNGEKYI
                     DGYDQLEDEAA"
gene                 32925..33302
                     /gene-"gp46"
CDS                  32925..33302
                     /gene-"gp46"
                     /codon_start-1
                     /translation-"-
                        SEAPPESMADGRFDFYDDARELYFWRDEGSDSDGLIFYRSYTEAERADKAKRLQLDGL
                        RAQADEAIPYLDERIDICLAYVERLEPASLEELAAQLKVVSDLAAYSAGTLKRLIVVL
                     GELTGRPV"
gene                 33380..35494
                      /gene-"gp47 - Reductase"
CDS                  33380..35494
                      /gene-"gp47 - Reductase"
                     /codon_start-1
                     /translation-"-
                        IDVPFGPTGELVYNRTYSRTLADGSKETWPDTVRRVAAGNLALVHGTDQTAWSDDVRA
                        EHDELVSYMDQFAIIPAGRHLWATGVKGRQYLFNCHVAPWGDKLSRHFEFTFMRLMEG
                        GGVGGNYSTKYLKEFGAPRRELDVHVVCDPMHQDYAEMKAAGLLSTEYDSDWAGAFEV
```

```
                    EDSREGWADALVDLIDTFMSDGEVKHRARVYDVSRVRCKGSRLKTFGGTASGPGPFAR
                    MLEEVAVVLNGAVGERHVEVGDGFGWSYGHLTPVEAMEIDHAIAECVVSGGVRRSARM
                    AICKWDDPFIDDFLGCKADGSKHWTTNISVEIDQDFIDYLSGDKHDDFGPGGNELAYM
                    VHKKVVEGMLRNGEPGYWNSTFSNEGEVNEVIATNPCGEIALPPTGACVLGHVNLDFF
                    APKVKGEHSDFDGLRRAHELMTRFLIRATYGDMTDDQQREVMHSERRVGVGHLGVQGF
                    LAKHGTRYSDAPYDPQFRNLLTFLADTVRDEAREYAFQLRIPEPVKVTTVAPTGSIAK
                    LPGVSEGIHPIYARHFLRRVRFSMTDPAQVKTIMAAEAAGLLVEECIYDKSGNTYVVA
                    YPTKEKLVAEVEDMGYDPAIVQSADEIDLHSMLAFQAMYQEDYADNAVSFTVNFPEGK
                    YSVEEAADIIEAWLPSLKGTTLMPDGTREQAPYERLTAEQFAQYEVTSVEDSTDEDCA
                TGACPVR"
gene                35597..35704
                    /gene="gp48"
CDS                  complement(35597..35704)
                    /codon_start=1
                       /translation="-IRRLIGRLAGPLLIFMISASIGVTLAAAYVLANP"
gene                35701..35943
                    /gene="gp49"
CDS                  complement(35701..35943)
                    /codon_start=1
                    /translation="-
                       SAGPRAIVRHVTWTLTPDREPDAEPQTYSIQCVVCSEESGRDEDWVEPQNWAISHSGK
                     NPSHHTYREQITRPWRTYMNE"
gene                35940..36188
                    /gene="gp50"
CDS                  complement(35940..36188)
                    /codon_start=1
                       /translation="MATTERDQAAVSPSDCELGSLVMDIATDKVGVVMGHYGGDRVQV
                      RPIKGGAEWDAFRVRPLTAREELSARNAARNAATRQGL"
gene                36266..37357
                    /gene="gp51"
CDS                  complement(36266..37357)
                    /codon_start=1
                    /translation="-
                       SLSLIRKLEQGERSDTRLETARQLAAALHVPTTRLVSDSSEDGATRETIDRWGAVRRA
                       LVTPTAADELEEPTLEGVTASLKAALPLFSGDRFAELGELLPALLRDANVLAELDPDG
                       RALRVKLMQLTGWLLTQTRQFEAAEMALERSLDDSADRLQGAATVNTQCWLLLRRGRL
                       SEARELATEWADETEPRLTKATPAELSAWGWLLLRMSAASVRDNRPGEAEDALRYANS
                       AAVALGREFAPPEDFLRTFGPVTVALKRAENASIIHRPDKVLELATKVPMDKLKPTSN
                       NRNRHLLDVADAYARTRQYGEAVEVLQNIRASSPQWLPNQRYARDIMGRIVARRRTLT
                    PEMRSLADLVGVPM"
gene                37707..38996
                     /gene="gp52 - Integrase"
CDS                 37707..38996
                    /codon_start=1
                       /translation="MEVGQVWVDNDLSATKKNVVRPDFEAMIASKPQAIVCWHTDRLI
                       RVTRDLERVIDLGVNVHAVMAGHLDLSTPAGRAVARTVTAWATYEGEQKAERQKLANI
                       QNARLGKPYTPGIRPFGYGEDHMTVVTAEADAIRDGAKMILDGWSLSAVARYWEELKL
                       QSPRSTAAGGKGWSLRGVKKVLTSPRYVGRSSYLGEVVGDALWPPILDPDVYYGVVAI
                       LNNPDRFSGGPRTGRTPGTLLSGIGQCGYDDCKDTVGGRGYRGVLVYGCKDTHTRTPR
                       SIADARASNATLARLMFPDFLPGLLASGQTESGESAASMHAEAQTLRERLDGLATAYA
                       EGAINLSQMTAGSEALRKKLEVVESELVGSAGIPPFDPVAGVAGLISGWPDLPLPTRR
                     AWVDFCCVVTLNPQRGRHLSNMTVDDHVTIEWRDVAE"
gene                39052..39252
                    /gene="gp53"
CDS                  complement(39052..39252)
                    /codon_start=1
                    /translation="-
                       SKPPQRRPEFLIERYPTLFNSEREKKLPVWVQVKLADLRLLLLQEAGNYEASLEEIDR
                    LESGHHN"
gene                39249..39956
                     /gene="gp54 - Thymidylate synthase"
CDS                  complement(39249..39956)
                    /codon_start=1
                       /translation="MIEIRNDVTVQLVKASATDSDVATAARVSTIGGSHEEVVDLAKD
                       QGLINYLMRDRHGSPFEHTTFTFYVEAPVFVAREFFRHRSGWSYNEESGRYKQLRPVF
                       YAPGHDRNLVQKGKPGAYQFTPGSPSQYAVTLSAMSDAYEQAYRSYEVMLEIGIAREV
                       ARMVLPVGIFTSFYATCNARSLMHFLGLRTISAIAQFPSFPQREIEMVAEKMEDHLAD
                    LMPITYTAFNKNGRVAP"
gene                39953..40216
                    /gene="gp55"
```

```
CDS             complement(39953..40216)
                /codon_start=1
                /translation="
                   TQNTESIAVVYGYHQTRKFAGLVKPENVIPFRLLHLLKDRRPEVIFRTGFGKSAPAWR
                 MLAALEDLAWETGAEIIHERQLREEPAA"
gene            40217..40708
                /gene="gp56"
CDS             complement(40217..40708)
                /codon_start=1
                /translation="
                   GVRMNGWNWIEEGQRIKEQAREIAGLTEEVMAESIVFNHQMIYEGELDSLKAAAASRP
                   ARPCREDGFCDAVYRENDRGGEAREIVGLLAKVVGVNLGASHEGPRPLDGNVWSPRNM
                  MRVVQAVQEVLGEHESLRKELHDRDKKDDRTVAALREALIHAGEAV"
gene            40693..40920
                /gene="gp57"
CDS             complement(40693..40920)
                /codon_start=1
                /translation="
                   NHNDWKSCTCGQKRDFMTRQNAEKALGRAQAKRSRRAGESRRGLRIEHRTYQCSEGGW
                HLTAESRAKYEGWVSA"
gene            complement(40917..41414)
                /gene="gp58"
CDS             complement(40917..41414)
                /gene="gp58"
                /codon_start=1
                /translation="
                   RLTPRAHEIKRVVDILEDPTFDNPEQLAKAVIKEVAEIVQMRDLFAMVHTWADGSKGL
                   NFGPFGSTAEAEAFAKKMAFGGTGRLVPLTSSGVMLANHDGKTSWPGYCWDEACGHAP
                  WMHALDGASRGKCHLESCDCVKFIKDDPSKTKRKPAAKRGAAKGVNEL"
gene            41411..41605
                /gene="gp59"
CDS             complement(41411..41605)
                /codon_start=1
                /translation="
                   KYLKVDIEGKFVGSDVTLYARLDGDEEYTDADLDEIAQNLVNEEYPWGQSIVDERDVP
               EGDRP"
gene            complement(41697..42101)
                /gene="gp60"
CDS             complement(41697..42101)
                /gene="gp60"
                /codon_start=1
                /translation="
                   SNVNSQIRALIKKYRAEEGYREAHVAKYRGEDGEPLPGKTREHDQACTDNAYETMDDH
                   DNLLLELAKLVGPELEVGTAVIVPAGAITVDGSAVNFEGEVTGEVYELEDSDGDVLVR
                SGGLLSYVDPRQLKLVE"
gene            complement(42098..42298)
                /gene="gp61"
CDS             complement(42098..42298)
                /gene="gp61"
                /codon_start=1
                /translation="
                   SIRTIRELDELPDGTEIVIEDKRDTPLFKRGGDWYSPSKAATQNMIAYVNTRRWGVRV
               IERGTQK"
gene            complement(42300..42887)
                /gene="gp62 - kinase"
CDS             complement(42300..42887)
                /gene="gp62 - kinase"
                /codon_start=1
                /translation="
                   SDSLIVGLSGYSRSGKNSAAEALVQYGWKQAAFADKLREFLLAVDPVIPGPYGAGLLR
                   LSVLIKDVGWEYAKDHYPEVRALLQRTGTEAGRKLLGSDVWVNALLAEFSDVPALVVT
                   DVRFPNEAQAVADRGGVLIRVRRPGVGPAKDRVGRVHESEVALDGWAFDHTLINDGSV
                RDLHLKLYGVADLVQLSQAV"
gene            complement(42884..43087)
                /gene="gp63"
CDS             complement(42884..43087)
                /gene="gp63"
                /codon_start=1
                   /translation="MSEPVPGTGRKSPAALRSKTGGKSGGGVSAGSHVVESRFADEGD
                KGGWIRRLVRRREARLWKKEAGE"
gene            complement(43080..43394)
```

```
                     /gene="gp64"
CDS                   complement(43080..43394)
                     /gene="gp64"
                     /codon_start=1
                     /translation="
                        TERVKIGMVDVDSGTVFVGDPCYTITGDASHHIKTWSEWCQKFPWDKKNYDVTEPAGS
                       GIGLSIPTMYGDGSYPVYAEIEEGRVARVTIQFNPEYDEEQDEDE"
gene                  complement(43394..43636)
                     /gene="gp65"
CDS                   complement(43394..43636)
                     /gene="gp65"
                     /codon_start=1
                        /translation="MARERDGWEYVDGGARWAPTVEGQIANMIYDKYGQEYDEARRQL
                      DDLVRAAERGAAERLSEAGYEEAARLIFPTYPEENE"
gene                  complement(43636..43968)
                     /gene="gp66"
CDS                   complement(43636..43968)
                     /gene="gp66"
                     /codon_start=1
                     /translation="
                        SEKRYNLGKSEVAGLGKLYAAGGDALKRDELGLENKEYTVFARLAWFGLAKREEEQRW
                       SITRLGIDFVEGIARVQRTAVTRDREFLTLAGELVRAHDVNQSFYFEGVPA"
gene                  complement(44061..44234)
                     /gene="gp67"
CDS                   complement(44061..44234)
                     /gene="gp67"
                     /codon_start=1
                        /translation="MTRIDLGPAGIVYAHAGWNRWLGTDRVWVHTYENGWERRVRVRP
                     DGRVLYARWIRPS"
gene                  complement(44234..44653)
                     /gene="gp68"
CDS                   complement(44234..44653)
                     /gene="gp68"
                     /codon_start=1
                        /translation="MKKIARTIAVGVTALLAYSLTSTPLASETSPVQVETVASATPSL
                        PTKPCRDDDSDSRNCFWDAGKRGNGKGHSYWVDKAGNVTYLDPKLNSPLKRKAWERKN
                      KALKREYWGEVFGHRLCWAKVGDTSYIYCFDGFKETS"
gene                  complement(44912..45139)
                     /gene="gp69"
CDS                   complement(44912..45139)
                     /gene="gp69"
                     /codon_start=1
                        /translation="MGRMKDLAIDLMTFEEGQLDPRETLELFSTLIASGMAWTLQGSY
                     GRAAQMMINEGWISPLGIITDDGAEMLEAVA"
gene                  complement(45142..45570)
                     /gene="gp70"
CDS                   complement(45142..45570)
                     /gene="gp70"
                     /codon_start=1
                        /translation="MNVLGQIKQYSAYRLSELADTGSPDDHGSDGAVFLTDVRDDLVD
                        RIQHHLNDGDELGDVLDSEGERIRNEVSDSAPAVSTHLKWKQFVDLTAYTEDLSDHGT
                      PTDNTPGGWADMALVQIAYRLIDVLMDEIEDSAETETEEV"
gene                  complement(45692..46105)
                     /gene="gp71"
CDS                   complement(45692..46105)
                     /gene="gp71"
                     /codon_start=1
                     /translation="
                        ITEQILAALTDQNVQDIVDIGAEGGITYWATEPTDEEFAGLPEGKTYTIVEGQGPCFY
                        FGGEREVEAVHYLSRDQIRVAYARLLDLGQQFVNREYHGYIVQSWIERDNKGIDASYI
                     DAGTADVIIQLAALGEIRYG"
gene                  complement(46134..46277)
                     /gene="gp72"
CDS                   complement(46134..46277)
                     /gene="gp72"
                     /codon_start=1
                     /translation="
                        TPKFRTHDLTIRDSKRKDRATTLARRELRQQKYEVSEAAVRVASNA"
gene                  complement(46359..46892)
                     /gene="gp73"
CDS                   complement(46359..46892)
```

```
                     /gene="gp73"
                     /codon_start=1
                     /translation="
                         RCLDDNGPILYPELGSPYRCATCGGTLKAGQRPDGYWTESYGYLVEVEVDDEGPDTEE
                         AIEELGGALRVLARVLGDGMTASGVGGHFTCTEADDLAQALMVGGHKHEAMTFLEGHA
                         EGDDDEDDLHGSVEDFEAYVLELAGLPVPELIEEPEPEIEPKGHDLPTVTVQELIVLI
                    GL"
gene                 complement(47014..47862)
                     /gene="gp47"
CDS                  complement(47014..47862)
                     /gene="gp47"
                     /codon_start=1
                     /translation="
                         SAQGGTWVLTMQPGEIMSTLDFETLDEIATNCLTRPSDAMFADDRLFETHGALFSWAE
                         LGDDILEESNYKSALSLIQGAAGDDADEHVIDGTARHWACGSLRTIYVQVYESYEDYE
                         CECEPTWEHEDECAQDEDDFYCQLYCRIECDGEKCLPEQEFTAAFIEAAELLVGLQQY
                         PIIDESDFSEREWKAFEDNCTEALDQAKREYSDDTIEEQTAIDNLIFQGGDLSELMGY
                        EANAGVDWDKVAEIYAGYRDTYFEELAYEVFLWNVLGYNPDQLTLDIAV"
gene                 complement(48111..48242)
                     /gene="gp75"
CDS                  complement(48111..48242)
                     /gene="gp75"
                     /codon_start=1
                        /translation="-DGYSITPEYSIEQLAELLAELEHHERRAVLELADKVAASHYE"
gene                 complement(48315..48503)
                     /gene="gp76"
CDS                  complement(48315..48503)
                     /gene="gp76"
                     /codon_start=1
                        /translation="MTGKELRQAHEAGKRIVGAVAPAIGREMQYSPRRPNDGLPWIEK
                     GQVHDWARYRSREVQVSQ"
gene                 complement(48566..48964)
                     /gene="gp77"
CDS                  complement(48566..48964)
                     /gene="gp77"
                     /codon_start=1
                     /translation="
                        TTMTTTATVYVHPASWSEYTAALGWAREESPRIAEATSKPEEMPRGSRYYLTPDFKSG
                        FGVAYDGTLIGLFSLVKGRGEGLWNDAVTHKGARRLDCFDGFLPEYYKRFGFVETERV
                     PNWTPGEPDVVFMAL"
gene                 complement(49057..49284)
                     /gene="gp78"
CDS                  complement(49057..49284)
                     /gene="gp78"
                     /codon_start=1
                     /translation="
                        RAFLAGLAGLGVGCVLTWGLVSSPQAPPYPVLDAVQTFNQGFADSKQDDCDQGFAPAC
                     EWVMVLNGIPLPPDGV"
gene                 complement(49281..49457)
                     /gene="gp79"
CDS                  complement(49281..49457)
                     /gene="gp79"
                     /codon_start=1
                        /translation="MVSLDKTPLSISYGRMRSNILNGVRRNNLRALSLEFGERLYLAE
                     VKAHEARLAAVYGK"
gene                 49568..49960
                     /gene="gp80"
CDS                  complement(49568..49960)
                     /codon_start=1
                        /translation="MICGICSEPATHYFATVSRGNAMRCQHHADQFGYPQYLVEINPP
                        KPAIRQPGYVIQFWDGSGNGDETAEIWPFEEELSEWWVQNISACDKRSAEWLVGDAAY
                      HPYQAQFLEAQDVHVEVTRVTCLALVQV"
gene                 complement(49957..50160)
                     /gene="gp81"
CDS                  complement(49957..50160)
                     /gene="gp81"
                     /codon_start=1
                     /translation="
                        TSEIHIRCTGCLEAFKVTDEGVIKVYDAEYGEMDWLVGATLADYIADPTCSNCFGIFD
                     FDTERVTR"
BASE COUNT       8685 a   17622 c   16648 g    8205 t
```

Figure 39: GenBank file of phiELB20