Practical considerations in developing *Clostridium difficile* bacteriophages as therapeutics: characterisation, evolutionary dynamics and model systems

Thesis submitted for the degree of Doctor of Philosophy at the University of Leicester

by

Tolis Panayi, BSc, MSc.
Department of Infection, Immunity and Inflammation
University of Leicester
June 2016

© 2016 Tolis Panayi
Abstract

*Clostridium difficile* is a major nosocomial pathogen responsible for epidemics associated with significant mortality and morbidity. The occurrence of *Clostridium difficile* infections is associated with the use of antibiotics to treat primary infections. Consequently, alternative treatment methods are being considered. Bacteriophages are bacterial viruses that have long been used as therapeutics, are a plausible alternative. This thesis investigates the potential of bacteriophages of *C. difficile* to be used a therapeutic. Eight novel bacteriophages were isolated, all of which were characterised in terms of their morphology, host range, adsorption and growth dynamics. One of these bacteriophages was sequenced on the basis of its broad host range, which included strains of the clinically relevant ribotype 027. Individual and combinations of phages were shown to significantly reduce the amount of bacterial counts *in vitro*, with combinations being more effective than individual bacteriophages. The therapeutic potential of bacteriophages was also tested *in vivo* in a hamster model. Treatment *in vivo* was shown to be able to significantly reduce the bacterial counts in a carriage model where the hamsters were infected with an avirulent strain of *C. difficile* and delay the onset of *C. difficile* infection and death in a hamster model infected with a highly virulent strain of *C. difficile*. The way that *C. difficile* and its bacteriophages co-evolve was tested *in vitro*. The results showed that time progresses bacterial become more resistant to bacteriophages and bacteriophages become more infectious, creating an evolutionary arms race. Genomic analysis of bacteriophage resistant bacteria showed the presence of SNPs mainly associated with cell surface proteins. In conclusion, this study has provided evidence that bacteriophages of *C. difficile* do indeed have the potential of being used as therapeutic agents. However this is not a straightforward process and more research needs to be invested in this area.
Acknowledgements

Completion of this PhD would not have been possible without the guidance, support and encouragement of my supervisor Professor Martha Clokie. I am extremely grateful for not only her support through her scientific expertise, but also for being a good group leader, providing motivation to the entire team and for always having a positive and optimistic outlook, which counterbalanced my own inherent pessimism.

I would also like to thank Dr Gillian Douce and Dr Janice Spencer and the rest of their team at the University of Glasgow for all the animal work done in collaboration with them. Also Dr Paul Monks, Dr Sharmilah Kuppusami and Dr Li Liu of the University of Leicester for their collaboration in the *C. difficile* VOC analysis.

Particular thanks go to Dr Andrew Millard of the University of Warwick for all the sequencing work that he has done that have form an integral part of this work. Also thanks to Natalie Allcock and the late Stefan Hyman for all the TEM work.

Completion of this work would also have been impossible without the love encouragement and moral (and financial!) support of my family. My amazing sister Ntina, my loving grandparents Apostolos, Kikis, Andriana and Ntina (who has passed away but would have given me her wholehearted support) and finally my wonderful parents Spyros and Yianna. I LOVE YOU ALL.

Special thanks to my friends from Cyprus, Giannis, Lytras, Yiota and Josephina for making these four years pleasant and for creating a home away from home

Last but not least, thanks to all past and present members of the great Lab 105 and 107, Jinyu, Ananthi, Janet, Neda, Rin, Guillermo, Ahmed, Ali, Swra, Depesh, Kate, Didier, Peter, Sophie, Sophie, Sam (or Jack) and Krusha, just to name a few, for providing an enjoyable work environment

Finally, a big thank you to my fellow musketeers Aramis (Anisha), Porthos (Jazz) and D’Artagnan (Chris). It has been a wonderful journey and it was good to have taken it with you and I hope we get the chance to work together again.

“As I grow older, I constantly learn more” Solon
List of Publications

Contents
List of Figures .................................................................................................................. 6
List of tables ...................................................................................................................... 8
Abbreviations .................................................................................................................... 9
Structure of the thesis ...................................................................................................... 10
Chapter 1: Introduction and Literature Review .................................................................. 11
1.0.0. Introduction ......................................................................................................... 12
1.1.0 Clostridium difficile ............................................................................................... 13
  1.1.1. Morphology and Ecology .................................................................................... 13
  1.1.2 Metabolism and genomics of C. difficile ........................................................... 14
  1.1.3 Epidemiology ...................................................................................................... 15
  1.1.4 Pathogenesis of C. difficile infection ................................................................... 16
  1.1.5 Toxins of C. difficile ........................................................................................... 17
  1.1.6 Other virulence factors ....................................................................................... 18
  1.1.7 Ribotyping of C. difficile .................................................................................... 19
  1.1.8 The animal model ............................................................................................... 21
1.2.0 Bacteriophages ..................................................................................................... 22
  1.2.1 Brief History of bacteriophages ......................................................................... 22
  1.2.2 Classification and morphology .......................................................................... 23
  1.2.3 Phages Life Cycle ............................................................................................... 24
  1.2.4 Role of Prophages in virulence ......................................................................... 27
  1.2.5 Bacterial Defences against bacteriophages ......................................................... 28
    1.2.5.1 Host receptors ............................................................................................... 28
    1.2.5.2 Restriction–modification systems ................................................................ 29
    1.2.5.3 CRISPR-Cas Systems ................................................................................. 29
  1.2.6 Bacteriophages counter defences .......................................................................... 32
    1.2.6.1 Adaptions to new receptors .......................................................................... 32
    1.2.6.2 Protections against Restriction–modification systems ................................. 32
    1.2.6.3 Protection against CRISPR-Cas Systems .................................................... 33
    1.2.6.4 Escaping abortive-infection mechanisms .................................................... 34
  1.2.7 Applications of Bacteriophages as therapeutic agents ........................................ 34
  1.2.8 Bacteriophages of C. difficile ............................................................................. 36
1.3.0 Volatile organic compounds .................................................................................... 37
1.4.0 Justification of study and, research aims and objectives ........................................ 38
1.4.1 Therapeutic applications of *C. difficile* phages .................................................................38
1.4.2 VOCs of *C. difficile* ........................................................................................................39

Chapter 2: Methods and materials .........................................................................................40
2.0.0 Methods and materials .................................................................................................41
2.1.0 General Methods ...........................................................................................................41
2.1.1 Media, Buffers and Solutions ......................................................................................41
2.1.2 Bacterial isolates and phages ....................................................................................41
2.1.3 Propagation of *C. difficile* .......................................................................................41
2.1.4 Plaque assays ................................................................................................................42
2.1.5 Small drop plaque assay system ................................................................................42
2.1.6 Enumeration of a phage titre .........................................................................................42
2.1.7 Bacterial Growth curve ..............................................................................................42
2.1.8 One-step Growth Curve .............................................................................................43
2.1.9 Measurement of the rate of Adsorption of phages to *C. difficile* cells ...............43
2.1.10 increasing the phage titre on solid media .................................................................44
2.1.11 Spore stock creation ..................................................................................................44

2.2.0 Isolation, Characterisation and development of bacteriophages for therapeutic applications .........................................................................................................................45
2.2.1 Isolation of new phages and *C. difficile* strains .......................................................45
2.2.2 Characterisation of new phages ..................................................................................45
2.2.3 Killing Assays ..............................................................................................................46
2.2.4 Host range expansion ..................................................................................................46
2.2.5 PFGE ...........................................................................................................................46
2.2.6 Extraction of Phage DNA ...........................................................................................47
2.2.7 Phage Sequencing .......................................................................................................48

2.3.0 Therapeutic applications of *C. difficile* phages in an *in vivo* model .......................48
2.3.1 Identifying phage from our current collection that infect the experimental strains and investigate their dynamics .................................................................48
2.3.2 Obtaining high stocks of Bacteriophages .................................................................49
2.3.3 PEG Concentration ....................................................................................................49
2.3.4 CsCl purification ........................................................................................................49
2.3.4 Washing of phage from toxins ....................................................................................50
2.3.5 Measuring Toxin Content ..........................................................................................50
2.3.6 Effects of phage treatment on *C. difficile* colonisation in an *in vivo* model ....50
2.3.7 Effects of phage treatment on C. difficile on morbidity and mortality in an in vivo model.................................................................51
2.3.8 Identification of the phage present in the colon and caecum washes of phage treated animals. .................................................................52
2.4.0 Antagonistic Co-evolution of C. difficile and its bacteriophages.................53
  2.4.1 Co-Evolution Assays ........................................................................53
  2.4.2 Measurement of bacterial resistance/sensitivity to phage ....................53
  2.4.3 Emergence of resistance ....................................................................53
  2.4.4 Measuring the effects of Co-evolution on phages ...............................54
  2.4.5 DNA Extraction from bacterial colonies ............................................54
  2.4.6 SNP Analysis ...................................................................................55
  2.4.7.1 CRISPR analysis .........................................................................55
  2.4.7.2 PCR for the detection of CRISPR arrays ......................................55
  2.4.8 Presence of infecting phage as a prophage .......................................56
2.5.0 Release of Volatile organic Compounds in C. difficile and the impact of phages ...................................................................................56
  2.5.1 Preparation of samples .....................................................................56
  2.5.2 Analysis by PTR-ToF-MS ................................................................56
Chapter 3: Isolation, characterisation and development of bacteriophages for therapeutic applications .................................................................58
  3.0.0 Introduction .......................................................................................59
  3.1.0 Aims and objectives ..........................................................................62
  3.2.0 Results ..............................................................................................63
    3.2.1 Isolation and characterisation of new phages .................................63
    3.2.2 Adsorption of Phages to Endospores ...........................................78
    3.2.3 Phage infection can significantly reduce number of C. difficile ..........79
  3.2.4 Genetic Sequencing of Phage ST .......................................................87
  3.3.0 Discussion .......................................................................................95
    3.3.1 Future Work ...................................................................................98
Chapter 4: Therapeutic applications of C. difficile phages in an in vivo model ....99
  4.0.0 Introduction .....................................................................................100
    4.1.0 Aims and objectives .......................................................................102
    4.2.0 Results ...........................................................................................103
    4.2.1 Identifying phages that infect strains CD1342 and CDB1 ...............103
List of Figures

Figure 1: Clostridium difficile under Electron microscope.................................................................15
Figure 2: Prevalence of C. difficile ribotypes in England by quarter (April 2007–March 2013)...........20
Figure 3: Diagram of a typical myovirus (left) and a myovirus under TEM (right).........................24
Figure 4: Lytic and Lysogenic life cycle of Bacteriophages...........................................................26
Figure 5: CRISPR-Cas System..........................................................................................................31
Figure 6: TEM of phages isolated....................................................................................................66
Figure 7: Adsorption assays of newly isolated phages .....................................................................68
Figure 8: One-step growth curve for phage Φ4N against CDB1 .....................................................69
Figure 9: One-step growth curve for phage Φ16o against CDB1 ....................................................70
Figure 10: One-step growth curve for phage Φ17 against CDB1 .....................................................71
Figure 11: One-step growth curve for phage ΦNE against CDB1 ...................................................72
Figure 12: One-step growth curve for phage ΦNW against CDB1 ....................................................73
Figure 13: One-step growth curve for phage ΦSE against CDB1 ....................................................74
Figure 14: One-step growth curve for phage ΦST against CDB1 ....................................................75
Figure 15: One-step growth curve for phage ΦSW against CDB1 ....................................................76
Figure 16: The plaquing efficiency of ΦST against ribotype 027 and 078 strains..............................77
Figure 17: Percentage adsorption of phages to spores .................................................................78
Figure 18: Using individual vs a combination of four bacteriophages to infect a culture in vitro .........81
Figure 19: Use of pairs of bacteriophages against a C. difficile culture in vitro .............................82
Figure 20: Use of combinations of three bacteriophages against a C. difficile culture in vitro .......83
Figure 21: Using individual bacteriophages vs using a combination of four bacteriophages to infect a culture in vitro.................................................................84
Figure 22: Use of pairs of bacteriophages against a C. difficile culture in vitro .........................85
Figure 23: Use of combinations of three bacteriophages against a C. difficile culture in vitro .......86
Figure 24: Comparison of Φ38-2 and ΦST .....................................................................................90
Figure 25: Genomic map of ΦST ....................................................................................................94
Figure 26: Plaquing efficiency of ΦCDHM1 on bacterial strains CD1342, CDB1 and CDT6 ..........105
Figure 27: One-step growth curve for ΦCDHM1 against CDB1 and CD 1342............................106
Figure 28: Counts of Vegetative cells and spores recovered from hamsters in group 1 ...............110
Figure 29: Counts of Vegetative cells and spores recovered from hamsters in group 2 .............111
Figure 30: Phage treatment reduces C. difficile colonisation in an in vivo model .....................112
Figure 31: High titres of phages were recovered from Caecum and Colon .................................113
Figure 32: Phage ΦCDHM1 is present in the faeces and caecum and colon of treated animals .................................................................115
Figure 33: Phage treatment causes delay in the onset of morbidity and mortality ......................116
Figure 34: Effect of phage treatment on toxin levels in an in vivo model ....................................117
Figure 35: The effects of toxins A and toxins B on a hamster colon...........................................117
Figure 36: Phage recovered from the phage treated hamsters ........................................... 118
Figure 37: The Co-Evolutionary assay ................................................................................. 127
Figure 38: C. difficile strain AIU exposed to ΦST at an MOI of 0.1 in BHI ......................... 132
Figure 39: C. difficile strain AIU exposed to ΦST at an MOI of 0.1 in GMGM ................. 133
Figure 40: C. difficile strain AIU exposed to ΦT9 at an MOI of 10 in BHI ....................... 134
Figure 41: C. difficile strain AIU exposed to ΦT9 at an MOI of 10 in GMGM ................. 135
Figure 42: C. difficile strain AIU exposed to ΦST at an MOI of 10 in GMGM ................. 136
Figure 43: C. difficile strain AIU exposed to ΦT9 at an MOI of 10 in GMGM ................. 137
Figure 44: Average bacterial resistance to phage across all transfers AIU-ST MOI 10 in BHI... 138
Figure 45: Plaquing efficiency of passaged phage ................................................................. 141
Figure 46: The genome of ΦST inserted in the chromosome of sample D2-ST .................. 146
Figure 47: The genome of ΦST inserted into the chromosome of sample A1-ST .............. 147
Figure 48: The plaquing efficiency of phage infected bacteria in vivo ............................... 153
Figure 49: Mass spectra from the VOC headspace analysis of different ribotypes of C. difficile (Kuppusami et al. 2014) .............................................................................. 164
Figure 50: A PCA biplot for the different C. difficile ribotypes ......................................... 165
Figure 51: Dendrogram of C. difficile ribotypes produced by cluster analysis .................. 166
Figure 52: Signal intensity chart of C. difficile ribotypes for selected mass peaks ............. 168
Figure 53: PTR--TOF-MS of VOCs with highest variance during measurement .............. 170
Figure 54: Gas Chromatography plot .................................................................................. 172
Figure 55: Flowchart for the development of a phage therapeutic ................................. 183
Figure 56: Alignment of WT and mutant cell wall hydrolase ........................................... 208
Figure 57: Alignment of WT and mutant membrane efflux pump 1 .............................. 209
Figure 58: Alignment of WT and mutant membrane efflux pump 2 .............................. 210
Figure 59: Alignment of WT and mutant putative amidohydrolase ............................... 211
Figure 60: Alignment of WT and mutant hypothetical protein 1 ..................................... 211
Figure 61: Alignment of WT and mutant hypothetical protein 2 ..................................... 212
Figure 62: Alignment of WT and mutant putative membrane protein .......................... 212
Figure 63: Alignment of WT and mutant purine nucleoside receptor ............................ 213
Figure 64: Alignment of WT and mutant RNA polymerase ........................................... 214
Figure 65: Total CFU and spore count recovered from animals that succumbed to CDI ...... 215
Figure 66: Average total CFU and spore count for the treated and untreated animals ....... 216
Figure 67: Average bacterial resistance to phage across all transfers AIU-T9 MOI 10 in BHI media ................................................................................................................... 217
Figure 68: Average bacterial resistance to phage across all transfers AIU-ST MOI 10 in GMGM .................................................................................................................. 218
Figure 69: Average bacterial resistance to phage across all transfers AIU-T9 MOI 10 in GMGM .................................................................................................................. 219
Figure 70: VOCs whose release was not affected by phage infection .............................. 220
Figure 71: Growth curves of CDB1 and CD1342 .............................................................. 221
Figure 72: Growth dynamics of wild type CDAIUI and CDAIU that underwent coevolution with phage ................................................................. 224
List of tables

Table 1: The host range of new bacteriophages................................................................. 67
Table 2: Genome annotation of ΦST ................................................................................. 91
Table 3: Five phages are able to lyse CDB1 and one phage able to lyse CD 1342 .......... 104
Table 4: Maximum dilution in which toxins were detectable in different phage preparations 108
Table 5: SNPs called for bacteria exposed to phage ............................................................... 143
Table 6: SNPs observed caused frameshift mutations in certain genes .......................... 144
Table 7: CRISPR arrays in the reference strain AIU ......................................................... 150
Table 8: CRISPR arrays in the phage infected clones ......................................................... 151
Table 9: Tentative VOC assignments.............................................................................. 171
Abbreviations

AAD: Antibiotic Associated Diarrhoea
ANOVA: Analysis of variance
BHI: Brain heart infusion
BHIS: Brain heart infusion supplemented agar
BLAST: Basic local alignment search tool
Blastp: protein BLAST
Blastn: nucleotide BLAST
CaCl2: Calcium chloride
CCEY: Cycloserine cefoxitin and egg yolk agar (Brazier's selective medium)
CDAD: C. difficile Associated Diarrhoea
CDI: Clostridium difficile infection
CFU/ml: Colony-forming unit per millilitre
DNA: Deoxyribonucleic acid
dsDNA: double stranded DNA
dsRNA: double stranded RNA
EDTA: Ethylenediaminetetraacetic acid
FAB: Fastidious anaerobe broth
FCS: Foetal calf serum
GMGM: Gut model growth medium
HPA: Health Protection Agency
MgCl: Magnesium chloride
MgSO4: Magnesium sulphate
MLVA: Multilocus variable-number tandem-repeat analysis
NaCl: Sodium chloride
NHS: National Health Service
OD550: (OD550 nm) Optical density at 550 nm wave length
PCR: Polymerase chain length polymorphism
PEG: Polyethylene glycol
PFGE: Pulsed-field gel electrophoresis
PFU/ml: Plaque-forming unit per millilitre
PMC: Pseudomembranous colitis
PFU/ml: Plaque-forming unit per millilitre
RFLPA: Restriction Fragment Length Polymorphism Analysis
RNA: Ribonucleic acid
rpm: Revolution per minute
rRNA: ribosomal RNA
SEM: Scanning electron microscopy
ssDNA: single stranded DNA
ssRNA: single stranded RNA
TEM: Transmission electron microscopy
ToxA: C. difficile toxin A
ToxB: C. difficile toxin B
UPH20: Ultra pure water
VOC: Volatile organic compound
Structure of the thesis

The thesis is comprised of seven chapters (plus appendices) that are briefly outlined here:

- Chapter 1 provides background information on *C. difficile*, outlines healthcare concerns regarding it and outlines the need for alternative therapeutics. It introduces phages as potential candidates for such an alternative, shows the advantages over antibiotics and discusses some of the considerations that need to be taken into account when using them as a therapeutic. It then outlines the aims and objectives of the thesis and lays out its structure.
- Chapter 2 shows the methods and materials used in this study, arranged with sections based on the results chapters 3-6.
- Chapter 3 explores the isolation of novel by bacteriophages by the screening of environmental sources, their characterisation and development. It also tests the dynamics of individual vs multiple phage in lysing bacteria *in vitro*.
- Chapter 4 shows the use of a phage to treat a CDI *in vivo* and the effects of the treatment on the colonisation and disease development in the hamster model.
- Chapter 5 investigates the co-evolution of *C. difficile* and its phages. It investigates how the co-evolution affects the development of resistance to phage by the bacteria and the infectivity of the phage. It also investigates the genotypic changes in the bacteria as a result of co-evolution and development of resistance.
- Chapter 6 investigates the possibility of using VOCs emissions by bacteria as a method for identifying their ribotype. It also investigates the effects of phage infection in the release of VOCs by bacteria.
- Chapter 7 summarises and concludes the work presented throughout the thesis and recommends future work.
Chapter 1: Introduction and Literature Review
1.0.0. Introduction

*Clostridium difficile* is the major causative agent of pseudomembranous colitis (PMC) and antibiotic associated diarrhoea. There were 14165 confirmed cases (26.3 cases per 100,000 population) of *C. difficile* infection (CDI) reported across the NHS in the period between April 2014 and March 2015 (Public Health England 2014). This is a 6% increase from the number of cases for the same period in 2013/2014. This is the first time an increase has been reported in CDIs since 2007. This figure however is still represents an overall 74.5% reduction in the number of cases reported for the period 2007/2008(55,498)(Public Health England 2014). This reduction is attributed to a number of measures implemented across the NHS for the management of CDIs as a result of increased awareness such as major hand hygiene programmes, enhanced cleaning with hypochlorite and the use of hydrogen peroxide vapour disinfection (Hughes et al. 2013). However, while all these preventative measures have led a reduction in the number of reported cases, treatment of CDIs remains a healthcare challenge. In 2012 there were 1646 deaths involving *Clostridium difficile* (*C. difficile*) infection in England and Wales (Office for National Statistics 2013). Treatment of CDIs in UK hospitals have an estimated cost of 20000 pounds per patient and there is an estimated cost of €3 billion in Europe (Bouza 2012; Nanwa et al. 2015). CDIs in nosocomial settings develop as secondary infections when the use of broad spectrum antibiotics to combat primary infections disrupts the natural microbiota of the colon (Crowther et al. 2013). Symptoms may vary from asymptomatic carriage to mild and severe diarrhoea to the life threatening pseudomembranous colitis, an inflammation of the colonic mucosa with the formation of pseudomembranes composed of fibrin, mucous, necrotic epithelial cells and leukocytes (Lyerly et al. 1988).

Despite the preventative measures and deep cleaning procedures that have been adopted in UK hospitals treatment of CDIs remains a healthcare challenge. There are currently three antibiotics used for the treatment of CDIs: vancomycin, metronidazole and fidaxomicin (Aslam et al. 2005). However failure of treatment and recurrence of infection can occur after use of any of these antibiotics (Surawicz et al. 2013). As a result alternative methods of treatment are being considered. Bacteriophages (or phages) are such an
alternative. Phages are viruses which infect bacteria and have been used as therapeutic agents ever since their discovery in 1915 (Summers 2001). Although there have been a number of phages isolated that are able to infect *C. difficile* strains, the amount of research that has gone into their therapeutic potential and application is limited. Therefore there is a need to investigate whether *C. difficile* phages are suitable for application for the treatment of *C. difficile*.

1.1.0 Clostridium difficile

1.1.1. Morphology and Ecology

*C. difficile* is a gram-positive, anaerobic, endospore forming bacterium, belonging to the Clostridiaceae family. The vegetative cells have an approximately size of 3-5 μm in length and between 1.5 to 2 μm wide. Colonies of *C. difficile* are morphologically variable and are 3-5 mm in diameter with irregular, lobate or rhizoidal edges when grown on BHI 7 % blood medium for 48 hours (Aktories & Wilkins 2000). The colonies may be grey or opaque and are non-haemolytic on blood agar. After a long incubation period of approximately 72 h, colonies may develop a light grey or whitish centre, a factor associated with sporulation in *C. difficile*. Endospores can be either terminal or sub-terminal (Aktories & Wilkins 2000).

*C. difficile* was first isolated in the stools of new born babies in 1935 and originally named *Bacillus difficileus* because of the difficulty encountered working with it (Hall & O’Toole 1935). A connection between the bacterium and antibiotic use was made when a high number of patients developed pseudomembranous colitis (PMC) after being treated with clindamycin (Bartlett et al. 1978). Since then *C. difficile* has been recognised as the major causative agent of antibiotic mediated diarrhoea and PMC (Voth & Ballard 2005; Lyerly et al. 1988).

The bacterium has been isolated from environmental sources such as soil, water and mud as well as from the stools and guts of animals such as dogs, cats, horses and pigs (Surawicz et al. 2013; Gerdin et al. 2013; Carroll & Bartlett 2011; Clements et al. 2010). Its ability to form endospores allows it to survive harsh environmental conditions that would kill vegetative cells.
1.1.2 Metabolism and genomics of C. difficile

*C. difficile* has a circular chromosome made of approximately 4 million base pairs and a GC content of about 30%. Some strains, like CD630 have found to have plasmids and mobile genetic elements such as conjugative transposons (Dridi et al. 2004). The *C. difficile* genome contains 426 (11%) CDSs that encode transcriptional regulators including 31 transcriptional antiterminators, 36 signalling proteins, 45 two-component regulatory systems, as well as five orphan histidine kinases and six orphan response regulators (Sebaihia et al. 2006). The large number of regulators most likely reflects the ability of *C. difficile* to monitor multiple environmental stimuli and trigger adaptation to a rapidly changing niche (Sebaihia et al. 2006).

Many strains of *C. difficile* carry prophages or prophage-like elements (Shan et al. 2012a; Sebaihia et al. 2006). A pathogenicity locus (PaLoc) is present in the disease causing strains, that can code for either toxin A or toxin B or both, depending on the strain (Cohen et al. 2000).

Similar with other anaerobes, *C. difficile* has a number of specific pathways by which they facilitate amino acid and sugar fermentation in order to create ATP as a source of energy (Jackson et al. 2006). Due to the bacterium’s anaerobic nature, energy is conserved mainly by substrate-level phosphorylation (Neumann-schaal et al. 2015).

During *C. difficile*’s exponential phase of growth (when there are amino acids and sugars available for metabolism) expression of toxin A and B is repressed but shows a significant increase once the bacteria enter stationary phase. By prioritizing growth over toxin production the bacteria can grow and colonize as much as possible before starting to produce toxins which also indicates that the genes for toxin A and B undergo some type of catabolite repression (Dupuy & Sonenshein 1998). The mechanism of toxin production and regulation in examined in detail in section 1.1.5

The endospore germination of *C. difficile* is occurs as a result of interaction with bile salts encountered after passage through the stomach (Sorg & Sonenshein 2008a). The bile salt germinants bind to the Csp family pseudoprotease CspC which triggers a proteolytic signalling cascade consisting of the Csp family protease CspB and the cortex hydrolase SleC (Kevorkian et
al. 2016) The cascade activated SleC degrades the protective cortex layer, allowing spores to resume metabolism and growth (Kevorkian et al. 2016).

![Clostridium difficile under Electron microscope](image)

**Figure 1: Clostridium difficile under Electron microscope**

Negative staining of *C. difficile* observed by TEM

### 1.1.3 Epidemiology

In recent years *C. difficile* rose from relative obscurity to the main etiological agent of antibiotic associated diarrhoea. It has been found responsible for a number of epidemics in Northern America and Europe. These epidemics have been associated with high instances of morbidity and mortality. One of the first major outbreaks occurred between 2003-2004 in hospitals in Quebec, Canada, where 14000 cases of CDI were reported and the mortality within 30 days of diagnosis rose to 13.8% from 4.7% it was in 1992 (Warny et al. 2005). In 2004-2005 the probably single most serious incident occurred in the Stoke Mandeville hospital in the UK where 334 cases of CDI were reported which resulted in 38 deaths (Buckley et al. 2011; Healthcare Commission 2006). Outbreaks have occurred in France in 2006-2007 and USA in 2008 (Birgand et al. 2010; Jump et al. 2010). A number of studies have also investigated the occurrence of CDI in hospitals across Europe (Bauer et al. 2011). In 2014 there was an outbreak of CDI in a residential care home in the
UK, where five people died within one month of the infection (Clayton & McHale-Owen 2014). The cause behind the majority of these outbreaks was determined to be highly virulent strains of ribotype 027. These strains were characterised by high toxin production, believed to be a result of a mutation in the \textit{tcdC regulator gene} as well as the presence of the binary toxin (O’Connor et al. 2009). Although similar strains were isolated in the past, they were not associated with any disease outbreak. The only difference between the older isolates and the epidemic ones was a resistance to fluoroquinolone antibiotics (O’Connor et al. 2009). This combined with the fact that the bacteria have the ability to form highly resistant endospores make treatment of the disease difficult.

1.1.4 Pathogenesis of \textit{C. difficile} infection

The disease begins when endospores of \textit{C. difficile} are ingested either though an environmental or nosocomial source. Unlike vegetative cells, these endospores are able to survive passage through the stomach. Endospores have evolved so that they can use bile salts such as taurocholate as germinants (Sorg & Sonenshein 2008b). These allow the germination of endospores into vegetative cells after passage through the gut. In the colon \textit{C. difficile} will remain in small numbers until a broad spectrum antibiotic to which it is resistant to is used to treat an unrelated infection. This event will disrupt the normal microbiota allowing \textit{C. difficile} to grow in its place (Baines et al. 2005). Growth of the bacteria into the stationary phase will cause the release of exotoxins which are the mediators of the disease. The toxin damage caused to the cells is the source of the diarrhoea observed in patients. The severity of the diarrhoea depends on factors such as age, clinical condition and toxin type of the bacterial strain (Voth & Ballard 2005). The toxins stimulate the release of cytokines from activated macrophages which in turn leads to neutrophil recruitment and an inflammatory response which leads to the formation of a pseudo-membrane in severe colitis (Burnham et al. 2013). Endospores are shed into the faeces of the patient can contaminate the environment allowing propagation of the disease (Paredes-Sabja et al. 2008).
1.1.5 Toxins of *C. difficile*

The pathogenicity of *C. difficile* strains is linked to a 19.6-kb pathogenicity locus which encodes five genes which contribute to expression of endotoxins and thus the development of disease (Voth & Ballard 2005). Two of these genes; *tcdA* and *tcdB* code for toxin A and toxin B respectively. The proximal locations and high homology between the two proteins suggest that their respective genes arose as a result of a gene duplication event (Voth & Ballard 2005). Genes *tcdR* and *tcdC* are putative positive and negative toxin gene regulators respectively. Gene *tcdE* has a putative holin function. Expression of toxin A and toxin B is controlled by the positive regulator *tcdR* and its antagonist *tcdE*, as well as the global regulator CodY (O’Connor et al. 2009). During the logarithmic phase of the bacterium’s growth, CodY binds in the presence of GTP and BCAA to the *tcdR* - depended promoter P_{tcdR}, which controls the expression *tcdR*. This prevents high level expression of toxin A and toxin B. At the same time *tcdC*, which is a membrane associated protein which is thought to sequester any tiny amount of *tcdR* that is expressed, which in turn prevents further expression (O’Connor et al. 2009). During the stationary phase, depravation of nutrients causes the expression of the toxin genes. In the absence of GTP and BCAA, CodY is unable to bind and repress the function of P_{tcdR} causing the optimal expression of *tcdR*. TcdR forms a complex with RNA polymerase holoenzyme(RNAP). This complex then binds to P_{tcdR}, P_{tcdA} and P_{tcdB} which leads to the maximal expression of *tcdR*, *tcdA* and *tcdB* which allows the production of toxin A and toxin B (O’Connor et al. 2009). TcdE, a holin like protein is thought to facilitate the release of the toxins into the extracellular environment by the formation of a pore on the cells surface (O’Connor et al. 2009).

Toxin A (308kDa), an enterotoxin was originally believed to be the sole mediator of the disease until tcdA-, tcdB+ strains were associated with outbreaks of CDIs (Carroll & Bartlett 2011; Freeman et al. 2010). Toxin B (270kDa) a cytotoxin is 1000 fold more toxic than toxin A to culture cells. Conflicting research exists as to the importance of the two toxins in the
development of disease. Lyras et al created *C. difficile* mutants showing that, in an *in vivo* model (hamster), the strains that were tcdA-, tcdB+ caused a higher degree of mortality than those strains that were tcdA+, tcdB- (Lyras et al. 2009). However research by Kuehne et al suggests that both toxins can cause significant disease (Kuehne et al. 2010). The authors hypothesise that these different observations are due to differences in *C. difficile* strains or the animal model (Carroll & Bartlett 2011; Lyras et al. 2009; Kuehne et al. 2010).

When the toxins are released into the extracellular environment, endocytosis by the host’s cells is required to elicit a cytotoxic effect. Both toxins enter the cell by receptor-mediated endocytosis using non-proteinatious receptors (Voth & Ballard 2005). An acidified endosome is required for translocation within the host cell. Both toxins are glucosyltransferases and they transfer a UDP-glucose to small GTPases such as Rho, Rac, and Cd2c42 in the cell, which are important proteins in regulating signal pathways. Glycosylation of these proteins, disrupt these pathways which result in morphological changes, inhibition of cell division, membrane trafficking and, eventually, cell death (Carroll & Bartlett 2011).

A third toxin, the binary toxin (98.8kDa) exists that is frequently seen in those *C. difficile* strains associated with increased severity of CDIs (Gerding et al. 2014). This toxin is an actin-specific ADP-ribosyltransferase that disrupts the cytoskeleton of the cell leading to excessive fluid loss, rounding of the cell and eventually cell death (Carroll & Bartlett 2011). It consists of two subunits: the enzymatic component *CDTa* and the binding component *CDTb* (Carter et al. 2007). The two subunits are encoded by two genes, *ctdA* and *ctdB*, which are transcriptionally linked to a single open reading frame on a locus separate from the *Pa*loc, called the *ctd locus* or *CtdLoc* (Carter et al. 2007; Gerding et al. 2014). Strains lacking the CtdLoc have a conserved 68bp sequenced instead (Carter et al. 2007; Gerding et al. 2014).

### 1.1.6 Other virulence factors

Although toxins are the major virulence factors in CDIs, other mechanisms exist that are significant in disease progression (Carroll & Bartlett 2011). Adhesins, flagella and heat shock proteins are all thought to be involved in development of disease (Fagan et al. 2009). *C. difficile*, and other gram-
positive bacteria, expresses proteinaceous layers on top of the peptidoglycan cell wall. In C. difficile these surface-layers (S-layers) consist of two different surface layer proteins (SLPs) with different molecular weights both of which are derived from a common precursor (Calabi et al. 2001). The two SLPs are called the high molecular weight (HMW) and low molecular weight (LMW) proteins. The coding gene slpA is translated into a single amino acid chain and then cleaved by post-translational modifications and translocated on the surface of the cell. There the HMW subunits for a single layer and the LMW subunits for a layer on top of that (Tam Dang et al. 2012). These layers along with other embedded proteins are collectively called S-layers. S-layers have been implicated in the adhesion of C. difficile to enteric cells (Fagan et al. 2009).

1.1.7 Ribotyping of C. difficile

There are several molecular subtyping methods that have been investigated for strain characterisation of C. difficile such as multilocus sequencing, multiple-locus variable-number tandem-repeats analysis (MLVA), toxinotyping and restriction endonuclease typing (Killgore et al. 2008). However PCR ribotyping has become the dominant method used in the UK and the rest of Europe (Indra et al. 2008). Ribotyping involves the identification and genomic classification of a bacterial species based on the sequenced of the 16S rRNA gene which is highly conserved within bacterial species (Bouchet et al. 2008). PCR-ribotyping uses specific primers complementary to the 3 prime end of the 16S rRNA gene and to the 5 prime end of the 23S rRNA gene to amplify the variable-length intergenic spacer region (Bidet et al. 2000). This yields a strain profile based on the number and sizes of the amplified interspaced regions is called a ribotype. All strains within a certain species that have the same number and sizes of amplified interspace regions are classified within the same ribotype (Bidet et al. 2000).

There are currently ~650 recognised ribotypes of C. difficile. In recent years, strains of ribotype 027 were found to be the causative agents of epidemics of (O’Connor et al. 2009). These epidemic strains were found to be virtually identical to historic isolates of 027, with the only difference being that they were resistant to fluoroquinolone antibiotics (O’Connor et al. 2009). The
epidemic strains also were found to have the gene for the binary toxin as well as a deletion in the negative toxin regulator gene \textit{tcdC}. This is believed to have led to increased toxin production, which coupled with the antibiotic resistances, led to these strains causing the epidemics they did (O’Connor et al. 2009).

In the UK the \textit{Clostridium difficile} Ribotyping Network (CDRN) has been documenting the distribution of the ribotypes across England and Northern Ireland since 2007 (Freeman et al. 2010). Ever since the monitoring began, there has been a drop in the prevalence of epidemic \textit{C. difficile} ribotypes, such as 001, 027 and 106 and the emergence in the prevalence of other ribotypes such as 078, 014/020, 015 and 005 (Public Health England 2014). However despite their reduction in prevalence, strains of these ribotypes can still be a significant threat.

![Diagram shows disruption of ribotypes of \textit{C. difficile} across England from 2007 to 2013. The prevalence of ribotype 027 has decline in favour of other ribotypes like 014/020, 001, 106, 015, 002 , 078 and 087 (Public Health England 2014)](image)

\textbf{Figure 2:} Prevalence of \textit{C. difficile} ribotypes in England by quarter (April 2007–March 2013)
1.1.8 The animal model

There have been a number of animal models developed to study different aspects of CDIs, including colonisation, disease pathophysiology, intoxication, transmission, recurrence, efficacy testing of potential therapeutics and the impact of strain variability on all of these factors (Hutton et al. 2014). A variety of small animals have been used like mice, hamsters, rats, rabbits, hares and guinea pigs (Hutton et al. 2014). Each of the different models has advantages and disadvantages. Choosing which one to use depends on multiple factors like the complexity of the experiment, the convenience of using a model and the scientific question that is being addressed.

The hamster model is one of the most commonly used animal models. It has been used extensively for studies relating to *C. difficile* pathogenesis, since the symptoms observed during hamster infection mirrors many clinical aspects of human CDI (Buckley et al. 2013). After administration of clindamycin and subsequent challenge with endospores of virulent strains of *C. difficile*, hamsters develop diarrhoea or ‘wet tail’ as well as fluctuations in body temperature, ruffled fur and lethargy, eventually resulting in death (Buckley et al. 2011). The timeframe for the development of symptoms depend on the virulence of the strain of *C. difficile* used (Goulding et al. 2009). For example one studied showed that when comparing CDI in hamster caused by CD630 and CDB1, the animals infected with CD630 developed a less severe disease than the ones infected with CD B1, despite both strains being the same toxinoype (Goulding et al. 2009). These differences were attributed to the different locations within the infected tissues the two strains were found at.

The mice model has been used to study many aspects of the disease such as colonisation, vaccine development and virulence factors. One limitation of this model is the lack of immunological reagents available to study host responses to CDI. Mice are relatively resistant to CDI and do not develop fatal symptoms (Lawley et al. 2009). This is probably due to colonisation resistance provided by the resident microbiota, although these mice can become asymptomatic carriers that persistently shed low numbers of spores (Lawley et al. 2009). However there has been an increase in its use to study CDI as a result of
improved methods of inducing disease susceptibility. Three different mouse infection models exist. The first utilises gnotobiotic (or germ) free mice, the second and third utilise a cocktail and individual antibiotics respectively to clear the microbiota and make the animal susceptible to CDI. In addition, a recent study showed that intrarectal administration of *C. difficile* toxins caused inflammation, upregulation of cytokines/chemokines and increased tissue damage (Lawley et al. 2009).

1.2.0 Bacteriophages

Bacteriophages (or phages) are viruses that infect bacteria. Phages were discovered independently by Frederick Twort (1915) in the UK and by Felix d’Herelle (1917) at the Pasteur Institute in France (Duckworth 1976). The term is derived from "bacteria" and the Greek: φαγεῖν (phagein), "to devour". Like the viruses of eukaryotic cells, phages are obligate intracellular parasites that have the ability to “hijack” their host’s internal processes to facilitate their own replication (Maura & Debarbieux 2011).

1.2.1 Brief History of bacteriophages

Felix De’relle, used phages to successfully treat salmonella infected chickens in pens, results that were successfully reproduced by other researchers of the time (Summers 2001). De’relle carried his research to the next level by testing the effectiveness of phage therapy in humans, after testing them on himself, to treat cholera patients in India. Many of those treated survived and recovered from the disease. These results led to an enthusiasm about phage therapy, with many independent researches and companies producing their own phage preparations (Duckworth 1976). However there were some reservations towards the application of phages. The main points of concern was the poorly understood nature of interaction between phages and bacteria and the lack of standardized quality criteria for the phage preparation (Lu & Koeris 2011). Also another concern was the apparent existence of phage resistant bacteria. Research into phage therapy was largely abandoned due to the onset of World War II and the discovery of antibiotics, which were considered more reliable, had a broader host range and their quality standards
were easier to establish. However important research centres continued to exist, the most important of which were the Pasteur institute in France and the Eliava institute in Tbilisi, Georgia. The majority of research by these facilities was largely unavailable in the English speaking world due to the language barrier and the limited amount of relationships between the Soviet Union and the western world. In recent years however a number of these publications have been translated to English, thus making them more widely available (Chanishvili et al. 2001).

Recent years have seen a renewal in the interest for therapeutic applications of phage, particularly in species of bacteria with known antibiotic resistant strains. A number animal studies have been carried out, involving the use of bacteriophages for the treatment of pathogens such as *Pseudomonas aureginosa*, *staphylococcus aureus*, *Klebsiella pneumonia* and *Enterococcus faecium* (Biswas 2002; Wills et al. 2005; McVay et al. 2007; Vinodkumar et al. 2005).

The results have largely been in favour for both the use and safety of the phage preparations. A number of human safety trials also indicate that phage treatments are generally effective with none to only minor side effects (Bruttin & Brüssow 2005; Rhoads et al. 2009; Wright et al. 2009).

### 1.2.2 Classification and morphology

Phages are classified according to their morphology, type of nucleic acid and mode of replication into 13 families (Hanlon 2007). Phages can have a variety of capsid sizes and shapes, can be tailed or non-tailed and can be enveloped or not. Also their nucleic acids can be ssDNA, dsDNA, ssRNA or dsRNA. However the majority phages belong to the *Myoviridae, Siphoviridae and Podoviridae* families of the order *Caudovirales* (Ackermann 1999; Rossmann et al. 2004). These phages have an icosahedral head which contains a dsDNA genome, with a tail that can be contractile or non-contractile and they have tail fibres at the distal end of the tail (Hanlon 2007). The *Myoviridae, Siphoviridae* and *Podoviridae* families are morphologically distinguished mainly based on their tail. Phages of the *Myoviridae* family have contractile tails, Siphoviridae phages have non-contractile tails and *Podoviridae* phages have very short non-contractile tails.
1.2.3 Phages Life Cycle

As is the case with eukaryotic viruses, phages replicate by infecting their host cells. The life cycle of a phage can be either lytic or lysogenic (Weinbauer 2004). In a lytic cycle, phage infection precipitates a series of events that results in the lysis of the host cell and the release of new virion particles. In the lysogenic cycle, the phage genome becomes integrated in the host genome and replicates along with it.

In either case, the life cycle begins when the phage adsorbs (or binds) to the surface of its host (Shao & Wang 2008). Adsorption occurs when adhesin proteins in the tail fibres at the distal end of the phage tail interact with receptors, which can be either proteins, glycoproteins or lipopolysaccharides, on the surface of the host cell (Shin, J. Lee, et al. 2012; Barr et al. 2013; Tétart et al. 1996). This allows the end of the phage tail to bind to the surface of the cell. Conformational changes cause the tail to contract (in those phages that have a contractile tail) and penetrate the surface of the cell. This allows the phage to inject its genetic material into the host cell.

During the lytic cycle, once the phage’s genetic material enters the cell, it...
may be transcribed and/or translated depending on the type of the nucleic acid. The phage’s genetic material overtakes the replication mechanisms of its host and diverts resources into its own replication. This leads to the production of large amounts of viral nucleic acids, structural proteins and enzymes necessary for release from the cell. The phage particles then begin to assemble and the nucleic particles are packed into the capsid. The combined action of the endolysin and holin proteins degrade the peptidoglycan cell wall causing lysis of the host cell (Govind et al. 2006)

However during the lysogenic cycle, upon entering the cell, the viral genome may become integrated into the host genome. This is facilitated by an integrase gene (Govind et al. 2006). When a phage is integrated into the host genome, it is referred to as a prophage. The prophage will replicate as the host’s genome replicates. A prophage may be exit the genome and go into the lytic cycle. The prophage excision may by spontaneous or as result of environmental stress such as antibiotics, UV light or nutrient depravation (Shao & Wang 2008).
Figure 4: Lytic and Lysogenic life cycle of Bacteriophages

Phage infection begins with the adsorption of the virion on the surface of the cell using the phage receptor (1). Conformational changes on the phage baseplate cause the tail to contract and puncture the cell wall and then inject the viral genome into the cell (2). From that point a phage can either enter the lysogenic cycle or the lytic cycle. In the lysogenic cycle, the viral genome becomes embedded into the genome of the bacterial host (3a). The viral genome continues to replicate as the host DNA replicates (4a). In the lytic cycle, the viral genome produces enzymes that degrade the host genome and begins replication of its own nucleic acid and proteins (3b). Near the end of the infection cycle the viral DNA and proteins assemble together (4b) and by the action of holin and endolysin, degrade the host membrane and release the new virions into the environment (5b). The new virions can then infect more bacterial cells to propagate the life cycle.
1.2.4 Role of Prophages in virulence

Genomic studies indicate that a substantial part of the accessory DNA is not acquired via vertical transfer (from parent cell to descendant) but through lateral gene transfer by transformation, conjugation or transduction (Spetz 2003). Temperate Phages provide an important mechanism for horizontal gene transfer between strains via transduction during the lysogenic cycle (Horgan et al. 2010). Therefore prophages constitute one of the main sources of genetic diversity and strain variation associated with the virulence of many bacterial pathogens including *Escherichia coli*, *Streptococcus pyogenes*, *Salmonella enterica* and *Staphylococcus aureus* (Ohnishi et al. 2001; Aziz et al. 2005; Cleary et al. 1998; Cooke et al. 2007; Rahimi et al. 2012). Many prophages code for extracellular toxins, effector proteins and other factors that can increase the virulence of the host cell. Lysogenic conversion occurs when a phage converts a non-virulent bacterial strain into a virulent one. Lysogenic conversion is thought to be responsible for the emergence of new virulent and epidemic strains such as the *Escherichia coli* O157:H7 that acquired two Shiga toxin encoding prophages (Sp5 and Sp15) and *Vibrio cholerae* that acquired the filamentous phage CTXϕ encoding the cholera toxin (Boyd et al. 2000; Waldor & Mekalanos 1996; Hayashi et al. 2001; Rahimi et al. 2012). Furthermore, prophage induction and mobility, which is often increased by the action of antibiotics, can shape bacterial communities and favour the dissemination of antibiotic resistance genes and other mobile genetic elements such as pathogenicity islands in *S. aureus* thereby promoting bacterial evolution (Duerkop et al. 2012; Zhang & LeJeune 2008; Chen & Novick 2009; Ubeda et al. 2005).

In some cases virulence genes carried within prophages are organized in discrete autonomous genetic elements called “morons”, which are flanked on one side by a σ70-like promoter region and on the opposite side by a factor-independent transcriptional terminator (Fortier & Sekulovic 2013). This allows for expression of the virulence genes during the lysogenic cycle when most of the prophage’s genes are repressed. This also allows for the transfer of
“morons” from one phage to another as a functional cassette terminator (Fortier & Sekulovic 2013).

The phage-mediated transfer of extracellular toxin genes is the classical type of lysogenic conversion. For example, phage \( \phi \text{Ce}\beta \) encodes one of the most powerful toxins ever described, the botulinum neurotoxin type C1 (Hundsberger et al. 1997). Prophage carriage has also been associated with increased toxin production even if they do not code any toxin genes themselves. \( C. \text{difficile} \) lysogens carrying the temperate phages \( \phi \text{C2}, \phi \text{C6}, \) and \( \phi \text{C8} \) have been reported to have increased virulence (Fortier & Moineau 2007). Also in \( C. \text{difficile} \), carriage of the \( \phi \text{CD38-2} \) was associated with increased mRNA levels of all 5 PacLoc genes (Sekulovic et al. 2011). And in \( \phi \text{CD119} \), it was shown that the encoded RepR transcriptional regulator repress toxin gene transcription in \( C. \text{difficile} \) (Govind et al. 2009).

### 1.2.5 Bacterial Defences against bacteriophages

Bacteria have developed an array of mechanisms that allow them to defend themselves against phages. Despite of this however many bacteria still succumb to phage infections. Phages also have mechanisms that allow them to overcome the bacterial defences.

#### 1.2.5.1 Host receptors

Phages utilise bacterial receptors in order to facilitate their adsorption to the host cell. If the phage is unable to adsorb to the host cell, it will be unable to inject its nucleic acid and infection of the cell will not occur (Shao & Wang 2008). Therefore it is not surprising that the first line of defence occurs at the level of the phage receptor. Phase variable expression of the phage receptor has been shown to make the bacterial cell resistant to phage infection. Examples of this include the glucosylation of O-antigen in \( \text{Salmonella enterica serovar Typhimurium} \) and methylation of capsular polysaccharide in \( \text{Campylobacter jejuni} \) (Shin, J. Lee, et al. 2012; Samson et al. 2013). Some bacteria have such as \( \text{E. coli} \) have shown the ability to limit phage infection by reducing the amount of available receptors through quorum-sensing (Høyland-Kroghsbo et al. 2013). Another study found that spontaneous mutations in the
spla gene of Caulobacter crescentus cause the S-layers, that contain the phage receptor to not properly attach to the cell surface. This causes the bacterium to be less susceptible to phage infection (Edwards & Smit 1991)

1.2.5.2 Restriction–modification systems

If a phage genome manages to gain entry into the cell, there are intracellular defence mechanisms it must overcome in order to facilitate its replication. Restriction–modification (R–M) systems are common in bacteria. R-M systems involve endonucleases that can cut DNA at specific recognition sites (Stern & Sorek 2011). However these endonucleases can target the host DNA as well. Therefore the bacterium uses methyltransferase enzymes to methylate its own DNA at the recognition sites and prevent its degradation as the endonucleases cannot recognise the methylated sites (Stern & Sorek 2011). The DNA methylation is sustained through multiple rounds of bacterial replication There are four types of R-M systems based on their structure, recognition site and mode of action (Roberts 2003; Tock & Dryden 2005)

1.2.5.3 CRISPR-Cas Systems

Clustered regularly interspaced short palindromic repeats (CRISPR) are a relatively newly discovered adaptive microbial immune system that that provides acquired immunity against viruses and plasmids (Heidelberg et al. 2009). A CRISPR array typically consists of several direct repeats (23-47 bp) separated by variable sequences called spacers (21-72bp) which are mostly derived from the sequences of viruses and plasmid the cell has come into contact with, and are often adjacent to cas genes (CRISPR-associated) (Heidelberg et al. 2009). The cas genes encode a large family of proteins that carry functional domains such as nucleases, helicases, polymerases, and polynucleotide-binding proteins. There are six main cas genes identified (Sternberg et al. 2014).

The mechanism by which CRISPR provides resistance against foreign genetic elements is not fully characterized. However it is thought that upon insertion of exogenous nucleic acid (either viral or plasmid), a complex of cas
proteins, including cas1 an endonuclease and cas2 an endoribonuclease, recognise the foreign DNA, cleaves it and inserts a novel repeat-spacer unit at the leader end of the CRISPR locus (Sternberg et al. 2014).

During subsequent infection with a viral genome or plasmid, the CRISPR locus is transcribed into pre-crRNA. It is then processed into specific small RNA molecules which correspond to a spacer flanked by two partial repeats, referred to as the mature crRNA, by a cas protein complex. A Cas interference machinery complex is then guided by the crRNA to interfere with the proto-spacer of the invading nucleic acid (Seed et al. 2013). The make-up of the cas complexes differ among bacterial species (Seed et al. 2013). Because of the large diversity of CRISPR arrays and cas proteins, it is believed that the CRISPR/cas system targets both DNA and RNA.

In Clostridium difficile the CRISPR/Cas system is diverse between strains (Hargreaves et al. 2014) In Hargreaves et al, multiple C. difficile spacers were identified that are homologous to known phage and prophage sequences. These include spacers found on plasmids, prophages and chromosomes of the bacterial strains in question. Host range analysis showed that ΦCD38-2 was able to infect strains CD196 and R20291 (Sekulovic et al. 2014). However this was done with a very low efficiency. This was attributed to the fact that these strains have a spacer which identically matches the phage in their sequence in its genome (Hargreaves et al. 2014).
Figure 5: CRISPR-Cas System

A CRISPR array typically consists of several direct repeats (23-47 bp) separated by variable sequences called spacers (21-72bp) and is adjacent to cas genes. When a phage or plasmid enters the cell, a complex of cas proteins, which include endonucleases and endoribonucleases, recognise the foreign DNA, cleaves it and inserts a novel repeat-spacer unit at the leader end of the CRISPR locus. During subsequent infection with the same phage or plasmid, the CRISPR locus is transcribed into pre-crRNA which is then processed into specific small RNA molecules. A Cas interference machinery complex is then guided by the crRNA to interfere with the proto-spacer of the invading nucleic acid. The make-up of the cas complexes differ among bacterial species (Samson et al. 2013).
1.2.6 Bacteriophages counter defences

1.2.6.1 Adaptions to new receptors

Bacteriophages are able to adapt to changes in the phage receptors of host cells. Tailed phages are able to modify their receptor binding proteins to acquire novel receptor tropism. For example the $\lambda$ phage is able to infect a new receptor (OmpF) when the expression of the cognate receptor in its *Escherichia coli* B host, is reduced as a result of mutation (Samson et al. 2013). In this case the phage will be able to infect the cognate receptor as well as the new receptor. There are also instances when a mutated form of the host receptor is expressed. In such a case, phages can evolve to recognize the altered receptor structure.

In those cases when the phage receptor is hidden by a surface component such as capsule, some phages have the ability to gain access to the receptor by hydrolysing such barriers. For example Coliphage K1F (*Podoviridae*) expresses an endosialidase that recognizes and degrades the K1 polysaccharide capsule of *Escherichia coli* (Scholl et al. 2005). Phages can acquire such enzymes by genetic transfer.

There are cases where the expression of the phage receptor occurs only under specific conditions. In such a case the chances of a phage infecting its host would decrease. Some phages have been shown to overcome this by encoding receptor binding proteins with variable specificities. This would allow them to bind to different receptors as they are expressed on the cell surface (Bertin et al. 2011). Phages that infect *Bordetella* spp. switch the specificity of their RBPs by mutating the major tropism determinant (*mtd*) gene, allowing them to adsorb to phase-variable receptors (Lampson et al. 2005)

1.2.6.2 Protections against Restriction–modification systems

When a phage with a dsDNA genome invades a host with a restriction modification system, the phage might be able to survive if the host’s methyltransferase modifies it before it is recognised by the restriction endonuclease. In such a case the nucleic acid of the virus would also be able to survive in other strains that have the same restriction modification system.
However the same DNA will be recognised as foreign in a cell with a different restriction modification system and will be cleaved. Conversely, some bacterial REases can recognize and cleave modified DNA (Samson et al. 2013).

Phages with fewer restriction sites in their genome are less prone to DNA cleavage by an REase, and these phages therefore have a selective advantage. Some restriction enzymes must recognize two sequences in a specific orientation (either direct or inverted) for efficient cleavage (Lampson et al. 2005)

Some phages incorporate modified bases into their genome, and this can also confer passive resistance to some REases (Jensen et al. 1998). However, some bacteria have restriction systems that specifically cleave modified bases.

Some phages use active mechanisms of evasion. In coliphage P1 (Myoviridae), the proteins DarA and DarB are co-injected into the host cell with the phage genome. Both proteins bind to phage DNA, thereby masking type I R–M recognition sites and preventing degradation of phage DNA (Samson et al. 2013).

1.2.6.3 Protection against CRISPR-Cas Systems

A recent study found that a subset of Pseudomonas aeruginosa temperate phages code for five different anti-CRISPR genes which are active against CRISPR subtype I-F systems (Hargreaves et al. 2014). These genes are located near the capsid genes. The anti-CRISPR activity appears to be conferred by a small phage protein (Hargreaves et al. 2014). It has been suggested that the protein interferes with the formation or action of the ribonucleoprotein CRISPR–Cas surveillance complexes (van der Oost et al. 2009). It is not known whether this protein is packaged into the capsid along with the viral genome or if it is produced immediately after entry into the cell (van der Oost et al. 2009). Similar genes have been found in a P. aeruginosa pathogenicity island and in putative mobile elements, suggesting that these anti-CRISPR genes are prone to horizontal transfer between phages (van der Oost et al. 2009).

Some phages like the ones that target Vibrio cholerae serogroup O1 encode their own CRISPR–Cas system. These phages have a cluster of six cas
genes and two CRISPR loci, resembling a CRISPR–Cas subtype I-F system. The fact that some of the spacers in this CRISPR locus are identical to sequences within a phage-inducible chromosomal island-like element (PLE) in the host genome, that has an antiphage function, suggests that the phage encoded CRISPR have the ability to inactivate PLE (Seed et al. 2013).

1.2.6.4 Escaping abortive-infection mechanisms

In addition to being able to overcome the bacterial defences that target viral DNA, phages are also able to cancel the effects of abortive infection (Abi) mechanisms(Samson et al. 2013). These Abi systems have the ability to inhibit the phage replication cycle but result in the death of the host cell. This would allow the non-infected bacteria to escape phage infection. Toxin–antitoxin (TA) systems, a subgroup of Abi systems, lead to bacterial death following activation by phage infection(Samson et al. 2013). These systems are composed of a toxin and a neutralizing antitoxin which keeps the toxin inactive during normal bacterial cell cycle. When a stress, such as a phage infection occurs, the antitoxin is degraded and thus the toxin is able to induce cell death (Samson et al. 2013). Some phages are able to avoid the activation of TA systems by a coding a protein that is functionally identical to the antitoxin. Thus when phage infection occurs the antitoxin does degrade but the toxin is still inactivated by the phage encoded “antitoxin”. In addition, mutations in the phage genome can prevent the degradation of the antitoxin. Other Abi systems are being encoded by lysogenized phages. For example, the defective prophage e14 codes for an Abi system that inhibits translation during T4 phage infection (Samson et al. 2013).

1.2.7 Applications of Bacteriophages as therapeutic agents

Ever since their discovery, phages have been used as a therapeutic agent for a variety of bacterial infections (Summers 2001). Being natural enemies of bacteria makes them suitable for such a role. The emergence of antibiotic resistant bacterial strains has created a demand for novel therapeutics. Phages have certain advantages over antibiotics that make them suitable as therapeutic agents. They have a relatively narrow host range, meaning that they can only infect some strains within one bacterial species.
This means that, in the case of *C. difficile*, they do not disturb the normal intestinal microbiota - one of the causes of CDIs. Secondly they have the ability to self-replicate at the site of the infection while antibiotics lose their efficacy over time. Thirdly being biological agents themselves they can, in contrast to antibiotics, evolve along with their host and try to overcome any resistance that arises to them (Summers 2001).

However there are also a number of considerations and limitations regarding phage therapeutic applications (Kutateladze & Adamia 2010). Firstly, it is essential to diagnose the causative agent of the disease, as the narrow host range of the phage will not allow it to infect outside its host species. Ideally, the infective agent should be grown *in vitro* and infected with a phage preparation to check if it will be effective. This however is not always possible or practical and will depend on the resources of the healthcare facility and the severity. Secondly, in order to develop an efficient phage cocktail it is essential to have a collection of well characterised phages with broad host ranges in order to allow them to infect as many bacterial strains as possible. Ideally, these phage should follow the lytic life cycle to prevent lysogeny and ensure optimal killing efficiency. They should also be sequenced to check whether they contain any genes that can affect the virulence of their host bacteria (Chan et al. 2013).

Phage therapy has been successfully used to treat leg ulcers, Staphylococcus aureus abscesses, Pseudomonas aeruginosa infections in both animals and clinical tests (McVay et al. 2007; Rhoads et al. 2009; Wills et al. 2005). The Eliava institute in Tbilisi, Georgia has a number of preparations including the Pyophage (Pyo) and Intestiphage (Intesti) mixes of phages produced commercially for patient use (Kutateladze & Adamia 2008). The Intralytix inc produces a number of phage preparations used for decontamination of food pathogens in commercial settings. AmpliPhi Biosciences, a USA based company has a number of phage preparations in pre-clinical studies for the treatment of lung infections caused in cystic fibrosis patients by *P. aeruginosa* and the treatment of skin and wound infections caused by *S. aureous* (Biosciences 2015).
1.2.8 Bacteriophages of *C. difficile*

The first reported set of *C. difficile* phages were isolated in 1983 by induction of *C. difficile* strains (Sell et al. 1983). There are currently no lytic phages identified for *C. difficile*. However prophage carriage is high and a number of diverse temperate phages have been identified. The sequences of some phages, including two prophages of CD630, are available on the NCBI database as part of *C. difficile* entries. A number of temperate phages have been identified that have lytic activity, but with different degrees of characterisation like particle morphology host range and molecular analysis.

To date there are eight temperate sequenced phages in published articles: six myoviruses and two siphoviruses (Govind et al. 2006; Goh et al. 2007; Mayer et al. 2008; Horgan et al. 2010; Sekulovic et al. 2011; Meessen-Pinard et al. 2012; Hargreaves et al. 2014). More individual sequenced phages exist on the NCBI database that has not been published in an article as of the writing of this work.

Some of these phages show some interesting characteristics. As already mentioned, carriage of ϕCD119, it was shown that the encoded RepR transcriptional regulator repress toxin gene transcription in *C. difficile* which was associated with reduce virulence while ϕCD38-2 is associated with increased virulence in bacterial strains (Govind et al. 2009; Sekulovic et al. 2011). While some these phages have been found to affect toxin production in lysogenic strains, none of the phages to date has been found to contain actual toxin genes. ΦC2 codes for a holin protein that has a high degree of similarity with the tcde gene in Pac Loc suggesting a phage origin for the latter (Goh et al. 2007). ΦC2 also carries an abiF gene which is involved in the abortive infection systems (Goh et al. 2007).

Some *C. difficile* phages, like the two prophages of CD630, are known to encode for CRISPR arrays (Monot et al. 2011). This would protect the bacterial host against infection and lysis from another phage. The purified endolysin of ϕCD27 was shown to be able to lyse *C. difficile* strains and has been patented for commercial use (Mayer et al. 2008). Hargreaves et al found that ΦCDHM1, a myovirus, contains genes that are homologous to quorum sensing genes in bacterial hosts.
1.3.0 Volatile organic compounds

Volatile organic compounds (VOCs) are organic chemicals that have a high vapour pressure in room temperature that have both natural and anthropogenic sources (Blake et al. 2009). Bacteria produce metabolic VOCs for various reasons, such as during growth, info chemicals for inter and intra-organism communication, and for cell-to-cell communication signals, growth-promoting or inhibiting agents to their own populations, or to other species (Kuppusami et al. 2014; Bunge et al. 2008). A large variety of VOCs are produced by bacteria, including fatty acids, aliphatic alcohols, ketones, dimethyl polysulfides, alkenes, nitrogen-containing compounds and volatile sulfur-containing compounds (Bunge et al. 2008). In recent years there have been a number of studies that have evaluated the potential of VOCs being used as a diagnostic in different biological specimens, particularly in breath, urine and stool to monitor metabolic disorders such as cancer, urinary tract infections, inflammatory lung disease, hepatic dysfunction, diabetes and chronic inflammatory bowel disease as well as for detection of illicit substances such as cocaine (Phillips et al. 2003; Miekisch et al. 2001; Manja & Rao 1983; Gallagher et al. 2008; Huestis et al. 1999; Buszewski et al. 2007; Angriman et al. 2007). In 2014 there was a study published, regarding the development of a method for the detection of *C. difficile* using gas chromatography mass spectrometry HS-SPME-GC-MS (Tait et al. 2014). This was based primarily on the detection of a specific VOC, 2-fluoro-4-methylphenol with isocaproic acid and p-cresol being used as secondary markers. The screening was done on cultures derived from stool samples (Tait et al. 2014).
1.4.0 Justification of study and, research aims and objectives

1.4.1 Therapeutic applications of C. difficile phages

The introduction has so far outlined the pathogenicity and epidemiology of C. difficile and its associated infections. Although the incidences of CDIs have declined in recent years, this is due to preventative measures implemented across the NHS for the such as major hand hygiene programmes, enhanced cleaning with hypochlorite and the use of hydrogen peroxide vapour disinfection (Hughes et al. 2013). However the actual treatment of C. difficile infections (CDIs) remains a healthcare challenge. The difficulty in treating CDIs is mainly due to antibiotic resistance acquired by C. difficile to many of the commonly used antibiotics. Coupled with the existence of strains that have a high production of the bacterial toxins create so called hypervirulent strains. Currently the treatment of C. difficile infection (CDI) relies on three antibiotics: vancomycin metronidazole and fidaxomicin (Aslam et al. 2005). However failure of treatment and recurrence of infection can occur after use of any of these antibiotics (Surawicz et al. 2013). Therefore alternative treatments are being considered.

Phages have the potential to be such an alternative. While there are a number of studies have investigated the application of phage therapy in other bacteria species, the research into the therapeutic applications of phages of C. difficile is limited. Therefore there is a need for increased research into this area. The main aim of this PhD was to assess the practical applications of phages of C. difficile as a therapeutic. However developing a phage as a therapeutic, as already outlined, is not a straightforward process and there are a number of considerations to be made and steps to be taken.

Therefore in order to address these issues, the main objectives of this thesis were as follows:

- The isolation of novel phages capable of infecting C. difficile by the screening of environmental samples
- The characterisation of novel phages based on their morphology, host range, adsorption dynamics and growth dynamics
- The expansion of the host range of exiting or novel phages
The evaluation of the ability of novel and existing phages to effectively infect, lyse and kill *C. difficile* in vitro for individual and multiple phages

- The sequencing of the genome of any phages with an interesting host range and a good therapeutic potential
- The evaluation of the ability of phages to lyse *C. difficile* and prevent the onset of CDI in an *in vivo* model
- Study the Co-evolution of *C. difficile* and its phages and assess how it affects the phenotype and genotype of both.

### 1.4.2 VOCs of *C. difficile*

In addition to the main work of the thesis, a side project was also carried out that aimed to investigate the release of VOCs by *C. difficile* and if they can be used as an identification for the different ribotypes as well as how phages of *C. difficile* affect the release of VOCs. This was done in collaboration with Dr Paul Monks of the department of Chemistry in the University of Leicester. Part of this work has been published in the journal of Metabolomics (Kuppusami et al. 2014). The objectives of this project were:

- To Investigate the VOCs released by different *C. difficile* ribotypes and determined whether said ribotypes can be classified based on their VOC profile by mass spectrometry
- To examine how phage infection affects the release of VOCs in *C. difficile*
Chapter 2: Methods and materials
2.0.0 Methods and materials

2.1.0 General Methods

2.1.1 Media, Buffers and Solutions

All reagents, buffers and solutions used in this study are described in Appendix 1.

2.1.2 Bacterial isolates and phages

Strains of *C. difficile* ribotypes 002, 005, 013, 14/020, 015, 023, 026, 027, 076, 078, 081, 087, 107 used in this study for the isolation and characterisation of novel phages, were previously isolated from CDI patients and environmental sources by Krusha Patel and Katharine Hargreaves in Prof Clokie’s research group. *C. difficile* strains CD1342 and CD B1 used in the *in vivo* studies were kindly provided by Dr Gillian Douce. Phages ΦCDHM1, ΦT9, ΦT15 and ΦX2 used in this study were isolated by Katharine Hargreaves and Swra Rashid.

2.1.3 Propagation of *C. difficile*

BHI 7% Blood 1% agar and Braziers CCEY Agar were used as the standard propagating solid media for *C. difficile*. BHI broth and FA broth were used as the standard propagating *C. difficile*. All liquid and solid media inoculated with *C. difficile* were grown by incubation in an anaerobic chamber (80% N₂, 10% CO₂, 10% H₂) at 37°C.
2.1.4 Plaque assays

The desired *C. difficile* strain was streaked BHI 7% Blood 1% agar plates and incubated in the anaerobic chamber for 48 hours at 37°C. A single colony was then picked and inoculated into 5 ml of FA broth and left to incubate overnight. Then 300μl of the overnight culture and 100μl of phage lysate were added to 5ml of 1:1 mixture of overlay BHI 0.8% Agar (2X) and overlay salts (2X). The mixture was then poured over an agar plate with BHI 1% agar. The plate was then incubated in an anaerobic chamber at 37°C overnight. The plate was then taken out of the chamber and observed for the presence of plaques, or zones of inhibition, in the bacteria lawn.

2.1.5 Small drop plaque assay system

The small drop assay works on the same principle as plaque assays but in this case, instead of the phage being added in the top agar and then poured on the bottom agar, the phage is pipetted unto an already poured top agar.

2.1.6 Enumeration of a phage titre

Serial dilutions of the phage lysate were made. They were then used to make plaque assays or spot tests as previously described in 2.1.4 and 2.1.5. The titre of the phage lysate was then determined by enumerating the plaques on a plate and multiplying it by the dilution factor.

2.1.7 Bacterial Growth curve

The media to be used were pre-reduced in the anaerobic chamber for at least 24 hours. The desired *C. difficile* strain was streaked on CCEY agar plates and incubated in the anaerobic chamber for 48 hours at 37°C. A single colony was then picked and inoculated into 5 ml of FA broth and left to incubate overnight. One millilitre of the overnight culture was added to 49ml of pre-reduced BHI broth, incubated in the anaerobic chamber and its growth observed by Optical Density measurement at 550nm. When the OD reached 0.1. The culture was divided into 5ml alliquotes, one for each time point used. At each time point the culture was used to take OD measurements at 550nm and to make CFU counts by spot tests on BHI 1% agar. The plates were
incubated in the anaerobic chamber for 24 hours at 37°C. Subsequently the CFU were enumerated.

2.1.8 One-step Growth Curve

The media to be used were pre-reduced in the anaerobic chamber for at least 24 hours. The desired C. difficile strain was streaked on CCEY agar plates and incubated in the anaerobic chamber for 48 hours at 37°C. A single colony was then picked and inoculated into 5 ml of FA broth and left to incubate overnight. One millilitre of the overnight culture was added to 49 ml of pre-reduced BHI broth, incubated in the anaerobic chamber and its growth observed by Optical Density measurement at 550nm. When the OD reached 0.2, the culture was infected with an MOI of 10 (this can vary according to experimental requirements) of the desired phage. The culture was left in the anaerobic chamber for 15 minutes with periodic inverting, in order to allow for phage adsorption. The culture was then centrifuged for 10 minutes at 4200g. The pellet was washed once (in order to remove free phages) and resuspended in 5 ml of BHI, which was subsequently added in 45 ml of fresh BHI. The culture was divided into 5 ml aliquotes, one for each time point used. At each time point, 3 ml were used for OD reading at 550nm (triplicate), 1 ml was centrifuged for 5 minutes at 20000g in order to use the supernatant to take PFU measurements and the remaining ml was used to take CFU measurements. All the plates used were incubated in the anaerobic chamber for 24 hours. Subsequently the CFU and PFU were enumerated. For the control, the culture was prepared in the same way, but instead of phage, filtered cultured media was used.

2.1.9 Measurement of the rate of Adsorption of phages to C. difficile cells

A culture of the bacterial Strain was grown in anaerobic conditions as previously discussed and infected in early exponential stage with the Phage strain to be tested with an MOI of 1. The culture was alliquoted to 1.5 ml Eppendorf tubes labelled 0 min, 0.5 min, 1 min, 5 min, 10 min, 15 min, 30 min and 60 min. At each time point the tube was taken out of the anaerobic chamber and
centrifuged at 20000g for 5 minutes. The supernatant was collected and a PFU count was taken. If there is any adsorption to the cell surface then there will be less free phage in the culture and its titre will fall.

2.1.10 increasing the phage titre on solid media

Plaque assays using the desired phage were carried out as described in 2.1.4. After the overnight incubation the soft top agar was scraped off into falcon tubes and the plates washed with BHI broth to retrieve any attached phages. The tubes were left in 4°C overnight to allow the phages to migrate into the broth. The tubes were then centrifuged at 4500xg for 30 minutes and then the supernatant filtered with a 0.22μm filter (Millipore). The titre of the lysate was then enumerated by spot tests.

2.1.11 Spore stock creation

In order to generate spore stocks the desired strain was streaked on a CCEY agar plate and incubated into the anaerobic chamber, for 48 hours in order to generate large district colonies. A single colony was subcultured on a BHI 7% Blood agar plate and incubated for 48 hours into the anaerobic chamber. A single colony was then subcultured one more time in the same way. This is done in order to ensure that the stock that will be produced will be pure. The culture was then streaked on BHI 7% Blood agar plates (~30) and left into the anaerobic chamber for 6 days. They were then removed from the chamber and left on the bench overnight. All culture was then scraped into a PBS:IMS (50% final concentration) solution. All plates were also washed with the same solution to ensure maximum spore recovery. The solution was left for an hour in room temperature in order to kill the vegetative cells and then centrifuged for 20 minutes at 45000g. The pellet was resuspended in 5ml of PBS and placed on a hotplate 60°C for 20 min to kill any remaining vegetative cells. The spores were alliquoted into 1.5ml Eppendorf tubes and stored in -20°C. The spores were then enumerated by CFU count on 1% BHI plates with 0.1% Sodium taurocholate.
2.2.0 Isolation, Characterisation and development of bacteriophages for therapeutic applications

2.2.1 Isolation of new phages and *C. difficile* strains

Samples of soil sediments and water from sewage outflows were collected from and around River Tame, West Midlands UK. Additional samples were kindly provided by Dr Julian Clokie.

The soil and water samples were enriched by adding 2ml/2g to 12ml of FA broth with 0.1% sodium terracholate and D-Cycloserine and Cefoxitin (250 and 8 mg/L). These were incubated for 10 days in the aerobic chamber at 37 °C. The cultures when then collected and centrifuged at 4500g for 30min. The supernatant was collected and filtered using a 0.2μm filter (Millipore). The filtered supernatant was used to carry out spot tests against a number of bacterial strains in order to isolate new phages. In order to isolate new phages against a specific strain of *C. difficile*, 1ml/g of filtered water/soil suspension was added to an exponentially growing culture of the chosen bacterial strain and left to incubate overnight in the anaerobic chamber. The cultures where then collected and centrifuged at 4500xg for 30min. The supernatant was collected and filtered using a 0.2μm filter. The filtered supernatant was used to carry out spot tests against a number of bacterial strains selected across different ribotypes in order to isolate new phages.

2.2.2 Characterisation of new phages

Phages isolated were visualised by Transmission Electron Microscopy (TEM) using negative staining. The morphology of the new phages was confirmed by using two sets of primers targeting holin gene sequences CDHmyoF-CDHmyoR and CDHsiphoF-CDHsiphoR for myoviruses and siphoviruses respectively (Shan et al. 2012b) See Appendix 2 and 3 for conditions and reagents). The host range of phages was determined by spot tests against a representative number of bacterial strains from different ribotypes. Adsorption assays and one-step growth curves were carried out for each new phage
2.2.3 Killing Assays

An exponentially growing culture of C. difficile of the appropriate strain was infected by the selected phage at MOI of 10 and 100. The culture was then aliquoted into 3ml tubes and one aliquot was taken out of the anaerobic chamber every 1 hour for seven hours and then at 24 hours. Each aliquot was used to carry out CFU counts in triplicate. This experiment had 2 biological replicates.

2.2.4 Host range expansion

In order to expand the existing host range of phages different methods were used. Plaques of the phages on a propagating host were picked using a pipette tip, transferred into 10μl of BHI and then used to carry out spot tests using a resistant host strain. The phage was used to infect an exponentially growing culture of the propagating host and left overnight. The culture was then centrifuged at 4500xg and the supernatant filtered using a 0.22μm (Millipore) filter. The supernatant was then used to infect an exponentially growing culture of the resistant host and left overnight. The culture was then centrifuged at 4500xg and the supernatant filtered using a 0.22μm filter and used. The phage was used to infect an exponentially growing co-culture of both the propagating and resistant host. Measurements were taken after 3, 8, and 24 hours of infection. The culture was then centrifuged at 4,500g and the supernatant filtered using a 0.22μm filter and used to carry out spot test using a lawn of the resistant host. The phage was passaged through exponentially growing co-cultures of both the propagating and resistant host. Every 24 hours the phage was recovered through centrifugation and filtration and used to infect another exponentially growing co-culture and to carry out spot test using a lawn of the resistant host. The phage was passaged 10 times.

2.2.5 PFGE

10μl DNAse (30mg/ml) and 2μl RNAse (100mg/ml) were added to 100μl sample of phage lysate of at least 10⁹ PFU/ml and incubated overnight at 37°C. Agarose plugs were made by dissolving 24mg of PFGE agarose (New England Biolabs) into 1.2 ml of 0.5x TBE buffer and heating the mixture at 100°C and then mixing 40μl of the enzyme treated phage lysate with 40μl of the dissolved
agarose. The phage-agarose mixture was then transferred into the plug mould and allowed to set at 4°C for 2-3 hours. Proteinase K was added to the prepared lysis buffer (50 mM EDTA, 50 mM TrisCl pH 9.0, 1% SDS) at a final concentration of 0.5mg/ml. The agarose plugs were removed from the mold and transferred to the lysis buffer mixture and left in a water bath at 55°C overnight. The plugs containing the DNA were then washed using 1x TE buffer. The 1 % PFGE running gel was prepared by mixing 200ml of 0.5xTBE buffer with 2g of PFGE agarose, melted it in a microwave oven, poured into the preassembled gel case and allowed to set. The plugs were then loaded into the gel along with the PFGE ladder (New England Biolabs) and sealed with more melted agarose. The gel was then placed into the PFGE machine and, using 0.5x TBE as running buffer, run for 19 hours at 6V (initial switch time 5s, final switch time 13s, 120° angle). The gel was then stained using 200ml of 0.5x TBE buffer and 10ml of EtBr for 1hr and a picture was taken at a UV cabinet.

2.2.6 Extraction of Phage DNA

The method from the phage hunting protocols from the University of Pittsburgh was used with certain modifications. Phage lysate of high titre (>1x10⁹) was incubated with 10μl of DNase (30mg/ml) and 2μl RNAse (100mg/ml) per ml and incubated overnight at 37°C. A PCR using the 16s universal primers 8F was run in order to determine if there was complete degradation of bacterial DNA. 12.5 μL 1 M MgCl₂ per ml of sample was added and mixed gently followed by an addition of 40 μl of 0.5 M EDTA, 5 μl of Proteinase K (10 mg/mL) and 50 μL of 10% SDS. The mixture was vortexed rigorously and incubated at 55oc for 60 minutes, with flicking and vortexing of the tubes in 20 minute intervals. The mixture was divided into 1ml aliquots into 2ml chloroform resistant tubes. An equal amount of Phenol:chloroform:isoamyl alcohol (25:24:1) was added to each tube. The tubes were inverted several times to mix and centrifuged for 15 minutes at 21000xg at 4°C. The top aqueous layer was recovered and an equal amount of PCL added to and the centrifugation step repeated. The aqueous layer was then added to 2 volumes of ice cold isopropanol and 0.1 volumes of 3M sodium acetate and left overnight at 4°C. The tubes were centrifuged for 20 minutes at 21000xg at 4°C.
The DNA should appear as a white precipitate. The supernatant was removed and the pellet washed with 500 μl of 70% ethanol. The tubes were centrifuged for a final time for 10 minutes at 2100xg at 4°C. The supernatant was carefully removed and the pellet briefly air-dried. The DNA was then dissolved in ~50 μl of 5mM TrisCl PH 8.5. The extracted DNA was then purified using a DNeasy Blood & Tissue Kit from Quigen and stored in -20°C. The extraction was confirmed by gel electrophoresis and quantified using the Qubit dsDNA HS assay Kit (Invitrogen).

2.2.7 Phage Sequencing

Genetic Sequencing was carried out by Dr Andrew Millard of Warwick University. Briefly, libraries were prepared utilising NexteraXT standard protocol with 1 ng of input DNA. Sequencing was carried out utilising 2 x 300 bp with V3 chemistry (Miseq Illumina). Reads were trimmed with sickle using the default parameters. Genome assembly was carried out with SPAdes 3 using --very-careful option. Sam and bam file manipulation was carried out with Samtools. Prediction of ORFs was done with PROKKA.

2.3.0 Therapeutic applications of C. difficile phages in an in vivo model

2.3.1 Identifying phage from our current collection that infect the experimental strains and investigate their dynamics

Endospores were grown on CCEY agar plates and used to make cryostocks as previously described. The growth profiles of both strains were characterised by growth curves as previously described. Spot tests using 20 phages with high titre from our current collection were performed against both strains in order to identify phages able to infect both strains. The dynamics of any phages identified were determined by means of one-step growth curves as previously described.

Any phages identified as being able to infect strains CDB1 and CD1342, after obtaining a high titre on solid and liquid media, had their plaquing efficiency compared. This was done by performing spot tests and plaque assays using those phage(s) stocks using lawns from strains B1 and 1342. A lawn with the propagating host of the selected phage was used as control.
2.3.2 Obtaining high stocks of Bacteriophages
Phages titres were made and determined on solid agar plates as previously described. Strains CD1342, CDB1 as well as the propagating host of each of the selected phages were used as a means of obtaining high titre stocks. The obtained phages were then titred using spot test and plaque assays as previously described.

2.3.3 PEG Concentration
The phage lysate was mixed with polyethylene glycol (PEG) 8000 addition to a final concentration of 10% and NaCl at a final concentration of 1 M. The mixture was put on a rotor for 2 hours at maximum speed in room temperature and then transferred at 4°C overnight. The mixture was then precipitated at 4,000 g for 1 hour at 4°C. The resulting pellet was resuspended in 5 ml of SM buffer and subsequently mixed with an equal volume of chloroform in micro centrifuge tubes. The organic and aqueous phases were separated by centrifugation at 20,000 g for 10 min at 4°C. The resulting top aqueous phase was collected and used to carry out spot tests and plaque assays against the phage’s propagating strain to ensure the concentration was successful.

2.3.4 CsCl purification
In order to purify phages by CsCl gradient, CsCl solutions of densities of 1.3, 1.4, 1.5 and 1.7 were prepared by dissolving CsCl powder with SM buffer (Appendix 1). Using a syringe with a long needle, the different densities of CsCl solutions were layered on top of each other with the higher densities at the bottom (1ml for 1.7, 3ml of 1.5, 2.5ml for 1.4 and 2ml for 1.3) using a 12ml ultracentrifuge tube. Then 1-3 ml of PEG concentrated phage was layered on top of the gradient. The tube was then centrifuged at 120,000x g for 3 hours at 4°C. Then the purified phage, which was visible as a bright blue-white band, was removed from the tube using a syringe with a needle. The purified phages were then transferred to a dialysis tube and placed into beaker with SM buffer and stirred overnight in order to allow the CsCl molecules to diffuse out of the tube and allow SM molecules to diffuse in to the tube. During the dialysis the SM buffer was changed three times every 12 hours for 36 hours. The phage
was then titred using spot tests.

2.3.4 Washing of phage from toxins
An alternative method was used to remove the majority of toxin content. The crude lysate was centrifuged for one hour at 21000g and the majority of the supernatant removed and the pellets resuspended in an equal volume of SM buffer. The phage was then titre using spot tests.

2.3.5 Measuring Toxin Content
The toxin content of the phage preparation was purified by the methods described was measured by using tissue cell culture cells. Two different cell lines were used. H29 cells, which are more sensitive to the presence of toxin A and vero cells which are more sensitive to the presence of toxin B (Carter et al. 2012). A sterile 96 well tissue culture plate was seeded with 100µl 1 x 10^5 cells/ml which were incubated overnight at 37°C with 5% CO₂ or until monolayer is confluent. 90µl of pre-warmed DMEM + 1% FCS + glutamine were added to each well. The phage samples were subsequently diluted 10-fold and the plate incubated at 37°C with 5% CO₂ overnight. The cells were washed with 100µl of room temperature PBS and fixed by adding 100µl of 1% formalin for 10 min and then washed again with 100µl of PBS and stained with 100µl Giemsa stain for 30 min. Excess stain was decanted and the relative toxin levels are calculated by observing the last dilution where cell rounding occurs i.e. where active toxin causes the cells to round, disrupting the monolayer which results in the loss of staining.

2.3.6 Effects of phage treatment on C. difficile colonisation in an in vivo model
Ten female outbred golden Syrian hamsters (Mesocricetus auratus) that weighted approximately 100g were obtained by Harlan Olac, UK. The animals were housed individually in sterile cages and given sterile food and water. Two groups of five animals were used in this experiment. All animals received an oral dose of clindamycin 24 hours before the beginning of the experiment in order to clear the intestinal flora and 400µl of 1M sodium bicarbonate 30 min before administration of bacteria and bacteriophage in order to neutralise the
PH of the stomach in order to prevent denaturation of the phages. All animals were inoculated with $1.15 \times 10^4$ spores. Three animals from each group also received $1 \times 10^8$ PFU of phage, while the other two were used as controls and received an equal amount of the phage dilutant (SM buffer). The treated animals received two further doses of phages at the 8th and 16th hour of the experiment, while the controls received the equal amount of the phage dilutant. All animals were culled 24 hours after the spore challenge and their colon and caecum removed aseptically. During the course of the experiment the animals were observed for signs of CDI.

Both organs were washed with sterile PBS in order to recover the total bacterial load (Lumen associated bacteria) and subsequently homogenised in a stomacher in order to recover the bacteria still attached to these organs (Tissue associated bacteria). The total CFU count was determined by streaking serial dilutions of the recovered bacterial suspensions on the selective CCEY agar. The amount of spores out of the total bacterial population was determined by boiling the samples at 70°C for 30 min to kill any vegetative cells and then streaking those dilutions on CCEY agar.

### 2.3.7 Effects of phage treatment on *C. difficile* on morbidity and mortality in an *in vivo* model

Five female outbred golden Syrian hamsters (*Mesocricetus auratus*) that weighted approximately 100g were obtained by Harlan Olac, UK. The animals were housed individually in cages and given food and water. Telemetry chips (Vital view Emitter) were inserted by laparotomy in the body cavities of the animals at least 3 weeks before challenge with *Clostridium difficile* in order to allow the wounds to heal. The animal cages were then placed on receiver pads that allow monitoring and recording of body temperature and activity (vital view software) (Buckley *et al.* 2011).

One group of five golden Syrian hamsters was used were used in this experiment. All animals received an oral dose of Clindamycin 24 hours before the beginning of the experiment in order to clear the intestinal flora. Following the administration of the antibiotic, the animals were transferred into sterile cages which were placed on the receiver pads and received sterile food water and bedding. All animals received 400μl of 1M sodium bicarbonate orally 30 min before administration of bacteria and bacteriophage in order to neutralise
the PH of the stomach to prevent denaturation of Phages. All animals were inoculated with $0.8 \times 10^3$ spores of CDB1. Three animals from each group also received $1 \times 10^8$ PFU of phage, while the other two were used as controls and received an equal amount of the phage diluant. All animals were monitored closely and culled when their body temperature dropped below 35ºC. The treated animals received further doses of phages every 8 hours post challenge until either reaching the clinical endpoint at or until 72hr after challenge. After culling all animals had their colon and caecum removed aseptically.

Both organs were washed with sterile PBS in order to recover the total bacterial load (Lumen associated bacteria) and subsequently homogenised in a stomacher in order to recover the bacteria still attached to these organs (Tissue associated bacteria). The total CFU count was determined by streaking serial dilutions of the recovered bacterial suspensions on the selective CCEY agar. The amount of spores out of the total bacterial population was determined by boiling the samples at 70ºC for 30 min to kill any vegetative cells and then streaking those dilutions on CCEY agar. During the course of the experiment the animals were observed for signs of CDI and spikes in body temperature.

2.3.8 Identification of the phage present in the colon and caecum washes of phage treated animals.

Using primers K12groupF and K12groupR, which target the capsid of phage CDHM1, a PCR was carried out using the stools, colon and caecum washes from phage treated animals (Appendix 2) (Hargreaves 2012). An initial denaturation temperature of 94ºC was used for 5 minutes followed by a denaturation step at 94ºC for 45 seconds, an annealing step at 48ºC for 45 seconds and an extension step for 1 minute at 72ºC. The denaturation, annealing and extension step were repeated for 30 cycles followed by a final extension temperature for 10 minutes at 72ºC. The PCR products were run on a 1% agarose gel at 100V for 1 hour against a 1Kbp DNA ladder (New England Labs). A 814bp product is expected.
2.4.0 Antagonistic Co-evolution of *C. difficile* and its bacteriophages

2.4.1 Co-Evolution Assays

The *C. difficile* strain AIU was streaked out on CCEY agar and incubated for 48 hours in the anaerobic chamber at 37°C. A single colony was then inoculated into FA broth and grown overnight (4 biological replicates). The culture was then added to pre-reduced BHI broth and grown to early exponential stage (0.2 OD$_{550nm}$). An appropriate amount of the bacterial culture was then inoculated into fresh media (BHI or gut model growth medium). Phage was then added at the desired MOI and allowed to incubate for 12 hours. Then 1% of the culture was inoculated into fresh broth and allowed in incubate for a further 12 hours. This was repeated for a total of 15 transfers. At each transfer the bacteria and phages were isolated and separated by centrifugation and filtration and cryopreserved in 25% glycerol at -80°C (Buckling & Rainey 2002) (Fig 6).

2.4.2 Measurement of bacterial resistance/sensitivity to phage

The sensitivity/resistance of bacteria isolated from a particular transfer to phage isolated from a particular transfer was determined binarily. Using a 10 μl loop, a line/pool of phage from the desired population was drawn across an 1% BHU agar plate, and a single colony was streaked across it. (20 colonies were measured for each of the 12 replicates). If there was inhibition of growth the colony was classed as sensitive, if not was classed as resistant (Buckling & Rainey 2002). During this, the phage titre was determined by plaque assays and normalised through each test to eliminate this variable.

2.4.3 Emergence of resistance

The bacterial populations from each of the twelve transfers was tested against the ancestral phage (phage that was not used in the assay), contemporary phage (phage isolated from the same transfer) and future phage (phage from two future transfers). At the same time the all the phage populations were tested against the ancestral bacteria (20 clones assayed for each replicate). In addition the average bacterial sensitivity/resistance to phage was determined by measuring within each transfer the resistance of bacterial
populations from all transfers to phage from all transfers (Buckling & Rainey 2002).

**2.4.4 Measuring the effects of Co-evolution on phages**

Using a colony that was previously determined to be resistant (see 2.4.2) to its contemporary phage, plaque assays were performed against the bacterial lawn as previously described.

**2.4.5 DNA Extraction from bacterial colonies**

The bacteria isolated from the final transfer for each of the 12 replicates were grown on CCEY agar for 48 hours. Half of the colony was used to determine whether or not it was resistant or sensitive to the ancestral phage as previously described (section 2.4.2). For those that were found to be resistant, the other half of the colony was used to extract DNA by mechanical lysis. The colony was inoculated in a tube containing 60μl of ultra-pure water and 1/3 of the volume 425-600 micron glass beads (Sigma-Aldrich). The tubes were vortexed on a vortex-genie with a microtube foam insert at a speed of 6 for 10 minutes. The tubes were then centrifuged at 20000g for 10 minutes. The supernatant was removed and run through the column of a DNeasy Blood & Tissue Kit (QIAGEN) according to manufacturer’s instructions in order to purify and concentrate the DNA. The Qubit dsDNA HS assay kit (Invitrogen) was used to quantify the amount of DNA. Any sample with a concentration more than 10μg/μl was sent for sequencing. Samples were extracted from the assay using phage ST at an MOI 10 in BHI and phage T9 at an MOI 10 in BHI for a total of 24 samples (Köser et al. 2014). The phage populations from the final transfer were collected and the DNA extracted as previously described (Section 2.2.6.) The DNA was run through the column of a DNeasy Blood & Tissue Kit (QIAGEN) in order to concentrate the DNA. The Qubit dsDNA HS assay kit (Invitrogen) was used to quantify the amount of DNA. Twelve samples for phage T9 and Twelve samples for phage ST with a concentration more than 10ng/μl was sent for sequencing (Köser et al. 2014). A previously sequenced clone of WT AIU (ancestor) was used as control.
2.4.6 SNP Analysis

Genetic Sequencing was carried out by Dr Andrew Millard of Warwick University. Briefly, libraries were prepared utilising NexteraXT standard protocol with 1 ng of input DNA. Sequencing was carried out utilising 2 x300 bp with V3 chemistry (MiSeq Illumina). Reads were trimmed with sickle using the default parameters. Genome assembly was carried out with SPAdes 3 using --very-careful option (http://bioinf.spbau.ru/spades/). Sam and bam file manipulation was carried out with Samtools (http://samtools.sourceforge.net/). SNPs were identified with VarScan (http://varscan.sourceforge.net/)

2.4.7.1 CRISPR analysis

The assembled FASTA sequences of the same 24 samples used in the previous section were obtained along with the sequence of the referenced strain. The CRISPRfinder program online (http://crispr.u-psud.fr/Server/) was used to identify CRISPR-like sequences in the tested samples. The nucleotide Basic Local Alignment Search Tool (BLAST) was used to search the NCBI database for homologues sequences to the identified CRISPR arrays. This was done to ensure that the identified arrays are “true” CRISPR and not just proteins with CRISPR-like sequences. Those verified CRISPR arrays from the phage exposed samples were aligned with the homologues arrays in the Reference strain using the ClustalW2 alignment tool and investigated for any changes.

2.4.7.2 PCR for the detection of CRISPR arrays

Primer sets for the amplification of the entire CRISPR were designed using Primer3Web (http://primer3.ut.ee/). Primers were designed to target regions upstream and downstream the CRISPR array. Primer pairs were designed for each of the ten CRISPR arrays identified (Appendix 2). DNA was extracted by a single colony from a bacterial population exposed to phage by boiling at 100°C. PCR reactions were performed in a total volume of 50μl containing template DNA, 0.2μM of forward and reverse primers, 200 μM of dNTPs, 1.25 units of OneTaq® DNA Polymerase (New England Biolabs) and 1x OneTaq® standard reaction buffer (New England Biolabs). An initial Denaturation temperature of 95°C was used for 5min followed by a denaturation step at 95°C
for 30 seconds, an annealing temperature of (see Appendix 2) for 30 seconds and an extension step at 68°C for 2 minutes. These three steps were repeated for 30 cycles. A final extension step at 68°C was carried out for 10 minutes. The PCR products were run along with a 1kb DNA ladder (New England Biolabs) on a 1% agarose gel at 100v for 1 hour. The expected products differed in size depending on the primers used and the CRISPR array targeted (see Appendix 2). The samples were purified using a QiAquick PCR purification kit and sent to GATC Biotech for sequencing. The sequence was aligned using Mega with the homologous sequence in the reference genome.

2.4.8 Presence of infecting phage as a prophage

The genomes of the phage infected samples were analysed for the insertions of lysogens by aligning them against the genome of the phage using Mauve and Geneious.

2.5.0 Release of Volatile organic Compounds in *C. difficile* and the impact of phages

2.5.1 Preparation of samples

Plates of the desired ribotypes were made by streaking the appropriate *C. difficile* strains on BHI 7% blood 1% agar plates and incubated in the anaerobic chamber for 48 hours at 37°C.

Liquid cultures were prepared by first streaking strain AIU on BHI 7% blood 1% agar plates and incubated in the anaerobic chamber for 48 hours at 37°C. One colony was then used to inoculate 6ml of FA Broth and left overnight. The FA culture was then used to inoculate pre-reduced BHI broth which was left in the anaerobic chamber until it reached early exponential stage (OD 0.2 at 550nm). The sample was then inoculated with phage ΦST at an MOI of 10. An uninfected culture was used as control.

2.5.2 Analysis by PTR-ToF-MS

VOC headspace samples were analysed by Proton transfer reaction time of flight mass Spectrometry (PTR-ToF-MS). For culture plates a custom made glass container which can accommodate a single culture plate was used. The upper section of the container has two outlets, one connected to the PTR–ToF–
MS instrument and the other acts as an inlet for gas flow into the sample container. A culture plate with the lid removed was placed into the glass container with the upper and lower sections of the container secured to form an air-tight chamber. The glass chamber was heated to 37°C by wrapping the chamber with a thermal blanket. Prior to the addition of any bacteria-laden culture plate, anaerobic gas (80 % N2, 10 % H2 and 10 % CO2) (BOC, UK) was supplied into the glass container and a background measurement was taken using PTR–ToF–MS.

For liquid cultures, the glass container was replaced by a rubber stopper that has two outlets, one connected to the PTR–ToF–MS instrument and the other acts as an inlet for anaerobic gas flow into the sample tube. The culture was kept at 37°C by placing it into a waterbath.

Agar plates were analysed for 10 min with a flow of 150 ml min-1 of the effluent fed to the PTR–ToF–MS instrument. Liquid cultures were analysed for 24 hours. Headspace of liquid cultures were also analysed using gas chromatography during the first, second, third, fourth and twentieth hours of the experiment. Samples were taken using Solid-phase microextraction (SPME)
Chapter 3: Isolation, characterisation and development of bacteriophages for therapeutic applications
3.0.0 Introduction

_Clostridium difficile_ is an important nosocomial pathogen, strains of which in recent years have acquired resistance to commonly used antibiotics (Rea _et al._ 2013). Currently the treatment of _C. difficile_ infection (CDI) relies on three antibiotics: vancomycin metronidazole and fidaxomicin (Aslam _et al._ 2005). However failure of treatment and recurrence of infection can occur after use of any of these antibiotics (Surawicz _et al._ 2013) Therefore phages are being considered as an alternative therapeutic option.

Bacteriophages (or phages) have certain advantages over conventional antibiotics that make them suitable as therapeutic agents. They have a relatively narrow host range, meaning that they can only infect some strains within one bacterial species. Therefore, in the case of _C. difficile_, they do not disturb the normal intestinal microbiota. This disruption is associated with CDIs. Secondly, they have the ability to self-replicate at the site of the infection while antibiotics lose their efficacy over time. Thirdly being biological agents themselves they can, in contrast to antibiotics, evolve along with their host and potentially overcome any resistance that arises to them. However, the development of bacteriophages as a therapeutic is not a straightforward process. There are many hurdles both technical and regulatory that must be overcome (Meader _et al._ 2013).

Applying phages as a therapeutic is a long process that has many stages. First and foremost suitable phages have to be isolated. Conventional wisdom dictates that phages will be found alongside their hosts. A previous study from Prof Clokie’s research group found that non-clinical isolates of _C. difficile_, that were isolated from soil and water samples, could be used to isolate phages using enrichment and induction cultures (Hargreaves 2012) Ideally any isolated phages must have a broad host range so it can be able to infect as many strains as possible. Sometimes though a phages’ host range can be very narrow, with phages infecting a single or a few strains within a species (W C Summers 2001). Previous studies have showed that it is possible to expand the host range of a phage by exposing it to a resistant host in the presence of a susceptible host (Koskella & Meaden 2013; Ferris _et al._ 2007)
Since they are being considered for therapy, these phages must be as efficient bacterial killers as possible. Therefore lytic phages are more desirable than temperate. Although it is possible for temperate phages to follow the lytic pathway under the right conditions, there is always the possibility that this type of phage infection will result in lysogeny as a result of the integrase gene. Even though lysogeny in itself means that a phage infection will not result in killing of the bacterium it has other downsides as well. It is possible that the genome of a temperate phage will encode for toxin or other genes that will increase the virulence of bacterium as a result of horizontal gene transfer. It is therefore essential that before any phage is applied as a therapeutic, to be fully sequenced to make sure that such genes are not present.

One of the issues with bacteriophage therapy in relation to *C. difficile* is the latter’s ability to form endospores. Although endospores are metabolically inactive (meaning that it is unlikely that phages are able to infect spores) whether a phage is able to adsorb to an endospore is important (Freeman et al. 2010). Since phage adsorption to a bacterium is irreversible, if a phage has the ability to bind an endospore, it would be excluded from binding to a vegetative cell and prevented from lysing it.

When all these issues have been addressed and the appropriate phages selected, considerations have to be made whether they are administered individually or as a mixture. A cocktail of phages can be more advantageous than individually administered viruses as it can allow coverage for a broad host range (Chan et al. 2013). Also using multiple bacteriophages with the same or similar host range can contribute in countering the development of resistance to one individual bacteriophage as it is less likely that a bacteria strain can evolve resistance to multiple infecting phages (Chan et al. 2013). Previous literature suggests that phages with different morphologies use different types of receptors on the surface of their hosts (Shin, J.-H. Lee, et al. 2012). Assuming that this holds true for *C. difficile*, it should be taken into account for when developing a therapeutic. One of the ways that bacteria can acquire resistance to an infecting phage is the modification or total loss of the receptor. Using a cocktail of phages with different morphologies can help to increase the effectiveness in bacterial killing by expanding the host range covered and accounting for the development of resistance to phage.
Our lab has isolated a large collection of phages from environmental sources able to infect nosocomial strains of *C. difficile*. These phages have been characterised in terms of their morphology, host range, adsorption and growth kinetics. Some of these phages have had their genome sequenced. However research into the therapeutic application *in vitro* has been limited. It was therefore essential to look into this aspect of phage research. In order to archive this it was also necessary to isolate more phages in order to expand the phage collection. This was needed in order to compare the therapeutic potential of individual and multiple phages against a single strain.

Strains of *C. difficile* ribotype 027 have been responsible for a number of epidemics that resulted in cases of high mortality and morbidity. Although the prevalence in UK hospitals has decline in recent years, it still remains a significant health risk. Because of this, it was decided to concentrate on the therapeutic applications of phages able to infect strains of ribotype 027, while at the same time searching for new phages against the other important ribotypes (078, 087, 106, 014/020).

In addition to cell death caused by new phages bursting out of the cell, the “lysis from without phenomenon” exists (Abedon 2011). When multiple phage infect a single cell, the large number of pores created by them can create a change in its osmotic content, causing it to burst, resulting in cell death without phage propagation (Abedon 2011).

In this study, it was attempted to isolate suitable bacteriophages able to infect *C. difficile* strains, characterise them and develop them into appropriate therapeutics. The most promising new phages were sequenced and annotated. The ability of individual vs multiple bacteriophages of infecting a *C. difficile* culture was tested *in vitro*. Described here is the isolation of eight novel phages that are capable of infecting strains of *C. difficile* by soil and water samples enrichment. The morphology of these phages was identified by Transmission Electron Microscopy (TEM) and PCR. The ability of phages to bind to their host and their growth parameters were identified by adsorption assays and one step growth curves. This was done in order to identify how efficiently a phage can bind to its host and the time course which the infection occurs, both of which are important if the phage in question is to be applied in therapy.
The host range of the phages was identified by spot tests and plaque assays. One of the phages identified had a good host range and was able to be propagated easily and was therefore chosen for further experimentation.

The phage selected above (ΦST) along with previously characterised phages were selected for investigation into effectively infecting a bacterial culture and reducing the CFU count within it by means of killing assays. The previously characterised phages chosen for these were ΦX2 which was a siphovirus, ΦT9 and ΦT15 which were both myoviruses. All these were able to infect the majority of the strains of R027 available to our collection. Therefore the ribotype 027 strain AIU was used as the host.

Attempts were made in order to expand the phage host range by exposure of the phage to a *C. difficile* strain it cannot currently infect with simultaneous exposure to the propagating strain of the virus. This was not straightforward and although different previous methods that worked were used, expansion of the phage host range was unsuccessful (Hall et al. 2013).

### 3.1.0 Aims and objectives

The aims of this study were the isolation of new phages able to infect *C. difficile* strains of the major nosocomial ribotypes (027, 078, 087, 106, 014/020) and their characterisation in terms of morphology, host range, adsorption and growth dynamics. Also to expand the host ranges of phages, the investigation of therapeutic potential of phages using strain AIU (R027) as a host and sequencing of the genomes of any new phages with a therapeutically promising profile.

This was archived by screening of soil and water samples in order to isolate new phages and their characterisation by spot tests, adsorption assays and growth curves. The expansion of the host range of phages was attempted by exposure of high titres to phage to phage resistant hosts in the presence of susceptible hosts. The therapeutic potential of phages was measured by investigating the ability of individual and multiple phages to effectively reduce the number of bacterial cells in an *in vitro* culture. Finally, the genome of a novel phage with a therapeutically promising profile was sequenced in order to check for the presence of virulence genes. This is important information to
know for any phage that is intended for therapeutic purposes since the presence of genes that can affect the virulence of bacteria can compromise the therapeutic potential of the phage. For example if a phage that codes for a toxin becomes embedded in the genome of a cell instead of killing it due to the action of the integrase gene, it can actually increase the virulence of the cell thus negating the therapeutic aspects of the phage in question. Examples of this change in virulence as result of lysogeny (termed lysogenic conversion as mentioned earlier) include the *Escherichia coli* O157:H7 that acquired two Shiga toxin encoding prophages (Sp5 and Sp15) and *Vibrio cholerae* that acquired the filamentous phage CTXφ encoding the cholera toxin (Boyd et al. 2000; Waldor & Mekalanos 1996; Hayashi et al. 2001; Rahimi et al. 2012)

### 3.2.0 Results

#### 3.2.1 Isolation and characterisation of new phages

Soil and water samples were enriched in conditions favouring the growth of *C. difficile*. The supernatant of the culture was subsequently tested by spot tests against strain T6 (R076) which was previously found to be a highly permissible host to phage infection, as well as strains representing the main disease causing ribotypes 027, 014/020, 016 and 078. This led to the isolation of eight bacteriophages. All these underwent five rounds of plaque purification to ensure that no mixed populations of phages exist. In all cases, small or medium sized clear plaques were produced. A high titre was subsequently obtained on agar plates by the double layer method, using CDT6 as a propagating strain as all the phage isolated were able to infect it. Their morphology was determined using TEM analysis by negative staining. All of the viruses’ morphology indicated that they belonged to the *Caudovirales* order, the non-enveloped tailed viruses. Six of them (Φ4N, Φ17, ΦSE, ΦSW, ΦNE, ΦNW) had the structure of myoviruses, with long contractile tails and two (ΦST, Φ16o) had the structure of siphoviruses with long non-contractile tails (Fig 7). In order to confirm whether the viruses belonged to the families observed in TEM, PCR analysis using different set of primers that target the holin genes of the phages was carried out. The results confirmed the observations. The host range of the phages was determined by carrying out spot tests against lawns made from a
representative number of strains from each ribotype available to our lab. Phages Φ16o, ΦNE and ΦSW were only able to infect strain T6 (R076). Strain Φ4N and Φ17 were able to infect T6 and one strain of ribotype 014/020. Strain ΦSE was able to infect strains T6 and two strains of ribotype 014/020 and one strain of ribotype 002. NW was able to infect strain T6 and one strain of ribotype 002. ST was able to infect strain T6 and all of the samples of ribotype 027 tested and one strain of ribotype 078 (Table 1).

The degree of phage adsorption was determined for all eight of the isolated phages against C. difficile strain T6. Phage SE showed 20% adsorption after 15 minutes of infection. Phages ΦNE, Φ4N and ΦNW reached about 40-50% of adsorption after 15 minutes. Phages Φ16o, Φ17, ΦSW and ΦST reached 60-70% adsorption after about 30 minutes (Fig 8).

The growth dynamics of the phages were characterised by one-step growth curves against strain T6. All eight of the novel phages showed an ability to inhibit the infected culture but the degree of inhibition varied among the different phages. In phage Φ4N there was inhibition in the growth of the phage infected sample from the onset of the assay which continued through the course of the assay with periodic spikes (Fig 9). Low levels of phage particles were observed through the assay but there did not appear to be any obvious burst. In samples Φ16o and ΦNE there was a low but noticeable degree of inhibition in the growth of the phage infected sample (Fig 10 and 12). There were no phage particles observed in either assay. However, the fact that both phages are able to form plaques, indicate that the phages are able to replicate, suggesting that the inability to see any phages in the culture is technical rather than biological. Samples Φ17, ΦNW, ΦSE, ΦST and ΦSW (Fig 11, 13, 14, 15, 16) showed inhibition of growth in the infected sample throughout the course of the assay. An exponential increase in phage count was observed between the third and fourth hour of infection.

With the exception of ΦST, the rest of the novel phages appear to have long eclipse phases (when a virus is inside a cell synthesising new viruses). However this could simply be due to the rather small number of phages released that were detectable in the 3rd or 4th hour of infection. The fact that
some of these phages have low adsorption rates would also account for the low amounts of phage seen.

Because of its ability to infect multiple strains of ribotype 027, ΦST was also tested for its plaquing efficiency across these strains. It was found to be able to infect all of the tested strains (25/25). The plaquing efficiency of the phage did not have statistically significant differences across the tested strains (2-way anova, p>0.05) (Fig 17).
A total of eight novel phages were isolated by the enrichment of soil and water samples. Phages Φ4N, Φ17, ΦSE, ΦSW, ΦNE, ΦNW appeared to be myoviruses with long tails while phages ΦST and Φ16o appeared to have the morphology of long tailed siphoviruses. These morphologies were confirmed by PCR using two different sets of primers targeting the holin gene of siphoviruses and myoviruses.

Figure 6: TEM of phages isolated
The eight new phages isolated were used to infect a number of *C. difficile* strains belonging to different ribotypes. Three of the phages were able to infect only one strain. Another three phages could infect two strains. One phage could infect four strains. The last phage was able to infect 25 strains, 23 of which belonged to ribotype 027. All the phages were able to infect strain T6 (R076) which was previously observed to be a permissible phage host.

### Table 1: The host range of new bacteriophages

| Ribotype | Strain | 4N | 17 | 16o | SW | SE | NW | ST | AHE | AIL | AIJ | ATH | ATG | AMP | AQB | ATB | APA | AIN | AUE | ASU | ARS | ASH | ASK | AQZ | AJZ | ANS | AME | AMR | ARN | AOE | ATJ | ATT | AHR | AOU | ARW | AQI | AQM | AQQ | AST | AIV | AHS | ATU | ATO | ATM | ANC | APL | ARD | AHK | ARB | ATR | ASO |
|----------|--------|----|----|-----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 002      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 006      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 015      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 016 / 020| AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 019      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 023      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 026      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 027      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 076      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 078      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 081      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 087      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| 307      | AMT    |    |    |     |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

The eight new phages isolated were used to infect a number of *C. difficile* strains belonging to different ribotypes. Three of the phages were able to infect only one strain. Another three phages could infect two strains. One phage could infect four strains. The last phage was able to infect 25 strains, 23 of which belonged to ribotype 027. All the phages were able to infect strain T6 (R076) which was previously observed to be a permissible phage host.
Figure 7: Adsorption assays of newly isolated phages

Adsorption assays were used to determine the time required and the degree to which adsorption of each of the phages isolated against *C. difficile* strain T6. Phage ΦSE showed 20% adsorption after 15 minutes of infection. Phages ΦNE, Φ4N and ΦNW reached about 40-50% of adsorption after 15 minutes. Phages 16o, Φ17, ΦSW and ΦST reached 60-70% adsorption after about 30 minutes (SEM, n=3).
Figure 8: One-step growth curve for phage Φ4N against CD AIU

Infection with phage caused an inhibition of the CFU count in relation to the control culture from the first time point. Fluctuations in the CFU numbers continued through the course of the experiment. There were no significant fluctuations observed in the number of PFU. No observable phage burst was detected (SEM, n=2).
Figure 9: One-step growth curve for phage Φ16o against CD AIU

The phage infected culture was slightly inhibited from hours 1 to 7 but after that the CFU numbers recovered to the levels of the control culture. No phage was detected in the infected culture (SEM, n=2).
Figure 10: One-step growth curve for phage Φ17 against CD AIU

The phage infected culture showed an inhibition in the CFU count 2 hours post infection in relation to the uninfected control. This drop continued until the third time point after which the culture began to recover. An exponential increase in PFU count was observed 3 hours post infection and it plateaued at the eight hours (SEM, n=2).
Figure 11: One-step growth curve for phage ΦNE against CD AIU

The phage infected culture was slightly inhibited thorough the course of the experiment in relation to the uninfected control. No phage was detected in the infected culture (SEM, n=2).
The phage infected culture showed an inhibition in the CFU count 2 hours post infection in relation to the uninfected control. This drop continued until the eighth hour of infection after which the culture began to recover. An exponential increase in PFU count was observed 3 hours post infection and it plateaued at the eighth hour (SEM, n=2).

Figure 12: One-step growth curve for phage ΦNW against CD AIU
Figure 13: One-step growth curve for phage ΦSE against CD AiU

The phage infected culture showed an inhibition in the CFU count 2 hours post infection in relation to the uninfected control. This drop continued until the fifth hour of infection after which the culture began to recover. An exponential increase in PFU count was observed 3 hours post infection and it plateaued at the ninth hour of infection (SEM, n=2).
The phage infected culture showed an inhibition in the CFU count two hours post infection in relation to the uninfected control. This drop continued until the fifth time point after which the culture began to recover. An exponential increase in PFU count was observed 2 hours post infection and it plateaued at the ninth hour of infection (SEM, n=2).

Figure 14: One-step growth curve for phage ΦST against CD AIU
Figure 15: One-step growth curve for phage ΦSW against CD AIU

The phage infected culture showed an inhibition in the CFU count 2 hours post infection in relation to the uninfected control. This drop continued until the seventh hour of infection after which the culture began to recover. An exponential increase in PFU count was observed 4 hours post infection and it plateaued at the ninth hour of infection (SEM, n=2).
The plaquing efficiency of ΦST was tested against 23 different ribotype 027 strains and one ribotype 078 strain. The phage was found equally effective against all the strains tested with no significant statistical difference occurring (SEM, n=3) (2-way anova, p>0.05).
3.2.2 Adsorption of Phages to Endospores

The ability of phages to adsorb to *C. difficile* endospores was tested *in vitro*. Phage ΦST was used to infect endospores of strain AIU *in vitro*. These adsorption assays against endospores showed that phages are not likely to bind to them. There was not significant observable reduction in PFU count in relation to the control (Fig 18). This inability of phages to bind to endospores can probably be attributed to the fact that the coating of the endospores is of different composition to the vegetative cells. This information is very important from the scope of therapy as had the phages been able to bind to endospores and, due to the nature of endospores, been unable to infect them that would mean that any phages that bind would be excluded from lysing a vegetative cell.

![Graph showing percentage adsorption of phages to spores](image)

**Figure 17: Percentage adsorption of phages to spores**

The absence of any reduction in the phage titre in the presence of endospores indicated that phages are unable to adsorb to endospores (SEM, n=2).
3.2.3 Phage infection can significantly reduce number of *C. difficile*

The ability of multiple and individual phages to lyse the bacteria was tested *in vitro*. In all cases where phages were used there was some inhibition of the infected culture in relation to the control. The degree of the inhibition depended on three factors a) the Multiplicity of infection (MOI) b) the phages used and c) whether the aforementioned phages were used in combination or not.

When an MOI of 10 was used, all the phages showed an ability to lower the CFU count in the infected culture when used individually. Some phages however were more effective than other. The siphovirus ΦX2 caused the least amount of reduction while the reduction caused by the siphovirus ΦST and myoviruses T9 and T15 caused greater amount of reduction. All phages caused the highest degree of reduction possible by the second hour of infection and prevented the regrowth of the phage seven hours post infection (Fig 19).

When the phages were used in combinations of two, they showed a greater ability to reduce the number of bacteria than their individual components did. The paired phages were able to reduce the number of bacteria by at least 2 logarithmic degrees. As it was the case with the individual phages, some combinations were more effective than others. The pairs of ΦX2-ΦT15 and ΦST-ΦT15 were not as effective as the other combinations. In this case the reduction in CFU count reached a plateau at the fifth hour of infection (Fig 20).

When a combination of three phages was used, the degree of reduction was similar to the one observed when two phages were used. No combination could be seen to be better than the others. The reduction in CFU count reached a plateau at the fourth hour of infection. A similar situation as above was observed when a four phage combination was used. It should be noted that the infected bacterial cultures were able to recover 24 hours post-infection but not to the degree of the control culture despite the fact that the latter had entered death phase (Fig 19 and 21).

At an MOI of 100, individual phages showed a greater ability to reduce the number of CFU *in vitro*. They were able to cause a reduction of between 2-4 logarithmic degrees, depending on the phage used. The reduction reached a plateau at the second hour of infection (Fig 22).

When combinations of two phages were used, a greater reduction in
CFU was observed when infecting the culture with the majority of the pairs. Pairs ΦST-Φ15 and ΦST-Φ19 were not as effective as the remaining pairs. The reduction seemed to reach a plateau between 3 and 6 hours post infection (Fig 23).

When three phages were used, the degree of reduction did not seem to increase. In fact in some cases a combination of three phages seemed to deliver a lower amount of reduction than some of combinations of two phages. The CFU reduction plateaued at the 5th hour of infection (Fig 24).

When all four phages were used, they caused a reduction of CFU into undetectable levels at about 4 hours post-infection. Moreover the culture had not recovered 24 hours post-infection (Fig 22). However when the culture was streaked 72 hours post infection on CCEY agar and 1% BHI agar, viable colonies were observed after 48 hours of anaerobic incubation on the CCEY plates but not the BHI agar plates. This suggests that endospores remained in the culture that were not killed by phages.

These results seem to suggest that the high MOI and the use of multiple bacteriophages are both required in order to effectively reduce or eradicate the number of bacteria in a culture. When used at an MOI of 10, a combination of two bacteriophages showed to, in some cases, reduce the number of CFU more effectively than a combination of three or four bacteriophages at the same MOI. However when the MOI was at 100, the combination of four bacteriophages was able to reduce the number of bacteria to non-detectable levels. These findings are not surprising, as the use of multiple bacteriophages is known from other studies to be used to combat the problem of bacterial strains becoming resistant to one bacteriophage (Kutateladze & Adamia 2010). The higher MOI increases the likelihood of successful interactions between host and phages or can induce lysis from without.
Figure 18: Using individual vs a combination of four bacteriophages to infect a culture in vitro

At an MOI of 10 ΦST, ΦT9 and ΦT15 were able to cause a reduction of ~2 logs in the infected culture compared to the uninfected control. ΦX2 caused a smaller reduction in the CFU count. The combination of four phages caused a greater reduction than any of the individually used phages by approximately 3 logs. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value=0.0001) (SEM, n=3).
Figure 19: Use of pairs of bacteriophages against a *C. difficile* culture *in vitro*

At an MOI of 10, combinations of two bacteriophages were able to effectively reduce the CFU count of an infected culture compared to the uninfected control. Pairs of ΦST-ΦX2, T15-T9, ΦX2-ΦT9 and ΦST-ΦT9 were able to reduce the CFU count by approximately ~3 log. Pairs of ΦST-ΦT15 and ΦX2-ΦT1 performed lower effect reducing the CFU count by ~2logs. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value= 0.0001) (SEM, n=3).
Figure 20: Use of combinations of three bacteriophages against a *C. difficile* culture *in vitro*

At an MOI of 10, combinations of three bacteriophages were able to effectively reduce the CFU count of an infected culture compared to the uninfected control. All possible combinations of ΦX2, ΦST, ΦT15 and ΦT9 were able to reduce the CFU count in the infected culture by ~3 logs in relation to the uninfected control. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value = 0.0001) (SEM, n=3).
Figure 21: Using individual bacteriophages vs using a combination of four bacteriophages to infect a culture *in vitro*

At an MOI of 100 ΦST, ΦT9, ΦT15 and ΦX2 were able to cause a reduction of ~2-4 logs in the infected culture compared to the uninfected control. The combination of four bacteriophages caused a greater reduction than any of the individually used phages by approximately 7 logs were no bacteria where detectable. The CFU numbers in all cultures apart from the 4 phage combination were able to recover 24 hours post infection. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value= 0.0001(SEM, n=3).
Figure 22: Use of pairs of bacteriophages against a C. difficile culture in vitro

At an MOI of 100, combinations of two bacteriophages were able to effectively reduce the CFU count of an infected culture compared to the uninfected control. Pairs of ΦST-ΦX2, T15-T9, ΦX2-ΦT and ΦST-ΦT9 were able to reduce the CFU count by approximately 6 logs. Pairs of ΦST-ΦT15 and ΦX2-ΦT15 had a lower effect reducing the CFU count by ~2logs. The CFU numbers in all cultures were able to recover 24 hours post infection. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value= 0.0001) (SEM, n=3).
Figure 23: Use of combinations of three bacteriophages against a *C. difficile* culture *in vitro*

At an MOI of 100, combinations of three bacteriophages were able to effectively reduce the CFU count of an infected culture compared to the uninfected control. All possible combinations of ΦX2, ΦST, ΦT15 and ΦT9 were able to reduce the CFU count in the infected culture by ~2-5 logs in relation to the uninfected control. The CFU numbers in all cultures were able to recover 24 hours post infection. The reduction caused by all phage combinations was statistically significant (1-way anova, p<0.05, P value= 0.0001) (SEM, n=3).
3.2.4 Genetic Sequencing of Phage ST

Phage ΦST is a temperate siphovirus with a dsDNA genome of 41318bp. It has predicted 51 ORFs the majority (44/51) of which are located on the sense stand when starting the genome with the small subunit terminase gene. Out of the 51 ORFs, 22 have been assigned a predicted function based on nucleotide homology, amino acid homology and identification of a protein domain. Another 8 ORFs were identified as conserved protein of unknown function and 21 ORFs identified as hypothetical proteins. Out of these 21 hypothetical proteins, one was found not to have any hits in the NCBI database (Table 2). When using Blastn tool to compare the phages nucleotide sequence on the NCBI database, the results indicate a high degree of similarity with the C. difficile siphovirus Φ38-2 which has been shown to have similar host range to ΦST (Sekulovic et al. 2011). When the two genomes aligned they showed 95.1 % identity (Fig 25). Parts of the minor capsid protein and tail fibre protein were the main regions of dissimilarity. The genome shows a high degree of mosaicism as some genes like the ones for lysis, lysogenic conversion, DNA replication and integration are more similar to the homologous genes in Φ 38-2 than the genes for DNA packaging, capsid and tail morphogenesis which show a higher degree of similarity with Φ111 and Φ 146. In addition to this mosaicism, many of the genes appear to be closely related to prophages in C. difficile as many (33/51 ORFs) of the top hits in BLASTn are C. difficile strains with identity range of 98-100%

The genome has an average G+C content of 30.8% but there are regions in the genome with above or below the average content. The DNA packaging, head morphogenesis and tail morphogenesis have an average content of 32.70% while the lysis, lysogenic control, DNA replication and integration have an average content of 28.4% The DNA plotter tool was used to construct a circular genomic map of the genome of ΦST (Fig 26).
The first cluster of genes (orf1-orf3) corresponds to head DNA packaging functions. All three genes have a high degree of similarity to the homologous proteins in Φ111. Orf1 and orf2 have domains corresponding to the small and large terminase subunits respectively (Phage_terminase and Phage_terminase2) while Orf3 had a Phage_portal domain. Terminase proteins facilitate the assembly of the capsid and tail proteins while the portal protein has the function of forming a hole, or portal, that enables DNA passage during packaging and ejection. It also forms the junction between the phage head and the tail proteins.

The second gene cluster lies between orf4 and orf11. Out these seven orfs, only 3 could be assigned a function based on the presence of conserved domains and top Blastn hits. Orf4 had a domain which corresponds to the minor capsid protein and had a high degree of homology with Φ111. Orf7 had a high degree of homology with Φ111 and has the Phage_GP20 domain which is present in minor structural proteins. Orf8 was assigned a phage coat function based on the homologous proteins in Φ38-2 and Φ111.

The third main gene cluster was the tail morphogenesis (orfs18-23). Orf18 had a tape_meas_TP901 domain which has a tail tape measure function. Orf19 and orf20 had the tail protein domains YomH and PblB respectively. Orf22 and orf23 had the PHA02515 tail fiber domain. Orf24 was assigned a putative holin function based on its similarity to the homologus protein in Φ38-2 but no conserved domains were found. Orf25 had a MurNAc-LAA domain which indicates function as an endolysin. These two proteins are responsible for lysing the cell wall at the end of the replication cycle allowing the new phages to be released into the environment.

There is a cluster of genes on the antisense strand downstream the lysis genes (orf30 to orf34). In other phage genomes, genes in that location were found to have lysogenic conversion domains, which are thought to be involved in increasing the virulence of the bacterial host when the phage in present as a lysogen. However with the exception of orf 32 which had a ParA domain which indicates function as an ATPase no other domains could be found on the rest of the genes therefore no function could be assigned.

The region between orfs35-49 appeared to be the cluster involved in DNA replication function. Orf35 was assigned as DNA helicase based on the
DEXDc it containe. Orf 40 appeared to be a replication termination protein based on BLASTn results but no conserved domains were found. Orf 46 appeared to code for a single strand DNA binding protein as it contained a SSB_OBF domains and orf 49 contained domains corresponding to sigma factors.

The last gene cluster was comprised of Orfs 50, 51, and 37 and appears to be involved in lysogeny control. Orf 50 and 51 had domains that gave them the functions of integrase and resolvase respectively (Phage_integrase and SR_ResInv). Orf37 appeared to be a CI phage repressor based on HTH_XRE domain present. All genes had a high degree of similarity either to *C. difficile* strains or to Φ38-2. The integrase gene facilitates integration of the phage genome into the host chromosome via site specific recombination between the phage attachment site, attP and bacterial host site, attB (Williams *et al.* 2013). During lysogeny, the CI phage repressor supresses expression of the viral genes by binding to the promoter regions (Ryoki *et al.* 2001). The resolvase can, once induced, excise the viral genome from the bacterial chromosome.

Previous studies have shown that phages can carry a CRISPR/Cas system that is used to counteract the phage inhibitory chromosomal island of the bacterial host (K. Hargreaves *et al.* 2014). The CRISPR finder tool was thus used to detect whether or not phage ST has any CRISPR arrays. The analysis showed that it has one sequence with CRISPR-like features comprised of a single spacer. The short length of the sequence and absence of any Cas like genes indicate that no functional CRISPR arrays are present.
Alignment of phages Φ38-2 and ΦST in Geneious. The two phages have a 95.1% pairwise identity. The two red squares show the two main regions where the two genomes have 0% pairwise identity. These correspond to the minor capsid protein and tail fibre protein in phage ΦST.
Table 2: Genome annotation of ΦST
CD at top Blast Hit denotes presence across many *C. difficile* strains. Otherwise homologous phage or *C. difficile* strains are shown

<table>
<thead>
<tr>
<th>ORF</th>
<th>start</th>
<th>end</th>
<th>AA size</th>
<th>Strand</th>
<th>Product Name</th>
<th>Domains</th>
<th>Top Blastp Hit</th>
<th>identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>125</td>
<td>736</td>
<td>203</td>
<td>+</td>
<td>terminase small subunit</td>
<td>Phage_terminase ,YjcR</td>
<td>Φ111</td>
<td>99</td>
</tr>
<tr>
<td>2</td>
<td>723</td>
<td>2045</td>
<td>440</td>
<td>+</td>
<td>terminase large subunit</td>
<td>phage_term_2,Terminase_3,Terminase_6</td>
<td>Φ111</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>2046</td>
<td>3458</td>
<td>470</td>
<td>+</td>
<td>portal protein</td>
<td>A118_put_portal,Phage_prot_Gp6</td>
<td>Φ111</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>3459</td>
<td>4682</td>
<td>407</td>
<td>+</td>
<td>minor capsid protein</td>
<td>Phage_min_cap2</td>
<td>Φ111</td>
<td>98</td>
</tr>
<tr>
<td>5</td>
<td>4684</td>
<td>5010</td>
<td>108</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>Phage_min_cap2</td>
<td>Φ111</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>5026</td>
<td>5307</td>
<td>93</td>
<td>+</td>
<td>hypothetical protein</td>
<td>Phage_min_cap2</td>
<td>Φ111</td>
<td>94</td>
</tr>
<tr>
<td>7</td>
<td>5439</td>
<td>6011</td>
<td>190</td>
<td>+</td>
<td>scaffolding protein</td>
<td>Phage_GP20</td>
<td>Φ111</td>
<td>98</td>
</tr>
<tr>
<td>8</td>
<td>6022</td>
<td>7095</td>
<td>357</td>
<td>+</td>
<td>Major capsid proteins</td>
<td>CD</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>7099</td>
<td>7431</td>
<td>110</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>CD</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7433</td>
<td>7867</td>
<td>144</td>
<td>+</td>
<td>hypothetical protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>7874</td>
<td>8221</td>
<td>115</td>
<td>+</td>
<td>phage capsid family protein</td>
<td>CD</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8218</td>
<td>8649</td>
<td>143</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>CD</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>8652</td>
<td>9452</td>
<td>256</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>CD</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>9510</td>
<td>9872</td>
<td>120</td>
<td>+</td>
<td>AP endonuclease family 2/con</td>
<td>CD</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>9872</td>
<td>10480</td>
<td>202</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>Phage_Gp15</td>
<td>Φ146</td>
<td>91</td>
</tr>
<tr>
<td>16</td>
<td>10452</td>
<td>10562</td>
<td>36</td>
<td>+</td>
<td>hypothetical protein</td>
<td>Phage_Gp15</td>
<td>Φ146</td>
<td>91</td>
</tr>
<tr>
<td>17</td>
<td>10552</td>
<td>11193</td>
<td>213</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>DUF4236</td>
<td>Φ111</td>
<td>100</td>
</tr>
<tr>
<td>#</td>
<td>Start</td>
<td>End</td>
<td>Size</td>
<td>Description</td>
<td>Additional Information</td>
<td>Source</td>
<td>Percentage</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>11299</td>
<td>16602</td>
<td>5334</td>
<td>tail tape measure</td>
<td>PhageMin_Tail,tape_meas_TP901,PKK</td>
<td>Φ111</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>16702</td>
<td>17457</td>
<td>756</td>
<td>tail component protein</td>
<td>YomH,Sipho_tail</td>
<td>Φ38-2</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>17787</td>
<td>19586</td>
<td>1818</td>
<td>Tail protein</td>
<td>PblB,put_anti_recept.Prophage_tail</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>19613</td>
<td>19856</td>
<td>243</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>20565</td>
<td>22031</td>
<td>1466</td>
<td>tail fiber protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>22031</td>
<td>23881</td>
<td>1447</td>
<td>tail fiber protein</td>
<td>PHA02515</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>23967</td>
<td>24230</td>
<td>264</td>
<td>holin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>24231</td>
<td>25037</td>
<td>804</td>
<td>endolysine</td>
<td>MurNAc-LAA,Amidase_3,Ami_3,spore_cwlD,AmiC</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>25086</td>
<td>25298</td>
<td>212</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>25477</td>
<td>26307</td>
<td>830</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>26386</td>
<td>26892</td>
<td>506</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>26912</td>
<td>27382</td>
<td>469</td>
<td>conserved protein of unknown function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>27615</td>
<td>27896</td>
<td>781</td>
<td>conserved protein of unknown function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>28040</td>
<td>28372</td>
<td>333</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>28337</td>
<td>29182</td>
<td>845</td>
<td>putative para</td>
<td>ParA,AAA_31,BcsQ,CbiA,PRK13705,PHA02519,</td>
<td>Φ111</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>29471</td>
<td>29653</td>
<td>182</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>30474</td>
<td>30626</td>
<td>152</td>
<td>hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>31125</td>
<td>33311</td>
<td>728</td>
<td>+</td>
<td>DNAse helicase</td>
<td>DEXDc, DEAD, DEXDc, SSL2, uvsW</td>
<td>CDP6</td>
<td>100</td>
</tr>
<tr>
<td>36</td>
<td>33358</td>
<td>33705</td>
<td>115</td>
<td>+</td>
<td>hypothetical protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>34519</td>
<td>34872</td>
<td>117</td>
<td>-</td>
<td>Cl phage repressor</td>
<td>HTH_XRE, HTH_31, HTH_XRE, XRE, HipB, PRK13355</td>
<td>CD</td>
<td>100</td>
</tr>
<tr>
<td>38</td>
<td>35045</td>
<td>35227</td>
<td>60</td>
<td>+</td>
<td>hypothetical protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>35294</td>
<td>35719</td>
<td>141</td>
<td>+</td>
<td>conserved protein of unknown function</td>
<td>CD</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>35794</td>
<td>36198</td>
<td>134</td>
<td>+</td>
<td>replication terminator protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>36221</td>
<td>36934</td>
<td>237</td>
<td>+</td>
<td>hypothetical protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>37011</td>
<td>37154</td>
<td>47</td>
<td>+</td>
<td>hypothetical protein</td>
<td>Φ38-2</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>37151</td>
<td>37462</td>
<td>103</td>
<td>+</td>
<td>hypothetical protein</td>
<td>Stk19</td>
<td>CD</td>
<td>100</td>
</tr>
<tr>
<td>44</td>
<td>37508</td>
<td>37657</td>
<td>49</td>
<td>+</td>
<td>hypothetical protein</td>
<td>Φ38-2</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>37688</td>
<td>37861</td>
<td>57</td>
<td>+</td>
<td>hypothetical protein</td>
<td>DUF3797</td>
<td>CD</td>
<td>100</td>
</tr>
<tr>
<td>46</td>
<td>37974</td>
<td>38300</td>
<td>108</td>
<td>+</td>
<td>ssDNA binding protein</td>
<td>SSB_OBF, SSB, ssb, PRK06751, Ssb</td>
<td>Φ38-2</td>
<td>100</td>
</tr>
<tr>
<td>47</td>
<td>38392</td>
<td>38757</td>
<td>121</td>
<td>+</td>
<td>hypothetical protein</td>
<td>CD</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>38741</td>
<td>38914</td>
<td>57</td>
<td>+</td>
<td>hypothetical protein ΦCD38-2</td>
<td>Φ38-2</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>38926</td>
<td>39498</td>
<td>190</td>
<td>+</td>
<td>sigma70/sigmaF-like protein</td>
<td>Sigma70_r4, Sigma70_r4_2, sigma70-ECF, RpoE</td>
<td>CD</td>
<td>100</td>
</tr>
<tr>
<td>50</td>
<td>39676</td>
<td>40284</td>
<td>202</td>
<td>+</td>
<td>integrase</td>
<td>INT_C_like_3, Phage_integrase, PRK09870, recomb_XerC, xerC, XerD</td>
<td>CD</td>
<td>98</td>
</tr>
<tr>
<td>51</td>
<td>40722</td>
<td>41318</td>
<td>198</td>
<td>+</td>
<td>resolvase</td>
<td>SR_ResInv, Resolvase, HTH_Hin_like, HTH_7, mpi, PinE</td>
<td>CD</td>
<td>99</td>
</tr>
</tbody>
</table>
Figure 25: Genomic map of ΦST

Double stranded Circular genomic map of ΦST with (starting with inner ring) GC skew, GC content and ORFs with putative functions. Ring 1 GC Skew: red is above average and blue below average. Ring 2 GC content: green is above average and black is below average. Ring 3 (antisense) and Ring 4 (sense) depict ORFs. Different colours depict different stages in phage life cycle: DNA packaging (red); head morphogenesis (dark green); tail morphogenesis (purple); lysis (blue); lysogenic conversion (grey); DNA replication (orange), and lysogeny control (light green). Aqua shows hypothetical proteins and yellow conserved proteins of unknown function. Black shows a hypothetical protein with no hits (made using DNA plotter).
3.3.0 Discussion

*Clostridium difficile* is a nosocomial pathogen that has in recent years become a major health concern. This is due to the difficulty of treating it as a result of the multiple antibiotic resistances it has acquired. The fact that it is indirectly caused by antibiotic treatment further exasperates the problem. An alternative therapeutic is therefore required.

Bacteriophages have several advantages over antibiotics which make them ideal candidates for such an alternative. Firstly, bacteriophages are the natural predators of bacteria and evolve along with them while antibiotics cannot adapt to evolving bacteria. Secondly phages can multiply at the site of infection increasing their efficacy while antibiotics lose their efficacy over time. Thirdly replacing a phage is less complicated and time-consuming than replacing an antibiotic. Lastly a bacteriophage only kills the bacterial strain it targets and not any other bacteria belonging to the local microbiota.

By screening about 30 samples of soil and water, eight bacteriophages were isolated and their morphologies were determined by TEM and PCR. All of them belonged to Caudovirales family of non-enveloped, tailed phage with an icosahedral head. Out of the eight, six were long tailed (contractile) myoviruses and two were long-tailed (non-contractile) siphoviruses. All of them had an approximate length of between 200-250nm with a 50nm wide head. All of them were identified by the enrichment of unfiltered soil samples while the enrichment filtered soil samples using of specific using specific hosts did not lead to the isolation of any phages. This lead to the conclusion that all the phages isolated were temperate phages. Despite of this however, all the phages isolated produced lytic (clear) small or medium sized plaques.

Host range analysis showed that with the exception of one phage (ΦST) , the phages were capable of infecting only a few nosocomial strains of clinically relevant ribotypes of *C. difficile* in addition to the environmental strain T6 (ribotype 076). On the other hand, phage ST was found to be capable of infecting 23/23 strains of the clinically relevant ribotype 027 and one strain of ribotype 087, all of which were isolated from patients with CDI. In addition
strains of these ribotypes have previously been associated with epidemics of CDI and are prevalent in UK hospitals (Arvand et al. 2009; Health Protection Agency 2012). Plaques efficiency assays showed that phage ST was capable of infecting all bacterial strains to the same degree. These two extreme examples of host range are not surprising. Previous studies have shown that it is possible for phages to have either a broad host range or to be only capable of infecting a few or even one strain within a single species. However, the latter example is not desirable when one needs to apply phages as a therapeutic where phages with broad host ranges are required. Therefore a number of phages were selected to undergo an attempt to broaden their host range.

This was done in a number of ways. First was the exposure of individually picked plaques to the resistant host in solid media. Second was the exposure of a high titre of phage to a resistant host in liquid media. Third was the exposure of phage to a co-culture of both the resistant and propagating host over a period of 24 hours. Lastly the phage was passaged through a co-culture of both the resistant and propagating host for 10 days every 24 hours. These assays were repeated multiple times. However there was only one occasion observed where a phage acquired the ability to infect a host it did not before; phage ST acquired the ability to infect one strain of ribotype 078 that it did not before. This inability to expand the host range could be attributed to the individual phages rather than the techniques used as these have been found to work by both previous published studies as well as from other members of our research group.

Since bacteriophages need to be applied as a therapeutic, they need to be able to effectively lyse their bacterial host. However bacteria can evolve resistances to phages. Therefore the ability of individual vs multiple phages to lyse *C. difficile* in vitro was tested. When used at an MOI of 10 all individual phages showed the ability to reduce the CFU numbers in the culture by at least 2-4 logarithmic degrees. When pairs of phages were used an even greater reduction was shown (~2-6 logs). The use of combinations of three and four phages did not cause a larger degree of reduction and in some cases showed to be counterproductive in relation to the use of pairs of bacteriophages. When used at an MOI of 100, the individually used phages delivered a bigger
reduction in CFU than when an MOI 10 was used. Like before, the paired bacteriophages delivered an even bigger reduction than both the individually used phages and the combination of three phages (~7 logs). However when the combination of four phages was used it reduced the CFU count to undetectable levels during the course of the assay. In all cases apart from the MOI 100-four bacteriophage combination the infected culture was able to recover and grow back. However even though the vegetative cells appeared to have been wiped out by the four phage combination, endospores were able to be germinated from the culture. This result indicates one of the difficulties of applying phage therapy for *C. difficile*. The hardiness of endospores is something that needs to be taken into account when applying phages as a therapeutic.

Even though none of the phages appeared to play a significantly more important role in the cell killing, the myoviruses ΦT9 and ΦT15 appeared to be slightly more effective than the siphoviruses ΦX2 and ΦST when used individually. In addition, when used in combination, the myoviruses appeared to deliver slightly higher killing than siphoviruses. The main determinant factor in the efficiency of killing appeared to be the MOI. It is possible that a great part of the CFU reduction when an MOI of 100 was used is lysis from without due to the sheer number of virions used. The use of multiple and different morphologies of phages would have accounted for the emergence of resistant strains. It should be considered however that these results seen *in vitro* are unlikely to be an accurate representation of what would happen *in vivo*. Given the fact that the conditions of the experiment are the optimal for the growth of the host, it is possible that it fairs *in vitro* much better than it would *in vivo*. It would be therefore appropriate to apply this experiment in an *in vivo* model to see exactly hoe the dynamics of phage–host interaction work there.

Out of the eight phages isolated, ΦST had the broadest host range and the better growth dynamics. It was therefore sequenced and annotated to gain more information about it. A total of 51 orfs were identified, 22 of which were assigned a putative function. The rest were either hypothetical proteins or conserved proteins of unknown function. The phage was found to have a high degree of similarity with previously sequenced *C. difficile* phages. The closest related phage was Φ38-2 with a 95.1% nucleotide identity which is another
siphovirus with the ability to infect strains of mainly ribotype 027. The genome shows a high degree of mosaicism, with the amino acids sequence of different genes have homologues in different proteins. The genome follows modular arrangement with genes corresponding to different stages of the virion morphogenesis clustered together. There is a cluster of hypothetical proteins and conserved proteins of unknown function downstream the lysis genes on the antisense strand that are believed to be associated with lysogenic conversion. These genes have in the past been associated with increased host cell virulence when the phage exists in the host as a prophage (Fortier & Sekulovic 2013). This, along with the fact that an integrase gene is present indicating that the phage has the ability to lysogenize, can be potentially problematic if the phage is to be applied as a therapeutic.

3.3.1 Future Work

Even though this project was successful in isolating and characterising eight phages, one of which was sequenced more will be needed for the development of an appropriate therapeutic. A future project should attempt to isolate more phages with an, ideally, good host range. Also, the host range expansion experiments should be repeated with any newly isolated phages. The killing assays should also be expanded and tested on different ribotypes and phage combinations.
Chapter 4: Therapeutic applications of *C. difficile* phages in an *in vivo* model
4.0.0 Introduction

In order for a phage to be approved for clinical trials in humans as a therapeutic against Clostridium difficile infection (CDI), it must first be tested in animal trials (Ryan et al. 2011; Vandersteegen et al. 2013; Paul et al. 2011; Lu & Collins 2009). These trials must demonstrate that phages, like any therapeutic, are able to eradicate or significantly reduce the bacterial load from a host without causing any harm to the host itself.

The golden Syrian hamster (Mesocricetus auratus) is currently acknowledged as the best small animal model to study CDI as, the animals develop many of the clinical symptoms such as diarrhoea, histological damage and relapse of the condition following removal of treatment (Buckley et al. 2013). Previous studies have demonstrated that when challenged with a toxigenic strain of *C. difficile* after administration of broad spectrum antibiotics that disrupt the colonic flora, the animals are colonised by the bacterium and develop the clinical symptoms of the disease that eventually lead to death (Buckley et al. 2011; Goulding et al. 2009). In contrast, infection in the mouse model results in a much milder disease in which the mice become transiently colonised but show limited pathology (Hutton et al. 2014). Although there have been several studies investigating the development of disease and the use of antibiotics to treat it in the *in vivo* hamster model, there are currently only two studies to date that have investigated the treatment of CDI in hamsters with phages. In Ramesh et al 1999, a study was carried out where phages were used to treat hamsters infected with *C. difficile*. In the experiment, the animals were challenged with $1 \times 10^3$ *C. difficile* spores following administration of Clindamycin. All phage treated animals were given a phage dose of $1 \times 10^8$ PFU after inoculation with spores while the control animals received an equal amount of the phage diluant. Phage treated animals also received a phage dose of $1 \times 10^8$ PFU per for every 8 hours after initial spore challenge. In that experiment the control animals reached the clinical endpoint between 72-96 hours post *C. difficile* challenge while the majority of the treated animals survived. When re-challenged however all the animals that previously survived died in the absence of phage treatment (Ramesh et al. 1999).

The same research group published another study in 2011 where, *C. difficile* phage ΦCD119 was shown to be able to lysogenize *in vitro* (Govind et
al. 2012). Similar to the experiment described before, phage treatment in hamsters challenged with \textit{C. difficile} caused an extension in their survival compared to the untreated animals. However in this case all animals eventually succumbed to the infection (Govind et al. 2012). The recovered \textit{C. difficile} from the faecal and Caecum samples of the phage treated animals showed increased resistance to phage ΦCD119. The control animals did not show any such increase in resistance. When tested, some of these resistant \textit{C. difficile} cells were shown to have phage ΦCD119 inserted within their genome. These lysogens were showed to have a reduced toxin production compared to the WT strain both \textit{in vitro} and \textit{in vivo} (Govind et al. 2012).

The research group of Dr Gillian Douce at the University of Glasgow has a working hamster model, with which they have used to study the pathology of CDI with several well characterised strains. The colonisation of the animal has been studied with the CD1342 (A- B- CTD-) strain (Buckley et al. 2013). This strain does not cause CDI because it lacks the three toxins linked with the pathogenesis of the disease, but it is still able to colonise the colon and caecum of the animal model following treatment with clindamycin. Among other strains, CDB1 (A+ B+ CTD +) has been used to study the pathology of the disease (Buckley et al. 2011). This strain is able to cause CDI along with many of the associated symptoms like diarrhoea, lack of activity and elevated temperature. The CDI caused by the strain will cause death in an untreated animal within 33 hours after \textit{C. difficile} spore challenge (Goulding et al. 2009).

In Professor Martha Clokie’s laboratory, a number of phages that are capable of lysing \textit{C. difficile} cells \textit{in vitro} have been identified. In order to test the potential of some of these phages as a therapeutic \textit{in vivo}, a collaboration was formed with Dr Gillian Douce in the University of Glasgow. It was decided that the therapeutic potential our phages would be tested against strains CD 1342 and CDB1 in a manner similar to the one described by Ramesh \textit{et al.} Strain CD1342 (A- B- CTD-) was chosen because it is toxin negative and therefore the infected animal will not develop CDI, allowing to test the effects of phage treatment on the colonisation of \textit{C. difficile}. Strain B1 (A+ B+ CTD+) was chosen because it was previously shown that infection with this strain in an animal will result in CDI along with many of the associated symptoms like diarrhoea, lack of activity and elevated temperature. The CDI caused by the
strain will cause death in an untreated animal within 33 hours after *C. difficile* spore challenge (Goulding et al. 2009).

### 4.1.0 Aims and objectives

In this study, I sought to find phages able to: a) prevent or inhibit the colonisation of the colon of the hamster model with *C. difficile* and b) prevent the development of CDI in the *in vivo* model and by extension prevent or reduce morbidity and mortality. This was done by using the phages from the phage library in my laboratory to identify those able to infect strains CD B1 and CD 1342. The dynamics of those phages would then be characterised using the two bacterial strains by means of spot tests, plaque assays and one-step growth curves. Those phages that were deemed promising were increased in titre up and purified to investigate the effects of phage treatment to prevent colonisation using strain CD1342 and the effects on morbidity and mortality using *C. difficile* strain CDB1 in an *in vivo* model.
4.2.0 Results

4.2.1 Identifying phages that infect strains CD1342 and CDB1

In order to identify phages able to infect strains CD1342 and CDB1, a collection of 20 phages was used to infect the two bacterial strains using spot tests. I was able to identify five phages able to infect the virulent strain CDB1 and one phage able to infect the avirulent strain CD1342 (Table 3). The myovirus ΦCDHM1, whose host range was previously found to include strains of ribotype 014/020, was able to infect both bacterial strains. Spot tests showed that ΦCDHM1 is able to form plaques, and thus replicate, on strain CDB1 at a reduced efficiency compared to its indicative strain. ΦCDHM1 was able to lyse lawns of strain CD1342 but not form plaques indicating that no replication occurs, suggesting that cell death is due to lysis from without. Due to the issue of not being able to replicate the selected phage on strain CD1342 it was decided that the standard manufacturing host of the phage, CDT6 (R076), would be used as the indicator strain for PFU enumeration (Fig 27).

One-step growth curves for the selected phage against both CDB1 and CD1342 were carried out. In strain CDB1 there was an inhibition in the number of CFU/ml from the first hour of phage introduction. At the third hour there was a steep drop in CFU which coincided with a steep increase in PFU count which indicated the occurrence of a phage burst. This was followed by a recovery of the CFU count at the next point. The CFU count of the infected culture remained lower than the control through the assay. In strain CD1342, the infected cultured was inhibited from the first hour of infection too, but unlike strain CDB1 the drop in the CFU count was less steep. In addition there were no phages detected in the culture through the course of the assay. This provides further evidence that CDHM1 is not able to replicate in this strain. As with the spot tests previously mentioned, the effects of the inhibition appear to be due to lysis from without (Fig 28).
Table 3: Five phages are able to lyse CDB1 and one phage able to lyse CD1342

Using the spot test and plaque assay methods, our 22 phages were tested against lawns of strains CDB1 and CD1342. Five phages were shown to be able to form plaques on strain CDB1. On the other hand only one phage was shown to be able to lyse strain CD1342. That phage -ΦCDHM1- was also able to lyse strain B1. However while phage CDHM1 was able to form plaques on CD B1, it was unable to do so on CD1342. This would suggest that ΦCDHM1 is unable to replicate in CD1342 and the killing of cells is due to lysis from without.

<table>
<thead>
<tr>
<th>Phage</th>
<th>B1</th>
<th>CD1342</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1/φS11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1/φS8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1/φ16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φCDHM1</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>10/φB2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S7/φ10</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>T6/φ7H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10b/φHb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φ12eh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/φ5bBH</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>S7/φ3</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>S7/φ7</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>10/φaH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIU/φX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATJ/φH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATJ/φF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φ4N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φ7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6/φ16C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 26: Plaquing efficiency of ΦCDHM1 on bacterial strains CD1342, CDB1 and CDT6

Phage ΦCDHM1 was via spot test and plaque assays against strains CD1342 and CDB1. The phage does not have plaquing efficiency on strain CD 1342 as it cannot replicate on it. When CDHM1 was tested against CDB1, its titre found to be ~2logs lower than when the same stock was tested against the propagating host T6. This means that phage CDHM1 has 1% plaquing efficiency compared to its host (SEM, n=2).
Figure 27: One-step growth curve for ΦCDHM1 against CD B1 and CD 1342

Strains CD B1 (A) and CD1342 (B) were both infected with phage CDHM1 at early log phase at an MOI of 10. Because phage CDHM1 was found not able to be propagated on strain CD1342, C. difficile strain CDT6 was used as a reference for both strains. In both cases, the infected culture (red line) showed inhibited growth in relation to the uninfected control (blue line). In CD B1 there was inhibition of the CFU count from the first hour of infection which culminated in a steep drop at the 3rd hour of infection. At the same time there was a rise in the number of the PFU count (green line), indicating the occurrence of a phage burst. In CD1342 there was inhibition of the infected culture but there was no steep drop in CFU count like CD B1. Furthermore there was no visible phage burst observed in the culture indicating that no phage replication occurs in CD1342 (SEM, n=2).
4.2.2 Obtaining a high titre and purification of the selected phages for *in vivo* experiments

Having identified a phage able to lyse strains CD1342 and CDB1, the next step was to make it suitable for *in vivo* administration. After obtaining high titres of phage CDHM1 on solid media, these stocks were then purified by Caecium chloride step gradient, by PEG concentration and by centrifugation and resuspension. The toxin content of these purified method suspensions was then measures by exposing them to two cell culture cell lines: H29 cells, which are more sensitive to the presence of toxin A and Vero cells which are more sensitive to the presence of toxin B (Kuehne et al. 2010). When comparing tissue cell rounding, CsCl purified phage did not cause any visible rounding while centrifuged and resuspended phage caused rounding at a concentration that was a thousand times less concentrated in Vero cells and 625 times less concentrated in H29 cells than the gut washes from an animal that died from CDI. PEG concentrated phage was found to have 10 times more toxin content than the crude lysate it was concentrated from (Table 4). When testing the titre of these suspensions, Caesium Chloride step gradient resulted to a loss of about 99% of the initial phage titre. PEG concertation and centrifugation and suspension of the phage stock resulted at a loss of 30% of the initial phage titre. Given these data, it was decided to use the centrifuged and resuspended phage titre because of the reduced toxin content with minimum reduction in phage titre.
Table 4: Maximum dilution in which toxins were detectable in different phage preparations

The highest dilution to which cell toxicity (cell rounding) was observed per phage preparation is recorded here. The lowest toxicity was observed in CsCl ultracentrifugation where no cell rounding was detected. The highest toxicity was observed in the PEG-concentrated preparation, presumably due to the concentrations of cell toxins along with the phage.

<table>
<thead>
<tr>
<th>Phage Preparation</th>
<th>Vero cells</th>
<th>H29 cells</th>
<th>Titre after purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude Lysate</td>
<td>100-fold</td>
<td>25-fold</td>
<td>100%</td>
</tr>
<tr>
<td>PEG-Concentrated</td>
<td>1000-fold</td>
<td>250-fold</td>
<td>70%</td>
</tr>
<tr>
<td>Washed Crude lysate</td>
<td>10-fold</td>
<td>5-fold</td>
<td>70%</td>
</tr>
<tr>
<td>CsCl ultracentrifugation</td>
<td>No rounding</td>
<td>No rounding</td>
<td>1%</td>
</tr>
<tr>
<td>Gut washes (control)</td>
<td>10000-fold</td>
<td>3125-fold</td>
<td>n/a</td>
</tr>
</tbody>
</table>

4.2.3 Phage treatment can significantly reduce *C. difficile* colonisation in an *in vivo* model

Having identified, obtained a high titre and purified one phage able to lyse CD 1342, the next step was to carry out the *in vivo* experiment to investigate the effects of phage treatment on colonisation of *C. difficile*. Two groups of hamsters were used. Each group consisted of 5 hamsters, 3 of which were treated with phage and 2 were used as controls.

During the course of the experiment, the hamsters did not exhibit any symptoms associated with CDI such as the development of diarrhoea, fluctuations in body temperature and overall reduced activity. Since strain CD 1342 lacks the pathogenicity locus (Pac Loc), this is not unexpected as the disease is toxin mediated. This would also indicate that the minute amounts of toxins in the phage lysate administered did not have any adverse effect on the animals. Upon completion of the 24 hour experimental period, washes of the contents of the caecum and colon were used to enumerate the spore and vegetative cell count. The recovery of total CFU and spore count was determined in both the lumen and the tissue of the colon and caecum.

Overall there were more cells recovered from the lumen of the organs than the tissue but with no significant difference (t-test, p>0.05). Both groups of
animals, the phage treated animals had a reduced amount of cells recovered in relation to the untreated controls. The total recovery of bacteria (vegetative cells and spores) per organ in the controls was in the range of $10^6$ and $10^8$ CFU/ml, depending on the organ, with spores making about 10-20% of the total recovered. From the treated animals, Hamster 3 in group 1 had the largest reduction of about 5 logs with no spores detected while hamster 5 in group 2 had the lowest with very little or no reduction compared to the controls (Fig 28 and Fig 29). Statistical analysis of the combined data of both groups indicates that this reduction is statistically significant (t-test, two-tailed, p value < 0.05). The treated colon and caecum are had significantly less total CFU count recovered compared to their untreated counterparts (p value 0.0148 and 0.0387 respectively). Similarly the spore count was also significantly lower for the treated colon and caecum (p value 0.0114 and 0.0405, respectively) (Fig 30). Phages were recovered in high titres from the lumen and tissue of colon and caecum of the phage treated animals. A higher amount of phages were isolated from the lumen of the colon and caecum than the tissue but the difference was not statistically significant (Fig 31) (t-test, two-tailed, p value > 0.05). Phages were also recovered from the stools of the phage treated animals at an average of $10^6$ PFU/ml (data not shown).
After the end of the incubation period, all the animals were sacrificed, their colon and caecum removed aseptically and washed to recover the vegetative cells and spores. The blue bars represent the total CFU count (spores and vegetative cells) and red bars represent the spore count. In both the caecum and the colon, the phage treated animals showed a significant reduction in both the overall CFU count and spore count recovered in comparison to the untreated animals. In both group 1 (Fig 28) and group 2 (Fig 29), phage treatment caused a reduction in the overall CFU and spore count compared to the untreated controls. There was however a degree of variance in this reduction among the treated animals, with animal H3 from group 1 having the greater reduction and animal H5 from group 2 being very similar to the controls.

Figure 28: Counts of Vegetative cells and spores recovered from hamsters in group 1
After the end of the incubation period, all the animals were sacrificed, their colon and caecum removed aseptically and washed to recover the vegetative cells and spores. The blue bars represent the total CFU count (spores and vegetative cells) and red bars represent the spore count. In both the caecum and the colon, the phage treated animals showed a significant reduction in both the overall CFU count and spore count recovered in comparison to the untreated animals. In both group 1 (Fig 28) and group 2 (Fig 29), phage treatment caused a reduction in the overall CFU and spore count compared to the untreated controls. There was however a degree of variance in this reduction among the treated animals, with animal H3 from group 1 having the greater reduction and animal H5 from group 2 being very similar to the controls.

Figure 29: Counts of Vegetative cells and spores recovered from hamsters in group 2
Figure 30: Phage treatment reduces *C. difficile* colonisation in an *in vivo* model

The combined data from both groups indicates that phage treatment causes a significant reduction in both the total CFU count (blue bars) and in the amount of sporulation (red bars). There was significantly less total CFU as well as spores recovered from the treated colon compared to the untreated one (p values 0.0148 and 0.0114 respectively). That was the case with the colon as well with p values 0.0387 and 0.0405 respectively (unpaired t-test p<0.5, SEM, n=3).
Figure 31: High titres of phages were recovered from Caecum and Colon

The washes from the Caecum and Colon were filtered and screen for phage by using strain T6 to form a lawn. The presence of phage CDHM1 was detected in both the tissue and lumen of the two organs. More phage was isolated from the lumen rather than the tissue (SEM, n=2).
4.2.4 Phage treatment on *C. difficile* can significantly delay the onset of morbidity and mortality in an *in vivo* model

Having obtained a high titre and purified one phage able to lyse CDB1, the next step was to carry out the *in vivo* experiment to investigate the effects of phage treatment on CDI. One group consisting of 5 hamsters, 3 of which were treated with phage and 2 were used as controls, was used in this experiment.

During the course of the experiments, the animals were monitored for signs of the disease such as development of diarrhoea, fluctuations in body temperature and lack of activity. The treated animals continued to receive a dose of phage every eight hours post challenge, while the control animals received an equal volume of the phage dilu tant. About 25 hours after challenge, the control animals, began showing signs of infection, while the treated animals showed no such signs. The body temperature of the control animals dropped below the clinical endpoint of 35ºC 4-5 hours after the development of symptoms and had to be culled. During that period both animals experienced one episode of diarrhoea. The treated animals continued to be monitored and receive phage doses. They began showing signs of infection between 42-44 hours after challenge. Hamster 1 and Hamster 3 body temperature dropped below the clinical endpoint and had to be culled 47 and 51 hours after challenge respectively. Hamster 2 which was culled at 57 hours due to the loss of body condition and dehydration related to the length of the severe diarrhoeal episode (Fig 33). The colon and caecum of all animals was aseptically extracted and CFU count of vegetative cells and spores carried out as previously described. Comparison between the counts in control and treated animals did not show any significant difference (Appendix 8).

The toxin levels in both control and treated animals were similar. The control animals seem to have slightly higher levels of toxin B in the caecum but due to the small number of animals it would be difficult to statistically confirm (Fig 34). Statistical analysis on the survival of both groups showed that the treated animals lived significantly longer than the controls (log-rank Mantel-Cox test p value = 0.0389). Phage was recovered from the colon and caecum of all phage treated animals at an average of $10^6$ PFU/ml (Fig 36). It was not possible to enumerate the PFU/ml in the stools of either the control or phage treated
animals as it was not possible to collect stools because of the acute diarrhoea.

4.2.5 Determination on the presence of the phage CDHM1 in the colon and caecum washes of phage treated animals

A PCR was conducted in order to determine whether the phage found in the colon and caecum washes of phage treated animals was ΦCDHM1. The primers used in this reaction targeted the sequence of the capsid of ΦCDHM1 (Hargreaves 2012). The size of the product obtained in all samples tested was consistent with the size expected with the primers used, indicating that the phage isolated was indeed ΦCDHM1 (Fig 32)(Hargreaves 2012).

Figure 32: Phage ΦCDHM1 is present in the faeces and caecum and colon of treated animals

The PCR using primers targeting the capsid of CDHM1 indicates the presence of the phages in the faeces and cecum and colon washes of phage treated animals (bands1-8 starting from the left). Bands 9-11 show the presence of the phage as a prophage in bacteria isolated from the treated animals.
Figure 33: Phage treatment causes delay in the onset of morbidity and mortality

a) All animals experienced one episode of diarrhoea, whose onset was accompanied by elevated body temperature followed by a drop that eventually fell below 35°C, which was the clinical endpoint with the exception of Hamster 2 (treated). In the phage treated animals (H1, H2, H3) this diarrhoea episode was delayed by 15-17 hours compared to the untreated animals (H4 and H5). b) All animals eventually developed symptoms of CDI and had to be sacrificed. However phage treated animals (squares) did not have to be culled until 17-21 hours after the untreated animals (circles) due to the absence of symptoms relating with the disease. Statistical analysis comparing the survival of untreated vs treated animals gave a p value of 0.0389 showing that the latter survived significantly more than the former (log-rank Mantel-Cox, p<0.5)
Figure 34: Effect of phage treatment on toxin levels in an *in vivo* model

The reciprocal toxin levels of the gut washes from the colon and caecum of both phage treated and untreated animals were measured as previously described. The results seem to indicate that the levels of toxins in both groups of animals are similar, which not surprising given the fact that all animals eventually succumb to infection. The control animals seem to have a higher level of toxin B in the caecum but as there are very small numbers of animals it would be difficult to confirm statistically.

Figure 35: The effects of toxins A and toxins B on a hamster colon

These images show the effects of the exotoxins A and B on an infected upper colon (right) from an animal that died from CDI compared to the colon of a healthy animal (left) (x 100 Magnification)
Figure 36: Phage recovered from the phage treated hamsters

The contents of the caecum and colon were washed with SM buffer and filtered. Plaque assays showed the presence of phage CDHM1 at a titre of $10^5$-$10^6$ PFU/ml in both the tissue (TA) and Lumen (LA) of Caecum and Colon (SEM, n=2).
4.3.0 Discussion

The Syrian golden hamster is recognised as the best small animal model for mimicking CDI (Buckley et al. 2013). There have been many studies citing the animal as an in vivo model of the disease. The many of those studies investigated the effect of antibiotic treatment in animals infected clostridium difficile. However there have been to date only two published papers investigating the effects of phage treatment on CID in vivo.

Ramesh et al in their 1999 paper, showed that one dose of phage treatment was sufficient to prevent the onset of disease in the in vivo model. The phage treated animals did not develop symptoms and did not succumb to the infection unlike the untreated controls which both developed symptoms and died within 72 hours after challenge with the C. difficile spores (Ramesh et al. 1999).

In this set of experiments, a relatively small number of animals was used in order to first establish a proof of concept on the effectiveness of the phage as a therapeutic before proceeding at a later stage with larger groups of animals. In addition, the experiment which infected animals with a virulent strain was carried out later in order to first demonstrate that phage infection has the possibility of clearing the disease in order to avoid putting the animals through unnecessary suffering.

In our experiment, we sought to assess if phage treatment is sufficient to completely clear an infected animal and prevent colonisation. We used a total of a 10 animals all of which were challenged with C. difficile spores. Six of those animals received 3 doses of bacteriophages and four animals were used as controls. At the end of the experiment, all the animals were culled, their colon and caecum was aseptically removed and C. difficile cells were recovered by washing and homogenisation of the organs. Colony counts showed that there were statistically significant less amount of cells (both spores and vegetative) recovered from the phage treated animals compared to the untreated controls. This indicates that phage ΦCDHM1 was able to effectively lyse the C. difficile cells in the colon and caecum of the infected animals. It has already been shown that phages are unable to be infect or adsorb to spores and that the phage ΦCDHM1 is unable to replicate on strain CD1342. This would mean that this reduction in CFU counts was due to the effects of lysis from without on
vegetative cells. Which leads to the question: If phages are only able to adsorb and therefore lyse vegetative cells why is there a reduction to the number of spores as well? Normally the phages would clear the vegetative cells while leaving the spores unharmed. The spores then continue to germinate as they do, the new vegetative cells are killed by the phages present and the cycle continues. This would also explain the large amounts of spores present in the control animals. Many cells would change to the spore stage in the case of heavy colonisation of the infected organ to prevent nutrient depravation.

In the next step of the experiment, we investigated the effects of phage treatment in an *in vivo* model infected with a hypervirulent *C. difficile* strain. It should be noted that the strain used, B1, was much more virulent than the strain used in the Ramesh *et al*, as it was previously demonstrated to able cause the development of CDI symptoms and eventually death within 30 hours in the *in vivo* model. The strain used by Ramesh *et al* caused the same effects after 72 hours. One group of five hamsters was used for these experiments, all of which were challenged with *C. difficile* spores. Three animals were treated with phages while the other two were used as controls. The treated animals continued to receive a phage dose every 8 hours for 72 hours or until they succumbed to infection. The body temperature and activity of the animals was monitored during the course of the experiment. The control animals developed symptoms at around 27 hours after *C. difficile* challenge and had to be culled at around 33 hours when their body temperature fell below the clinical end point. The phage animals did not develop the clinical symptoms until 44 hours after challenge and their body temperature fell below the clinical end point between 47 and 51 hours, after which they were culled. All the animals’ colon and caecum were extracted and the *C. difficile* cells recovered as previously described. Comparison of the CFU counts of infected and treated animals did not show the existence of any significant statistical significance. This is not surprising or unexpected since all animals eventually developed symptoms and died as a result of the disease. Statistical analysis of the cull time of the treated animal’s vs the untreated controls showed that the former survived significantly longer than the latter. The delay in the onset of symptoms in the phage treated animals would suggest that the phage was successful in reducing the number
of bacteria enough to delay them from reaching numbers where the toxin production delivered delirious effects to the host.

Some of the animals showed a large amount of variation in the amount of *C. difficile* recovered between them and survival time. This can probably be attributed to individual differences between the animals and the relatively small number of animals used, and the fact that the animals were outbred.

Phages were recovered in high amounts from the colon and caecum washes in the phage treated animals infected with CD 1342 and B1. Phages were only observed in the stools of those animals infected with CD 1342. However this was because it was not possible to recover stools from the animals infected with CD B1 due to the acute diarrhoea. Nevertheless these results indicate that phage can survive passage through the gut (after stomach acid neutralisation) and remain infectious. In addition the high amount of phage recovered from the animals infected with CD 1342 suggest that phage CDHM1 might be able to replicate on the bacterial strain *in vivo* even though it is unable to do so *in vitro*.

Despite the fact that phage treatment was able to significantly reduce the amount of bacteria *in vivo*, it was unable to cause total clearance in neither the carriage nor the disease model. As a result the bacteria eventually overcome the phages and grow back allowing them to establish an infection despite the fact that the administration of phage in the treated animals was continuous. This would suggest that while the phage was able to kill a large number of bacteria, enough to delay the onset of symptoms and disease, the remaining bacteria survived by acquiring resistance to the infecting phage.

The inability of the phage to completely clear the bacteria culture can be due to several reasons. As already mentioned, there is the possibility that the bacteria acquired resistance to the phage. This resistance might originate from lysogeny of the infective phage. In the case the phage lysogenized, not only would it excluded from lysing the bacteria but also confer resistance to the host to other virions. If a non-lysogenic phage, which is a phage lacking the integrase gene was used this issue would not have arisen. Another explanation could be that the phage is simply not efficient enough at killing the bacteria. Evidence of this exists from the low plaquing efficiency observed in the *in vitro*
studies. The ability of the bacteria to form endospores could also be a contributing factor. As already mentioned in chapter 3, phages are unable to infect or adsorb to endospores of *C. difficile*. This means that those *C. difficile* that have formed endospores will not be affected by the phage dosage in the treated animals. The endospores might then germinate at a point where the phage amount is low and therefore manage to establish an infection.

### 4.3.1 Conclusion and Future work

It has been shown here that phage CDHM1 was able to reduce the number of *C. difficile* cells in *in vivo* as well as delay the onset of symptoms of CDI but has not been able to eliminate the bacteria population and the disease completely. This can be to a variety of reasons discussed above. We have demonstrated that phage treatment can be somewhat effective against an extremely virulent strain of *C. difficile*. These results led to cooperation between my lab with AmpliPhi BioSciences Corporation, an American phage research company, which funded a research project for the testing of more bacterial strains and phage combinations in an *in vivo* model that is being currently conducted by Dr Janet Nale. Also the data obtained here went part for the patent application for phage ΦCDHM1.
Chapter 5: Antagonistic Co-evolution of C. difficile and its bacteriophages
5.0.0 Introduction

Bacteria are known to be able to adapt in changes in their environment in order to ensure their continued survival (Modi et al. 2013). When an antibiotic is used to treat a bacterial infection, this applies evolutionary pressure on the bacteria as a result of the population bottleneck which, over many bacterial generations, could result in a wide spread strain that is either more resistant, or immune to the antibiotic used (Modi et al. 2013). When this occurs, new antibiotics need to be developed that are capable of efficiently killing the new strain. This process of sourcing novel antibiotics is both costly and time-consuming.

Phages are natural predators of bacteria, with which they have been locked in an evolutionary struggle over millions of years (Samson et al. 2013). In contrast to antibiotics, bacteriophages can also evolve along with their hosts, enabling them to eventually overcome any resistances the bacteria might develop which make them ideal candidates as an alternative to antibiotics (Gu et al. 2012).

When a population of phages attacks a population of bacteria, killing most of them, the latter must develop a resistance to the former or face extinction. If the bacterial population becomes resistant, then the phage population must evolve in order to attack the emerging strain, and the cycle goes on (Gómez & Buckling 2011).

The evolutionary aspect of the emergence of resistance to phages has been studied on many bacterial systems for example Salmonella, Pseudomonas, Escherichia and cyanobacteria (Barbosa et al. 2013; Lennon et al. 2007; Crill et al. 2000). In all of these studies it has been shown that whenever a viral strain invades a bacterial population, viral-resistant bacteria quickly emerge and persist through the population. This sets up a “paradox” (Buckling & Rainey 2002). If such viral resistant populations emerge so quickly and persist, what prevents them from of overcoming the sensitive bacteria and driving the viruses extinct? Previous studies suggest that in such a case, increased resistance to phages will come at the cost of fitness (Lennon et al. 2007; Poullain et al. 2008). A newly emerged viral-resistant bacterial population might be less able to complete for resources than sensitive bacteria (Buckling &
Rainey 2002). This will depend of course on the mechanism of the acquirement of resistance such as the loss or mutation of the bacterial phage receptor.

Phages also have the ability to evolve and adapt as a means of avoiding extinction. If a bacterial population becomes resistant to the local bacteriophage population, the bacteria must produce a new strain that is able to infect the emergent bacterial resistance population.

Exposure to an increasing number of bacteriophages could have additive fitness costs to bacterial population assuming that resistance to one parasite does not correlate with resistance to another parasite. Furthermore, a diverse bacteriophage population could select for a more general resistance mechanism which is more costly than single specific resistance mechanism.

There are different mechanisms through which bacterial resistance to phages can emerge, which can vary in specificity. For example, clustered regularly interspaced short palindromic repeats (CRISPR), which supress viral reproduction, and modifications of the phage receptor are typically specific. On the other hand, the complete loss of the receptor is likely to be more general.

Antagonistic co-evolution is the reciprocal evolution of host resistance and parasite infectivity. It plays an important role in the dynamics of host and virus as well as the phage infectivity (Morgan et al. 2009). Antagonistic co-evolution and the resultant population genetic differentiation is thought to operate on numerous host-parasite systems such as interactions between agricultural plants and associated insect and fungal pathogens, snails and their trematode parasites as well as bacteria and their associated phages. Antagonistic coevolution may be driven by directional or fluctuating selection or a combination of the two (Gómez & Buckling 2011). Directional selection, a mode of natural selection in which an extreme phenotype is favoured over other phenotypes, would favour hosts that are resistant to all parasites encountered and parasites that are able to infect all hosts encountered (Poullain et al. 2008). Fluctuating selection, a mode of natural selection where alternative phenotypes will be favoured, on the other hand may lead to different rather than greater resistance ranges being favoured through time. Fluctuating selection might be able to drive antagonistic coevolution indefinitely. Directional selection is less likely to do that as resistance and infectivity ranges must ultimately be constrained either genetically or metabolically. It is assumed that coevolution is
largely driven by fluctuating selection in natural populations and by directional selection in laboratory microbial populations (Buckling & Rainey 2002; Morgan et al. 2005; Paterson et al. 2010).

Dr Angus Buckling’s group published a series of papers investigating antagonistic coevolution using a bacteria and bacteriophage model. They were able to demonstrate antagonistic coevolution using *P. fluorescens* and a phage infecting it over the course of 400 bacterial generations. The results demonstrated largely the effects of directional selection as the antagonistic coevolution made the bacteria better at surviving phage infections and the phages better at infecting and killing bacteria (Paterson et al. 2010; Gómez & Buckling 2011; Morgan et al. 2007).

Although this model for studying antagonistic coevolution has been used in other bacteria-phage model, there is no published study investigating the model using *C. difficile* and its phages. Understanding how antagonistic coevolution affects the dynamics of *C. difficile* and its associated phages will provide important background data for the in development of phages as a therapeutic to treat CDIs more efficiently.

Therefore it was essential to explore the co-evolution between *C. difficile* and its phages. In addition to investigating the effects of antagonistic co-evolution on the phenotype of bacteria and phages it was decided to also investigate the effects on the genotype.

The virulent *C. difficile* strain AIU (ribotype 027) was used in this study. This strain was isolated from a stool sample of a patient with CDI by my lab. This was used in the experiments where it was subjected individually to two bacteriophages, the siphovirus ST, which was isolated and characterised in this study (chapter 3) and the myovirus T9, which was isolated by Swra Rashid, who’s host range includes strains of ribotype 027 (Rashid et al in prep). These two phages were chosen specifically for their different morphologies. This was done to determine whether and how the morphology of the phage plays a factor in the evolutionary struggle. The effects of variable MOI were also tested. Finally the assay was carried out both in BHI, the standard medium for growing *C. difficile* as well as gut model growth medium (GMGM), which simulates the colonic conditions. This was done in order to have a more representative antagonistic coevolution model of the colonic conditions (Baines et al. 2005).
In addition to carrying out assays in vitro, it is important to see what are the effects of exposing bacteria and phages in vivo. Samples of bacteria that had been exposed to phage in vivo were provided by Dr Gillian Douce. This exposure occurred in the hamster model, in the same way it has already been discussed in chapter 4. The bacterial strain used T6 was exposed to phage CDHM1 in the hamsters and subsequently extracted from the caecum and colon of the animals that succumbed to the infection.

**Figure 37: The Co-Evolutionary assay**

Strain AIU (R027) was streaked on a CCEY selective agar plate and incubated for 48 hours in an anaerobic chamber at 37°C. A single colony was then inoculated into FA broth and incubated overnight in the anaerobic chamber at 37°C. The Fa broth was then used to inoculate pre-reduced BHI or Gut model growth media and grown into early exponential stage (OD 0.2 at 550nm). After transferring an appropriate amount of bacteria to fresh pre-reduced media, the culture was then infected with phage T9 or ST at an MOI of 0.1 or 10. The culture was incubated for 12 hours. Then 1% of the culture was transferred to fresh broth which was also incubated for 12 hours. This was repeated for 15 transfers. At every transfer phage and bacteria were cryopreserved with glycerol (final concentration 25%).
5.1.0 Aims and objectives

The purpose of this series of experiments was to investigate how antagonistic coevolution affects the phenotype and genotype of *C. difficile* cells and their phages. In order to archive this, phages ΦT9 and ΦST were individually passaged along with strain CD AIU every 12 hours for 15 transfers (~200 bacterial generations). Bacteria and phages were collected at every transfer. Phages from all transfers were used to infect all bacteria from all transfers. This assay was carried out at an initial MOI of 0.1 and 10 and BHI media and GMGM for each phage.

The genotypic differences that arose during the course of the experiment were investigated as well. The DNA of bacteria and phage were extracted at the end of the coevolution experiment and sequenced. SNP analysis was performed in order to detect any genotypic changes that would explain the rise of resistance in the bacterial host. CRISPR analysis was carried out to determine if phage infection resulted in the insertion of new arrays in the host. The hosts were also investigated for the possibility of phage lysogeny by the phage used. The cost of acquiring resistance to bacteria was determined by comparing the growth dynamics of bacteria that had been exposed to phages against bacteria had not been exposed to phages. In order to determine whether or not antagonistic coevolution can take place in vivo, bacteria that had been exposed to phage in the hamster model, were infected with the ancestral strain of the phage used on plates.

5.2.0 Results

5.2.1. Co-evolution of bacteria and their phages affect their resistance and infectivity

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future from within their own replicate with a sample 20 bacterial colony clones for each of the 12 replicates. This experiment was carried out using different phages (ST and T9), different MOIs (0.1 and 10) and different media (BHI and gut model growth media). Statistical analysis (G-test of goodness-of-fit) indicates that the emergence of resistance/sensitivity, dependent mainly on
the MOI multiplicity used in each experiment, is of statistical significance. In addition to the importance of MOI, the media used also have an effect on the statistical significance as the use of GMGM moves the G value closer to the null hypothesis. Contrary to that BHI, moves the G value further away from the null hypothesis (that there are no differences in the number of times that future phage were more infectious than contemporary phage, and the number of times that contemporary bacteria were more resistant than past bacteria).

The data obtained indicate that there is as time-lagged antagonistic co-evolution between phage infectivity and bacterial resistance. This is indicated by the fact that generally speaking bacteria were more susceptible to phages from future transfers than their own contemporary phage, as well as the ancestral phage. Furthermore, phages from the past were less efficient in their ability to infect bacteria from future transfers.

5.2.1.1 Co-evolution of Phage ΦST/ΦT9 with AIU at MOI 0.1

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future from within their own replicate with a sample 20 bacterial colony clones for each of the 12 replicates. When phage ΦST was co-evolved with AIU at MOI of 0.1 in BHI, the emergence of the overall resistance of bacterial populations to phage was visible from the first transfer, where it increased from 0% to about 40% (Fig 38). Bacteria from subsequent transfers showed fluctuations in the degree of resistance gained and lost against the ancestral, contemporary and future phages. No phage isolated from any transfer appeared to be significantly more or less infectious than the other against the bacteria isolated from any time point although the contemporary phage did have a tendency to be slightly more infectious than the other two. When comparing the sensitivity-resistance of the bacteria against their contemporary phage and phage from two transfers into the future, the infectivity of the contemporary phage was observed to be statistically significantly higher (G=32.024, p<0.001). Similar results were obtained when the assay was repeated with the same phage and MOI, but instead of BHI the assay was carried out in GMGM (Fig 39).
5.2.1.1 Co-evolution of Phage ΦST with AIU at MOI of 10

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future from within their own replicate with a sample 20 bacterial colony clones for each of the 12 replicates. When phage ΦST was co-evolved with AIU at MOI of 10 in BHI, the emergence of overall resistance within bacterial populations to phage was visible from the first transfer, the degree of which varied depending on the transfer the phage used was isolated from with 80% to ancestral phage, 60% resistant to contemporary phage and 40% resistant to the future phage. Resistance of bacterial populations fluctuated depending on the phage population used. When infected with ancestral phage, bacteria showed the highest degree of resistance with the least amount of fluctuations. Bacteria were less resistant to their contemporary phage with more fluctuations in the degree of resistance. Future phage, were shown to have the highest degree of infectivity against bacteria from two transfers into the past with a high degree of fluctuations existing among the transfers (Fig 40). The resistance generated by bacteria to phage from two transfers into the future was statistically more significant than the resistance to the contemporary phage (G=7.114, p=0.00752). When the same assay was repeated but instead of BHI, the bacteria were grown in GMGM the initial overall resistance (20 bacterial colony clones for each of the 12 replicates) generated from the first to the ancestral phage was ~83% and ~60% resistant to the contemporary and future phage (Fig 42). The resistance generated by bacteria to phage from two transfers into the future was statistically more significant than the resistance to the contemporary phage (G=4.54, p= 0.033)

5.2.1.2 Co-evolution of Phage ΦT9 with AIU at MOI of 10

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future from within their own replicate with a sample 20 bacterial colony clones for each of the 12 replicates. When phage ΦT9 was co-evolved with AIU at MOI of 10 in BHI, the emergence of overall resistance within bacterial populations to phage was visible from the first transfer, the degree of which
varied depending on the transfer the phage used was isolated from with 90% resistance to ancestral phage and 75% resistant to contemporary and future phage (Fig 41). Resistance of bacterial populations fluctuated depending on the phage population used. When infected with ancestral phage, bacteria showed the highest degree of resistance with the least amount of fluctuations. Bacteria were less resistant to their contemporary phage with more fluctuations in the degree of resistance. Future phage, were shown to have the highest degree of infectivity against bacteria from two transfers into the past with a high degree of fluctuations existing among the transfers. The resistance generated by bacteria to phage from two transfers into the future was statistically more significant than the resistance to the contemporary phage (G=15.687, p= 0.000075). When the same assay was repeated but instead of BHI, the bacteria were grown in GMGM the initial overall resistance generated from the first to the ancestral phage was ~95% and ~80% resistant to the contemporary and future phage (Fig 43). The resistance generated by bacteria to phage from two transfers into the future was statistically more significant than the resistance to the contemporary phage (G=8.318, p= 0.003925)

5.2.1.3 Average resistance across all replicates and transfers

The phages isolated from all transfers within each of the 12 replicates were used to infect the bacteria from all transfers within that particular replicate (20 bacterial colony clones for each replicate). In the assay where phage ΦST was used at MOI 10 in BHI, the average resistance to phage at the first transfer was between 20% and 30%. This increased exponentially as transfers progressed with resistance in transfer 15 peaking at ~90 % (Fig 44). The same analysis was carried out for ΦST at MOI 10 in gut model growth media and ΦT9 at MOI 10 in BHI and GMGM with very similar results (Appendix 8; Fig 67-69)
Figure 38:  *C. difficile* strain AIU exposed to ΦST at an MOI of 0.1 in BHI

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage (green circles), contemporary phage (red squares) and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. From the first transfer, there was an increase of about 40% in overall phage resistance (12 replicates) with resistance in later bacterial transfers peaked at about 60%. Statistical analysis showed that there was significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=32.024, p<0.001) (SEM, n=20).
Figure 39: *C. difficile* strain AIU exposed to ΦST at an MOI of 0.1 in GMGM

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage (green circles), contemporary phage (red squares) and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. From the first transfer, there was an increase of about 40% in overall phage resistance (12 replicates) with resistance in later bacterial transfers peaked at about 60%. Statistical analysis showed that there was significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=32.024, p<0.001) (SEM, n=20).
Figure 40: *C. difficile* stain AIU exposed to ΦST at an MOI of 10 in BHI

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. There was an overall increase in phage resistance (12 replicates) from the first transfer, the extent of which varied with the phage used. The bacteria in the first transfer were ~80% resistant against the ancestral phage, ~60% resistant against the contemporary phage and ~40% resistant against the future phage. The resistance of bacteria from subsequent transfers fluctuated, the degree of which depended on the transfer the phage used to infect it was isolated from. The resistance against the ancestral phage was most constant whereas the resistance against the future phage had the most variability. Statistical analysis showed that there was significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=7.114, p=0.00752) (SEM, n=20).
Figure 41: *C. difficile* stain AlU exposed to ΦT9 at an MOI of 10 in BHI

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage, contemporary phage and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. There was an increase in the overall resistance (12 replicates) from the first transfer, the extent of which varied with the phage used. The bacteria in the first transfer were ~90% resistant against the ancestral phage, ~75% resistant against the contemporary phage and ~75% resistant against the future phage. The resistance of bacteria from subsequent transfers fluctuated, the degree of which depended on the transfer the phage used to infect it was isolated from. The resistance against the ancestral phage was most constant whereas the resistance against the future phage had the most variability. Statistical analysis showed that there significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=15.689, p= 0.000075) (SEM, n=20).
The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage (green circles), contemporary phage (red squares) and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. There was an increase in the overall resistance (12 replicates) from the first transfer, the extent of which varied with the phage used. The bacteria in the first transfer were ~83% resistant against the ancestral phage, ~60% resistant against the contemporary phage and ~60% resistant against the future phage. The resistance of bacteria from subsequent transfers fluctuated, the degree of which depended on the transfer the phage used to infect it was isolated from. The resistance against the ancestral phage was most constant whereas the resistance against the future phage had the most variability. Statistical analysis showed that there was significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=4.54, p= 0.033) (SEM, n=20).
Figure 43: *C. difficile* strain AlU exposed to ΦT9 at an MOI of 10 in GMGM

The bacteria sampled from each transfer (15 in total) were infected with the ancestral phage (green circles), contemporary phage (red squares) and phage from two transfers into the future (blue Triangles) within each of the 12 replicates with 20 bacterial clones for each replicate. There was an increase in overall resistance (12 replicates) from the first transfer, the extent of which varied with the phage used. The bacteria in the first transfer were ~95% resistant against the ancestral phage, ~80% resistant against the contemporary phage and ~80% resistant against the future phage. The resistance of bacteria from subsequent transfers fluctuated, the degree of which depended on the transfer the phage used to infect it was isolated from. The resistance against the ancestral phage was most constant whereas the resistance against the future phage had the most variability. Statistical analysis showed that there was significant difference in the resistance generated by the bacteria against the contemporary and future phage (G=8.318, p= 0.003925) (SEM, n=20)
Figure 44: Average bacterial resistance to phage across all transfers AIU-ST MOI 10 in BHI

Phage isolated from across all transfers was used to infect bacteria from each transfer for each of the 12 replicates. The average resistance to phage at the first transfer was between 20% and 30%. This increased exponentially as transfers progressed with resistance in transfer 15 peaking at ~90%
5.2.2 Co-Evolution affects the Degree of resistance and the phage plaquing efficiency

Colonies that were found to be resistant were grown in liquid media as previously discussed and used to measure the sensitivity of the ancestral phage by means of plaque assays. The same phage preparation was used to infect lawns of bacteria derived from the first, fifth, tenth and final transfer. The ancestral strain was used as control. The results showed that there was significant drop in the ability of the ancestral phage to infect the bacteria previously exposed to phage compared to the ancestral strain. From the first transfer there was a drop of about 3 logs in the PFU. Exposure of the phage in subsequent transfers showed a slight drop in the PFU obtained which was however not significant. Exposure to phage from the first transfer showed it had higher infectivity against the ancestral and first transfer bacteria and lower infectivity against the fifth, tenth and final transfers. Phage from the fifth transfer were had higher infectivity against ancestral first transfer bacteria, reduced infectivity against fifth transfer bacteria and lower infectivity against tenth and final transfer bacteria. Phage from transfer ten had increased infectivity against the ancestral, first and fifth transfer bacteria, reduced infectivity against tenth transfer bacteria and lower against final transfer bacteria. Phage from the final transfer had increased infectivity against ancestral, first, fifth and tenth transfer bacteria and reduced infectivity against final concentration bacteria (Fig 45). Statistical analysis indicated that the phages from the future were more significantly infectious when compared to the contemporary phage or phage from the past (2-way anova, p<0.05)

These results are consistent to the ones observed in the previous section, with future phages being better at infecting bacteria from the past (and newer bacteria being more resistant to old phages). However in this case, the culture used for the lawns was derived from a single colony that was previously deemed -to be resistant by streaking half a colony across a line of phage as previously described. These results would suggest that either the culture “lost” total resistance or that the resistance previously observed was actually a reduced ability of the phage to infect the bacteria (or a greater ability of the
bacteria to resist infection). It has been noted in previous research that acquisition of resistance comes at a cost of fitness (Koskella et al. 2012). Assuming that the mode of resistance is reversible (like lysogeny) or controllable (like the CRISPR/CAS system), when the bacteria are grown in the absence of phage, are possible to lose resistance as this would make them more biologically fit. If resistance is due to an irreversible event such as the loss of the receptor, this could suggest that the phage was able (or evolved to) infect an alternative receptor, without the same efficiency as the original.

It is also possible that the phage resistance previously observed was actually due to a reduced ability of the phage to infect the bacteria, due for example to a modified receptor and it is observable here due to the higher titre of phage used in the assay than the one isolated from the transfers. A combination of both factors might be also a more plausible explanation.
Figure 45: Plaquing efficiency of passaged phage

Phages that were isolated later in the co-evolution assay were more efficient in plaque formation against older bacteria than older phages against newer bacteria populations. A total of 10 colonies, that were previously found to be resistant, were used from each transfer. Phages isolated from the 15th transfer had a higher PFU count against ancestral bacteria as well as against bacteria from the first, fifth, tenth and fifteenth transfer. These results were obtained despite the fact that the colonies from which all the lawns were derived were tested to be resistant against their contemporary phage. Statistical analysis indicated that the phages from the future were more significantly infectious when compared to the contemporary phage or phage from the past (2-way anova, p<0.05) (SEM, n=3).
5.2.3 Single Nucleotide Polymorphism analysis

Twelve clones from the co-evolution assay of AIU and phage ST at MOI of 10 in BHI, one for each of the 12 replicates, were used to call SNPs against the WT bacterial strain. The same was done for the co-evolution assay for AIU and phage T9 at MOI of 10 in BHI. Of the total 24 clones, 3 of them had no SNPs detected. This could possibly be attributed to the very low coverage in those clones which was between 0.02-0.04x. In another sample there were 7850 SNPs detected. The large number of SNPs along with the discrepancy with the other clones suggests that contamination must have occurred. The same can be said for another sample where 251 SNPs were detected and had coverage of 0.7x (Table 5). The other 19 clones had an average of 12 SNPs detected. Using BLAST, the genes affected by the detected SNPs were checked on the NCBI database. The function and conserved domains of the affected ORFs were determined by using Blastp (Table 6). The amino acid sequences of the identified genes from the phage infected clones were aligned against the homologous sequences in the WT bacteria (Appendix 6).

Some of the amino acid substitutions in the putative membrane protein, cell wall hydrolase, hypothetical protein 2 and Purine Nucleoside receptor had different properties. These could potentially affect the functionality of the protein. The amino acid substitutions in both the membrane efflux pumps caused the insertion of stop codons that would make the polypeptide chain shorter, non-functional or completely absent from its location. The substitutions of amino acids would result in a modified protein, while the insertion of a stop codon would create a shorter, non-functional protein. If one or more of these proteins are involved in the adsorption of the phage on the surface of the cell, this would explain the emergence of resistance to phage infection. The fact that Purine Nucleoside receptor was mutated in 15 of the 19 clones analysed suggests that it could be particularly important in the adsorption stage.
Table 5: SNPs called for bacteria exposed to phage

A total of 24 sample exposed to phage (12 exposed to ΦST and ΦT9) at an MOI of 10 were sequenced and SNPs called again the reference genome (bacteria not exposed to phage). This showed the presence of a number of SNPs, some of which were common in a number of clones.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Coverage</th>
<th>SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1_ST</td>
<td>40.64X</td>
<td>17</td>
</tr>
<tr>
<td>A2_ST</td>
<td>44.05X</td>
<td>2</td>
</tr>
<tr>
<td>A3_ST</td>
<td>72.72X</td>
<td>5</td>
</tr>
<tr>
<td>B1_ST</td>
<td>32.67X</td>
<td>6</td>
</tr>
<tr>
<td>B2_ST</td>
<td>15.58X</td>
<td>7850</td>
</tr>
<tr>
<td>B3_ST</td>
<td>0.04X</td>
<td>0</td>
</tr>
<tr>
<td>C1_ST</td>
<td>119.1X</td>
<td>7</td>
</tr>
<tr>
<td>C2_ST</td>
<td>89.68X</td>
<td>3</td>
</tr>
<tr>
<td>C3_ST</td>
<td>85.7X</td>
<td>3</td>
</tr>
<tr>
<td>D1_ST</td>
<td>0.03X</td>
<td>0</td>
</tr>
<tr>
<td>D2_ST</td>
<td>27.22X</td>
<td>39</td>
</tr>
<tr>
<td>D3_ST</td>
<td>0.02X</td>
<td>0</td>
</tr>
<tr>
<td>A1_T9</td>
<td>62.93X</td>
<td>5</td>
</tr>
<tr>
<td>A2_T9</td>
<td>69.65X</td>
<td>5</td>
</tr>
<tr>
<td>A3_T9</td>
<td>19.86X</td>
<td>3</td>
</tr>
<tr>
<td>B1_T9</td>
<td>50.55X</td>
<td>5</td>
</tr>
<tr>
<td>B2_T9</td>
<td>75.45X</td>
<td>5</td>
</tr>
<tr>
<td>B3_T9</td>
<td>112.45X</td>
<td>4</td>
</tr>
<tr>
<td>C1_T9</td>
<td>0.7X</td>
<td>251</td>
</tr>
<tr>
<td>C2_T9</td>
<td>32.27X</td>
<td>54</td>
</tr>
<tr>
<td>C3_T9</td>
<td>50.81X</td>
<td>5</td>
</tr>
<tr>
<td>D1_T9</td>
<td>39.87X</td>
<td>5</td>
</tr>
<tr>
<td>D2_T9</td>
<td>43.81X</td>
<td>6</td>
</tr>
<tr>
<td>D3_T9</td>
<td>46.17X</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 6: SNPs observed caused frameshift mutations in certain genes

The SNPs observed in the 24 phage exposed clones were found to affect 9 genes. In 4 of those genes there was substitution of some amino acids with ones with different properties. In two of those genes, there were insertions of stop codons. In the first case this could result in a protein with altered properties and in the second, in a shorter or absent protein. If any of these proteins are involved in the adsorption of the phage, this could explain the resistance to phage observed in the clones.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of SNPs</th>
<th>Number of Samples where the SNPs were detected</th>
<th>Effect of Frameshift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putative Membrane Protein</td>
<td>17</td>
<td>3</td>
<td>Amino Acids with different properties</td>
</tr>
<tr>
<td>Cell Wall Hydrolase</td>
<td>20</td>
<td>2</td>
<td>Amino Acids with different properties</td>
</tr>
<tr>
<td>Swarming motility protein</td>
<td>3</td>
<td>4</td>
<td>Insertion of Stop Codons</td>
</tr>
<tr>
<td>Putative multidrug efflux pump</td>
<td>1</td>
<td>3</td>
<td>Insertion of Stop Codons</td>
</tr>
<tr>
<td>RNA polymerase</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Hypothetical Protein</td>
<td>4</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Purine Nucleoside Receptor</td>
<td>1</td>
<td>15</td>
<td>Amino Acids with different properties</td>
</tr>
<tr>
<td>Putative penicillin binding protein</td>
<td>2</td>
<td>2</td>
<td>Amino Acids with different properties</td>
</tr>
<tr>
<td>putative amidohydrolase</td>
<td>1</td>
<td>18</td>
<td>-</td>
</tr>
</tbody>
</table>
5.2.4 Co-evolution causes Lysogeny with the infecting phage

The phage infected clones were investigated for lysogeny using Mauve Qualimap to compare reads in the genomes of phage ST and phage T9 against the genome of the phage infected clones. All of the clones appeared to be positive for some reads that matched to the genome of the phages. However this is likely due to the presence of sequences common in the bacteria and phage or due to other prophages present. Only two clones had the reads of the full phage genome present: clones A1-ST and D2-ST. The genomes of these two clones were aligned against the sequence of phage ΦST using Geneious. Phage ST was inserted intact in sample D2-ST but appears to have been inserted as two fragments in sample A1-ST (Fig 46 and 47). No lysogeny was detected in those sample infected with ΦT9. This can be attributed to the small number of clones sequenced.

An insertion of a phage genome into the bacterial chromosome is facilitated by the integrase gene and occurs via site specific recombination between the phage attachment site, attP and bacterial host site, attB (Williams et al. 2013). The attP site, which is usually located downstream the integrase gene, can be identified by inverse PCR using divergent primers which are complementary to the integrase gene (Govind et al. 2006) Due to time constrains it was not possible to carry out identification of this site and it is something that should be carried out for future work.
Figure 46: The genome of ΦST inserted in the chromosome of sample D2-ST

Alignment the genome of ΦST and part of the genome sample D2-ST shows 100% identity indicating the presence as a prophage.
Figure 47: The genome of ΦST inserted into the chromosome of sample A1-ST

Alignment the genome of ΦST and part of the genome sample D2-ST shows 100% identity apart from one region (red square) indicating the presence as a prophage.
5.2.5 Effect of Co-evolution on CRISPR arrays

Phage-exposed clones were analysed to discern whether or not the phage infection resulted in alteration of their CRISPR arrays. Using the CRISPR-finder, analysis of the reference genome (not exposed to phage) identified 19 candidate CRISPR arrays. Six of these were marked as “questionable” due to the small size of spacers (one or two) and were therefore excluded from consideration. The rest were checked using BLAST on the NCBI database. This revealed that 3 of the sequences were actually parts of 3 different genes cdtA (Toxin A) cdtB (toxin B) and a hypothetical protein and thus no “true” CRISPRs. This left 10 sequences as candidates for CRISPR arrays (Table 7). Analysis also identified the CRISPR associated cas proteins, which are located near the largest CRISPR array (CRISPR 7). The cas proteins identified were cas1, cas2, cas3, cas4, cas5, cas7 and cas8al. These showed 99% identity with the homologues proteins in CD630.

Of the 24 phage exposed clones analysed, three were unable to assemble a genome. CRISPR analysis of sample C1-T9 identified the CRISPR as Pseudomonas spp, providing further evidence that the sample in question is contaminated as previously discussed. Analysis of sample B2-ST provided a totally different CRISPR profile than the reference strain also suggesting that this sample was contaminated.

When CRISPR analysis was performed on the remaining 19 clones, the homologous CRISPR arrays to the ones identified in the reference genome were aligned using Mega6. These ten CRISPR arrays were investigated for any insertions or deletions of spacers. Ten deletions of individual spacers were identified across seven clones. CRISPR 2 was found to have deletions in clones A1-ST, A3-ST, B1-ST, B1-T9, C2-T9, and D1-T9. CRISPR 4 was found to have deletions in sample D2-ST. CRISPR5 was found to have deletions in 2 clones. CRISPR 6 was found to have deletions in sample A1-ST. There were two spacers present in the clones that were not present in the Reference strain. One in CRISPR 2 in sample B1-ST and one in CRISPR 2 of sample C2-T9. However this did not appear to be new spacers due to insertion but rather due to mutation of the CRISPR array as there were only one and three nucleotides apart respectively from the homologous region in the reference strain identified.
in any of the clones (Table 8).

Deletions and SNPs at the distal end of the arrays have been documented to frequently occur in the CRISPR. The spacers at the end of the array are the “oldest” and therefore the ones most likely to be lost (K. Hargreaves et al. 2014). No new spacers were found inserted in any of the CRISPR arrays across all the clones investigated. From the results obtained we can infer that the emergence of resistance to phage infection did not occur as a result of the CRISPR/Cas system due to the absence of new spacers derived from the phages used in the experiment. The deletions and mutations observed in some of the clones do not appear to have occurred as a result of phage infection, as previous literature suggests that this can occur without any selection pressure. However, due to the relatively small number of clones analysed, neither of these hypotheses can be excluded from consideration. In addition to using CRISPR finder to identify the CRISPR arrays in the sequenced clones, primer pairs were designed for each of the 10 CRISPR arrays identified. Each of the primer pairs was designed to amplify the entire CRISPR array. These primers were used to carry out PCR assays in order to detect whether or not there has been an insertion in the array. Ten colonies were chosen from the final transfer, collected across different replicates and the DNA extracted used as a template with all the primer pairs. The product sizes for each primer pair varied between 500bp and 2000bp depending on the targeted CRISPR array. After being sequenced by GATC Biotech, the products were aligned using Mega against the homologous region in the reference genome. No insertions of novel spacers were detected either.
Table 7: CRISPR arrays in the reference strain AIU

CRIPSR finder identified 19 candidate CRISPR arrays in the reference strain AIU. Nine of these were excluded from consideration either because they were actually parts of genes or because they were too short (one or two spacers). This left 10 CRISPR arrays whose homologues from other strains were identified from the CRISPRs Database.

<table>
<thead>
<tr>
<th>CRISPR Arrays</th>
<th>Number of Spacers</th>
<th>BLAST</th>
<th>Homologous CRISPR in CRISPRs Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIU_Crispr_1</td>
<td>14</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_15</td>
</tr>
<tr>
<td>AIU_Crispr_2</td>
<td>9</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_20</td>
</tr>
<tr>
<td>AIU_Crispr_3</td>
<td>3</td>
<td>conserved hypothetical protein</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_4</td>
<td>4</td>
<td>conserved hypothetical protein</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_5</td>
<td>2</td>
<td>Questionable CRISPR</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_6</td>
<td>9</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_16</td>
</tr>
<tr>
<td>AIU_Crispr_7</td>
<td>14</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_10</td>
</tr>
<tr>
<td>AIU_Crispr_8</td>
<td>11</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_11</td>
</tr>
<tr>
<td>AIU_Crispr_9</td>
<td>1</td>
<td>Toxin B</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_10</td>
<td>1</td>
<td>Toxin A</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_11</td>
<td>1</td>
<td>Toxin A</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_12</td>
<td>4</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_12</td>
</tr>
<tr>
<td>AIU_Crispr_13</td>
<td>5</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_13</td>
</tr>
<tr>
<td>AIU_Crispr_14</td>
<td>4</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_14</td>
</tr>
<tr>
<td>AIU_Crispr_15</td>
<td>3</td>
<td>putative collagen-binding protein</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_16</td>
<td>26</td>
<td>Confirmed CRISPR</td>
<td>NC_013316_19</td>
</tr>
<tr>
<td>AIU_Crispr_17</td>
<td>5</td>
<td>Confirmed CRISPR</td>
<td>NC_013974_15</td>
</tr>
<tr>
<td>AIU_Crispr_18</td>
<td>1</td>
<td>Questionable CRISPR</td>
<td>-</td>
</tr>
<tr>
<td>AIU_Crispr_19</td>
<td>1</td>
<td>Questionable CRISPR</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 8: CRISPR arrays in the phage infected clones

The table shows the CRISPR arrays in the tested clones. Clones in grey are those where no sequence was able to be constructed. Clones in pink are those that were based on sequence of the CRISPR arrays were found to be contaminated. Green squares indicate that the sequence of the CRISPR arrays were 100% identical to the homologous arrays in the reference strain. Yellow squares indicate that the CRISPR arrays have deletions in relation to the homologous CRISPR arrays in the reference strains (this number excludes those CRISPR that were found to be shortened as a result of them being located at the end of a contig). Blue squares indicate the alteration of a spacer as a result of SNPs.

<table>
<thead>
<tr>
<th>Sample</th>
<th>CRISPR Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1-ST</td>
<td></td>
</tr>
<tr>
<td>A1-T9</td>
<td></td>
</tr>
<tr>
<td>A2-ST</td>
<td></td>
</tr>
<tr>
<td>A2-T9</td>
<td></td>
</tr>
<tr>
<td>A3-ST</td>
<td></td>
</tr>
<tr>
<td>A3-T9</td>
<td></td>
</tr>
<tr>
<td>B1-ST</td>
<td></td>
</tr>
<tr>
<td>B1-T9</td>
<td></td>
</tr>
<tr>
<td>B2-ST</td>
<td></td>
</tr>
<tr>
<td>B2-T9</td>
<td></td>
</tr>
<tr>
<td>B3-ST</td>
<td></td>
</tr>
<tr>
<td>B3-T9</td>
<td></td>
</tr>
<tr>
<td>C1-ST</td>
<td></td>
</tr>
<tr>
<td>C1-T9</td>
<td></td>
</tr>
<tr>
<td>C2-ST</td>
<td></td>
</tr>
<tr>
<td>C2-T9</td>
<td></td>
</tr>
<tr>
<td>C3-ST</td>
<td></td>
</tr>
<tr>
<td>C3-T9</td>
<td></td>
</tr>
<tr>
<td>D1-ST</td>
<td></td>
</tr>
<tr>
<td>D1-T9</td>
<td></td>
</tr>
<tr>
<td>D2-ST</td>
<td></td>
</tr>
<tr>
<td>D2-T9</td>
<td></td>
</tr>
<tr>
<td>D3-ST</td>
<td></td>
</tr>
<tr>
<td>D3-T9</td>
<td></td>
</tr>
</tbody>
</table>
5.2.7 Effects of Antagonistic coevolution in vivo

As the data have shown here, *C. difficile* cells and their phages undergo antagonistic coevolution in vitro, where bacteria quickly acquire resistance to phages. But the next question is, does this occur in vivo as well? In order to answer this question, bacteria that were exposed to phage in vivo were infected with the ancestral strain of the same phage in vitro.

Having obtained bacteria that had been exposed to phage CDHM1 in vivo, plaque assays were then carried out using lawns derived from these bacteria and infected with the ancestral strain phage CDHM1 (Fig 48). The bacteria chosen for these assays were previously determined to be resistant to the ancestral phage with the plate streak method as previously discussed. The bacteria that were exposed to phage in the infected animals were found to be more resistant to the ancestral phage than the bacteria that were not exposed to phage in vivo. The plaquing efficiency of the phage in the treated colon fell by 3 logs compared to the untreated colon. The equivalent reduction in the caecum was about 2 logs. The plaquing efficiency of the phage is statistically significant lower for the treated samples from the colon and caecum compared to their untreated counterparts (t-test, \( p=0.0092 \) and 0.0245 respectively. These data suggest that antagonistic coevolution occurs in vivo in a similar way as it occurs in vitro.
Figure 48: The plaquing efficiency of phage infected bacteria *in vivo*

Phage ΦCDHM1 was used to infect lawns of *C. difficile* (strain T6) that had been exposed *in vivo* with the same phage. Bacteria collected from the colon and caecum of infected hamsters was used. The bacteria isolated from the treated colon and caecum (blue and purple) had a higher resistance to phage than the bacteria collected from the untreated colon and caecum (red and green). The plaquing efficiency of the phage is statistically significant lower for the treated samples from the colon and caecum compared to their untreated counterparts (t-test, $p=0.0092$ and $0.0245$ respectively, $p < 0.05$) (SEM, n=3).
5.3.0 Discussion

5.3.1 Findings

Bacteriophages are natural predators of bacteria with which they have been locked in an evolutionary struggle for millions of years. Therefore their nature makes them ideal candidates for therapeutic agents.

Understanding and identifying the mode and manner by which resistance to bacteriophages arises by bacterial hosts is important when it comes to bacteriophage therapy. For therapeutic purposes, an ideal phage would be one that infects a bacterial population, multiplies at the site of infection and wipes out the bacterial population. However in evolutionary terms such a phage would not be fit for survival as wiping out its host would ensure its own extinction.

In this case I present, to my knowledge, the first study that investigates the effects of antagonistic coevolution in the emergence of resistance to bacteriophages in the bacterium *C. difficile*. Furthermore, I did only investigate the phenotypic changes that have been investigated in other systems but the genotypic changes that lead to a different phenotype.

In the experiments performed here, a phage population was used to infect a bacterial population and then they were subsequently grown together over the course of ~200 bacterial generations in rich media and over the course of ~100 bacterial generations in GMGM. Then the sensitivity of bacterial populations isolated from different time points in the assay was by infecting them with the ancestral phage, contemporary phage and phage from two transfers into the future from within their own replicate with a sample 20 bacterial colony clones for each of the 12 replicates.

Phages from the past, present and future were tested against the bacteria. It was observed that a significant degree of resistance to the phages arose very early in the assay (1st transfer). It was also observed that the bacterial populations isolated from different time points were more resistant to phages isolated from earlier transfers as well as their own contemporary phages than to phages isolated from future transfers. It was also shown that there was an incremental increase in the degree of bacterial resistance to phage within each of the replicates when compared to phage populations from all time points.
Evidence has also been provided, indicating that individual colonies that were thought to be immune to phage infection were in fact several orders of magnitude more resistant than those colonies that were determined to be sensitive. This provides further evidence that a modified rather than a lost phage receptor might be the main determinant in the generation of resistance.

The main determinant in the degree of resistance in the assays carried out was the bacterial-phage ratio (MOI). When a high MOI was used, the degree of resistance generated was much higher than when a lower MOI was used in the assay. Statistical analysis (G-test of goodness-of-fit) indicated that the increase in the generation of resistance was of statistical significance. These results make sense as a higher initial MOI would create a larger bottleneck that would kill a larger number of sensitive bacteria, or a higher selective pressure on bacteria to evolve resistance, allowing the resistant bacteria to grow better. These results are in contrast to the work presented by Hall et al., which indicated that MOI does not have an effect on infectivity and resistance evolution during coevolution (Hall et al. 2012). There are a number of reasons why these discrepancies might occur. First, the MOI used by Hall et al was $1^<$ for high and $1^>$ for low. In this experiment the high MOI was 10 and the low 1. This much higher MOI in our experiment might have delivered the bottleneck effect discussed above. Second the experimental design differs. While in Hall et al the MOI was kept consistent throughout the course of the experiment for every transfer by administration of phage at every transfer, in the experiments described here both the high and low MOI were administered at the beginning of the experiment. A future experiment could be designed to test the effects of the MOI used here with the method utilised by Hall et all.

In addition to the MOI, the media in which the assays were carried out had a statistically significant effect on the generation of bacterial resistance. This applied for all the phages across the transfer clones used (ancestral, contemporary, and future). This trend could be attributed to the fact that the *C. difficile* bacteria multiply slower in the GMGM rather than in BHI (generation time 120min vs 40min), the phages can kill them more efficiently and thus the evolutionary pressure to generate resistance is higher. With a longer generation time, the bacteria populations are much slower to increase their numbers, thus
more of them are killed by phages which creates a smaller bottleneck, allowing the resistant strains to grow to a larger extent. Statistical analysis (G-test of goodness-of-fit) indicated that when GMGM was used, the G value moved closer to the null hypothesis, while when BHI was used the G value moved further away from the null hypothesis (that there are no differences in the number of times that future phage were more infectious than contemporary phage, and the number of times that contemporary bacteria were more resistant than past bacteria).

In this experiment two different morphologies of phages were used in the coevolution assays: a siphovirus and a myovirus. The difference in morphologies did not have any significant effect in the degree of bacterial resistance to phage. This would be surprising if the bacterial resistance to phage is due to modification or loss of receptor as previous literature suggests that different phage morphologies use different receptor systems (Shin et al. 2012)

Bacterial colonies that were tested to have increased resistance to their contemporary phage were found to have several SNPs in genes mainly associated with the surface of the cell. In some cases the existence of SNPs caused a change in the reading frame of the amino acid sequence. This resulted in the substitution of some amino acids with either ones with different properties or with stop codons. In the first case this could cause to a new protein with altered properties. If said protein is involved in the adsorption of phages to the surface of the cell, an altered protein could potentially reduce the binding efficiency of the phage. In the case where a stop codon is inserted, this would lead in the formation of a shorter amino acid sequence which could cause a non-functional protein or one that is absent altogether. When this occurs for a protein that is involved in phage adsorption, phage binding could stop altogether. Thus these mutant proteins could be a factor in the bacterial resistance to phages.

Although CRISPR arrays have been shown to be involved in resistance to bacteriophages in other bacteria-phage systems, the research involving *C. difficile* and its phages has been limited. Recently however Hargreaves et al published a paper investigating the interaction of CRISPR and phages in strain
The analysis outlined here was not able to detect any insertions of new spacers derived from phages into the CRISPR arrays of the bacterial population. This would imply that the bacterial resistance to phages observed did not arise as a result of the CRISPR/Cas system. However this possibility cannot be totally discounted due to the relatively small number of clones analysed. Existing research suggests that *C. difficile* CRISPR arrays have received insertions of spacers originating from phages (K. Hargreaves *et al.* 2014). The phage-resistant bacterial colonies were also investigated to determine whether the phage used to infect the bacteria had become lysogenized into their genome. Two clones exposed to phage ΦST were shown to be lysogens. The presence of a prophage is possible to have contributed to the increase in bacterial resistance to phage as it is known that the existence of a prophage can inhibit re-infection by the same phage (Williamson *et al.* 2001). It should be noted that there was no significant difference between the resistance generated in the lysogenic and non-lysogenic sample that were exposed to phage. However due to the relatively small number of clones sequenced it is difficult to say that the presence of a prophage as well as SNPs have a cumulative effect on the degree of resistance, in contrast to those clones that acquired only SNPs and prophages. Furthermore there is also the possibility that the reason phage ΦST was not detected as a prophage in other clones might be because the phage can exists as a plasmid. There are evidence that suggest that phage Φ38-2, which is 95.4% similar to phage ΦST can exit as a plasmid in the host cell (Sekulovic *et al.* 2011). Further experimentation is required to determine if that is the case with phage ΦST however if that is the case it might explain why some of the resistant strains sequenced did not have ΦST inserted as a prophage.

While some of the bacteria were able to acquire resistance to phage infection and survive because that, it would appear that this resistance has come at cost. Bacteria that were exposed to phage showed reduced amount of growth *in vitro* compared to bacteria that had not been exposed to phage. While the difference seemed to be small and not statistically significant *in vitro* where nutrients are plentiful, in a nutrient deprived environment where phage is not present, the phage resistant bacteria would be at a disadvantage compared to
the sensitive bacteria.

When bacteria that had been exposed to phage in vivo were reinfected with the ancestral strain of the same phage in vitro, they were showed to be more resistant to the phage than bacteria that were passaged in vivo but not exposed to phage. This indicates that antagonistic coevolution can occur in vivo and the effects as similar to those observed in this study in vitro.

5.3.2 Conclusions

In this case coevolution would appear to be largely driven by directional selection with bacterial populations becoming more resistant to a wide range of bacteriophages and with bacteriophages able to infect a wide range of bacterial hosts. These findings are consistent with previous studies investigating microbial laboratory systems. On the other hand, studies investigating antagonistic coevolution in the wild found that fluctuating selection was the main driving force with the existence of more extensive polymorphisms for resistance and infectivity. Previous literature would suggest that coevolution is first driven by directional selection, which exhibits arms race dynamics, followed by fluctuating selection later (Hall et al. 2011).

I was able to provide evidence as to the ultimate cause of emergence of resistance, which data suggest to be associated with the acquiring of mutations on the proteins located on the surface of the bacterial cell. However whether or not those proteins are involved directly or indirectly to phage adsorption remains to be proven. I also provided evidence that this phage can exist as a prophage. There was no insertion of new CRISPR spacers detected in the clones either by the sequencing of the whole genome or the sequencing of PCR product. This however can be attributed to the small number of clones analysed, as both processes have been previously observed in the C. difficile-phage system.

The findings provide insight into the coevolution of C. difficile bacteria and phages in a laboratory setting. Although some work was done analysing the effects of antagonistic coevolution in vivo, this was not extensive and needs to be investigated further.
5.3.3 Future Work

One of the limitations in this study was the inability to isolate sufficient DNA for sequencing from the phages isolated from the different transfers. This would have allowed us to examine the genotypic changes in the phages as well as the ones in the bacteria. This limitation was attributed to the relatively low titre of phage isolated from each transfer as well as the difficulty in the DNA isolation process as a whole. If this experiment is repeated in the future, considerations should be made on how to overcome, this limitation.

Due to time constraints it was not possible to carry out identification of the attP site in the phage. This could be done inverse PCR using divergent primers which are complementary to the integrase gene (Govind et al. 2006). Any future work involving lysogens of this phage should aim to do this.

More work should be invested in investigating the effects of lysogeny in ΦST. Lysogeny with Φ38-2, with whom ΦST shares 95% identity, has been shown to increase the toxin production of the cell. It should therefore be investigated to see if lysogeny with ΦST has the same effects. Although there are no obvious genes associated with lysogenic conversion ΦST such is the case with Φ38-2 as well.

As already mentioned, there was limited amount of work done of investigating the effects of coevolution in vivo. Ideally an experiment that investigates antagonistic coevolution for C. difficile and its phages could be carried out in an in vivo model such as a hamster. Alternatively the artificial gut model could be used to replicate colonic conditions. Both models have prons and cons. While an in vivo model would give a more accurate picture of the degree of resistance and how this alters the generic makeup of both host and phage while the artificial gut model will allow more measurements to be taken to determine the timing of the generation of resistance.

Another possible experiment that could be carried out is the use of multiple phages in the assay to see how that affects the generation of resistance in bacteria. Similarly the use of multiple bacterial hosts can provide inside into how that affects evolution of the phage used to infect them.
Chapter 6: Release of Volatile organic Compounds in *C. difficile* and the impact of phages
6.1.0 Introduction

In recent years there have been a number of epidemics of *C. difficile* infections (CDIs), with high morbidity and mortality. These were caused, to a large extent, by strains of ribotype 027. The decline of ribotype 027 has given rise of other ribotypes like 014/020, 116, 078, 087 and 001. The *Clostridium difficile* ribotyping network (CDRN) has been monitoring the disruption of ribotypes across England and Wales since 2007. Although ribotyping is useful for categorising *C. difficile*, it is not practical for identification in a nosocomial setting, where a quick diagnosis is essential.

Identification of *C. difficile* toxins or toxin genes in the faeces of CDI patients currently depend on enzyme immunoassay (EIA) tests or tests based on PCR detection (Grein et al. 2014). While these tests are rapid, issues exist with both sensitivity and specificity (Probert 2011). PCR based tests require a priori knowledge gene sequences of strains being detected thus different strains may be missed. In contrast, a test that is dependent on metabolic Volatile Organic Compounds (VOCs) would be free from such constraints.

Having previously been shown to be effective as a diagnostic for the in different biological specimens, particularly in breath, urine and stool metabolic disorders such as cancer, urinary tract infections, inflammatory lung disease, hepatic dysfunction, diabetes and chronic inflammatory bowel disease, VOCs have the potential of being used as a diagnostic for the detection of *C. difficile* (Phillips et al. 2003; Miekisch et al. 2001; Manja & Rao 1983; Gallagher et al. 2008; Huestis et al. 1999; Buszewski et al. 2007; Angriman et al. 2007) Tait et al showed that gas chromatography mass spectrometry (HS-SPME-GC-MS) can be used to detect *C. difficile* based on the VOC emissions by cultures grown from the stools of patients with CDI (Tait et al. 2014).

The purpose of this study was to see if it would be possible to group strains of the same ribotype together based on their VOC emission profile and to investigate how phage infection in *C. difficile* affects VOC expression. This study was carried out in collaboration with Dr Paul Monks of the Department of Chemistry at the University of Leicester. Described here is the profiling of *C. difficile* ribotypes R027, R014/R020, R002, R013, R005, R107, R026, R087, R078, and R076 based on the release of VOCs using Proton-transfer reaction
mass spectrometry (PTR-MS). Handling of the instrument and data analysis was done by Dr Sharmilah Kuppusami and Dr Li Liu. The purpose of the investigation was the exploration of multiple VOC “fingerprints” as a metabolome for *C. difficile* ribotypes and to assess the potential of their application as a rapid and non-invasive method to diagnose *C. difficile* infections without the need for sample preparation. The release of VOCs was also observed in cultures of ribotype R027 strain AIU that was infected with phage $\Phi$ST. This was done in order to determine how phage infection would affect the release of VOCs in relation to an uninfected culture and the potential of developing a diagnostic for whether or not phage treatment is successful. Part of these data have been published in the journal of Metabolomics (Kuppusami et al. 2014).

6.2.0 Aims and objectives

The aims of the experiment was the determination of the VOC profile of ribotypes R027, R014/R020, R002, R013, R005, R107, R026, R087, R078, and R076 of *C. difficile*. This was done by streaking the relevant strains on Blood agar and their VOC profile determined using PTR-TOF-MS. The data obtained were used to determine the mass spectra accumulated over one minute. Principle component analysis (PCA) was used to generate a visual representation of the discrimination between the ribotypes by their metabolite profiles. In order to explore the similarities and differences between the VOC emissions from ribotypes, a dendrogram was produced using cluster analysis. Finally signal intensities of peaks arising from several m/z values for the various ribotypes were determined and given a tentative compound assignment.

Cultures of ribotype 027 strain AIU were infected with phage and their VOC profile was determined using PTR-TOF-MS.
6.3.0 Results

6.3.1 Clustering of Ribotypes based on VOCs

Strains of ribotypes R027, R014/R020, R002, R013, R005, R107, R026, R087, R078, and R076 of *C. difficile* were analysed using PTR-TOF-MS. Individual mass spectra for each sample were accumulated over one minute by measuring the VOCs in the headspace of the streaked sample. The measurements reveal complex mass spectra for six of the ribotypes: R027, R014/R020, R002, R013, R005 and R107. In contrast less complex mass spectra were found for R026, R078, R087 and R076. After the subtraction of the blank measurements, only a few peaks are observed in the 200–300 atomic mass unit (amu) mass regions (Fig 49). Statistical analysis indicates that 69 signals in the mass range between 15 and 200 were significantly different from the blank medium measurements (Mann–Whitney, P<0.05).

Principle component analysis (PCA) was used to represent the discrimination between the ribotypes by their metabolite profiles using the 69 mass peaks identified as being significant(Fig 50). It was performed using PLS toolbox which operates in MATLAB. The first two principle components accounted for 68.97% of the variance. Most of the ribotypes can be separated from each other with the exception of ribotypes R014/R020, R002 and R013, which almost overlap each other. Ribotypes R107 and R005 were clustered very closely but there is a clear separation between them. All of the other ribotypes can be clearly distinguished from each other.

A dendrogram was constructed in order to explore the similarities and differences between the VOC emissions from ribotypes using cluster analysis (Fig 51). The dendrogram shows that, based on the emission of VOCs, strains of the same ribotypes are clustered together with an overall relatedness at <30 on the variance weighted distance between clustered centres. Ribotypes 013, 014/020 and 002 have the largest differences in terms to their VOC emissions compared to the other ribotypes. R107 and R005 have a high degree of relation to each other. Ribotypes R027, R026, R078, and R087 are more similar when compared with the other ribotypes.
Figure 49: Mass spectra from the VOC headspace analysis of different ribotypes of *C. difficile* (Kuppusami et al. 2014)

There are complex mass spectra for six of the ribotypes: R027, R014/R020, R002, R013, R005 and R107. One the other hand R026, R078, R087 and R076 had less complex spectra. Only masses between 15 and 200 amu were measured.
Figure 50: A PCA biplot for the different *C. difficile* ribotypes

The first principal component has been plotted against the second principal component. The PCA was performed using peaks at 69 distinct m/z values. The oval lines surrounding each class depict a 95% confidence level. The results show that all the strains within the same ribotypes are clustered together based on the 69 mass peaks identified as significant. The results also indicate that certain ribotypes were more similar between themselves than with other ribotypes (Kuppusami *et al.* 2014).
Figure 51: Dendrogram of *C. difficile* ribotypes produced by cluster analysis

The dendrogram shows the similarity between the different ribotypes and their respective strains used according to the 69 mass peaks selected from a Mann–Whitney test (Kuppusami *et al.* 2014)
6.3.2 Metabolite identification

A tentative identification of some of the most significant peaks seen in the mass spectra was achieved. This was done for a total of six peaks (Fig 52). One of the most prominent peaks was m/z 33 which appears to be protonated methanol. Methanol appeared to be at highest concentrations in ribotypes 026, 078, 087 and 076. It was detected at significantly lower levels at ribotypes 027, 014/020 and 013. There was no methanol detected in ribotypes 107 and 005. Peak m/z 46 was assigned to dimethylamine based on the fact that the mass is consistent with a nitrogen-containing molecule and the fact that a previous study detected dimethylamine in *C. difficile* (Pons et al. 1985). Significant quantities of dimethylamine were only detected in ribotypes 005 and 107. Peak m/z 109 was assigned to p-cresol based on the fact that this compound was previously detected in *C. difficile* in other studies (Berg et al. 1985; Dawson et al. 2011). The compound was detected in significant quantities in ribotypes 014/020, 002 and 013. Ethylene sulfide was assigned to m/z 61 based on the fact it was previously reputedly released by *Clostridium spp* and was detected in high quantities in ribotypes 107 and 005 (Rimbault et al. 1986). Peak m/z 63 was assigned as dimethyl sulphide and was detected in high quantities in ribotypes 014/020, 013 and 002. It was detected in lower quantities in ribotypes 027, 026, 087 and 076. It was previously detected in *Clostridium spp* (Stotzky & Schenck 1976). Peak m/z 91 was assigned as methyl thioacetate and was detected in ribotype 027 in a concentration of three times more than other ribotypes and not at all in ribotype 076. The compound was previously detected in *C. difficile* (Rimbault et al. 1986).
Figure 52: Signal intensity chart of *C. difficile* ribotypes for selected mass peaks

Tentative assignments of the selected peaks are for protonated versions of the following compounds: m/z 33 = methanol; m/z 46 = dimethylamine; m/z 61 = ethylene sulfide; m/z 63 = dimethyl sulfide; m/z 91 = methyl thioacetate; m/z 109 = p-cresol (Kuppusami et al. 2014)
6.3.3 The effect of phage on the release of VOCs

Cultures of *C. difficile* strain AIU (R027) were analysed using PTR-TOF-MS as previously described. A total of six cultures of exponentially growing bacteria, three of which were infected with phage ΦST and three were used as controls. The analysis was done over the course of 24 hours. Figure 54 shows eight out of the top 16 protonated product ions (neutral analyte molecule plus 1 amu detected) with highest variance during measurement (Fig 53). The abundance of these VOCs fluctuated throughout the 24 hour period, with some increasing and others decreasing. While the release of some of these compounds was similar in both infected and uninfected culture, in other compounds it would appear that their release was affected by phage infection. Compounds m/z 57, m/z 41, m/z 39, m/z 27 and m/z 47 showed an initial increase in abundance followed by a decline in abundance in the uninfected cultures. On the contrary phage infected cultures had an initial decrease in abundance followed by an increase in abundance. In compound m/z 33 there is with an initial increase in the infected sample followed by a decrease. The uninfected cultures had a steady decline throughout the time course. Compound m/z 55 and m/z 45 had an initial increase in the uninfected samples followed by a decrease while the infected samples showed a decrease thought the experiment. The other compounds did not show distinct difference in the release of the VOCs based on the fact of whether the culture was infected or not. Although there was some variation in the initial abundance, this can be attributed to differences in growth among the replicates (Appendix 9).

The trend observed in compounds m/z 57, m/z 41, m/z 39, m/z 27 and m/z 47 can be attributed to the action of the phages. As the phages lyse the bacterial cells they inhibit their overall growth and thus delay the expression of these VOCs which are part of the cells metabolism. As the cells recover from the action of the phages so does the expression of these compounds. The trend seen in compound m/z 33 could be due to the lysis of the cells. As the cells are killed the compound is released from them and detected by the instrument. The drop in compound m/z 55 in infected samples but not uninfected samples could be a result of phage infection pre-lysis. Somehow phage infection prevents the expression or release of that particular compound.
The total ion count ratio of uninfected samples (red, green and black) and phage infected samples (purple, blue and light blue) is plotted against time. For some of the protonated compounds (m/z 57, m/z 41, m/z 39, m/z 27 and m/z 47) phage infection caused a delay in their expression or release from the cell. On the contrary m/z 33 increased in infected culture, likely as a result of cell lysis and release from the cell. The decreased expression of m/z 55 and m/z 45 in infected cultures could be attributed to a pre-lysis infected cell metabolism. One of the three measurements of the uninfected cultures (black) shows abnormal readings that do not agree with the rest after ~700mins. This is very likely to an instrument malfunction that occurred overnight during the course of the reading.
### 6.3.4 Identification of VOCs compounds

During the course of the measurement of VOCs abundance, the samples were analysed by gas chromatography (GC) during the first, second, third, fourth and twentieth hours of the experiment. Samples were taken using Solid-phase microextraction, (SPME). The data were analysed using Enhanced ChemStation chromatography software and compared with NIST database. The results were plotted on a GC chromatogram using Igor (Fig 55). While as of the writing of this work the analysis of the results is still underway, it was possible to have tentative assignments on some of the more abundant VOCs. Compound m/z 27 was assigned as C$_2$H$_3$+, m/z 29 was assigned as C$_2$H$_5$+, m/z 39 C$_3$H$_3$+, m/z 41 as C$_3$H$_5$+, m/z 45 as protonated as CO$_2$, m/z 47 as protonated ethanol, m/z 18 as NH$_4$+, m/z 19 as H$_3$O+, m/z 59 as protonated acetone and m/57 as 2-methyl-propan-2-ol(Table 9). The release of the more abundant VOCs appears to have taken place in the first 30 minutes of the analysis (each analysis step takes about 50min)

#### Table 9: Tentative VOC assignments

<table>
<thead>
<tr>
<th>Protonated Compound m/z</th>
<th>Tentative assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>NH$_4^+$</td>
</tr>
<tr>
<td>19</td>
<td>H$_3$O+</td>
</tr>
<tr>
<td>27</td>
<td>C$_2$H$_3^+$</td>
</tr>
<tr>
<td>29</td>
<td>C$_2$H$_5^+$</td>
</tr>
<tr>
<td>39</td>
<td>C$_3$H$_3^+$</td>
</tr>
<tr>
<td>41</td>
<td>C$_3$H$_5^+$</td>
</tr>
<tr>
<td>45</td>
<td>protonated CO$_2$</td>
</tr>
<tr>
<td>47</td>
<td>Protonated ethanol</td>
</tr>
<tr>
<td>59</td>
<td>Protonated acetone</td>
</tr>
<tr>
<td>57</td>
<td>Dehydrated 2-methyl-propan-2-ol</td>
</tr>
</tbody>
</table>
Figure 54: Gas Chromatography plot

The three figures show zoomed in section of the GC plot. Peaks in the plot show the VOS released by the culture and their comparison in the NIST library with identities of candidate compounds.
6.4.0 Discussion

The rapid detection and identification of *C. difficile* is a concern for healthcare facilities. While some of the diagnostics are fast, some lack specificity and some like PCR assays require priori knowledge of the genetic makeup of the target organisms. Volatile organic Compounds provide an alternative to tackle these issues.

Data obtained provide evidence that PTR-ToF-Ms analysis is capable of detecting VOCs emitted from *C. difficile* cultures within minutes. Some ribotypes were shown to have more complex mass spectra than other. Analysis of the 69 more prominent mass peaks using a PCA plot and a dendrogram indicated that it is possible to differentiate between some ribotypes and their strain based on their VOCs emissions. Some of the mass peaks were assigned a tentative function based on their mass and the detection of VOCs in *C. difficile* in previous studies. Analysis indicates that different ribotypes have markedly different emissions of methanol, p-cresol, dimethylamine, ethylene sulfide, dimethylsulfide and methyl thioacetate. The results indicate that detection of VOCs have the potential of being used in the future as not only as a diagnostic for *C. difficile* but also differentiating between different ribotypes.

The ability to differentiate between different ribotypes in a nosocomial setting it can be very important, particularly if a phage therapeutic, whose host range can be relatively short, is to be applied.

The effects of phage infection in the VOCs emissions on cultures of the clinically relevant *C. difficile* strain AIU (R027). The cultures were infected with the siphovirus ΦST at an MOI of 10. The VOC released from the cultures were detected over 24 hours. The results indicated that infection with phage caused a delay in the release of some VOCs detected (m/z 57, m/z 41, m/z 39, m/z 27 and m/z 47). This was attributed to the initial reduction in the numbers of bacteria. The subsequent regrowth of bacteria is used to explain the later increased abundance in these particular VOCs. In another compound (m/z 33) an increase in the abundance of the compound was linked to lysis of the cells and release of the particular VOC to the environment. The drop in compound m/z 55 in infected samples but not uninfected samples could be a result of phage infection pre-lysis. Somehow phage infection prevents the expression or release of that particular compound. Samples were also analysed by GC and
the results were compared against NIST database, which led the assignment of tentative IDs to some of the more prominent VOCs.

These data provide clear evidence that phage infection affects the expression of certain VOCs in a culture of *C. difficile*. This knowledge has the potential to set a foundation for a way to detect whether or not phage treatment of *C. difficile* infections are successful. This diagnostic could potentially be applied by screening the stools of patients currently treated.

### 6.4.1 Future work

As already mentioned, the data presented here studying the effects of phage treatment on VOCs is tentative and more analysis needs to be done in order to determine the exact VOCs affected by the treatment. In addition although evidence have been provided that VOCs have the potential of being used for the diagnosis and differentiation of *C. difficile* and its ribotypes more work and research needs to be carried out if this knowledge is to applied in a practical fashion in a nosocomial setting.
Chapter 7: Discussion, Future work and conclusion
7.0.0 Final Discussion

7.1.0 Summary

The anaerobic bacterium *Clostridium difficile* is the major causative agent of antibiotic associated diarrhoea. *Clostridium difficile infections* (CDIs) have been associated with a number of epidemics across the world, primarily in nosocomial settings, with high incidences of morbidity and mortality. Although number of cases has decline in recent years, as a result of preventative measures and decontaminations techniques adopted in NHS hospitals, treatment of existing infections is challenging (Johnson 2009). This is due to the resistance to multiple commonly used antibiotics and the reoccurrence of infection following treatment. There are currently three antibiotics used for the treatment of CDIs but resistance has been reported for all of them and two these have a high monetary cost (Surawicz et al. 2013; Bouza 2012; Gerding et al. 2014). Therefore alternative methods of treatment must be considered.

Bacteriophages are natural predators of bacteria, which makes them an ideal alternative therapeutic (Weinbauer 2004). Bacteriophages have been used for a century to treat bacterial infections, although their prominence declined since the discovery of antibiotics. However with the emergence of antibiotic resistant bacteria there has been a renewed interested in bacteriophage therapy (Kutter et al. 2015). A number of bacteriophages have been isolated that are able to infect a range of *C. difficile* ribotype strains, some of which have been sequenced and annotated. However there are to date only two studies investigating the effects of phage treatment on CDIs in an *in vivo* model (Ramesh et al. 1999; Govind et al. 2012) The studies in this thesis were carried out in order to consider the practical applications of developing phages of *C. difficile* as a therapeutic.

Soil sediments and water samples were enriched, and the supernatant of those enriched cultures was tested against bacteria lawns grown from important nosocomial strains of *C. difficile*. This led to the isolation of 8 novel bacteriophages. TEM and PCR assays indicated that all 8 of them belonged to the *Caudovirales* family of tailed viruses. Six of them belonged to the *Myoviridae* order and the remaining two to the *Siphoviridae* order. All of them
had a tail length of about 200nm and a head diameter of about 50nm. They were characterised by adsorption assays and one-step growth curves. This showed that they had an adsorption rate between 20% and 70% depending on the phage used. Similarly the one-step growth curves showed diversity into the effect of CFU/ml reduction as well as the release of PFU/ml. Host range analysis showed that seven out of eight of these phages were able to infect between one and four bacterial strains belonging to ribotypes 014/020, 001, 078 and 076. One phage (ΦST) however was able to infect 23/23 strains of ribotype 027, and one strain of ribotype 087, all of which were isolated from patients with CDI. In addition strains of these ribotypes have previously been associated with epidemics of CDI and are prevalent in UK hospitals (Arvand et al. 2009; Health Protection Agency 2012)

Killing assays were carried out in order to test the ability of phages to lyse cells in an in vitro culture. Four different phages, two myoviruses and two siphoviruses were used in different combinations of one, two, three and four were used to infect a culture of CD AIU (ribotype 027). This combination was tested in MOIs of 10 and 100. The results showed that both a high MOI and a combination of phages is needed to cause a significant reduction in the CFU count of a C. difficile culture.

The expansion of the host range of both existing and novel bacteriophages was attempted exposure of the phage to a C. difficile strain it cannot currently infect with simultaneous exposure to the propagating strain of the virus. This was not straightforward and although different previously methods that work were used, expansion of the phage host range was unsuccessful (Hall et al. 2013)

One of the novel phages isolated, ΦST, was chosen for sequencing because of its host range which included all 23 strains of ribotype 027 tested. Phage ST is a temperate siphovirus with a dsDNA genome of 41318bp. Out of the 51 ORFs, 22 have been assigned a predicted function based on nucleotide homology, amino acid homology and identification of a protein domain. The phage genome follows a modular arrangement with genes corresponding to different stages of the virion morphogenesis clustered together. Blastn analysis showed that the genome of the phage has a 95.1 % identity with phage 38-2 also a siphovirus with a similar host range (Sekulovic et al. 2011). The phage
has the integrase gene, suggesting that it has the ability to intergrade into the genomes of *C. difficile* cells. No toxin genes were detected in the genome of the cell. As shown in chapter 5, the phage has the ability to form lysogens but it was not possible to identify the attP site due to time constraints. It should be noted that sequencing of this phage proved challenging. The sequencing failed several times during the library preparation process and as a result there was considerable delay in acquiring the phage sequence and analysing it.

In order to test the therapeutic applications of *C. difficile in vivo*, I worked with Dr Gillian Douce in the University of Glasgow who has an established hamster model for *C. difficile* (Goulding et al. 2009; Buckley et al. 2011; Buckley et al. 2013). I identified that phage CDHM1, which was isolated by our lab, was able to infect bacterial strains CD 1342 and CD B1, which were both used in the hamster model to study colonisation and development of disease respectively. Infection of the avirulent CD 1342 *in vivo* with phage CDHM1 resulted in a statistically significant reduction in the amount of vegetative cells and spores extracted from the colon and caecum of the animal model. Infection of the virulent CD B1 *in vivo* with phage CDHM1 resulted in a statistically significant extension in the survival time and a delay in the onset of symptoms of treated animals.

Previous studies have shown that bacteria and their associated phages undergo antagonistic coevolution, the reciprocal evolution of host resistance and parasite infectivity, *in vitro* (Scanlan et al. 2013; Buckling & Rainey 2002; Morgan et al. 2009). These studies have shown that as time progresses, bacteria become more resistant to the infecting phage. Subsequently, the infecting phage evolves to become more infectious to the now resistant bacteria. While this system has been studied extensively using *Pseudomonas fluorescense* and in other bacterial species, no such study has been carried out for *C. difficile*. In addition, while these studies have explored the phenotypic changes in bacteria and phages, genotypic characterisation has been limited.

Having shown that phage treatment *in vivo* can reduce the CFU count and delay the onset of CDI in the model animal, the next step was to investigate why the bacteria are able to recover from phage and re-establish an infection even though phage treatment is continues. It was hypothesised that the
bacteria acquired resistance to the phage used to treat them and were therefore able to overtake them. This has huge implications on the application of phages as a therapeutic for *C. difficile*. It was therefore decided to examine how *C. difficile* coevolves with its phages. To test this, a series of assays similar to the ones described by Rainy *et al* were carried out.

This was done using phages ST, a *siphovirus*, and phage T9, a *myovirus*, both of which host ranges include strains of ribotype 027. The phages were individually passaged with *C. difficile* strain CD AIU (R027) over 100 and 200 bacterial generations in minimal and rich media respectively at varying MOIs. Cryostocks of both phages and bacteria were made for each of the 15 transfers. Phages from all transfers were used to infect the bacteria from all transfers. The results showed that bacteria were more susceptible to infection to phages from future transfers than to phages from their contemporary transfer or to the ancestral phage. This was true for both the *myovirus* and *siphovirus*. The initial MOI had significant effects on the degree of resistance generated. A higher initial MOI led to bacteria that had a higher degree of resistance to phages compared to bacteria that were exposed to a lower initial MOI. These results however are in contrast to previous studies that found that MOI does not affect the generation of resistance (Hall *et al*. 2012). Further experiments could be conducted in order to validate the findings.

Apart from studying the effects of the co-evolutionary process on the phenotype of the bacteria, the genotype was also analysed in order to determine the genetic causes of resistance to phages. Bacteria that were exposed to phage ST and bacteria that were exposed to phage T9 and isolated from the final transfer were sequenced and had SNPs called against a bacterial strain that had not been exposed to phage. This analysis showed that there were a number of proteins, mainly associated with the surface of the cell that had either amino acid substitutions or insertion of stop codon both of which could affect the function of the protein. If one or more of these proteins are involved in the adsorption of phage to the cell, this would explain the emergence of resistance in the cells exposed to phage.

Analysis also indicated that two of the samples exposed to phage became lysogens, which could also be the cause of resistance. However the fact that the other samples that were exposed to phage did not lysogenise,
suggest that if lysogeny is involved in the generation of resistance to phages it is not the only source of resistance.

Analysis of the CRISPR arrays of the sequenced phage infected samples did not indicate the insertion of any novel spacers associated with either the phage used. Analysis of the resistant colonies using PCR with primers targeting the arrays did not indicate any insertions either. However this could be due to the relative small number of samples analysed as well as the relative short timeframe the bacteria were allowed to coevolve with their phages.

Profiling of VOCs in *C. difficile* by Proton transfer reaction time of flight mass Spectrometry (PTR-ToF-MS) indicates that the ribotypes of the bacterial strains can be differentiated based on these emissions. In addition the more abundant of these VOCs were able to be assigned tentative identities based on their molecular weight. This indicates that VOCs have the potential to be used as not only a diagnostic for *C. difficile infection* but also for differentiating between different strains based on their ribotypes. Furthermore the effect of phage infection on the release of VOCs by *C. difficile* was investigated also by PTR-ToF-MS. This analysis showed that the release of some of the VOCs with the highest variance overtime were inhibited by phage infection of the *C. difficile* culture, while others actually increased as a result of the infection. Since these results indicate that the VOC of phage infected cultures are different from non-infected cultures, VOC have the potential of being developed as a diagnostic to determine whether phage treatment is working in a patient or not.
The development of a phage therapeutic

The development of a phage therapeutic is not a straightforward process, as it entails technical and regulatory difficulties. The first obvious step is to isolate appropriate phages. However, isolating suitable (ideally lytic) phages with a suitable host range and developing them can be challenging.

The next step is to test the ability of these phages to effectively lyse bacteria in an *in vitro* model. Choosing or developing an *in vitro* model which is representative of the conditions *in vivo* can be difficult, mainly because conditions *in vitro* are optimised to allow the growth of bacteria, and as a result provide conditions that are more favourable than their actual niche. This is particularly relevant for *C. difficile*, which can be challenging to grow even with rich media. Because of this, phage may be in a disadvantage in relation to bacteria in an *in vitro* model.

Similarly, difficulties also exist when choosing a suitable animal model to test the efficiency of a phage therapeutic following the *in vitro* studies. In the case of *C. difficile*, animal models tend to be either immune (mice) or too sensitive to the bacterial toxins.

After all the *in vitro* and *in vivo* assays have been completed, the next step is to carry out clinical studies. Before this is done however, the phage has to be purified in order to remove all traces and make it suitable for human consumption, and consider an efficient delivery method. For phage purification, Caecium Chloride or Sucrose gradient is regarded standard. The delivery method would depend on the site of infection. In the case of *C. difficile* phages, where the phages must survive passage through the gut, lyophilised phages in the form of tablets would be appropriate.

Even after all these issues have been addressed and a phage therapeutic passes the clinical trials and is approved for clinical use, there are still pitfalls that might appear. The most important of these is the development of resistant bacterial populations to the phages present in the phage cocktail. Another issue would be the inability to control the longevity of the phage injected by a patient. A way to address that would be continual application of phage doses. A *C. difficile* phage therapeutic is not expected to have adverse effects on human health since other phage therapeutics administered orally have not been shown to cause any such effects. Another important aspect of
this is to determine whether the phages that are to be used as a therapeutic would be administered as a cocktail or be given sequentially. Prior studies have shown that sequential use of antibiotics can be as effective or more effective than when given as a cocktail (Fuentes-hernandez et al. 2015).

This thesis has covered some of the steps in the development of a phage therapeutic. It covers the isolation, identification and characterisation of novel phages. It made unsuccessful attempts to expand the host ranges of said phages. It investigated their ability to kill bacteria in vitro and also carried small scale testing in vivo. Although this thesis has explored the use and effectiveness of cocktails of phages, it has not explored the possibility of using the phages sequentially. It might be useful to explore this possibility as a future goal.
Figure 55: Flowchart for the development of a phage therapeutic
7.3.0 Future Work

This study describes the isolation of eight novel phages, one of which has been sequenced. Even though our laboratory has a large collection of phages, more are needed if a suitable therapeutic against \textit{C. difficile} is to be developed. These phages must have a broad host range so that they can infect and kill as many bacterial strains as possible. Ideally these phage should be lytic (lacking the integrase gene), to ensure that they will not lysogenise and therefore be more efficient bacterial killers.

As the result shown here indicated, phages have more effective bactericidal properties when used as a cocktail rather as when used individually. Therefore obtaining bacteriophages with a wide host range capable of infecting clinically relevant ribotypes of \textit{C. difficile} is essential. Similar to this study these new phages must be characterised in terms of their morphology, adsorption and sequenced to check for the presence of toxin genes. In addition to the isolation of novel phages, the expansion of the host range of existing phage is also another possibility to explore. Even though multiple methods were attempted, it was not possible to do this. Any future research project into the therapeutic application of phages of \textit{C. difficile} must attempt to tackle this.

More work should be invested in investigating the effects of lysogeny in \(\Phi\)ST. Lysogeny with \(\Phi38\)-2, with whom \(\Phi\)ST shares 95% identity, has been shown to increase the toxin production of the cell. It should therefore be investigated to see if lysogeny with \(\Phi\)ST has the same effects. Although there are no obvious genes associated with lysogenic conversion \(\Phi\)ST such is the case with \(\Phi38\)-2 as well.

Having demonstrated that the use of multiple phages can cause a higher amount of reduction of bacterial cell counts \textit{in vitro} than the use of individual phages for ribotype 027 strains, this experiment should be carried out using strains of other clinically relevant ribotypes such as 106, 087, 078, 014/020 and 001. If possible this should be carried out using as many phages as possible, taking into consideration however that regulatory constrains might exist as to the number of phages that can go into a therapeutic.

Data in this study have shown that phage CDHM1 can effectively reduce the amount of bacterial counts \textit{in vivo}. This led to the patenting of this phage
among other ones for use as a therapeutic. The results generated also led to a collaboration with between our lab, Dr Douce’s lab and AmpliPhi BioSciences for use of the in vivo model with different combinations of bacterial and phage strains.

My work shows that bacteria of C. difficile and their phages undergo co-evolution in vitro, studying both the phenotypic and genotyping changes, neither of which has been investigated in previous studies. The next step in this would be creating a model to study the effects of co-evolution to both the phenotype and genotype of C. difficile and its phages in vivo, to see if the same observations will be seen as in vitro. If this is to be carried in a hamster model then an avirulent strain like CD1342 could be used so as to be able to study the long term effects of Co-evolution. The hamster model could be infected with spores of C. difficile and then inoculated with phage, as described in this study and then have their stools screened for bacteria and phage. The resistance to phage of the bacteria and the infectivity of phages could then be tested as described here.

7.4.0 Conclusion
This study has provided evidence that phages of C. difficile indeed have the potential of being used as therapeutic agents. Phages have been shown to be able to significantly reduce bacterial counts in vitro and in vivo as well as prevent the onset of disease in vivo. Phages have also been shown to be able to become better infectious agents by passaging. However applying phages as a therapeutic is not a straightforward process. Multiple phages might be needed to be able to effectively fight an infection, and they need to be genetically characterised to make sure they do not contain toxin genes. Lytic phages would be ideal for treatment but none has been isolated for C. difficile so far. Also while phages can evolve to become better infectious agents, bacteria can also evolve to become resistant to these phages, creating an evolutionary arms struggle. In conclusion, while it is not a straightforward process with a number of variables to be considered and more research needs to be invested into this, C. difficile phages do indeed have the potential to be used as therapeutic agents.
8.0.0 Appendices

8.1.0 Appendix 1: Media and Buffers Used

**BHI Broth (Oxoid):** 18.5g of BHI (Oxoid) were dissolved in 500ml of distilled water and autoclaved at 121°c for 15min.

**BHI 1%Agar:** 18.5g of BHI(Oxoid) and 5g of Bacteriological agar No 1 (Oxoid) were dissolved in 500ml of distilled water and autoclaved at 121°c for 15min.

**CCEY Agar:** 24g of CCEY agar (BioConnections) was dissolved in 470ml of distilled water and autoclaved at 121°c for 15min. When the media cooled down to 45°c it was supplemented with 1 vial of S2093 containing Cycloserine and Cefoxitin (BioConnections) dissolved in 5ml of ultra-pure water and 25ml of Egg yolk emulsion (BioConnections).

**FA Broth:** 14.8g of FA broth(BioConnections) was dissolved in 500ml of distilled water and autoclaved at 121°c for 15min.

**Overlay BHI 0.8% Agar(2X):** 18.5g of BHI(Oxoid) and 2g of Bacteriological agar No 1 (Oxoid) were dissolved in 500ml of distilled water and autoclaved at 121°c for 15min.

**Overlay Salts(2X):** 40.6g of Magnesium chloride hexahydrate (Acros Organics) and 0.75g of Calcium Chloride anhydrous (Acros Organics).

**Blood Agar:** 18.5g of BHI(Oxoid) and 5g of Bacteriological agar No 1 (Oxoid) were dissolved in 465ml of distilled water and autoclaved at 121°c for 15min. When the media cooled down to 50°c, 35ml of defibrianted horse blood (Oxoid) was added.

**Gut model Growth medium:** The medium consisted of (g/L): peptone water 2.0, yeast extract 2.0, NaCl 0.1, K2HPO4 0.04, KH2PO4 0.04, MgSO4 7H2O 0.01, CaCl2 2H2O 0.01, NaHCO3 2.0, haemin 0.005, cysteine HCl 0.5, bile salts 0.5, arabinogalactan 1, pectin 2, starch 3, vitamin K 10 µL/L and Tween 80 0.2%. After autoclaving, glucose (0.4 g/L) and starch (3 g/L) solutions were added to the medium through a sterile filtration device. Resazurin anaerobic indicator was also added at 0.005 g/L(Freeman et al. 2003)
5x TBE (Tris-Borate-EDTA): 54g Tris Base and 27.5g boric acid were mixed with 20ml 0.5M EDTA pH8 and then made up to 1L with UP H$_2$O

10x TE pH8 (100 mM Tris. Cl, 10 mM EDTA): Mix 10ml of 1M Tris.Cl and 2ml of 0.5 M EDTA and make up to 100ml with UP H$_2$O

PFGE Lysis buffer: Mix 20ml 1M Tris Cl, 20ml 10% SDS and 50ml 500mM EDTA and make up to 200ml with UP H$_2$O

CsCl solutions (25ml): Density 1.3 (10.1 g), Density 1.4(13.47g), Density 1.5(16.87g) and Density 1.7(23.57g). Each amount was made up to 25ml with SM buffer
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
<th>Product size(bp)</th>
<th>Annealing Temperature(°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>8F</td>
<td>AGAGTTTGATCCTGGCTCAG</td>
<td>1500</td>
<td>55</td>
<td>(Eckburg et al. 2005)</td>
</tr>
<tr>
<td>1391R</td>
<td>GACGGGCGGTTGTTGTRCA</td>
<td>227</td>
<td>48</td>
<td>(Eckburg et al. 2005)</td>
</tr>
<tr>
<td>CDHmyoF</td>
<td>TATACCAGAGCAGTTRCTRA</td>
<td>1500</td>
<td>48</td>
<td>(Shan et al. 2012)</td>
</tr>
<tr>
<td>CDHmyoR</td>
<td>CMTCCTTCAAYTGGTTGTA</td>
<td>227</td>
<td>48</td>
<td>(Shan et al. 2012)</td>
</tr>
<tr>
<td>CDHsiphoF</td>
<td>TATAGCGCTTTGCTRTTYA</td>
<td>150</td>
<td>46</td>
<td>(Shan et al. 2012)</td>
</tr>
<tr>
<td>CDHsiphoR</td>
<td>MGTTTCATTGCTCCCATTT</td>
<td>227</td>
<td>48</td>
<td>(Shan et al. 2012)</td>
</tr>
<tr>
<td>CRISPR 1 F</td>
<td>AGAGCTTCAAAATCCACTCCA</td>
<td>954</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 1 R</td>
<td>GGGCTCTCAGTTAATTTTC</td>
<td>625</td>
<td>47</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 2 F</td>
<td>CCCCAGTTTCTATCACC</td>
<td>625</td>
<td>47</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 2 R</td>
<td>GGGGAAACATAAGACAAAGA</td>
<td>622</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 3 F</td>
<td>GTCCTTTGATGCTTCAGGAGT</td>
<td>957</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 3 R</td>
<td>GGGGTCCCATCTTACATTAA</td>
<td>957</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 4 F</td>
<td>GGGGGTTCACCTGCAATTTAAG</td>
<td>1391</td>
<td>52</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 4 R</td>
<td>GGGGCACCTACTATTTAAGTAA</td>
<td>1391</td>
<td>52</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 5 F</td>
<td>GGGGTCTTGGAAATACAAAT</td>
<td>752</td>
<td>48</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 5 R</td>
<td>GGGACATTATAATCAGTCTTTC</td>
<td>752</td>
<td>48</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 6 F</td>
<td>GCGGGTCTGATATATGCT</td>
<td>293</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 6 R</td>
<td>GGGCTGTATATTATTAGTTG</td>
<td>293</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 7 F</td>
<td>GGGGAATCAGAAACAATAA</td>
<td>356</td>
<td>46</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 7 R</td>
<td>GCGGTATGAATATTCCAAAGGA</td>
<td>356</td>
<td>46</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 8 F</td>
<td>TGGGAATGCTTCTCTGATGG</td>
<td>290</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 8 R</td>
<td>TTGCTGATATGGCTATGCTT</td>
<td>290</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 9 F</td>
<td>GGGGGCAGTTTATTTTGG</td>
<td>1746</td>
<td>49</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 9 R</td>
<td>GGGGTCTTGGTTGACAG</td>
<td>1746</td>
<td>49</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 10 F</td>
<td>GGGGATAATCAATAGAATACACTATTTAAG</td>
<td>325</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>CRISPR 10 R</td>
<td>GGGTACTAAGTTTATTGTTGAGG</td>
<td>325</td>
<td>45</td>
<td>This study</td>
</tr>
<tr>
<td>For ST det</td>
<td>CAGGAGCAGGGAGGAGAAAC</td>
<td>1300</td>
<td>52</td>
<td>This study</td>
</tr>
<tr>
<td>Rev ST det</td>
<td>TCCATTCGGCCTGTAACCCT</td>
<td>1300</td>
<td>52</td>
<td>This study</td>
</tr>
<tr>
<td>K12groupF</td>
<td>GAACTCAGGGATTTGTATGC</td>
<td>814</td>
<td>48</td>
<td>(Hargreaves 2012)</td>
</tr>
<tr>
<td>K12groupR</td>
<td>CAACAAATTGTATTGCACAG</td>
<td>814</td>
<td>48</td>
<td>(Hargreaves 2012)</td>
</tr>
</tbody>
</table>
8.3.0 Appendix 3: PCR conditions and master mix components

**PCR conditions for detection of Holin gene in Myoviruses and siphoviruses**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 uM F primer</td>
<td>1 μl</td>
</tr>
<tr>
<td>50 uM R primer</td>
<td>1 μl</td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>1 μl</td>
</tr>
<tr>
<td>10x PCR buffer</td>
<td>2.5 μl</td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>1.5 μl</td>
</tr>
<tr>
<td>water</td>
<td>16.9 μl</td>
</tr>
<tr>
<td>5U/ul Taq polymerase</td>
<td>0.1 μl</td>
</tr>
<tr>
<td>DNA</td>
<td>10-100 ng (1 ul)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>25 ul</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature(°C)</th>
<th>Time</th>
<th>Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>94</td>
<td>5 min</td>
<td>-</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94</td>
<td>45 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Annealing</td>
<td>48 (myo) or 46 (sipho)</td>
<td>45 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Extension</td>
<td>72</td>
<td>1 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Final extension</td>
<td>72</td>
<td>10 min</td>
<td>-</td>
</tr>
<tr>
<td>Hold</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**PCR conditions for amplification of the CRISPR arrays**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM Forward Primer</td>
<td>2.5 μl</td>
</tr>
<tr>
<td>10 mM Reverse Primer</td>
<td>2.5 μl</td>
</tr>
<tr>
<td>One taq buffer</td>
<td>10 μl</td>
</tr>
<tr>
<td>UP H₂O</td>
<td>32.9 μl</td>
</tr>
<tr>
<td>One Taq polymerase</td>
<td>0.1 μl</td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>1 μl</td>
</tr>
<tr>
<td>DNA template</td>
<td>1 μl</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>50 μl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature(°C)</th>
<th>Time</th>
<th>Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95</td>
<td>5 min</td>
<td>-</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95</td>
<td>30 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Annealing</td>
<td>See Appendix 2</td>
<td>30 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Extension</td>
<td>68</td>
<td>2 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Final extension</td>
<td>68</td>
<td>10 min</td>
<td>-</td>
</tr>
<tr>
<td>Hold</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**PCR Conditions for the detection of bacterial DNA**
<table>
<thead>
<tr>
<th>50 uM F primer</th>
<th>1 µl</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 uM R primer</td>
<td>1 µl</td>
</tr>
<tr>
<td>2.5 mM dNTP</td>
<td>1.25 µl</td>
</tr>
<tr>
<td>10x PCR buffer</td>
<td>5 µl</td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>2 µl</td>
</tr>
<tr>
<td>water</td>
<td>38.7 µl</td>
</tr>
<tr>
<td>5U/µl Taq polymerase</td>
<td>0.1 µl</td>
</tr>
<tr>
<td>DNA</td>
<td>1 µl</td>
</tr>
<tr>
<td>Total</td>
<td>50 µl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature(°C)</th>
<th>Time</th>
<th>Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95</td>
<td>5 min</td>
<td>-</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95</td>
<td>30 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Annealing</td>
<td>55</td>
<td>1 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Extension</td>
<td>72</td>
<td>2 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Final extension</td>
<td>72</td>
<td>10 min</td>
<td>-</td>
</tr>
<tr>
<td>Hold</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**PCR Conditions for the detection of ΦST**

<table>
<thead>
<tr>
<th>50 uM F primer</th>
<th>1 µl</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 uM R primer</td>
<td>1 µl</td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>1 µl</td>
</tr>
<tr>
<td>10 PCR buffer</td>
<td>2.5 µl</td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>1.5 µl</td>
</tr>
<tr>
<td>water</td>
<td>16.9 µl</td>
</tr>
<tr>
<td>5U/µl Taq polymerase</td>
<td>0.1 µl</td>
</tr>
<tr>
<td>DNA</td>
<td>10-100 ng (1 ul)</td>
</tr>
<tr>
<td>Total</td>
<td>25 µl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature(°C)</th>
<th>Time</th>
<th>Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95</td>
<td>5 min</td>
<td>-</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95</td>
<td>30 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Annealing</td>
<td>52</td>
<td>30 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Extension</td>
<td>72</td>
<td>2 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Final extension</td>
<td>72</td>
<td>10 min</td>
<td>-</td>
</tr>
<tr>
<td>Hold</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
**PCR Conditions for the detection of the capsid protein of CDHM1**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>50 uM F primer</td>
<td>1 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 uM R primer</td>
<td>1 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 mM dNTP</td>
<td>1 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 PCR buffer</td>
<td>2.5 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 mM MgCl₂</td>
<td>1.5 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>water</td>
<td>16.9 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5U/μlTaq polymerase</td>
<td>0.1 μl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>10-100 ng (1 μl)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>25 μl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature(°C)</th>
<th>Time</th>
<th>Repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>94</td>
<td>5 min</td>
<td>-</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94</td>
<td>45 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Annealing</td>
<td>48</td>
<td>45 sec</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Extension</td>
<td>72</td>
<td>1 min</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Final extension</td>
<td>72</td>
<td>10 min</td>
<td>-</td>
</tr>
<tr>
<td>Hold</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
8.4.0 Appendix 4: Nucleotide sequences of annotated orfs of phage ST

>small terminase subunit - 125: 736 MW: 23278.527

MLEEKSCNISRWKIDRWDYRLGINNKVGAPAGNQNALGHEGGAPANNQNA
RTHGFFSKLLPARTYEIVKYIEKDGNSLDILWNSIVVQYAKILDSKITHVKNK
KDHTTDIKNGNKNKEYIEQHWSDVKVTNAMKAETAFKALSKMKIDYEELLHKN
WDLATEEQKSRDNIKARTNKLTGDLEIEDIEEAEIYG

>large terminase subunit - 723: 2045 MW: 49996.17

MAVTKKTIPFRFGDKHKEYIKCAINTYIAEGAVRAGKTVDVNVFAF
MEQYDEGIWYKRDILGRCVAEGLYEYFANNKEQFKTEDPLMDIVIGV
DFGGNKSYHAFVATGITYNYKVKVIALASERPSADTPDLNALLINFIKRVINLY
GKVDYVYCVDAEQLIRGKiNAVEEKKNVSNILKSAINDRILTDLISQG
RFETYTKHSETLNVNLCSAVWDSKETNEDVRLDDGSSDIDSDFAYTIERYL
KFIRGD

>portal protein - 2046: 3458 MW: 53506.094

MFEGKNAIRGVINLKNSSIQNELKVDIAVSDKMSRAIDSWMAMYKNKAPWL
NETTKSLNPSPAISSEVARLVTLELEVELSEIVGNDFLNEQYEVLEKDIRKYCEYAC
VAGGLVFKPYGRNKEIVDQVQANNNFPVEYSSTDNTSAIFSEIKVKGDIKYTR
LEYHNTSTSTSYSTNAYKSTNARILGYSYDNLGKRILSDVPWAEELSEELVLE
GIERPLFAYFQANSLGAVSVSFRAVDLIKEADQYSRILWEYEAETI
AIDASSMFRNNANGEKEYEVPHQIKRLYRMLEDGDNKWNVFPAPRDSSL
YAGLNQLRLKIEFCNGLSYGIISDAQEIEKTAEIKSKQRLYSTVLDQKLESDEA
LEGLIYSIDKWILAGFTPSSLKYETTFNWDDSVIIDKDSELLAMQDDVAGLIRP
ELYIMKGYGVTEEALKMMPDTGELVEDNPFKDGI
>Minor capsid protein - 3459: 4682 MW: 45827.01
MLTPNELQEPKYGFLGQTLLEDVFADIARRIAKAGNITDMAEWQLIRAEIGMA
DVKIKMIKAELGISFEAVDKLFEESAIKSIESDSALYEHAKLTPIHLNSSEELKNY
VSSAKEQTKGKLKMTGTLGCTIKGKVISSKLDTSYIESLDLAFQVSTGVLDYKTAVVKAKVLADSLGRLFYDTGWANRIDVATRRAVLTGQNMSQONINN
KVIEDLDTDIVEVTASHGARCEGAGIKNKKKWKQKWYSLSKSTKPSLRAV
TGWQGQDGLGGWNCSHQFHAVVPNAMEYKTEQLKNIDPPDGYKGRVYT
HYQALQYQRKIEAMRTQKRQLIAMYASEAGLKDIFETNASIKLQRQKQEQYREFSK
VAKLRLQNDRHQVLGYNKSISQKAVYAGKKEK

>conserved protein of unknown function - 4684: 5010 MW: 12669.785
MDVEILGTKYTILKNCTKKEKEPLMVKDFGFMDDTVNKIIIAKMEHSDESLDLNF
YEKQVRHEVIHAFLSESGLKSNSDWARNEEMVDYFAIQFPKMLKVFMELDAL

>hypothetical protein - 5026: 5307 MW: 10795.592
MGKYRKKPVTEAFKWLGSIEQKEEPKWVAIESGNVWWIWOQLGELSPRVMYIK
TLEGHEANVGDYIIQGIKGEYPCKADIFRETYEEVLY

>scaffolding protein - 5439: 6011 MW: 21783.746
MKREFLKDNLDEDEVIEKIMSENGREDIEKYKKEVEKKKEELESKNTELETANNK
IKDLEKIDVESIKKEVDDWKSKEAQKDKKLEINQMSEQTYLNLNDNLNSFK
FSSNLSEKAVKIKMKKEKGLYKNGAFEQADDYIKLEQANDPGAFINGNSVPMKV
VSSSGGDAEGSKVLMEQMIAKNREMLNI

>phage coat protein - 6022: 7095 MW: 39369.566
MSLFDSKIFGENEGFKGYVETVPNLNRDELIKSGAIRISSDLDKAMFSAQSGANY
ATIPMLGLDDDPVNYGDGQTIVATTTLKTFSQSVIVVGRAKVSMKDFSDQVT
GGVPFMCNVEQVGKAYWDNIDQKTLLAVLEGIFAMTGAKNLEFVDNHFTFDTK
EAETENQKVKAETLNSATQKACDNEKFALVIMHISTNLENLNLVARLKYTD
DPQGIQREMKLGTWGGKLVLIDDNMPTREVAESEGSNQGYTEYTTYVLGEN
SIDFEDVGAOKPYEMDRAPGKNGGEDLYRSRQKVAPFYGISFTKKAATAS
PTDAELKNANWELVNSNSKTYINHKAIPARIISRG
> conserved protein of unknown function - 7099: 7431 MW: 12808.485
MNNMYIDYEYSTLGGEPQDKFNYAIRATKYIDYNTFNRIEEVTEEIKIATCEIA
DLIYKSNLEGDKEIQSESIGSHSVTSANNKTIEQKAYDLKMYLDSDLLRYGV

> hypothetical protein - 7433: 7867 MW: 16817.482
VFFKDNITLYNKYDKSHTLYKRAYLIGVDYQGSKNIAVTDKGGLSADSVKVI
VPFSANARGKKYIDPFKYYELEDEIEKEKFYTFKVGIDIVKEVDIEITSVKPYTL
RELQSKFDDVSIKSVIKCDIFGSIRMRHIELEAE

> large capsid protein - 7874: 8221 MW: 13094.052
MNVHVIDVARIMRERGGERGGKAOQKFTHETVRRLSDYSVPQKTGKLNTAR
EEVDKIKYIQPYAKIHYYNNSRGMSEGSGGAGGKRGKYWDKRMALDQKELEIG
SVARFIGGSRE

> conserved protein of unknown function - 8218: 8649 MW: 16503.7
VTIVESVREFIKKCPYLDEFAKSINVEFLAEEFTSYTETVPAETIVKVFVGDSI
KQFVFIFASRESYGSVMQNIENSFQYEQFADWIYRENLENGNLVPEVLQNKES
MSLEVSTTGYPVQTSIDTAQYQIQLKLKYFEKGEI

> conserved protein of unknown function - 8652: 9452 MW: 28484.78
MAGIETIQRYKIAYDLGLGDDDQNSWELMGAGFNTLDENPSAQSDSKTYIND
RNSTSAIKSYYQTQFPFESDLIKSEKAMELYKIGRDQATGADAERDYLRVELFL
PVEGKDNTFKARKFRVSVEVSGLAGGETIIVSGNLNVGNYFVQDFGTQTT
KIFTEVGEAAPPILKLTNVAIGTNTGDTKVTVPGLESNGSYKYKGTGAV
SVPALDADCDSVGYTTWDGTEDITANEDKILIVEVNSNGNKAKKAGATVTVKE

> ap endonuclease family protein - 9510: 9872 MW: 13940.922
MLINNIELKDFDIYDADEYENIEKTIGLITITETGKVEKEATKGYEMIRDICNLIFKC
FDDIFGEGTNSKMKFGDRNLKVCSCAFFELIEEVNRQRQEGEEFFASAVSKYTPNRAQRRACK
> conserved protein of unknown function - 9872: 10480 MW: 23224.742

MNMLIDVLPENVEIDGEEYKINTDFRISILFEMLIQDSDISDEEKGNALSLYYPV
IPINTTVAIDKIIWFYSICGKREINDCPGQGAGGSGKQIYSDHDDDDDEYISAF
LGQYGIDLQDIKLHWWKFKALFKSKLEDEIVKIMGYRAMDISGDMPKAQKD
FYRKMKIIHAIPLPKNEVKEKINEIERALLNGGDISNLL

> hypothetical protein - 10452: 10562 MW: 4353.214

MVEILVIYYKIIYSTPSELYNYINILIIWGDICYGI

> conserved protein of unknown function - 10552: 11193 MW: 23647.22

MGFKFRKSNLQLGLKLNINKNSVSVSGGKVSRVSNKKGKATTTLGVPUGS
GLYYYQETTQLGGGKKKNKGSKAGGSESSRQAQQQLKIINDCANIVNTIKDPKTF
FYRYNMLLDKSYALVAIEDLNFEQPSQSMQLDSIIAKKTDTDINDFLTRYYNEM
SSKLKELKTEKARMNNAYKFSLAIDYEKRLNNDNKTKLNRLYKQLLDVKKQD

> tail tape - 11299: 16602 MW: 194998.17

MADGSIIIIDITLNTDAEQLQLNLADTMKKAKSALAIAGIGASLGAIGKQAIDF
GDEYQKAMNGFESATGNAEAKEAKGFGEALQQVYANNFGEAGMDSIEVMSL
SQQKLDGIDASNQEVTESAIMMRTDFDMVGESNLNGVNSMMRFQGIGSAKESYN
LNIAGQGAQQQLQNDDLGDQLAETYSTYYAQMFGSAEMNMKNGAEGG
YQIDYLNDALKENARDGSNKGTEAFELGNADELTKFANGGSSAEKAF
VEVTTALNNLDNVLKNOIGVQLGFTFKFEDLEADAVALTNEIESSSSKDKLE
EINKIKYNSFGAIGHTGIMQMQLPIALAEGLPWLNLANKFPAQFNENTKNAIRCSTAESIGALIMGVNLALPLTINVITFLHGSLVAPLIGVAFAGLKGAIAG
IVRGIKSFTDAQLAVRLFVTGMAESGTTTLTLYETIVGLVTKIISILAQVATNLWKGALTALGGPGLTIVAGALIGFVSLKNNEGFRDAINWNIKENIEVGEKIFGQ
IAKFTETIPQAMSIFDFKENWQGLALLVNPFIFGKLIYDENCEFRNTINK
LFNSIKDALSKTMSFKSNWQLLLILNPNFIFGKLIYDENCEFRNTVNSIFKNI
VGIFKSVSFKNNWSKISGMLLSPFVNGFKAIGNCFSNKVFSSFSSSVKGG
FVTGFKNUIIAGINLLSSGGILVVDTVPNINLGAVTFKBVIHSPSKVTTELGE
FAGIGFANGKNTSDKVVEATKELISAMEKLNKSKFKEICAITALKQNYE
SQKDIQIKALDERLVVEALKQFYVEKEYNEKLKYLDTNEKTAIQEQIODAIKDKQIEEQQKAEEAKYNEKISDLDRKLATAKNQEQKEREKIIMEAQAQQSDRQ
KLWLDEORQQENRLENKETEASNKLKKEOYKEEYDNOKKEAESKLELEK
SNASQKEEIEKYFSELLEETNIQEARRLLLQKNSEIIKLLSEYNPHWQDAGQ
SLADSLNGVNSKKQSIQEAVKEAINLKEIIPAQEKEKLDRLKKLEEEYKLEKA
NTSSASGGDVSASSAKSSSLDGSACKLDAGELEEAYADGIDDVKTSEVLELSV
DELVTGTVPKRDASKLSESIIKGFNGIGGGFFGGEIKWLRDTKDNINEWLAS
AKNSISDDSKVKDKFKEGIENIGGFGGGLGNKISEGFAAGAGDFFTEFGDKAQ
QGFSSIKDKEAIEFGLGSLSWFKETGSKIGDAIGSLKSGLDFFTSTVPKWFAGI
GDKIKECGKIGEIKISSFFNETMPAIIINNEVEFVKQIPYINIFVIGISAGFFVDLG
ANYAWATETLPEIINSIVEFVSLPEKISEFFSIILQNIQTWGENFKVSVSEFFG
AIWENITQFFTELPERISEFWNTVETISGWGENKLQALDFNFQDNVLIPIS
ELPGBKLEKFTDIIMGKVEWGSNIQKALETGSKFLIMKFFSLEPGKWEK
LTNIMDQKVMEEWVLKSKALETAKNFNIYNTVSOLPGKFREWLDNISNVIS
WGTNLVKQAGEAGKNNMANKDALKDDPSKIMSIGKDVVRGLWEGITGMGG
WLDQVLDFAASNVIDFGRDFGFGVHSPSIIIMDRLIGRNVKGVGVGIDVETPGL
KEKIEKNISETLGGKATVNFETSKIQANIVASTDFKAGKETAIMSNNKADEANE
NQAGIKLNIENFVNNRQQDIENLFDEILFKRKGAL

>tail component - 16702: 17457 MW: 28845.121
MRGGHKAMFFVYNGRDSREIPIYINLDSAPQMEVERVSVPGKDGDLLLK
KGFENFTLTIEDICDARQSNIEEEVATEIKKWQLGDISYKLFLNSDFYLYASCN
NKLDITRNKFNFASCSTLTFDCYFPFRAYECSIISLNVLNKSAITITNYRESKPVLY
IEAGDISIKINTQSIVLRGVAENGILSDLIDSEQMNVRINKANIIVNENKLF
SDFPILEEEGENQISWEGDIKSINPRWNIL

>tail protein - 17787: 19586 MW: 68568.75
LLNNFIESLENDITCEALERVFRACCCQEONRFRGHSDIKARNTIEMQSPYSA
ICEGENSLIQKPFETAKLFDDFNFNVLNYQRGESKNVLLAYKKNITGLEAEYDT
QDIYKDYPIYAYEDEMILTEKYIVISPNMKATQKIVAIYDFSDSEVSEEELR
EKCKDYFKYQVDPLKVLKYKNDVLSTTINYKDYKMLETVNLGDEV1RDFNL
NINTARVVVDSPINRKYYSIEVDGLNHDFLNNKFNNIENKETVKKAIADNV
KVDSEFDPDLPNTPALEAGLFANLNTWTFTNKSYYKYEYVASQIKDFVPDT
INFTNRFVGQASAYVHQAEPFQTWTYFRAVSNHSVTEFNOVAETTKID
DGMWIEKGAADALIGKLRDQWFQGLKMGMYNAREMVNNVNDVKMTMEI
DSFGDVHNFNPNTFKAIVSNPVYLLDEEGKLISGKFGFTKMDSGLYTKFTYG
SDVTESYLYLQSESTGYFYDQKDSKIIIDLPKKFAKGNNKATCAIRSTSTNGSCLSY
FSTSAATKSDKDTNPQLELSASVRGVSYEMSGSQTFIGYWLGMNFFSEEFITVQ
YYVYA
hologenic protein - 19613: 19855 MW: 9661.229

MVTIFYSKRTGEIYNCIKESEMQDYTCFADREEDYRQILDKIVVEDIPEILNPREDYKIENGKFVLKEIEEKEFTPKMVE

tail fibre protein - 20565: 22031 MW: 54702.1

MISDVNTKTTLIDDEVNTVKAEVTTAKNTMISEVTATAKETMTQEVTDAINAIPTKEELKGVGIEIKGSDLNISALPVNPTLSDAYFVKSATIENQIDLYVVDNTNWVKVPDIKIKGENGDGLGEFNWGDGTRLGIRIEGQENYTYTDLKGQKGDGDSIEFNVNGTRTLIGIKEIGEQQENYSYTELKGEQGYTPIEGNGNWWINNIDTQKPARGASRLRIGKLDISDNPLPLDPTIGDCWIIGRNIIYIYQTKWEDLGSLAGDVGGNKLEFNDWGTOQLGVRQQYQELDYKI1DKGDNIEFAWGDGTRLGVRIEGQNYTYTDLKGQKGDGIEFSWDGETLGVRIEGQEDSYTNNLKGATGNKLEFNNWSQIGIGREEGQTEYIYTELREGQGYTPAEGNGNWINGDGETGKASKGVTMWNELEKPEKDLGYIKKSTTFNSDGSTIILDSVSKTIKTFSNADGTVDEKYIDNVLVSJVTTFKGQIKEEVI

tail fibre protein - 22031: 23881 MW: 68239.66

MSWAETYKVNSDLQGPEPLNLQYLDQLKINGLDSYVLFIIGNARIWEEYLNSLYLFSDRGIRETSVTYAFSETDIDNLFKSTKLGQNAFRTIDFSLGNADNVKEMTIEHYNSLEEKFKAGYDQYTREQEKSTIGAWFNSTFSLDNTDLENTTIEEILANVEATNAILNSSNAIVALTMCKSSMDAVVASSNAMLGLQYI1RVTTESPVIRAILKNNVIRDAINSDEAMTQIQSNENSVMEIFNDLEATKVLVQONQNSINKILTNNVTEKVIPNLLEMKYNLQTSNLNYINTIKSNIASGKGGIMAITYNEYEEIFPILKNAVKNYDGMTTRNISQRDIEEIKISSDAISEIAMAFANAVSNKVGDRVGGIESIFSKTUSLNAFMKSTTAAILVINHKNKTTAFTKIANNSTAFNAMLTISENNVTDANNTAMGIIANNAQAMSTVANNDSISVFVNNTTAMGIIANSTAMTI1LGTGLALNRMVKSNATKISLKKNQTLQYKNIQNTHIQGSTAYFRTITGFADADDNPPTQINTYVGITYCQYKGSNYGYIYGHYNTSEIAAGRNGKYKDETTKFITLGGARYDQSGDGFTYAMYQAI

holin - 23967: 24230 MW: 9967.7295

MEEILITQLSGLGAVGILCALLFKKNTMQEKKEDRDMYKKTVENFIELSTQQEINKNILVEMGAMKTDVEEIKEDVTDIKDMLQKEGL
>endolysine - 24231: 25037 MW: 29834.406
MRVALTAGHTLTGKTGATGYINEGETNRILMDLVLKGGLETGAVVSESGKVD
KSNNYLAEQCQIANKQNVQDVAVQHIHFADHTTLNVMGTETIYKTNNGKTYAER
VNEKLATIFKNRGAKSedarLYWLSHTKAPAILIEVCFTIDSDKADTDYIRHKDIV
AKLIAEGILNKTIDNENSEDKMYKHTIVYGDVEVKVIPATVVGWGVNDGKILIC
DIKDYVPGQTQNLVYVGGGACEKIGSMTEKFTMIKGNDRFDTLYKALDFIDR

> hypothetical protein - 16021: 16233 MW: 8391.818
MTWKNEIKSLVKENISMNENLNNLMNEYNNKNNTVENLROQKINNETMYSEILT
IANLIGYEVVWKKKEK

> hypothetical protein - 25477: 26307 MW: 32838.973
MKTLFKEAHKIAREIKEKYNDVYKVQFSLCLSFLNKGDICKMKELKGSEQII
WANDIRTDILNLTNELEKSKIERIKNEDYKIKDMSIEEMTERCRKFERIREAISN
IEDAKFFIDFRNVLYKNSLUNQKAFQINFIRESQFTEEIGNLKFKGETQASHK
LRKVEGSISYEEAKKIAEINFNDNWSNEIKEVEIVEIFDDAVKQATSTEYIEK
AKKMAILTIQKNAQFYKNGFADIKKQIENYRKEAERKIEIAYQIFRDGVEC

> hypothetical protein - 26386: 26892 MW: 19553.484
MNIDFSTILNLKVNKENEMIEKEINKIINDIFDISLKDCAANSHTGIVKETVGKLS
NIETWGHGEINSSIGIRLFNFFYYSKGDKVIDIENILYIVNSLCSKTHEEDILIA
SVRPGFYLKLEEKSREVEKCEIRHEFTGINKQEIKEFLEYLTEIEKERTVMYKN

> conserved protein of unknown function - 26912: 27382 MW: 18492.938
MNKQKARRFLRVIDMNIDKIEEEAIKAFESCLIKEIKNKIIYIDIQGKVEAIAVQI
WAKLGDDKEINFTLQAPTHLNMLGEICYVNDYEEFNWENWENLDW
DSYKKFNKENFEEAERNIDSTSVFLEELQGIESCKQQLQVNVIEN

> conserved protein of unknown function - 13423: 13704 MW: 10716.461
MNVYLQIGKELSEKMTKSENLKFSQEVTNLYRFAHGHDNDMKKLJQTTIEAIIAR
KRVTLPADNEFTQGIMSNELEMTTNVGFIIEGLYHNN
> hypothetical protein - 12947: 13279

MPSFNRLDLVDKEKPILKNGTIVQNDSTTTDNKTILNYNLVNNKSKKVVTMSATL
DEELVNLKLSFSDMNSDVSKLLSDILTQILSDVTIKEENLNIYNERNRRNKTKK
K

> Para protein - 12137: 12946

MANTKLLTYFNKGGIIYKTTSIMAYELAKDDKDKKLLWLDLVQANLTQYYVEI
NHNDNTTLDILKGISANDAIKSPENENYNVDLIPDSIQMARFEQESPLPAPEK
FLARWYMQNFNTLSEYDIICDLSRYDTAKNVFLADSIIIPQDKNISSLRGA
ELFKQLWVIENTRYFDKE LinkedIn NSTILVGFEKKTQISDDTFSYLEGFDNMDRDMI
L DTYIRKNEFIEKALLKLSLTDYKTITKEHFSRQETFNTMLEELKRKVGL

> hypothetical protein - 11666: 11848

MEKEKRDLKFHKGGSFPRMSPVKWADEMGITTERPNVTATFEDRI
IEEKAKNE

> hypothetical protein - 10693: 10845

MDNFLNVLQGVIALSFYIISKLFRQVKNHSQKSGWEFDLRIFKFKFK

> Dnase helicase - 31125: 33311

VEIGGIKMKNVKNKNDCDMDLEECASAERAEHEKENYSFNRHYEIKYINEK
KLGICKAIAINAINGGKTLFIAPTAGKAGSYFINTLKLKTKLALFIPLNASVQAM
HEYDIPGAYDKIPVSEAFKNGNLVMTWDKTEKILNTDLSEYIIIVDEIHQYTD
AYRGKAIKLNNSMKCKGRDLITATPTKLEFEIYDIEYQTQKTEYKWLYND
FDKNNFTEMINIKSKNSAMLNDISTLFIERSDVNMKVGGVYAGKEENELY
NRIVKNSDMKDYYETLNTTTLAGVNNIINKIDIIDIVNIKDVAIAKQYVARFRNLK
KVNVHIFMNKIECNVYKIEWLVNKNKIEKATILKDANVSKHMLQETFVGINA
TPIRDSNVNYTCKDNYIKSYQYVSNYYNTRTIQSFKVLLEEFYVENNI
TDKEVGTNEKELKEYKKVVEIKDATREILKEHKEILVGYQIKDSKSFSLM
QYHNDMKLSSKGCLEAYRAYDIHNLVKSKSNMLYSNVLVYLDNFKDLDL
AWKLANTANKRGRGIFNKINTLIIYLREKKEEEPEFLNDELIQVTVFNKIDLKFGGITS
YKELHQELSDNLRTILGENWDLTTKLSVTLNIGIFIEKKSYSNWHDITEFMFFY
KNINPISCOFQEKKIKRQLNNIKRYITVDIJKDLEDEKDSLENAIKTDRMLNS
LDEDEKLLLLKGMF
> hypothetical protein - 33358: 33705 MW: 13865.892
MFELDIKVKNKYIVSVSYEFSELLIRAFTANDYDNFIKDKIKNKDKFIPGINLGYRELL
EENSAYHEYNFYYDEEVKDYFFELQYQEOQDLREELSKQLLLELIKINLPAAGELIVVD

> CI phage repressor - 6447: 6800 MW: 13602.939
MNTRIRDLRIAEGLSQIEFGNKLGMTRHEIYNLESGRTRIKESDLKLITSTFKVNEKWLKTGEQDMYTFSNNNLITAEVLCAIDKNERLAKAMLKFSKLNDKELEAMEKLLLELFQKE

> hypothetical protein - 35045: 35227 MW: 7060.3145
MTREERILRNAILYLDMSEEKQRIVDCILNAKTENIVENDILDVNKLNEAVKEIQIIC

> conserved protein of unknown function 35294: 35719 MW: 15507.861
MIKVILCSTIDEYIEYEDVIGAKDSYLMDVIIGECSEIDAIETLEILKNSKARIVSEAETCLKRGEFQSKLRRHCDNCGAIVDSSEELTSMPVLENYNNNSVSDQLGSIDVCPNCLLEITDLL

> replication terminator protein - 35794: 36198 MW: 15007.063
MINLETLCNGETKEKIEGFMEIFKNIQDPNTPATSTRSLTVKVTLPNKGNRSH
VNTQIQVIPKLASVLPSBFLIIVEKDFRTGEVNAEYNQLPGQVKLGDLKESKE
TSVTEETVEKESEEDNVRKFKSLKDL

> hypothetical protein - 36221: 36934 MW: 26518.996
MLTDLFNVRMEVGARKIETREINGFVYTTTDPLIEVHPRAEMVRSTLDGLIGYIKSSIDSLKSGKLMVQVESPKSVILGCELSEGKRDIYESRALPHVNFDCSLDAEQFNVMQSFVDTEDKKLLLKVSNGVKEKSTVQDDGVSVQVTTMSTGVA
SVEDVILPNVRKLPYRTFAEVDQPESEFIRVKEGIAMKLEADGGAWRLEAIKNIKEYLEKKGKLEIGNIDIAA
> hypothetical protein - 37011: 37154 MW: 5320.149
MGAIERGEYVGNVAELKEEENPHLSKIDLIELALKKIREEEKEQEAGQ

> hypothetical protein - 37151: 37462 MW: 12358.621
MKIGKYLIVSEQVDRNIECLKEENKVRDDKIRYYERIMSNNEELIQILEEQLKE
YKSELDVYSLFIDCILNEKETRKAIAIYRTKSIRIKQKCLNSIPIKKG
>PROKKA_00029 - 37508: 37657 MW: 5771.62
MSEFVQCCCCCERTINIEENNHVQYEKEALGLVFTLYFCLNCVDELSEME

> hypothetical protein - 37688: 37861 MW: 6510.61
MKMKNVLKLSAMYFCPKCGSDELGEDEGSIVDEYTFHRKCKCGFDIIVIDER
NDEI

>ssDNA binding protein - 37974: 38300 MW: 12170.15
MNNTTLIGRLTRDPELKYIPGSGTAVSTFTIAVDRDYVKKDGTETDFIPIEVMG
KLAEVCANNLDKGRLVAEGSIRVNLYEKDGKRTYTKVHTNKKFLDYKKRG

> hypothetical protein - 38392: 38757 MW: 14053.612
MIKKVEKIFSGSVNLNFEDDLTFLKNYLALCICIKPVKHYKNKVLASSTKACS
ALKIREHQQKERRKVRVLNAKTGEIKEMIACLEAENFLNVKDLVPIIRGRPTRT
GWYVKDSNEERY

> hypothetical protein - 38741: 38914 MW: 6902.31
MKKDIKYNVIYFNSRTRLDKESFSTLIDARVCKNEIKIKEYENVEIIEKLM
V
>sigma factor - 38926: 39498 MW: 22606.371
MAKKKRICSWCGKSFFTESKNFGFCCKSCERKYYKDKKEKVWSSCLVEETKEYL
IELGNKKRRVIDLRKEIDFKDFKREINFNELGFKIHSPKGIDEIMNEDKISFKESIDFIEYRRLMDIVISELTEEQKIIKFMYFYDGASKLSNEEIAASKIKCSR
SNVHYKHEALKELAEMLFESGLL

> integrase - 39676: 40284 MW: 23574.797
VGSKKPNPVKDKRMVLNIQEYLKEKSVRNYVLFVLGATGYRAGDLVQLQVRDVRNAIDEGYFLIMESKKEKTNIRKKNKPKRPAPIVPNLERVLKSYIRDKDYEYMFPISQKSVTPHIGVERTVILKEAGRYFGLKHITAHSMRKTYAYTIYEESGFDIIIRVKEMLGHSSIEETKVYLGLNEEQQYQEYSMFLNDLIG

>resolvase - 40722: 41318 MW: 22855.758

MVKSINGCFMFIFGCRVSSKVQIDNSLEQEQEIKRHYSDAKIFREQHTGSTTHRPIFEMVKQMRGDKLVVTKLDLARNATEGINLQNILFDSGISVHNLVGLLENTTLGKFFITTLLAVEMERNLILERTFAGKEIAKQNPNFREGRPKKYTKMQMEHAELLLKTNSYTQVERMTGISKLQLQRAKRQED
Appendix 5: CRISPR Sequences of reference genome

- **CRISPR 1**
  - CRISPR start position: 294953
  - CRISPR end position: 295907
  - CRISPR length: 954

```
ATTTACATACCACGATGTTAATATAAAAAACAAATTTCGATTTGGAACACG
AATAAAAATTGCAATTATATATCATGTAGTTAATATAAAGCAGATAATTGTTG
CATATATAGTTTCAACTAAGTATAAGATTACATTACCATCATTAGTTAAATAA
ACAATGTGTCGACACACATAGAAAGGAAAAGTAAAATTTATACATAC
CACAATGGTAACTGATATTGTTATTATATATTATATTATATTATATTCTGTT
AAAATTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
TCTAGGCACACACATTTTCTGTTAATATAAAACACATTGTTAATGTTG
AAAATTTACATACCACATAGTTAATAATTAAACGATTTTTTCATGGGAGGAA
AATTGATGTTGGGACATTTAACATACCACATAGTTAATAAAGCAGATAAA
```

- **CRISPR 2**
  - CRISPR start position: 462762
  - CRISPR end position: 463387
  - CRISPR length: 625

```
ATTTATTTACATATAATAATAATAAAAACTATAATGAAGTGGGTGTTATGAAAGA
GAAAACACAAATTAAAATTACATTTTCATATAGTTAATAAAAACAGTTTTTT
TATTTTTTCTAAGAGAAGAAACTAGAGACTTTTACATCCATATAGTTAATAAA
AACTTTTCTGTATATTTTAATTGTTATTTAATTATAAACACTTTTACATTCCATA
TAGTTAATATAAAAACATGCGAGTAACTAAGAAGAAAACTATACCTTTTTTTATT
CTTTACATTCCATATATAAGTATAAAAACACGGAAATTATATACATATGACTG
ATGCTAGATTTAACTTTTACATTTCCATATAGTTAATATAAAAACATTTATTAAAG
GAATTATACGACCAAAGCAGCAAGACGCTTTTACATTCATATTAGTTAATAATAA
AAACACAGCAGCAAGAGTGCCGCAACACATTTTACCTTTACACATTCCATAT
CATATGTTAATAATAAAAACCTTTTATTTTTCTCTATATCGAGTGTCGGC
TCTACCTTTTACATTTCCATATAGTTAATATAAAAACAGTCCAGCTTTGTGT
CGTAACACACAAATCTACTTTTACATTTCCATATAGTTAATATAAAAAC
```
• >CRSIPR 3  CRISPR start position : 1465530  ------- CRISPR end position : 1466152  ------- CRISPR length : 622

GTTTAAATTACACTAAGTTAGTTATAAAATATTCCGGAAAGTACTTTTGTATT
TGATTATTCTACCCGGTTAATTACACTAAGTTAGTTATAAAATTTCTACTAATT
TTTAGTTACACTATTACAAGTCCACGCCGTTAATTACACTAAGTTAGTTA
TAAATTAGTGAATGATTTTTAATGTTTTTCTAGCAACAAATTGTTTTTAAATAC
CTAAATTAGTTATAAAATTTCTACTAAGTTAGTTATAAAATGTAACAGCA
CGAATGATGACCACACCCCTATCTGAATTGTTAAATTACACTAAGTTAGTTA
TAAAAGAACGACTTTAAAGGAATAGACACACCGTACAGAAGCTATTAAAC
TAACACTAAGTTAGTTATAAAATTTCTGACTATATGGAATGTTATTTTAGTA
GAGAGTTTAAATTACACTAAGTTAGTTATAAAATTTTTAAGGGGTGAGAG
TTGTAAGTACAGCTCAGTTTATTTTTGATTTGAAAATTTT

• >CRSIPR 4  CRISPR start position : 1537629  ------- CRISPR end position : 1538586  ------- CRISPR length : 957

GTTTTATTTAATTTAATTTAAGTTTAGTTAAGTGAATGATTTGCACCGCTTTGG
AAATCTACAAAGGATTATTTTCTCCACAGCTTTATATTACACTAAGTTAGTT
GTAAAGAAATAAAAGGATTATTTTGCATAACTCTCTCTTTTTCTTGTGTTTTTAATAG
CTAAGTTAGTTAAGTACAGCAGGTGTTTTATATTACACTAAGTTAGTTA
GGAATGATGATGTTATTTGACTATATGGAATGTTATTTTAGTA
GAGAGTTTAAATTACACTAAGTTAGTTATAAAATTTTTAAGGGGTGAGAG
TTGTAAGTACAGCTCAGTTTATTTTTGATTTGAAAATTTT
- >CRISPR 5  CRISPR start position : 1746975  ------- CRISPR end position : 1747727  ------- CRISPR length : 752

GTTTTATATTAACCTAAGTGTTAGGAAAGAGTTTTTTCAGAAGTATATGTAC
GTCTATATGGAAGTTTATATATTAACCTAAGTGTTAGGAAATAGGAAAGT
GGAAGTCATATGCATGCACTCTTTAATATTTAACCTAAGTGTTAGGAAATG
TAAATACGAAATTAGAGTTTGAATATATGACTACATAATAGTTTTTATAT
CTAAGTGATGATAATTTTGGTCACATCATTGTCACCTCTTTTGAATATCAA
AAAGTTTTATATAACTAAGTGTTAGGAAATTTTTAACGCAAGTGATTAA
TTAAAAAGTGAAGTTTTATATATTAACCTAAGTGTTAGGAAATTTAAAA
CTTGTCAAGTTGATATTATTGCAACAACTCTTTACTTTTTTATATACTAAGT
GTATGTAATGAAATTATATATATGAAAAATATGGGCTACAGACAGGGTTTT
ATATTAACCTAAGTGTTAGGAAATTTTGCAGTTTTTTCTATTAGTTTTGCATA
TGTTCCCTGGTTTTATACTAAGTGTTGATAATTACTCTTAGCACTAGT
AAAGGTGATAATTATATATAGTTTTATATTAAACTAAGTGTTAGGAAATTT
AATTTTTACAAATTCTTCCAACTTTTTGGTTTTTATATATACTAAGT
GTATGTAATGATGATTAGAGTAAATCCATCTTTTTGTTTACAATATTTATAT
ACTATACAACTAAGTTTTTACGGAATATAAAT

- >CRISPR 6  CRISPR start position : 2649732  ------- CRISPR end position : 2650025  ------- CRISPR length : 293

GTTTTATATTAACCTAATGGAATGTAAATCCTCGGAGAATTTCCTAATGGGA
TTTAATAGGAAAGTTTTATATATGGAATGTAAATTTTAGCAGAATTT
TTTGTTAAGTGCTCATTATTTTTATATTTTTATATATATTATAGGAAATG
TAATACATAGAAACTTTTTCTTCTGTTATATTAAAACGTATTATTATATTAA
CTAATGGAATGTAATCCACTTTTTGTTTACAAATTTTACATACCTACG
GTTTTATATTAACCTAATGGAATATAAAT

- >CRISPR 7  CRISPR start position : 2742409  ------- CRISPR end position : 2742765  ------- CRISPR length : 356

GTTTAGATTAACCTAATGGAATGTAAATGTTGTAAGAAGGTATCATCATTCTT
TTAATCTTTCTGGTTTTTGAATATTAACCTAATGGAATGAAATTTTCCAGTG
AATTAAGTTTATTGCTCGATGAACTACCTGTTTTTAGATATTAACCTATAGG
AAATATGACTATAATGAGCTTTTACATCGATTATGAAAGTTTTTAGATATAC
TTATATGGAATGTAATGATTGATCATTATTTACAGCGTCATCAGCTTTTGTCT
ATGCCTTTGATATTAAACCTAATGGAATGTAAATTTTACAGTGAAACATTACG
TTACTTTGGAATGTTTTATATTAACTATATGGAATCAAAAT
- **CRISPR 8**
  - CRISPR start position: 2744630
  - CRISPR end position: 2744920
  - CRISPR length: 290

  GTTTTATTTAATTTGTTAATTTTATTTTTGAATATTTTGTATATTGCTATTGTT
  ATTCTACAATTTTATATATGCTTTTATTAATATAAATAAGTTTTCTAAATAC
  TCGTGTGATTTGCTATTTATTCTTATCTACGTATTTTGTATATAAAACCT
  GACACTGTTGTCTTGTGAGTCTTTTCTTTTATCTCTATTTACATACGTCTTAG
  TAAATATAAAACAGATTAAGCCGATATAAAAGACTGCAATCTCTATTGCATT
  TACATACACTCTTAGTTATAATAAATACCAATTTAGTAAACACGTGATTTTTGGG
  ATTACATACACTCTTAGTTATAAATACCAATTTAGTAAACACGTGATTTTTGGG

- **CRISPR 9**
  - CRISPR start position: 3028032
  - CRISPR end position: 3029778
  - CRISPR length: 1746

  ATTTAGATGATTTAGTTAATATGATGATTTTAGTTAATGCTATTGTT
  ATTCTACAATTTTATATATGCTTTTATTAATATAAATAAGTTTTCTAAATAC
  TCGTGTGATTTGCTATTTATTCTTATCTACGTATTTTGTATATAAAACCT
  GACACTGTTGTCTTGTGAGTCTTTTCTTTTATCTCTATTTACATACGTCTTAG
  TAAATATAAAACAGATTAAGCCGATATAAAAGACTGCAATCTCTATTGCATT
  TACATACACTCTTAGTTATAATAAATACCAATTTAGTAAACACGTGATTTTTGGG
  ATTACATACACTCTTAGTTATAAATACCAATTTAGTAAACACGTGATTTTTGGG
ATAATGGGATGGTCAGGTTTCTGATAAGGTTAATAAATTTACATACCACCTT
AGTTAATATAAAAACGTGTTTTTTAGTTTTCGATTCAATGGTTCTTTGTTAT
TTACATACCACCTTAGTTAATATAAAAACAACCTCATGTAAACAAAGTAAGACCT
GCACCTTTAAATTACATACCACTTAGTTAATATAAAAACCTTACATATTTCCCTT
TCATAGCAGCTTATGGTTAACCATAAAAATTACATACCACCTTAGTTAATATAAAAACCT
TTAGTTGTGTCGCTTTTGCTTCTACTATAGTTGCTTATTACATACCACCTTA
GTTAATATAAACC

- >CRISPR 10 CRISPR start position : 3065535 ---------- CRISPR end
  position : 3065860 ---------- CRISPR length : 325

ACAGGATGGAAAGCAATTAATAACAAAAAAATATTATTTAGGAACTGATGGA
GCAAGAGTATCTGGATGGAAAGTAATAGACGGAAAAACTTATTACCTTTAAC
TCAGATGGAGTAATGCCAAACAGGATGGCAGAAAAAAGAATGGGAAAGAAGAATGG
TAATAAGTATTACTTTGACAAAAACAGGAAATAATGGCGACAGGATGGAACA
AATAGATGTAAGTCTTACTATTAGGACAAAGAGTCAGATTCTTACGG
TTCTCAAAAAATAGATGGGAA
8.6.0 Appendix 6: Alignments of WT and Mutant proteins of *C. difficile* that coevolved with phage.

**Figure 56: Alignment of WT and mutant cell wall hydrolase**

The mutant protein had amino acid substitutions at positions 201 (serine for proline), 233 (isoleucine for valine), 236 (alanine for glycine) and 269 (asparagine for aspartic acid)
Figure 57: Alignment of WT and mutant membrane efflux pump 1
The mutant protein had stop codons inserted in amino acid positions 511, 513, 585.
The mutant protein had an insertion of a stop codon in position 364 in place of a glutamic acid.
Figure 59: Alignment of WT and mutant putative amidohydrolase
No frameshift occurred as a result of the SNPs

Figure 60: Alignment of WT and mutant hypothetical protein 1
No frameshift occurred as a result of the SNPs
Figure 61: Alignment of WT and mutant hypothetical protein 2

There were amino acid substitutions in positions 7 (arginine for serine) and 9 (asparagine for lysine).

Figure 62: Alignment of WT and mutant putative membrane protein

There were amino acid substitutions in positions 60 (glutamine for glutamic acid) and 63 (isoleucine for alanine).
There was one amino acid substitution, were a proline was replaced by a serine in amino acid position 113.
Figure 64: Alignment of WT and mutant RNA polymerase

No frameshift occurred as a result of the SNPs
8.7.0 Appendix 7: CFU counts for CDB1 infected animals treated with phage

Figure 65: Total CFU and spore count recovered from animals that succumbed to CDI
All animals that were challenged with CDB1 eventually succumbed to infection, with the exception of treated hamster 2 which was culled at 57 hours. The caecum and colon of the animals was aseptically removed and washed to recover the vegetative cells and spores. There was no observable difference between the phage treated (H1, H2, H3) and untreated (H4, H5) hamsters.

Figure 66: Average total CFU and spore count for the treated and untreated animals
8.8.0 Appendix 8: Average bacterial resistance to phage generated by co-evolution.

Figure 67: Average bacterial resistance to phage across all transfers AIU-T9 MOI 10 in BHI media

Phage isolated from across all transfers was used to infect bacteria from each transfer for each of the 12 replicates. The average resistance to phage at the first transfer was between 20% and 30%. This increased exponentially as transfers progressed with resistance in transfer 15 peaking at 90%
Figure 68: Average bacterial resistance to phage across all transfers Alu-ST MOI 10 in GMGM

Phage isolated from across all transfers was used to infect bacteria from each transfer for each of the 12 replicates. The average resistance to phage at the first transfer was between 20% and 30%. This increased exponentially as transfers progressed with resistance in transfer 15 peaking at 90%
Figure 69: Average bacterial resistance to phage across all transfers AlU-T9 MOI 10 in GMGM

Phage isolated from across all transfers was used to infect bacteria from each transfer for each of the 12 replicates. The average resistance to phage at the first transfer was between 20% and 30%. This increased exponentially as transfers progressed with resistance in transfer 15 peaking at 90%
8.9.0 Appendix 9: VOC whose release was not affected by phage infection

Figure 70: VOCs whose release was not affected by phage infection
Figure 71: Growth curves of CDB1 and CD1342

The growth patterns of CD B1(A) and CD 1342(B) were recorded in order to allow for the determination of the optimal time for phage infection. The optical density at 550 nm (blue line) and CFU per ml (red line) was recorded. Both strains showed similar growth patterns with a lag phase of ~2 hours, followed by a log phase lasting about 4 hours before entering into stationary. In both cases, the number of CFU did not rise above $5 \times 10^8$ per ml. This growth profile is consistent with other *C. difficile* strains previously investigated.
8.11.0 Appendix 10 : The cost and effects of resistance for bacteria

Acquiring resistance to phage is essential for the survival of the cell, however acquiring such resistance would come at a cost of fitness, which would depend on the manner of which resistance was acquired.

The growth dynamics of bacteria that had undergone antagonistic coevolution (phage ST, MOI 10, BHI) were compared to the growth dynamics of bacteria that had not undergone antagonistic coevolution. Bacteria collected from the 5th, 10th and 15th transfer were characterised by bacterial growth curves. Bacteria from the three transfers showed reduced ability to grow compared to the bacteria that had not undergone antagonistic coevolution. In addition bacteria from the later transfers had a reduced ability to grow compared to bacteria from the earlier transfers (Fig 48). Although this reduction in growth was not statistically significant (t-test, p value >0.5) it was nevertheless consistent.

This reduction in growth, the cost of acquiring resistance to phage, the primary cause of which would appear to be the accumulation of SNPs in several proteins mainly associated with the surface membrane of the phage exposed cell. As the results already shown indicate, some proteins have amino acid substitutions that could have altered the active site of the protein or have stop codons inserted in the middle of the protein which would lead to a shorter protein. Both cases can adversely affect the function of the protein leading to the reduced growth observed. There is also the possibility that the lysogeny of the phage used could also cause a reduction in growth. However, this needs to be investigated further with bacteria that have been proven to be lysogens.
8.11.0 Appendix 11: Expansion of host range

Attempts were made to expand the host range of newly isolated and existing phages. Several methods to do this were attempted.

Over 300 plaques from each of the phages used, the eight novel phages isolated in this study as well as the phages ΦX2 and ΦCDHM1, were picked from a lawn of their propagating host and after being dissolved in BHI, were spotted on a lawn of a resistant host. With the exception of phage ΦST, none of the phages tested was able to infect a new host. Phage ΦST was able to infect one new strain belonging to ribotype 078. It was previously observed to be able to infect one other strain within the same ribotype.

The same principle was tested in liquid culture where a phage was used to infect an exponentially growing culture of its propagating strain, harvested and then used to infect the exponentially growing culture of a resistant host. The harvested phage was then used to infect a lawn of the resistant host. However, none of the phages showed an expanded host range either.

A stock of the phage was also used to infect an exponentially growing culture that contained two bacterial hosts: the propagating and a resistant one. After harvesting the phage across different time points (3 hours, 8 hours and 24 hours) it was tested on lawns of the resistant host to see if an expansion of host range occurred. None of the phages showed an expanded host range.

The phages were also passaged in a co-culture of the propagating and resistant host with the phage collected at each stage and tested against a lawn of the resistant host. Once again however none of the phages showed an expanded host range.
Figure 72: Growth dynamics of wild type CDAIU and CDAIU that underwent coevolution with phage

Bacteria that had been previously exposed to phage had a reduced ability to grow compared to bacteria that had not been exposed to phage. Bacteria from the 5th, 10th and 15th transfers showed reduced amount of growth compared to the wild type bacteria. Every successive transfer also showed reduced amount of growth compared to the previous one. However this difference in growth between the samples was no statistical significand (t-test, p value >0.5)
Bibliography


Dredi, L., Tankovic, J. & Petit, J.-C., 2004. CdeA of Clostridium difficile, a new multidrug efflux...


pp.390–402.


Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C. & Doudna, J. a, 2014. DNA interrogation by


