Expanding the Genome Editing Repertoire in *Clostridium difficile* for Improved Studies of Sporulation and Germination

Patrick Samuel Ingle, BSc. (Hons)

Thesis Submitted to the University of Nottingham for the Degree of

Doctor of Philosophy

September 2017

Abstract

Clostridium difficile is an anaerobic, Gram-positive, endospore-forming, pathogenic bacterium which is the leading cause of antibiotic-associated diarrhoea, and causes a significant burden to healthcare facilities and communities, worldwide. Bacterial endospores are one of the most resilient forms of life, able to withstand exposures to wet-heat, desiccation, UV radiation, oxygen, and some disinfectants, which would otherwise kill the vegetative cell form. Thus, endospores of *C. difficile* are able to persist in the environment and contaminate surfaces within healthcare settings. Once ingested, these spores pass into the anaerobic lower intestines and in susceptible individuals find favourable conditions in which to germinate, generating the toxin-producing vegetative cells responsible for *C. difficile* associated disease. Consequently, spores are the infectious agent of this disease and both sporulation and germination processes are essential for disease. Whilst these processes have been well studied in *Bacillus subtilis*, it is only recently, with the development of appropriate reverse genetics tools for clostridia, that the mechanisms of sporulation and germination have begun to be described for *C. difficile*. This study uses the currently available mutagenesis tools of ClosTron and allelic exchange to generate mutant spores lacking spore-specific proteins, and through a range of assays characterises the sporulation, germination and resistance properties of these mutants, to understand the roles of these proteins in *C. difficile* endospores. Furthermore, these genetics tools are established in a novel *C. difficile* strain with beneficial properties for studying the processes of sporulation and germination *in vitro*. Finally, this study establishes CRISPR/Cas9 genome editing in *C. difficile* for the first time to overcome the major pitfalls associated with the previously available reverse genetics tools. This mutagenesis method was found to offer fast, highly efficient genome editing of two different *C. difficile* strains and will be the method of choice for future studies in *C. difficile*.

Acknowledgements

I wish to thank my supervisors Professor Nigel Minton and Dr Sarah Kuehne for giving me the opportunity to study towards this PhD and their support and mentoring throughout the duration of my studies. I would also like to thank the Swiss National Science Foundation (SNSF) for providing the funding for this research and our project collaborators at the University of Bern and École Polytechnique Fédérale de Lausanne for their support and discussions throughout this project. Thank you also to all the members of the SBRC for their advice and for creating such an enjoyable working environment. Finally, I wish to thank all my friends and family for their support, particularly during the write-up period.

Table of Contents

Abstract 1
Acknowledgements 2
List of Figures 9
List of Tables 11
Abbreviations 12
Declaration 14
1 Introduction 16
1.1 *Clostridium difficile* 17
1.1.1 An emergent pathogen 17
1.1.2 *C. difficile* associated disease 20
1.1.3 Virulence factors 21
1.2 Endospores of *C. difficile* 23
1.2.1 Sporulation 24
1.2.2 Endospore structure 26
1.2.3 Resistance properties of bacterial endospores 28
1.3 Germination of endospores 29
1.3.1 Germination of *B. subtilis* endospores 29
1.3.2 Germination of *C. difficile* endospores 30
1.4 Current tools for precise genome editing in *C. difficile* 34
1.4.1 ClosTron insertional mutagenesis 34
1.4.2 Allelic exchange mutagenesis 37
1.5 CRISPR/Cas9 genome editing 39
1.5.1 CRISPR/Cas9 systems for prokaryotic adaptive immunity 39
1.5.2 CRISPR/Cas9 systems for genome editing 42
1.6 Aims of this project 44
2 Materials and Methods 46
2.1 Materials 46
2.2 Bacterial Strains 46
2.3 Plasmids 48
2.4 Oligonucleotide Primers 51
2.5 Bioinformatics Tools 54
2.5.1 Sequence data analysis 54
2.5.2 Plasmid maps 54
2.5.3 Identification of promoter and terminator sequences 54
2.5.4 BLAST 55

2.5.5 Sequence alignments 55
2.5.6 Oligonucleotide design 55
2.5.7 Whole genome sequencing analysis 56
2.5.8 Data analysis 56
2.6 General Microbiological Techniques 57
2.6.1 Growth media 57
2.6.2 Growth conditions 58
2.6.3 Growth supplements 58
2.6.4 Strain storage 58
2.6.5 Measurements of bacterial growth 59
2.6.6 Preparation of electro-competent *E. coli* 60
2.6.7 Transformation methods 60
2.6.8 Preparation of *C. difficile* spores 61
2.6.9 Measuring *C. difficile* colony formation after heat treatment 62
2.6.10 Measuring spore titres using phase-contrast microscopy 63
2.6.11 De-coating *C. difficile* spores 63
2.6.12 Heat-treatment of *C. difficile* spore preparations 63
2.6.13 UV-treatment of *C. difficile* spore preparations 64
2.6.14 Ethanol treatment of *C. difficile* spore preparations 64
2.6.15 Measurements of the Dipicolinic acid content of *C. difficile* spores 64
2.6.16 Assessment of *C. difficile* cytotoxicity against Vero cells 65
2.6.17 Assessment of bacterial motility 66
2.7 General Molecular Techniques 66
2.7.1 Genomic DNA extraction from *Clostridium difficile* 66
2.7.2 Extraction of plasmid DNA from *E. coli* 68
2.7.3 Quantification of DNA preparations 68
2.7.4 Polymerase Chain Reaction amplification of DNA 68
2.7.5 SOEing PCR 69
2.7.6 Colony PCR 70
2.7.7 Restriction endonuclease digestion of DNA 70
2.7.8 De-phosphorylation of DNA 70
2.7.9 Agarose gel electrophoresis 71
2.7.10 DNA extraction from agarose gels and reaction mixtures 71
2.7.11 DNA ligation 71
2.7.12 Dialysis of DNA ligation products 72
2.7.13 Nucleotide sequencing 72

2.8 ClosTron mutagenesis 72
2.8.1 Intron retargeting for pMTL007C-E2 72
2.8.2 Synthesis of retargeted pMTL007C-E2 73
2.8.3 Generation of insertional mutants in *C. difficile* 73
2.8.4 Confirmation of single Group II intron insertion 74
2.9 Allele Coupled Exchange 75
2.9.1 Generation of DH1916Δ*pyrE* 75
2.9.2 Correction of *pyrE* negative strains 75
2.10 Allelic Exchange Mutagenesis 76
2.10.1 Construction of knockout cassettes 76
2.10.2 Construction of allelic exchange vectors 77
2.10.3 Generation of in-frame deletion mutants via allelic exchange 78
2.10.4 Complementation of *C. difficile* 630*ΔermΔpyrEΔspoVA* 79
2.11 CRISPR/Cas9 mutagenesis 79
2.11.1 Identification of sgRNA seed sequences within target sequences 79
2.11.2 Selection of sgRNA seed sequences for CRISPR/Cas9 genome editing 80
2.11.3 Construction of CRISPR/Cas9 genome editing vectors 80
2.11.4 Generation of in-frame deletion mutants via CRISPR/Cas9 genome editing 81
3 Elucidating the role of SpoVA proteins in *C. difficile* 630Δ*erm* 83
3.1 Introduction 83
3.1.1 Aims of this project 87
3.2 The *C. difficile spoVA* operon 87
3.2.1 630Δ*erm spoVA* operon structure 87
3.2.2 Analysis of the 630Δ*erm* SpoVA protein sequences 88
3.3 Generation of *C. difficile* 630Δ*erm spoVA* insertional mutants using ClosTron 90
3.3.1 Confirmation of single insertion of the Group II intron into 630Δ*erm* 92
3.4 Characterisation of *C. difficile* 630Δ*erm*::*spoVA* insertional mutants 93
3.4.1 Development of heat-resistant CFU over five days 94
3.4.2 Spore titres after five days 96
3.5 Generation of in-frame *spoVA* deletion mutants using Allelic Exchange 98
3.5.1 Complementation of the 630Δ*ermΔpyrEΔspoVA* mutant 103
3.6 Characterisation of in-frame deletion *spoVA* mutant and complemented strains 105
3.6.1 Development of heat-resistant CFU over five days 105
3.6.2 Five-day spore titres 107
3.6.3 Lysozyme treatment of de-coated spores 108
3.7 Obtaining pure, countable spore preparations 109

3.7.1 Sonication of spore preparations 110
3.7.2 Spore separation using Tween-80 114
3.8 Assessing the dipicolinic acid content of *spoVA*$^-$ spores 117
3.8.1 Spore calcium content 117
3.8.2 Spore DPA content measurements using Terbium chloride fluorescence 120
3.9 Generation of single s*poVA* deletion mutants via complementation 123
3.9.1 Construction of dual *spoVA* complementation vectors 123
3.9.2 Generation of strains lacking single *spoVA* genes 127
3.10 Characterisation of strains lacking single *spoVA* genes 130
3.10.1 Growth in BHIS broth over 24 hours 130
3.10.2 Colony formation after heat-treatment 131
3.10.3 Resistance of *C. difficile* spores to wet-heat 132
3.10.4 Resistance of *C. difficile* spores to UV radiation 135
3.10.5 Resistance of *C. difficile* spores to ethanol 138
3.11 Discussion 140
3.12 Summary 144
4 Implementing the gene editing roadmap in *C. difficile* DH1916 146
4.1 Introduction 146
4.1.1 Aims of this study 148
4.2 Comparative studies between R20291 and DH1916 149
4.2.1 Growth in BHIS broth 149
4.2.2 Colony formation after heat-treatment 150
4.2.3 Spore titres after five days 151
4.2.4 Assessment of spore clumping following purification 152
4.2.5 Cytotoxicity against Vero cells 153
4.3 Assessment of conjugation efficiency into DH1916 154
4.3.1 Conjugation efficiency into DH1916 from CA434 donor cells 154
4.3.2 Effect of heat-treatment on conjugation efficiency into DH1916 156
4.3.3 Restriction-modification systems in DH1916 158
4.3.4 Conjugation efficiency from NEB sExpress donor cells 160
4.4 Generation of a DH1916Δ*pyrE* strain 162
4.4.1 Confirmation of 5-Fluoroorotic acid susceptibility 162
4.4.2 Construction of a DH1916 *pyrE* ACE vector 163
4.4.2 Generation of DH1916Δ*pyrE* 164
4.5 Establishing allelic exchange in DH1916 166
4.5.1 Construction of allelic exchange vectors targeting *spo0A* and *cspC* 166

4.5.2 Generation of sporulation/germination control strains in DH1916 167
4.5.3 Confirmation of sporulation/germination deficient phenotypes 170
4.6 Genome resequencing of DH1916Δ*pyrE* strains 172
4.7 Discussion 175
4.8 Summary 180
5 Establishing CRISPR/Cas9 genome editing in *C. difficile* 182
5.1 Introduction 182
5.1.1 CRISPR/Cas9 genome editing in clostridia 182
5.1.2 Aims of this study 183
5.2 Preliminary testing of CRISPR/Cas9 systems in *C. difficile* 184
5.2.1 CRISPR/Cas9 mediated mutagenesis of *spo0A* in *C. difficile* 630 186
5.2.2 Replacement of the P_{araE} sgRNA promoter 189
5.3 Generation of *C. difficileΔerm** using CRISPR/Cas9 191
5.3.1 Erythromycin sensitive derivatives of *C. difficile* 630 191
5.3.1 Construction of vectors targeting CD2008 in *C. difficile* 630 194
5.3.2 Deletion of both *erm(B)* genes from *C. difficile* 630 195
5.4 Phenotypic characterisation of *C. difficile* 630Δ*erm** 197
5.4.1 Growth in BHIS broth 198
5.4.2 Colony formation after heat-treatment 198
5.4.3 Cytotoxicity against Vero cells 199
5.4.4 Cell Motility 200
5.4.5 Reversion to erythromycin resistance 201
5.5 Generation of a *pyrE* deficient *C. difficile* 630*Δerm** strain 203
5.5.1 Construction of CRISPR/Cas9 vectors targeting *pyrE* 204
5.5.2 Generation of *C. difficile* 630Δ*erm**Δ*pyrE* 205
5.5.3 Sanger sequencing of pPSI_cas_132 from colonies 19 and 21 207
5.6 Whole genome sequencing of *C. difficile* 630*Δerm** 208
5.7 CRISPR/Cas9 mutagenesis in *C. difficile* DH1916 210
5.7.1 Construction of CRISPR/Cas9 vectors targeting DH1916_*hsdR* 210
5.7.2 Generation of DH1916Δ*pyrE*Δ*hsdR* using CRISPR/Cas9 mutagenesis 213
5.7.3 Assessment of conjugation efficiency in DH1916Δ*pyrE*Δ*hsdR* 215
5.7.3.1 Determining the feasibility of suicide plasmid use in DH1916Δ*hsdR* 217
5.8 Discussion 219
5.8 Summary 224
6 General Discussion 226
6.1 Functions of the SpoVA proteins 226

6.2 Endospore Clumping 228
6.3 CRISPR/Cas9 genome editing 229
7 Bibliography 233

List of Figures

Figure 1-1 Reports of *C. difficile* positive faecal specimens in England, Wales and Northern Ireland from 1990-2014. 18
Figure 1-2 Deaths attributed to *C. difficile* infection in England and Wales, 2001-2012. . 18
Figure 1-3 Diagrammatic representation of the *C. difficile* pathogenicity locus. 22
Figure 1-4 Diagrammatic representation of the sporulation cycle in Gram-positive bacteria. 25
Figure 1-5 Ultrastructure of *C. difficile* endospores. 27
Figure 1-6 Model of bile acid mediated in vivo germination of *C. difficile* endospores. .. 33
Figure 1-7 Diagrammatic representation of the generation of erythromycin resistant insertional mutants using ClosTron mutagenesis. 35
Figure 1-8 Diagrammatic representation of the two-step allelic exchange mutagenesis system 38
Figure 1-9 Diagrammatic representation of the prokaryotic CRISPR/Cas adaptive immune systems..... 41
Figure 1-10 Diagrammatic representation of CRISPR/Cas9-mediated genome editing. 43
Figure 2-1 Generation of allelic exchange knockout cassettes via SOE PCR. 77
Figure 3-1 Diagrammatic representation of the molecular structure of Dipicolinic acid. 84
Figure 3-2 Diagrammatic representation of *spoVA* operons present within Firmicutes. . 85
Figure 3-3 Diagrammatic representation of the 630Δ*erm spoVA* 'operon'. 88
Figure 3-4 PCR screening for insertion of the ClosTron group II intron into each *C. difficile* 630Δ*erm spoVA* gene 91
Figure 3-5 Southern Blot hybridisation using an intron-specific probe. 92
Figure 3-6 Diagrammatic representation of the phenotypic characterisation of *C. difficile* mutants defective in sporulation and/or germination 93
Figure 3-7 Development of heat resistant CFU of DH1916 strains over five days. 95
Figure 3-8 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy. 97
Figure 3-9 Diagrammatic representation of allelic exchange vector pMTL-YN3_*spoVA*. . 99
Figure 3-10 Screening PCR to confirm *spoVA* deletion and *pyrE* restoration. 102
Figure 3-11 Diagrammatic representation of the complementation vector pMTL-YN1C-*spoVA* 104
Figure 3-12 Screening PCR to confirm complementation of *spoVA* deletion mutants. 105
Figure 3-13 Development of heat resistant CFU of *C. difficile* strains over five days. 106
Figure 3-14 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy. 107
Figure 3-15 Outgrowth of de-coated spores on BHIS agar supplemented with lysozyme. 109
Figure 3-16 Identification of a suitable sonication protocol for spore separation. 111
Figure 3-17 Effects of heat-treatment on sonicated spore preparations. 113
Figure 3-18 Effects of Tween80 on spore clumping assessed under phase-contrast microscopy 115
Figure 3-19 CFU development following spore washing with PBS and Tween80. 116
Figure 3-20 STEM-EDS detection of C. difficile spore elemental composition..... 119
Figure 3-21 Total DPA content of spores measured using Terbium fluorescence..... 121
Figure 3-22 Standard curve of Terbium fluorescence with defined DPA concentrations. 122

Figure 3-23 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVACD*……125
Figure 3-24 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVACE*……126
Figure 3-25 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVADE*……127
Figure 3-26 PCR screening for dual complementation and *pyrE* repair of 630Δ*erm*Δ*pyrE*Δ*spoVA*……129
Figure 3-27 Growth of *C. difficile* strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}……130
Figure 3-28 Development of heat resistant CFU of *C. difficile* 630 strains over five days.132
Figure 3-29 Colony development after exposure of *C. difficile* spores to wet heat……133
Figure 3-30 Heat-resistant CFU development in various spore suspension media……135
Figure 3-31 Resistance of 630Δ*erm* spores to various amounts of UV radiation……136
Figure 3-32 Resistance of 630Δ*erm spoVA* spores to UV radiation……138
Figure 3-33 Resistance of 630Δ*erm spoVA* spores to ethanol treatment……139
Figure 4-1 Growth of *C. difficile* strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}……149
Figure 4-2 Development of heat resistant CFU of *C. difficile* strains over five days.……150
Figure 4-3 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy.……151
Figure 4-4 Assessment of spore clumping in R20291 and DH1916 spore preparations.153
Figure 4-5 In vitro cell cytotoxicity of *C. difficile* strains against Vero cell cultures.……154
Figure 4-6 Conjugation efficiency into DH1916 from CA434 donor cells.……156
Figure 4-7 Effect of heat-treatment on conjugation efficiency into DH1916.……157
Figure 4-8 Identification of Restriction-Modification (RM) systems within *C. difficile* R20291……159
Figure 4-9 Conjugation efficiency into DH1916 from NEB sExpress donor cells……161
Figure 4-10 PCR screening for the generation of DH1916Δ*pyrE* strains.……165
Figure 4-11 Colony PCR screening for *C. difficile* DH1916 in-frame deletion mutants.……169
Figure 4-12 Development of heat resistant CFU of DH1916 strains over five days.……170
Figure 4-13 Spore titres of DH1916 strains enumerated by phase-contrast microscopy.171
Figure 5-1 Diagrammatic representation of the ORFs present on the *C. difficile* strain 630 mobilizable transposon Tn 5398.……192
Figure 5-2 Diagrammatic representation of the construction of the CRISPR/Cas9 vectors used in this study.……185
Figure 5-3 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_011.……188
Figure 5-4 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_111.……190
Figure 5-5 Colony PCR screening for *C. difficile* 630Δ*erm** mutants following CRISPR/Cas9 mutagenesis.……196
Figure 5-6 Growth of *C. difficile* 630 strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}……198
Figure 5-7 Development of heat resistant CFU of *C. difficile* 630 strains over five days.199
Figure 5-8 In vitro cell cytotoxicity of *C. difficile* strains against Vero cell cultures.……200
Figure 5-9 Motility of *C. difficile* 630 strains after 48 hours on 2YTG plates……201
Figure 5-10 Reversion of erythromycin sensitive *C. difficile* 630 derivatives to resistant phenotypes following repeated subculture.……202

Figure 5-11 Colony PCR screening for *pyrE* truncated 630Δ*erm** strains following CRISPR/Cas9 mutagenesis. 206
Figure 5-12 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_141. 212
Figure 5-13 Colony PCR screening for *hsdR* deletion from DH1916Δ*pyrE* strains following CRISPR/Cas9 mutagenesis. 215
Figure 5-14 Effect of hsdR on conjugation efficiency into DH1916. 216
Figure 5-15 Diagrammatic representation of pMTL-SC1 constituents. 218

List of Tables

Table 2-1 List of bacterial strains used in this study. 48
Table 2-2 List of plasmids used in this study. 51
Table 2-3 List of oligonucleotides used in this study. 54
Table 2-4 Growth media components used in this study. 57
Table 2-5 Growth media supplements used in this study. 59
Table 2-6 Typical components of a PCR amplification mixture used in this study. 69
Table 2-7 Typical thermocycling conditions for DNA amplification via PCR. 69
Table 3-1 Homology of *spoVA* genes within selected Firmicutes. 89
Table 4-1 Gram-positive replicons in the pMTL80000 modular vector series. 155
Table 4-2 Homologues of R20291 Restriction-Modification system ORFs present in DH1916. 160
Table 4-3 Susceptibility testing of DH1916 to 5-FOA. 163
Table 4-4 Alignments of R20291 and DH1916 *pyrE* nucleotide sequences. 164
Table 4-5 Resequencing analysis of DH1916Δ*pyrE* strains. 173
Table 4-6 Additional variants identified following resequencing of DH1916Δ*pyrE* 174
Table 5-1 sgRNA seed sequences identified within CD12140 (*spo0A*). 187
Table 5-2 sgRNA seed sequences identified within CD2008. 195
Table 5-3 Efficiency of each sgRNA in the generation of *C. difficile* 630Δ*erm** via CRISPR/Cas9 mutagenesis. 197
Table 5-4 sgRNA seed sequences identified within the *pyrE* target sequence. 204
Table 5-5 Efficiency of each gRNA in the generation of *C. difficile* 630Δ*erm**Δ*pyrE* via CRISPR/Cas9 mutagenesis. 207
Table 5-6 Resequencing analysis of *C. difficile* 630 (CRG856) and 630Δ*erm** strains. 208
Table 5-7 Protein alignments of CD630_26850 and CD630_34900 gene products against *B. subtilis*. 210
Table 5-8 sgRNA seed sequences identified within the *hsdR* target sequence. 212
Table 5-9 Conjugation efficiencies using CRISPR/Cas9 vectors. 213
Table 5-10 Transposon mutagenesis efficiency from the suicide vector pMTL-GL15 into DH1916Δ*pyrE* and DH1916Δ*pyrE*Δ*hsdR*. 219
Table 5-11 Editing efficiencies reported using CRISPR/Cas9 genome editing systems in various organisms. 220

Abbreviations

°C	Degrees Celsius
µg	Microgram
µl	Microlitre
ACE	Allele-Coupled Exchange
BHIS	Brain Heart Infusion Supplemented
BLAST	Basic Local Alignment Search Tool
bp	base pair
CDAD	*Clostridium difficile* Associated Disease
CDI	*Clostridium difficile* Infection
CDMM	*Clostridium difficile* Minimal Media
CFU	Colony Forming Units
CSPD	Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
DPA	Pyridine-2,6-dicarboxylic acid; Dipicolinic Acid
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
FOA	Fluoroorotic Acid
g	gram
g	Gravity
kb	Kilobase pair
KEGG	Kyoto Encyclopaedia of Genes and Genomes
kV	Kilovolt
LB	Luria-Bertani
M	Molar
ml	millilitre
MOPS	4-Morpholinepropanesulfonic Acid
NEB	New England Biolabs
OD	Optical Density

ORF	Open Reading Frame
PBS	Phosphate-buffered Saline
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
SOE	Splicing by Overlap Extension
SSC	Saline-sodium citrate
TAE	Tris-Acetate-EDTA
UV	Ultra-violet

Declaration

Unless otherwise stated, the work presented in this thesis is my own. No part has been submitted for another degree at the University of Nottingham or any other institute of learning.

Chapter One

Introduction

1 Introduction

Members of the *Clostridium* genus, consisting of over 130 physiologically diverse species, are Gram-positive, rod-shaped bacteria which are obligately anaerobic and are capable of forming endospores. The vast majority of these species are non-pathogenic and many have been exploited for industrial applications, such as the production of ethanol from lignocellulosic biomass by *Clostridium thermocellum* (Akinosho *et al*., 2014), and the production of acetone, butanol and ethanol in ABE fermentation of sugars using *Clostridium acetobutylicum* (Sreekumar *et al*., 2015) and *Clostridium beijerinckii* (Qureshi & Blaschek, 2001). Other non-pathogenic clostridia are key commensal organisms within the human gut microbiota, comprising 10-40% of the total bacteria present (Lopetuso *et al*., 2013). Furthermore, *Clostridium butyricum* is marketed as a probiotic and has been shown to promote regulatory T-cell generation in the intestine and prevent infection with enterohaemorrhagic *Escherichia coli* O157:H7 (Cassir *et al*., 2016). The potential therapeutic use of *Clostridium sporogenes* spores as a delivery vehicle for cancer treatments is an active area of research based on findings that intravenously injected spores exclusively germinate within hypoxic tumours (Kubiak & Minton, 2015). However, clostridia have gained notoriety through the actions of a small number of species which are responsible for clostridia producing higher numbers of toxins than any other group of bacteria, leading to severe disease and mortality in both humans and animals (Popoff & Bouvet, 2009). Despite its use in the cosmetics industry, *Clostridium botulinum* produces a neurotoxin which is the most potent natural toxin ever discovered and is associated with food-poisoning and concerns over its use in bio-terrorism (Schantz & Johnson, 1992). *Clostridium perfringens* causes gastrointestinal disease in humans and animals, and is a common causative agent of the necrotizing infection and gas gangrene (Stevens & Bryant, 2002). However, perhaps of most

recent concern to healthcare providers worldwide is the enteropathogenic *Clostridium difficile*, the leading cause of antibiotic-associated diarrhoea and colitis in the developed world (Ghose, 2013). Two recent studies have proposed a phylogenetic reorganisation of the clostridia and subsequently renamed *Clostridium difficile* to *Peptoclostridium difficile* (Yutin & Galperin, 2013) and *Clostridioides difficile* (Lawson *et al*., 2016). For the purposes of clarity, until a consensus amongst the scientific community is reached, this study will refer to this organism as *Clostridium difficile*.

1.1 *Clostridium difficile*

1.1.1 An emergent pathogen

In 1935, a study of stool samples from healthy neonates identified a strictly anaerobic, Gram-positive, rod-shaped bacterium with subterminal, non-bulging, elongate spores (Hall & O'Toole, 1935). The authors named this bacterium *Bacillus difficilis*, due to difficulties encountered in isolating and culturing the organism, which they found to be pathogenic to guinea pigs. The first study to suggest a role for the then renamed *Clostridium difficile* in human disease was by Smith & King (1962), who isolated eight strains from patients suffering from a variety of illnesses. However, it was not until 1978 that *C. difficile* was identified as the causative agent of pseudomembranous colitis in humans (George *et al.*, 1978). In the following years, increasing numbers of *C. difficile* infections (CDIs) were reported worldwide, particularly in the developed world. Voluntary surveillance of positive *C. difficile* laboratory samples was introduced in England and Wales in 1990, whilst in 2007 mandatory reporting of all CDI cases in NHS Trusts in patients over the age of two years in England was introduced. During this period, the number of CDI cases reported increased from 1,193 in 1990 to a peak of 56,270 in 2007 (Figure 1.1). Both the number of death certificates mentioning CDI and identifying CDI as the underlying cause of death also

increased over this time, reaching a peak of 8324 and 4056 deaths respectively in England and Wales in 2007 (Figure 1.2).

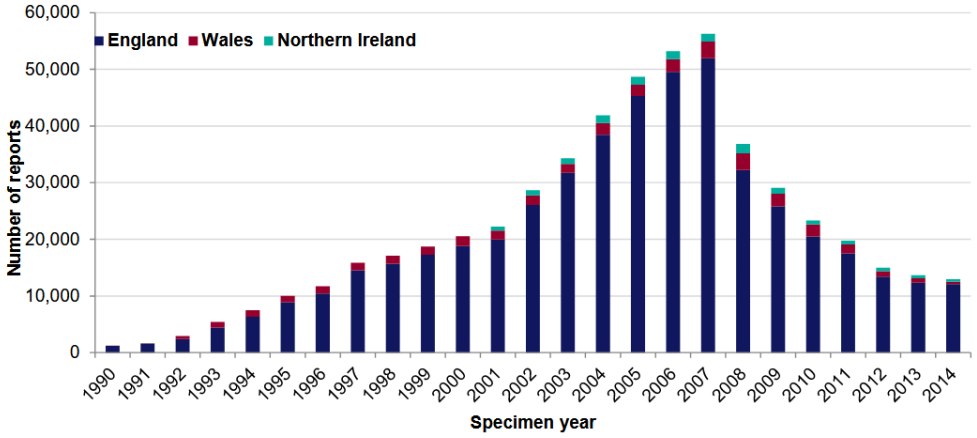


Figure 1-1 Reports of *C. difficile* positive faecal specimens in England, Wales and Northern Ireland from 1990-2014.

All voluntary or mandatory reporting of confirmed CDI cases in NHS Trusts in England, Wales and Northern Ireland from 1990-2014. Data from Northern Ireland included from 2001 onwards (Public Health England, 2015).

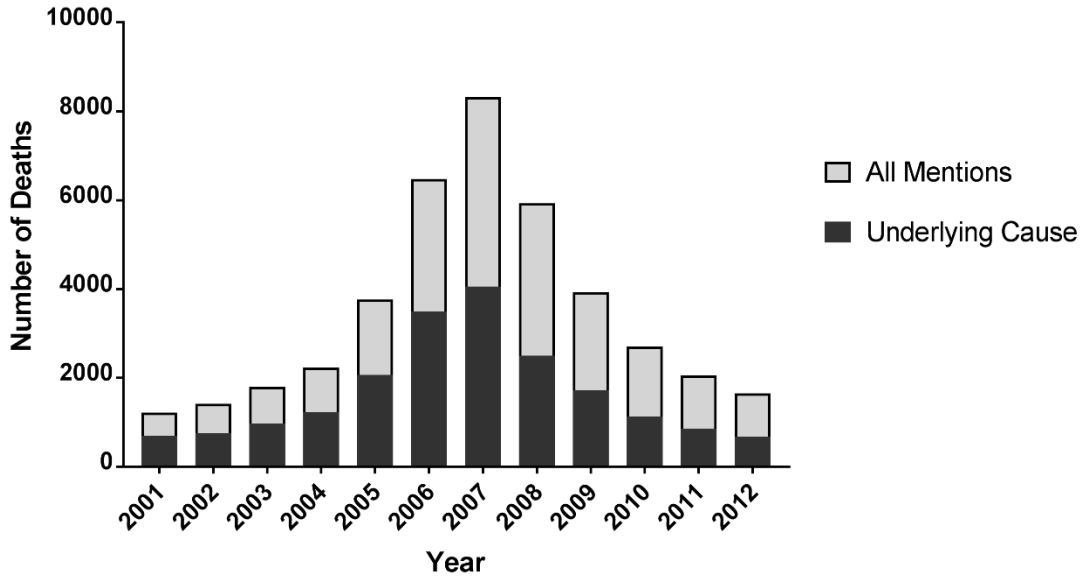


Figure 1-2 Deaths attributed to *C. difficile* infection in England and Wales, 2001-2012.

Proportion of death certificates in England and Wales identifying *C. difficile* as the underlying cause of death against all mentions per year. Graph produced using data from (Office for National Statistics, 2013).

The peak number of CDI cases in England and Wales in 2007 also coincided with observations of increased severity of CDAD symptoms, with more patients developing fulminant disease requiring surgical interventions, such as colectomy, leading to increased mortality (Cole & Stahl, 2015). The increased severity of CDAD symptoms and subsequent mortality rates have been attributed to the increased prevalence of the 'hypervirulent', epidemic, NAPI/B1/027 type strains isolated from CDI patients during major outbreaks worldwide during this time.

From 2007 onwards, enhanced infection control measures were introduced into hospitals and other healthcare settings in the UK in an attempt to curtail the rise of CDI cases. Such measures included the isolation of CDAD patients, alterations in antibiotic prescription policy and improved personal hygiene (Public Health England, 2015). In recent years, a decline in the number of CDI cases and deaths mentioning *C. difficile* have been reported, however CDI remains the number one cause of antibiotic-associated diarrhoea and a significant burden on healthcare providers. In the United States alone in 2011 there were 500,000 cases and around 29,000 deaths associated with CDAD (Lessa *et al*., 2015). Furthermore, the number of CDI cases has not declined as much as those of Methicillin-resistant *Staphylococcus aureus* (MRSA) infections, another nosocomial infection. A potential explanation for this is the increased prevalence of community-acquired CDI cases which accounted for 41% of cases in a recent study of 385 CDI cases (Khanna *et al*., 2012). As well as the human cost, CDAD also represents an economic burden with estimated annual costs of $2 billion in Europe and $3 billion in the US, through ward closures, extended lengths of stay within healthcare settings and increased treatment costs (Dubberke & Olsen, 2012).

1.1.2 *C. difficile* associated disease

Around 60-70% of new-borns are asymptomatically colonised by *C. difficile* and it has been hypothesised that these infants show apparent immunity to CDI due to an absence of toxin-receptors (Kelly & LaMont, 1998). Meanwhile, approximately 3% of the adult population carries *C. difficile* in their gastrointestinal (GI) tract asymptomatically due to colonisation resistance provided by the commensal gut microbiota. Under favourable conditions, *C. difficile* will proliferate and lead to *C. difficile* associated disease (CDAD). Symptoms of CDAD can range from mild, self-limited diarrhoea to severe disease, including pseudomembranous colitis, toxic megacolon, sepsis and in some cases ultimately death (Sayedy *et al.*, 2010). The main risk factor for CDAD is prior exposure to broad spectrum antibiotics and this link is well established. Shortly after the identification of *C. difficile* as the causative agent of pseudomembranous colitis in humans, it was shown to also cause disease in hamsters which had previously been treated with clindamycin (Chang *et al*., 1978). Other risk factors for CDAD include increased age, prolonged stays within healthcare settings, treatment with more than one antibiotic, prolonged duration of antibiotic course and feeding via nasogastric tube (Bignardi, 1998) (Henrich *et al*., 2009).

Routine treatment of CDAD occurs via the withdrawal of any previously supplied antibiotics either alone or with the administering of metronidazole or vancomycin antibiotics, or in rare cases the last resort antibiotic, fidaxomicin (Louie *et al*., 2011). Genome sequencing of *C. difficile* revealed several mobile genetic elements harbouring antibiotic resistance genes, yet so far, few reports have identified any acquisition to these front-line antibiotics. Another increasingly common treatment for CDAD is faecal microbiota transplantation, which aims to replace the gut microbiota lost during antibiotic treatment by supplying faecal bacteria from 'healthy' individuals to the CDI patient to restore colonisation resistance against *C. difficile*.

Successful treatment rates averaging 90% have been achieved using such methods (Rohlke & Stollman, 2012), although subsequent adverse effects have been reported in several instances (Wang *et al*., 2016).

The burden of CDI is worsened by the prevalence of relapse; the recurrence of CDI after the initial course of antibiotic treatment has been completed. Typically, CDI recurrence occurs in 20-25% of patients successfully treated after the initial episode (Figueroa *et al*., 2012). However, various studies have reported relapse rates of between 6-50% (Eyre *et al*., 2012), and evidence suggests that individuals experiencing one recurrent CDI episode are more likely to experience further such episodes. Such recurrences may occur via re-emergence of the original strain or re-infection with the original or a novel strain. Higher rates of relapse have been reported for the 'hypervirulent' PCR-ribotype 027 strains (Marsh *et al*., 2012).

1.1.3 Virulence factors

The principle virulence factors of *C. difficile* are two large clostridial toxins, the enterotoxic Toxin A (TcdA) and the cytotoxic Toxin B (TcdB). Whilst the relative importance of both toxins in CDAD was the source of some recent debate (Kuehne *et al*., 2010; Lyras *et al*., 2009), it is now accepted that both toxins are individually capable of causing disease. TcdA and TcdB have similar protein domain structures and biochemical function, with both possessing an N-terminal glycosyltransferase domain, a central putative membrane translocation domain and a C-terminal receptor binding domain. Upon their release from vegetative *C. difficile* cells, TcdA and TcdB enter host epithelial cells via receptor-mediated endocytosis, potentially through interactions with the host gp96 glycoprotein (Na *et al*., 2008). After exiting the endosome, these toxins are responsible for inactivation of the Rho-family GTPases, Rho, Rac and Cdc42 via monoglucosylation (Gerhard *et al*., 2008). This leads

to disruption of the actin cytoskeleton and tight junctions, resulting in excessive fluid accumulation, cell rounding and apoptosis (Voth & Ballard, 2005).

Expression of *tcdA* and *tcdB* occurs during late log and stationary growth phases of *C. difficile* from a 19.6 kb genomic locus referred to as the pathogenicity locus (PaLoc), shown in Figure 1.3. The PaLoc also contains three regulatory genes; *tcdR*, a sigma factor which up-regulates toxin transcription, *tcdC*, a putative anti-sigma factor which represses toxin transcription during exponential growth phases, and *tcdE*, which encodes a putative holin involved in lysis of the cytoplasmic membrane to allow toxin release (Tan *et al*., 2001). Variability in PaLoc sequences between strains forms the basis of toxinotyping, for example, some strains have been isolated which lack *tcdA*, whilst others display point mutations in catalytic domains. Non-toxigenic strains possess a 115 bp intergenic sequence instead of the PaLoc (Rupnik & Janezic, 2016). The PaLoc is located upon a mobile genetic element and transfer of the PaLoc from toxigenic to non-toxigenic strains via horizontal gene transfer has been reported (Brouwer *et al*., 2013).

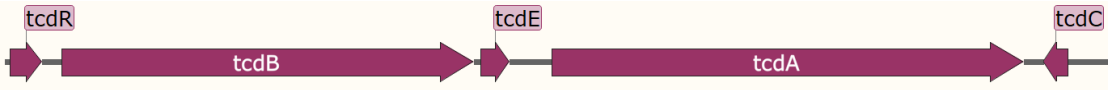


Figure 1-3 Diagrammatic representation of the *C. difficile* pathogenicity locus.
Depiction of a typical PaLoc from *C. difficile* R20291 showing genes encoding the two large clostridial toxins, *tcdA* and *tcdB*, and the three regulatory genes, *tcdR*, *tcdE* and *tcdC*.

In addition to the main virulence factors TcdA and TcdB, some strains including around 6% of clinical isolates also produce the binary toxin, *C. difficile* transferase (CDT). This toxin is related in structure and function to the *C. perfringens* iota toxin and contains two functionally distinct components, CDTa and CDTb. CDTa causes actin depolymerisation via its

ADP-ribosyltransferase activity, whilst CDTb is responsible for binding to a host lipoprotein receptor leading to translocation of CDTa into the cytosol (Papatheodorou *et al*., 2011). Studies challenging hamsters with *C. difficile* strains which produce CDT but neither TcdA nor TcdB have shown that CDT alone is not able to cause CDAD symptoms (Kuehne *et al*., 2014). Nonetheless, the presence of binary toxin is associated with more severe disease symptoms (Bacci *et al*., 2011).

Additional *C. difficile* virulence factors have been shown to assist in gut colonisation, which is essential for disease onset. Inactivation of flagella genes *fliC*, *fliD* and *flgE* in strain R20291 produced mutants which lacked flagella and showed decreased adherence to intestinal epithelial cells *in vitro* (Baban *et al*., 2013). Furthermore, ClosTron-mediated inactivation of CD0873 in strain 630Δ*erm*, which encodes a surface expressed lipoprotein, significantly reduced the ability of the mutant to bind to Caco-2 cells *in vitro* relative to the wildtype, suggesting CD0873 functions as an adhesin with a key role in host colonisation (Kovacs-Simon *et al*., 2014).

1.2 Endospores of *C. difficile*

One of the defining characteristics of clostridia is the ability of these bacteria to form endospores from their original vegetative cell state. This occurs in a process termed sporulation and produces an endospore which ensures long-term survival of the organism against agents which would otherwise have killed the vegetative cell form or where nutrients required for growth are absent. Bacterial endospores are extremely resistant structures and can survive exposure to oxygen, heat, alcohol, antibiotics and certain disinfectants (Setlow *et al*., 2002). Hence, when *C. difficile* spores are shed from the body in

faecal matter they can persist on contaminated surfaces within healthcare settings for prolonged periods. These spores can then lead to CDAD if they are inadvertently ingested, migrate into the anaerobic environment of the large intestine and germinate under appropriate stimuli to form the toxin-producing vegetative cells which can proliferate in susceptible individuals. Hence, spores are the infectious agent of CDAD and are an incredibly important stage of the *C. difficile* lifecycle.

1.2.1 Sporulation

Sporulation is the process by which bacteria produce endospores and has been extensively studied in the model spore-former, *Bacillus subtilis*, although recent studies have identified numerous differences in the regulation of sporulation between this organism and *C. difficile*. The exact environmental stimuli that triggers sporulation in *C. difficile* have not been identified but is hypothesised to be nutrient starvation, other stress factors, or the vegetative cell reaching a specific point in its growth cycle (Nawrocki *et al*., 2016). In *B. subtilis*, nutrient starvation leads to the phosphorylation of Spo0A which in turn inhibits transcription of the *abrB* repressor, allowing expression of *spoVG* and other genes essential for sporulation (Greene & Spiegelman, 1996). Once a threshold of phosphorylated Spo0A (Spo0A-P) is reached, a sporulation cascade is activated. Spo0A is the master transcriptional regulator of sporulation in all endospore-forming bacteria studied, and is essential for the formation of endospores in *C. difficile* (Heap *et al*., 2007). Generation of Spo0A-P in *C. difficile* is likely to occur via direct transfer of the phosphoryl group from a histidine kinase to Spo0A, due to the absence of *B. subtilis* phosphorelay protein homologues (Galperin *et al*., 2012). The sporulation process can be divided into six distinct stages, illustrated in Figure 1.4. First, (i) the asymmetrically positioned spore septum begins to form and isolate a newly replicated chromosome copy and a small section of the cytoplasm, then (ii) the septum

divides the cell into a larger mother cell compartment and a smaller pre-spore compartment. In stage (iii) this pre-spore compartment is engulfed by the mother cell compartment forming the forespore within a double membrane (Errington, 2010). Stage (iv) sees the synthesis of the spore peptidoglycan cortex between these two membranes, followed by sequential gene expression leading to the formation of a protein coat surrounding the outer membrane of the forespore (v). During this period, the mother cell compartment produces large amounts of pyridine-2,6-dicarboxylic acid (dipicolinic acid; DPA) which in *B. subtilis* is transported into the developing forespore by products of the *spoVA* operon (Tovar-rojo *et al*., 2002). During the final stage of sporulation (vi), spore maturation occurs leading to further dehydration of the spore and increased density of the spore coat, after which programmed cell death in the mother cell results in the release of the mature endospore into the surrounding environment (Higgins & Dworkin, 2013).

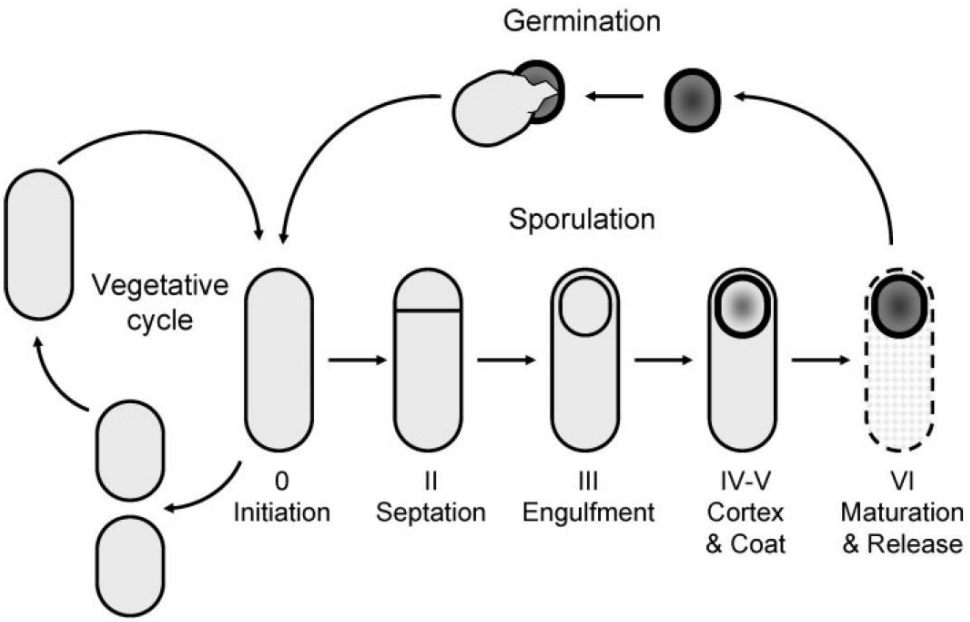


Figure 1-4 Diagrammatic representation of the sporulation cycle in Gram-positive bacteria.

Roman numerals denote the named morphologically distinct stage of sporulation depicted. Stage I, consisting of DNA replication and the asymmetrically positioned septum beginning to form, has been omitted. Adapted from (Errington, 2010).

A recent study by Fimlaid *et al*., (2013) found that 314 genes are induced during the sporulation process in *C. difficile*, the expression of which are primarily controlled temporally (different stages of sporulation) and spatially (different cell compartments) by sigma factors σ^E, σ^F, σ^G and σ^K. Given these four sigma factors are present in both *B. subtilis* and *C. difficile*, their functions were presumed to be largely similar in both organisms. Inactivation of any of these sigma factors in *B. subtilis* (Hilbert & Piggot, 2004) or *C. difficile* (Fimlaid *et al*., 2013) abrogates spore formation. However, the morphological phenotypes of the individual sigma factor mutants in *C. difficile* differed from equivalent *B. subtilis* mutants, in most cases, indicating differences in the regulation of sporulation between these two Firmicutes (Fimlaid *et al*., 2013).

1.2.2 Endospore structure

Bacterial endospores are multi-layered structures, as shown in Figure 1.5. The central core region, enclosed by the spore inner membrane, contains all the spore's DNA, RNA, ribosomes and the majority of its enzymes (Setlow, 2003). This core is relatively dehydrated compared to the rest of the spore, with water constituting around 30% of the core by weight but around 80% of the remaining spore regions (Knudsen *et al.*, 2016). The spore inner membrane has a similar lipid composition to that of growing cells but has very low permeability to small molecules, including water (Knudsen *et al.*, 2016). A germ cell wall surrounds the spore inner membrane and is composed of peptidoglycan as found in vegetative cells. Surrounding this germ cell wall is an altered spore-specific peptidoglycan layer, termed the spore cortex. Unique properties of this cortex peptidoglycan include reduced cross-linking and the substitution of every second muramic acid residue with a muramic-δ-lactam entity. The cortex is itself enclosed by the proteinaceous spore coat, which in *B. subtilis* spores consists of over 50 highly cross-linked polypeptides arranged into

distinct, ordered layers, but which does not pose a significant permeability barrier (Driks, 1999). Of these identified spore coat proteins in *B. subtilis*, fewer than one quarter have homologues in *C. difficile*, highlighting further differences between endospores of *C. difficile* and the model spore-former. Some bacterial endospores including *C. difficile* but not *B. subtilis*, contain an additional outermost layer termed the exosporium, consisting of a basal layer made from several different proteins from which hair-like protrusions, predominantly composed of the collagen-like glycoprotein BclA, extend (Sylvestre *et al*., 2002). The function of the exosporium is unclear, however studies showing decreased adherence of *C. difficile* spores to human Caco-2 cells following either trypsin treatment or sonication suggest the exosporium is a putative virulence factor (Paredes-Sabja & Sarker, 2012).

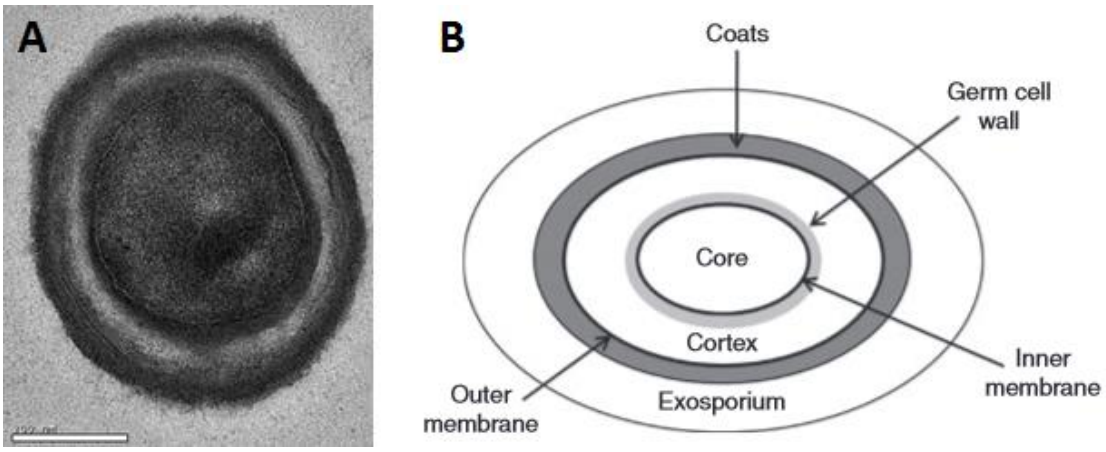


Figure 1-5 Ultrastructure of *C. difficile* endospores.

The ultrastructure of *C. difficile* endospores is shown via (A) an electron microscopy image of a *C. difficile* endospore (Permpoonpattana *et al*., 2011), and (B) a diagrammatic representation of the layers within a typical bacterial endospore (not to scale; Legget *et al*., 2012).

1.2.3 Resistance properties of bacterial endospores

Bacterial endospores are extremely resilient structures which function to protect the DNA located in the core from a variety of toxic chemical and physical agents, including wet and dry heat, UV radiation, ethanol, formaldehyde and several disinfectants (Setlow *et al*., 2002; Setlow, 2006). Many of the ultrastructure features of endospores identified in the previous section contribute to this highly resistant phenotype. Whilst the functions of its individual protein components are still being elucidated, together, the spore coat confers resistance to exogenous enzymes capable of degrading the spore cortex, to some chemicals, and protects against predation by protozoa (Klobutcher *et al*., 2006), but does not confer spore resistance to heat or radiation (Nicholson *et al*., 2000). The thick layer of spore-specific peptidoglycan comprising the spore cortex provides resistance against heat treatments and functions to maintain spore dormancy (Driks, 1999). Moving further inwards, the spore inner membrane contains lipid molecules with restricted mobility and provides a strong permeability barrier against foreign molecules including DNA-damaging chemicals (Young & Setlow, 2003). The reduced water content of the spore core is the major provider of resistance against wet-heat treatments (Setlow, 2006). This dehydrated spore core also contributes to enzyme inactivity and the metabolic dormancy of the endospore. The presence of DPA exclusively in the spore core, predominantly in a 1:1 chelate with calcium or other divalent cations, aids in this dehydration of the spore core and helps protect against UV radiation (Setlow *et al*., 2006). The major determinant for spore resistance to UV radiation is the saturation of DNA in the spore core with α/β-type small acid soluble proteins (SASPs). These proteins comprise 10-20% of the spore's total protein and, through binding to spore DNA ,dramatically change the structure and photochemistry of this DNA to provide such protection (Raju *et al*., 2006).

The precise mechanism for spore killing by wet heat is not clear, but is not due to DNA damage and often coincides with perforation of the spore inner membrane (Zhang *et al*., 2010). Spore killing by dry heat is due to the accumulation of DNA damage which overwhelms the *recA*-dependant DNA repair system. Furthermore, DNA damage is the cause of spore killing via exposure to chemicals such as formaldehyde and nitrous acid (Loshon *et al*., 1999), whilst strong acids and hydrogen peroxide kill spores via damage to the spore inner membrane causing spores to rupture (Setlow *et al*., 2002). DNA damage is also the method by which UV radiation kills spores with the most damaging wavelength of UV radiation for spores being 254 nm (Setlow, 2006).

1.3 Germination of endospores

Whilst bacterial endospores are metabolically dormant and can remain in this resistant state for many years, they are constantly ready to return to their vegetative cell state, in a process termed germination and defined as the irreversible loss of spore specific characteristics. The process of germination has been extensively studied in *B. subtilis* and shown to be initiated upon the binding of specific effectors, or 'germinants', to receptor proteins located in the spore inner membrane.

1.3.1 Germination of *B. subtilis* endospores

Germination is induced in *B. subtilis* endospores upon the binding of L-alanine, or a combination of L-asparagine, glucose, fructose and potassium ions, to germination receptors. These receptor proteins belong to the GerA family and homologues of these receptors have been found in all members of the *Bacillales* and *Clostridiales* orders studied thus far, except from in *C. difficile* (Paredes-Sabja *et al*., 2011). Following this, there is a lag

phase which can last from a few minutes to over 24 hours for a small proportion of 'superdormant' spores, prior to the commitment stage. Commitment to spore germination is defined as the point at which the addition of strong competitive inhibitors or acidification to pH 4.5 can no longer prevent germination completion. Within minutes of commitment the spore loses its wet heat resistance properties and releases monovalent cations (H^+, K^+ and Na^+) followed by all of the DPA from the spore core (Tovar-rojo *et al*., 2002). Stage II of the germination process begins with the degradation of the spore peptidoglycan cortex by the cortex lytic enzymes (CLEs) CwlJ and SleB. During this stage, the core expands and core water content returns to that present in vegetative cells, leading to the end of spore dormancy and the return of metabolism. Macromolecular synthesis then begins and the germinated spore enters outgrowth which leads to the creation of a vegetative cell. The above describes the nutrient-dependent germination of *B. subtilis* spores, however, exposure of these spores to exogenous DPA, dodecylamine and high pressures over 100 megaPascals has also been shown to induce germination of *B. subtilis* spores (Paidhungat *et al*., 2000), although these mechanisms are not likely to hold much relevance for *in vivo* spore germination.

1.3.2 Germination of *C. difficile* endospores

C. difficile endospores must undergo the same processes of cortex hydrolysis, DPA release and core hydration as described above for *B. subtilis* spores, but in some cases these events occur in a different order and are carried out by different proteins. *C. difficile* endospores uniquely lack any homologues to any of the GerA, GerB and GerK germinant receptor proteins, suggesting these spores respond to different germinants and/or utilise different germinant receptor proteins (Paredes-Sabja *et al*., 2011). The specific germinants which trigger germination vary between species. For example, *C. perfringens* spores germinate in

the presence of sodium ions and inorganic phosphate, *Bacillus megaterium* spores respond to L-proline, whilst purine ribonucleosides and various amino acids are co-germinants of *Bacillus anthracis* spores. Initially, it was demonstrated that various bile salts stimulated the germination of *C. difficile* spores (Wilson, 1983), and subsequently that glycine and the bile salt taurocholate are co-germinants of these spores *in vitro* (Sorg & Sonenshein, 2008). The importance of bile salts for endospore germination was further demonstrated by Giel *et al*. (2010), who showed that the addition of the bile salt binding resin, cholestyramine, to intestinal and cecal extracts from mice resulted in a drastic reduction in the number of *C. difficile* spores able to form colonies. However, the relationship between bile salts and spore germination is more complex, given findings that the bile salt chenodeoxycholate is a competitive inhibitor of germination via cholic acid derivatives (Sorg & Sonenshein, 2010), although this inhibitory effect of chenodeoxycholate has not been observed for all *C. difficile* strains (Heeg *et al*., 2012).

Bile acids are amphipathic molecules defined as 'saturated, hydroxylated C-24 cyclopentanephenanthrene sterols' which aid in the absorption of dietary lipids and lipid soluble vitamins (Ridlon *et al*., 2006). The primary bile acids cholic acid and chenodeoxycholic acid are produced from cholesterol in the liver and can be conjugated with glycine, or more commonly taurine, to produce bile salts with increased solubility (Hofmann, 1963). Each day, bile salts complete several cycles of enterohepatic circulation and up to 800 mg of these bile salts escape the active transport mechanisms involved in this process and enter the large intestine instead of being recycled back to the liver. In the large intestine, these bile salts can undergo biotransformation by certain members of the intestinal anaerobic microbiota. Bacteria possessing bile salt hydrolases, including *C. perfringens* and *Bacteroides fragilis*, can deconjugate the bile salts by cleaving the C-24 *N*-

acyl amide bond connecting the bile acid and amino acid (Coleman & Hudson, 1995). This reaction is a pre-requisite for 7α-dehydroxylation conversion of the primary bile acids cholic acid and chenodeoxycholic acid into the secondary bile acids deoxycholic acid and lithocholic acid, respectively. These secondary bile salts differentially affect *C. difficile* endospore germination with lithocholate being shown to inhibit germination whilst deoxycholate appears to stimulate germination yet also inhibit the growth of the subsequent vegetative *C. difficile* cells (Sorg & Sonenshein, 2008). Organisms capable of this 7α-dehydroxylation reaction, including *Clostridium scindens* and *Clostridium hiranonis*, transport unconjugated primary bile acids into the cytosol. Here, a series of sequential enzymatic reactions catalysed by gene products of the *bai* operon occur to generate secondary bile acids, which are then exported from the cell (Ridlon *et al*., 2006). These *in vitro* effects of bile acids on spore germination have been replicated *in vivo*, with Francis *et al*. (2013) demonstrating the importance of bile acid-mediated germination in establishing CDI in hamsters. This same study identified the germination-specific protease CspC as the germinant receptor involved in bile-acid mediated germination, since mutations in this gene either abrogated *C. difficile* spore germination in response to bile acids or altered the specificity of bile acid recognition by these spores. However, it should also be stated, that the endospores of some *C. difficile* strains have also been observed to germinate in rich (BHIS) medium lacking any bile salts (Heeg *et al*. 2012), potentially indicating a more complex or alternative method of spore germination.

From these recent studies, a model of *C. difficile* endospore germination has emerged, which is summarised below and in Figure 1.6. Highly resilient endospores persist in the environment and within healthcare settings until they are inadvertently ingested and migrate to the anaerobic lower intestine. In healthy individuals, commensal bacteria convert

primary bile acids into the secondary bile acids which are inhibitory to spore germination and/or outgrowth, and any spores which do complete germination and outgrowth are out-competed for nutrients by other commensals and therefore cannot proliferate. However, in susceptible individuals pre-treated with broad-spectrum antibiotics, the 7α-dehydroxylating bacteria which previously provided colonisation resistance are now depleted, providing *C. difficile* spores with favourable conditions in which to germinate. Furthermore, these resulting toxin-producing vegetative cells can proliferate, given the reduction in numbers of commensal bacteria competing for nutrients, and lead to CDAD.

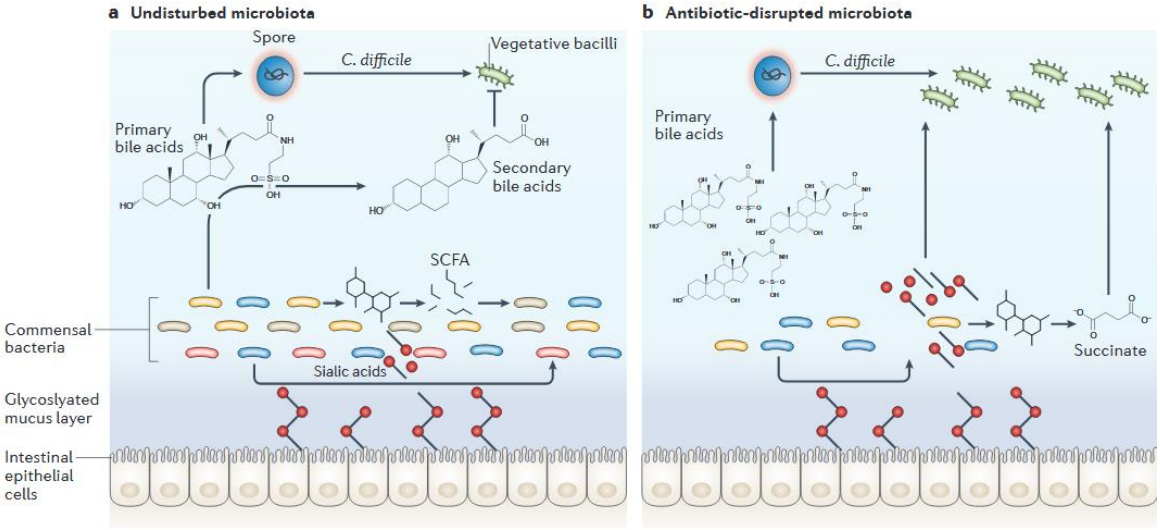


Figure 1-6 Model of bile acid mediated *in vivo* germination of *C. difficile* endospores. Diagrammatic representation of the role of commensal gastrointestinal bacteria on the regulation of *C. difficile* endospore germination. A) The intact microbiota inhibits endospore germination by producing secondary bile acids, and fermentative commensals convert carbohydrates into short chain fatty acids (SCFAs) which are consumed by other commensals. B) Endospore germination and outgrowth can occur in the disrupted microbiota and vegetative *C. difficile* cells can proliferate and more-effectively compete for succinate and other nutrients (Abt *et al*., 2016).

1.4 Current tools for precise genome editing in *C. difficile*

Historically, our understanding of the precise functioning of genes within the *C. difficile* chromosome has been hindered by a lack of genetic tools for creating precise, targeted mutants. Within the last fifteen years, such tools have been developed and utilised to provide much of our current understanding of *C. difficile* gene function.

1.4.1 ClosTron insertional mutagenesis

The ClosTron insertional mutagenesis tool is based upon a mobile group II intron isolated from the *Lactococcus lactis ltrB* gene which can be engineered to target a specific site within a given gene via modifications to the intron RNA (Mohr *et al*., 2000; Heap *et al*., 2007). Mobile Group II introns are catalytic RNAs which self-excise from RNA transcripts, splice directly into double-stranded DNA and are then reverse transcribed by an intron-encoded protein (IEP) with reverse transcriptase activity (Mohr *et al*., 2000). The IEP also contains maturase, which promotes RNA splicing, and site-specific DNA endonuclease activities, both of which are required for successful translocation. Recognition of the DNA target site sequence occurs primarily via base pairing between the group II intron RNA and the DNA target. Thus, the group II intron can be retargeted to any given gene simply by altering the intron RNA (Mohr *et al*., 2000; Karberg *et al*., 2001). The ClosTron mutagenesis system, outlined in Figure 1.7, utilises a vector carrying the LtrA IEP and a retargeted group II intron. Delivery of the IEP on a plasmid containing a deficient Gram-positive replicon, ensures plasmid loss and removal of the IEP from the mutant strain following successful intron insertion, thus creating stable insertional mutants by preventing further mobility of the group II intron. These stable, insertionally inactivated mutants are selected on the basis of conferred erythromycin resistance, given the presence of a retrotransposition-activated

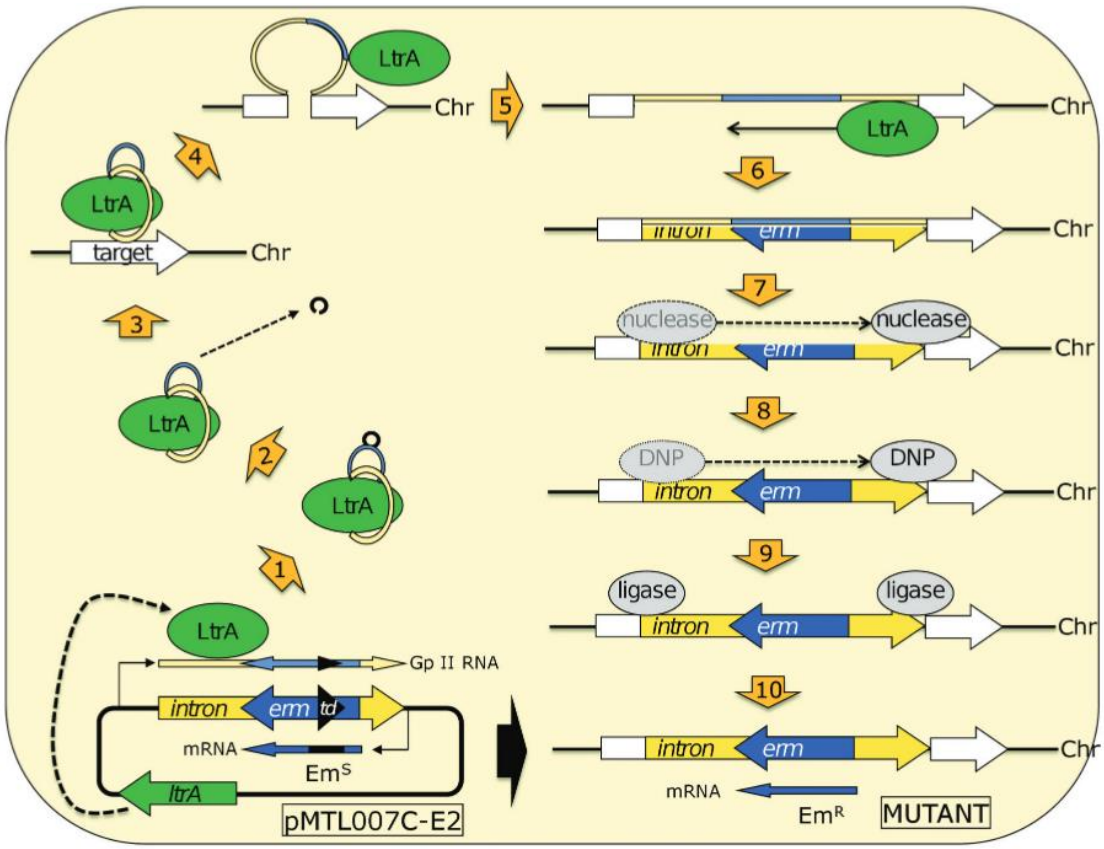


Figure 1-7 Diagrammatic representation of the generation of erythromycin resistant insertional mutants using ClosTron mutagenesis.

The ClosTron plasmid pMTL-007C-E2 containing *ltrA*, encoding the *Lactococcus lactis* group II intron encoded protein, and the group II intron (yellow arrow) with inserted *ermB* retrotransposition-activated marker (blue arrow) inactivated by the phage td group I intron (black arrow). (1) A ribonuclear protein complex forms following binding of the LtrA protein to the group II intron RNA transcript; (2) The td group I intron in its correct orientation self-excises and (3) the retargeted region of the group II intron recognises and binds to the chromosomal DNA target. (4) The multi-functional LtrA protein then nicks the DNA via endonuclease activity, (5) inserts the group II intron RNA and (6) synthesises the complementary DNA strand via reverse transcriptase activity. Host factors are then utilised for (7) degrading the RNA insert, (8) synthesising the opposite DNA strand via DNA polymerases (DNP) and (9) sealing the gaps using ligases. The result (10) is an erythromycin resistant mutant strain inactivated in the target gene via insertion of the retargeted group II intron (Heap *et al.*, 2007; Kuehne *et al*., 2011).

marker (RAM) based on the *ermB* gene, which is inserted into the group II intron and is itself inactivated by the presence of a group I intron. This group I intron self-excises from the RNA transcript in an orientation specific manner, restoring the functionality of the *ermB* gene concomitant with integration of the group II intron (Heap *et al*., 2007). In establishing the ClosTron mutagenesis tool in *C. difficile*, Heap *et al*., (2007) generated five insertionally inactivated mutant strains in the five chosen target genes (*spo0A*, *pyrF*, CD0153, CD0552 and CD3563), exceeding the total number of previously generated mutant strains of *C. difficile* at the time. This system has subsequently been used by many to investigate and improve our understanding of this organism, including the pathogenicity of both large clostridial toxins (Kuehne *et al*., 2010), the role of the cysteine protease Cwp84 in surface layer maturation (Kirby *et al*., 2009) and the role of histidine kinases in sporulation regulation (Underwood *et al.,* 2009). Advantages of the ClosTron mutagenesis tool include its relative ease of use and the speed at which mutants can be obtained (generally less than two weeks). However, there are some limitations to this system, including the insertional nature of the mutagenesis and the chromosomal integration of the RAM preventing further mutagenesis. Insertional mutagenesis for reverse genetic approaches is not ideal, since the observed mutant phenotype may be wholly or partially due to downstream effects on proximal genes instead of the insertionally inactivated target gene (Ciampi & Roth, 1988). Whilst the selection of insertional mutants using the antibiotic-resistance conferring RAM is advantageous for single mutants, it precludes the generation of strains with insertions in more than one target gene. Multiple insertionally inactivated strains can be generated via ClosTron plasmids lacking a RAM, but this increases the amount of mutant screening required to obtain the desired strain, thus slowing down the time taken to obtain mutants (Kuehne *et al*., 2011).

1.4.2 Allelic exchange mutagenesis

Within the last five years, two variant allelic exchange based mutagenesis methods have been exemplified in *C. difficile*, permitting both the generation of clean, in-frame deletion and large or single nucleotide polymorphism (SNP) mutants (Cartman *et al*., 2012; Ng *et al*., 2013). These methods, outlined in Figure 1.8, are based upon homologous recombination and utilise the chloramphenicol acetyltransferase gene, *catP*, as a positive selection marker and either a heterologous *codA* allele from *E. coli* or *pyrE* allele from *C. sporogenes* as the counter-selection marker.

Prior to conducting *pyrE*-mediated allelic exchange, it was necessary to delete the native *C. difficile pyrE* gene and this was done via another homologous recombination technique called allele-coupled exchange (Heap *et al*., 2012; Ng *et al*., 2013). The order of recombination events in ACE is largely controlled via the use of a longer homology arm (1200 bp) consisting of the region downstream of the *pyrE* gene, in which crossovers occur preferentially, and a second smaller homology arm (300 bp) consisting of an internal *pyrE* fragment. Thus, *pyrE* is not inactivated in most cells until the second recombination event has occurred. The resulting strain harbours a truncated, inactivated allele of the *pyrE* gene, which encodes orotate phosphoribosyl transferase, a key enzyme in the uracil biosynthesis pathway. Since *pyrE*-negative strains require uracil supplementation for growth and *pyrE*-positive strains convert the innocuous compound 5-fluoroorotic acid (5-FOA) into the toxic compound 5-fluoro-uridine monophosphate, both forms of the *pyrE* allele are positively or negatively selectable in appropriate media.

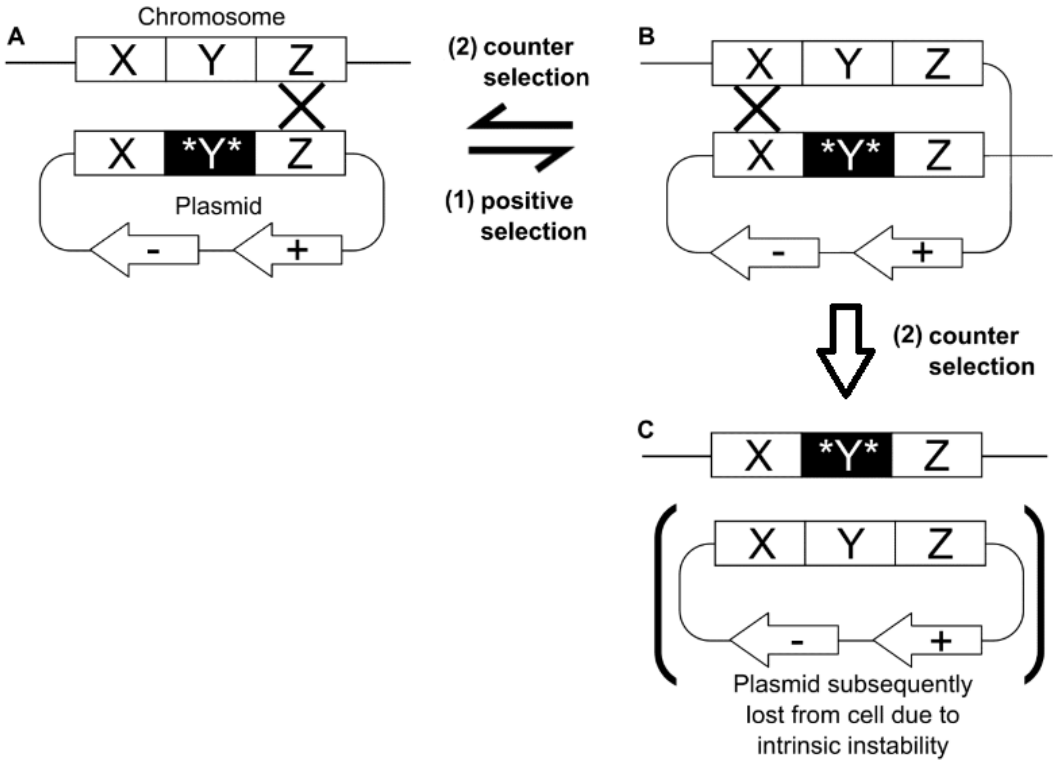


Figure 1-8 Diagrammatic representation of the two-step allelic exchange mutagenesis system.

A) Allelic exchange vectors harbouring a recombinant allele, in which the *Y* sequence differs from that of the chromosome (e.g. to generate a SNP) or is absent (to generate an in-frame deletion mutant), are conjugated into *C. difficile*. These vectors are replication deficient via the presence of an unstable Gram-positive replicon and also harbour a positive (+) selection marker (antibiotic resistance marker) and a counter-selection (-) marker. B) Single crossover recombinant clones resulting from homologous recombination at either X or Z, and resulting in plasmid integration into the chromosome, are enriched for using positive selection. C) Counter-selection is then used to select for clones in which a second crossover event has occurred at either the same or opposite site (X or Z) to that which occurred in (B), excising the plasmid from the chromosome and producing either wild-type revertant or double crossover recombinant clones. Adapted from (Cartman *et al.*, 2012).

For the creation of precise, in-frame deletions, a PCR-generated recombinant allele is produced with homology to chromosomal regions upstream and downstream of the region to be deleted. The appropriate allelic exchange vector harbouring this recombinant cassette, a deficient Gram-positive replicon and a heterologous *pyrE* allele from *C. sporogenes* is then transferred into a *pyrE*-negative strain to allow counter-selection in the presence of 5-FOA. Once the desired mutant has been isolated and plasmid loss confirmed, the *pyrE*-negative mutant background can be reverted to wildtype *pyrE* via ACE with the appropriate repair vector. This process of allelic exchange mutagenesis is advantageous, when compared with ClosTron mutagenesis, as it overcomes the issues of downstream polar effects which arise from insertional methods, can be used to create clean, in-frame deletion mutants or strains with point mutations, and allows the generation and selection of strains harbouring multiple deletions/genome alterations. However, there are several disadvantageous features to allelic exchange mutagenesis including the increased length of time taken to generate the desired modification and the increased risk of SNPs accumulating within the genome due to the increased numbers of passages required during selection steps.

1.5 CRISPR/Cas9 genome editing

1.5.1 CRISPR/Cas9 systems for prokaryotic adaptive immunity

Prokaryotes possess a number of defence mechanisms to protect against invading foreign DNA, including restriction-modification systems and toxin-antitoxin modules. Recently, another such method termed clustered regularly interspaced palindromic repeats (CRISPR) has been characterised. CRISPR sequences, and CRISPR-associated (Cas) genes, were identified within many diverse bacteria and archaea as whole genome sequencing became increasingly available (Jansen *et al*., 2002). These direct repeat sequences vary in size

between 21-37 bp and are interspaced with similarly sized non-repetitive sequences (spacers), all of which are flanked on one side by a 300-500 bp leader sequence. The *cas* genes, invariably located adjacent to the identified CRISPR loci, were shown to contain helicase and exonuclease motifs which led to the hypothesis that these genes are involved in gene expression control (Jansen *et al*., 2002). Key to elucidating the function of the CRISPR system were observations that the DNA sequences between the identical CRISPR repeats showed homology to invasive DNA sequences including bacteriophages and conjugative plasmids (Bolotin *et al*., 2005; Mojica *et al*., 2005). These results, together with findings that specific CRISPR spacer sequences correlated with phage resistant phenotypes in *Streptococcus thermophilus*, suggested the CRISPR system may function to protect cells against foreign, invasive DNA (Makarova *et al*., 2006). This was confirmed by Barrangou *et al*. (2007), who sequenced CRISPR loci from historical *S. thermophilus* strains and their phage-resistant derivatives and were able to identify the CRISPR spacer sequences responsible for resistance to specific bacteriophages. Furthermore, it was shown that these phage-resistance phenotypes could be altered by changes to the CRISPR spacer sequences, that CRISPR loci are able to incorporate phage DNA as novel spacers, and that *cas* genes were involved in this acquisition process and in providing the corresponding phage resistant phenotype (Barrangou *et al*., 2007). Subsequent work has characterised CRISPR systems in a diverse range of prokaryotes, such that *cas* genes have been found in around half of all sequenced bacterial genomes and almost all sequenced archaeal genomes (Boudry *et al*., 2015). Further work identified the protospacer adjacent motif (PAM) sequence which is dependent on the variant CRISPR/Cas system present and which determines the target site of interference (Mojica *et al*., 2009). These systems can be divided into two broad classes according to the type of nuclease which cleaves recognised foreign DNA; class 1 which use a multi-protein complex and class 2 which utilise a single protein. A summary of the CRISPR systems identified to date, and their mechanisms of action, is provided in figure 1.9.

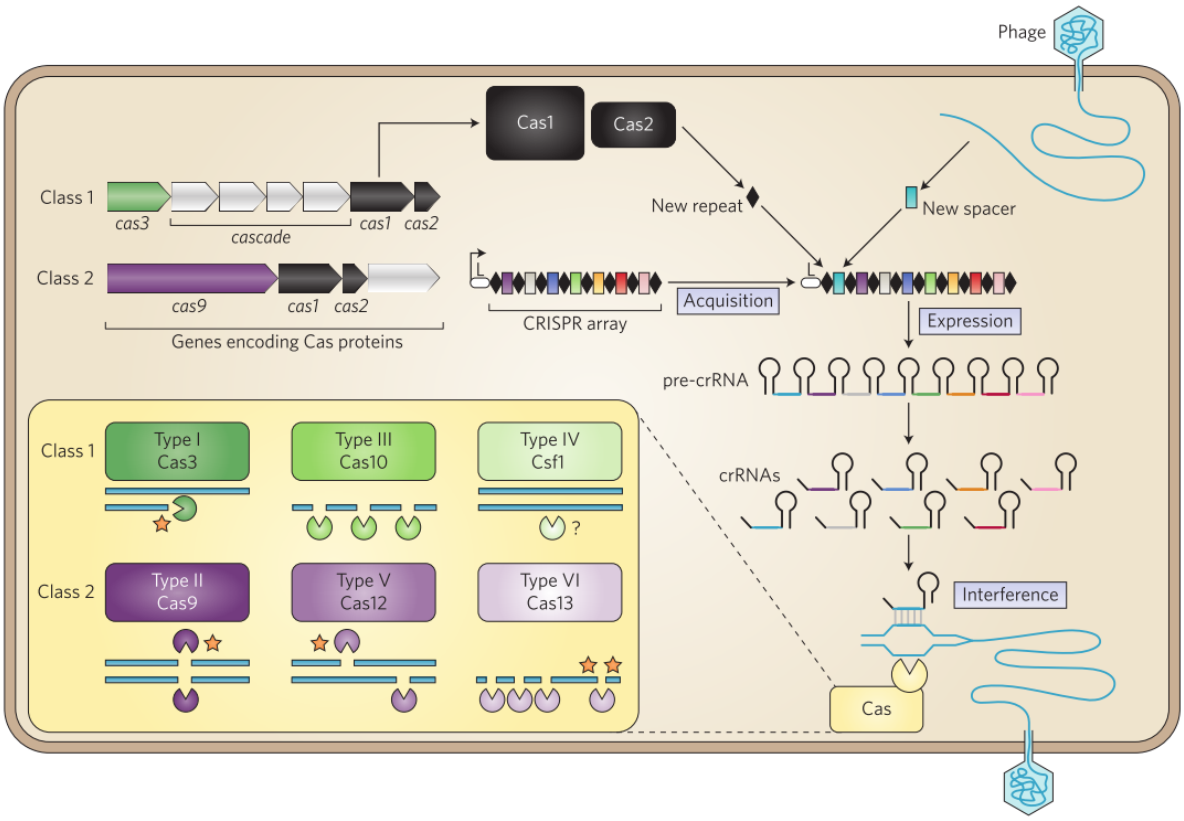


Figure 1-9 Diagrammatic representation of the prokaryotic CRISPR/Cas adaptive immune systems

Acquisition: *cas* genes are expressed from the *cas* operon adjacent to the CRISPR array, which is transcribed from the leader (L) promoter sequence. Foreign, viral DNA is injected into the cell via a bacteriophage (top) from which a novel short spacer sequence (blue rectangle) is incorporated into the CRISPR array, at the leader-end, with a new CRISPR repeat (black diamond), in a Cas1-Cas2-dependent process. Expression: The CRISPR array is transcribed into pre CRISPR-RNA (pre-crRNA) before being processed into small, mature crRNAs, each containing a unique targeting sequence (guide; coloured) derived from an individual CRISPR spacer. Interference: The Cas effector nuclease (yellow rectangle) consisting of either a protein complex (class 1) or single protein (class 2) is guided to the foreign DNA (blue line) target by the crRNA via complementary base-pairing. The precise nucleolytic activity depends on the organism-specific Cas protein(s) present. For example, class 2, type II systems use the Cas9 endonuclease to generate a double strand DNA break 5' to the PAM sequence (star symbol). Diagram from (Barrangou & Horvath, 2017).

1.5.2 CRISPR/Cas9 systems for genome editing

The CRISPR/Cas adaptive immune systems described above can be applied to facilitate genome editing in heterologous organisms thanks to recent pioneering studies in this field. The Cas9 protein is a component of type II CRISPR/Cas systems and is guided to target DNA upon forming a complex with a crRNA:tracrRNA (trans-activating CRISPR RNA) hybrid (Jinek *et al*., 2012). These studies demonstrated that Cas9 is an endonuclease which nicks both strands of target dsDNA via RuvC and HNH motifs (Gasiunas *et al*., 2012), and that this can be reprogrammed using a chimeric single guide RNA (sgRNA) which mimics the crRNA:tracrRNA complex (Jinek *et al*., 2012). Consequently, the site of Cas9-mediated double strand breaks was retargeted by supplying a sgRNA with complementary sequence to the target site. The *Streptococcus pyogenes* Cas9 PAM is 5'-NGG-3', therefore this Cas9 can cut any target DNA region with a 5'-N_{20}-NGG-3' sequence, where N is any nucleotide and N_{20} is the protospacer, flanked on one side by the NGG PAM (Karvelis *et al*., 2015). Thus, having generated a dsDNA break at the desired target site, genome editing can be achieved using DNA repair systems, non-homologous end joining or homology directed repair. A flurry of recent publications have exemplified this CRISPR/Cas9 genome editing method in a wide range of organisms. As shown in figure 1.10, the *S. pyogenes* Cas9 (or derivative) is supplied on a plasmid with an editing template and sgRNA sequence retargeted to a specific chromosomal 5'-N_{20}-NGG-3' target sequence. Hence, when generating deletion mutants, the endonuclease activity of Cas9 is used as a highly efficient counter-selection marker for cells which have successfully undergone two-step homologous recombination. Advantages of CRISPR/Cas9 genome editing are; i) precise genome editing can be performed on a similar timescale to ClosTron mutagenesis, ii) this method does not generate insertional mutants, thus there are no risks of polar effects on downstream genes, and iii) this method can be performed in a wild-type background, unlike *pyrE*-mediated allelic exchange.

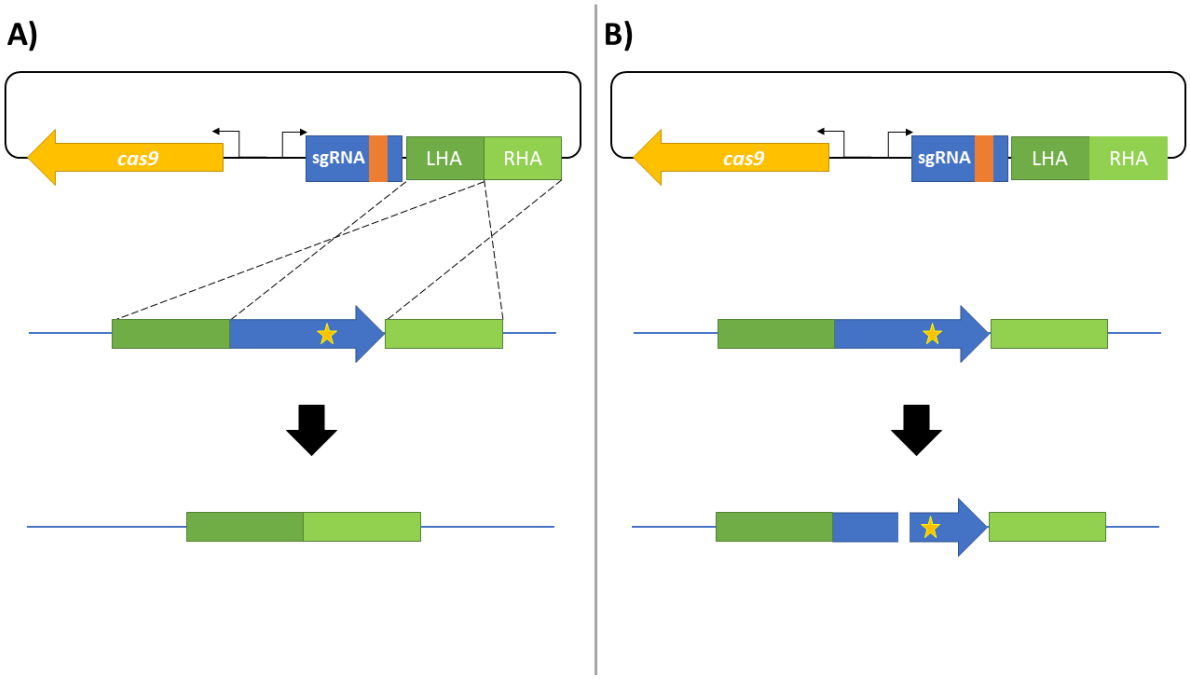


Figure 1-10 Diagrammatic representation of CRISPR/Cas9-mediated genome editing.

The CRISPR/Cas9 vector harbouring the *S. pyogenes cas9* gene, a single guide RNA (sgRNA) component with target protospacer sequence (orange rectangle), and editing template for the generation of an in-frame deletion of the chromosomal gene (blue arrow) is transferred into the target organism. A) Double homologous recombination occurs in this cell, removing the sgRNA guide sequence and PAM (yellow star) resulting in an edited chromosome and no Cas9-mediated nucleolytic activity. All colonies obtained should be mutants lacking the target gene. B) Double homologous recombination has not occurred, allowing the sgRNA to guide Cas9 to the target sequence and generate a lethal, double strand DNA break. No colonies should be obtained in this case.

1.6 Aims of this project

The broad aim of this project is to gain a deeper understanding of the molecular mechanisms involved in sporulation and germination in *C. difficile*, based upon our detailed current knowledge of these processes in the model spore-former *B. subtilis*, and to develop new strains and mutagenesis tools to facilitate this aim. Specifically, this study aims to use the currently available mutagenesis tools to elucidate the role of the SpoVA proteins in *C. difficile*. Furthermore, this study will establish these mutagenesis tools in a novel 'hypervirulent' PCR-ribotype 027 *C. difficile* strain and seek to establish CRISPR/Cas9 genome editing in *C. difficile*.

Chapter Two

Materials and Methods

2 Materials and Methods

2.1 Materials

Unless otherwise stated, all chemicals were supplied by Sigma-Aldrich (US), enzymes for molecular biology and their buffers were provided by New England Biolabs (NEB; US) and components of bacterial growth media were supplied by Oxoid (UK).

2.2 Bacterial Strains

Strain	Relevant Properties	Source/Reference
Escherichia coli TOP10	F– *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
E. coli DH5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NEB
E. coli CA434	Conjugation donor into 630Δ*erm*	Purdy *et al*., (2002)
E. coli NEB sExpress	Conjugation donor into DH1916	Craig Woods (unpublished)
C. difficile 630 (CRG 856)	PCR-ribotype 012 (Zurich, Switzerland)	Peter Mullany, UCL
C. difficile 630Δ*erm*	Erythromycin-sensitive derivative of *C. difficile* 630	Hussain *et al*., (2005)
C. difficile 630 NCTC 13307	PCR-ribotype 012 (Zurich, Switzerland)	NCTC, Public Health England
C. difficile R20291	PCR-ribotype 027 (Stoke Mandeville, UK)	Anaerobe Reference Laboratory, Cardiff

C. difficile DH1916	PCR-ribotype 027 (Torbay, UK)	Anaerobe Reference Laboratory, Cardiff
C. difficile 630Δ*erm*Δ*pyrE*	*pyrE* in-frame deletion mutant created using ACE	Ng *et al*., (2013)
C. difficile 630Δ*erm*Δ*spo0A*	*spo0A* in-frame deletion mutant created using allelic exchange	Ng *et al*., (2013)
C. difficile 630Δ*erm*Δ*cspC*	*cspC* in-frame deletion mutant created using allelic exchange	Ng *et al*., (2013)
630Δ*erm*::*spoVAC*	*C. difficile* 630Δ*erm spoVAC*::intron *ermB*	Lorna Finch (unpublished)
630Δ*erm*::*spoVAD*	*C. difficile* 630Δ*erm spoVAD*::intron *ermB*	This study
630Δ*erm*::*spoVAE*	*C. difficile* 630Δ*erm spoVAE*::intron *ermB*	Lorna Finch (unpublished)
C. difficile 630Δ*erm*Δ*pyrE*Δ*spoVA*	*spoVA* in-frame deletion mutant created using allelic exchange	This study
C. difficile 630Δ*erm*Δ*spoVA* (CRG5234)	*pyrE*-corrected, *spoVA* in-frame deletion mutant created using allelic exchange	This study
C. difficile 630Δ*erm*Δ*spoVA*::*spoVA* (CRG5235)	CRG5234 with *spoVA* operon inserted at corrected *pyrE* locus	This study
PSI_7374	CRG5234 with *spoVA* operon, minus *spoVAE*, inserted at corrected *pyrE* locus	This study
PSI_7375	CRG5234 with *spoVA* operon, minus *spoVAD*, inserted at corrected *pyrE* locus	This study
PSI_7475	CRG5234 with *spoVA* operon, minus *spoVAC*, inserted at corrected *pyrE* locus	This study
C. difficile R20291Δ*pyrE*	*pyrE* in-frame deletion mutant created using ACE	Ng *et al*., (2013)
C. difficile DH1916Δ*pyrE*	*pyrE* in-frame deletion mutant created using ACE	This study
C. difficile DH1916Δ*pyrE*Δ*cspC*	*cspC* in-frame deletion mutant created using allelic exchange	This study

C. difficile DH1916Δ*cspC*	*pyrE*-corrected, *cspC* in-frame deletion mutant created using allelic exchange	This study
C. difficile DH1916Δ*pyrE*Δ*spo0A*	*spo0A* in-frame deletion mutant created using allelic exchange	This study
C. difficile DH1916Δ*spo0A*	*pyrE*-corrected, *spoVA* in-frame deletion mutant created using allelic exchange	This study
C. difficile 630Δ*erm**	CRISPR/Cas9 genome edited *C. difficile* 630, lacking *erm1(B)* and *erm2(B)*; erythromycin-sensitive	This study
C. difficile 630Δ*erm**Δ*pyrE*	CRISPR/Cas9 genome edited *C. difficile* 630* with *pyrE* truncation	This study
C. difficile DH1916Δ*hsdR*	CRISPR/Cas9 genome edited *C. difficile* DH1916 lacking *hsdR*	This study

Table 2-1 List of bacterial strains used in this study. Strain isolation locations are indicated in parentheses.

2.3 Plasmids

Plasmid	Description	Source/Reference
pMTL81151	Modular plasmid containing ColE1+*tra* and *catP*. No Gram-positive replicon	Heap *et al*. (2009)
pMTL82151	pMTL81151 with pBP1 Gram-positive replicon	Heap *et al*. (2009)
pMTL83151	pMTL81151 with pCB102 Gram-positive replicon	Heap *et al*. (2009)
pMTL84151	pMTL81151 with pCD6 Gram-positive replicon	Heap *et al*. (2009)
pMTL85151	pMTL81151 with pIM13 Gram-positive replicon	Heap *et al*. (2009)
pMTL86151	pMTL81151 with pPIP404 Gram-positive replicon	Ed Farries (unpublished)
pMTL007C-E2	ClosTron plasmid containing *catP* and group II intron containing *ermB* RAM	Heap *et al.* (2007)
pMTL007C-E2::Cdi-*spoVAC*-155/156a	pMTL007C-E2 retargeted to CD630_0773	This study

pMTL007C-E2::Cdi-*spoVAD*-585/586s	pMTL007C-E2 retargeted to CD630_0774	This study
pMTL007C-E2::Cdi-*spoVAE*-167/168a	pMTL007C-E2 retargeted to CD630_0775	This study
pMTL-YN1	ACE vector for correction of the *pyrE* truncation in *C. difficile* 630Δ*erm*	Ng *et al*. (2013)
pMTL-YN1C	ACE complementation vector to restore *pyrE* truncation and insert cargo DNA at this locus	Ng *et al*. (2013)
pMTL-YN2	ACE vector for correction of the *pyrE* truncation in *C. difficile* R20291 or DH1916	Ng *et al*. (2013)
pMTL-YN3	Allelic exchange vector containing heterologous *pyrE* from *C. sporogenes* and pCB102 Gram-positive replicon	Ng *et al*. (2013)
pMTL-YN4	pMTL-YN3 with pBP1 Gram-positive replicon	Ng *et al*. (2013)
pMTL-YN3-*spoVAC*	Allelic exchange vector for the creation of an in-frame *spoVAC* deletion	This study
pMTL-YN3-*spoVAD*	Allelic exchange vector for the creation of an in-frame *spoVAD*deletion	This study
pMTL-YN3-*spoVAE*	Allelic exchange vector for the creation of an in-frame *spoVAE* deletion	This study
pMTL-YN3-*spoVA*	Allelic exchange vector for the creation of an in-frame *spoVA* operon deletion	This study
pMTL-YN1C-*spoVA*	ACE complementation vector to insert the 630Δ*erm spoVA* operon at *pyrE* locus	This study
pMTL-YN1C-*spoVACD*	pMTL-YN1C-spoVA lacking the *spoVAE* ORF	This study
pMTL-YN1C-*spoVACE*	pMTL-YN1C-spoVA lacking the *spoVAD* ORF	This study
pMTL-YN1C-*spoVADE*	pMTL-YN1C-spoVA lacking the *spoVAC* ORF	This study
pMTL-JH18::λ6.5	ACE vector used to generate 630Δ*erm*Δ*pyrE*	Heap *et al*. (2012)
pMTL-PSI18	ACE vector for generation of DH1916Δ*pyrE*	This study
pMTL-YN4-DH_*cspC*	Allelic exchange vector for generating in-frame deletion of *cspC* from DH1916	This study

pMTL-YN4-DH_*spo0A*	Allelic exchange vector for generating in-frame deletion of *spo0A* from DH1916	This study
pMTL-Cas9-Caethg-*pyrE*-1	CRISPR/Cas9 genome editing vector with pMTL83151 backbone; *S. pyogenes cas9* under P_{thl} control, sgRNA component under P_{araE} control, and an editing template	Pete Rowe (unpublished)
pPSI_cas_001	pMTL-Cas9-Caethg-*pyrE*-1 except for replacement of the *pyrE* sgRNA component for the spo0A1 sgRNA	This study
pPSI_cas_005	pPSI_cas_001 with spo0A5 sgRNA component	This study
pPSI_cas_006	pPSI_cas_001 with spo0A6 sgRNA component	This study
pPSI_cas_011	pPSI_cas_001 with a *spo0A* editing template replacing the previous *pyrE*-based sequence	This study
pPSI_cas_015	pPSI_cas_011 with spo0A5 sgRNA component	This study
pPSI_cas_016	pPSI_cas_011 with spo0A6 sgRNA component	This study
pPSI_cas_111	pPSI_cas_011 with P_{tcdB} replacing P_{araE}	This study
pPSI_cas_120	pPSI_cas_111 replacing the editing template with one to generate a deletion within Tn *5398*	This study
pPSI_cas_121	pPSI_cas_120 with CD2008A sgRNA	This study
pPSI_cas_122	pPSI_cas_120 with CD2008B sgRNA	This study
pPSI_cas_123	pPSI_cas_120 with CD2008C sgRNA	This study
pPSI_cas_130	pPSI_cas_121 replacing the editing template with one to generate a truncated *pyrE*	This study
pPSI_cas_131	pPSI_cas_130 with *pyrE*1 sgRNA component	This study
pPSI_cas_132	pPSI_cas_130 with *pyrE*2 sgRNA component	This study
pPSI_cas_133	pPSI_cas_130 with *pyrE*3 sgRNA component	This study
pPSI_cas_140	pPSI_cas_111 replacing the editing template with one to generate a DH1916_*hsdR* deletion	This study
pPSI_cas_141	pPSI_cas_140 with *hsdR*1 sgRNA component	This study
pPSI_cas_142	pPSI_cas_140 with *hsdR*2 sgRNA component	This study
pPSI_cas_143	pPSI_cas_140 with *hsdR*3 sgRNA component	This study

pMTL-GL15	Transposon mutagenesis vector harbouring the *Himar1* transposase, no Gram-positive replicon	Cartman & Minton, (2010)

Table 2-2 List of plasmids used in this study.

2.4 Oligonucleotide Primers

Name	sequence (5' – 3')	Function
0773_sF1	ggatagagaagaaagtattgtctaag	Screening and sequencing of putative *spoVA* ClosTron mutants
0773_sR1	tactatagttccagttgatattatggttg	
0774_sF1	caatggcagaaggtctttgtatagg	
0774_sR1	gccatagctggtcccatattattac	
0775_sF1	tcaccgacaagtacactacaaaagc	
0775_sR1	ttgtcaacttcaattatgtaaacgcttc	
EBS Universal	cgaaattagaaacttgcgttcagtaaac	Junction PCR screens
EBS2	tgaacgcaatttctaatttcgatttaacttcgatagaggaaagt	Probe for Southern Blot
Intron-Sal-R1	attactgtgactggtttgcaccaccctcttcg	
spoVA_LF1	ttttttcctgcaggtaagcatgaaaggagaaagtg	SOE PCR primers for the generation of knockout cassettes for deletion of *spoVACDE* via allelic exchange
spoVA_LR1	tttatggtttatccatactaaacacctctctaaag	
spoVA_RF1	tagtatggataaaccataaatctttattaaaagaataataag	
spoVA_RR1	aaaaaaggcgcgccactaatgttattattaggctg	
YN3_sF1	ctccatcaagaagagcgac	Sequencing of knockout cassettes
YN3_sR1	ctttctattcagcactgttatgcc	
CD630_pyrD_sF1	agagaaggaataaaaagtttagacgaaataagagg	PCR primers flanking CD630 *pyrE* gene
CD630_0189_sR3	ccaagctctatgacagacagctcattgtttagaac	
m13f	tgtaaaacgacggccagt	Sequencing primer
VACDE_1C_F3	aaaaaagcggccgcataacttttatgttaagtttttaaattattaaaag	*spoVACDE* complementation cassette
VACDE_1C_R1	ctccaaaagcaaaaccataactcgagaaaaaa	
VACDE_1C_F1	ttttttctcgagttatggttttgcttttggag	PCR and subsequent HiFi assembly of
VACD_1C_R1	agcttgcatgtctgcaggcctcgagttatggttttgcttttggag	

YN1C_pro_F1	gaattagggatgtaataagcggccgctaatacttatgatatgtagaataataaaatataataaatatattac	*spoVA* complementation vectors lacking individual *spoVA*
YN1C_pVAC_F1	gaattagggatgtaataagcggccgctaatacttatgatatgtagaataataaaatataataaatatattac	
VAE_pVAC_R1	atatcacttcattttcacacctttttaattttgc	
VADE_pro_R1	ataaaaatcatactaaacacctctctaaagtattatttg	
pro_VADE_F1	ggtgtttagtatgatttttatttaatttaaatttttattaattatgaagaag	
VADE_YN1C	agcttgcatgtctgcaggcctcgagttatggttttgcttttggag	
DH1916_LF1	cctgcaggggagggacatttttttattatcttcag	Generation of ACE *pyrE* truncation cassette
DH1916_LR1	gcggccgcacaacgtcttcagcaattattatctttg	
DH1916_RF1	gctagcaaacttaattatttatagtgttacttaaaaaatg	
DH1916_RR1	ggcgcgccatagtatataacattaataaaatttaaaatcaataattatc	
DH1916_spo0A_LF1	ttttttcctgcaggtgtattggtcttaggatatgaggaagagg	SOE PCR primers for generation of *spo0A* knockout cassette
DH1916_spo0A_LR1	tttatttaacccccattaaaaaacatcttcttattacag	
DH1916_spo0A_RF1	tttaatgggggttaaataaacaagacataaaaagtaagg	
DH1916_spo0A_RR1	aaaaaaggcgcgcccaatactggctttacctcttctatataag	
DH1916_cspC_LF1	ttttttcctgcaggaatatgtctttaggtaccagtagcttgg	SOE PCR primers for generation of *cspC* knockout cassette
DH1916_cspC_LR1	tctatagagtttccataaatccctcctatcttaaaac	
DH1916_cspC_RF1	atttatggaaactctatagataagaacctatgtaaaattataattg	
DH1916_cspC_RR1	aaaaaaggcgcgccgtactgttctttttctctgtacaattc	
DH1916_cspC_sF1	tctacatccagattttatttatccagatgg	Screening and sequencing of putative *spo0A* or *cspC* deletions
DH1916_cspC_sR1	tgtattgcataggcagctattgatcttactc	
DH1916_spo0A_sF1	atcatatattccaacaggtacccaagc	
DH1916_spo0A_sR1	tcctggtgacaatgatgcaactacc	
RNg630spo_MauBI_F	atatatcgcgcgcgggtaaaataaaaggagattttaatgacagc	Amplification of spo0A knockout cassette with flanking *Mau*BI / *Asc*I sites
RNg630spo_AscI_R	atatatggcgcgccctccaacattatcaattattagtatattatttttcag	
sgRNA_2.6R	atatatggcgcgccatatatgcgatcgcataaaaataagaagcctg	Amplification of sgRNA
sgRNA_sF1	atatattctagatttatatttagtcccttgcc	

RNg630spo_AsiSI_F	atatatgcgatcgcctccaacattatcaattattagtatattatttt cag	Amplification of *spo0A* knockout cassette
PtcdB_*Xba*I	atatattctagattaatgaatttaaagaaatatttacaatag	Amplification of promoter sequences for cloning into CRISPR/Cas9 vectors
PtcdB_*Sal*I	atatatcagctgattttctcctttactataatatttttattg	
Pfdx_*Xba*I	atatattctagagtgtagtagcctgtgaaataagtaagg	
Pfdx_*Sal*I	atatatcagctgcacaactttatacgatttatattttacc	
Tn5398_sF1	atacgtcttatttcccagatatgc	Screening of putative erythromycin resistant mutants
Tn5398_sF2	agttgaatgtaggaaatgaaaagg	
Tn5398_sR1	tctgatggcgtttcataagc	
630erm_asiSI_LF1	atatatgcgatcgcaaccattcttaccgcattgc	SOE PCR primers for generation of Tn5398 knockout cassette
630erm_LR1	tatcatcaacaatcacagtaatcactcctgaagtgattacatc	
630erm_RF1	ttactgtgattgttgatgataaaataagaataagaag	
630erm_ascI_RR1	atatatggcgcgccatcgtccttgtgaagcacaagc	
CD2008A_HiFi	ttatagtaaaggagaaaatgtcgacaacatcaaatcaaacgtgccgttt tagagctagaaatagcaagtt	HiFi assembly of CRISPR/Cas9 vectors targeting CD630_20080
CD2008B_HiFi	ttatagtaaaggagaaaatgtcgacttgcaagaacagtttaacccgtttt agagctagaaatagcaagtt	
CD2008C_HiiFi	ttatagtaaaggagaaaatgtcgacaatggcattacagaacacaagttt tagagctagaaatagcaagtt	
630pyrE_*asi*SI_LF1	atatatgcgatcgcttaagtgtgaaatttggaagtgtag	SOE PCR primers for generation of *pyrE* knockout cassette
630pyrE_LR1	aagtttttattaaacgtcttcagcaattattatctttgc	
630pyrE_RF1	taattgctgaagacgtttaataaaaacttaattatttatagtgttac ttaaaaaatg	
630pyrE_*asc*I_RR1	atatatggcgcgccatcttgaagcattgatgttcttcc	
pyrE1_HiFi	ttatagtaaaggagaaaatgtcgacagagtattagaagccttag ggttttagagctagaaatagcaagtt	HiFi assembly of CRISPR/Cas9 vectors targeting *pyrE*
pyrE2_HiFi	ttatagtaaaggagaaaatgtcgacgagtgtcctttatgtaagga gttttagagctagaaatagcaagtt	
pyrE3_HiFi	ttatagtaaaggagaaaatgtcgacacctacaacttctccaccta gttttagagctagaaatagcaagtt	
DH_hsdR_ *asi*SI_LF1	atatatgcgatcgcaaggagctactgaaggatttacaagg	SOE PCR primers for generation of *hsdR* knockout cassette
DH_hsdR_LR1	tcattagcctatattacccaaggcttttctcc	
DH_hsdR_RF1	agccttgggtaatataggctaatgatgttatttagtgc	

DH_hsdR_ascI_RR1	atatatggcgcgccattaattgagggtcatcttcacc	
hsdR1_HiFi	ttatagtaaaggagaaaatgtcgactattgaaaaagcaatgacaggttttagagctagaaatagcaagtt	HiFi assembly of CRISPR/Cas9 vectors targeting *hsdR*
hsdR2_HiFi	ttatagtaaaggagaaaatgtcgacactagtgcgaacatagagaggttttagagctagaaatagcaagtt	
hsdR3_HiFi	ttatagtaaaggagaaaatgtcgactcactatatgaagttatgaggttttagagctagaaatagcaagtt	
spoIIE_sF1	atggcagttataaaatgaagaatttagtgg	Sequencing of *spoIIE*
spoIIE_sR1	aaatccatcaatactcttattaagg	
hsdR_sF1	aagctgatttacaatgtaactataagc	Screening of putative *hsdR* CRISPR/Cas9 deletion mutants
hsdR_sF2	attttggttcttgcttatatgacc	
hsdR_sR1	taacacaaagtttcttactctatcagc	

Table 2-3 List of oligonucleotides used in this study.

2.5 Bioinformatics Tools

2.5.1 Sequence data analysis

Sequence data were routinely viewed using Artemis and SnapGene Viewer software packages, or using the Benchling online resource accessible at www.benchling.com.

2.5.2 Plasmid maps

Plasmid maps were designed and edited using the SnapGene Viewer software package and the Benchling online resource.

2.5.3 Identification of promoter and terminator sequences

Nucleotide sequences were inputted into the bacterial promoter prediction resource, Softberry, BPROM (Solovyev & Salamov, 2011) accessible at

http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb. For the identification of putative bacterial terminator sequences, nucleotide sequences were inputted into the ARNold online resource (Gautheret & Lambert, 2001; Macke *et al*. 2001), accessible at http://rna.igmors.u-psud.fr/toolbox/arnold/index.php.

2.5.4 BLAST

Searches of protein sequence databases using the Basic Local Alignment Search Tool (BLAST) were performed using the BLASTp algorithm (Altschul *et al.* 1990) accessible at https://blast.ncbi.nlm.nih.gov/Blast.cgi.

2.5.5 Sequence alignments

Alignments of one or more sequences against a reference sequence were performed using the alignment tool available on the Benchling online resource, which utilises version 7 of the MAFFT programme (Katoh & Standley, 2013) with the following settings; max refinement iterations = 0; Tree rebuilding = 2; Gap open penalty = 1.53; Gap extension penalty = 0; Adjust Direction = No.

2.5.6 Oligonucleotide design

Oligonucleotides for PCR primers were manually designed using SnapGene viewer or Benchling. Analysis of the designed primers for secondary structures, self-dimerization and melting temperatures was performed using the OligoAnalyser tool accessible online at https://www.idtdna.com/calc/analyzer.

2.5.7 Whole genome sequencing analysis

Whole genome sequencing was conducted by the DeepSeq facility (University Of Nottingham), using Illumina MiSeq 500bp V2 SBS chemistry. Analysis of reads generated from whole genome sequencing was performed using the CLC Genomics Workbench v6.0 (Qiagen). Raw Illumina paired end reads data were mapped to a reference genome using local alignment with the following settings; mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.5, similarity fraction = 0.8. Reference genomes utilised were *C. difficile* 630 AM180355.1 (Sebaihia *et al*., 2006) or DH1916 (draft genome). Determination of genomic variants was performed using the Quality-based variant detection tool using quality parameters (neighbourhood radius = 5, maximum gap and mismatch count = 2, minimum neighbourhood quality = 15, minimum central quality = 20) and significance parameters (minimum coverage = 10, minimum variant frequency = 70, require presence in both forward and reverse reads = yes).

2.5.8 Data analysis

All data analyses and the graphical presentation of these data were performed using GraphPad Prism. Calculations of statistical significance were performed using Graphpad Prism with Student's t-tests, one-way analysis of variance (ANOVA) with Dunnett's or Tukey's multiple comparisons tests, or two-way ANOVA with Sidak's multiple comparisons tests. Calculated *P*-values <0.05 were deemed statistically significant.

2.6 General Microbiological Techniques

2.6.1 Growth media

Media	Components	Quantity $g.l^{-1}$
Luria-Bertani (LB)	Sodium chloride	10.0
	Tryptone	10.0
	Yeast extract	5.0
Brain Heart Infusion Supplemented (BHIS)	Brain Heart Infusion	37.0
	Yeast extract	5.0
	L-cysteine	1.0
2xYTG	Tryptone	16.0
	Yeast extract	10.0
	Glucose	5.0
	Sodium chloride	4.0
	Gelzan	2.4
CDMM	Cas-amino acids	10.0
	Tryptophan	0.5
	Cysteine	0.5
	Sodium chloride	0.9
	KH_2PO_4	0.9
	Na_2HPO_4	5.0
	$NaHCO_3$	5.0
	Glucose	5.0
	$CaCl_2.2H_2O$	0.026
	$MnCl_2$	0.01
	$CoCl_2$	0.001
	$MgCl_2$	0.02
	$(NH4)_2SO_4$	0.04
	$FeSO_4.7H_2O$	0.0013
	Ca-D-pantothenate	0.001
	Pyridoxine	0.001
	D-biotin	0.001

Table 2-4 Growth media components used in this study.

2.6.2 Growth conditions

2.6.2.1 Aerobic growth conditions

Escherichia coli strains were cultured aerobically in Luria-Bertani (LB) media at 37°C, in liquid broth with horizontal shaking at 200 revolutions per minute, or on solid agar plates.

2.6.2.2 Anaerobic growth conditions

Clostridium difficile strains were cultured anaerobically at 37°C within an MG1000 Mark II anaerobic workstation (Don Whitley Scientific Ltd, UK), with an internal atmosphere of Nitrogen (80%), Carbon dioxide (10%) and Hydrogen (10%). Prior to use, all culture media were pre-reduced within the anaerobic workstation, for a minimum of four hours for agar plates or sixteen hours for liquid media. *C. difficile* strains were routinely cultured on Brain Heart Infusion Supplemented (BHIS) media with appropriate supplements, or *C. difficile* minimal media (CDMM) when required.

2.6.3 Growth supplements

Where necessary, the supplements listed in Table 2.5 were added to growth media in the concentrations indicated.

2.6.4 Strain storage

E. coli strains were grown on appropriate solid media, before the resulting growth was harvested using a 10 µl plastic loop and resuspended in Microbank Long Term Storage tubes (Pro-Lab Diagnostics), for storage at -80°C. *C. difficile* strains were grown on appropriate solid media, before the resulting growth was harvested using a 10 µl plastic loop and

resuspended in screw cap tubes containing BHIS broth with 10 % (v/v) glycerol, for storage at -80°C.

Supplement	Stock concentration $mg.ml^{-1}$	Solvent	Working concentration in *E. coli* $\mu g.ml^{-1}$	Working concentration in clostridia $\mu g.ml^{-1}$
Chloramphenicol	25	EtOH (100%)	Broth: 12.5 Agar plates: 25	-
Thiamphenicol	15	EtOH (50%)	-	15
Erythromycin	50	EtOH (100%)	500	10
D-cycloserine	50	dH_2O	-	250
Cefoxitin	50	dH_2O	-	8
5-Fluoroorotic acid (5-FOA)	100	DMSO	-	4000
Uracil	1	dH_2O	-	20
Sodium taurocholate	100	dH_2O	-	1000

Table 2-5 Growth media supplements used in this study.

2.6.5 Measurements of bacterial growth

Growth of *C. difficile* strains was measured via monitoring the changes in optical density at 600 nm (OD_{600}) over a 24-hour period. The optical density of overnight *C. difficile* cultures was measured and the volume required to produce sub-cultures (50 ml) with OD_{600} = 0.01 calculated. OD_{600} readings of 1 ml aliquots from the resulting subcultures were taken immediately (T=0) and at hourly intervals for 12 hours (T=12), before a final reading 24-hours after inoculation (T=24) was taken.

2.6.6 Preparation of electro-competent *E. coli*

E. coli TOP10, CA434 or NEB sExpress LB broth cultures (5 ml) were inoculated with growth from an LB agar plate and incubated at 37°C with horizontal shaking (200 rpm) for a minimum of 16 hours. 1 ml of subsequent culture was used to inoculate LB broth (200 ml, 37°C) and the resulting sub-culture incubated (37°C, 200 rpm horizontal shaking) until an optical cell density at 600 nm of 0.5-0.7 was obtained. The culture was then divided between four 50 ml falcon tubes and cells harvested via centrifugation (4000 x *g*, 10 minutes, 4°C). Obtained cells were resuspended gently in sterile, ice-cold distilled water (40 ml) and centrifugation repeated as previously. The resulting cell pellets were resuspended in ice-cold distilled water (20 ml), combined, centrifuged as before and resuspended in sterile, ice-cold distilled water (10 ml). A final centrifugation step was performed as previously and the washed pellet was resuspended in sterile, ice-cold MOPS (1 ml, 1mM) with glycerol (10% v/v), before being pooled together and dispensed into 50 µl aliquots and stored at -80°C.

2.6.7 Transformation methods

2.6.7.1 Transformation of E. coli by electroporation

An aliquot of electro-competent *E. coli* cells (50 µl) was thawed gently on ice prior to the addition of plasmid DNA (2-4 µl) or dialysed ligation reaction products (4 µl) and mixing by aspiration. This DNA-cell mixture was transferred into a chilled electroporation cuvette (0.2 cm gap, Bio-Rad) and pulsed with an electroporator (Bio-Rad MicroPulser) using pre-set conditions (2.5 kV, 200 Ω, 25 µF capacitance). Immediately, cells were resuspended in SOC broth (400 µl, Invitrogen), transferred to a 1.5 ml microcentrifuge and incubated (37°C, 200 rpm horizontal shaking) for one hour. Following incubation, 100 µl of the cell mixture was

spread plated onto LB agar medium containing the appropriate antibiotic and incubated at 37°C overnight.

2.6.7.2 Conjugative transfer of plasmid DNA into Clostridium difficile

LB broth (5 ml) supplemented with appropriate antibiotics was inoculated with *E. coli* CA434 or NEB sExpress conjugal donor cells harbouring the plasmid to be transferred and incubated (37°C, 200 rpm horizontal shaking) overnight. BHIS broth (1 ml) was inoculated with the recipient *C. difficile* strain and incubated anaerobically at 37°C overnight. The next day, stationary phase *E. coli* donor cells (1 ml) were harvested via centrifugation (1800 x *g*, 1 minute, 25°C), the supernatant discarded, and the pellet washed twice in sterile PBS (500 µl). This pellet was transferred into the anaerobic workstation and resuspended in recipient overnight culture (200 µl). The resulting cell suspension was spotted onto a BHIS agar plate in ten spots, each approximately 20 µl in volume, and incubated anaerobically for a minimum of 8 hours. Subsequent growth was harvested via scraping using a 10 µl plastic loop and resuspended in sterile, pre-reduced PBS (500 µl). This suspension was spread plated onto BHIS agar supplemented with D-cycloserine, cefoxitin and the plasmid-appropriate antibiotic in 100 µl aliquots of 10^0 and 10^{-1} dilutions in PBS. Plates were incubated anaerobically at 37°C for 24-72 hours, until distinct transconjugant colonies had appeared.

2.6.8 Preparation of *C. difficile* spores

C. difficile was grown overnight in BHIS broth, spread in 100 µl aliquots onto BHIS agar plates and incubated anaerobically at 37°C for at least five days. After incubation, plates were removed from the anaerobic workstation and incubated overnight at 4°C. Following this, the

spore/vegetative cell mixture was harvested from three agar plates via cell scraping using 10 µl plastic loops and resuspended in ice-cold, sterile distilled water (1 ml) and incubated at 4°C overnight. The resulting mixture was gently resuspended via pipette aspiration and centrifuged (16000 x *g*, 4°C, 4 minutes). Samples were placed on ice whilst the supernatant and as much of the cell debris suspended within it as possible was then gently removed via pipette without disturbing the spore pellet and discarded. The spore pellet was then washed in 1 ml of sterile, ice-cold distilled water, PBS or PBS with Tween80 (0.1%, 0.05% or 0.01% v/v) and centrifugation repeated as previously. A minimum of 10 such wash steps were performed until spore preparations >99% free-from vegetative cells and cell debris were obtained and stored at 4°C.

2.6.9 Measuring *C. difficile* colony formation after heat treatment

C. difficile strains to be assayed were grown in BHIS broth (1 ml) overnight. The following day, BHIS broth (990 µl) was inoculated with the resulting growth (10 µl) and incubated (5 hours). Following incubation, sporulation cultures were set up with the resulting growth (150 µl) added to BHIS broth (14.85 ml). At various times, two 500 µl aliquots were transferred to microcentrifuge tubes from these sporulation cultures; one was then heat-treated (65°C, 30 minutes) whilst the other incubated on the bench (30 minutes). Both samples were then serially diluted in sterile PBS down to 10^{-6} and spotted (20 µl) in triplicate onto BHIS agar plates supplemented with taurocholate (0.1% w/v, Sigma). After 24 hours incubation, these plates were removed from the anaerobic workstation and CFU counted on dilutions which produced between 5-50 CFU per 20 µl spot, from which $CFU.ml^{-1}$ values were calculated.

2.6.10 Measuring spore titres using phase-contrast microscopy

Sporulation cultures were set up as described in 2.6.9 and after five days incubation, aliquots were removed for enumeration of spore titres using an improved Neubauer counting chamber (Sigma). A coverslip was positioned on top of the counting chamber using water droplets and sporulation culture (10 µl) loaded onto the counting chamber. *C. difficile* spores were observed under phase-contrast microscopy and counted in ten, 200 µm squares, averaged and the spores.ml^{-1} value calculated using the appropriate dilution factors.

2.6.11 De-coating *C. difficile* spores

Sporulation cultures (50 ml) were set up using the protocol outlined in 2.6.9 and incubated in the anaerobic workstation for a minimum of five days and spore pellets obtained by centrifugation (16000 x *g*, 4°C, 15 minutes). As described by Popham *et al.* (1995), spore pellets were resuspended in 1 ml of de-coating solution containing Tris-HCl (50 mM, pH 8.0), urea (8 M), SDS (1% w/v) and dithiothreitol (50 mM), and incubated (37°C, 90 minutes). De-coated spores were washed twice in sterile PBS (400 µl), heat-treated (65°C, 30 minutes), serially diluted and spotted in triplicate onto BHIS agar with and without supplemented lysozyme (1 µg.ml^{-1}). These plates were incubated in the anaerobic workstation for 24 hours and resulting CFU.ml^{-1} values calculated as described in 2.6.9.

2.6.12 Heat-treatment of *C. difficile* spore preparations

Purified and counted *C. difficile* spores (1x10^{7}) were aliquoted into microcentrifuge tubes and incubated at various temperatures or left on the bench (not heat-treated; NHT) for 30

minutes. After this incubation, spores were serially diluted in PBS, plated on BHIS agar supplemented with taurocholate (0.1%, Sigma) and CFU.ml^{-1} values calculated as previously.

2.6.13 UV-treatment of *C. difficile* spore preparations

Purified and counted *C. difficile* spores (1x10^7) were aliquoted into microcentrifuge tubes and exposed to varying amounts of UV energy at 254 nm in a CX-2000 Ultraviolet Crosslinker (UVP, Germany), or left on the bench (not treated; NT) After this incubation, spores were serially diluted in PBS, plated on BHIS agar supplemented with taurocholate (0.1%, Sigma) and CFU.ml^{-1} values calculated as previously.

2.6.14 Ethanol treatment of *C. difficile* spore preparations

Purified and counted *C. difficile* spores (1x10^7) were aliquoted into microcentrifuge tubes, centrifuged (16000 x *g*, 2 minutes) and spore pellets resuspended in various ethanol concentrations or sterile, distilled water and incubated at 37^{o}C for 30 minutes. After this incubation, spores were washed twice in distilled water, serially diluted in PBS, plated on BHIS agar supplemented with taurocholate (0.1%, Sigma) and CFU.ml^{-1} values calculated as previously.

2.6.15 Measurements of the Dipicolinic acid content of *C. difficile* spores

Purified and counted *C. difficile* spores (1x10^8) were centrifuged (16000 x g, 2 minutes) and washed twice in sterile distilled water (500 µl) and boiled (100^{o}C) or incubated on the bench for one hour. After incubation, spores were centrifuged (16000 x *g*, 2 minutes) and supernatant (190 µl) was added to wells of an opaque-walled 96-well plate containing either

sterile distilled water (10 µl) or terbium chloride (10 µl, 10 mM, Acros Organics, Belgium). This mixture was gently mixed by pipette aspiration and incubated on the bench for 15 minutes. The plate was then transferred to an Infinite M1000 Pro plate reader (Tecan, Switzerland) and the fluorescence intensity measured (excitation 270 nm, emission 545 nm, gain 100).

2.6.16 Assessment of *C. difficile* cytotoxicity against Vero cells

African green monkey kidney (Vero) cells were grown to a confluent monolayer in Dulbecco's modification of Eagle medium (DMEM) supplemented with heat-inactivated foetal bovine serum (FBS; 10% v/v) and penicillin-streptomycin (1% v/v) at 37°C with 5% carbon dioxide. Confluent growth was washed twice in sterile PBS (10 ml) and incubated (37°C, 5% CO_2, 10 minutes) with trypsin-EDTA (5 ml). The resulting Vero cell suspension was centrifuged (500 x *g*, 25°C, 10 minutes) and the pellet resuspended gently in DMEM (30 ml) before being aliquoted into flat-bottomed 96-well plates (100 µl per well) and incubated overnight (37°C, 5% CO_2). Overnight cultures of the *C. difficile* strains to be assayed were grown in BHIS (1ml), after which the optical density of the growth of all strains was determined and adjusted in fresh BHIS to the lowest recorded value. The subsequent *C. difficile* growth dilutions were centrifuged (16000 x *g*, 5 minutes) and the resulting supernatants filter sterilised (0.2 µm membrane, Minisart), serially diluted (1:4) in sterile PBS and gently aliquoted (20 µl) onto the overnight growth of Vero cells. Supernatant-Vero cell mixtures were incubated (37°C, 5% CO_2, 24 hours) after which cell rounding was observed under light microscopy. Endpoint toxin titres were determined semi-quantitatively by observing under light microscopy the dilution at which Vero cell rounding was indistinguishable from the control wells, in which Vero cells were supplemented with PBS alone.

2.6.17 Assessment of bacterial motility

Overnight cultures of *C. difficile* strains grown in BHIS broth were spotted in triplicate, in 2 µl volumes, onto pre-reduced 2xYTG agar plates (25 ml) with 0.24% Gelzan. Plates were incubated anaerobically (37°C, 48 hours) after which they were removed from the anaerobic cabinet and photographed, and the motility of each strain was assessed by eye.

2.7 General Molecular Techniques

2.7.1 Genomic DNA extraction from *Clostridium difficile*

2.7.1.1 Purification of chromosomal DNA using GenElute kits

C. difficile cultures grown overnight in BHIS (1 ml) were removed from the anaerobic workstation and cells harvested via centrifugation (16000 x *g*, 3 minutes, 25°C). The protocol within the GenElute Bacterial Genomic DNA Kit (Sigma) was then followed. Briefly, supernatants were discarded and cell pellets resuspended in Lysis buffer (200 µl, PBS containing 10mg.ml^{-1} lysozyme) and incubated at 37°C for 30 minutes, with gentle agitation after each 10-minute period. RNase A Solution (20 µl, Sigma) was added and the mixture incubated (on bench, 2 minutes), before the addition of Proteinase K (20 µl, Sigma) and Lysis Solution (200 µl, Sigma) followed by thorough mixing of the suspension via vortex and a further incubation step (55°C, 10 minutes). Ethanol precipitation of the chromosomal DNA was performed via the addition of ethanol (200 µl, 100%) prior to transfer of the mixture to a silica membrane binding column and centrifugation (6500 x *g*, 1 minute). The membrane was then washed twice using the provided buffers via centrifugation (16000 x *g*, 3 minutes), before chromosomal DNA was eluted from the column in a Tris-EDTA elution buffer (100 µl, Sigma) and stored at -20°C.

2.7.1.2 Purification of genomic DNA via Phenol:Chloroform extraction

C. difficile cultures grown overnight in BHIS (5 ml) were removed from the anaerobic workstation and cells harvested via centrifugation (16000 x *g*, 3 minutes, 25°C). Resulting supernatants were discarded and cell pellets resuspended in Lysis buffer (180 µl, PBS containing 10 $mg.ml^{-1}$ lysozyme) and incubated at 37°C for 30 minutes, with gentle agitation after each 10-minute period. RNase A Solution (20 µl, Sigma) was added and the mixture incubated (on bench, 2 minutes), before the addition of Proteinase K (25 µl, Sigma), sterile, distilled water (85 µl) and SDS solution (110 µl, 10% v/v). Samples were mixed thoroughly by inversion and incubated at 65°C for 30 minutes, again with periodic agitation. Phenol:Chloroforom:Isoamyl alcohol (25:24:1, 400 µl, Sigma) saturated with Tris (10 mM, pH 8.0) EDTA (1mM) was added to samples, mixed thoroughly by inversion and transferred to a Phase lock gel tube (Eppendorf). Phase separation occurred following centrifugation of samples (18000 x *g*, 3 minutes, 25°C) after which the upper supernatant layer was transferred via pipette to a fresh Phase lock tube and the extraction step repeated twice further. The subsequent upper supernatant layer was transferred into a microcentrifuge tube (1.5 ml, Eppendorf), to which ice-cold sodium acetate buffer (40 µl, 3M) and ice-cold ethanol (800 µl, 100%) were added followed by gentle yet thorough mixing via inversion and incubation of the resulting homogenous mixture (-80°C, 30 minutes). DNA was then pelleted via centrifugation (18000 x *g*, 15 minutes, 4°C), the supernatant was immediately removed and the pellet resuspended gently in ethanol (1 ml, 70%) and centrifuged (18000 x *g*, 3 minutes, 4°C). Again, the supernatant was quickly removed and the pellet air-dried for 45 minutes, before being resuspended in sterile, distilled water (50 µl) and stored at -20°C.

2.7.2 Extraction of plasmid DNA from *E. coli*

Cells harbouring the desired plasmid were harvested from overnight *E. coli* cultures via centrifugation (16000 x *g*, 2 minutes, 25°C). Plasmid DNA was extracted from cell pellets using either the GenElute HP Plasmid Miniprep Kit (Sigma) or the Monarch Plasmid DNA Miniprep Kit (NEB), in accordance with the manufacturer's instructions in each case. Plasmid DNA was eluted in sterile, distilled water (50 µl) and stored at -20°C.

2.7.3 Quantification of DNA preparations

The concentration of extracted plasmid DNA was determined by measuring the OD_{260} using a nanodrop ND-1000 spectrophotometer (Thermo Scientific), which then calculated the DNA concentration given that 50 $\mu g.ml^{-1}$ double-stranded DNA has an OD_{260} value of 1 with a 1 cm path length. The concentration of purified genomic DNA was measured using a Qubit 3.0 Fluorometer (Thermo Scientific) according to the manufacturer's instructions.

2.7.4 Polymerase Chain Reaction amplification of DNA

Oligonucleotide primers were designed as described in section 2.5.6 and synthesised by Sigma. Primers were diluted in dH_2O to stock concentrations of 100 µM and subsequently to working concentrations of 10 µM, both of which were stored at -20°C. The volumes of PCR amplification components in a typical 25 µl reaction are outlined in Table 2.6. Reactions were conducted with typical thermocycling conditions outlined in Table 2.7. The annealing temperature for each individual reaction was calculated using the online tool provided by NEB, accessible at https://tmcalculator.neb.com, and the elongation duration determined according to the expected product size, with 30 secs allocated per 1 kb.

PCR Component	Volume (µl)	Manufacturer/Source
Q5 High Fidelity DNA polymerase	0.25	NEB
Q5 HF Reaction buffer (5X)	5	NEB
Oligonucleotide forward primer (10 µM)	1.25	Sigma
Oligonucleotide reverse primer (10 µM)	1.25	Sigma
Deoxynucleotide (dNTP) mix (10 mM)	0.5	NEB
Sterile, nuclease-free water	15.75	Sigma
DNA template (10-20 ng)	1	This study

Table 2-6 Typical components of a PCR amplification mixture used in this study.

Reaction Step	Temperature (°C)	Duration	Programme
1 – Initial denaturation	98	30 secs	
2 – Denaturation	98	10 secs	
3 – Annealing	X	30 secs	
4 – Elongation	72	Y	Return to step 2 (29x)
5 – Final elongation	72	5 minutes	
6 – Hold	15	-	

Table 2-7 Typical thermocycling conditions for DNA amplification via PCR.

Annealing temperature (X) ranged from 50-72°C depending upon the specific oligonucleotide primers used. Elongation duration (Y) was dependent upon the expected product length, with 30 secs per 1 kb allocated.

2.7.5 SOEing PCR

Knockout cassettes were generated by the splicing together of two PCR products via splicing by overlap extension (SOEing) PCR, using appropriately designed primers (Horton *et al*., 1990). These cassettes consist of two distinct regions of homology with the target genome, spliced together and flanked by appropriate restriction sites for downstream cloning.

2.7.6 Colony PCR

Where many strains were to be screened by PCR, a more time-efficient colony PCR protocol was followed which replaced the need for extraction of genomic DNA from these strains. Single colonies were scraped with a sterile toothpick and resuspended in sterile, distilled water (10 µl). This resuspension was gently mixed using pipette aspiration and 1 µl of the resulting homogenous suspension used as template for subsequent PCR amplifications with appropriate primers. The typical PCR protocol was followed, as listed in Table 2.7, except that the initial denaturation step lasted for five minutes

2.7.7 Restriction endonuclease digestion of DNA

DNA was cleaved at specific sites using appropriate restriction endonucleases (NEB) in accordance with the manufacturer's instructions. Typically, reactions were performed in 20 µl volumes with incubation at 37^{o}C for at least one hour.

2.7.8 De-phosphorylation of DNA

5'-phosphate groups were removed from linearized DNA fragments using Antarctic Phosphatase (NEB) in accordance with the manufacturer's instructions. Briefly, following the completion of a restriction endonuclease reaction outlined in section 2.7.7, Antarctic Phosphatase buffer (10X; 2.3 µl) and Antarctic Phosphatase (1 µl) was added to the products of a restriction endonuclease reaction (20 µl) and incubated at 37^{o}C for a minimum of 30 minutes.

2.7.9 Agarose gel electrophoresis

Separation of PCR products, restriction fragments and plasmid DNA was performed via electrophoresis at 100 V for 60-120 minutes through 0.8-1.0% agarose (Sigma) gels in TAE buffer containing 0.01% (v/v) SYBR Safe DNA Gel Stain (Thermo Scientific). Separated DNA was subsequently viewed under blue light.

2.7.10 DNA extraction from agarose gels and reaction mixtures

Agarose gel pieces containing the desired DNA fragments were excised under blue light using a scalpel, from which DNA was extracted using the Zymoclean Gel DNA Recovery Kit (Zymo Research) according to the manufacturer's instructions. DNA extraction from reaction mixtures was performed using the GenElute PCR Clean-Up Kit (Sigma) according to the manufacturer's instructions. All extracted DNA was eluted in sterile, distilled water and stored at -20^{o}C.

2.7.11 DNA ligation

Purified, restriction digested, DNA fragments with compatible ends were ligated together using T4 DNA ligase (NEB) in accordance with the manufacturer's instructions. All reactions were performed in 20 µl volumes and left to incubate in ice water for a minimum of 16 hours.

2.7.12 Dialysis of DNA ligation products

Products of ligation reactions were dialysed through a nitrocellulose membrane (0.025 µm, Millipore) over sterile, distilled water for a minimum of 30 minutes. The dialysed products were carefully removed from the nitrocellulose membrane via pipette and stored on ice prior to being transformed into the desired *E. coli* strain.

2.7.13 Nucleotide sequencing

Sanger sequencing of PCR products and plasmids was performed by Source Bioscience (UK), using provided oligonucleotide primers.

2.8 ClosTron mutagenesis

ClosTron mutagenesis was performed to insertionally inactivate target genes using the modular ClosTron vector pMTL007C-E2 (Heap *et al*., 2010), retargeted to genes of interest to this study.

2.8.1 Intron retargeting for pMTL007C-E2

Target sites for intron insertion within genes to be inactivated were searched for using the Perutka algorithm (Perutka, *et al*., 2004), accessible at http://www.clostron.com. For each target gene, the most favourable intron insertion site was chosen, regardless of the sense/anti-sense orientation.

2.8.2 Synthesis of retargeted pMTL007C-E2

Synthesis of the 353 bp retargeting region for each target gene and incorporation of these regions into separate pMTL007C-E2 vectors was performed by DNA 2.0 (USA). Retargeted plasmids arrived on filter paper and were named according to the established nomenclature (Heap *et al*., 2007), such that pMTL007C-E2::Cdi-*spoVAD*-585/586s denotes a ClosTron vector retargeted to insert after nucleotide number 585, in the sense direction of the *C. difficile* 630Δerm *spoVAD* open reading frame (ORF).

2.8.3 Generation of insertional mutants in *C. difficile*

Retargeted pMTL007C-E2 plasmids were eluted in sterile, distilled water from the supplied filter paper and transformed into *E. coli* TOP10 (for storage) and CA434, using the method outlined in 2.6.7.1. Transformed *E. coli* colonies harbouring pMTL007C-E2 plasmids were selected for based upon acquisition of chloramphenicol resistance. These plasmids were transferred into *C. difficile* 630Δ*erm* from chloramphenicol-resistant *E. coli* CA434 cells, via conjugative transfer outlined in 2.6.7.2, and transconjugants selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. Primary transconjugant colonies were streaked to purity on the same media, before being plated onto BHIS agar supplemented with D-cycloserine, cefoxitin and erythromycin, to select for strains in which the group II intron had inserted into the chromosome. Genomic DNA was extracted from resulting erythromycin resistant colonies and insertion of the Group II intron at the intended chromosomal locus confirmed via a flanking PCR, using primers which anneal to chromosomal regions flanking the insertion site, and a junction PCR, using the EBS Universal primer and the appropriate flanking primer, depending upon the directionality of insertion. Where possible, three independently generated, PCR confirmed mutant strains were stored for each target.

2.8.4 Confirmation of single Group II intron insertion

Southern Blot hybridisation was performed to confirm single insertion of the Group II intron within the 630Δ*erm* chromosome, for each mutant. An intron-specific probe was PCR generated from pMTL007C-E2 plasmid DNA template and EBS2 and Intron-SalR1 primers and labelled with DIG-High Prime (Roche, 4 µl) via incubation (37°C, 16 hours). Chromosomal DNA extracted from PCR-confirmed mutants and pMTL007C-E2 plasmid DNA were individually digested with *Eco*RI and *Hin*dIII restriction endonucleases (37°C, 16 hours). Digest products were separated by agarose gel electrophoresis (0.8%, 120 V, 90 minutes), transferred to a nitrocellulose membrane (on bench, 4 hours), UV-fixed and incubated with the prepared, labelled probes (42°C, 16 hours). The membrane was then washed several times in SSC with 0.1% SDS (100 ml, 68°C, 15 minutes), equilibrated in maleic acid buffer with 0.3% Tween20 (100 ml, 1 minute) and blocked using 1X EasyHyb blocking buffer (Roche, 25 ml, 30 minutes, 25°C). The membrane was then incubated with anti-DIG antibody probe (Roche, 30 minutes, 25°C), washed twice in maleic acid buffer with 0.3% Tween20 (100 ml, 15 minutes, 25°C) and equilibrated in Tris-HCl (100 mM, pH 9.5) NaCl (100 mM) detection buffer (25 ml, 2 minutes, 25°C). CSPD (Roche, 1.5 ml) was then applied to the membrane and incubated at 37°C for 10 minutes. The membrane was exposed to photographic film in a dark room (15 minutes), which was then developed and fixed in the appropriate buffers under red light.

2.9 Allele Coupled Exchange

To facilitate allelic exchange mutagenesis in *C. difficile* DH1916, it was first necessary to generate an in-frame deletion in the *pyrE* gene, encoding orotate phosphoribosyltransferase, via the ACE method developed previously (Heap, *et al*. 2012).

2.9.1 Generation of DH1916Δ*pyrE*

Plasmid pMTL-PSI18 was transferred into DH1916 via conjugative transfer from *E. coli* NEB sExpress and transconjugants selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. Larger thiamphenicol-resistant transconjugant colonies, indicating a single crossover plasmid integration event had occurred, were streaked to purity on the same media. These strains were then passaged twice on CDMM agar supplemented with 5-FOA and uracil. Serial passage in non-selective medium was then performed. Three 5-FOA-resistant colonies, generated from independent conjugations were then PCR screened for the desired 234 bp truncation at the 3' end of *pyrE*.

2.9.2 Correction of *pyrE* negative strains

Reversion of the truncated, *pyrE*-negative strains to wildtype *pyrE* was performed via conjugative transfer of correction vectors pMTL-YN1 (in the case of 630Δ*erm*) or pMTL-YN2 (in the case of R20291 and DH1916). This allows mutants generated via allelic exchange in a *pyrE*-negative background, to be analysed in a wildtype background. Subsequent transconjugant colonies were selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. Larger, faster growing, colonies were streaked to purity on the same medium to obtain single crossover integrants at the larger homology arm. Single thiamphenicol-resistant colonies were then passaged twice on CDMM agar to screen for the

ability of these strains to grow without uracil supplementation. Successful restoration of the *pyrE* locus to wildtype was then confirmed via PCR screens with appropriate primers.

2.10 Allelic Exchange Mutagenesis

All allelic exchange mutagenesis in this study was performed in a *pyrE* negative *C. difficile* strain, following the recently established protocol in *C. difficile* (Ng *et al*., 2013). Following the creation of clean, in-frame deletion mutants using this method, the *pyrE* negative strain was then restored to prototrophy (*pyrE*+) using the appropriate correction vectors.

2.10.1 Construction of knockout cassettes

Sequences of each gene targeted for deletion, including 1 kb up- and down-stream, were identified using Artemis software and *in silico* deletions of each gene, leaving only the first two and final three codons of the ORF, were performed using SnapGene software. Using the resulting recombinant sequence, oligonucleotide primers were designed to amplify left and right homology arm fragments, with 20 bp overlapping homology between the primers 2 and 3 depicted in Figure 2.1.

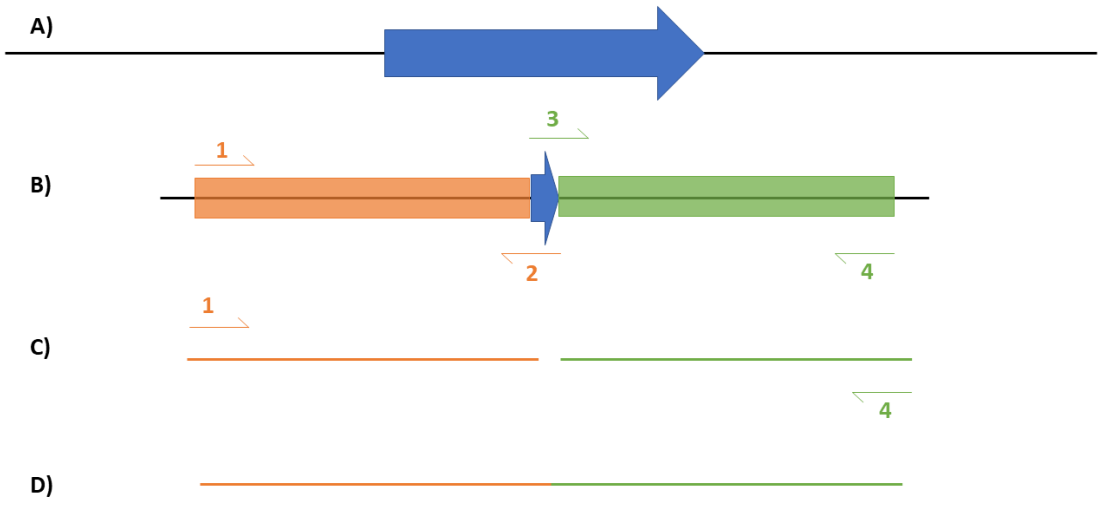


Figure 2-1 Generation of allelic exchange knockout cassettes via SOE PCR.

Target genes (blue arrow) were identified along with 1kb regions up- and down-stream *in silico* (A), an internal gene fragment deleted and oligonucleotides designed (B) to amplify left (orange box; primers 1 and 2) and right homology arms (green box; primers 3 and 4). Purified homology arm fragments (orange and green lines) were then used as template for a SOEing PCR with flanking primers 1 and 4 (C) to generate the recombinant knockout cassette (D).

2.10.2 Construction of allelic exchange vectors

Knockout cassettes were cloned into allelic exchange vectors pMTL-YN3 (630Δ*erm*) or pMTL-YN4 (DH1916) via restriction digestion and subsequent ligation using the restriction sites *Asc*I and *Sbf*I flanking the knockout cassette, introduced during SOEing PCR via their incorporation at the 5' of primers 1 and 4. Ligation products were dialysed and transformed into *E. coli* TOP10 cells and transformants selected on LB agar supplemented with chloramphenicol. Plasmids were extracted from at least six independent chloramphenicol-resistant colonies and Sanger sequenced using the appropriate oligonucleotides to determine successful allelic exchange vector generation.

2.10.3 Generation of in-frame deletion mutants via allelic exchange

Sequence confirmed allelic exchange vectors were transformed into *E. coli* conjugative donors, CA434 for 630Δ*erm* or NEB sExpress for DH1916, and then transferred into the relevant *pyrE* negative *C. difficile* strain via conjugative transfer. Transconjugant colonies harbouring the allelic exchange vectors were selected for on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. Single crossover integrant strains were identified after 48 hours as larger, faster growing, colonies and streaked to purity on the same media. Genomic DNA was extracted from these larger colonies and PCR screening performed using a flanking primer (annealing to a chromosomal region approximately 150 bp upstream and downstream of each homology arm) and a primer annealing within the allelic exchange vector. Two PCRs were performed to check for integration at either homology arm. Confirmed single crossover integrants were then streaked onto BHIS agar and incubated anaerobically for 48 hours to allow a second recombination event to occur. A 10 µl loop of resulting growth was then resuspended in sterile, pre-reduced PBS (200 µl), serially diluted down to 10^{-6} and 100 µl aliquots of each dilution spread plated onto CDMM supplemented with 5-FOA and uracil. At least 50 resulting 5-FOA-resistant colonies were patch plated onto BHIS agar supplemented with D-cycloserine and cefoxitin, with and without thiamphenicol, to screen for colonies which had lost the allelic exchange vector. Genomic DNA was extracted from resultant 5-FOA-resistant, thiamphenicol-sensitive strains and PCR screened for the expected deletion using chromosomal-annealing primers which bind outside of the two homology arms. Confirmed in-frame deletion strains were restored to wildtype *pyrE* alleles using the protocol outlined in 2.9.2.

2.10.4 Complementation of *C. difficile* 630*ΔermΔpyrEΔspoVA*

A *spoVA* complementation cassette was generated by PCR amplification of all three *spoVA* genes including the 279 bp 5'UTR upstream of *spoVAC* using primers *spoVAC*_1C_F3 and *spoVAE*_1C_R1. The 2.276 kb DNA fragment generated was cloned via restriction digestion with *Not*I and *Xho*I enzymes into the 630Δ*erm* complementation vector pMTL-YN1C. The generated vector, pMTL-YN1C_*spoVA* was Sanger sequence confirmed and transformed into *E. coli* CA434 prior to conjugative transfer into 630*ΔermΔpyrEΔspoVA*. Transconjugant colonies were selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol and complemented, *pyrE*-repaired strains were obtained by following the *pyrE*-repair protocol outlined in 2.9.2.

2.11 CRISPR/Cas9 mutagenesis

The CRISPR/Cas9 genome editing method utilised in this study is based upon the work of Huang *et al*. (2016) and Pete Rowe and Chris Humphreys (unpublished). The CRISPR/Cas9 vectors generated in this study all derive from pMTL-Cas9-Caethg-*pyrE*-1, created by Pete Rowe (unpublished).

2.11.1 Identification of sgRNA seed sequences within target sequences

Sequences targeted for deletion were searched for 20 bp seed sequences directly upstream of protospacer associated motifs, 5'-NGG-3' in the case of the *S. pyogenes cas9* gene used in this study, using the Benchling CRISPR guide design tool, accessible at https://benchling.com.

2.11.2 Selection of sgRNA seed sequences for CRISPR/Cas9 genome editing

All identified seed sequences were analysed using the Benchling CRISPR guide design tool to generate on- and off-target scores for each sequence, based upon the Doench *et al*. (2016) algorithm, having provided chromosomal coordinates of the target sequence region. This generates a score out of 100 for the on-target and off-target activities of each sgRNA, with on-target scores representing the probability of each sgRNA being within the top 20% for cleavage activity and the off-target score representing the inverse probability of the sgRNA binding to other genomic sequences. Identified sequences were then ranked by highest on-target score. Three sgRNA guide sequences were then chosen for each target satisfying the following criteria; i) high on-target scores, ii) off-target scores above 99.0 where possible, iii) the three chosen guides are spaced throughout the target sequence, iv) all three chosen guides are not present on the same DNA strand.

2.11.3 Construction of CRISPR/Cas9 genome editing vectors

2.11.3.1 Insertion of sgRNA seed sequences

Seed sequences were incorporated into CRISPR/Cas9 vectors either by PCR amplification using overlapping primers containing the 20 nt seed sequence within the region of overlapping homology between SOEing PCR primers, followed by restriction digests to insert the generated cassette, or via HiFi assembly using a 70 nt oligonucleotide harbouring the seed sequence, flanked by 25 nt sequences with homology to the CRISPR/Cas9 vector sgRNA seed region. Prior to HiFi assembly, a linearized CRISPR/Cas9 fragment was generated via *Sal*I restriction digest, treated with Antarctic phosphatase (2 hours, 37^{o}C), separated by agarose gel electrophoresis and purified. This purified, linearised vector fragment (2 µl) was added to HiFi DNA assembly master mix (10 µl, NEB), seed sequence containing

oligonucleotide (2 µl, 1mM in NEBuffer 2) and sterile, distilled water (6 µl) and this HiFi assembly reaction mix incubated in a thermocycler (1 hour, 50°C). Assembly reaction products (2 µl) were transformed into competent *E. coli* DH5α cells (NEB) according to the manufacturer's instructions.

2.11.3.2 Generation of editing templates

Editing templates for the generation of in-frame deletions via homologous recombination with the chromosome were designed and created as described in sections 2.10.1 and 2.10.2, using primers containing flanking *Asi*SI and *Asc*I restriction sites. Resulting editing templates were cloned into CRISPR/Cas9 vectors via appropriate restriction digests and ligation reactions.

2.11.4 Generation of in-frame deletion mutants via CRISPR/Cas9 genome editing

CRISPR/Cas9 vectors were transferred into *C. difficile* from the relevant *E. coli* donor strains via conjugative transfer and transconjugant colonies selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. Thiamphenicol-resistant colonies appeared after 48-96 hours and were streaked to purity on the same media. Resulting colonies were screened via colony PCR using appropriate primers. PCR-confirmed mutants were grown overnight in BHIS broth (1 ml) and subsequent growth serially diluted in sterile PBS and spread plated onto BHIS agar supplemented with D-cycloserine and cefoxitin. Distinct resulting colonies were patch plated onto BHIS agar supplemented with D-cycloserine and cefoxitin, with and without thiamphenicol to screen for plasmid loss. Thiamphenicol-sensitive colonies were stored in BHIS with glycerol (10% v/v) at -80°C.

Chapter Three

Elucidating the role of the SpoVA proteins in *C. difficile* 630Δ*erm*

3 Elucidating the role of SpoVA proteins in *C. difficile* 630Δ*erm*

3.1 Introduction

Bacterial endospores are remarkably resilient structures, capable of surviving exposure to a variety of physical and chemical agents, including wet-heat, desiccation, ultraviolet radiation, antibiotics, ethanol and detergents. Previous studies have identified factors which contribute to this highly resilient phenotype as low spore core water content, saturation of DNA in the spore core with small acid-soluble proteins (SASPs), a relatively impermeable spore inner membrane, and high levels of dipicolinic acid present in the spore core. During sporulation, the spore-specific molecule pyridine-2,6-dicarboxylic acid (dipicolinic acid; DPA) enters and accumulates within the spore core (Figure 3.1), to the extent that DPA comprises around 10% of the dry weight of *B. subtilis* spores (Paidhungat *et al*., 2000). This DPA is synthesised in the mother cell compartment solely during sporulation from the lysine biosynthesis pathway, with the final reaction step performed by DPA synthetase, encoded by *dpaA* and *dpaB* within the *spoVF* operon in *B. subtilis* (Daniel & Errington, 1993). From the mother cell compartment, DPA is then transported into the forespore where it accumulates in the spore core and is present in a 1:1 chelate with divalent cations, typically calcium (Setlow *et al*., 2006). Following subsequent commitment to germination, DPA is released from the spore core in the first minutes of germination, resulting in increased spore water content. Furthermore, in *B. subtilis*, DPA activates the cortex-lytic enzyme (CLE), CwlJ, initiating hydrolysis of the peptidoglycan spore cortex which allows the core to expand, returning the core water content to levels found in vegetative cells, thus eventually allowing

the resumption of enzyme activity and metabolism. DPA is also considered to play a role in spore stability, since in *B. subtilis* strains harbouring mutations in *spoVF*, rendering them unable to synthesise DPA, are unable to complete sporulation due to lysis of the spores prior to maturation.

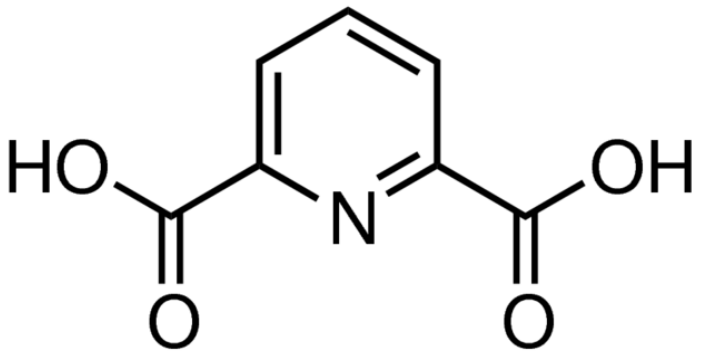


Figure 3-1 Diagrammatic representation of the molecular structure of Dipicolinic acid.

Genes encoding SpoVA proteins have been found in the genomes of almost all Firmicutes studied thus far. Where present, the minimal number of s*poVA* genes exists as a tricistronic operon consisting of *spoVAC*, *spoVAD* and s*poVAE*, as is found in *C. difficile*. However, considerable variation exists amongst members of the spore-forming Firmicutes in the form of the *spoVA* operon(s) present (Figure 3.2). For example, *B. subtilis* contains a heptacistronic *spoVAABCDEbEaF* operon, whilst *Clostridium carboxidivorans* contains two *spoVA* operons, a tricistronic *spoVAC*DE operon and a heptacistronic *spoVA* operon (Paredes-Sabja *et al.*, 2011).

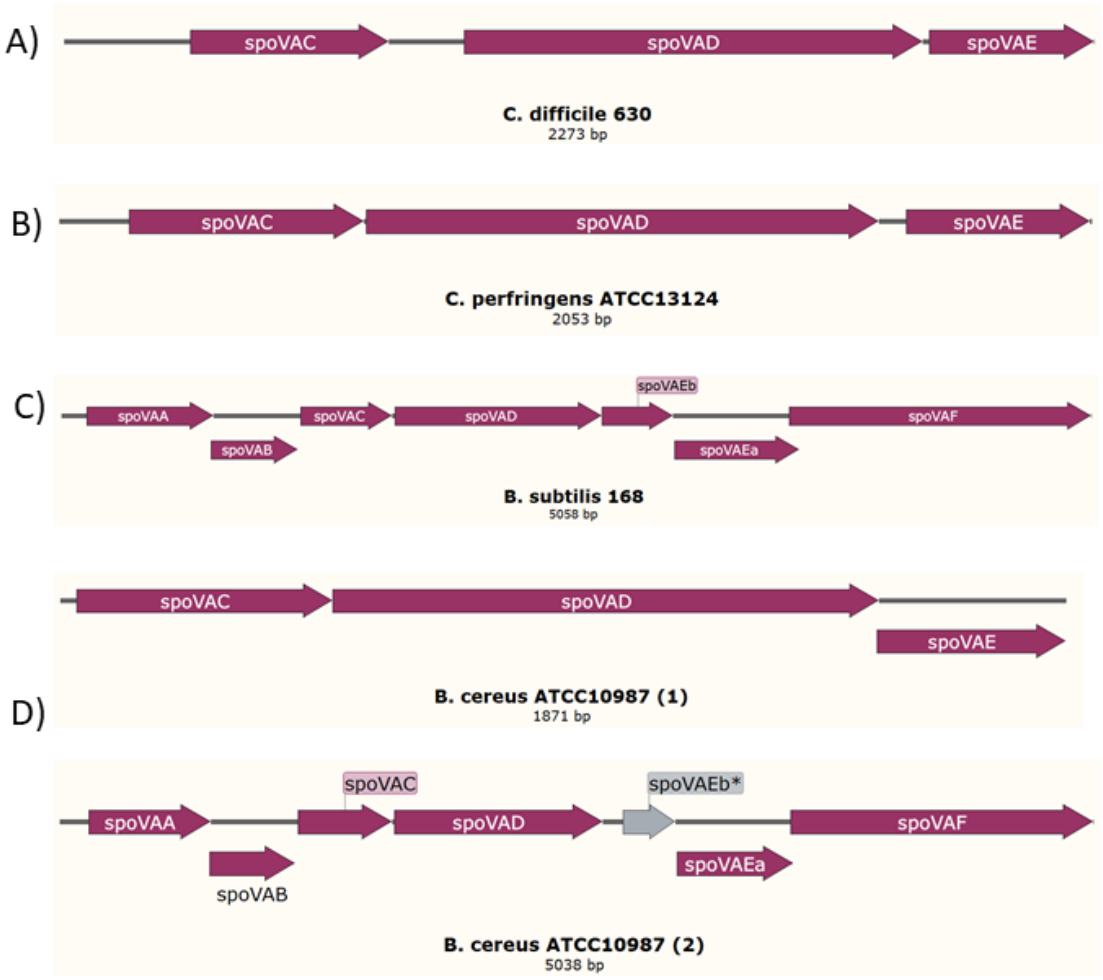


Figure 3-2 Diagrammatic representation of *spoVA* operons present within Firmicutes.

Structure of the *spoVA* operons present in *C. difficile* 630Δ*erm* (A) and three representative Firmicutes with either a tricistronic *spoVACDE* operon (*C. perfringens*, [B]), heptacistonic *SpoVAABCDEbEaF* operon (*B. subtilis*, [C]) or both operons (*B. cereus*, [D]). Purple arrows indicate ORFs, grey line denotes intergenic sequence. Offset arrows indicate overlapping ORFs. The grey arrow denotes the frameshifted *spoVAEb* gene present in *B. cereus* ATCC10987.

The transcription of *spoVA* genes solely within the developing forespore occurs immediately prior to DPA synthesis in the mother cell compartment (Tovar-Rojo *et al.*, 2002). Evidence for the role of SpoVA proteins in transporting DPA into the forespore during sporulation include; i) *B. subtilis* spores lacking *spoVA* and either *sleB* or *ger3* (deletion of *sleB* or all three

germinant receptors allows for stable *B. subtilis* spores lacking *spoVA* to be produced) were unable to accumulate DPA and had the same lower level of wet heat resistance as spores lacking *ger3* and *spoVF* (DPA synthetase), ii) supplementation of this *spoVA ger3* strain with exogenous DPA during sporulation did not lead to an increase in spore DPA levels, however the same experiment with *spoVF ger3* spores restored spore DPA levels to 53% of the wildtype level (Tovar-Rojo *et al*., 2002) and iii) spores of *C. perfringens* lacking *spoVA* were stable, but were unable to accumulate DPA nor calcium ions, and had two-fold higher water content in the spore core compared with wildtype spores (Paredes-Sabja *et al*., 2008). The link between SpoVA proteins and the *B. subtilis* germination receptors is an interesting one, however its relevance to this study is not clear, given the lack of *ger* receptor homologues within *C. difficile* (Ross & Abel-Santos, 2010). Previous work in *C. perfringens* is perhaps more pertinent to this study, given *C. perfringens* possesses a single tricistronic *spoVACDE* operon as is found in *C. difficile*. *C. perfingens spoVA* spores, which lacked DPA, were more susceptible to treatments with wet heat, UV radiation, hydrochloric acid, formaldehyde and hydrogen peroxide (Paredes-Sabja *et al.*, 2011). Furthermore, this study found that levels of the α/β-type SASPs in spores lacking *spoVA* were almost identical to levels for wildtype *C. perfringens* spores. This is particularly interesting given the previous findings in *B. subtilis* which identified SASPs as the key determinant for spore resistance to formaldehyde and hydrogen peroxide (Setlow *et al*., 2000). The non-specific binding of α/β-type SASPs to spore DNA, saturating the spore chromosome, protects spores against wet heat, UV radiation (Mason & Setlow, 1986), formaldehyde and hydrogen peroxide (Setlow & Setlow, 1993) treatments. Previous works detailing the resistance properties of spores and the mechanisms of spore death in response to various physical and chemical agents are summarised in Chapter One. In summary, DPA is a highly important molecule with roles in spore stability, resistance against various physical and chemical agents, and in some species direct activation of CLEs required to complete germination.

3.1.1 Aims of this project

Sporulation and germination are essential stages in the life cycle of *C. difficile*, yet much of our understanding of these processes comes from studies in the model spore-former *Bacillus subtilis*. However, several recent studies have suggested key differences in the regulation and timing of components within these processes exist between *B. subtilis* and *C. difficile*. This study aims to elucidate the role of the SpoVA proteins in *C. difficile* through the creation of strains lacking individual, and all three, *spoVA* gene(s).

3.2 The *C. difficile spoVA* operon

3.2.1 630Δ*erm spoVA* operon structure

C. difficile 630Δ*erm* contains three *spoVA* genes arranged in a putative tricistronic operon, namely, *spoVAC* (CD630_0773), *spoVAD* (CD630_0774) and *spoVAE* (CD630_0775), shown in Figure 3.3. This *spoVA* locus is described as a putative operon despite it containing a 166 bp intergenic region between *spoVAC* and *spoVAD*, which is larger than is found in other Firmicutes. To investigate whether an additional promoter sequence is located in this intergenic region, the DNA sequence of the entire *spoVA* operon plus the 276 bp intergenic region upstream of *spoVAC* was inputted into Softberry BPROM, an online bioinformatics tool for the identification of bacterial promoter sequences. Two intergenic promoters were subsequently identified, the location of the -35 and -10 elements of which are highlighted in Figure 3.3. Hence, the 630Δ*erm spoVA* genes do not appear to be arranged within a single operon.



Figure 3-3 Diagrammatic representation of the 630Δ*erm spoVA* 'operon'.

Purple arrows indicate ORFs, orange and yellow boxes indicate -35 and -10 boxes of the predicted promoter regions identified using Softberry BPROM, respectively. Grey box represents the terminator sequence predicted by the ARNold online tool. Grey line denotes intergenic regions. All elements are annotated to scale.

3.2.2 Analysis of the 630Δ*erm* SpoVA protein sequences

Amino acid (aa) sequences for each of the 630Δerm *spoVA* genes were entered into the EMBL-EBI InterPro online tool for the detection of protein domains within each sequence. Both the 145 aa SpoVAC and 120 aa SpoVAE proteins from 630Δ*erm* possess four transmembrane (TM) helix domains interspersed regularly throughout the protein sequence. Meanwhile, the 336 aa SpoVAD protein sequence contains a central domain, with predicted catalytic activity, belonging to the thiolase-like superfamily. No transmembrane domains were predicted within the SpoVAD protein sequence. These results are consistent with studies in *C. perfringens*, which also found no TM-domains in SpoVAD and four each in SpoVAC and SpoVAE (Paredes-Sabja *et al*., 2011). These SpoVA translated protein sequences were then aligned with SpoVA sequences from selected other Firmicutes to investigate the homology of the 630Δ*erm* sequences with those from other spore-forming organisms (Table 3.1).

Organism	*SpoVA* operon(s)	Protein	Length (aa)	Homology
B. subtilis 168	*ABCDEaEbF*	SpoVAC	150	69/145 (48%)
		SpoVAD	338	155/336 (46%)
		SpoVAEb	116	49/116 (42%)
C. perfringens ATCC	*CDE*	SpoVAC	155	75/145 (52%)
		SpoVAD	339	183/336 (54%)
		SpoVAE	121	59/120 (49%)
Bacillus cereus ATCC10987	(1) *CDE*	SpoVAC	158	64/145 (44%)
		SpoVAD	338	155/336 (46%)
		SpoVAE	116	48/116 (41%)
	(2) *ABCDEb*EaF*	SpoVAC	152	67/145 (46%)
		SpoVAD	338	160/336 (48%)
		SpoVAEb*	85	38/85 (45%)

Table 3-1 Homology of *spoVA* genes within selected Firmicutes.

Translated peptide sequences of *spoVACDE* genes from selected Firmicutes possessing either a tricistronic, heptacistronic or both *spoVA* operons were aligned with the equivalent *spoVACDE* translated peptide sequences of 630Δ*erm* using the pBLAST algorithm to determine the number and percentage identities (homology) between these sequences, the lengths of which are measured in amino acids (aa). Asterisked annotation denotes the frameshift mutation present in a *B. cereus* ATCC10987 *spoVAEb* gene.

3.3 Generation of *C. difficile* 630Δ*erm spoVA* insertional mutants using ClosTron

The ClosTron system of insertional mutagenesis was utilised to generate mutants of each of the three *spoVA* genes (CD630_0773, CD630_0774 & CD630_0775) within *C. difficile* 630Δ*erm* as described in Materials and Methods. Putative, erythromycin-resistant mutants were PCR screened using flanking primers which anneal to chromosomal regions upstream and downstream of the intended insertion sites within each gene. A junction PCR screen was performed to amplify across the insertion site using the EBS universal primer, which anneals within the Group II intron, and either the forward or reverse flanking primer depending upon the sense/anti-sense orientation of the intron insertion (Figure 3.4). Observed PCR products for the junction PCR screen for putative mutants of each *spoVA* gene were of the expected sizes in each case. Similarly, the observed flanking PCR products for the putative ClosTron mutants were 4.366 kb for *spoVAC*, 2.768 kb for *spoVAD* and 2.624 kb for *SpoVAE*, again as expected.

Insertion of the Group II intron at the intended site within each *spoVA* gene was confirmed via Sanger sequencing of purified junction PCR products with the same primers used to generate these PCR products. In each case, the group II intron had inserted within the *spoVA* gene at the intended site.

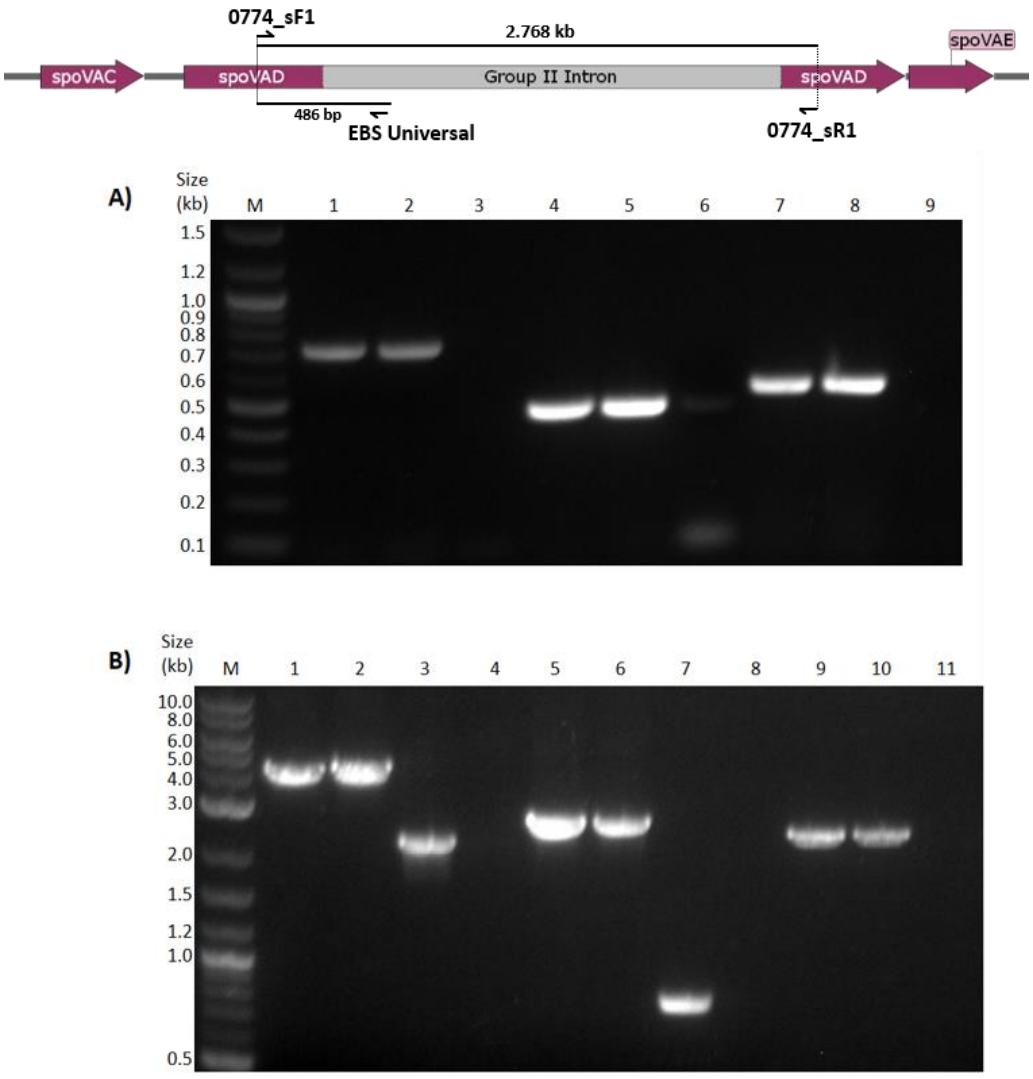


Figure 3-4 PCR screening for insertion of the ClosTron group II intron into each *C. difficile* 630Δ*erm spoVA* gene

Top: Diagrammatic representation of PCR primer binding sites for screening of insertion of the Group II intron into *spoVAD*. A) Junction PCR screen using gDNA isolated from two erythromycin-resistant colonies following ClosTron insertional mutagenesis into *spoVAC* (lanes 1 and 2), *spoVAD* (lanes 4 and 5) and *spoVAE* (lanes 7 and 8). Negative controls for each PCR reaction containing distilled water instead of gDNA template are present in lanes 3, 6 and 9. M: DNA marker (2-log ladder; NEB). The faint band present in the negative control lane 6 is presumed to be due to an overspill from lane 5. B) Flanking PCR screen using the same gDNA as in (A) to screen for insertion of the 2.056 kb Group II intron into *spoVAC* (lanes 1-4), *spoVAD* (lanes 5-8) and *spoVAE* (lanes 9-11). Controls using the 630Δ*erm* gDNA are present in lanes 3, 7 and 11. Negative controls using distilled water are present in lanes 4 and 8. M: DNA marker (2-log ladder; NEB).

3.3.1 Confirmation of single insertion of the Group II intron into 630Δ*erm*

Genomic DNA was prepared from each *spoVA* ClosTron mutant and separately digested with *Eco*RI and *Hin*dIII restriction enzymes, in preparation for Southern Blot hybridisation using an intron-specific probe to confirm single insertion of the Group II intron within each strain (Figure 3.5). Whilst the resulting Southern Blot was far from perfect, it was of sufficient quality to confirm that each *spoVA* ClosTron mutant contained only a single insertion of the Group II intron, as indicated by the presence of single bands in each genomic DNA digest. A very faint band could also be observed, as expected, in the pMTL-007C-E2 control digest lanes.

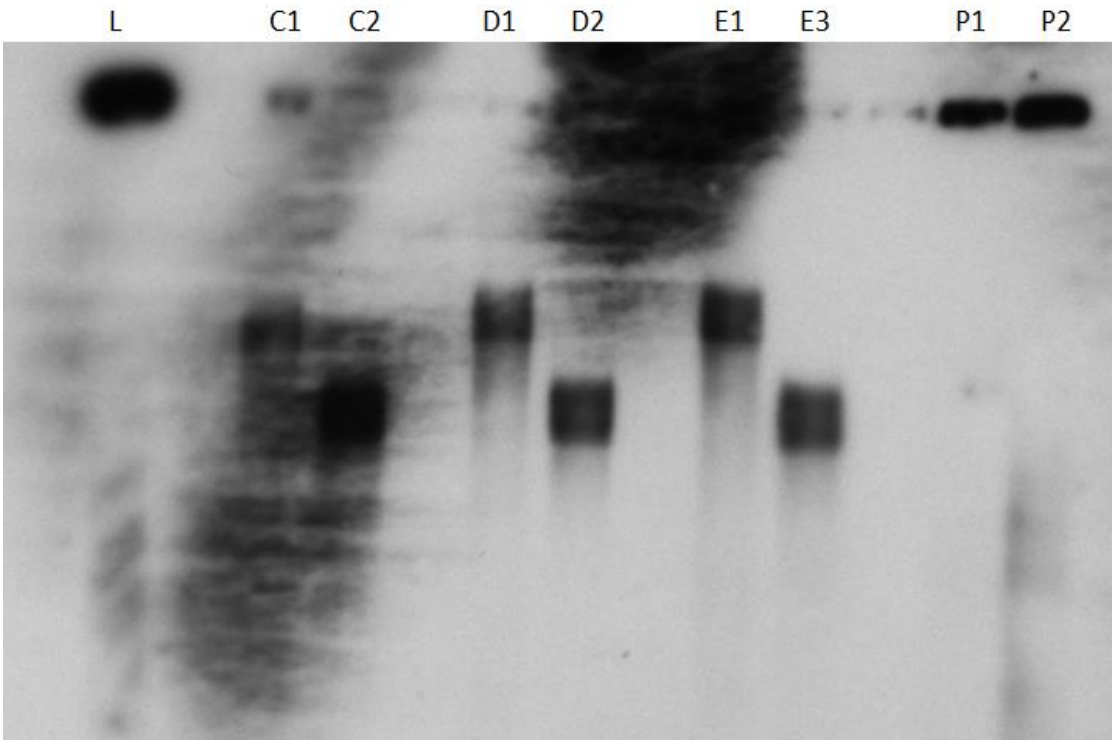


Figure 3-5 Southern Blot hybridisation using an intron-specific probe.

The control plasmid pMTL007C-E2 (P) and genomic DNA of ClosTron insertional mutants within *spoVAC* (C), *spoVAD* (D) and *spoVAE* (E) digested with *Eco*RI (1) or *Hin*dIII (2) were verified by Southern Blot using an intron-specific probe, as described in Materials and Methods. L denotes the λ DNA ladder.

3.4 Characterisation of *C. difficile* 630Δ*erm*::*spoVA* insertional mutants

The impact of insertional mutagenesis of the three *C. difficile spoVA* genes was assessed according to the established protocol developed in our research group, summarised in the flow diagram in Figure 3.6.

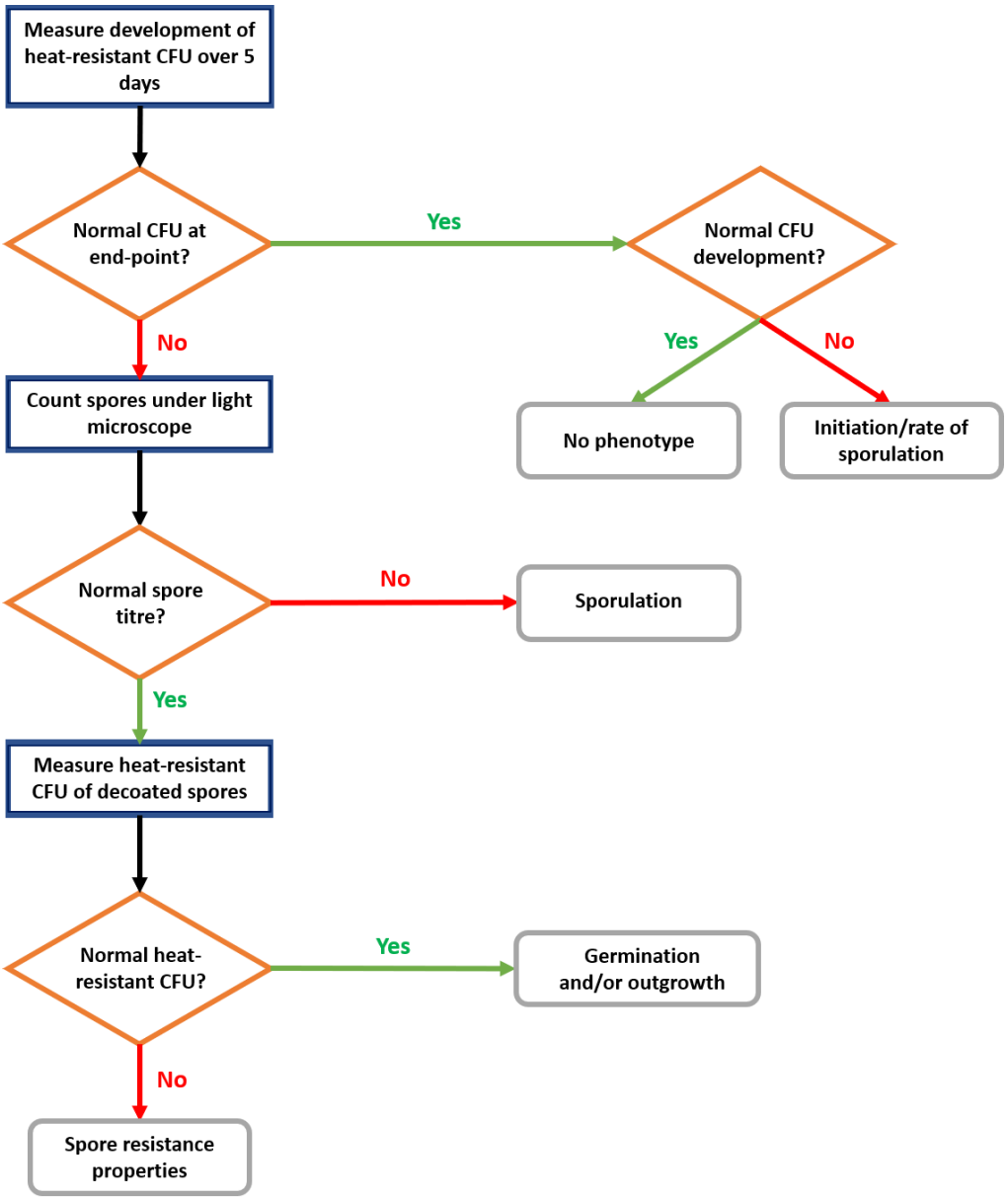


Figure 3-6 Diagrammatic representation of the phenotypic characterisation of *C. difficile* mutants defective in sporulation and/or germination

Performing the outlined experiments will identify defects in rate of sporulation, total sporulation, spore germination/outgrowth and spore viability. Adapted from (Burns, 2011).

3.4.1 Development of heat-resistant CFU over five days

The ability of each *spoVA* Clostron insertional mutant, along with positive (parental *C. difficile* 630Δ*erm*) and negative (630Δ*erm*Δ*spo0A*) control strains, to form heat-resistant CFU on BHIS agar supplemented with the germinant taurocholic acid over five days was assessed using the method described in 2.6.9 (Figure 3.7). Each resulting CFU corresponds to a spore which has survived the heat-treatment step and successfully completed germination and outgrowth to produce the observed vegetative cells. All three *spoVA* insertional mutants were severely impeded in their ability to form heat-resistant CFU, with such colonies only appearing after 96 hours, 72 hours later than the parental positive control. Furthermore, the levels of heat-resistant CFU.ml^{-1} for all three *spoVA* ClosTron mutants after 120 hours were at least 13,000-fold lower relative to 630Δ*erm*. At this 120 hour time-point, mean levels of heat resistant CFU.ml^{-1} were 2.26x10^{6} for 630Δ*erm*, and 105.56, 77.78 and 172.22 for the ClosTron insertional mutants of genes *spoVAC*, *spoVAD* and *spoVAE*, respectively. No heat-resistant CFU were observed as expected for the 630Δ*erm*Δ*spo0A* negative control during this assay, as this in-frame deletion strain lacks a functional copy of the *spo0A* gene, the master regulator of sporulation.

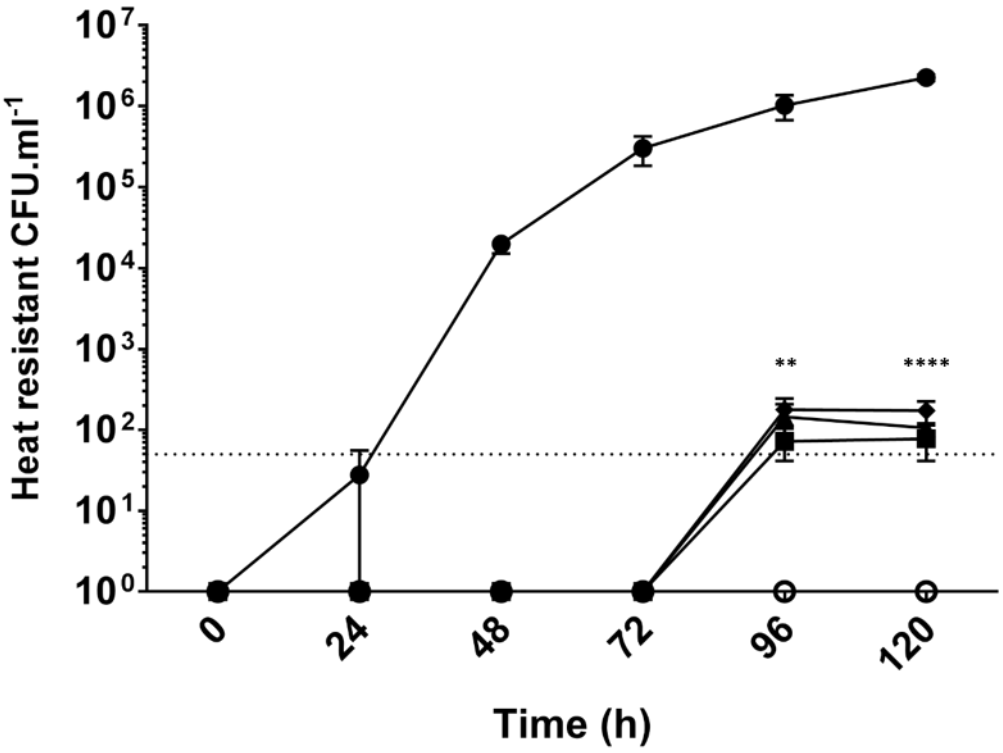


Figure 3-7 Development of heat resistant CFU of *C. difficile* strains over five days.

●, *C. difficile* 630Δ*erm*; ▲, 630Δ*erm*::*spoVAC*; ■, 630Δ*erm*::*spoVAD*; ◆, 630Δ*erm*::*spoVAE*; **O**, 630Δ*erm*Δ*spo0A*. Symbols represent the mean values from three independent experiments. Values depicted as 10^0 should be interpreted as zero. Error bars indicate standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (**) denotes a *P*-value <0.01 and (****) denotes a *P*-value <0.0001.

3.4.2 Spore titres after five days

The reduction in heat-resistant CFU observed for each *spoVA* ClosTron mutant could result from a number of deficiencies, outlined in figure 3.6, including; i) a reduced number of spores produced; ii) a deficiency affecting germination of the spores produced; and/or iii) a reduction in tolerance of the spores to heat treatment. To determine whether this observed reduction in developed heat-resistant CFU over five days was the result of *spoVA* mutants producing fewer spores, aliquots were taken from the sporulation cultures after five days growth assayed in 3.4.1 and spores of each strain counted under phase contrast microscopy (Figure 3.8). The parental 630Δ*erm* strain produced a spore titre of 2.27×10^7 spores.ml^{-1} after five days growth in BHIS agar, whilst the ClosTron insertional mutants in *spoVAC*, *spoVAD* and *spoVAE* produced spore titres, expressed as a percentage of the parental value, of 35.2%, 53.0% and 76.8%, respectively. No spores were observed for the sporulation control strain 630Δ*erm*Δ*spo0A* as expected. These results indicate that whilst differences in spore titre were observed, the number of spores produced by the individual *spoVA* ClosTron mutants cannot explain the scale of reduction of heat-resistant CFU observed in Figure 3.7. Furthermore, these results suggest the *spoVAC* insertional mutant produced fewer spores than did the *spoVAD* ClosTron mutant, which in turn produced fewer spores than the *spoVAE* mutant. Whilst these results may be genuine, due to the insertional nature of ClosTron mutagenesis, polar effects of insertions in the upstream gene may also be a factor.

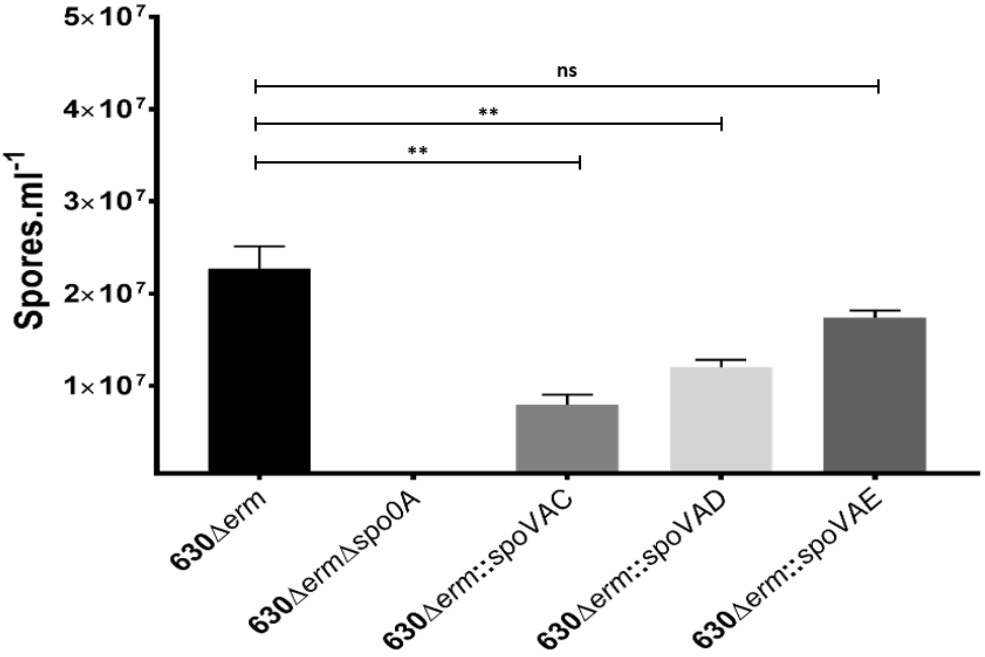


Figure 3-8 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy.

C. difficile strains grown in BHIS broth for five days, after which aliquots were loaded onto a haemocytometer, spore counts performed and the numbers of spores.ml^{-1} calculated. Bars indicate mean values from three independent experiments. Error bars indicate the standard errors of the means. The detection limit for spore counts was 5x10^3 spores. Zero spores were observed for the sporulation control strain 630Δ*erm*Δ*spo0A*. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (**) denotes a *P*-value <0.01, whilst (ns) denotes non-significant *P*-values >0.05.

3.5 Generation of in-frame *spoVA* deletion mutants using Allelic Exchange

One drawback of the ClosTron mutagenesis system is the potential for polar effects to be observed, particularly following insertion of the Group II intron to an operon. The arrangement of the *spoVA* genes within a putative operon suggests that the insertion in s*poVAC* may also impact upon *spoVAD* and *spoVAE*, reducing our ability to determine the role of each SpoVA protein. To overcome this issue, this study aimed to generate in-frame deletion mutants of the entire *C. difficile spoVA* operon and each gene within it via allelic exchange.

The allelic exchange vector pMTL-YN3, previously used to generate in-frame deletion mutants in strain 630Δ*erm* (Ng *et al*., 2013), was modified to target the three individual *C. difficile spoVA* genes and the whole operon. For each target, homology arms were designed to leave only the first two and final three codons of the gene to be deleted, or the first two codons of *spoVAC* and final three codons of *spoVAE* in the case of the whole operon deletion. This design, leaving the start and stop codons intact, results in the production of a non-functional 4-aa peptide and prevents downstream effects from the native promoter of the deleted gene. Designed homology arms were PCR amplified from 630Δ*erm* genomic DNA using the primers listed in Table 2.3. PCR generated left and right homology arms for each target were then combined using SOE PCR, as outlined in 2.7.5, to produce knockout cassettes with flanking 5' *Sbf*I and 3' *Asc*I restriction sites. All four knockout cassettes underwent separate restriction digests with *Sbf*I and *Asc*I enzymes, ligation reactions with similarly digested pMTL-YN3 vector backbone fragments and the ligation products were transformed into electrocompetent *E. coli* TOP10 cells. The generated vectors, pMTL-YN3_*spoVA* (whole operon deletion; figure 3.9), pMTL-YN3_*spoVAC*, pMTL-YN3_*spoVAD* and

pMTL-YN3_*spoVAE* were purified from *E. coli* TOP10 and the sequences of the knockout cassette confirmed correct in each case via Sanger sequencing, using primers YN3_sF1 and YN3_sR1.

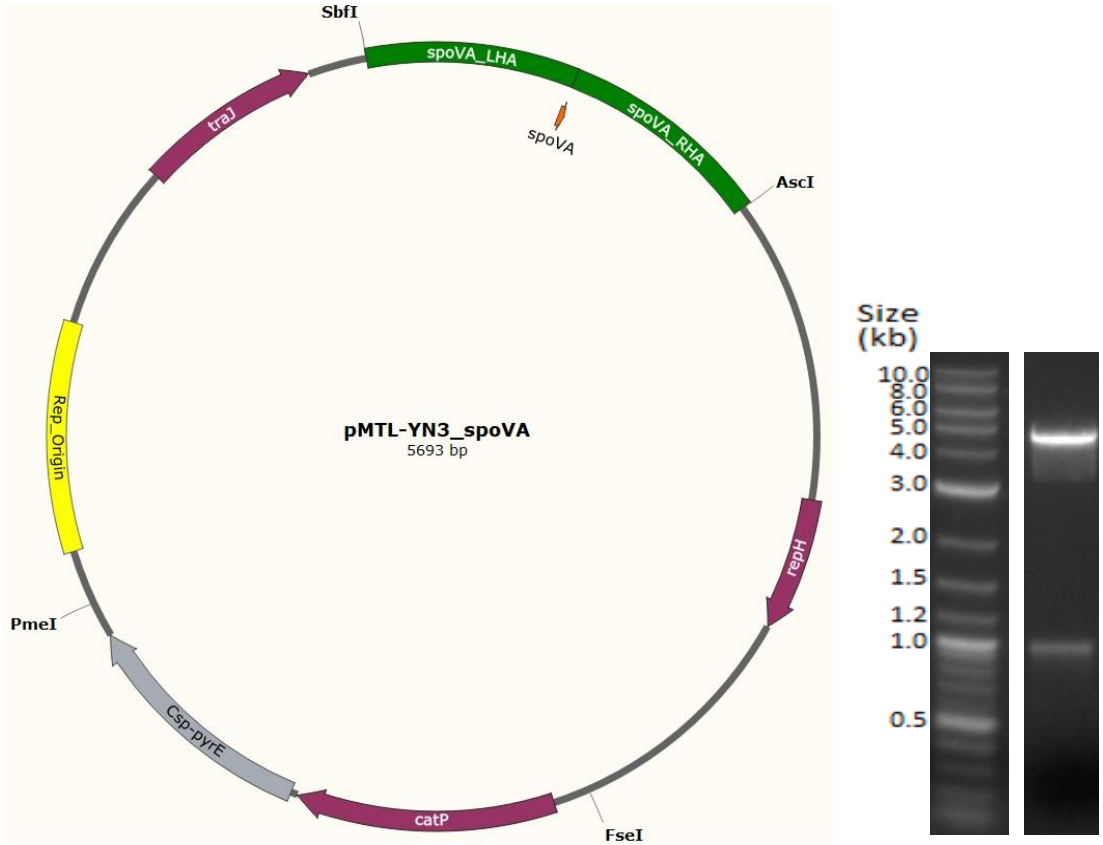


Figure 3-9 Diagrammatic representation of allelic exchange vector pMTL-YN3_*spoVA*.

Left: Plasmid map for pMTL-YN3_*spoVA*, the allelic exchange vector designed to generate an in-frame deletion of the *C. difficile* 630Δ*erm spoVA* operon. *spoVA*_LHA and *spoVA*_RHA denote the left and right homology arms of the *spoVA* knockout cassette, respectively; *spoVA*, fifteen bp sequence fusion of first two codons of *spoVAC* and final three codons of *spoVAE*; *repH*, defective Gram-positive replicon pCB102; *catP*, chloramphenicol acetyltransferase gene; csp-*pyrE*, heterologous *pyrE* allele from *C. sporogenes* necessary for negative/counter selection; *traJ*, plasmid transfer regulator. Vectors pMTL-YN3_*spoVAC*, pMTL-YN3_*spoVAD* and pMTL-YN3_*spoVAE* differ from this vector sequence only in the knockout cassette sequence present between *Sbf*I and *Asc*I restriction sites. Right: Agarose gel electrophoresis following restriction digestion with *Asc*I and *Sbf*I of pMTL-YN3_*spoVA*.

To generate the desired *spoVA* deletions, the four allelic exchange vectors were separately transformed into the conjugation donor strain *E. coli* CA434 and chloramphenicol-resistant strains harbouring each vector were obtained. Individual conjugation reactions were performed using the generated strains with overnight cultures of the *C. difficile* recipient, 630Δ*erm*Δ*pyrE*. Plating of the conjugal cell suspensions on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol was used to select against *E. coli* cells and *C. difficile* cells not containing an allelic exchange vector. For each conjugation reaction, thiamphenicol-resistant transconjugant colonies were obtained and passaged onto the same media to allow additional time for the allelic exchange vectors to integrate into the *C. difficile* chromosome. Subsequent larger colonies, indicative of a single crossover chromosomal integration having occurred, were re-streaked to purity and genomic DNA extracted from overnight cultures. To confirm whether single crossover integrations had occurred in these pure, faster-growing colonies, extracted genomic DNA was used as template for screening PCRs. Two PCRs per strain were performed using primer pairs YN3_sF1/Y and YN3_sR1/X, where X and Y refer to primers which anneal to flanking chromosomal regions upstream and downstream, respectively, of the specific knockout cassette being studied (Table 2.3; data not shown). Despite multiple attempts, no single crossover integrants were obtained following conjugations with vectors pMTL-YN3_*spoVAC* nor pMTL-YN3_*spoVAE*. However, chromosomal integration was observed for three of the four strains harbouring pMTL-YN3_*spoVA* and two of the five 630Δ*erm*Δ*pyrE* strains harbouring pMTL-YN3_*spoVAD*. To allow the second crossover to occur, resulting in excision of the plasmid from the chromosome, pure, PCR-confirmed, single crossover strains were plated onto non-selective BHIS agar plates. After four days, growth was harvested and plated in serial dilutions onto *C. difficile* minimal media (CDMM) agar plates supplemented with 5-fluoroorotic acid (5-FOA). Resulting colonies, which lack a functional copy of *pyrE* and therefore are resistant to 5-FOA, were patch plated onto BHIS agar with and without supplemented thiamphenicol, to confirm

plasmid loss had occurred. Genomic DNA was extracted from true, double crossover (thiamphenicol-sensitive) colonies and used to screen for the desired deletion using primers annealing to chromosomal regions immediately flanking the respective knockout cassette sequence. Unfortunately, no double crossover strains for pMTL-YN3_*spoVAD* single crossovers could be obtained, despite repeated attempts. Of the ten 5-FOA-resistant, thiamphenicol-sensitive, putative whole *spoVA* operon deletion mutants screened using primers 0773_sF1 and 0775_sR1, two produced the 1.906 kb sized band corresponding to an in-frame deletion of *spoVA* (data not shown).

The two PCR products of 1.906 kb were purified and Sanger sequenced using primers 0773_sF1 and 0775_sR1, which confirmed the in-frame deletion of the *spoVA* operon had occurred as intended, generating two 630Δ*erm*Δ*pyrE*Δ*spoVA* strains. Before this strain could be characterised, it was necessary to convert it back to prototrophy (*pyrE*+) to allow for comparisons with the well-studied 630Δ*erm* strain. This was performed using the *pyrE* repair vector pMTL-YN1, which uses allele coupled exchange to replace the truncated *pyrE* with the wildtype *pyrE* allele. pMTL-YN1 was transferred into a Sanger sequencing confirmed 630Δ*erm*Δ*pyrE*Δ*spoVA* strain via conjugation and thiamphenicol-resistant, transconjugant colonies obtained on BHIS agar plates supplemented with D-cycloserine, cefoxitin, thiamphenicol and uracil. Faster-growing transconjugants were passaged twice onto CDMM agar plates without supplemented uracil to isolate colonies containing a wildtype *pyrE* allele and which no longer require uracil supplementation to grow. Genomic DNA was isolated from four such colonies and used as the template for PCR screens to confirm the presence of a wildtype *pyrE* allele, using primer pair CD630_pyrD_sF1/CD630_0189_sR3, and PCR screens to confirm these strains were *spoVA* deletion mutants, using primer pair 0773_sF1 and 0775_sR1 (Figure 3.10).

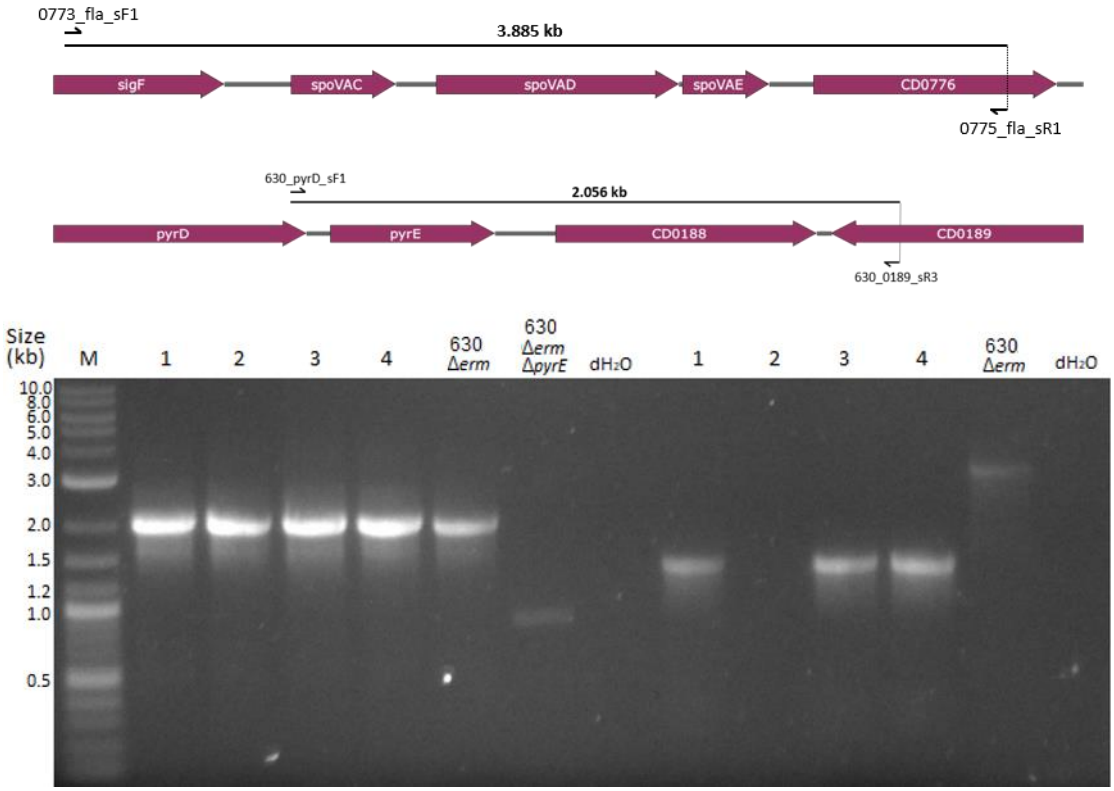


Figure 3-10 Screening PCR to confirm *spoVA* deletion and *pyrE* restoration.

Four thiamphenicol-sensitive 630Δ*erm*Δ*spoVA* colonies PCR screened following allele-coupled exchange with *pyrE* repair vector pMTL-YN1. M: DNA Marker (2-log ladder; NEB); Numbered lanes, gDNA extracted from thiamphenicol-sensitive strains; 630Δ*erm***,** 630Δ*erm* gDNA; 630Δ*erm*Δ*pyrE*, 630Δ*erm*Δ*pyrE* gDNA; dH_2O, distilled water. Left seven lanes: screening PCRs using primer pair CD630_pyrD_sF1/CD630_0189_sR3; Right six lanes: screening PCRs using primer pair 0773_sF1 and 0775_sR1.

Following PCR screening with primers CD630_pyrD_sF1/CD630_0189_sR3, all four colonies generated a 2.056 kb fragment identical in size to that generated from the 630Δ*erm* genomic DNA, indicating successful repair of *pyrE*. PCR screening with primers 0773_sF1 and 0775_sR1 confirmed three of these strains lacked a functional copy of the *spoVA* operon, as these strains produced a 1.606 kb sized band compared with the 3.885 kb band generated from 630Δ*erm* genomic DNA. This *spoVA* screening PCR failed for the second strain (lane 2, figure 3.10) failed, hence the other three strains were taken forward for further characterisation.

3.5.1 Complementation of the 630Δ*erm*Δ*pyrE*Δ*spoVA* mutant

The complementation of mutants is essential in attributing resulting phenotypes to the generated mutation. To complement the 630Δ*erm*Δ*pyrE*Δ*spoVA* in-frame deletion mutant, the *spoVA* operon, including the native promoter upstream of *spoVAC* was PCR amplified using primers VACDE_1C_F3 and VACDE_1C_R1. This generated a 2.276 kb DNA fragment with *Not*I and *Xho*I restriction sites at the 5' and 3' ends respectively. Restriction digests with *Not*I and *Xho*I and subsequent ligations were performed to clone this fragment into the complementation vector pMTL-YN1C, generating pMTL-YN1C_*spoVA* (Figure 3.11). This *spoVA* complementation vector was transformed into *E. coli* CA434 and resulting chloramphenicol-resistant strains were used as conjugal donors to transfer pMTL-YN1C_*spoVA* into 630Δ*erm*Δ*pyrE*Δ*spoVA*. The *pyrE* repair protocol outlined in the previous section was then repeated with the subsequent thiamphenicol-resistant, transconjugant colonies. To confirm successful complementation, that the *spoVA* operon had been inserted at the *pyrE* locus concomitant with *pyrE* repair, genomic DNA was prepared from five colonies able to grow on CDMM plates without uracil supplementation and PCR screens performed using primer pair CD630_pyrD_sF1/CD630_0189_sR3 (Figure 3.12). DNA fragments 4.332 kb in length were amplified from all five strains tested, indicating successful complementation had occurred. Three of these 4.332 kb DNA fragments were purified from the agarose gel and Sanger sequenced using primers 0773_sF1, 0774_sF1, 0775_sF1 and 0774_sR1. All resulting sequences obtained for all three fragments were as expected, thus confirming successful integration of the *spoVA* operon at the *pyrE* locus of 630Δ*erm*Δ*pyrE*Δ*spoVA* and therefore the generation of the complemented strain, CRG5235 (*C. difficile* 630Δ*erm*Δ*spoVA*::*spoVA*).

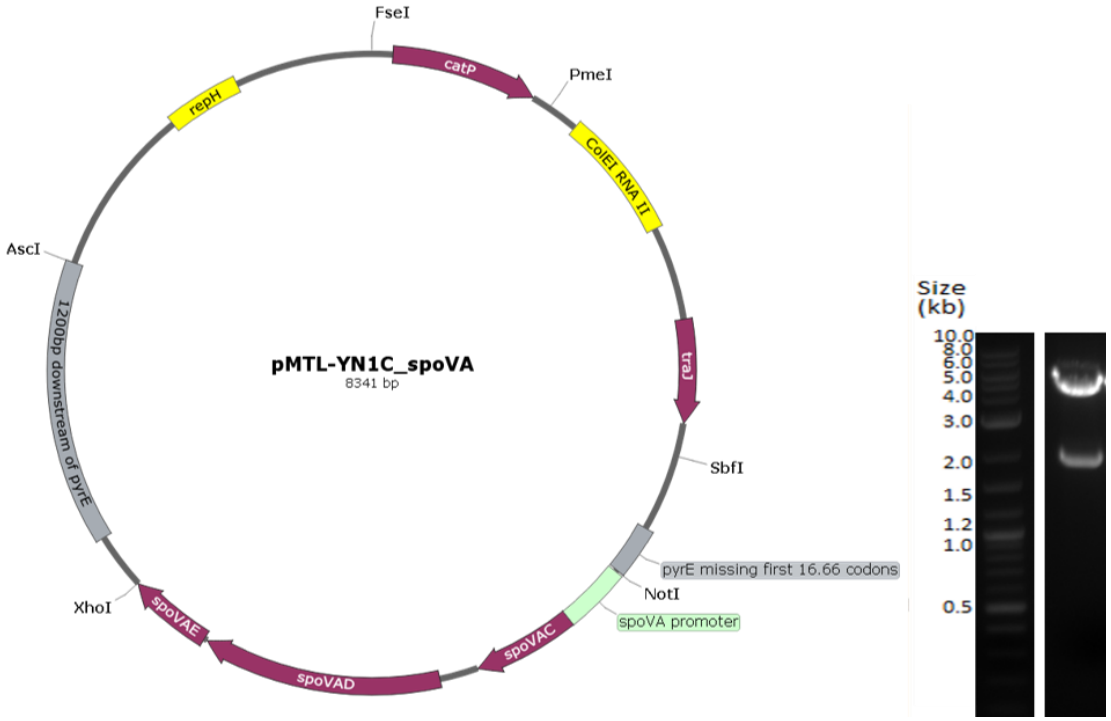


Figure 3-11 Diagrammatic representation of the complementation vector pMTL-YN1C_*spoVA*.

Left: The native *spoVA* operon from 630Δ*erm* is cloned between *Not*I and *Xho*I sites of the complementation vector pMTL-YN1C, in between the short and long homology arms responsible for repair of *pyrE*. Right: Agarose gel electrophoresis following restriction digestion with *Not*I and *Xho*I of pMTL-YN1C-*spoVA*.

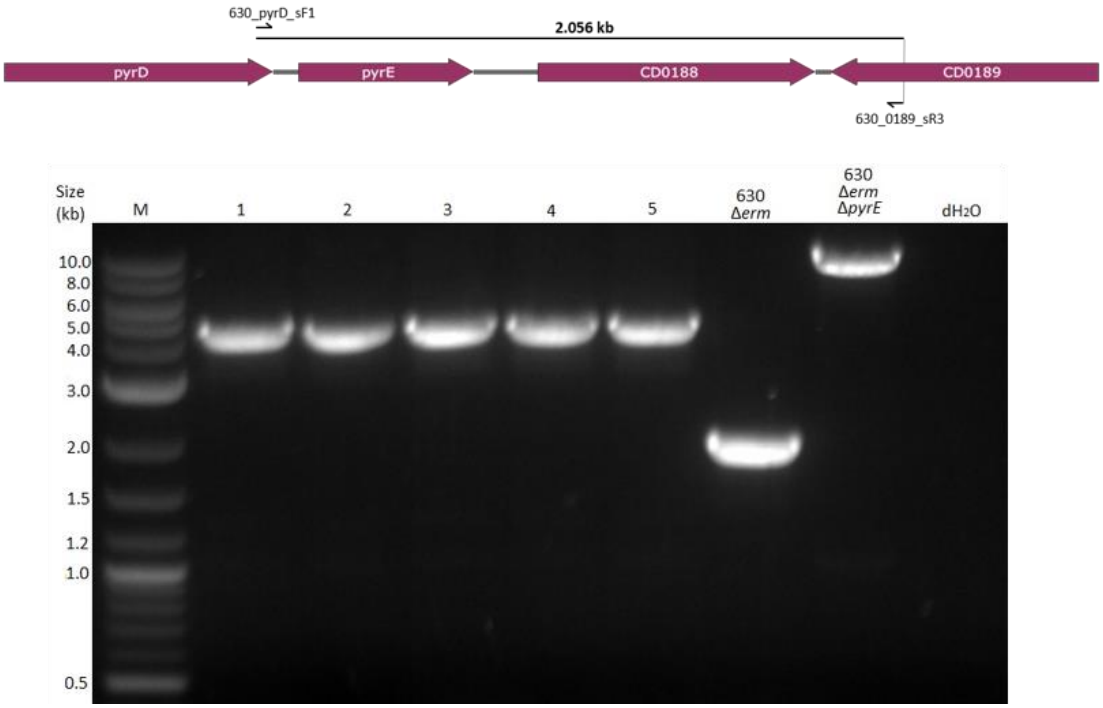


Figure 3-12 Screening PCR to confirm complementation of *spoVA* deletion mutants.

Five thiamphenicol-sensitive 630Δ*erm*Δ*spoVA* colonies PCR screened using primer pair CD630_pyrD_sF1/CD630_0189_sR3 following allele-coupled exchange with the complementation vector pMTL-YN1C_*spoVA*. M: DNA Marker (2-log ladder; NEB); Numbered lanes, gDNA extracted from thiamphenicol-sensitive strains; 630Δ*erm*, 630Δ*erm* gDNA; 630Δ*erm*Δ*pyrE*, 630Δ*erm*Δ*pyrE* gDNA; dH_2O, distilled water.

3.6 Characterisation of in-frame deletion *spoVA* mutant and complemented strains

3.6.1 Development of heat-resistant CFU over five days

The ability of the *pyrE*-repaired, whole operon *spoVA* in-frame deletion mutant (CRG5234) and its complemented counterpart (CRG5235) to develop heat-resistant CFU over five days was determined (Figure 3.13). Mean heat-resistant CFU after 120 hours were $1.28x10^6$ for the parental 630Δ*erm* strain, $9.33x10^5$ for the complemented mutant CRG5235 and 94.44 for the *spoVA* operon mutant, CRG5234. As was observed for the individual *spoVA* ClosTron

mutants, CRG5234 was able to form heat resistant CFU, but these only appeared after 96 hours and at levels over 13000-fold lower than for 630Δ*erm*. Meanwhile, the levels of heat-resistant CFU observed for CRG5235 almost exactly mirror levels for 630Δ*erm* at each timepoint, indicating successful complementation of the *spoVA* deletion.

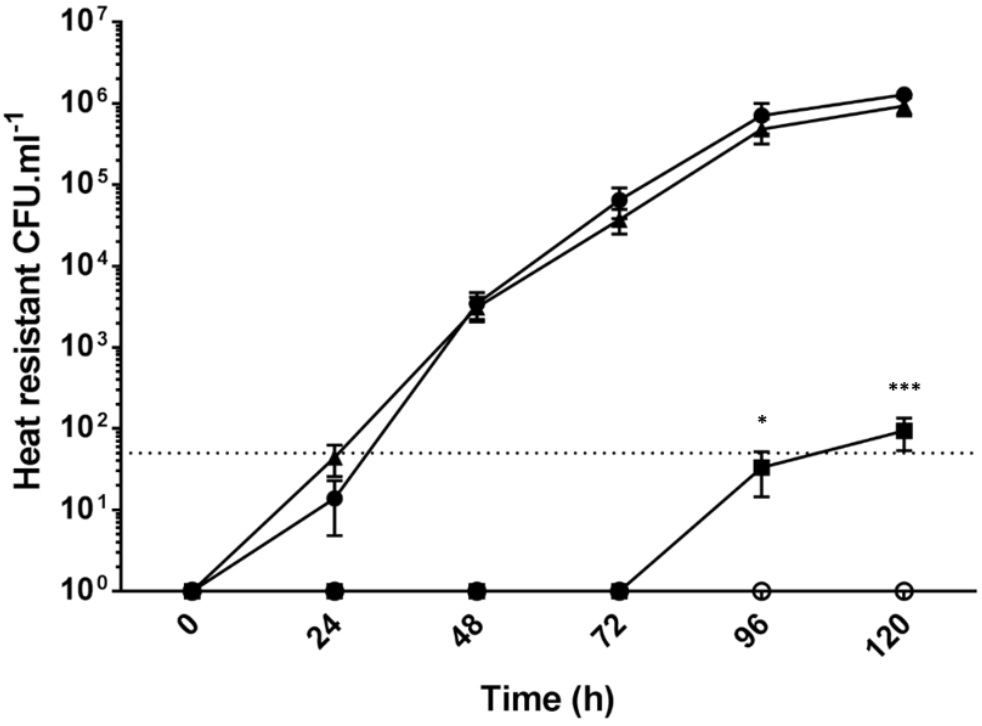


Figure 3-13 Development of heat resistant CFU of *C. difficile* strains over five days.

●, *C. difficile* 630Δ*erm*; ■, CRG5234; ▲, CRG5235; **O**, 630Δ*erm*Δ*spo0A*. Symbols represent the mean values from three independent experiments, performed on two separate occasions (n=6). Values depicted as 10^0 should be interpreted as zero. Error bars indicate standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (*) denotes a *P*-value <0.05 and (***) denotes a *P*-value <0.001.

3.6.2 Five-day spore titres

To determine whether the observed reduction in heat-resistant CFU after five days for CRG5234 relative to 630Δ*erm* was due to a reduction in the number of spores produced by this strain, spores present in five-day sporulation cultures for each strain were counted under light microscopy (Figure 3.14). The number of spores produced by each strain was almost identical with mean spores.ml^{-1} values of 1.83x10^{7}, 1.77x10^{7} and 1.76x10^{7} calculated for strains 630Δ*erm*, CRG5234 (630Δ*erm*Δ*spoVA*) and CRG5235 (630Δ*erm*Δ*spoVA*::*spoVA*), respectively. These results show that spore production is not hindered in CRG5234, which lacks the *spoVA* operon, and are consistent with the results obtained from the individual *spoVA* ClosTron mutants.

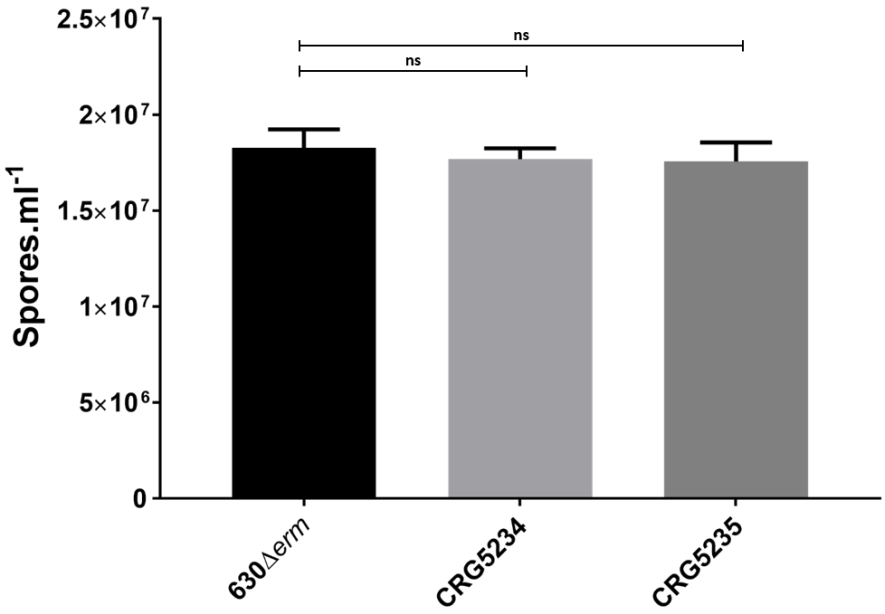


Figure 3-14 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy.

C. difficile strains grown in BHIS broth for five days, after which aliquots were loaded onto a haemocytometer, spore counts performed and the numbers of spores.ml^{-1} calculated. Bars indicate mean values from three independent experiments. Error bars indicate the standard errors of the means. The detection limit for spore counts was 5x10^{3} spores. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05.

3.6.3 Lysozyme treatment of de-coated spores

Given that differences in the number of spores produced cannot explain the differences in heat-resistant CFU development between CRG5235 and 630Δ*erm*, the germination proficiency of spores of these strains was investigated. Five-day sporulation cultures of strains 630Δ*erm*, CRG5234, CRG5235 and a germination deficient control strain, 630Δ*erm*Δ*sleC*, were pelleted and the spore coats chemically removed as described in 2.6.11. De-coated spores were then heat-treated at 65°C for 30 minutes and plated in serial dilutions onto BHIS agar plates with and without supplementation with lysozyme (Figure 3.15). No colonies were observed on media lacking lysozyme for all strains tested, confirming BHIS alone cannot germinate de-coated spores. In the presence of lysozyme, heat-resistant CFU were recovered for all strains except the whole *spoVA* operon mutant, CRG5234. This result indicates that the *spoVA*- spores are susceptible to heat-treatment.

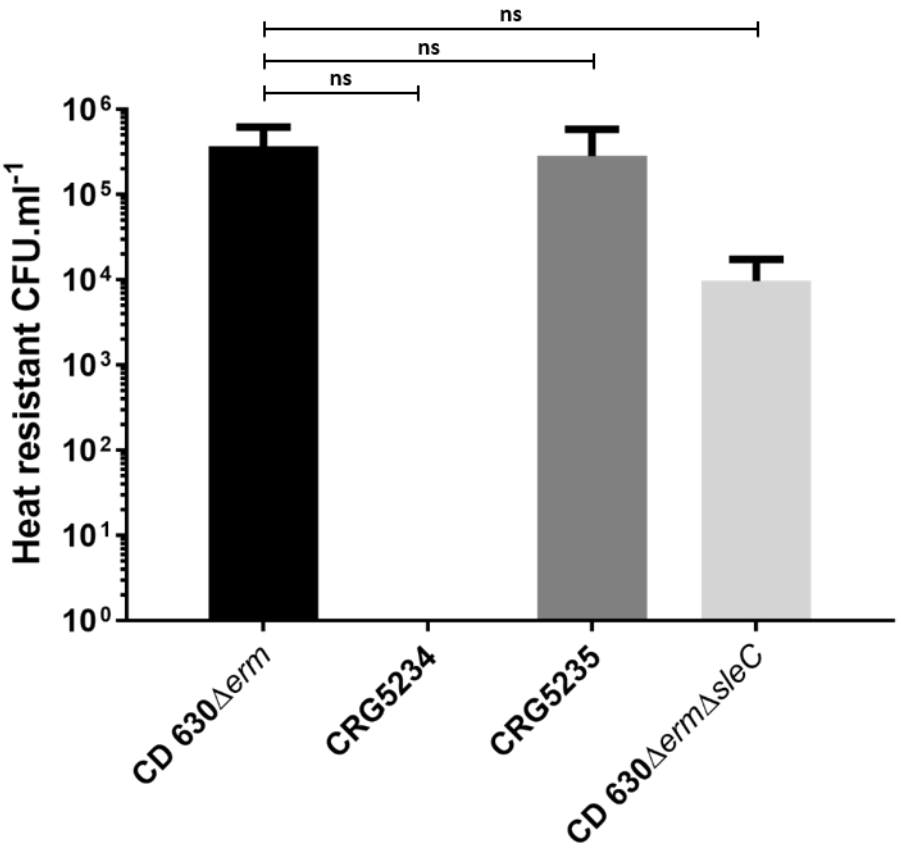


Figure 3-15 Outgrowth of de-coated spores on BHIS agar supplemented with lysozyme.

C. difficile 630Δ*erm* spores from five-day sporulation cultures were de-coated as described in Materials and Methods, heat-treated and plated onto BHIS agar supplemented with lysozyme. Resulting heat-resistant CFU.ml^{-1} values denote the means from three independent experiments. Error bars indicate the standard errors of the means. Due to the logarithmic scale, values of 10^0 should be interpreted as zero. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05.

3.7 Obtaining pure, countable spore preparations

To test the hypothesis that the reduction in heat-resistant CFU developed over five days for CRG5234 is due to reduced heat-resistance properties of spores of this strain, it was first necessary to generate pure spore preparations. Consistent with methods from previous studies, spores of 630Δ*erm*, CRG5234 and CRG5235 were harvested from BHIS agar plates

after five days growth. To obtain pure spore preparations, the harvested cells were washed repeatedly in distilled water at 4°C. After a minimum of ten washes, spores were viewed under light microscopy to confirm spore preparations were >99% free from vegetative cells. Whilst it was possible to obtain pure spore preparations, the majority of spores from all strains were found to be present in clumps of varying sizes (Figure 3.16). This clumping of spores was observed following spore washing from cells harvested from liquid sporulation cultures and from plates. Consequently, it was not possible to enumerate these spores, with the aim of generating aliquots containing equal numbers of spores. Several studies use optical density measurements of spore preparations to approximate the number of spores present, however the clumping of spores observed in this study would inevitably greatly affect the accuracy of such optical density measurements. As such, methods for producing un-clumped spore preparations were investigated.

3.7.1 Sonication of spore preparations

Whilst sonication has previously been shown to separate clumped spores to produce a homogenous spore suspension, it has also been used to remove the exosporium and to lyse spores. Hence, a sonication protocol which was sufficient for spore separation but did not result in spore damage nor lysis was sought. Using the sonication protocol outlined in Materials and Methods, spore preparations were exposed to sonication for 30 seconds at 4°C after which spore clumping was assessed by viewing aliquots under a phase-contrast microscope. This process was repeated until the minimum amount of sonication to fully separate spore clumps was identified as 4 cycles of sonication on low setting at 4°C, with 30 seconds per cycle (Figure 3.16). These phase-contrast microscopy images show that this sonication treatment was sufficient to separate all clumping for spores of each strain and that the resulting spores remained phase-bright, indicating spore lysis had been avoided.

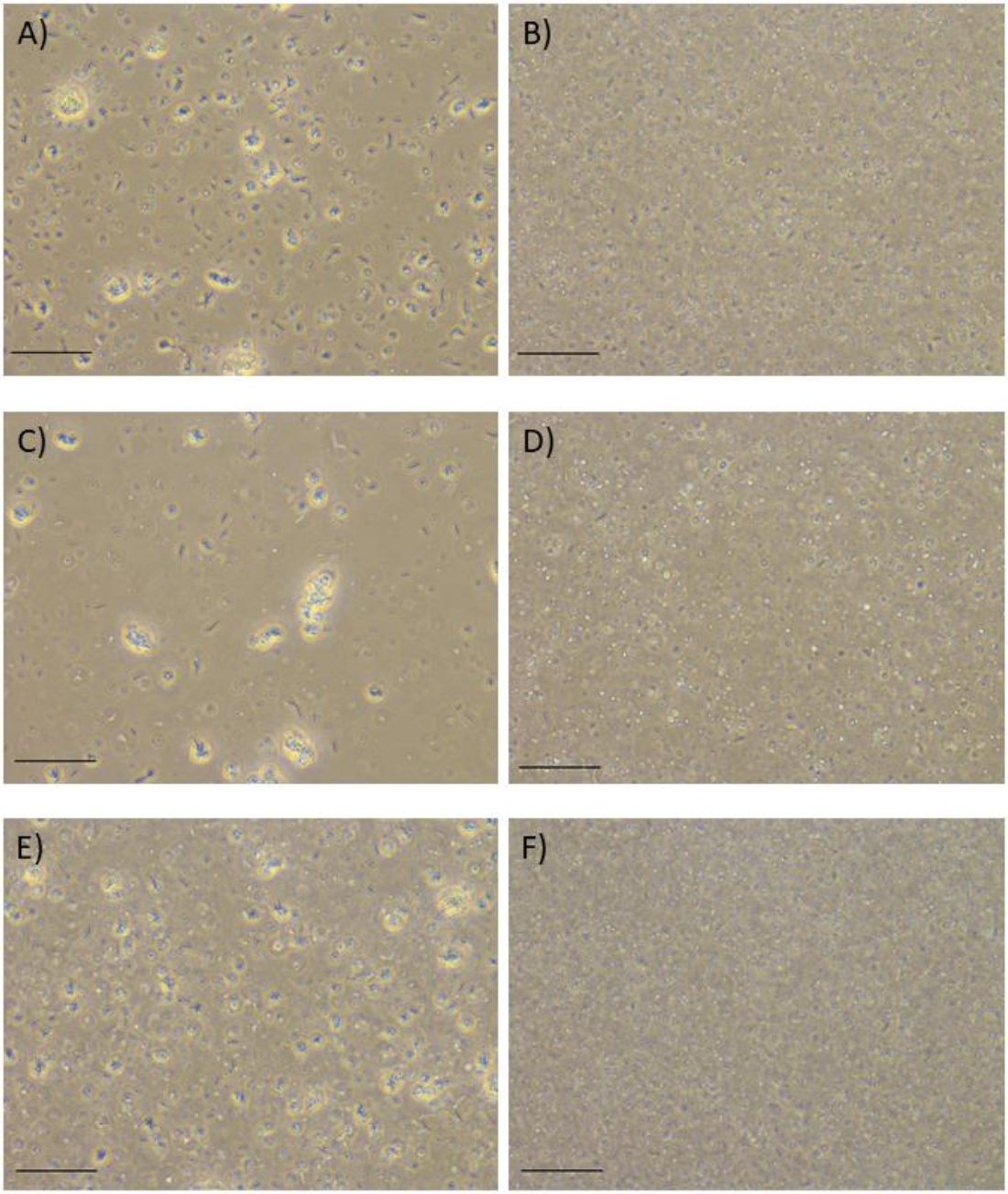


Figure 3-16 Identification of a suitable sonication protocol for spore separation.

Images captured during phase contrast-microscopy of crude spore preparations following sonication, from which the extent of spore clumping was assessed by eye. Crude spore preparations of 630Δ*erm* [(A) and (B)], CRG5234 [(C) and (D)] and CRG5235[(E) and (F)] before sonication [(A), (C) and (E)], and after 4 cycles of sonication for 30 seconds on, 30 seconds off at 4°C [(B), (D) and (F)]. Scale bar represents 200 µm.

Having identified a minimum sonication protocol using crude spore preparations, this was repeated on pure spore preparations. Sonicated spores were then enumerated under phase-contrast microscopy using a haemocytometer and $1x10^7$ spores.ml^{-1} stocks prepared in triplicate for strains 630Δ*erm*, CRG5234 and CRG5235. Preliminary experiments were performed using these stocks to investigate the heat-resistance properties of each, in which aliquots of $1x10^7$ spores.ml^{-1} were heated at 65^{o}C or incubated on the bench for 30 minutes. Both samples for each spore stock triplicate were then plated in serial dilutions onto BHIS agar supplemented with taurocholate and CFU.ml^{-1} values calculated after 24 hours (Figure 3.17). The effects of heat treatment on the parental 630Δ*erm* strain and complemented *spoVA* mutant strain were minimal, whilst after heat-treatment the *spoVA*$^{-}$ strain exhibited 300-fold fewer CFU.ml^{-1} relative to the non-heat-treated sample. These results appear to confirm the hypothesis that spores of CRG5234, lacking a functional *spoVA* operon, are susceptible to heat-treatment. However, it was also noticed that the non-heat-treated CFU.ml^{-1} values for CRG5234 were approximately fifteen-fold lower than equivalent values for the 630Δ*erm* and CRG5235 strains.

The percentage germination of spores of each strain plated onto BHIS agar supplemented with taurocholate without prior heat-treatment was calculated by dividing the mean observed CFU.ml^{-1} values by the starting number of spores, $1x10^7$, and multiplying by 100. The calculated percentage germination values were 23.8% for 630Δ*erm*, 1.43% for CRG5234 and 21.7% for CRG5235. This suggests that either exposure to sonication has negatively affected the subsequent ability of the spores lacking *spoVA* to germinate and outgrow to produce CFU in the presence of germinants, or that spores of CRG5234 are deficient in germination relative to their parental counterparts. Hence, an alternative method of spore washing which generates countable spore preparations free from clumping was sought. To

determine whether sonication affected spore heat resistance properties, a control consisting of approximately 1x10^7 non-sonicated 630Δ*erm* spores was included in the heat-treatment assay. These results show the proportion of heat-resistant CFU to non-heat-treated CFU were highly similar for sonicated 630Δ*erm* spores (74.59%) as for non-sonicated 630Δ*erm* spores (68.75%). Hence, sonication does not appear to have altered spore heat-resistance properties for the parental 630Δ*erm* strain.

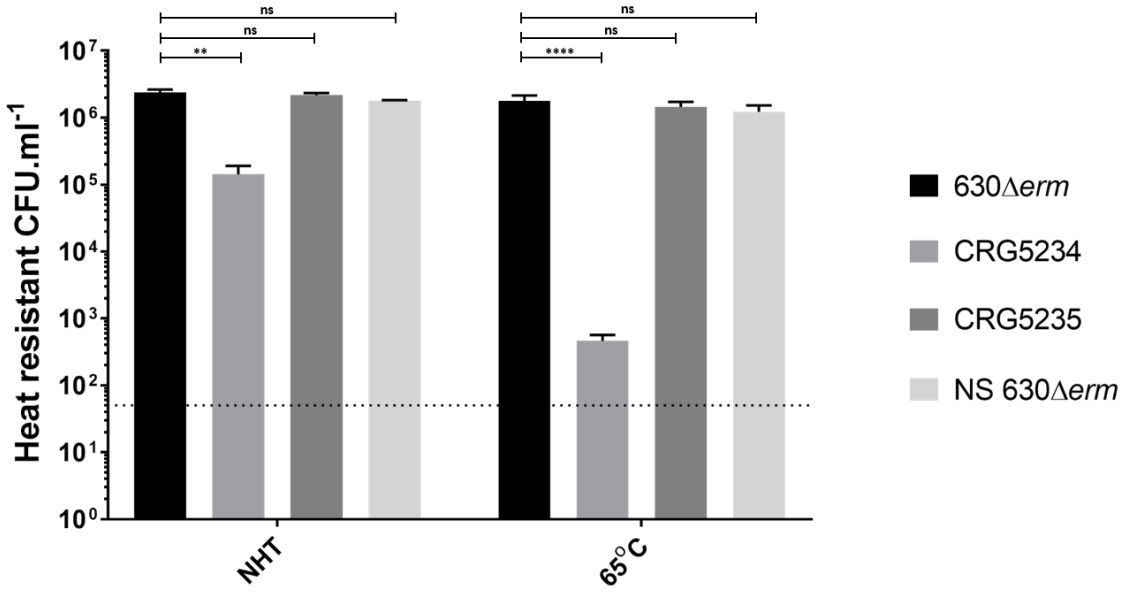


Figure 3-17 Effects of heat-treatment on sonicated spore preparations.

1x10^7 *C. difficile* spores separated by sonication were exposed to heat treatment at 65°C or incubated on bench (not heat-treated; NHT), then plated onto BHIS agar supplemented with taurocholate. Resulting CFU.ml^{-1} values were calculated after 24 hours incubation. Approximately 1x10^7 630Δ*erm* spores which had not been sonicated (NS) were included as a control. Bars denote mean values from three independent experiments. Error bars indicate the standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit of the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (**) denotes a *P*-value <0.01 and (****) denotes a *P*-value <0.0001. Not significant (ns) denotes a *P*-value >0.05.

3.7.2 Spore separation using Tween-80

Having observed that even the minimum amount of sonication necessary to generate a homogenous spore suspension was potentially detrimental to the outgrowth of *spoVA*$^{-}$ spores without heat-treatment, an alternative method was sought to separate clumped spores. A study by Warda *et al*. (2016) used chilled phosphate buffer supplemented with gradually decreasing concentrations of Tween80 to prevent clumping of *Bacillus cereus* spores during spore washing. To test whether this method was suitable for generating homogenous *C. difficile* spore suspensions, fresh sporulation cultures of 630Δ*erm*, CRG5234 and CRG5235 were established, cells harvested as previously, and spore washing performed using either PBS alone, or PBS supplemented with decreasing concentrations of Tween80, as detailed in 2.6.8. The resulting spore preparations were assessed for clumping and purity under a phase-contrast microscope (Figure 3.18). Spore washing in chilled PBS alone was not sufficient to prevent spore clumping, however the addition of Tween80 in concentrations ranging from 0.1% to 0.01% prevented spore clumping in all strains prepared. The resulting concentrated spore stocks were stored in PBS with 0.01% Tween80 at 4^{o}C and the number of spores.ml^{-1} in each spore stock triplicate was enumerated under phase-contrast microscopy, from which diluted stocks containing 1x10^{8} spores.ml^{-1} were prepared.

Aliquots containing 1x10^{7} spores of each strain were plated onto BHIS agar supplemented with taurocholate to determine whether spores lacking *spoVA* are deficient in germination, or whether the previously observed reduction in CFU for spores of CRG5234 was an artefact resulting from sonication treatment (Figure 3.19). Mean CFU.ml^{-1} values were 2.43x10^{6} for 630Δ*erm*, 2.62x10^{6} for CRG5234 and 3.68x10^{6} for CRG5235, indicating that *spoVA*- spores are not deficient in germination relative to the parental and complemented mutant strains. Furthermore, these values correspond to germination efficiency percentages of 24.3% for

630Δ*erm*, 26.2% for CRG5234 and 36.8% for CRG5235, which are consistent with equivalent values for sonicated spore preparations for strains 630Δ*erm* and CRG5235, but for CRG5234 this represents an 18-fold increase relative to the value for sonicated spore preparations.

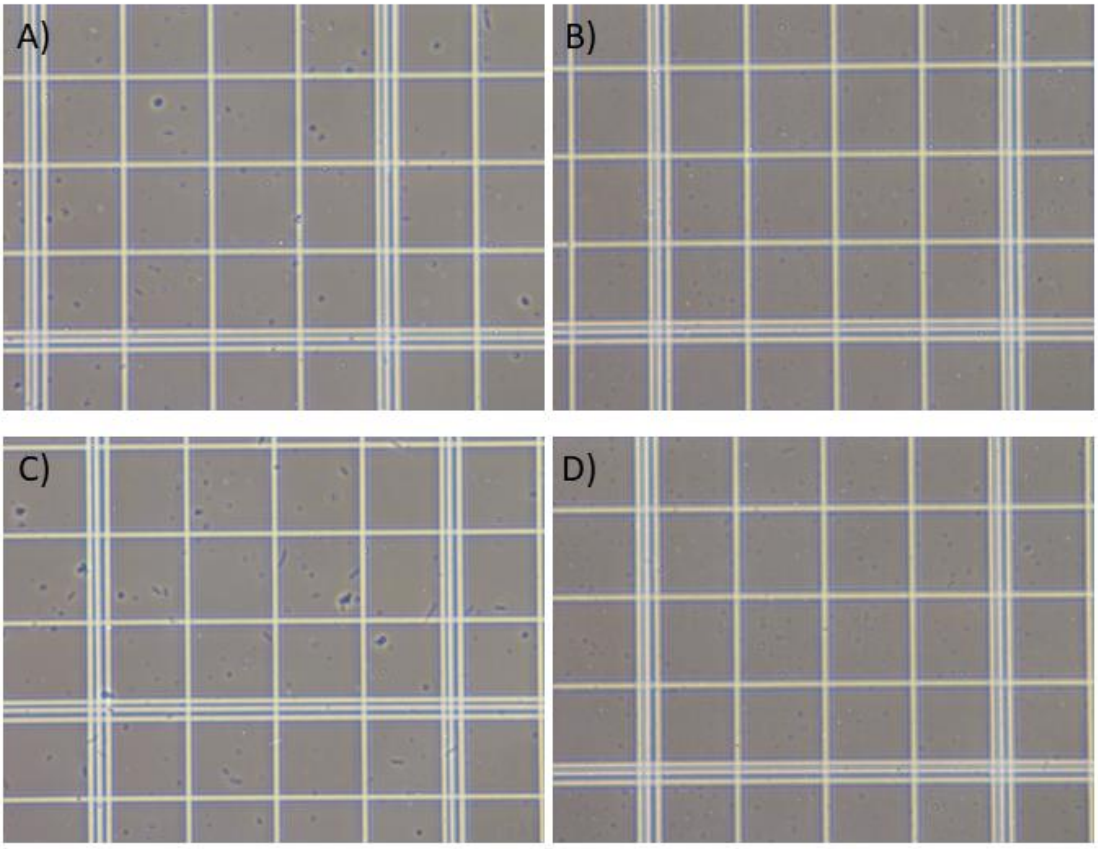


Figure 3-18 Effects of Tween80 on spore clumping assessed under phase-contrast microscopy.

C. difficile 630Δ*erm* [(A) and (B)] and CRG5234 [(C) and (D)] spores washed in PBS [(A) and (C)] or PBS supplemented with gradually decreasing Tween80 concentrations [(B) and (D)], from 0.1% down to 0.01% Tween80, diluted and viewed under phase-contrast microscopy. The diameter of each small box in the haemocytometer is 200 µm.

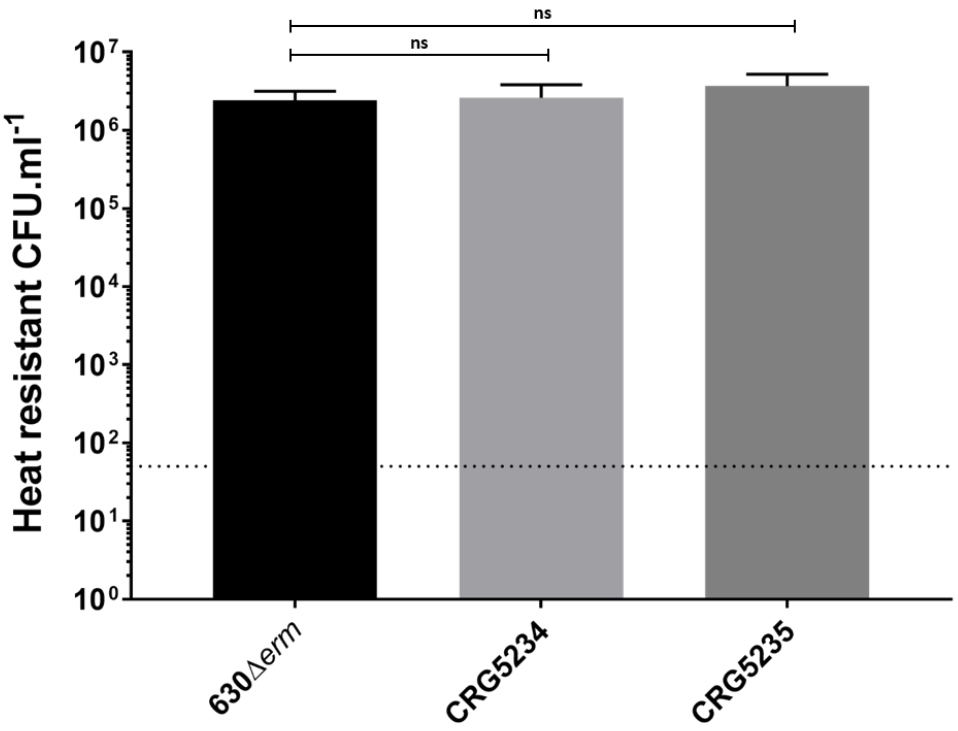


Figure 3-19 CFU development following spore washing with PBS and Tween80.

1x10^7 *C. difficile* spores suspended in PBS with 0.01% Tween80 were serially diluted and spotted onto BHIS agar supplemented with taurocholate. CFU.ml^{-1} values were calculated after 24 hours incubation. Bars represent mean values from three independent experiments. Error bars indicate the standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05.

3.8 Assessing the dipicolinic acid content of *spoVA*$^{-}$ spores

Previous studies have identified that products of the *spoVA* operon in *B. subtilis* and *C. perfringens* are essential for the accumulation of dipicolinic acid (DPA) in the spore core during sporulation. Hence, assays were performed using the whole operon and individual *spoVA* mutant strains to determine the DPA content of these spores, relative to the parental and fully complemented strains, to assess whether the three *spoVA* genes of *C. difficile* perform a similar function.

3.8.1 Spore calcium content

The majority of DPA in the spore core is present chelated with calcium ions, hence measuring the calcium content of the spore can be used as an indirect measurement of the amount of DPA present. Separated spores were sent for analysis by our collaborators Solenne Marion and Barbora Bartova in Lausanne, Switzerland who used STEM-EDS to measure the atomic percentages of various elements in spores of the parental strain 630Δ*erm*, CRG5234 and CRG5235 (Figure 3.20; Table 3.2). CRG5234 spores were found to contain a reduced mean atomic percentage of calcium relative to 630Δ*erm* and the complemented mutant, CRG5235. This reduction in the intensity of calcium signals following STEM-EDS of CRG5234 spores is in stark contrast to the concentrated signals observed in the cores of 630Δ*erm* and CRG5235 spores. From this, we can suggest that spores of CRG5234 also contain a lower concentration of DPA, given previous findings which showed that levels of DPA mimic the mean atomic concentration of calcium in spores of selected Firmicutes (Jamroskovic *et al*., 2016). Whilst these results give an indication that CRG5234 spores contain less calcium than their 630Δ*erm* and CRG5235 counterparts, these results must not

be overstated, given the large variation in the atomic percentages of carbon and nitrogen also observed between spores of these strains.

Strain	Mean Atomic Percentage					
	Carbon	Nitrogen	Sodium	Phosphorus	Calcium	Magnesium
630Δ*erm*	22.99 (± 3.04)	18.69 (± 1.38)	6.43 (± 0.46)	5.85 (± 0.84)	4.20 (± 0.80)	0.53 (± 0.13)
CRG5234	12.69 (± 3.45)	27.45 (± 2.12)	3.79 (± 0.49)	5.28 (± 0.75)	1.89 (± 0.46)	0.53 (± 0.04)
CRG5235	33.58 (± 6.62)	18.99 (± 2.95)	3.68 (± 0.84)	3.66 (± 0.46)	4.51 (± 1.25)	0.64 (± 0.18)

Table 03-2 Elemental composition of *C. difficile* spores determined using STEM-EDS.

C. difficile spores were loaded onto carbon grids prior to analysis using Scanning Transmission Electron Microscopy to Energy Dispersive Spectroscopy (STEM-EDS). Values represent mean atomic percentages from measurements of six spores of each strain. Values in parentheses denote the standard errors of the means. All measurements were performed by Solenne Marion and Barbora Bartova of the Swiss Federal Institute of Technology in Lausanne (EPFL), Lausanne, Switzerland.

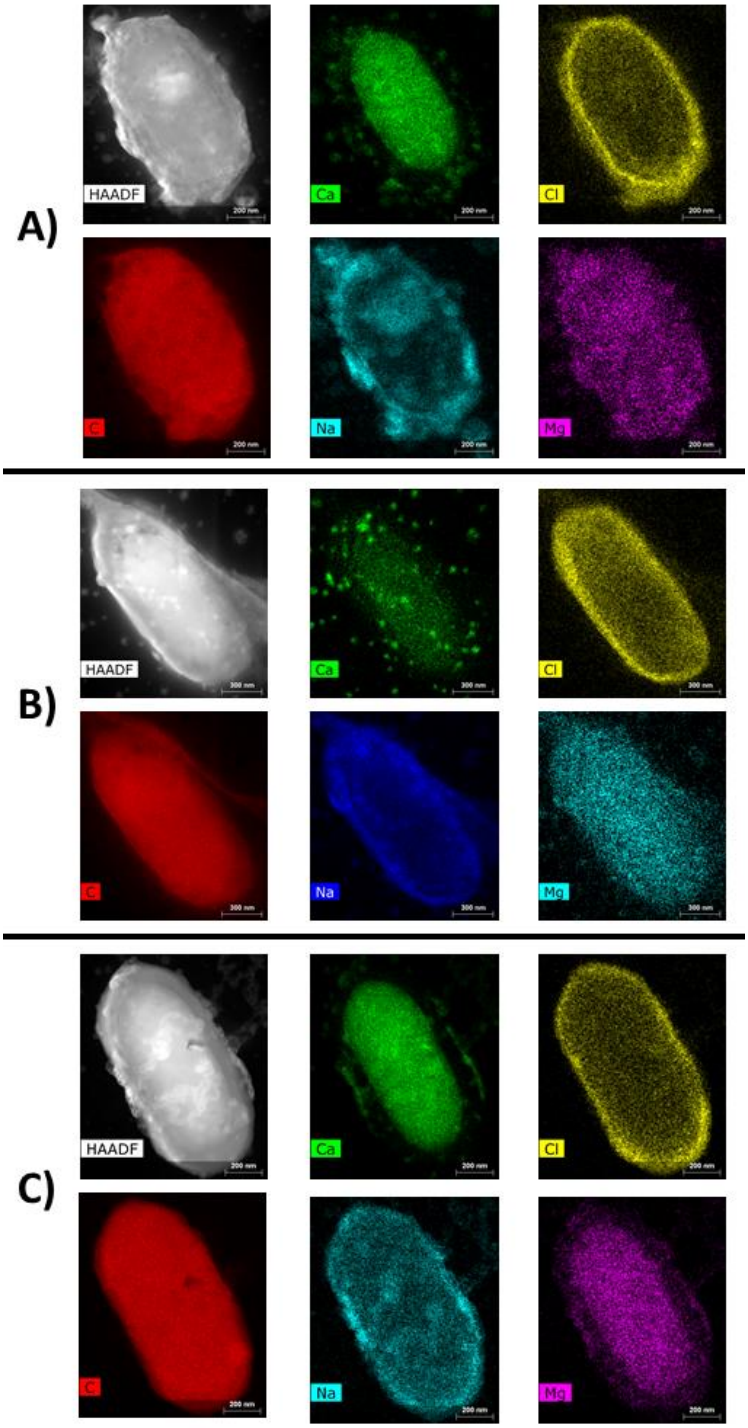


Figure 03-20 STEM-EDS detection of *C. difficile* spore elemental composition.

Scanning Transmission Electron Microscopy to Energy Dispersive Spectroscopy (STEM-EDS) imaging of *C. difficile* 630Δ*erm* (A), CRG5234 (B) and CRG5235 (C) spores, including high angle angular dark field microscopy (HAADF-STEM). All data and images captured by Solenne Marion and Barbora Bartova of the Swiss Federal Institute of Technology in Lausanne (EPFL), Lausanne, Switzerland.

3.8.2 Spore DPA content measurements using Terbium chloride fluorescence

Previous studies investigating the DPA content of spores have measured the absorbance at 270 nm, given that around 80% of the absorbance at this wavelength is a consequence of DPA (Paredes-Sabja *et al*., 2008). Recently however, an improved, more specific method of quantifying the amount of DPA in spores has been developed which relies on fluorescence spectroscopy. In the presence of Terbium (III) chloride, DPA will bind to Terbium which results in a 10,000-fold increase in fluorescence emission at 545 nm for the terbium-DPA complex, relative to when DPA is absent (Ammann *et al*., 2011). Quantification of the DPA content of spores of the parental strain 630Δerm, CRG5234 and all complemented derivatives was performed as outlined in 2.6.15 (Figure 3.21). Mean RFU values expressed as a percentage of the equivalent value for 630Δ*erm* spores were 5.7% for CRG5234 and PSI7475 (Δ*spoVAC*), 4.1% for PSI7375 (Δ*spoVAD*) and PSI7374 (Δ*spoVAE*) and 173.6% for CRG5235. These results show that *C. difficile* 630Δ*erm* spores lacking either one or all three *spoVA* gene(s) are unable to accumulate dipicolinic acid in their spore cores at levels of the parental strain. CRG5235 spores were fully complemented for spore DPA content, containing almost double the amount of DPA compared to 630Δ*erm* spores.

To attempt to determine the DPA content within spores of each strain, a standard curve was plotted from RFU values measured using defined concentrations of dipicolinic acid (Figure 3.22). Between DPA concentrations of 0.1 µM to 100 µM the coefficient of determination (R^2) value was found to be 0.9996, indicating extrapolations between this range would be a good approximation of their true values. Consequently, the concentration of DPA present in $1x10^7$ spores of each strain was approximated via extrapolation from the standard curve, from which the amount of DPA in each spore was then calculated (Table 3.3).

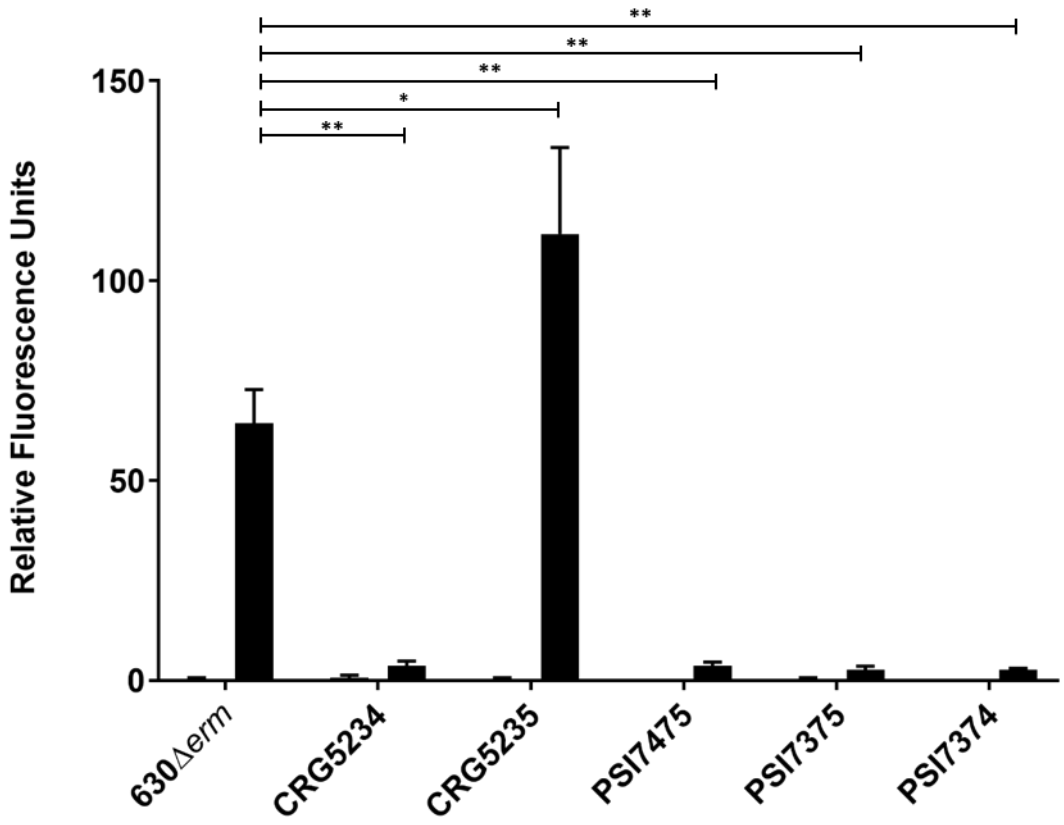


Figure 03-21 Total DPA content of spores measured using Terbium fluorescence.

Terbium fluorescence measured using the supernatants of spore preparations heated for one hour at 100°C (filled bars) or 37°C (empty bars). Bars indicate the relative fluorescence units (RFU) after background fluorescence from spores of each strain not supplemented with terbium was deducted from the respective value obtained with terbium. Bars represent mean RFU values from three independent experiments. Error bars indicate the standard errors of the means. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. One asterisk (*) denotes a *P*-value <0.05, while two asterisks (**) denotes a *P*-value <0.01.

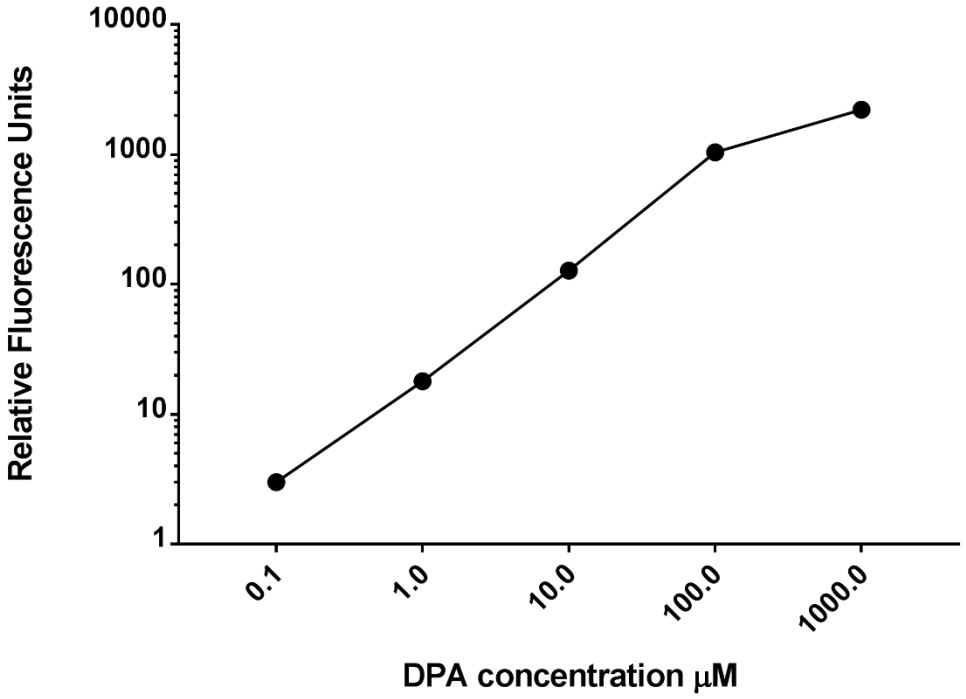


Figure 03-22 Standard curve of Terbium fluorescence with defined DPA concentrations.

Terbium fluorescence measured from the concentrations of dipicolinic acid shown. Symbols indicate the relative fluorescence units (RFU) after background fluorescence from aliquots of each DPA concentration not supplemented with terbium was deducted from the respective value obtained with terbium.

C. difficile strain	DPA concentration (nM per $1x10^8$ spores)	DPA content (pg.spore^{-1})
630Δ*erm*	517	8.64
CRG5234	72	1.21
CRG5235	976	16.31
PSI7475 (Δ*spoVAC*)	72	1.21
PSI7375 (Δ*spoVAD*)	82	1.37
PSI7374 (Δ*spoVAE*)	82	1.37

Table 03-3 Calculations of DPA content per spore.

The mean concentration of DPA per $1x10^8$ spores of each strain was determined using the respective terbium fluorescence data from Figure 3.21 and extrapolating the corresponding DPA concentration from the standard curve in Figure 3.22. These values were then used to calculate the approximate amount of DPA per spore.

3.9 Generation of single s*poVA* deletion mutants via complementation

Having now identified a method of spore purification which prevents clumping and enables accurate enumeration, renewed efforts were made to generate strains lacking individual *spoVA* genes. Whilst individual in-frame deletions could not be generated for each of the 630Δ*erm spoVA* genes via allelic exchange mutagenesis, a strain with the entire operon deleted was produced. The successful complementation of this strain with the native *spoVA* operon at the *pyrE* locus prompted an alternative approach; to complement the 630Δ*erm*Δ*pyrE*Δ*spoVA* mutant with the native promoter and combinations of two of the three *spoVA* genes, thus generating strains lacking each of the individual *spoVA* genes.

3.9.1 Construction of dual *spoVA* complementation vectors

Conventional cloning was used to create a vector capable of complementing 630Δ*erm*Δ*pyrE*Δ*spoVA* with the *spoVAC* and *spoVAD* genes. A 1.851 kb DNA fragment consisting of the native *spoVA* promoter upstream of *spoVAC* through to the 13 bp 5'-UTR upstream of *spoVAE*, with flanking 5'-*Not*I and 3'-*Xho*I restriction sites, was PCR amplified from 630Δ*erm* genomic DNA using primers VACDE_1C_F1 and VACD_1C_R1. This fragment was cloned into pMTL-YN1C via restriction digests with *Not*I and *Xho*I enzymes followed by appropriate ligation reactions, and transformed into *E. coli* TOP10 cells. Plasmids were extracted from the resulting chloramphenicol-resistant colonies and Sanger sequenced with primers m13F and VACDE_1C_F1 to confirm correct PCR amplification of the complementation cassette, generating the plasmid pMTL-YN1C_*spoVAC*D (Figure 3.23).

To generate complementation vectors for the production of strains lacking either *spoVAC* or *spoVAD*, a HiFi assembly protocol was followed, which allows multiple DNA fragments to be joined in a single reaction. To generate a complemented strain lacking only *spoVAD*, an 883 bp DNA fragment containing the native *spoVA* promoter, *spoVAC* gene and 166 bp intergenic region between the *spoVAC* stop codon and *spoVAD* start codon was PCR amplified from 630Δ*erm* genomic DNA using primers YN1C_pVAC_F1 and VAE_pVAC_R1, containing 5' overlapping homology with pMTL-YN1C and *spoVAE*, respectively. Similarly, a 376 bp DNA fragment containing *spoVAE* and the 13 bp intergenic sequence immediately following the *spoVAD* stop codon was PCR amplified using primers pVAC_VAE_F1 and YN1C_VAE_R1, containing 5' overlapping homology to the 3' end of the *spoVAC*-containing fragment, and to pMTL-YN1C, respectively. Both DNA fragments were purified and combined in a 2:1 molar ratio with a purified pMTL-YN1C fragment which had been linearised via restriction digestion with *Not*I and *Xho*I enzymes. HiFi assembly master mix (NEB) was added to this mixture and incubated at 50^{o}C for 1 hour, during which the activity of the supplied exonuclease, polymerase and ligase enzymes assembled vector pMTL-YN1C_*spoVACE*. Successful assembly of pMTL-YN1C_*spoVACE* (Figure 3.24) was confirmed, via restriction digestion and Sanger sequencing with primers m13F and VACDE_1C_F1, with plasmids extracted from chloramphenicol-resistant *E. coli* TOP10 colonies harbouring the HiFi assembly products.

Similarly, for the complementation of 630Δ*erm*Δ*pyrE*Δ*spoVA* with *spoVAD* and *spoVAE,* a 279 bp promoter region was PCR amplified from 630Δ*erm* genomic DNA using primers YN1C_pro_F1 (with 5' overlapping homology to YN1C) and VADE_pro_R1 (with 5' overlapping homology to the VADE fragment). A second fragment, 1.556 kb in length and containing the *spoVAD* and *spoVAE* genes plus the 166 bp immediately following the *spoVAC* stop codon, was PCR amplified using primers pro_VADE_F1 (with 5' overlapping homology to

the *SpoVA* promoter fragment) and VADE_YN1C_R1 (with 5' overlapping homology to pMTL-YN1C). The HiFi assembly reaction was performed using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB) according the manufacturer's instructions, using both of these purified PCR products in a 2:1 molar ratio with a purified pMTL-YN1C DNA fragment linearized via restriction digestion with *Not*I and *Xho*I enzymes. Subsequent assembly products were transformed into *E. coli* TOP10 and from the resulting chloramphenicol colonies, plasmids were extracted and successful assembly of pMTL-YN1C_*spoVAD*E confirmed via restriction digestion and Sanger sequencing with primers m13F and VACDE_1C_F1.

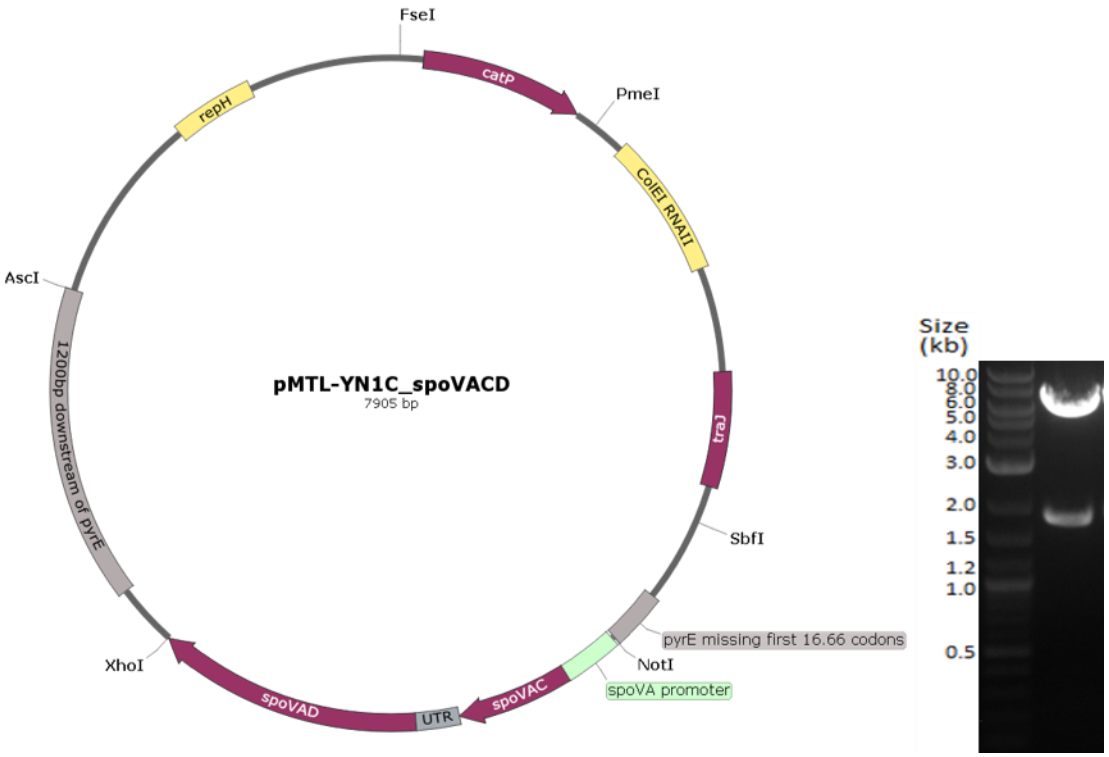


Figure 3-23 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVACD*.

Left: Plasmid map detailing features of the complementation vector, pMTL-YN1C_*spoVACD*, for generation of a *C. difficile* 630Δ*erm* strain lacking only *spoVAE*. Vector backbone identical to pMTL-YN1C. Complementation cassette cloned between *Not*I and *Xho*I restriction sites consists of the *spoVA* promoter, *spoVAC*, untranslated region downstream of *spoVAC*, and *spoVAD*. Right: Agarose gel electrophoresis following restriction digestion with *Not*I and *Xho*I of pMTL-YN1C-*spoVACD*.

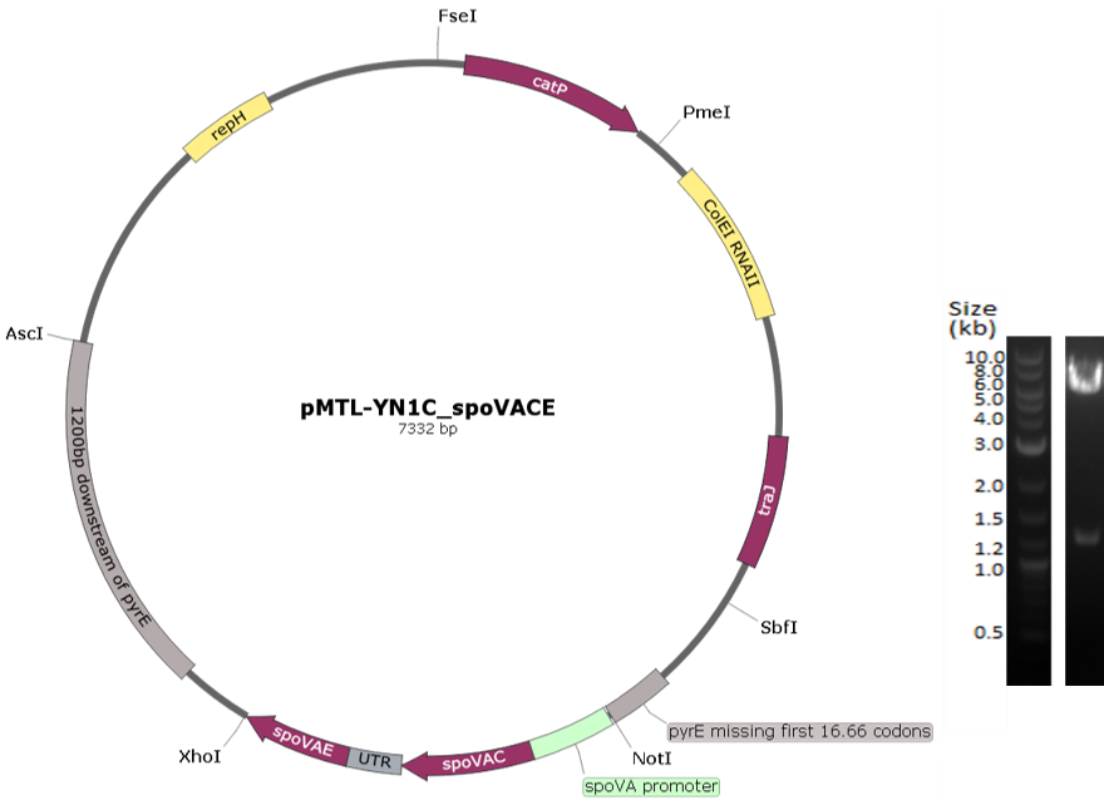


Figure 3-14 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVACE*.

Left: Plasmid map detailing features of the HiFi-assembled complementation vector, pMTL-YN1C_*spoVACE*, for generation of a *C. difficile* 630Δ*erm* strain lacking only *spoVAD*. Vector backbone identical to pMTL-YN1C. Complementation cassette cloned between *Not*I and *Xho*I restriction sites consists of the *spoVA* promoter, *spoVAC*, untranslated region (UTR) downstream of *spoVAC*, UTR downstream of *spoVAD*, and *spoVAE*. Right: Agarose gel electrophoresis following restriction digestion with *Not*I and *Xho*I of pMTL-YN1C-*spoVACE*.

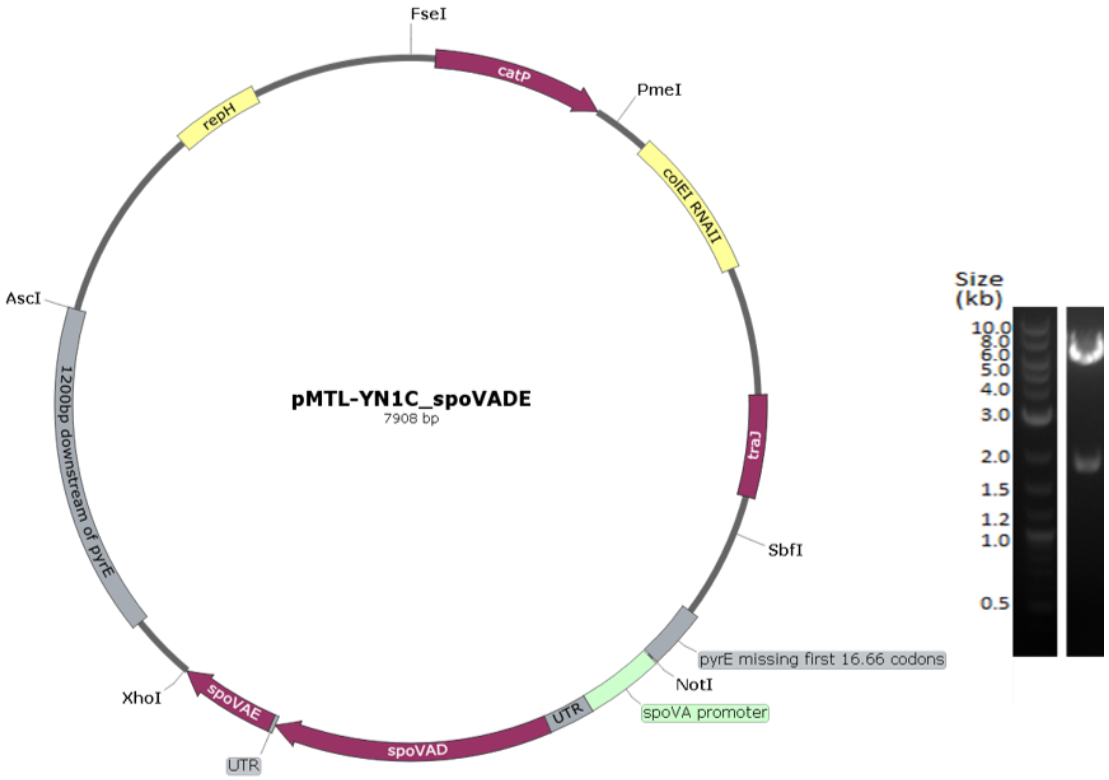


Figure 3-25 Diagrammatic representation of the dual complementation vector, pMTL-YN1C_*spoVADE*.

Left: Plasmid map detailing features of the HiFi-assembled complementation vector, pMTL-YN1C_*spoVADE*, for generation of a *C. difficile* 630Δ*erm* strain lacking only *spoVAC*. Vector backbone identical to pMTL-YN1C. Complementation cassette cloned between *Not*I and *Xho*I restriction sites consists of the *spoVA* promoter, untranslated region (UTR) downstream of *spoVAC*, UTR downstream of *spoVAD*, and *spoVAE*. Right: Agarose gel electrophoresis following restriction digestion with *Not*I and *Xho*I of pMTL-YN1C-*spoVADE*.

3.9.2 Generation of strains lacking single *spoVA* genes

The assembled, Sanger sequence confirmed dual complementation vectors pMTL-YN1C_*spoVACD*, pMTL-YN1C_*spoVACE* and pMTL-YN1C_*spoVADE* were separately transformed into *E. coli* CA434 and the resulting chloramphenicol-resistant strains used as conjugation donors for individual transfer of these vectors into 630Δ*erm*Δ*pyrE*Δ*spoVA*. The complementation protocol outlined in 2.10.4.2 was followed for subsequent thiamphenicol-

resistant colonies harbouring each dual complementation vector. Consequently, thiamphenicol-sensitive, putative dual-complemented colonies able to grow on CDMM without uracil supplementation were obtained; one for complementation with *spoVAC* and *spoVAE*, four for complementation with *spoVAC* and *spoVAD*, and one for complementation with *spoVAD* and *spoVAE*. Genomic DNA was extracted from each of these colonies and used as template for PCR screening with primers CD630_pyrD_sF1 and CD630_0189_sR3 to confirm the desired complementation and *pyrE* repair had occurred (Figure 3.26A). The expected PCR product sizes for strains individually lacking *spoVAC*, *spoVAD* and *spoVAE* were 3.897 kb, 3.315 kb and 3.893 kb, respectively. With the exception of the band in lane 5 of Figure 3.26A, which appeared slightly larger than anticipated, all PCR products were of the expected size. The strain which produced the larger than expected PCR product was not taken forward for further analysis.

To confirm the apparent successful complementation, one PCR product from each newly generated strain was Sanger sequenced using primers m13F, 1C_pro_F1 and an appropriate internal primer. Alignments of the generated DNA sequences with the expected sequence for each strain confirmed the successful generation of 630Δ*erm*Δ*spoVA*::*spoVAC_spoVAD* (PSI7374) and 630Δ*erm*Δ*spoVA*::*spoVAD_spoVAE* (PSI7475). However, alignment of the sequence generated from the putative 630Δ*erm*Δ*SpoVA*::*SpoVAC_SpoVAE* (PSI7375) with the expected sequence identified a 13 bp deletion in the untranslated region downstream of *spoVAC*. To remedy this, a further six thiamphenicol-sensitive colonies able to grow on minimal media without supplemented uracil were generated, from which genomic DNA was extracted and used in PCR screens with primers CD630_pyrD_sF1 and CD630_0189_sR3 (Figure 3.26B). All six PCR products were of the expected size, 3.315 kb. Furthermore, of the three PCR products subsequently Sanger sequenced, all aligned exactly to the expected

sequence, thus confirming the generation of strain 630Δ*erm*Δ*SpoVA*::*SpoVAC_SpoVAE* (PSI7375).

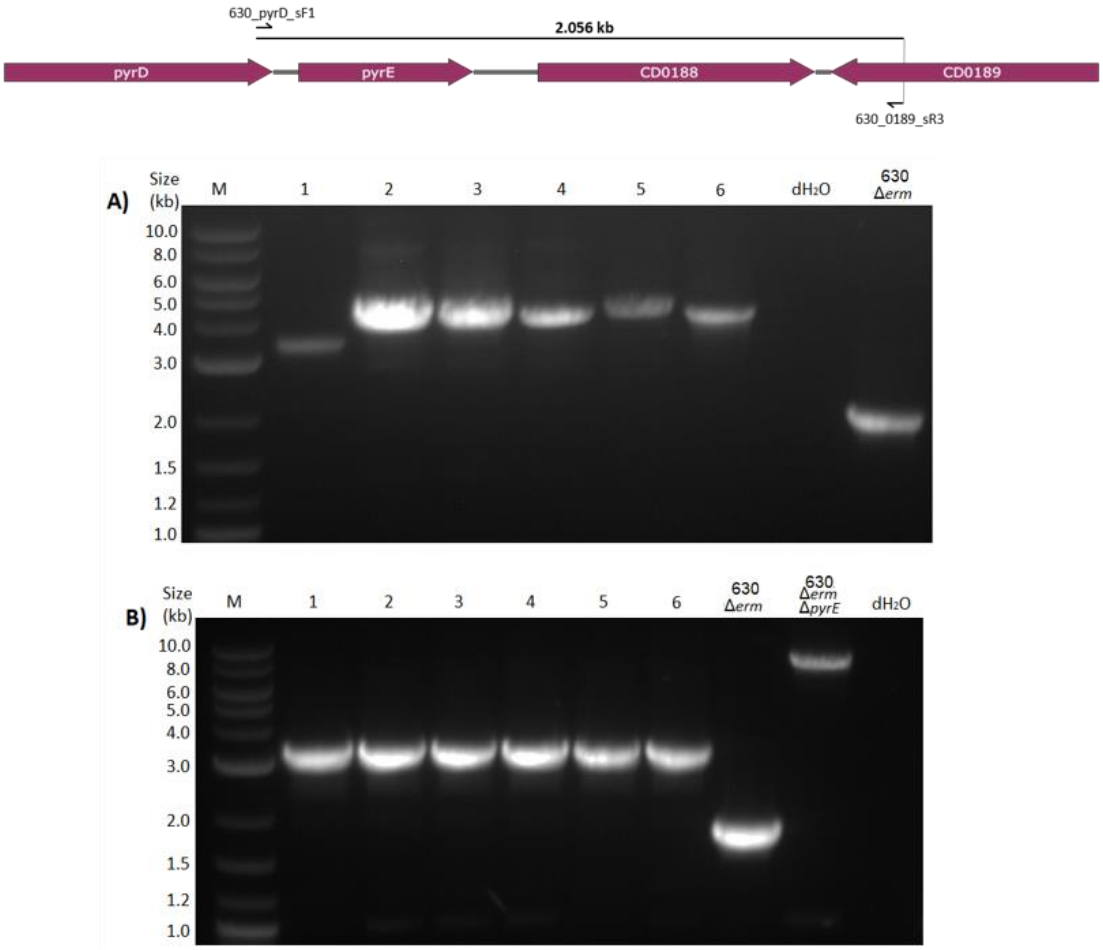


Figure 3-26 PCR screening for dual complementation and *pyrE* repair of 630Δ*erm*Δ*pyrE*Δ*spoVA*.

Thiamphenicol-sensitive colonies PCR screened using primer pair CD630_pyrD_sF1/CD630_0189_sR3 following allele-coupled exchange with each of the *spoVA* dual complementation vectors. (A) Numbered lanes correspond to gDNA isolated from individual thiamphenicol-sensitive colonies complemented with 1) *spoVAC* and *spoVAE*; 2)-5) *spoVAC* and *spoVAD*; and 6) *spoVAD* and *spoVAE*. (B) Numbered lanes correspond to a further six individual thiamphenicol-sensitive colonies complemented with *spoVAC* and *spoVAE*. M, DNA Marker (2-log ladder; NEB); 630Δ*erm***,** 630Δ*erm* gDNA; 630Δ*erm*Δ*pyrE*, 630Δ*erm*Δ*pyrE* gDNA; dH_2O, distilled water.

3.10 Characterisation of strains lacking single *spoVA* genes

3.10.1 Growth in BHIS broth over 24 hours

The growth of each *spoVA* mutant strain along with parental 630Δ*erm* and fully complemented control strains was assessed in BHIS broth using changes in measured OD_{600} over 24 hours (Figure 3.27). Over the 24-hour assay period the growth of all strains was indistinguishable from that of the parental 630Δ*erm* strain, indicating that deletion of all or individual *spoVA* genes does not affect the growth characteristics of these strains. Furthermore, any differences between these strains cannot be attributed to differences in the growth rates of these strains.

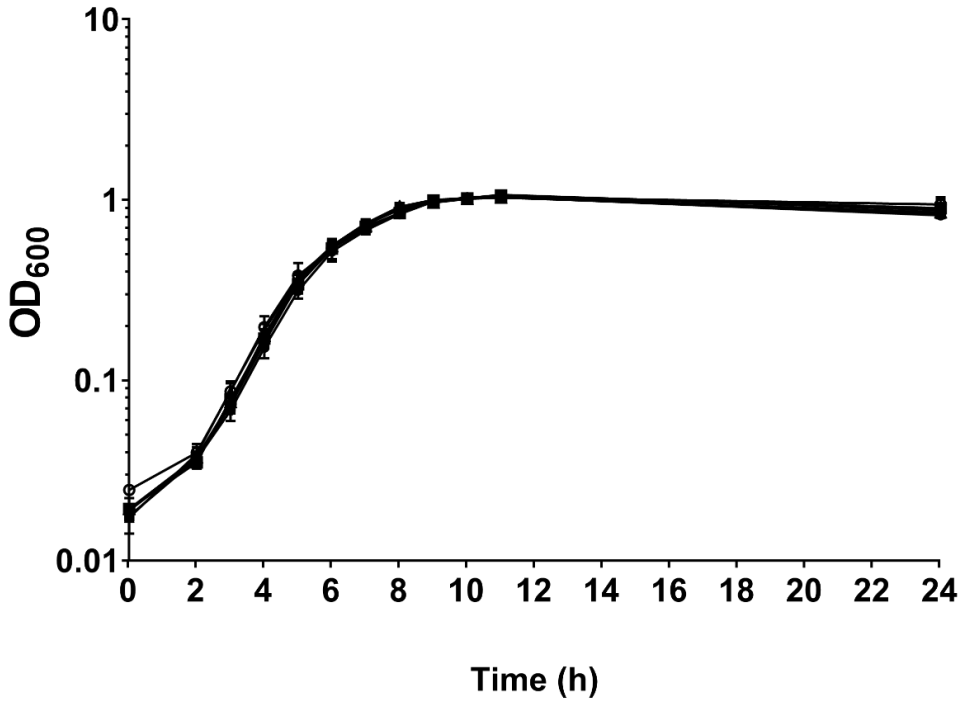


Figure 03-27 Growth of *C. difficile* strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}.

● *C. difficile* 630Δ*erm*; ■, CRG5234; ▲, CRG5235; ○, PSI7475; □, PSI7374; △, PSI7375. Symbols represent the averages of three independent experiments. Error bars indicate the standard errors of the mean. Statistical significance determined using two-way ANOVA with Sidak's multiple comparisons test; *P*-value=0.4287 for variance between strains.

3.10.2 Colony formation after heat-treatment

The ability of each of these strains to develop heat-resistant CFU over five days in BHIS agar supplemented with taurocholate was determined (Figure 3.28). Mean heat-resistant CFU observed after 120 hours were $7.56x10^6$ for 630Δ*erm*, $8.22x10^3$ for CRG5234, $6.22x10^6$ for CRG5235, $1.39x10^4$ for PSI7475, $4.89x10^3$ for PSI7375 and $4.56x10^3$ for PSI7374. Similar to previous findings (Figure 3.13), the development of heat-resistant CFU for 630Δ*erm* and the fully complemented *spoVA* deletion strain (CRG5235) are almost identical at each timepoint measured. After 120 hours, the whole *spoVA* operon deletion mutant (CRG5234) and each individual *spoVA*- strain displayed similar levels of heat-resistant CFU, showing the complementation of two out of the three *spoVA* genes could not recover these spores to the levels of heat resistance observed for 630Δ*erm*. Interestingly, heat-resistant CFU were detected from CRG5234 and strains lacking *spoVAD* and *spoVAE* after 48 hours, while the strain lacking *spoVAC* developed detectable heat-resistant CFU after 72 hours. These values appear to contradict the results of the previous assay (Figure 3.13), in which heat-resistant CFU from strain CRG5234 appeared only after 96 hours. Furthermore, after 120 hours in the previous assay of colony formation after heat-treatment, the mean number of heat-resistant CFU for CRG5234 was 94.44, yet the equivalent value in this assay was $8.22x10^3$, almost 100-fold higher. Nonetheless, these levels of heat-resistant CFU observed for CRG5234 after 120 hours in this assay were still 920-fold lower than equivalent values for the parental 630Δ*erm* strain.

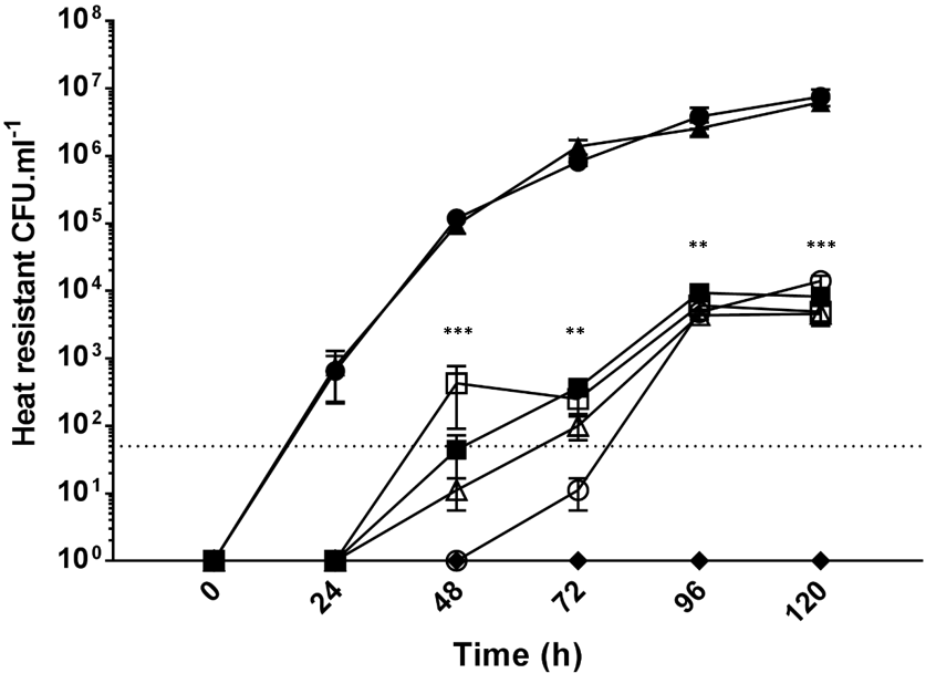


Figure 03-28 Development of heat resistant CFU of *C. difficile* 630 strains over five days.

●, *C. difficile* 630Δ*erm*; ■, CRG5234; ▲, CRG5235; O, PSI7475; □, PSI7375; △, PSI7374; ◆, 630Δ*erm*Δ*spo0A*. Symbols represent the mean values from three independent experiments. Error bars indicate standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (**) denotes a *P*-value **<**0.01 and (***) denotes a *P*-value **<**0.001.

3.10.3 Resistance of *C. difficile* spores to wet-heat

Spores of the *spoVA* whole operon deletion mutant and the fully and dually complemented strains were heated at various temperatures to determine the resistance of these spores to wet heat (Figure 3.29). Variation between strains was observed in the CFU obtained from non-heat-treated spore preparations, therefore the effects of wet heat on individual strains were determined from comparisons to these non-heat-treated values for each strain. All strains tolerated exposure to 50°C wet-heat well, with only minor changes from the non-heat-treated samples. Following thirty minutes exposure to wet-heat at 65°C, mean CFU.ml^{-1}

for each strain expressed as a percentage of the corresponding non-heat-treated value was 49.30% for 630Δ*erm*, 3.69% for CRG5234, 89.81% for CRG5235, 6.19% for PSI7475, 4.84% for PSI7375 and 3.74% for PSI7374. No spores lacking either all, or just one, *spoVA* gene(s) formed any detectable colonies after exposure to 80°C wet-heat, whilst 630Δ*erm* and CRG5235 produced 252.8 (0.005%) and 263.9 (0.006%) mean CFU.ml^{-1} respectively. These results confirm that spores lacking either one or all *spoVA* gene(s) are less resistant to wet-heat at 65°C, and are completely susceptible to wet-heat at 80°C, compared to parental and fully complemented spores.

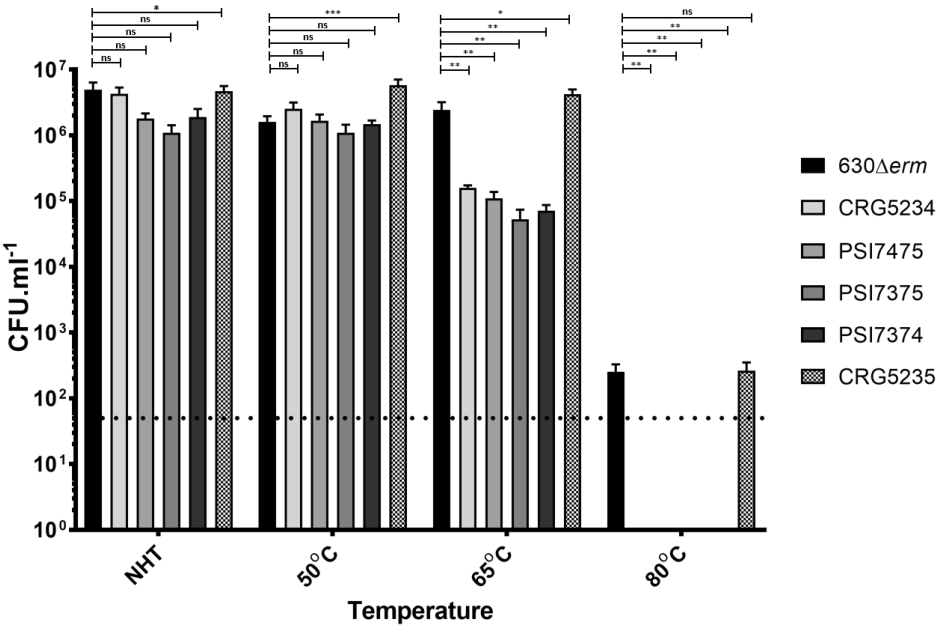


Figure 03-29 Colony development after exposure of *C. difficile* spores to wet heat.

1x10^7 spores of each 630Δ*erm* strain were incubated on the bench (not heat treated; NHT) or heated at the temperatures indicated for 30 minutes, then serially diluted and spotted onto BHIS agar plates supplemented with taurocholate. CFU.ml^{-1} values were determined after 24 hours anaerobic incubation at 37°C. Bars represent mean values from three independent experiments performed in duplicate (n=6). Error bars indicate the standard errors of the means. Line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, one asterisk (*) denotes a *P*-value <0.05, two asterisks (**) denotes a *P*-value <0.01, while three asterisks (***) denotes a *P*-value <0.001.

The level of reduction in observed CFU.ml^{-1} following heat-treatment of spore preparations at 65^{o}C for spores lacking one or all *spoVA* gene(s) ranged from 16- to 27-fold reductions. However, this level of reduction is far smaller than that observed in the rate of sporulation assays in section 3.6.1 for *spoVA*- strains compared to the parental and fully complemented strains. This would appear to suggest that changes in heat-resistance properties were not solely responsible for the reduction in CFU.ml^{-1} following heat treatment. Several differences in the method of heat-treatment exist between these two studies, including the medium in which cells/spores are heat-treated. In the rate of sporulation assay, cells and spores were heat-treated in the growth medium, BHIS broth, whilst in this heat resistance assay pure spores were heated in the washing buffer, PBS with 0.01% Tween80. It was hypothesised that a component of BHIS broth may accentuate the susceptibility of *spoVA*- spores to wet-heat and/or the PBS and Tween80 medium may confer some protection to these spores against heat treatments. To investigate this, pure spore preparations in PBS with Tween80 were washed twice in distilled water and then resuspended in either distilled water, PBS, PBS with Tween80 or BHIS broth and incubated at 65^{o}C or on the bench for 30 minutes (Figure 3.30). The resulting heat-resistant CFUs for each strain in each medium were then expressed as a percentage of the corresponding non-heat-treated values. For CRG5234, the percentage of heat-treated to non-heat-treated CFUs observed in each buffer was 2.58% for distilled water, 0.95% for PBS, 2.08% for PBS with Tween80 and 0.67% for BHIS. Furthermore, similar analysis with 630Δ*erm* and CRG5235 revealed that both strains produced their highest percentages of heat-treated to non-heat-treated CFUs in PBS with Tween80, and their lowest such values in BHIS broth. Thus, these results appear to show that the medium in which heat-treatment occurs does affect the subsequent proportion of heat-resistant CFUs observed, with spore resistance to wet-heat increasing in PBS with Tween80 and decreasing in BHIS broth.

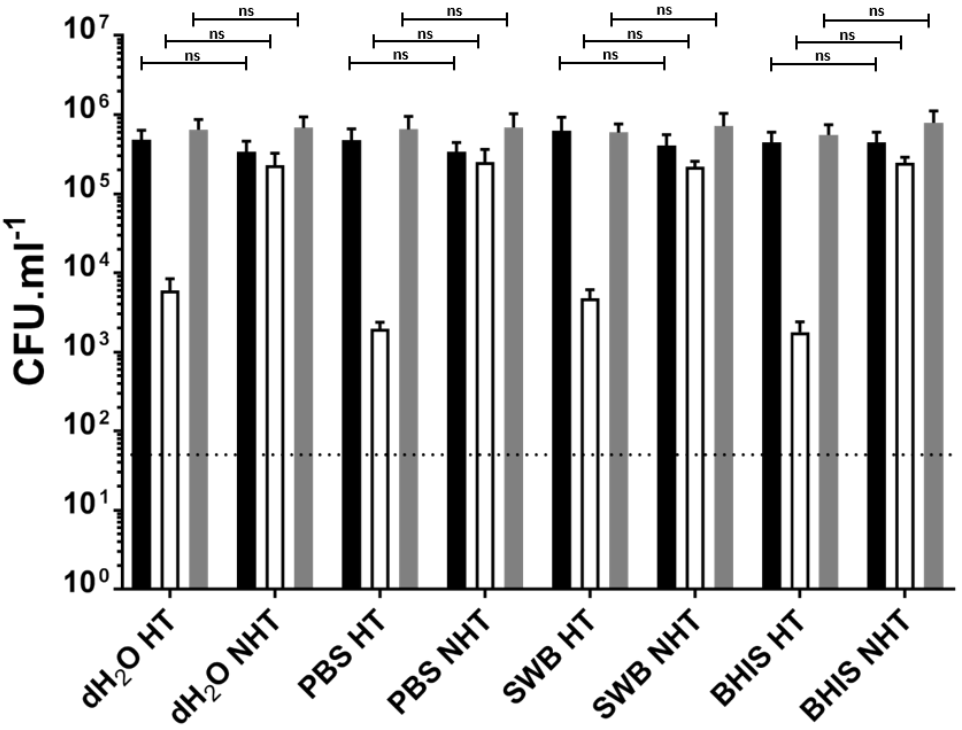


Figure 03-30 Heat-resistant CFU development in various spore suspension media.

1x10^7 spores of each 630Δ*erm* strain were resuspended in distilled water (dH$_2$O), PBS, PBS with Tween80 (spore washing buffer; SWB) or BHIS broth and incubated on the bench (not heat treated; NHT) or heated (HT) at 65°C for 30 minutes, then serially diluted and spotted onto BHIS agar plates supplemented with taurocholate. Black bars, 630Δ*erm*; white bars, CRG5234; grey bars, CRG5235. CFU.ml^{-1} values were determined after 24 hours anaerobic incubation at 37°C. Bars represent mean values from three independent experiments. Error bars indicate the standard errors of the means. Line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using two-way ANOVA with Sidak's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05.

3.10.4 Resistance of *C. difficile* spores to UV radiation

Having shown *C. difficile spoVA*$^-$ spores lacking DPA are more susceptible to wet-heat exposure than parental spores, the resistance properties of these spores to ultraviolet radiation was then investigated. To determine an appropriate range of UV radiation to expose spores to, 1x10^7 spores.ml^{-1} of the parental strain 630Δ*erm* were exposed to varying

amounts of UV radiation, then serially diluted and plated onto BHIS agar supplemented with taurocholate (Figure 3.31). Mean UV-resistant CFU.ml^{-1} values decreased from the non-treated sample, 3.21x10^{6}, to 1.07x10^{6} after exposure to 100,000 uJ.cm^{-2}, 1.92x10^{4} after exposure to 200,000 uJ.cm^{-2}, and to 5.22x10^{2} after exposure to 300,000 uJ.cm^{-2} UV energy. This observed 2-log decrease in CFU.ml^{-1} from the non-treated 630Δ*erm* sample to the equivalent value for spores exposed to 200,000 uJ.cm^{-2} UV energy was deemed an appropriate range of UV energy with which to expose all spore preparations to.

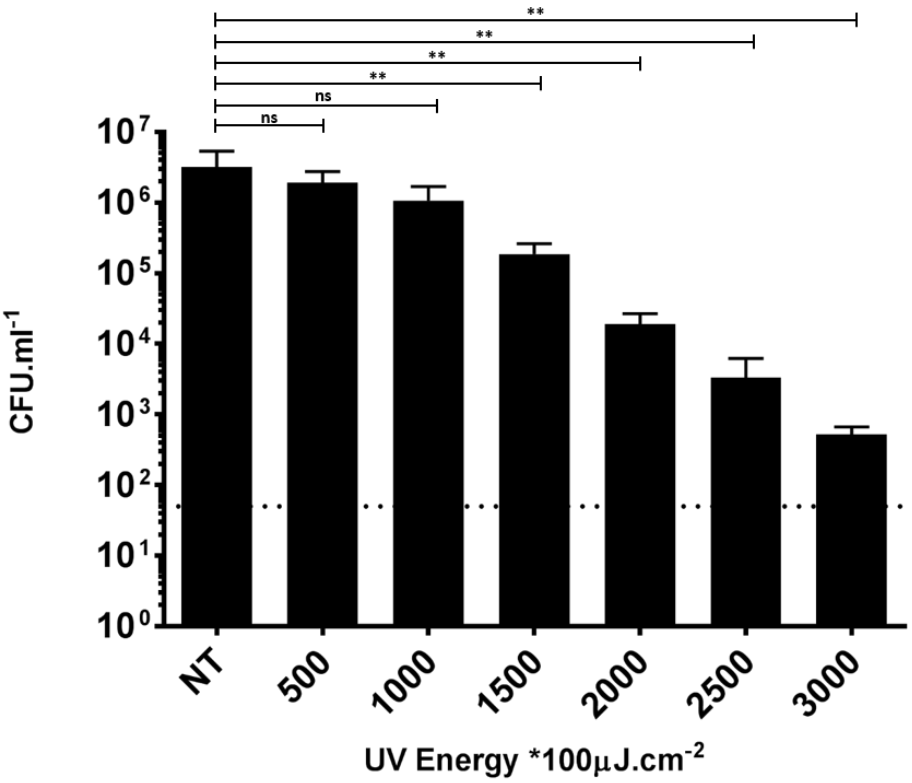


Figure 03-31 Resistance of 630Δ*erm* spores to various amounts of UV radiation.

Colony forming units observed following the exposure of aliquots of 1x10^{7} *C. difficile* 630Δ*erm* spores to the amounts of UV energy indicated or incubated on the bench (not treated; NT). Bars indicate the mean values from three independent experiments. Error bars indicate the standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, two asterisks (**) denotes a *P*-value <0.01.

Having determined an appropriate range of UV Energy with which to cause a 2-log reduction in the number of CFU.ml^{-1} of 630Δ*erm* spores on BHIS agar supplemented with taurocholate, this assay was repeated by exposing spores of all strains to up to 200,000 uJ.cm^{-2} UV energy (Figure 3.32). Even at the lowest amount of UV energy tested in this assay, 50,000 uJ.cm^{-2}, all spores lacking either one or all *spoVA* gene(s) were severely affected when compared to the equivalent non-treated values. For example, untreated CRG5234 spores produced 2.62x10^{6} CFU.ml^{-1} which decreased to 3.94x10^{3} CFU.ml^{-1} following exposure to 50,000 uJ.cm^{-2} UV energy. *spoVA*- spores were affected by UV exposure in a dose-dependent manner with CRG5234 spores producing only 294.44 CFU.ml^{-1} after treatment with 100,000 uJ.cm^{-2} UV energy and the strains lacking individual *spoVA* genes similarly affected. Comparatively, after the same UV exposure the parental 630Δ*erm* and fully complemented CRG5235 spores produced 8.83x10^{5} and 9.61x10^{5} CFU.ml^{-1}, respectively. Furthermore, no CFU.ml^{-1} were observed from any spores lacking either one or all *spoVA* gene(s) following exposure to 150,000 uJ.cm^{-2} UV energy or higher. Meanwhile, at the highest amount of UV energy tested, 200,000 uJ.cm^{-2}, spores of 630Δ*erm* and CRG5235 produced 2.06x10^{4} and 1.16x10^{4} CFU.ml^{-1}, respectively. As with previous assays, the fully complemented mutant CRG5235 behaved in an almost identical manner to the parental 630Δ*erm* strain.

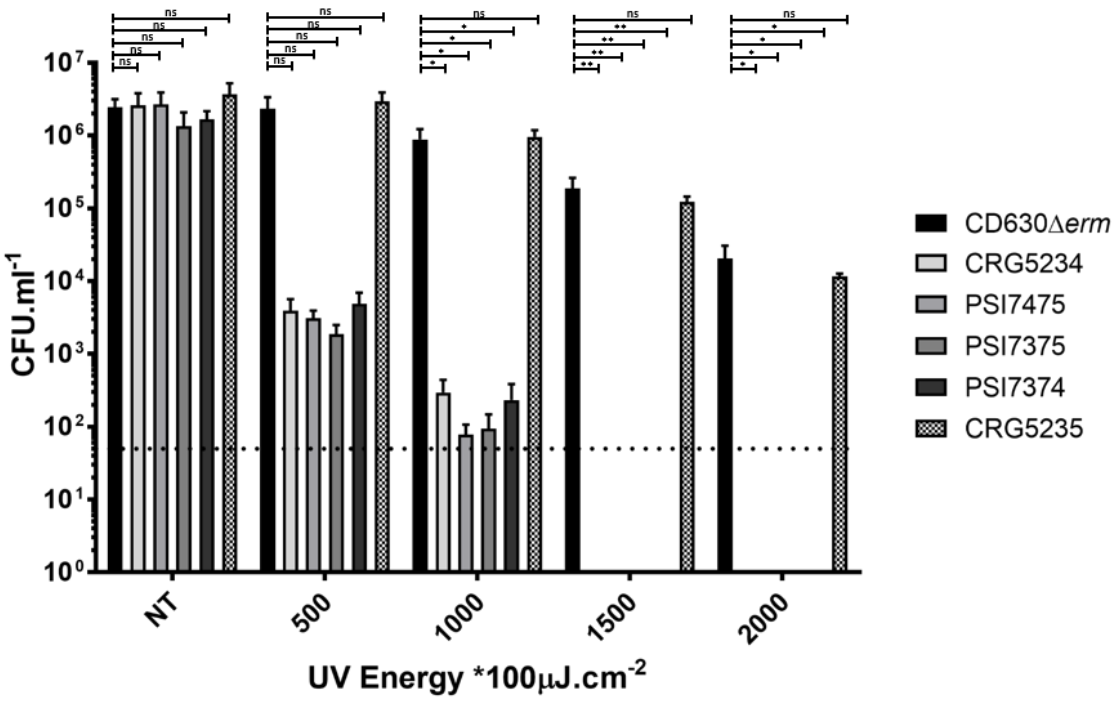


Figure 03-32 Resistance of 630Δ*erm spoVA* spores to UV radiation.

Colony forming units observed following the exposure of 1x10^7 *C. difficile* spores to the amounts of UV energy indicated or incubated on the bench (not treated; NT). Bars indicate the mean values from three independent experiments. Error bars indicate the standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, one asterisk (*) denotes a *P*-value <0.05, while two asterisks (**) denotes a *P*-value <0.01.

3.10.5 Resistance of *C. difficile* spores to ethanol

Having identified the susceptibility of *spoVA*- spores to exposure from wet-heat and UV radiation, the resistance of these spores to ethanol was then investigated. Previous studies determining the effect of ethanol on *B. subtilis* spore viability incubated spores with 70% ethanol at 65°C (Setlow *et al*., 2002). In this study, spores were incubated at 37°C to ensure only the effects of ethanol treatment were investigated. Aliquots containing 1x10^7 spores of each strain were resuspended in 70% or 100% ethanol or PBS, incubated at 37°C for one

hour, then serially diluted and spotted onto BHIS agar supplemented with taurocholate, after which CFU.ml^{-1} values were determined (Figure 3.33). Whilst ethanol treatment affected all strains to some extent, again it was the spores lacking either one or all *spoVA* gene(s) which were most severely affected. Suspension in 70% ethanol resulted in CFU.ml^{-1} values, expressed as a percentage of the equivalent values for suspension in PBS, were 56.4% for 630Δ*erm*, 8.67% for CRG5234 and 26.8% for CRG5235. Similarly, after suspension in 100% ethanol CFU.ml^{-1} values relative to the PBS control for 630Δ*erm*, CRG5234 and CRG5235 were 10.2%, 0.43% and 7.47%, respectively. Hence, it appears that *C. difficile* spores lacking DPA are also more susceptible to ethanol treatment, with spore killing increasing proportionally with increasing ethanol concentration.

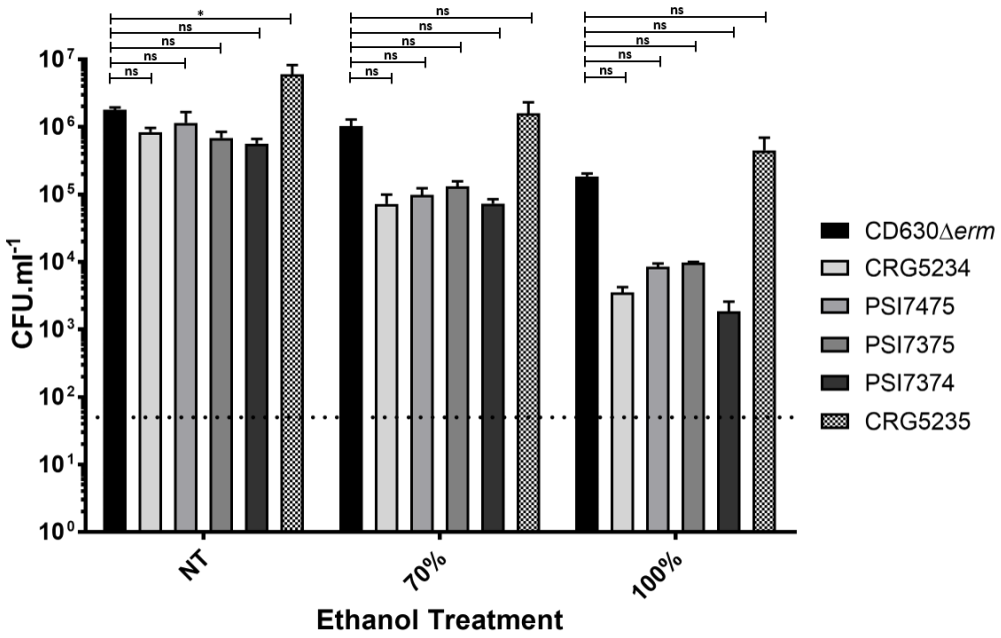


Figure 03-33 Resistance of 630Δ*erm spoVA* spores to ethanol treatment.

Colony forming units observed following the resuspension of 1x10^7 *C. difficile* spores in either 70% or 100% ethanol, or PBS (not treated; NT), and incubated for one hour at 37^{o}C. Bars indicate the mean values from three independent experiments. Error bars indicate the standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, while one asterisk (*) denotes a *P*-value <0.05.

3.11 Discussion

This study utilised ClosTron and allelic exchange mutagenesis methods to generate *C. difficile* 630Δ*erm* strains lacking either individual, two or all three *spoVA* genes found in *C. difficile*. Unlike *spoVA*- spores of *B. subtilis* which lyse during the sporulation process, *C. difficile* 630Δ*erm*Δ*spoVA* strains produced spores in similar numbers to the parental 630Δ*erm* strain and which could complete germination and outgrowth into vegetative cells in the presence of glycine and taurocholate. However, as with *spoVA*- spores of *C. perfringens*, these spores had lost the characteristic resistance properties associated with bacterial endospores. Complementation of the 630Δ*erm*Δ*pyrE*Δ*spoVA* strain lacking all three *spoVA* genes was performed *in situ*, by inserting either all three, or combinations of two, *spoVA* genes under the control of the native *spoVA* promoter at the *pyrE* locus, concomitant with *pyrE* repair. Of the resulting strains, only the fully complemented mutant, CRG5235, was restored to wildtype levels of DPA in the spore core, and subsequently similar levels of resistance to wet-heat, UV and ethanol treatments as 630Δ*erm* spores. Throughout this study, spores from all three of the dually complemented strains had similar resistance profiles and contained similar levels of DPA to spores of CRG5234. This suggests that all three *spoVA* gene products are required for successful import of DPA into the spore core during sporulation. Future work could add more evidence to this finding by using RT-PCR to confirm the dual complemented *spoVA* strains produced in this study exhibit similar transcription levels as the wildtype.

This study identified two putative promoters in intergenic regions within the hypothetical *spoVA* operon; one upstream of *spoVAC* and another upstream of *spoVAD*, in the 166 bp intergenic region between *spoVAC* and *spoVAD*. For this reason, care was taken when designing the dual complementation cassettes to ensure that only the gene of interest was

lacking in each single *spoVA* deletion strain generated. This increased intergenic region between *spoVAC* and *spoVAD*, complete with predicted promoter sequence, is not found in *spoVA* operons of *B. subtilis* or *C. perfringens* and may be unique to *C. difficile*. Future investigations should perform a transcriptional linkage assay (Cooksley *et al*., 2010) to determine whether the *C. difficile spoVA* genes are expressed as a single transcript or as two distinct transcripts. Only after the results of such an assay can one begin to ponder the advantages of having additional regulation of the *spoVA* genes in *C. difficile*.

The roles of the SpoVA proteins in *C. difficile*, and other Firmicutes, have been the focus of several studies published over the duration of this work. A study by Francis *et al*. (2015) analysed DPA release in *C. difficile* spores inactivated via TargeTron insertions in either the bile acid receptor, *cspC*, or cortex hydrolase, *sleC*, and found that cortex hydrolysis and DPA release from the spore core was prevented in each mutant. This was a significant finding as it highlighted further differences in the germination processes between *C. difficile* and *B. subtilis*, in which DPA release from the spore core activates CwlJ-mediated hydrolysis of the cortex. Hence, in *B. subtilis* DPA release precedes cortex hydrolysis, whilst the opposite is true in *C. difficile*.

Other studies attempted to determine the roles of individual SpoVA proteins, in particular SpoVAC. Previous work in *B. subtilis* had already determined that SpoVAD localises to the spore inner membrane (Vepachedu & Setlow, 2005) and through analysis of its crystal structure and subsequent genetic studies identified a binding site for DPA within SpoVAD, which had similar affinity for Ca-DPA (Li *et al*., 2012). Firstly, Velásquez, *et al*. (2014) found that expressing the *B. subtilis* SpoVAC in *E. coli* conferred similar levels of protection of these cells against osmotic downshift, which occurs with increased membrane tension due to

excessive water intake, as did the mechanosensitive protein, MscL. Mechanosensitive channel proteins prevent membrane damage in such instances by releasing osmolytes upon activation of these proteins (Kung *et al*., 2010). Velásquez *et al*. (2014) were also able to show that SpoVAC forms transient pores with a symmetrical cylindrical shape and an estimated diameter of 4.6 Å. Building upon this hypothesis that SpoVAC is a mechanosensitive channel protein, Francis & Sorg (2016) assessed the ability of *C. difficile* R20291 spores containing wildtype or truncated forms of SpoVAC to germinate in the presence of high osmolyte concentrations by monitoring subsequent DPA release over time. When the concentration of osmolytes was high, cortex degradation was unaffected but release of DPA was delayed and could be triggered by sufficient dilution of the germination buffer. These results appear to confirm the role of SpoVAC as a mechanosensitive protein which the authors suggest becomes activated following local changes in the spore inner membrane resulting from cortex degradation, leading to release of DPA via SpoVAC.

Further characterisation of the role of SpoVAC in *C. difficile* spores came from a study by Donnelly *et al.* (2016) in which TargeTron insertional mutagenesis was used to generate two strains of 630E inactivated in either *spoVAC* or *dpaAB* (DPA synthetase). Analysis of the subsequent spores lacking functional *spoVAC* and *dpaAB* found that spores of both strains; i) were less dense than wildtype or plasmid complemented strains, ii) contained significantly lower levels of DPA, determined using the terbium fluorescence assay with boiled spore preparations, iii) had expanded core regions, determined via measurements of spore TEM images, and iv) exhibited a significant reduction in resistance to wet heat treatments relative to wildtype and complemented spores. They also found that supplementing both strains with exogenous DPA could restore the phenotype of *dpaAB*-minus spores, but not those lacking *spoVAC*.

Comparing the recent literature with the work presented in this study, all studies agree that varying *C. difficile* strains lacking *spoVAC* retain the ability to form phase-bright spores which complete sporulation and are capable of returning to vegetative cell growth via germination and outgrowth. Furthermore, all studies agree that *C. difficile* strains lacking functional SpoVAC proteins are unable to accumulate DPA in the resulting spores. This study, through the creation of strains lacking individual *spoVA* genes, has additionally shown that strains lacking any of the three *spoVA* genes produce spores which cannot accumulate DPA and are as susceptible to wet heat, ethanol and UV treatments. The study by Donnelly *et al*. (2016) aimed to solely investigate the role of SpoVAC, but due to the insertional mutagenesis method employed, polar effects on SpoVAD and SpoVAE cannot be ruled out, as the authors themselves highlight. Nonetheless, this study confirms their findings that spores lacking functional SpoVAC proteins are more susceptible to wet heat treatments.

This study measured the DPA content of spores of the parental 630Δ*erm* strain, the *spoVA* whole operon deletion strain, CRG5234, and the fully and dually complemented derivatives using terbium fluorescence. From these measurements, the DPA content per spore was approximated as 8.64 pg for 630Δ*erm* and 16.34 pg for CRG5235, whilst spores of strains lacking one or all three *spoVA* gene(s) contained approximately 1.21-1.37 pg DPA per spore. This value of 8.64 pg.spore^{-1} for parental 630Δ*erm* is consistent with previous studies in other spore-forming organisms. For example, using similar methods Jamroskovic *et al*. (2016) determined that spores of *C. acetobutylicum* DSM792, *C. beijerinckii* DSM791, *B. subtilis* PY79 and *Clostridium collagenovorans* DSM3089 contained approximately 1.98, 0.088, 1.35 and 9.25 pg.spore^{-1} of DPA. Whilst at lower levels than was observed for wildtype spores, DPA was still detected in spores of 630Δ*erm* lacking one, or all, SpoVA proteins in this study. Assuming the terbium fluorescence assay used in this study is specific

for DPA, this result could be explained by leakage of DPA into the forespore across the spore inner membrane. Alternatively, non-specific interactions with other spore inner membrane proteins could be responsible for the low-level transport of this molecule into these spores from the site of DPA synthesis in the mother cell. This could be investigated further via the creation of individual *spoVA* mutants in a *spoVF* background, unable to synthesise DPA, and assessing the DPA content of spores produced in media with and without supplementation of exogenous DPA.

3.12 Summary

- *C. difficile* 630Δ*erm* s*poVA*$^{-}$ strains produce spores which do not lyse during the sporulation process and can complete germination and outgrowth into vegetative cells.
- Spores of *C. difficile* 630Δ*erm spoVA*$^{-}$ strains contain lower amounts of calcium and the spore specific molecule Dipicolinic Acid than their wild-type counterparts.
- The presence of all three *C. difficile spoVA* gene products are required for transport of DPA in the spore core.
- DPA-less 630Δ*erm*Δ*spoVA* spores are more susceptible to wet heat, UV radiation and ethanol treatments than the wild-type or fully complemented counterparts.
- Spores of 630Δ*erm* display significant levels of clumping following repeated centrifugation steps during spore washing, which can be abated by washing spores in PBS with decreasing concentrations of the detergent Tween80.

Chapter Four

Implementing the gene editing road map in *C. difficile* DH1916

4 Implementing the gene editing roadmap in *C. difficile* DH1916

4.1 Introduction

At the onset of this work, allelic exchange mutagenesis was established in two *C. difficile* strains, 630Δ*erm* (an erythromycin-sensitive derivative of *C. difficile* 630 isolated from a CDI patient in Zurich, Switzerland in 1982 and belonging to PCR ribotype 012) and R20291 (a PCR ribotype 027 strain isolated from a CDI patient in Aylesbury, UK in 2006 [Stabler *et al*., 2009]). The subsequent generation of in-frame deletion mutants via allelic exchange, including multiple deletions, in the *pyrE*-mutated derivatives of these strains has facilitated valuable insights into the biology of *C. difficile* (Ng *et al*., 2013). R20291 belongs to the BI/NAP1/027 (restriction endonuclease type B1, North American pulsed-field type 1, PCR-ribotype 027; hereafter referred to as ribotype-027) group of *C. difficile* strains, which have been associated with more severe disease symptoms, higher rates of relapse and an increased repertoire of antibiotic resistance. Consequently, these *C. difficile* strains are commonly referred to as being 'hypervirulent'. Furthermore, ribotype-027 strains have also been shown to produce higher titres of toxins A and B, the main virulence factors associated with CDI, in the laboratory (Warny *et al*., 2005). The molecular basis for these observations has been studied and was thought to arise from a single base pair deletion within *tcdC*, the negative regulator of toxins A and B, which generates a frame-shift mutation and results in a truncated, non-functional TcdC protein (Matamouros *et al*., 2007). Other potential factors which may contribute to this 'hypervirulence' include the production of an additional ADP-ribosyltransferase toxin (binary toxin) and increased motility of the ribotype-027 strains. The

high level of resistance to fluoroquinolone antibiotics demonstrated by ribotype-027 strains is due to point mutations in DNA gyrase genes and is thought to be a major contributing factor to the predominance of these strains within health care facilities (Stabler *et al*., 2009).

A dramatic shift in the epidemiology of CDI was experienced worldwide from the turn of the millennium, a shift which was characterised by the increased prevalence of ribotype-027 strains in CDI cases. For example, before 2000, ribotype-027 strains accounted for fewer than 1% of recorded *C. difficile* isolates in North America (Freeman *et al*., 2010). Subsequent assessment of eight healthcare facilities in the US between 2000 and 2003 found ribotype-027 strains were responsible for 51% of CDI cases (McDonald *et al*. 2005), whilst across twelve hospitals in Quebec in 2004 ribotype-027 strains accounted for 84% of all CDI cases (Loo *et al*., 2005). During this period, mortality rates linked to CDI in the US also increased from 5.7 deaths per million in 1999 to 23.7 deaths per million of the population in 2004 (Redelings, *et al*., 2007). This observation was repeated in the UK, where ribotype-027 strains accounted for 41.3% of clinical isolates from English hospitals sampled between 2007 and 2008 (Brazier *et al*., 2008). Similar to results from the US, the increased prevalence of ribotype-027 strains in England and Wales also coincided with increased mortality, with CDI-related deaths per million of the population increasing from 23.3 in 2004 to 82.9 in 2007 (Carter, 2013).

In response to studies claiming increased spore production amongst ribotype-027 strains was responsible for the 'hypervirulent' nature of these strains, Burns *et al*. (2011) compared the sporulation characteristics of fifty-three *C. difficile* strains, including twenty-eight ribotype-027 isolates. This study concluded that whilst substantial differences in *C. difficile* sporulation rates occurred between individual isolates, there was no correlation between PCR-ribotype and rate of sporulation. Of the ribotype-027 strains tested in this study, the

isolate which produced the most heat-resistant CFU after five days growth in BHIS broth was DH1916, which was isolated from a CDI patient in Torbay, UK. Preliminary work with DH1916 showed that its spores did not form clumps during purification via repeated washing in distilled water.

Initial molecular genetics investigations into these ribotype-027 strains were performed by Clostron insertional inactivation of genes, for example, Deakin *et al.*, (2012) inactivated *spo0A* in strain R20291 and showed that this gene is essential for *C. difficile* spore formation and is required for persistence and transmission in a mouse model. Advanced genetics investigations into ribotype-027 strains were facilitated by Ng *et al*. (2013), who established allelic exchange mutagenesis in the R20291 strain. Subsequent characterisation of sporulation and germination in genome edited R20291, and 630Δ*erm*, strains are hampered by difficulties in obtaining large numbers of high purity spores which are free from clumping. Spores produced by both strains exhibit clumping during the repeated centrifugation steps of spore washing required to obtain pure spore preparations. As discussed in the previous chapter, this clumping limits the downstream applications of these spore preparations since clumping impedes the accuracy of both spore counting under light microscopy and assays measuring changes in optical density. Hence, DH1916, a ribotype-027 strain which produces higher spore titres than R20291 and spores which do not clump during purification unlike R20291, has the potential to be developed into a model ribotype-027 laboratory strain.

4.1.1 Aims of this study

An improved laboratory strain of a ribotype-027 *C. difficile* for studying the processes of sporulation and germination is one which generates higher spore titres relative to other ribotype-027 strains, and whose spores do not clump during spore purification washes.

Hence, this study aims to further characterise, and establish allelic exchange mutagenesis in, *C. difficile* DH1916, with the aim of assessing the suitability of this strain for use as a model ribotype-027 strain.

4.2 Comparative studies between R20291 and DH1916

4.2.1 Growth in BHIS broth

The growth characteristics of both ribotype-027 strains were assessed in BHIS broth over 24 hours using changes in measured OD_{600} (Figure 4.1). R20291 exhibited a longer lag phase than was observed for DH1916 or 630Δ*erm*, as indicated by lower OD_{600} values for R20291 during the first 5 hours of growth. From 6 hours onwards, the growth of DH1916 was indistinguishable from that of R20291.

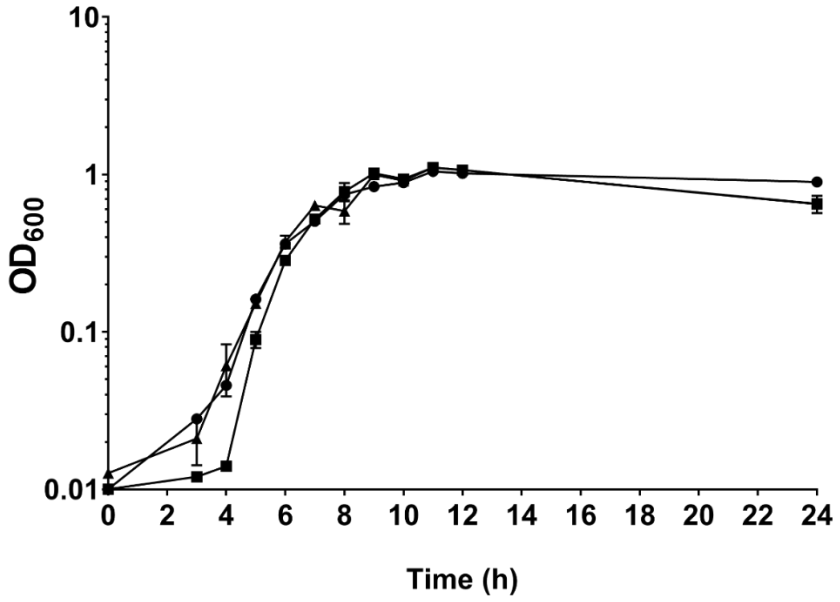


Figure 4-1 Growth of *C. difficile* strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}

● *C. difficile* 630Δ*erm*; ■, R20291; ▲, DH1916. Symbols represent the averages of three independent experiments. Error bars indicate the standard errors of the mean. Statistical significance determined using two-way ANOVA with Sidak's multiple comparisons test; *P*-value=0.8361 for variance between strains.

4.2.2 Colony formation after heat-treatment

To confirm the findings of Burns, *et al* (2011), that DH1916 exhibits greater sporulation over five days than R20291, both these strains were assessed for their ability to form heat-resistant CFU over five days on BHIS agar supplemented with the bile salt taurocholate (Figure 4.2). Heat-resistant CFU appeared after 24 hours for DH1916, but appeared 24 hours later and at a lower level for R20291, suggesting a delayed onset and/or rate of sporulation in R20291. Furthermore, the observed levels of heat resistant CFU for DH1916 were higher than for R20291 on each of the five days tested. After five days, the mean heat-resistant CFU observed were 1.44×10^6 for 630*Δerm*, for 9.57×10^4 R20291 and 1.31×10^6 for DH1916.

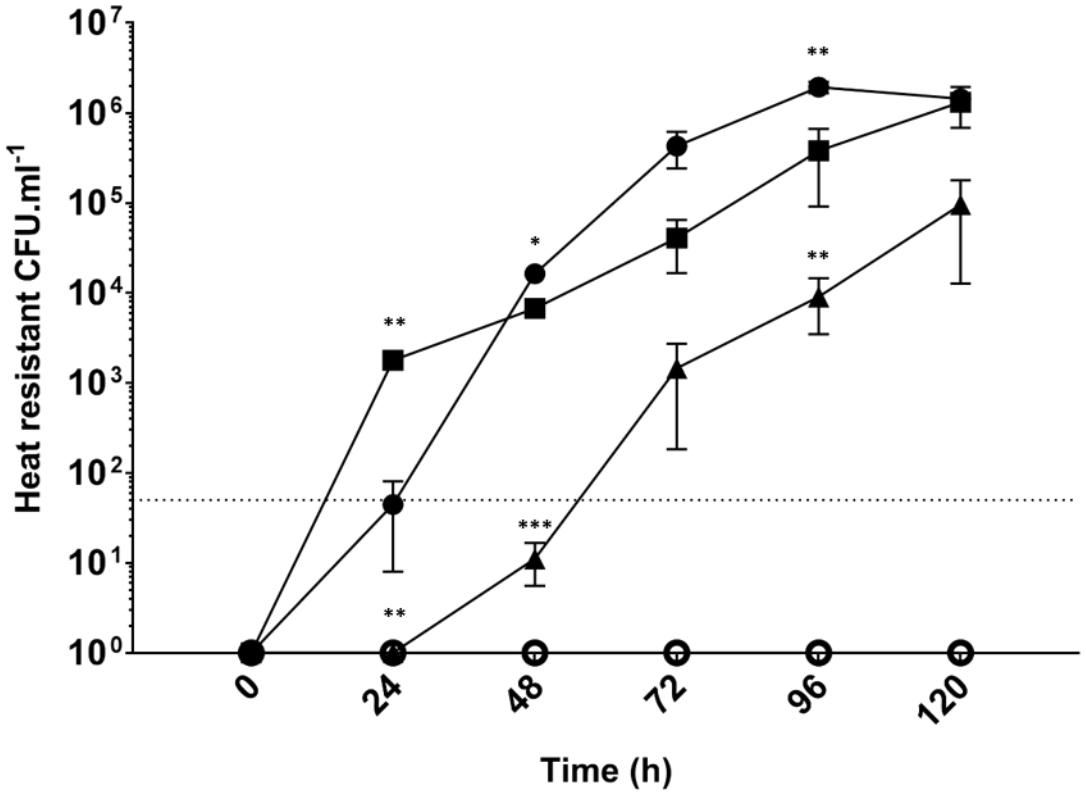


Figure 4-2 Development of heat resistant CFU of *C. difficile* strains over five days.

●, *C. difficile* 630Δ*erm*; ▲, R20291; ■, DH1916; **O**, *C. difficile* 630Δ*erm*Δ*spo0A*. Symbols represent the mean values from three independent experiments. Error bars indicate standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay.

4.2.3 Spore titres after five days

Burns *et al*. (2011) also observed an increased spore titre after five days growth in BHIS broth for DH1916 relative to R20291. To confirm this finding, aliquots from the five-day sporulation cultures in BHIS broth used to assess the development of heat resistant CFU (section 4.2.2) were taken, loaded onto a haemocytometer and spores counted under a phase-contrast microscope (Figure 4.3). Spores/ml values were calculated according to the protocol in Materials and Methods.

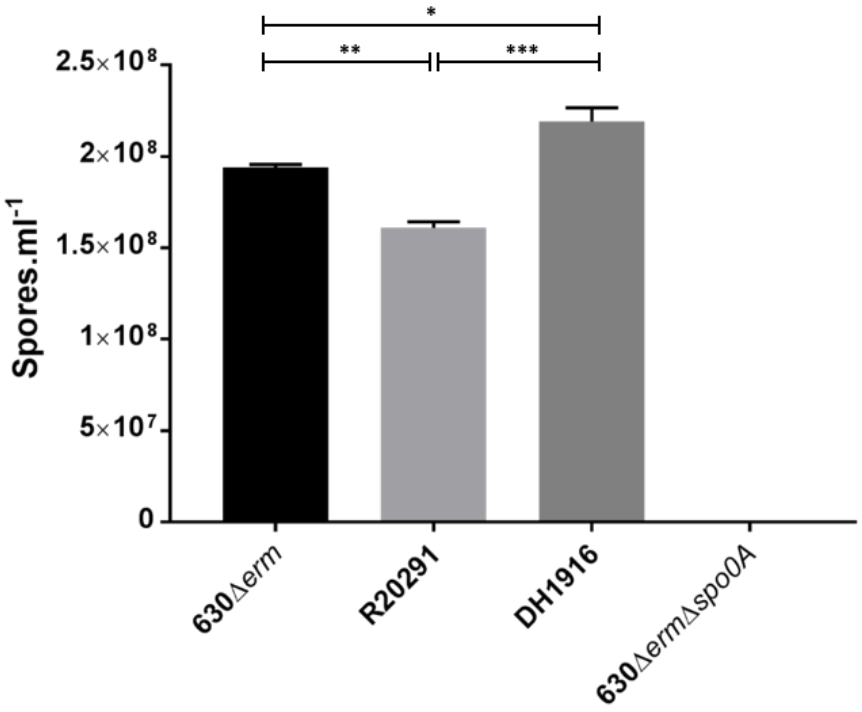


Figure 4-3 Spore titres of *C. difficile* strains enumerated by phase-contrast microscopy.
C. difficile strains grown in BHIS broth for five days, after which aliquots were loaded onto a haemocytometer, spore counts performed and the numbers of spores.ml^{-1} calculated. Bars indicate mean values from three independent experiments. Error bars indicate the standard errors of the means. The detection limit for spore counts was 5x10^3 spores. Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test. A single asterisk (*) denotes a *P*-value <0.05, two asterisks (**) denotes a *P*-value <0.01, and three asterisks (***) denotes a *P*-value <0.001.

Mean spore titres observed were $1.94x10^8$ for 630Δ*erm*, $1.61x10^8$ for R20291 and $2.19x10^9$ for DH1916, whilst no spores were observed for the 630Δ*erm*Δ*spo0A* control strain. These results, showing a >25% reduction in R20291 spores produced relative to the DH1916 spore titre, confirm the findings of the previous study by Burns *et al*. (2011).

4.2.4 Assessment of spore clumping following purification

Previous studies have found that spores of R20291 clump together in a similar manner to 630Δ*erm* spores observed in section 3.7 of this study, whilst DH1916 spores did not, following repeated centrifugation steps during spore washing. To confirm these observations, R20291 and DH1916 spores were harvested from five-day sporulation plates and purified via repeated washes in distilled water, as described in Materials and Methods. The extent of spore clumping within these purified spore preparations of each strain was then assessed by eye under a phase-contrast microscope (Figure 4.4). Spores of R20291 were found to clump in aggregates of various sizes, whilst no clumping was observed with DH1916 spores. Furthermore, the spore pellet of DH1916 was larger than that for R20291 spores (not shown), which made purification of DH1916 spores a far simpler task.

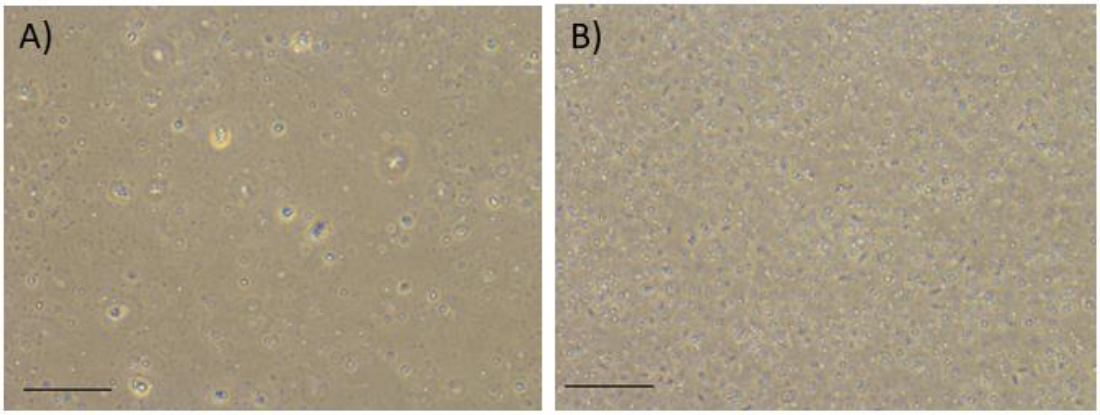


Figure 4-4 Assessment of spore clumping in R20291 and DH1916 spore preparations

Images captured during phase contrast-microscopy of undiluted, pure spore preparations of (A) R20291 spores and (B) DH1916 spores, from which the extent of spore clumping was assessed by eye. The scale bar represents 200 µm.

4.2.5 Cytotoxicity against Vero cells

A key characteristic of the ribotype-027 strains is their increased levels of toxin production relative to members of other ribotypes. Previous studies in R20291 have shown that ribotype-027 strains produce 16-fold and 23-fold higher levels of toxin A and toxin B, respectively, than do strains belonging to other PCR-ribotypes *in vitro* (Warny *et al*., 2005). It was therefore necessary to confirm, prior to establishing DH1916 as a new model ribotype-027 strain, that this isolate produced similar levels of toxin to R20291 *in vitro*. Hence, cell-free supernatants were obtained from overnight cultures of 630Δ*erm*, R20291 and DH1916 adjusted to the same OD_{600} value, and used to assess the *in vitro* cytotoxicity of these strains against Vero cells. After 24 hours incubation, the end-point titres were determined under light microscopy by identifying the dilution at which there was no visible change in cell rounding relative to the supernatant-free controls (Figure 4.5).

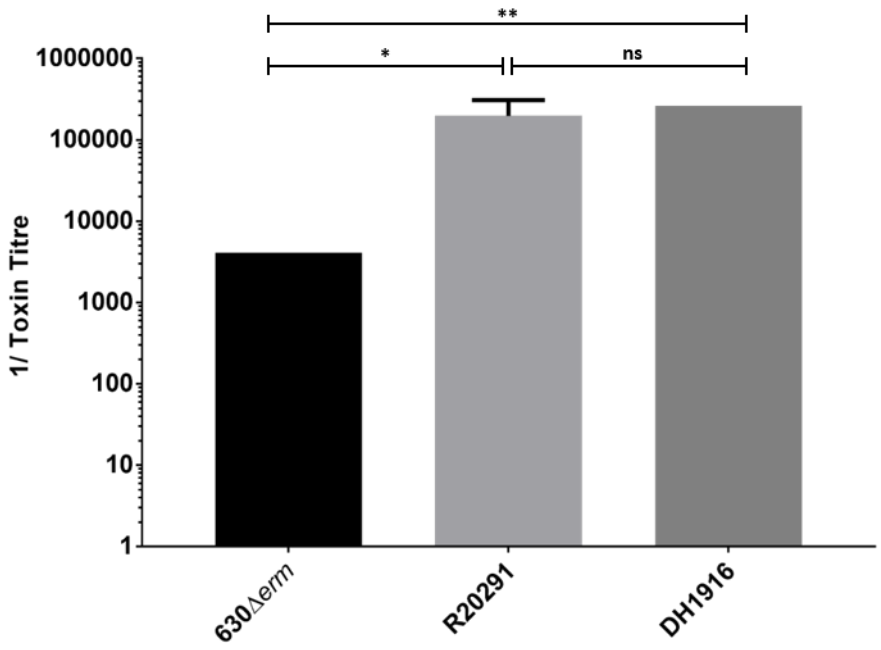


Figure 4-5 *In vitro* cell cytotoxicity of *C. difficile* strains against Vero cell cultures.

C. difficile supernatants were adjusted for OD_{600} and four-fold serial dilutions incubated with cultured Vero cells for 24 hours. Subsequent rounding of Vero cells was observed under light microscopy and end-point titres determined. Bars represent the mean end-point titres from three independent reactions. Error bars indicate the standard errors of the means. Error bars indicate the standard errors of the means. Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, single asterisk (*) denotes a *P*-value <0.05, and two asterisks (**) denotes a *P*-value <0.01.

4.3 Assessment of conjugation efficiency into DH1916

4.3.1 Conjugation efficiency into DH1916 from CA434 donor cells

The ability to transfer DNA into a particular strain is the critical barrier to overcome in order to establish genome editing. Since *C. difficile* cannot readily be transformed, heterologous DNA must be supplied via conjugation reactions from suitable conjugal donor strains. To determine the efficiency of DNA transfer into DH1916, a series of conjugations were performed into DH1916 recipient cells from donor *E. coli* CA434 cells harbouring pMTL-

80000 modular vectors containing various Gram-positive replicons (Figure 4.6). The bacterial source of each replicon is detailed in Table 4.1. Mean conjugation efficiencies into DH1916 from *E. coli* CA434 were $4.94x10^{-8}$ for pMTL82151 (pBP1), $1.09x10^{-9}$ for pMTL83151 (pCB102) and $3.02x10^{-8}$ for pMTL84151 (pCD6). In all cases, these values corresponded to fewer than 10 transconjugant colonies per conjugation. No thiamphenicol-resistant, transconjugant colonies were obtained following conjugations with vectors pMTL81151 (no Gram-positive replicon), pMTL85151 (pIM13) nor pMTL86151 (pIP404). The low conjugation efficiencies observed for the pMTL80000 modular vectors into DH1916 are consistent with previous studies in R20291, where conjugation efficiencies for vectors pMTL82151, pMTL83151, pMTL84151 and pMTL85151 from *E. coli* CA434 donors into R20291 was $2.61x10^{-7}$, $3.40x10^{-8}$, $4.48x10^{-7}$ and zero, respectively (Michelle Lister, unpublished). Furthermore, the same study showed conjugation efficiencies of vectors pMTL82151, pMTL83151, pMTL84151 and pMTL85151 from *E. coli* CA434 donors into 630Δ*erm* was $3.36x10^{-6}$, $2.23x10^{-6}$, $7.00x10^{-6}$ and $4.18x10^{-7}$, respectively (Michelle Lister, unpublished), suggesting low conjugation efficiency is a characteristic of ribotype-027 strains.

pMTL80000 vector	Gram positive replicon	Source Organism
pMTL81151	None	N/A
pMTL82151	pBP1	*Clostridium botulinum*
pMTL83151	pCB102	*Clostridium butyricum*
pMTL84151	pCD6	*Clostridium difficile*
pMTL85151	pIM13	*Bacillus subtilis*
pMTL86151	pIP404	*Clostridium perfringens*

Table 4-1 Gram-positive replicons in the pMTL80000 modular vector series.

Gram-positive replicons within the pMTL80000 modular vector series and the organisms these were sourced from.

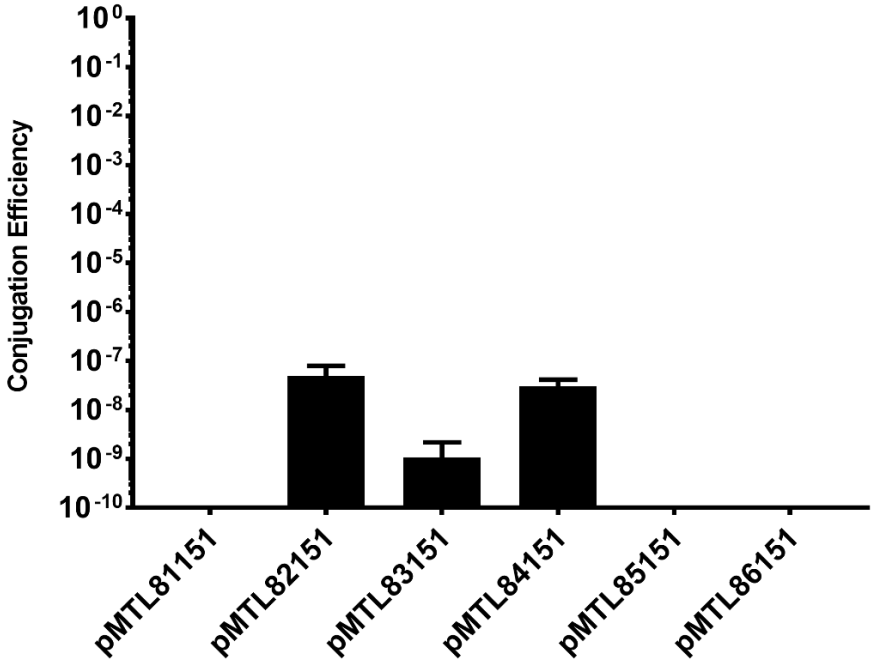


Figure 4-6 Conjugation efficiency into DH1916 from CA434 donor cells.

Conjugations into DH1916 from *E. coli* CA434 cells harbouring the various pMTL-80000 modular vectors listed were performed as described in Materials and Methods. Conjugation efficiency was calculated as the proportion of thiamphenicol resistant CFU obtained relative to the recipient *C. difficile* DH1916 CFU. Bars represent the means of three independent experiments. Error bars indicate the standard errors of the means. Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test (*P*-value = 0.1085).

4.3.2 Effect of heat-treatment on conjugation efficiency into DH1916

A sub-lethal exposure of bacterial strains to heat, immediately prior to transformation, has been shown to increase the efficiency of DNA transfer. Kirk & Fagan (2016) recently demonstrated such an increase in DNA transfer via conjugation after heating recipient R20291 cells at 52^{o}C for 5 minutes. To determine whether such an effect could be replicated in DH1916, this heat treatment step was conducted using overnight cultures of DH1916 immediately prior to conjugation from CA434 donor cells harbouring either pMTL82151, pMT83151 or pMTL84151 vectors, and conjugation efficiencies calculated (Figure 4.7). Mean

unheated conjugation efficiencies were $1.60x10^{-8}$, $4.90x10^{-9}$ and $2.87x10^{-8}$ for vectors pMTL82151, pMTL83151 and pMTL84151 respectively. Meanwhile, equivalent values of conjugation efficiency of pMTL82151 and pMTL84151 into heat-treated DH1916 were $5.86x10^{-8}$ and $1.63x10^{-8}$ respectively. No thiamphenicol-resistant colonies were obtained following conjugations with pMTL83151 into heat-treated DH1916. Hence, this study was unable to replicate for DH1916, the increase in conjugation efficiency observed following heat-treatment of R20291.

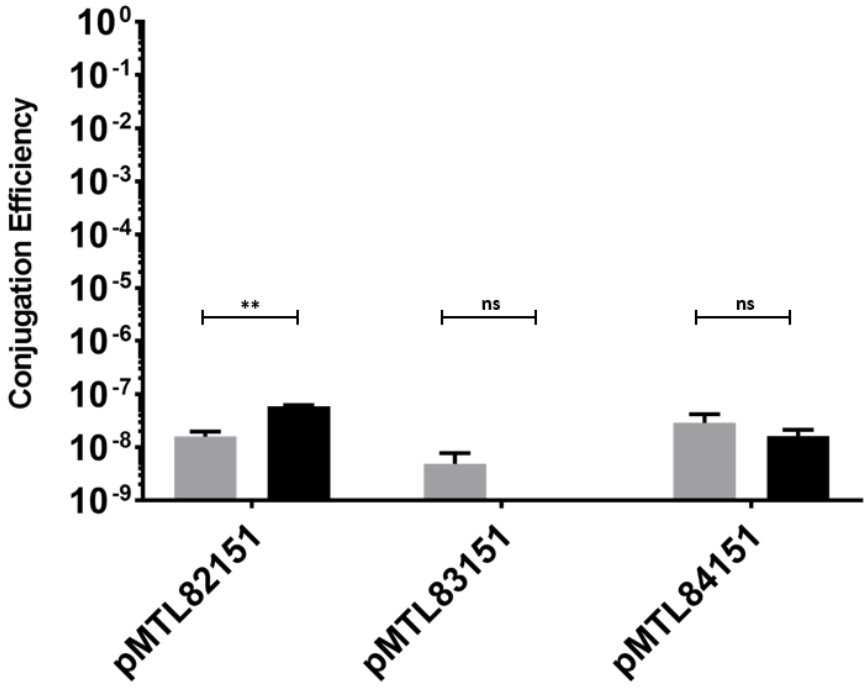


Figure 4-7 Effect of heat-treatment on conjugation efficiency into DH1916.

Conjugations into untreated (grey bars) or heat-treated (black bars) DH1916 overnight cultures from *E. coli* CA434 cells harbouring various pMTL-8000 modular vectors were performed as described in Materials and Methods. Conjugation efficiency calculated as the proportion of thiamphenicol resistant CFU obtained relative to the recipient *C. difficile* DH1916 CFU. Bars represent the means of three independent experiments. Error bars indicate the standard errors of the means. Statistical significance determined using multiple unpaired t-tests. Not significant (ns) denotes a *P*-value >0.05, two asterisks (**) denotes a *P*-value <0.01.

4.3.3 Restriction-modification systems in DH1916

The potential for the observed low conjugation efficiencies into DH1916 being due to the presence of restriction-modification (RM) systems was investigated. The online NEB Rebase database was accessed to find all RM systems present in R20291 (Figure 4.8), and the DNA sequences of the identified genes were searched for within the DH1916 draft genome (Table 4.2). Three RM systems are present in R20291, one type I, one type II and one type IV, of which only the type I and type IV systems contain restriction subunits. The type I system present contains a restriction subunit (CDR20291_2909), specificity subunit (CDR20291_2911) and modification subunit (CDR20291_2912) all with specificity for CA-N_7-TAAAG sequences. The single ORF within the type II system present encodes a modification subunit with specificity for C-A_5 sequences. The type IV system present contains two restriction subunit genes *mcrBP* (CDR20291_2001) and *mcrCP* (CDR20291_2000), which are identified as 'putative Type IV methyl-directed restriction enzyme of unknown recognition sequence' (Roberts *et al*., 2015). Each of these R20291 restriction-modification system ORFs identified was also found within the DH1916 genome, with 100 % DNA sequence identity in each case.

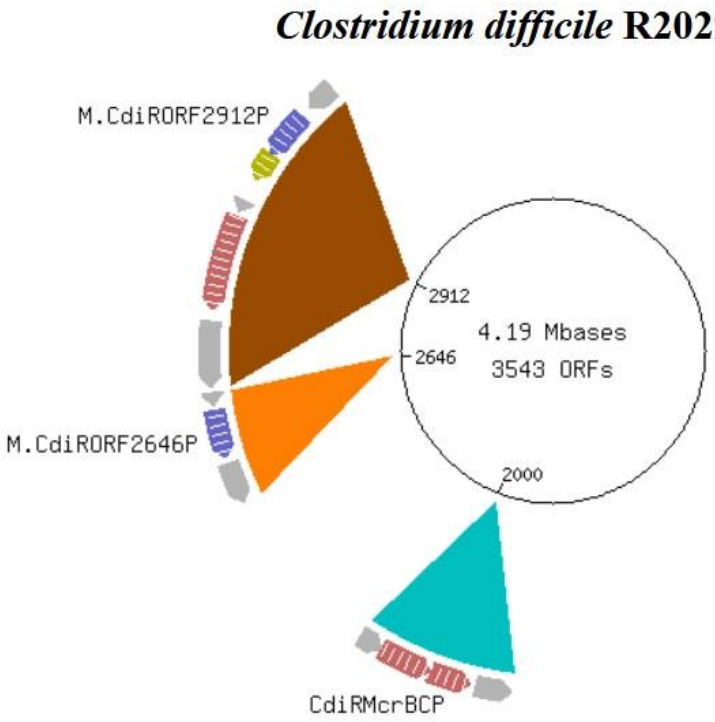


Figure 4-8 Identification of Restriction-Modification (RM) systems within *C. difficile* R20291.

ORFs (arrows) within RM systems present at the positions indicated along the circular R20291 genome shown using the online REBASE database (NEB). Modification subunits (purple arrows), restriction subunits (red arrows) and specificity subunits (green arrows) are shown within each system. RM system classes denoted by colour of expansion fragments; Type I (brown), Type II (orange) and Type IV (blue).

RM System	RM component	R20291 ORFs	DH1916 homologue	% identity
Type I	hsdM	CDR20291_2912	CDDH1916_00407	100 (1464/1464)
	hsdS	CDR20291_2911	CDDH1916_00408	100 (1185/1185)
	hsdR	CDR20291_2909	CDDH1916_00412	100 (3336/3336)
Type II	(M)	CDR20291_2646	CDDH1916_00686	100 (1734/1734)
Type IV	McrCP	CDR20291_2000	CDDH1916_01369	100 (1269/1269)
	McrBP	CDR20291_2001	CDDH1916_01368	100 (1935/1935)

Table 4-2 Homologues of R20291 Restriction-Modification system ORFs present in DH1916.

R20291 ORFs encoding restriction (R), modification (M) or specificity (S) subunits of RM systems identified by REBASE were searched against the DH1916 draft genome sequence in Artemis. BLAST alignments were used to determine the % identity of DH1916 homologues to each R20291 RM ORF.

4.3.4 Conjugation efficiency from NEB sExpress donor cells

In an attempt to circumvent the barrier to DNA transfer posed by Type IV restriction systems, which cleave methylated DNA, a new conjugation donor strain of *E. coli* with fewer methylation capabilities was sought. One such strain, *E. coli* Express, has the genotype *dam+ dcm-* and is therefore unable to generate 5-methylcytosine nucleotides which can be targeted by some Type IV restriction systems. Work by Craig Woods (unpublished) converted NEB Express into a conjugation donor strain by transferring the R702 conjugal plasmid from CA434 into NEB Express. The resulting strain was named NEB sExpress and into which the six previously used pMTL80000 modular vectors containing various Gram-positive replicons (pMTL81151-86151) were individually electroporated. Conjugations into DH1916 using the resulting chloramphenicol resistant NEB sExpress strains were performed and the conjugation efficiencies calculated as previously (Figure 4.9). Successful transfer of DNA into

DH1916 from sExpress conjugal donors was achieved for all modular vectors tested, with the exception of the pMTL-81151 control vector lacking a Gram-positive replicon. This compares favourably with the equivalent experiment using *E. coli* CA434 conjugal donor strains, from which no transconjugant colonies were obtained for vectors pMTL85151 nor pMTL86151. Furthermore, increases in conjugation efficiency for the transfer of vectors pMTL82151, pMTL83151 and pMTL84151 were observed when using NEB sExpress relative to CA434 in each case.

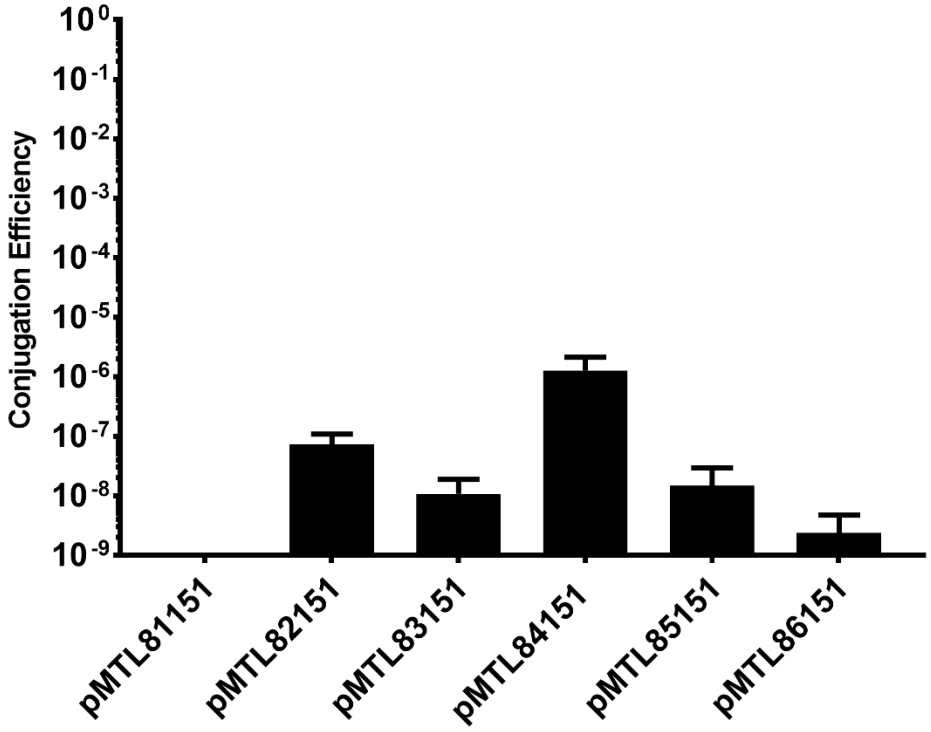


Figure 4-9 Conjugation efficiency into DH1916 from NEB sExpress donor cells.

Conjugations into DH1916 overnight cultures from *E. coli* NEB 'sExpress' cells harbouring various pMTL-80000 modular vectors were performed as described in Materials and Methods. Conjugation efficiency calculated as the proportion of thiamphenicol resistant CFU obtained relative to the recipient *C. difficile* DH1916 CFU. Bars represent the means of three independent experiments. Error bars indicate the standard errors of the means. Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test (*P*-value = 0.1623).

Using NEB sExpress increased conjugation efficiency from $4.94x10^{-8}$ to $7.41x10^{-8}$ for pMTL82151, from $1.09x10^{-9}$ to $1.08x10^{-8}$ for pMTL83151, and from $3.02x10^{-8}$ to $1.27x10^{-6}$ for pMTL84151 compared with CA434. Following these results, all future conjugations into DH1916 were performed using NEB sExpress as the conjugal donor strain.

4.4 Generation of a DH1916Δ*pyrE* strain

4.4.1 Confirmation of 5-Fluoroorotic acid susceptibility

A pre-requisite to establishing allelic exchange mutagenesis in DH1916 is the generation of a *pyrE*-truncated strain which is auxotrophic for uracil and resistant to 5-fluoroorotic acid (FOA). This is performed using Allele Coupled Exchange (ACE) and allows a heterologous *pyrE* allele to be used as a counter-/negative selection marker with supplemented 5-FOA. Hence, it was first necessary to confirm that 5-FOA is suitably able to counter select against DH1916 strains carrying a functional copy of the *pyrE* gene. Equivalent counter selection for allelic exchange mutagenesis of R20291 is performed using *C. difficile* minimal media (CDMM) agar plates supplemented with 2 $mg.ml^{-1}$ 5-FOA (Ng *et al*., 2013). Hence, DH1916 colonies were passaged onto CDMM plates supplemented with or without 2 $mg.ml^{-1}$ 5-FOA and subsequent growth assessed following incubation in the anaerobic cabinet for 48 hours (Table 4.3). DH1916 grew reasonably well on CDMM plates, whilst no colonies were observed on the same media supplemented with 2 $mg.ml^{-1}$ 5-FOA, confirming this as a suitable concentration to use for counter selection against cells containing functional *pyrE* alleles.

	Growth on BHIS	Growth on CDMM	Growth on CDMM + 5-FOA
R20291	✓✓✓	✓✓	0
DH1916	✓✓✓	✓✓	0

Table 4-3 Susceptibility testing of DH1916 to 5-FOA.

Growth of R20291 (control) and DH1916 on BHIS or CDMM agar plates with or without supplementation with 2mg.ml^{-1} 5-FOA was assessed by eye after 48 hours. Growth was scored from zero, representing no visible colonies, to three ticks, representing normal, uninhibited growth.

4.4.2 Construction of a DH1916 *pyrE* ACE vector

Alignments of the nucleotide sequences of the *pyrE* genes within R20291 and DH1916 were performed and found to be an exact match (Table 4.4). The ACE vector pMTL-YN18 was used by Ng *et al*. (2013) to generate R20291Δ*pyrE*, which lacks 234 bp from the 3'-end of the 585 bp R20291_0188 (*pyrE*) gene. The 1200 bp long homology arm and 300 bp short homology arm sequences of pMTL-YN18 were searched for in the DH1916 genome and 100 % identity was found in each case. Unfortunately, the pMTL-YN18 vector was not stored in our culture collection and therefore needed to be reconstructed. The long homology arm comprising 1200 bp downstream of the DH1916 *pyrE* gene, and short homology arm, comprising an internal fragment of the *pyrE* gene lacking 50 nucleotides from the 5'-end and 234 bp from the 3'-end of the *pyrE* gene, were both PCR amplified from DH1916 gDNA using primer pairs DH1916_LF1 and DH1916_LR1, and DH1916_RF1 and DH1916_RR1, respectively. The long homology arm was cloned via restriction digests with *Asc*I and *Nhe*I into similarly digested pMTL-JH18::λ6.5, the *pyrE* ACE vector used to generate 630Δ*erm*Δ*pyrE*. The short homology arm was then cloned into the resulting vector via restriction digests with *Sbf*I and *Not*I enzymes. Sanger sequencing confirmed the successful replacement of both 630 homology arms with their DH1916 counterparts and thus the generation of pMTL-PSI18. This plasmid

varies from pMTL-YN18 in the Gram-positive replicon present; pB1 in pMTL-YN18, pCB102 in pMTL-PSI18.

	pyrE locus ID	Gene length (bp)	Identity to R20291_0188
R20291	R20291_0188	585	N/A
DH1916	DH1916_03289	585	100%

Table 4-4 Alignments of R20291 and DH1916 *pyrE* nucleotide sequences.

DNA sequences of *pyrE* genes from R20291 and DH1916 were aligned and found to share 100% identity.

4.4.2 Generation of DH1916Δ*pyrE*

ACE vector pMTL-PSI18 was conjugated into DH1916 and thiamphenicol resistant colonies arose on BHIS agar supplemented with thiamphenicol after 48-72 hours. Three colonies from independent conjugations were subsequently passaged once more on the same media and then twice further on *C. difficile* minimal media supplemented with 5-FOA and uracil. Three independent 5-FOA-resistant colonies were then cured of pMTL-PSI18 via growth in BHIS broth followed by patch plating onto BHIS agar with and without supplemented thiamphenicol. Three resulting independent, 5-FOA-resistant, thiamphenicol-sensitive colonies were PCR screened for the 234 bp truncation of *pyrE* using primers CD630_pyrD_sF1 and m13F (Figure 4.10). These PCR primers anneal to the DH1916 chromosomal *pyrD* and to *lacZ* within pMTL-PSI18, respectively. The PCR products from each of the three, independent, putative DH1916Δ*pyrE* strains are the 474 bp size expected and match the PCR product from the R20291Δ*pyrE* positive control. Also as expected, no PCR products were observed using the pMTL-PSI18 vector, DH1916 gDNA or the water control as PCR templates. To confirm the truncation of *pyrE*, these junction PCR products were excised and Sanger sequenced using the m13F primer. In each case the sequence was as intended confirming the successful generation of three independent DH1916Δ*pyrE* strains.

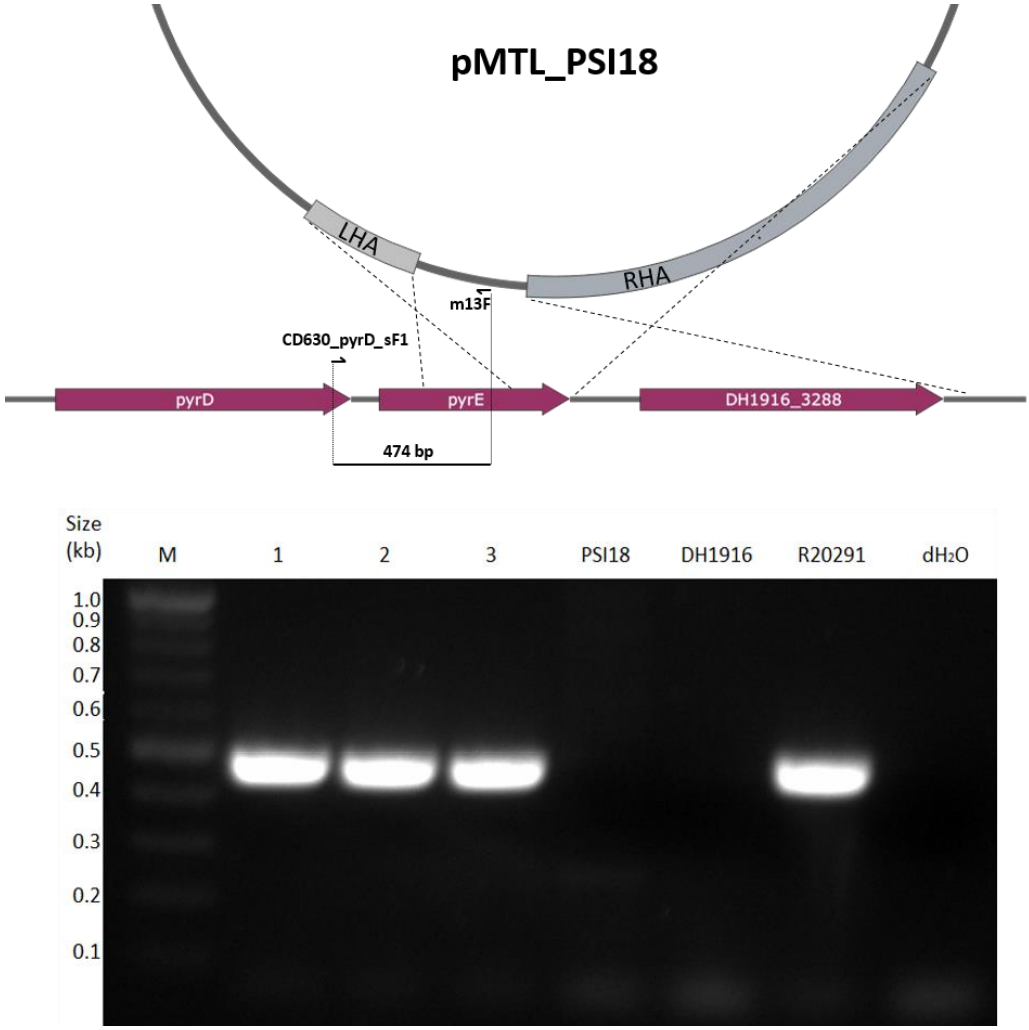


Figure 4-10 PCR screening for the generation of DH1916Δ*pyrE* strains.

Top: Diagrammatic representation of PCR primer binding sites within the DH1916 chromosome and pMTL-PSI18 plasmid. Bottom: The three independent, 5-FOA-resistant, thiamphenicol-sensitive DH1916 strains with putative *pyrE* truncations were PCR screened using primers CD630_pyrD_sF1 and m13F and the products resolved via agarose electrophoresis. M: DNA marker (2-log ladder; NEB); lanes 1-3, PCR products using 5-FOA-resistant, thiamphenicol-sensitive DH1916 as template; PSI18, pMTL-PSI18 PCR product; DH1916, DH1916 gDNA PCR product; R20291, R20291Δ*pyrE* positive control PCR product; dH2O, water control.

4.5 Establishing allelic exchange in DH1916

Two target genes were chosen for deletion, and therefore exemplification of allelic exchange mutagenesis, in DH1916; *spo0A* and *cspC*. These genes are essential for sporulation and germination in *C. difficile*, respectively. Hence, deletions of these genes will generate useful sporulation and/or germination control strains for subsequent genetic studies using this strain.

4.5.1 Construction of allelic exchange vectors targeting *spo0A* and *cspC*

The allelic exchange vector pMTL-YN4 has been used previously to generate in-frame deletion mutants in R20291, and on this basis, was chosen for allelic exchange mutagenesis in DH1916. In silico deletions of the DH1916 *spo0A* and *cspC* genes were performed and these sequences used for the designing of overlapping primers. Homology arm fragments consisting of approximately 1 kb up- and down-stream of each of the DH1916 *spo0A* and *cspC* genes were PCR amplified. These purified PCR products were used as templates in individual SOE PCRs to generate the requisite knockout cassettes targeting *spo0A* and *cspC* in DH1916, as detailed in Materials and Methods. The *spo0A* and *cspC* knockout cassettes were separately cloned into pMTL-YN4 via restriction digestions with *Asc*I and *Sbf*I enzymes followed by the appropriate ligation reactions. Successful cloning was confirmed by restriction digests and the two allelic exchange vectors produced were named pMTL-YN4::DH_*cspC* and pMTL-YN4::DH_*spo0A*.

4.5.2 Generation of sporulation/germination control strains in DH1916

Both pMTL-YN4::DH_*cspC* and pMTL-YN4::DH_*spo0A* were transformed into *E. coli* NEB SExpress via electroporation. The resulting chloramphenicol-resistant strains were used as donors for conjugations into *C. difficile* DH1916Δ*pyrE*, performed in triplicate for both vectors. Thiamphenicol-resistant colonies appeared after 48-72 hours on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol, and were passaged onto the same media to confirm the presence of the respective allelic exchange vector. Larger resulting thiamphenicol-resistant colonies, indicative of faster growth and, therefore, of genomic integration of the allelic exchange vector, were colony PCR screened to confirm these strains were single crossover mutants (as described in Materials and Methods; Results not shown). Confirmed single crossover strains were re-streaked twice onto *C. difficile* minimal media supplemented with FOA and uracil. Subsequent FOA-resistant colonies were patch plated onto BHIS agar with and without thiamphenicol to isolate thiamphenicol-sensitive colonies which were cured of their respective allelic exchange vectors. All resulting FOA-resistant, thiamphenicol-sensitive colonies were PCR screened using primer pairs DH1916_*cspC*_sF1 and DH1916_*cspC*_sR1, or DH1916_*spo0A*_sF1 and DH1916_*spo0A*_sR1, for the expected deletions within the *cspC* or *spo0A* genes respectively (Figure 4.11). Fourteen of the fifteen putative *spo0A* mutants yielded a 3.008 kb sized band indicative of successful in-frame deletion of the *spo0A* gene, whilst the other colony produced a 3.810 kb sized band consistent with the continued presence of the wild-type *spo0A* gene. Similarly, of the sixteen putative *cspC* mutants screened, ten produced a 2.719 kb fragment indicative of successful deletion of the *cspC* gene, five yielded a 4.379 kb fragment expected for an unmodified wild-type *cspC* allele, and one PCR did not generate any product. Three PCR confirmed strains of DH1916Δ*pyrE*Δ*spo0A* and DH1916Δ*pyrE*Δ*cspC*, each obtained from independent conjugation reactions, were Sanger sequenced and the in-frame deletions of *spo0A* or *cspC* were confirmed in each instance. All three strains of both mutants (n=6) were

then converted back to uracil prototrophy by restoring the *pyrE* locus back to wild-type. For this, the R20291 *pyrE* repair vector pMTL-YN2 was transformed into *E. coli* NEB sExpress via electroporation, and the resulting strain used as a conjugation donor for transfer of pMTL-YN2 into each DH1916 allelic exchange mutant. Colonies capable of growing on *C. difficile* minimal media without uracil supplementation were obtained for each triplicate of both mutants. Hence, DH1916Δ*spo0A* and DH1916Δ*cspC* strains were obtained, allowing phenotypic comparisons of the effects of these mutations with the wild-type DH1916.

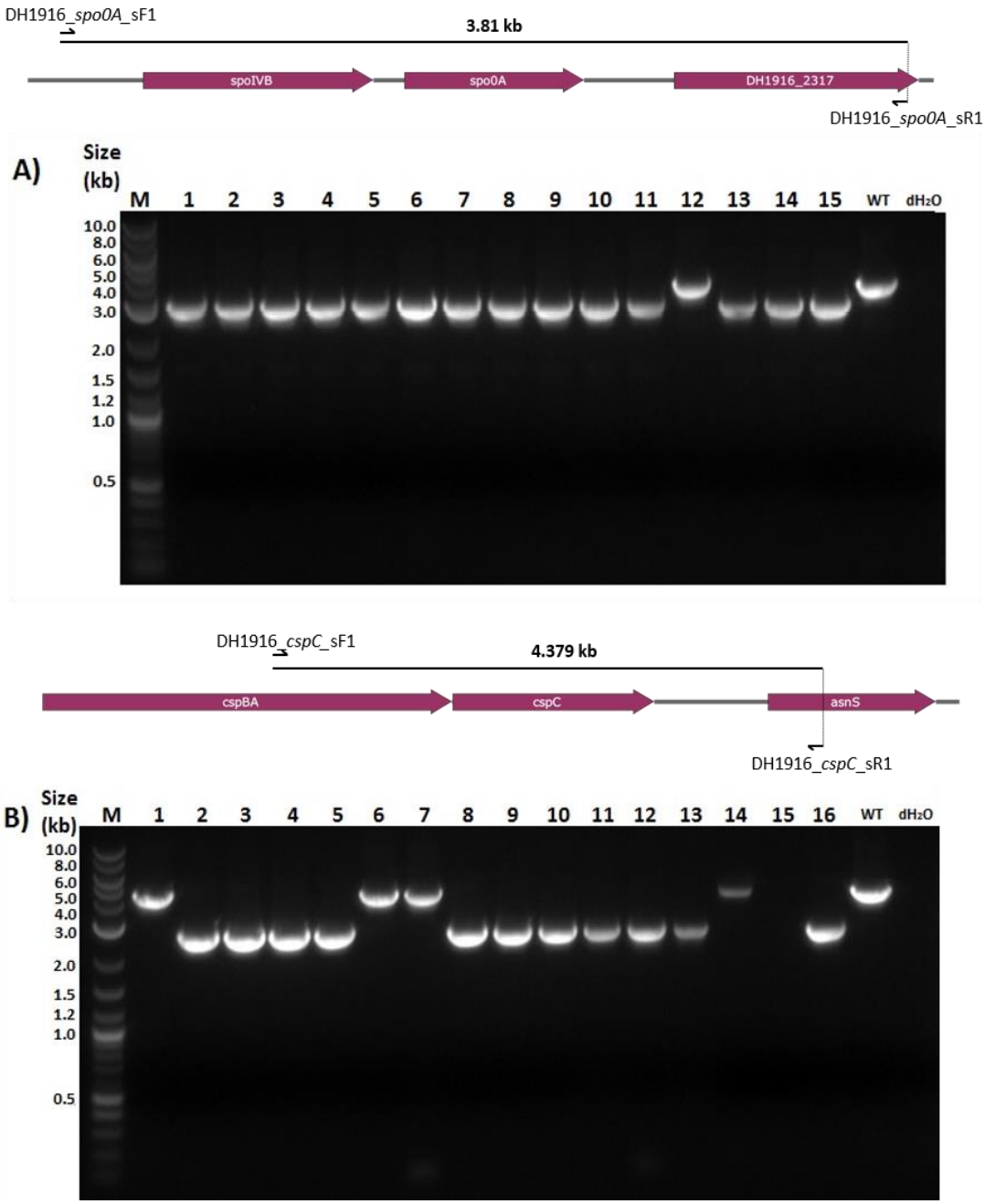


Figure 4-11 Colony PCR screening for *C. difficile* DH1916 in-frame deletion mutants.

Top: Diagrammatic representation of PCR primer binding sites. Bottom: Agarose gel electrophoresis; FOA-resistant, thiamphenicol-sensitive DH1916Δ*pyrE* colonies PCR screened following allelic exchange mutagenesis of A) *spo0A* (DH1916_2318) and B) *cspC* (DH1916_1210). M: DNA Marker (2-log ladder; NEB); Numbered lanes, FOA-resistant, thiamphenicol-sensitive DH1916Δ*pyrE* colonies; WT, DH1916 gDNA; dH_2O, distilled water.

4.5.3 Confirmation of sporulation/germination deficient phenotypes

To confirm the DH1916 mutants lacking either *spo0A* or *cspC* were deficient in sporulation or germination, respectively, these strains were compared to the DH1916 parental strain for their ability to form heat resistant CFU over five days (Figure 4.12). Neither the DH1916Δ*spo0A* nor DH1916Δ*cspC* strains produced any detectable heat-resistant CFU over the course of this assay, whilst after 120 hours the parental DH1916 strain produced a mean of 5.22×10^5 heat-resistant CFU.ml^{-1}. This suggests the *spo0A*$^-$ and *cspC*$^-$ DH1916 strains were indeed deficient in sporulation and/or germination.

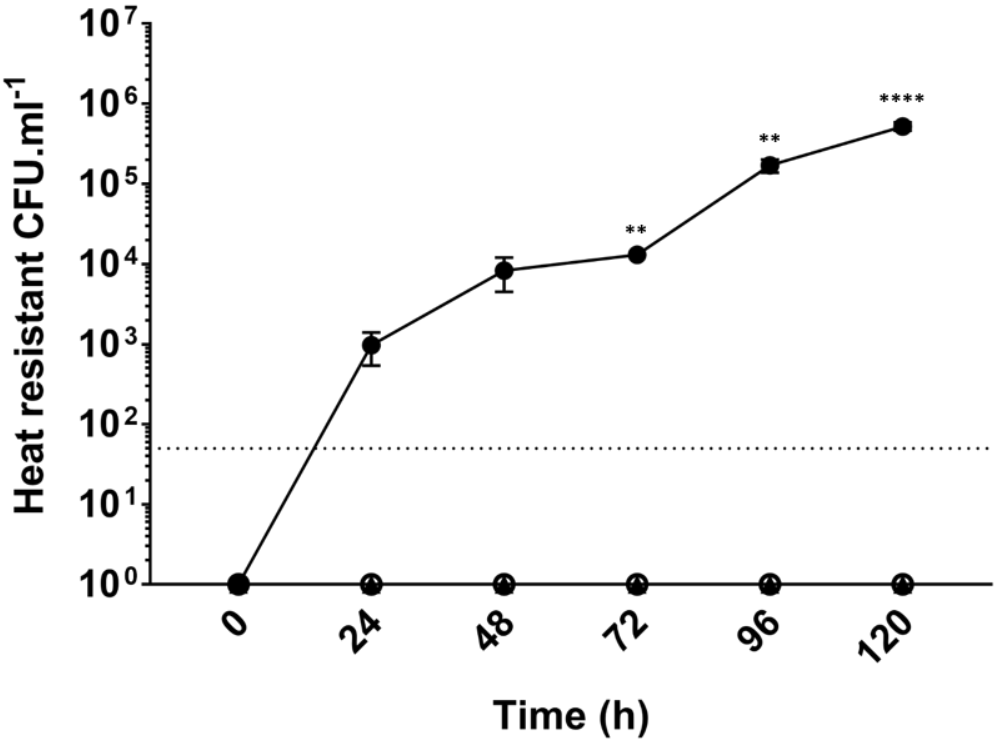


Figure 4-12 Development of heat resistant CFU of DH1916 strains over five days.

●, *C. difficile* DH1916; ▲, DH1916Δ*cspC*; **O**, DH1916Δ*spo0A*. Symbols represent the mean values from three independent experiments. Symbols for DH1916Δ*cspC* and DH1916Δ*spo0A* are overlaid as neither strain produced any heat-resistant CFU throughout the five days assayed. Error bars indicate standard errors of the means. The dotted line at 50 CFU.ml^{-1} denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Significant values are indicated by asterisks, where (**) denotes a *P*-value <0.01 and (****) denotes a *P*-value <0.0001.

Further characterisation was conducted to confirm the exact process in which these DH1916 *spo0A* and *cspC* mutants were deficient. Hence, the spore titres of each mutant after five days growth in BHIS broth were enumerated under light microscopy and compared to equivalent measures of the parental DH1916 strain (Figure 4.13). As expected, DH1916Δ*spo0A* did not produce any detectable spores whilst DH1916Δ*cspC* produced a mean spore titre of 1.11×10^8 spores.ml^{-1}. By comparison, the mean spore titre observed for DH1916 was 1.37×10^8 spores.ml^{-1}. These results confirm the suitability of DH1916Δ*spo0A* and DH1916Δ*cspC* mutants for use in subsequent studies as sporulation-negative or germination-negative control strains, respectively.

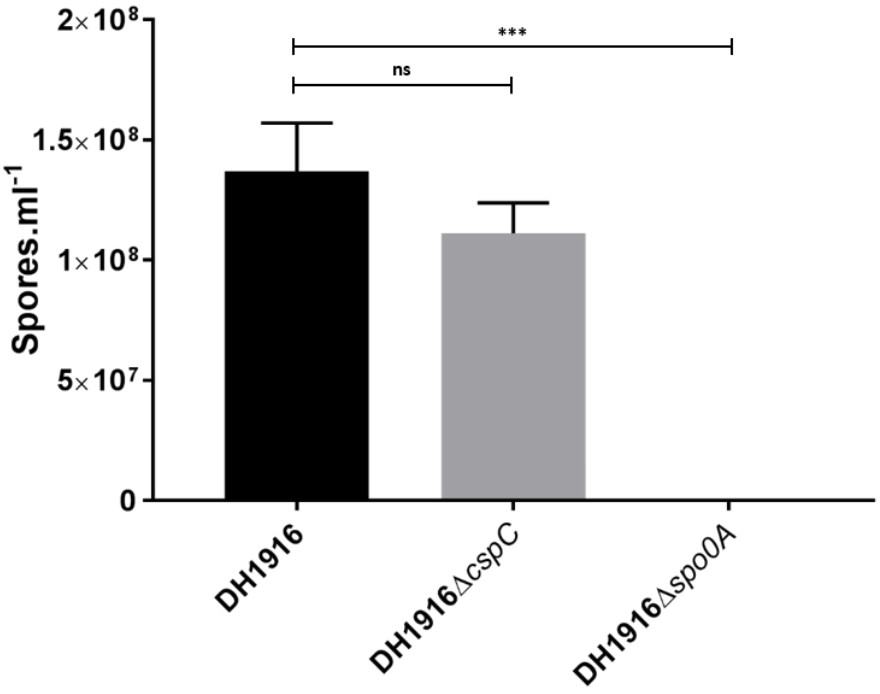


Figure 4-13 Spore titres of DH1916 strains enumerated by phase-contrast microscopy.

C. difficile DH1916 strains grown in BHIS broth for five days, after which aliquots were loaded onto a haemocytometer, spore counts performed and the numbers of spores.ml^{-1} calculated. Bars indicate mean values from three independent experiments. Error bars indicate the standard errors of the means. The detection limit for spore counts was 5×10^3 spores. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05, three asterisks (***) denotes a *P*-value <0.001.

4.6 Genome resequencing of DH1916Δ*pyrE* strains

Whilst efforts were underway to establish allelic exchange mutagenesis using the newly generated DH1916Δ*pyrE* strains, genomic DNA was prepared from these three, independently-generated mutants via phenol-chloroform extraction and sent for whole genome resequencing using the Illumina MiSeq platform (DeepSeq, Nottingham, UK). This was to determine whether any SNPs or other genomic variants had been accumulated during allele-coupled exchange used to generate these *pyrE* truncated mutants. The generated reads were aligned with the DH1916 draft genome annotation using the CLC Genomics Workbench and variants were called according to the guidelines stated in Materials and Methods (Tables 4.5 and 4.6).

Across all three DH1916Δ*pyrE* sequence alignments a total of 38 variants were identified, all of which were either SNPs or deletions. Of these 38 variants, 16 were common to all DH1916Δ*pyrE* strains, including a SNP at locus 967280 resulting in a non-synonymous mutation and premature stop codon in a putative membrane protein. WGS was not performed on the DH1916 culture from which these *pyrE*-truncation strains were created, making it impossible to conclusively state that these variants were present in the original culture. However, this seems the most likely explanation for the presence of these variants in all three mutants, each generated from independent conjugation reactions. Fifteen variants are detailed in Table 4.5, of which twelve are common to all strains, two SNPs within transposon-like protein b (DH_01477) are found only in DH1916Δ*pyrE* strains #1 and #3, whilst the remaining variant is a non-synonymous SNP within a hypothetical protein (DH_02771), unique to DH1916Δ*pyrE* strain #3.

In addition to the 15 variants listed in Table 4.5, a further 23 variants, across all three DH1916Δ*pyrE* strains, were called within a 418 bp intergenic region between loci 3530577 and 3530684 (Table 4.6). This intergenic region lies upstream of DH_03226 (*flgB*) which encodes a flagellar basal body protein and downstream of DH_03227, which is annotated as a miscellaneous RNA GEMM cis-regulatory element and is likely to be the cyclic-di-GMP riboswitch, Cd1. Of these, two are common to all strains, a further thirteen are unique to DH1916Δ*pyrE* strain #2 and the remaining eight are shared only by DH1916Δ*pyrE* strains #1 and #3. These 23 variants all occur in a 418 bp intergenic region

Gene	Description	Position	DH1916 reference	ΔpyrE #1			ΔpyrE #2			ΔpyrE #3			AA
				S	F	C	S	F	C	S	F	C	
00815	Putative membrane protein	967280	A	-	100	56	-	100	41	-	100	34	Asp182fs
01477	Transposase-like protein b	1701479	C	T	81.3	16	T	76	25	T	90.5	21	Val292Ile
01477		1701501	A	T	73.3	15				T	76.2	21	
01477		1701508	A	G	73.3	15				G	72.2	18	Leu282Ser
01477		1701529	C	T	93.3	15	T	100	25	T	100	21	Arg275Lys
IG		2668599	A	G	95.4	22	G	95.7	23	G	94.1	17	
02771	Hypothetical protein	3089504	T							C	100	10	Lys37Arg
IG		3089619	C	T	100	45	T	100	31	T	100	29	
03383	23S rRNA	3692233	A	G	99.2	121	G	98.7	77	G	100	77	
IG		4006823	AC	CA	100	40	CA	100	46	CA	100	27	
IG		4006839	T	-	100	29	-	100	34	-	92.6	27	
IG		4006851	T	A	100	32	A	100	30	A	95	20	
IG		4006867	T	-	100	32	-	96.2	26	-	100	21	
IG		4006886	TC	CT	92.1	38	CT	94.3	26	CT	92	25	
IG		4006927	A	-	96.9	32	-	100	23	-	100	21	

Table 4-5 Resequencing analysis of DH1916Δ*pyrE* strains.

SNPs and other variants identified following resequencing analysis of three DH1916Δ*pyrE* strains (#1, #2 and #3) in the genes identified or in intergenic (IG) regions compared to the draft annotated DH1916 reference sequence. Position, genomic locus of variant; S, variant sequence at given locus; F, frequency of variant in generated reads; C, coverage of reads at given locus. Any resulting amino acid (AA) changes are also listed. Blank spaces denote no change from the DH1916 reference sequence. For clarity, a large number of variants occurring in a 418 bp intergenic region are omitted from this table and are listed in Table 4.6 instead.

Gene	Description	Position	DH1916 reference	ΔpyrE #1			ΔpyrE #2			ΔpyrE #3		
				S	F	C	S	F	C	S	F	C
IG	418 bp intergenic region upstream of DH_03226 and downstream of DH_03227	3530575	A				T	81.3	48			
		3530577	A	G	100	57				G	100	46
		3530579	AA				TT	78.1	64			
		3530587	AA				TC	75.8	62			
		3530590	G	T	85.7	14	T	100	33	T	100	10
		3530591	CC	-	100	59				-	100	48
		3530599	A	T	80.8	73				T	82.8	58
		3530603	A				C	82.6	46			
		3530608	A				C	82.2	45			
		3530614	A				C	81.8	44			
		3530616	A				T	81.4	43			
		3530619	A	T	79.4	63				T	84.9	53
		3530635	C	-	84.4	64				-	83.3	54
		3530641	ATT				-	78.3	46			
		3530646	A	C	98.2	55	C	83.3	12	C	100	48
		3530652	C	-	83.6	67				-	84.2	57
		3530661	G				-	80.9	47			
		3530670	AAT	-	84.4	64				-	80	50
		3530684	A	C	100	54				C	100	39
		3530699	G				T	80.5	41			
		3530705	G				T	80.5	41			
		3530710	G				T	82.5	40			
		3530714	T				A	78	41			

Table 4-6 Additional variants identified following resequencing of DH1916Δ*pyrE*.

SNPs and other variants identified following resequencing analysis of three DH1916Δ*pyrE* strains (#1, #2 and #3) in the 418 bp intergenic (IG) region between DH_03226 (*flgB*) and a small noncoding RNA, DH_03227, compared to the draft annotated DH1916 reference sequence. Position, genomic locus of variant; S, variant sequence at given locus; F, frequency of variant in generated reads; C, coverage of reads at given locus. Blank spaces denote no change from the DH1916 reference sequence. These variants are in addition to those present in Table 4.5.

4.7 Discussion

Whilst the diverse sporulation characteristics observed between isolates belonging to ribtoype-027 (Burns *et al*., 2010) highlights the fact that no one strain can completely describe all members of a ribotype, model strains are still highly valuable for molecular genetics studies and for phenotypic comparisons between ribotypes. Two such model strains are routinely used in our group in which allelic exchange mutagenesis has been established; the 012-ribotype 630Δ*erm* and the 027-ribotype R20291. This work, together with the results presented in Chapter Three, has confirmed previous observations that following purification via repeated washing in distilled water, spores of 630Δ*erm* and R20291 can be seen clumped together in aggregates of varying sizes, whist spores of DH1916 were free from clumping.

Phenotypic characterisation of DH1916 in this study showed it produces similarly high toxin titres as R20291 in relation to 630Δ*erm*, and grows in a similar manner to R20291 in rich media over 24 hours. Further assays confirmed previous observations that DH1916 produces a greater number of spores and develops more heat-resistant CFU over five days than does R20291, however the number of additional spores produced cannot account for this observed increase in heat-resistant CFU development alone. DH1916 displayed an almost 14-fold increase in heat-resistant CFU after five days compared with R20291, yet the counting of spores in these five-day sporulation cultures showed R20291 produced 73.5% of the equivalent DH1916 spore titre. Future studies should aim to identify the cause of the increased heat-resistant CFU of DH1916 relative to R20291. The germination proficiencies of both strains should be assessed using assays which monitor the rate of germination and through plating of spores of each strain onto rich agar plates supplemented with lysozyme. If this still cannot account for the observed differences in heat-resistant CFU development,

together with the previously observed reduction in spore titres of R20291, then the resistance properties of spores of both R20291 and DH1916 to wet heat treatment should be investigated.

A major barrier to establishing genome editing in DH1916 was the low frequency of DNA transfer into this strain via conjugation from *E. coli* CA434 donor cells harbouring pMTL-80000 modular vectors with various Gram-positive replicons. Under such conditions, the highest conjugation frequency observed in this study was 4.94×10^{-8} with the vector pMTL-82151 (pBP1), which was surprising given pB1 is a non-native replicon. Such low conjugation frequency was previously observed for another ribotype-027 strain and was improved by around 3-logs, in a study by Kirk & Fagan (2016), in which *E. coli* CA434 cells harbouring a vector containing the CD6 Gram-positive replicon were used as conjugal donors into R20291 overnight cultures heated for 5 minutes at 52^{o}C. Enhancement of conjugation efficiency following heat shock has also been described for several other organisms including *Campylobacter jejuni* (Zeng *et al*., 2015) and *Corynebacterium glutamicum* (Schafer *et al*., 1994). Heat shock has also made strains of *Staphylococcus aureus* susceptible to phage infection (Asheshov & Jevons, 1963), indicating that heat treatment leads to a transient inactivation of bacterial RM-systems, permitting the entry of foreign DNA. However, this study was unable to replicate the findings of Kirk & Fagan (2016) using DH1916 overnight cultures heated at 52^{o}C for 5 minutes; a protocol which was chosen on the basis of this being the exposure to heat which yielded the greatest increase in conjugation efficiency for R20291. This heat treatment was shown to be sub-lethal to DH1916 vegetative cells, hence killing of the vegetative cells by wet heat cannot explain the lack of expected increases in conjugation efficiency. It is possible that this level of heat-treatment, whilst successful in R20291, is not of sufficient strength to result in a temporal disruption of the restriction machinery in DH1916. Future work could seek to optimise the heat-treatment of DH1916

conjugation recipient, testing different temperatures and incubation lengths. Furthermore, conjugations could be performed with BHI agar, since Kirk & Fagan (2016) also demonstrated an increase in conjugation efficiency using this medium compared with BHIS.

Due to the inability to replicate in DH1916 the increases in conjugation efficiency observed in R20291 following heat-treatment of *C. difficile* conjugation recipient cells, this study sought an alternate approach. Transfer of genetic elements via electroporation into DH1916 was not attempted, despite previous descriptions of successful transformation of *C. difficile* (Ackermann *et al*., 2001), since this method is considerably more complex and lengthy than for conjugative transfer. However, electroporation is widely used in other clostridia, including *C. acetobutylicum* (Oultram *et al*., 1988) and *C. beijerinckii* (Birrer *et al*., 1994), therefore future studies could seek to optimise the protocol for transfer into *C. difficile* via electroporation.

An alternative method to circumvent the bacterial restriction barrier is to supply DNA with methylation patterns which mimic that of the recipient strain. For example, conjugations into *C. acetobutylicum* are routinely performed using DNA which has undergone *in vivo* methylation in an *E. coli* strain harbouring the pAN2 plasmid, containing the ϕ3TI methyltransferase from phage ϕ3TI of *B. subtilis* (Dong *et al*., 2010). A similar concept was employed by this study, which utilised a *dcm- E. coli* strain (NEB sExpress), which does not generate 5-methylcytosine nucleotides, as a conjugal donor strain. This allows circumvention of the two type IV RM-systems found in DH1916, which were originally identified in R20291, since type IV RM-systems cleave methylated DNA. Conjugations into DH1916 from *E. coli* NEB sExpress cells resulted in an increase of up to 42-fold in conjugation frequency, depending upon the Gram-positive replicon of the pMTL-80000 used. However, this analysis of RM-systems present in DH1916 was performed by searching for homologues to RM-

systems previously identified in R20291. Therefore, PacBio WGS of DH1916 with incorporated analysis of the restriction-modification systems present in DH1916 by REBASE would allow any additional RM-systems to be identified, instead of relying on information specific to R20291, and could lead to more DH1916-specific methylation strategies being developed for plasmid transfer into this strain.

Successful plasmid transfer using vectors carrying pseudo-suicide replicons was achieved into DH1916 from NEB sExpress cells and led to the generation of three, independent DH1916Δ*pyrE* strains via allele-coupled exchange. Allelic exchange was subsequently exemplified in these *pyrE*-deficient strains and used to create DH1916 strains with clean, in-frame deletions of *spo0A* or *cspC*. Furthermore, this allelic exchange was highly efficient, with 14/15 (93%) putative *spo0A* mutants and 11/15 (73%) putative *cspC* mutants containing the desired, respective deletions. Given that Spo0A is the master regulator of sporulation and CspC is a germination-specific protease which recognises specific bile acids (Francis, *et al.*, 2013), these generated mutants can be used as controls in subsequent assays of sporulation and germination in DH1916.

The ACE mutagenesis protocol employed in this study to generate the three, *pyrE*-truncated strains required multiple passaging of strains on various media, risking the accumulation of SNPs and other genomic variants with each re-streak. These three DH1916Δ*pyrE* strains underwent whole genome resequencing to determine whether any such variants were accumulated during ACE. Alignment of the resulting reads with the draft annotated DH1916 genome identified a non-synonymous SNP unique to one triplicate, in a hypothetical protein (DH_02771). This results in an amino acid change from lysine to arginine, both of which are charged amino acids, suggesting the impact of this substitution on protein function might be minimal. The impact of this SNP can be assessed in future work by comparing the

phenotypes of DH1916Δ*pyrE* strains #1 and #3, as this SNP is the only source of variation between these two strains. Whilst 16 other variants were also identified, these were present in all three DH1916Δ*pyrE* strains, relative to the DH1916 reference genome, suggesting these were present in the parental DH1916 strain prior to ACE mutagenesis. A series of PCR amplifications flanking these variants followed by Sanger sequencing of the generated PCR products would confirm or reject this hypothesis in each case. The large number of variants detected in the 418 bp intergenic region upstream of *flgB* identified in this study, leading to genetic differences between the DH1916Δ*pyrE* re-sequenced strains, can most likely be explained by a DNA inversion occurring in this region. Recent work by Anjuwon-Foster & Tamayo, (2017) has identified a 154 bp invertible DNA element flanked by inverted repeat sequences and located between Cd1 and the start codon of *flgB* in R20291. This element was named a 'flagellar switch' since when in the orientation described in the published genome transcription of the *flg* operon occurs (*flg* 'on') producing peritrichous flagella, whilst in the inverse orientation (*flg* 'off') the resulting bacteria are non-flagellated and cannot perform swimming motility (Anjuwon-Foster & Tamayo, 2017). All variants identified in this study upstream of DH_03226 (*flgB*) occur within a 140 bp region between loci 3530575 and 3530714, hence an inversion of the 154 bp DNA element identified in R20291 would well explain the resequencing conflicts between DH1916Δ*pyrE* strains. This hypothesis of DNA inversion would explain the two-state differences between the three DH1916Δ*pyrE* strains, but does not explain why neither state fully matches the sequence of the DH1916 reference sequence. Future investigations should PCR amplify this intergenic region upstream of *flgB* from the parental DH1916 reference strain and the three *pyrE*-minus derivatives, and confirm the sequences of each strain via Sanger sequencing. Furthermore, if DNA inversions at this locus are observed, swimming motility assays could be performed with these four strains to confirm that the orientation of this 'flagellar switch' also controls flagellar formation in DH1916, as with R20291.

4.8 Summary

- DH1916 produces a statistically significant, higher number of spores over five days in rich media than does R20291.
- DH1916 spores washed to purity in distilled water do not clump together to form aggregates of various sizes.
- Very low conjugation efficiencies into DH1916 can be increased by using a *dcm- E. coli* conjugation donor strain to circumvent Type IV restriction systems, but no increase in conjugation frequency was observed following heat treatment of the *C. difficile* recipient
- Three, independent DH1916Δ*pyrE* strains were produced and used to establish allelic exchange mutagenesis. This resulted in the creation of sporulation and germination control strains.
- Whole genome resequencing of the DH1916Δ*pyrE* strains suggests few, if any, variants were incorporated into these strains during ACE.

Chapter Five

Establishing CRISPR/Cas9 genome editing in *C. difficile*

5 Establishing CRISPR/Cas9 genome editing in *C. difficile*

5.1 Introduction

At the beginning of this project, two methods existed for precise genome editing in clostridia; insertional mutagenesis, using retargeted ClosTron plasmids to insert an antibiotic resistance gene at a chosen site, and allelic exchange, the *pyrE*-mediated form of which utilises homologous recombination from a vector carrying a heterologous *pyrE* allele as a negative/counter selection marker into mutant strains auxotrophic for uracil. Both methods have been used to greatly enhance our understanding of *C. difficile* via studies identifying the role of numerous genes involved in sporulation, germination and virulence. Both insertional mutagenesis and allelic exchange methodologies have advantageous features but neither is appropriate for high throughput genome editing. ClosTron insertions, whilst rapid, are unfavourable due to the risk of polar effects downstream of the insertion site. In contrast, allelic exchange requires numerous passages on rich and minimal media, which risks the accumulation of SNPs and limits the speed at which mutants can be obtained. A preferable system is one which combines the speed of mutant generation using ClosTron technology with the clean and precise genome editing of allelic exchange mutagenesis.

5.1.1 CRISPR/Cas9 genome editing in clostridia

CRISPR/Cas9 genome editing, outlined in section 1.5, has the potential to quickly generate in-frame deletion mutants without requiring multiple passaging. In prokaryotes lacking non-homologous end joining (NHEJ) to repair double-strand breaks, this mutagenesis method

relies on the nucleolytic activity of Cas9 to counter-select against cells which have not undergone chromosomal homologous recombination with the plasmid-borne editing template. Thus, double crossover mutants are obtained in a single step.

Furthermore, CRISPR/Cas9 genome editing has recently been described for a wide range of bacteria including three biotechnologically relevant clostridia; *Clostridium beijerinckii* (Wang *et al*., 2015), *Clostridium cellulolyticum* (Xu *et al*., 2015) and *Clostridium ljungdahlii* (Huang *et al*., 2016). In each of these three previous studies in Clostridia the authors used the *Streptococcus pyogenes* Cas9 in either its native form, or a nickase derivative (Cas9n) which has only one nucleolytic domain and generates a single strand cut at the target site. Cas9n was used to generate a deletion of *pyrF* in *Clostridium cellulolytcium* as the native *S. pyogenes* Cas9 was found to be too lethal under the conditions tested (Xu *et al*., 2015). All three studies used a single vector harbouring the requisite components for CRISPR/Cas9 genome editing; a *cas9* gene and a suitable single guide RNA component, each under the control of individual promoters, and an editing template for homologous recombination with the host chromosome. The single guide RNA (sgRNA) component contains a 20 bp protospacer sequence corresponding to the sequence immediately upstream of the protospacer-associated motif (PAM) and guides the Cas9 protein to the chromosomal target.

5.1.2 Aims of this study

CRISPR/Cas9 genome editing has been reported in a number of bacterial species, including several clostridia. This study aims to exemplify this mutagenesis method in *C. difficile* by generating an erythromycin sensitive 630 strain which lacks the SNPs accumulated during repeated passage present in the current laboratory strains 630*Δerm* and 630E.

5.2 Preliminary testing of CRISPR/Cas9 systems in *C. difficile*

The starting point for this study was the pMTL-Cas9-Caethg-*pyrE*-1 vector developed by Pete Rowe in collaboration with Huang *et al*. (2016). This vector, shown in Figure 5.1, contains the requisite components for CRISPR/Cas9 genome editing cloned between the *Not*I and *Asc*I restriction sites within pMTL83151, and was used to generate *pyrE* mutants of *Clostridium autoethanogenum* (Pete Rowe, unpublished). The *cas9* gene present is a variant of the *cas9* found in *S. pyogenes* in that it contains an additional adenosine nucleotide at position 137, generating a frame-shift mutation. *In silico* analysis suggests the transcription of this *cas9* occurs at a downstream start codon resulting in a truncated product lacking 267 bp from the 5' end of the *S. pyogenes cas9* gene (Chris Humphreys, personal communication). Determining the effects of this frame-shift mutation on the nucleolytic activity of Cas9 was beyond the scope of this project, which simply sought a functioning Cas9 protein.

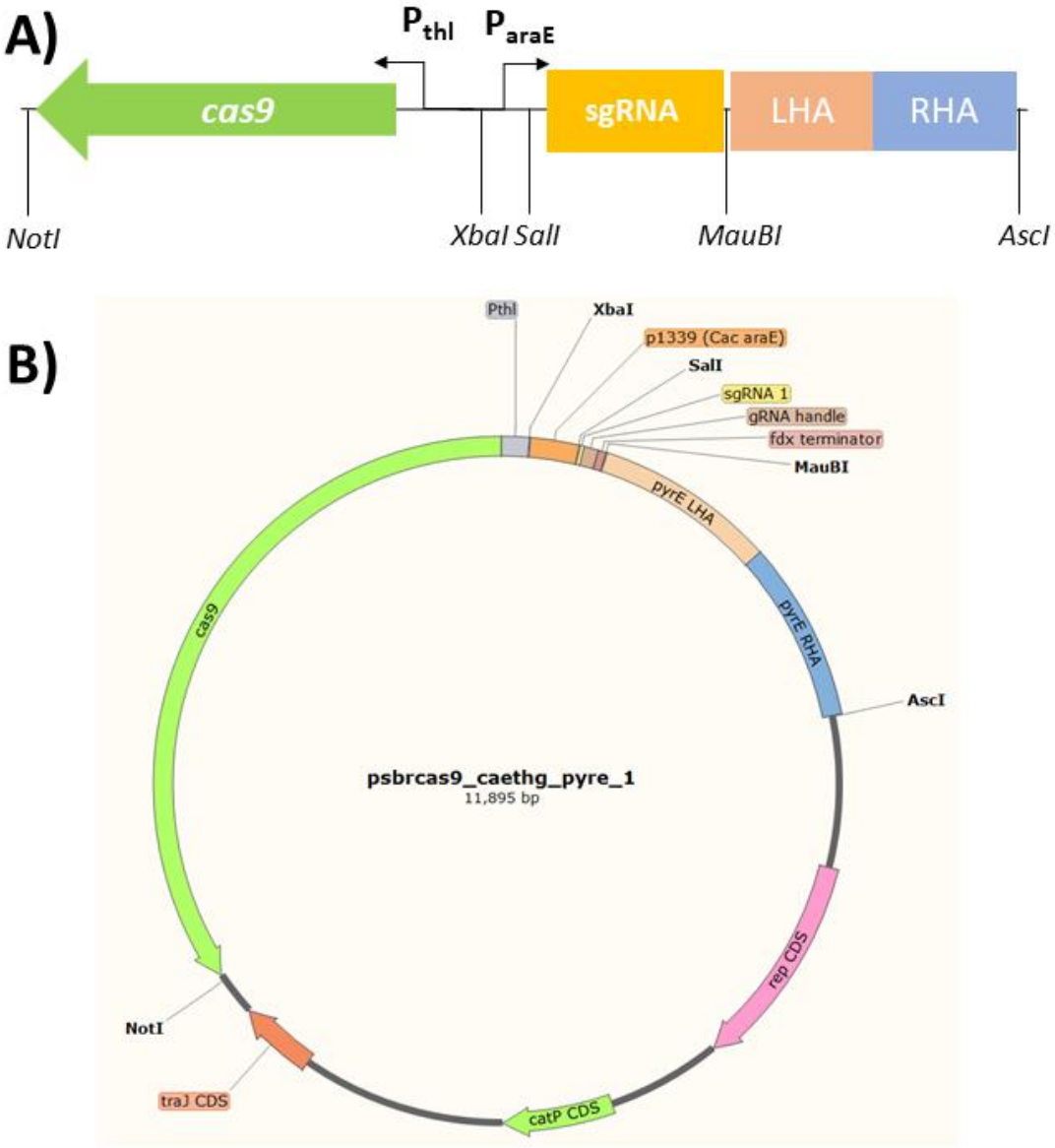


Figure 5-1 Diagrammatic representation of the construction of the CRISPR/Cas9 vectors used in this study.

Diagrams show the cloning strategy utilised to insert the CRISPR/Cas9 genome editing machinery (A) between *Not*I and *Asc*I restriction sites in the modular plasmid pMTL83151 resulting in the pSBRCas9_caethg_*pyrE*1 vector (B). *cas9* refers to the frame-shift version of the native *S. pyogenes cas9* gene; sgRNA, single guide RNA component consisting of 20 bp seed and 83 bp handle regions; P_{thl}, thiolase promoter from *C. acetobutylicum*; P_{araE}, CAC1339 promoter from *C. acetobutylicum*; the editing template is generated by SOE PCR of left homology arm (LHA) and right homology arm (RHA) DNA fragments upstream and downstream of the chromosomal target region.

5.2.1 CRISPR/Cas9 mediated mutagenesis of *spo0A* in *C. difficile* 630

As a quick test of the CRISPR/Cas9 system present in pMTL-Cas9_Caethg_*pyrE*_1, the *spo0A* gene of *C. difficile* 630 was to be deleted. The *spo0A* gene was chosen as it has previously been deleted in 630*Δerm* using allelic exchange (Ng *et al*., 2013), the *spo0A* mutant gives a clear phenotype and the necessary editing template was present in pMTL_YN3_*spo0A*.

5.2.1.1 Construction of CRISPR/Cas9 vectors targeting C. difficile 630 spo0A

This *spo0A* knock-out cassette was PCR amplified from pMTL_YN3_*spo0A* using primers RNg630spo_MauBI_F and RNg630spo_AscI_R, and the purified PCR products digested with *Mau*BI and *Asc*I restriction enzymes. The ligation of these *Mau*BI/*Asc*I double digest products into similarly digested pMTL-Cas9_Caethg_*pyrE*_1 proved troublesome, hence an alternate cloning strategy was devised. This method involved SOE PCR generation of the sgRNA component with the flanking reverse primer sgRNA_2.6R containing flanking *Asi*SI and *Mau*BI restriction sites. For the sgRNA components, the Benchling Guide Design tool was used to search for 5'-N_{20}NGG-3' sequences within the *spo0A* gene (CD_12140) of *C. difficile* 630, corresponding to the *S. pyogenes* Cas9 PAM recognition sequence. Three 20 nt sgRNA seeds, upstream of 5'-NGG-3' sequences, from the 35 identified within the *spo0A* gene were selected (Table 5.1), according to the criteria stated in Materials and Methods. These sgRNA seed sequences were incorporated into overlapping SOE PCR primers and DNA fragments containing the sgRNA promoter, P_{araE}, each sgRNA 20 nt seed sequence and sgRNA handle sequence were generated by SOE PCR as outlined in Materials and Methods.

Name	Sequence	PAM	Position	Strand	On score	Off score
***Spo0A*1**	ACTCAAAGCGCAATAAATCT	AGG	312	+	51.1	100
***Spo0A*5**	ATAATAGTACTATCAGCAGT	AGG	276	+	63.4	97.7
***Spo0A*6**	ACACCTAATCCATCTAGATG	TGG	199	-	61.5	98.4

Table 5-1 sgRNA seed sequences identified within CD12140 (*spo0A*).

sgRNA seeds and the corresponding PAM sequences identified using the Benchling sgRNA design tool at the indicated positions within *C. difficile* 630 CD12140. On- and Off-target scores refer to rules set by Doench *et al*., (2016).

Each of the three sgRNA containing fragments were then cloned into pMTL-Cas9_Caethg_*pyrE*_1_*spo0A* via restriction digest at *Xba*I and *Mau*BI sites, generating vectors pPSI_cas_001 (spo0A1), pPSI_cas_005 (spo0A5) and pPSI_cas_006 (spo0A6). Sanger sequencing with primer sgRNA_sF1 was performed to confirm the correct insertion of each fragment. The *spo0A* editing template from pMTL-YN3-spo0A was then re-amplified using primers RNg630spo_AsiSI_F and RNg630spo_AscI_R, the products of which were cloned into each of pPSI_cas_001, pPSI_cas_005 and pPSI_cas_006 via restriction digestion with *Asi*SI and *Asc*I enzymes and subsequent ligation reactions. Correct insertion of the *spo0A* editing templates was confirmed in each case via Sanger sequencing and the generated plasmids named pPSI_cas_011, pPSI_cas_015 and pPSI_cas_016 corresponding to sgRNAs spo0A1, spo0A5 and spo0A6 respectively (Figure 5.2).

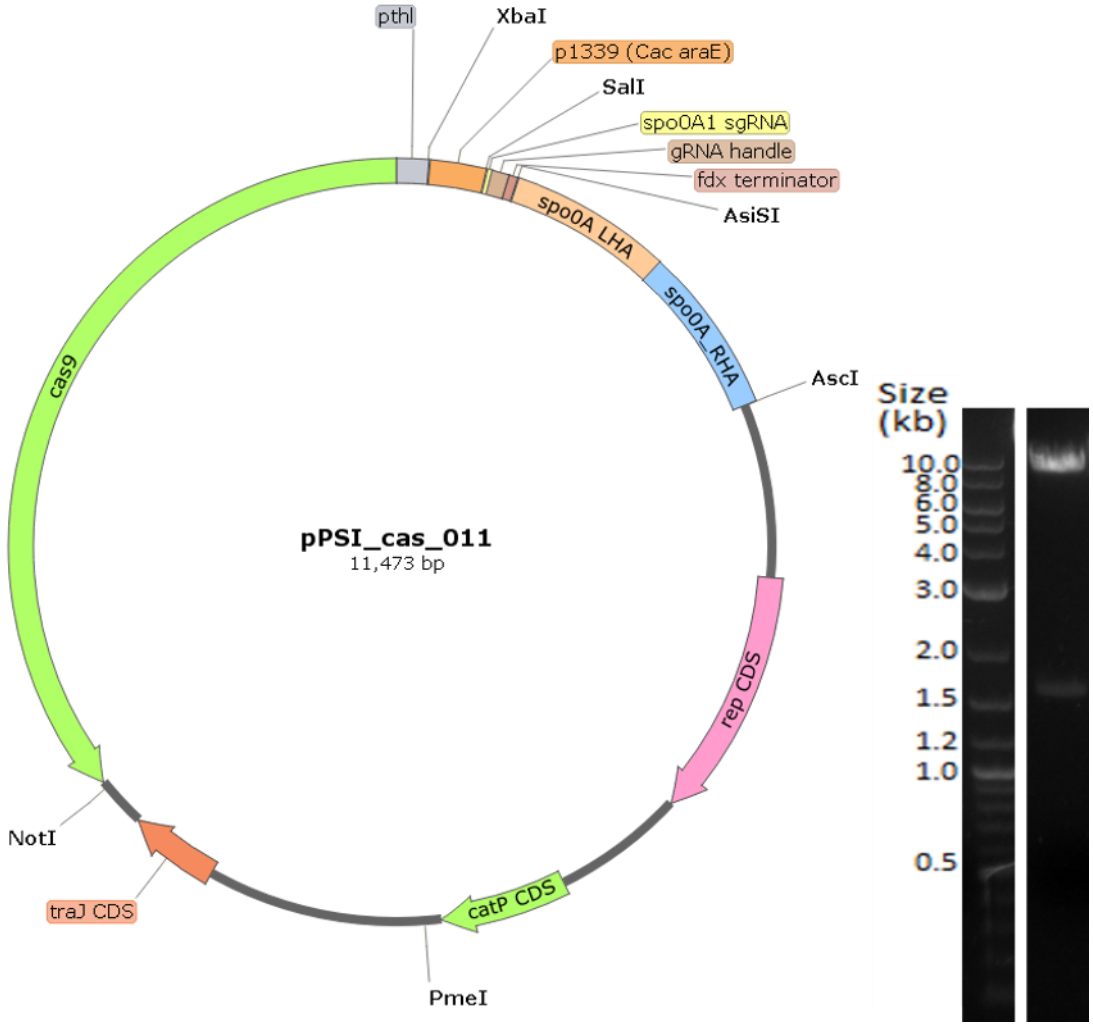


Figure 5-2 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_011.

Left: Plasmid map for pPSI_cas_011, a CRISPR/Cas9 vector targeting *C. difficile* 630 *spo0A.* Features are as for pSBRCas9_caethg_*pyrE*1 except for the replacement of the *pyrE* editing template with a *spo0A* editing template and the replacement of the *Mau*BI restriction site for an *Asi*SI site. Right: Agarose gel electrophoresis following restriction digestion with *Asi*SI and *Asc*I of pPSI_cas_011.

5.2.1.2 Generation of C. difficile 630Δspo0A using CRISPR/Cas9

The three CRISPR/Cas9 vectors targeting *spo0A* in *C. difficile* 630, pPSI_cas_011, pPSI_cas_015 and pPSI_cas_016, were transformed into *E. coli* CA434 and the resulting strains used as conjugation donors for transferring these vectors into *C. difficile* 630.

Thiamphenicol resistant colonies appeared 48-72 hours after plating onto BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol and were re-streaked onto the same media to confirm the presence of the respective CRISPR/Cas9 vector. Over 50 subsequent thiamphenicol resistant colonies were screened via colony PCR using primers spo0A_sF1 and spo0A_sR1, none of which possessed the desired deletion of *spo0A*. This indicated a lack of function of the CRISPR/Cas9 system, since cells which had not undergone homologous recombination with the editing template were not killed by Cas9-mediated double strand breaks.

5.2.2 Replacement of the P_{araE} sgRNA promoter

Troubleshooting the inability to generate a *C. difficile* 630 *spo0A* mutant using CRISPR/Cas9 led to the hypothesis that the sgRNA promoter (P_{araE}) used may not be functional in *C. difficile*. This was on the basis that; (i) our variant Cas9 had been successfully used to generate mutants in other clostridia, (ii) the promoter driving *cas9* (P_{thl}) had previously been shown to work in *C. difficile*, (iii) three independent sgRNAs were utilised and (iv) the editing template was previously used to generate 630*Δerm*Δ*spo0A* via allelic exchange. Accordingly, two promoters known to function in *C. difficile*, the native Toxin B (P_{tcdB}) promoter and the *C. sporogenes* ferrodoxin (P_{fdx}) promoter, were PCR amplified using primer pairs PtcdB_*XbaI* and PtcdB_*Sal*I, and Pfdx_*Xba*I and Pfdx_*Sal*I, respectively. These PCR products underwent restriction digest with enzymes *Xba*I and *Sal*I and ligation reactions into similarly digested pPSI_cas_011 were performed. Despite repeated attempts, the replacement of P_{araE} with P_{fdx} could not be achieved. The replacement of P_{araE} with P_{tcdB} was achieved however and was confirmed via Sanger sequencing. The resulting vector containing the P_{tcdB} promoter driving sgRNA expression was termed pPSI_cas_111 (Figure 5.3). The P_{tcdB} promoter is an attractive choice as this promoter is recognised by TcdR, a sigma factor which is only found in several

toxin-producing clostridial species, therefore minimising expression of the sgRNA component in *E. coli* which lacks an analogous sigma factor.

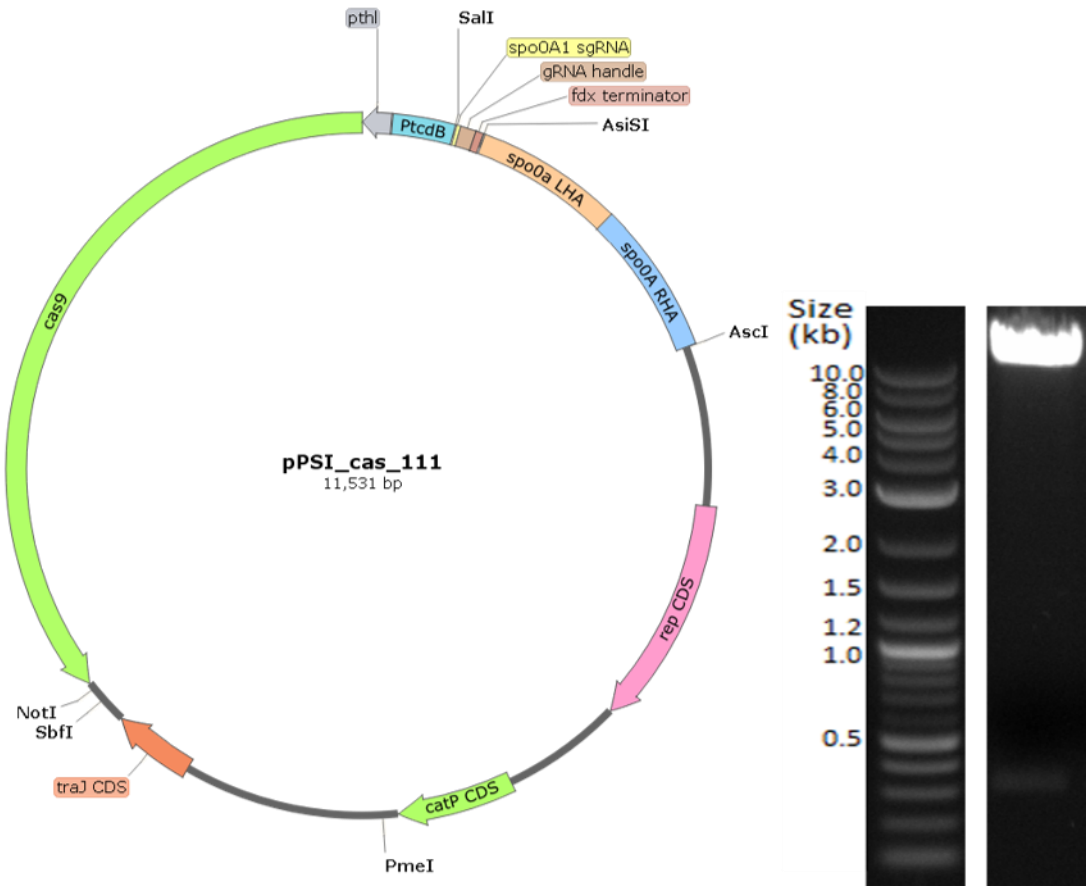


Figure 5-3 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_111.

Left: Plasmid map for pPSI_cas_111, a CRISPR/Cas9 vector targeting *C. difficile* 630 *spo0A.* Features are as for pPSI_cas_011 except for the replacement of the CAC_1339 P_{araE} promoter with that of P_{tcdB}, the promoter of the *C. difficile* toxin B gene, *tcdB*. Right: Agarose gel electrophoresis following restriction digestion with *Sal*I and *Xba*I of pPSI_cas_111.

Conjugations into *C. difficile* 630 from *E. coli* CA434 donor cells harbouring pPSI_cas_111 were performed. Again, thiamphenicol resistant transconjugant colonies appeared after 48-72 hours but again these colonies all produced wild-type sized bands following colony PCR screens as before. Sanger sequencing of pPSI_cas_111, extracted from the conjugation

donor cells, was performed to ascertain whether any SNPs were present which might explain the continued lack of functioning of the CRISPR/Cas9 system. A single SNP within the editing template of vector pPSI_cas_111 was subsequently identified. Consequently, attempts to generate a *spo0A* mutant using CRISPR/Cas9 were halted and focus turned to a different target to attempt to exemplify this method of mutagenesis in *C. difficile*.

5.3 Generation of *C. difficileΔerm** using CRISPR/Cas9

5.3.1 Erythromycin sensitive derivatives of *C. difficile* 630

C. difficile strain 630 was isolated from a CDI patient with severe pseudomembranous colitis during a notable outbreak in Zurich, Switzerland in 1982 (Wust *et al*., 1982). Early studies of *C. difficile* used this strain, belonging to the PCR-ribotype 012, and this was the first *C. difficile* strain to have its genome sequenced (Sebaihia *et al*., 2006). Analysis of the genome of strain 630 by Sebaihia *et al*. (2006) showed that 11% of the 4.29 Mbp genome consists of mobile genetic elements, including seven conjugative transposons, a mobilizable transposon (Tn *5398*), two prophages and other mobile elements. A number of these mobile elements contain accessory genes responsible for the high level of resistance to antibiotics exhibited by strain 630. For example, the Tn *5397* conjugative transposon contains *tet(M)* which confers resistance to tetracycline whilst the mobilizable transposon Tn *5398* contains two sequence identical *ermB* genes, termed *erm1(B)* and *erm2(B)*, which confer resistance to erythromycin. Other antibiotics which strain 630 displays resistance against include clindamycin, gentamycin and ampicillin. The transfer of erythromycin resistance from strain 630 to susceptible strains of *C. difficile* (Wust & Hardegger, 1983) and other Gram-positive bacteria, including *Staphylococcus aureus* (Hachler *et al*. 1987) and *B. subtilis* (Mullany *et al*. 1995), has been reported.

Erythromycin is a member of the MLS (macrolides, lincosamide and streptogramins) antibiotic superfamily and irreversibly binds to the 50S ribosomal subunit to inhibit protein synthesis. Resistance to erythromycin most commonly occurs via N6-dimethylation of a specific adenine residue within 23S rRNA by an *erm* gene encoded rRNA methyltransferase. Homologues of the two *erm(B)* genes present in *C. difficile* 630 have been found in a wide range of bacterial genera, in contrast to alleles from other *erm* classes which are generally limited to one genus (Farrow, *et al*., 2001). The structure of Tn *5398*, including other ORFs present, is shown in Figure 5.4.

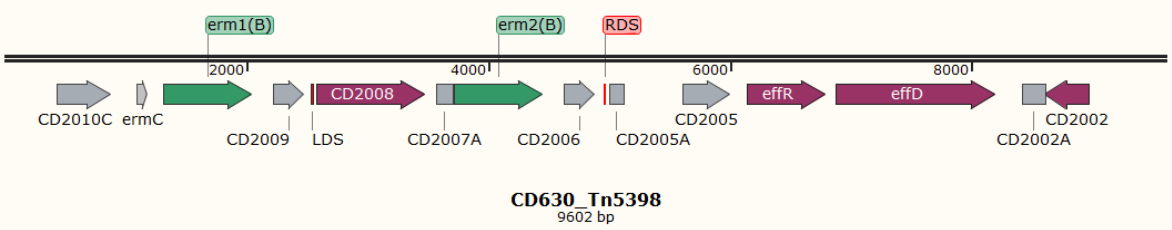


Figure 5-4 Diagrammatic representation of the ORFs present on the *C. difficile* strain 630 mobilizable transposon Tn *5398*.

The 2.4 kb region deleted from 630Δ*erm* and 630E is flanked by the left- (LDS) and right-deletion sites (RDS). Sequence identical *erm(B)* genes are shown in green. *ermC* encodes a 23S rRNA methylase leader peptide; CD2008, a putative hydrolase; *effD*, a putative antiporter; *effR*, a putative regulator; CD2002, a transcriptional regulator. All other annotations denote pseudogenes or gene remnants.

Prior to the development of allelic exchange, the only method to precisely edit the genome of *C. difficile* was the insertion of an antibiotic resistance marker at a given site via ClosTron mutagenesis. The use of ClosTron vectors containing the *ermB* marker first required the generation of an erythromycin sensitive *C. difficile* 630 strain. Two such strains were independently generated by the Mullany Laboratory (UCL, London, UK) and Rood Laboratory (Monash, Australia). These erythromycin sensitive derivatives of *C. difficile* 630 were both generated through repeated subculture of the strain in non-selective media. The Mullany

laboratory performed 30 such subcultures to isolate 630*Δerm* (Hussain *et al.*, 2005), whist the Rood laboratory performed an undisclosed number of subcultures to generate JIR8094, also referred to as 630E (O'Connor *et al*., 2006).

Sanger sequencing of 630*Δerm* and 630E revealed both had lost an identical 2.4 kb region within the mobilizable transposon Tn *5398*, shown in Figure 5.1. Interestingly, only *erm2(B*) is present within the 2.4 kb region deleted from both 630*Δerm* and 630E, which suggested this gene alone is responsible for conferring erythromycin resistance to strain 630. However, subsequent results showed that 630*Δerm* is capable of reverting to an erythromycin resistant phenotype with no detected change in Tn *5398* sequence, suggesting *erm1(B)* is functional in some instances. The authors claim the observed level of reversion to erythromycin resistance of 2.79×10^{-8} is not a barrier for use of erythromycin as a selection marker, since this level of reversion is lower than conjugation efficiency into 630*Δerm* (Hussain *et al.*, 2005).

Both 630*Δerm* and 630E strains were subsequently used to investigate *C. difficile* via reverse genetics approaches utilising insertional mutagenesis with the *ermB* selectable marker. However, discrepancies between studies soon emerged, including two directly contradictory reports on the role of TcdB in CDI. Lyras *et al.* (2009) generated isogenic *tcdA* and *tcdB* mutants of 630E and found the strain lacking a functional TcdB to be significantly less virulent than the wild type or *tcdA* mutant in the hamster model, leading them to the conclusion that Toxin B, and not Toxin A, is essential for virulence of *C. difficile*. By contrast, Kuehne *et al*., (2010) created isogenic *tcdA* and *tcdB* mutants, and a double mutant, of strain 630*Δerm*, and observed similar levels of virulence *in vitro* and in the hamster infection model for both isogenic mutants lacking Toxin A or Toxin B and the wild type. Furthermore, the

tcdA^{-}*tcdB*$^{-}$ double mutant was completely attenuated for virulence. This study concluded that there is a role for both toxins A and B in CDI.

These two contradictory accounts of the role of Toxin B in CDI were investigated by Collery *et al*. (2016) who hypothesised that the observed differences may result from secondary mutations caused by repeated sub-culture during the generation of the two erythromycin sensitive derivatives of strain 630. Whole genome sequencing revealed 630E to have 9 unique SNPs (7 of which are non-synonymous) and an inversion when compared to strains 630 and 630*Δerm*, whilst 7 unique SNPs (6 of which are non-synonymous) and an insertion were found in 630*Δerm* when compared to strains 630 and 630E. Phenotypic comparisons of all three 630 strains showed that 630*Δerm* behaves in a more similar manner to strain 630 than does 630E. Compared to strain 630, 630E grew slower and to a lower cell density, produced lower levels of Toxins A and B, was less virulent in the hamster infection model and yielded lower spore titres after 120 hours (Collery *et al*., 2016).

5.3.1 Construction of vectors targeting CD2008 in *C. difficile* 630

Design of the editing template for the generation of a new erythromycin sensitive strain 630 was performed such that both sequence identical *erm(B)* genes would be deleted from Tn *5398*. This was with the aim of overcoming the reversion to erythromycin resistance exhibited by strain 630*Δerm*. Hence, the 2.4 kb region deleted from both 630*Δerm* and 630E was extended to include the *erm1(B)* gene. An editing template consisting of 1 kb up- and down-stream of this *in silico* deletion was generated by SOE PCR, as detailed in Materials and Methods, and cloned into pPSI_cas_111 via restriction digestion with *Asc*I and *Asi*SI enzymes followed by ligation reactions. The generated vector, pPSI_cas_120, was Sanger sequenced to confirm the correct insertion of the editing template.

Of the forty-two 5'-N_{20}NGG-3' sequences identified within the CD2008 gene of *C. difficile* 630 using the Benchling sgRNA design tool, three were chosen for use (Table 5.2). The N20 protospacer sequences of each were separately inserted into pPSI_cas_120 via *Sal*I restriction digestion followed by HiFi assembly as described in Materials and Methods. Successful replacement of the *spo0A*1 seed sequence with those of CD2008A, CD2008B and CD2008C was confirmed by Sanger sequencing and the resulting vectors named pPSI_cas_121, pPSI_cas_122 and pPSI_cas_123 respectively.

Name	Sequence	PAM	Position	Strand	On score	Off score
CD2008A	AACATCAAATCAAACGTGCC	AGG	361	+	49.1	100
CD2008B	TTGCAAGAACAGTTTAACCC	TGG	627	-	59.2	99.4
CD2008C	AATGGCATTACAGAACACAA	AGG	804	+	74.7	100

Table 5-2 sgRNA seed sequences identified within CD2008.

sgRNA seeds and the corresponding PAM sequences identified using the Benchling sgRNA design tool at the indicated positions within *C. difficile* 630 CD2008. On- and Off-target scores refer to rules set by Doench *et al*., (2016).

5.3.2 Deletion of both *erm(B)* genes from *C. difficile* 630

CRISPR/Cas9 vectors pPSI_cas_121, pPSI_cas_122 and pPSI_cas_123 were conjugated into *C. difficile* 630 from *E. coli* CA434 donor cells. Thiamphenicol-resistant, transconjugant colonies appeared 48-72 hours after plating of the conjugal mixture onto BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol and were re-streaked onto the same media to confirm the presence of the respective plasmids. Twenty-five subsequent thiamphenicol resistant colonies were screened via colony PCR using primers Tn*5398*_sF1 and Tn*5398*_sR1, to confirm the desired 3.648 kb deletion from within Tn *5398* had occurred (Figure 5.5).

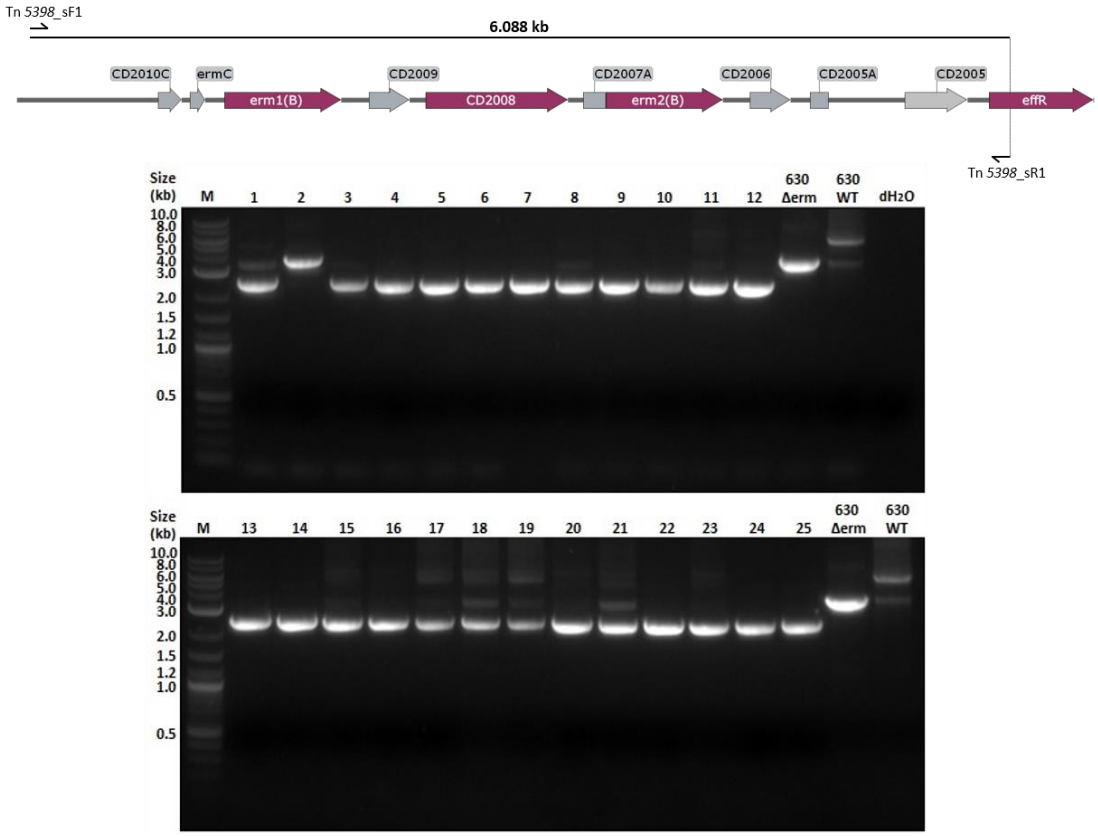


Figure 5-1 Colony PCR screening for *C. difficile* 630Δ*erm mutants following CRISPR/Cas9 mutagenesis.**

Top: Diagrammatic representation of PCR primer binding sites. Bottom: Agarose gel electrophoresis; M: DNA Marker (2-log ladder; NEB); Lanes 1-25: 25 thiamphenicol resistant colonies following conjugation of pPSI_cas_121 (lanes 1-8), pPSI_cas_122 (lanes 9-14) and pPSI_cas_123 (lanes 15-25) into *C. difficile* 630. WT: 630 gDNA, 630*Δerm*: 630*Δerm* gDNA; dH_2O: distilled water.

Of the 25 thiamphenicol resistant colonies screened, 24 (96%) produced the 2.44 kb sized band expected for deletion mutants, either alone (pure mutant) or together with the 6.088 kb sized band expected for wild-type colonies. The efficiency of each sgRNA in generating the desired deletion was determined (Table 5.3). CRISPR/Cas9 mutagenesis was successfully performed with all three guides tested, of which CD2008B was deemed the most efficient by means of generating the highest proportion of pure mutants. To confirm the 2.44 kb sized bands corresponded to the desired deletion within Tn *5398*, DNA fragments from lanes 5, 9 and 24 (Figure 5.5) were purified and Sanger sequenced with primers Tn*5398*_sF1,

Tn*5398*_sF2 and Tn*5398*_sR1. In all three cases, the expected deletion of *erm1(B)* and *erm2(B)* was confirmed. The thiamphenicol resistant colonies which corresponded to these PCR products were grown overnight in BHIS broth to facilitate the loss of their respective CRISPR/Cas9 plasmids. Serially diluted overnight cultures were spread onto BHIS agar and resulting colonies patch plated onto BHIS agar with and without thiamphenicol to screen for plasmid loss. A single thiamphenicol sensitive colony from each independently obtained mutant was carried forward and these triplicate strains lacking both *erm(B)* genes were named *C. difficile* 630Δ*erm** 1-3.

sgRNA	Wild-type	Mutant	Both
CD2008A	**1/8**	**4/8**	**3/8**
CD2008B	**0/6**	**5/6**	**1/6**
CD2008C	**0/11**	**4/11**	**7/11**
Total	**1/25 (4%)**	**13/25 (52%)**	**11/25 (44%)**

Table 5-3 Efficiency of each sgRNA in the generation of *C. difficile* 630*Δerm via CRISPR/Cas9 mutagenesis.**

Numbers indicate the proportion of pure mutant, wild-type or strains displaying a mixture of both alleles present following colony PCR screening of thiamphenicol resistant colonies resulting from conjugations into *C. difficile* 630 of pPSI_cas_121 (CD2008A), pPSI_cas_122 (CD2008B) and pPSI_cas_123 (CD2008C).

5.4 Phenotypic characterisation of *C. difficile* 630Δ*erm**

Several phenotypic assays were performed to compare the newly generated *C. difficileΔerm** strains to the parental strain, *C. difficile* 630 and the previous erythromycin sensitive strain, *C. difficile* 630*Δerm*. These experiments were performed to ensure that the absence of *erm1(B)* from *C. difficileΔerm** does not affect the key traits of growth, sporulation, cytotoxicity or cell motility.

5.4.1 Growth in BHIS broth

The growth of each 630 strain in BHIS broth was assessed using changes in measured OD_{600} over 24 hours (Figure 5.6). *C. difficile* 630*Δerm** growth was indistinguishable from that of the parental strain 630 and also 630*Δerm*, indicating there were no growth defects resulting from the deletion of *erm1(B)*.

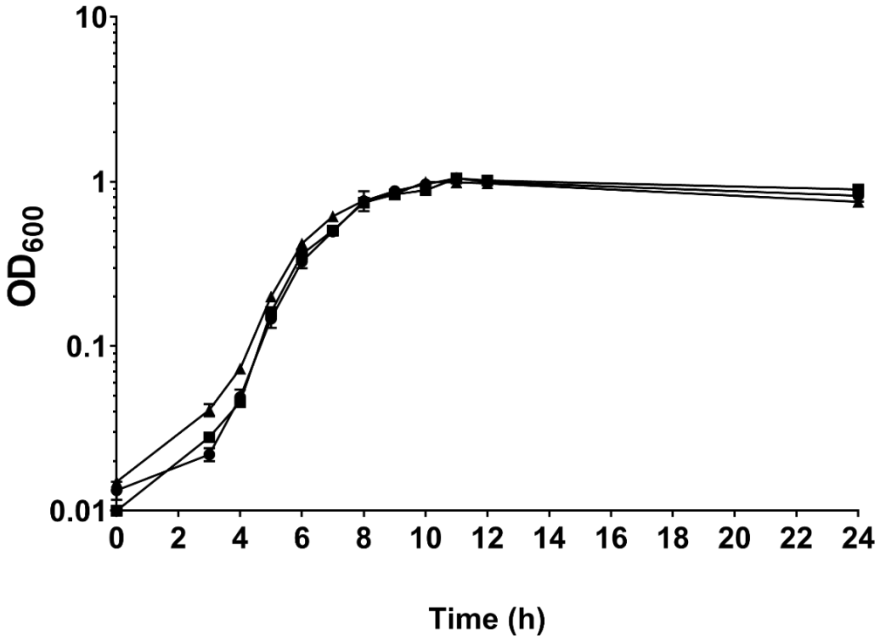


Figure 5-2 Growth of *C. difficile* 630 strains in BHIS broth over 24 hours, indicated by changes in measured OD_{600}.

● *C. difficile* 630; ■, 630*Δerm*; ▲, 630*Δerm**. Symbols represent the averages of three independent experiments. Error bars indicate the standard errors of the mean. Statistical significance determined using two-way ANOVA with Sidak's multiple comparisons test; *P*-value=0.1617 for variance between strains.

5.4.2 Colony formation after heat-treatment

The ability of each 630 strain to form heat-resistant CFU over five days on BHIS agar supplemented with taurocholate was determined (Figure 5.7). After five days, the mean heat-resistant CFU observed were $3.66x10^6$ for 630*Δerm*, $1.49x10^6$ for 630 (CRG 856) and

$7.78x10^5$ for 630*Δerm**. This indicates that the additional loss of *erm1(B)* did not have a substantial effect upon the proficiency of the strain to develop heat resistant CFU.

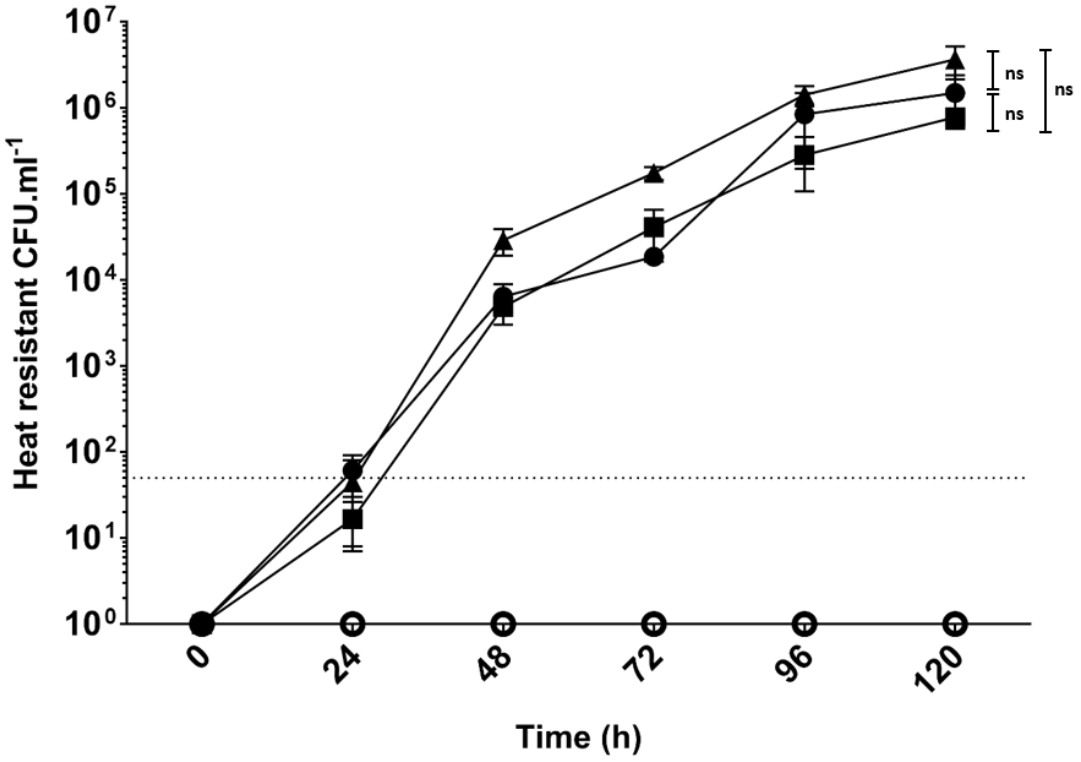


Figure 5-3 Development of heat resistant CFU of *C. difficile* 630 strains over five days.

●, *C. difficile* 630 (CRG 856); ▲, *C. difficile* 630*Δerm*; ■, *C. difficile* 630*Δerm**; **O**, *C. difficile* 630*ΔermΔspo0A*. Symbols represent the mean values from three independent experiments. Error bars indicate standard errors of the means. The dotted line at 50 $CFU.ml^{-1}$ denotes the detection limit for the assay. Statistical significance determined using one-way ANOVA with Dunnett's multiple comparisons test. Not significant (ns) denotes a *P*-value >0.05.

5.4.3 Cytotoxicity against Vero cells

In vitro cell cytotoxicity assays using Vero (African green monkey kidney) cells were performed to determine whether the *erm1(B)* deletion had any impact on the cytotoxic activity of 630Δ*erm** (Figure 5.8). Identical end-point titres were observed for all 630 strains

tested, suggesting no association between the *erm(B)* genes and *in vitro* cytotoxicity. These values were considerably lower than for the ribotype-027 strain, R20291, as expected.

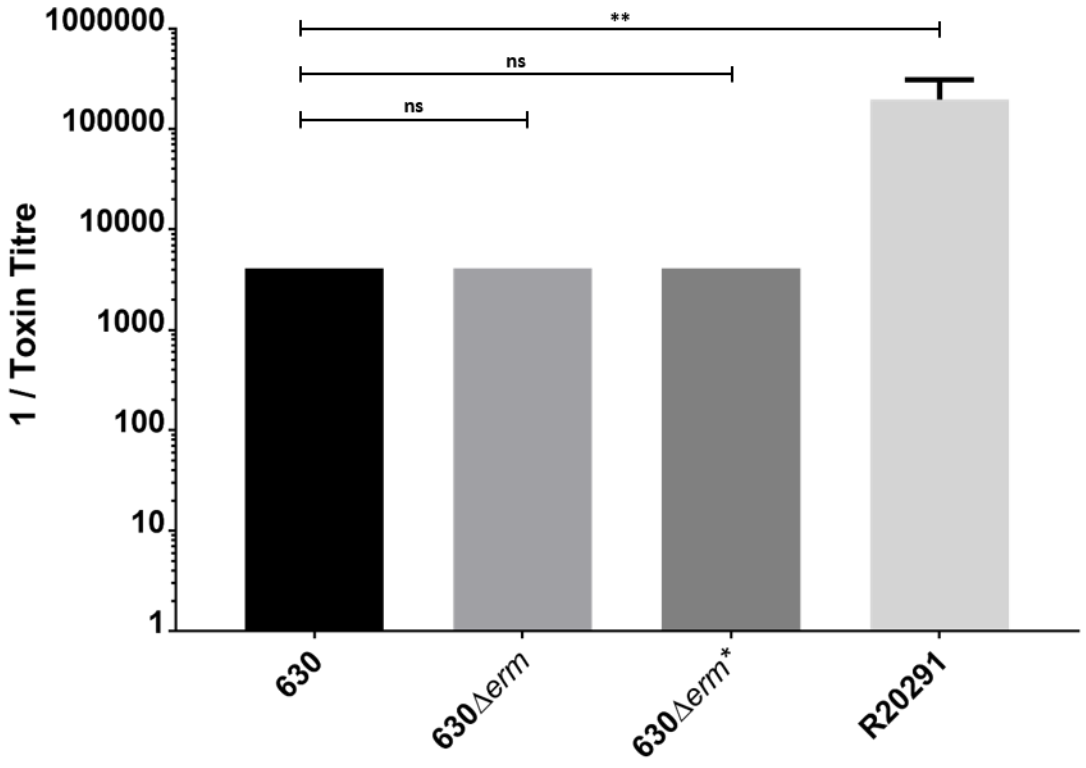


Figure 5-4 *In vitro* cell cytotoxicity of *C. difficile* strains against Vero cell cultures.

C. difficile supernatants were adjusted for OD_{600} and four-fold serial dilutions incubated with cultured Vero cells for 24 hours. Subsequent rounding of Vero cells was observed under light microscopy and end-point titres determined. Bars represent the mean end-point titres from three independent reactions. Error bars indicate the standard errors of the means. Statistical significance determined using multiple unpaired t-tests. Not significant (ns) denotes a *P*-value >0.05, two asterisks (**) denotes a *P*-value <0.01.

5.4.4 Cell Motility

The motility of each 630 strain in this study was determined by spotting overnight cultures of each strain onto 2YTG media containing 0.24% Gelzan (Figure 5.9). Strains 630*Δerm*, 630 (CRG 856) and 630*Δerm** all exhibited similar patterns of motility compared with that of

630E, and all formed pseudopod-like structures characteristic of bacterial swarming motility. Conversely, the non-motile strain 630E did not form such structures over the course of this assay.

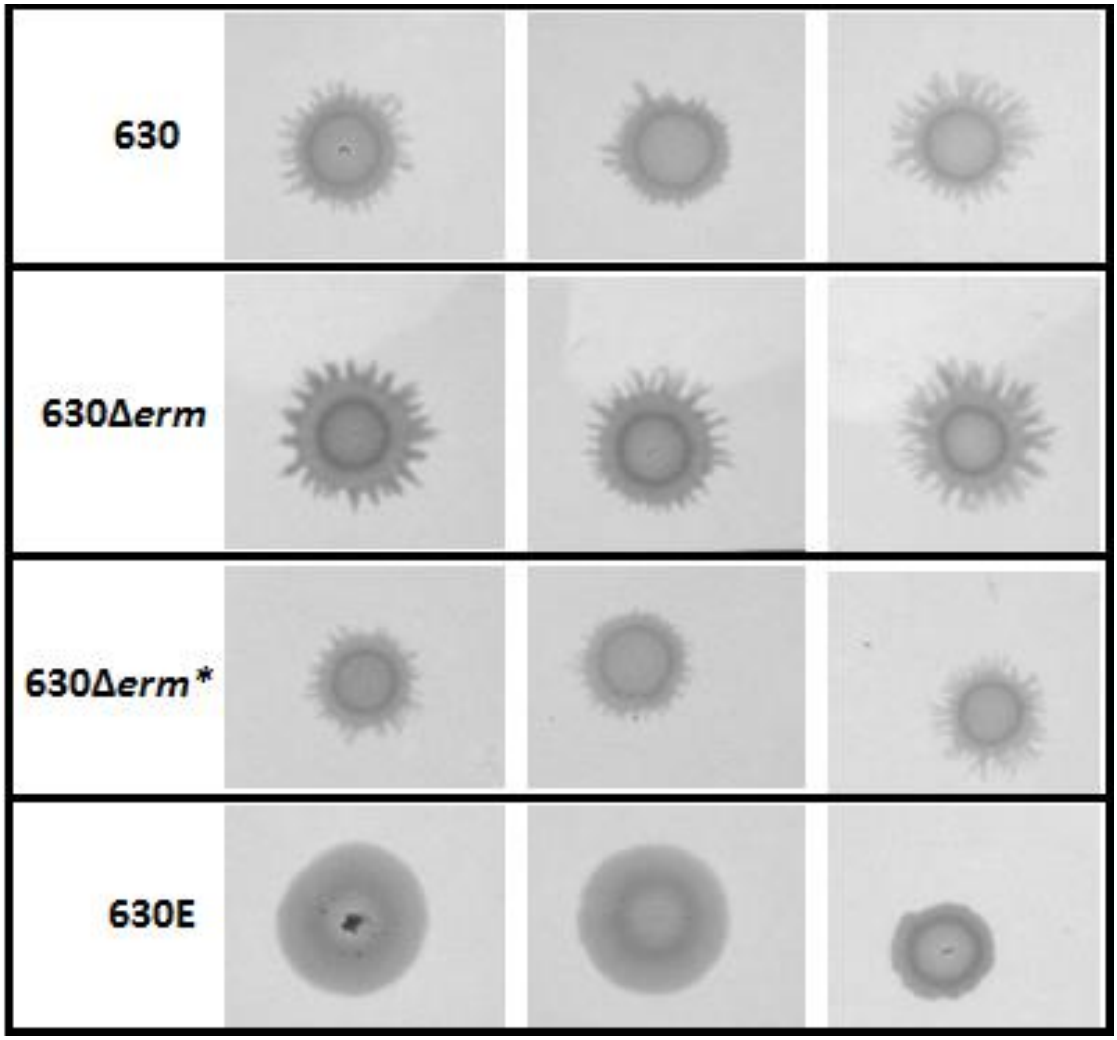


Figure 5-5 Motility of *C. difficile* 630 strains after 48 hours on 2YTG plates

C. difficile overnight cultures were spotted onto 2YTG +0.24% Gelzan plates and incubated anaerobically for 48 hours. Resultant growth was imaged and the proficiency for motility of each strain was assessed by eye. Images depict the results of three independent experiments.

5.4.5 Reversion to erythromycin resistance

Since 630*Δerm* still contains the *erm1(B)* gene, it was hypothesised that the newly generated strain 630*Δerm**, which lacks both *erm*(B) genes, may revert to an erythromycin resistant

phenotype at a reduced frequency to that shown for 630*Δerm* (2.79 x 10^{-8}; Hussein *et al*., 2005). To test this hypothesis, strains of 630, 630*Δerm* and 630*Δerm** were sub-cultured in BHIS broth for a total of 30 passages. After 0, 10, 20 and 30 passages, growth was serially diluted in PBS and plated onto BHIS agar with and without erythromycin. Colonies present after anaerobic incubation for 24 hours at 37°C were counted and CFUs calculated. The frequency of reversion to erythromycin resistance was calculated by dividing the CFU from media supplemented with erythromycin by the CFU from media lacking erythromycin (Figure 5.10).

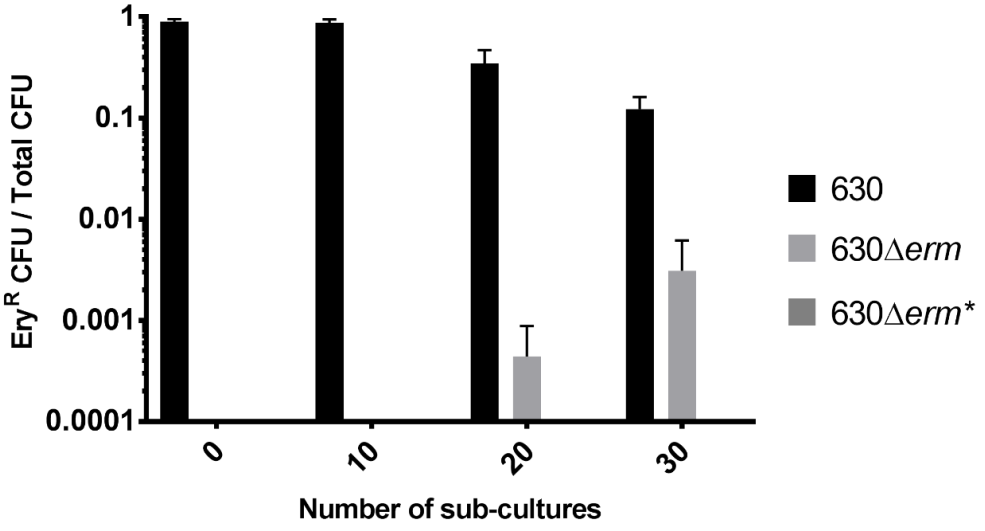


Figure 5-6 Reversion of erythromycin sensitive *C. difficile* 630 derivatives to resistant phenotypes following repeated subculture.

C. difficile 630 strains were grown in BHIS broth and sub-cultured 1/100 every 24 hours for 30 days. After 0, 10, 20 and 30 sub-cultures, cells were serially diluted in PBS, spread onto BHIS agar with and without erythromycin and the proportion of erythromycin resistant (Ery^R) CFU to total CFU for each strain was determined. Bars represent the mean of three independent experiments. Error bars indicate the standard errors of the means.

No erythromycin resistant CFU were observed for 630*Δerm** strains throughout the course of this assay. Erythromycin resistant CFU were observed in one 630*Δerm* triplicate from sub-culture number 20 onwards. However, in a colony PCR screen of 18 erythromycin resistant 630*Δerm* colonies from subculture 20, twelve produced a 6.088 kb WT-sized band consistent with possessing two *erm(B)* genes, whilst 6 produced a 3.682 kb sized band expected of 630*Δerm* strains. Repeat PCR screens on erythromycin resistant colonies of this 630*Δerm* triplicate following 30 subcultures produced only the 6.088 kb sized band expected of WT strain 630. Hence, whilst genuine 630Δ*erm* revertants were observed, due to a likely contamination of this culture with strain 630 the true level of reversion to erythromycin resistance by 630*Δerm* was not able to be determined.

5.5 Generation of a *pyrE* deficient *C. difficile* 630*Δerm** strain

Having exemplified CRISPR/Cas9 genome editing in *C. difficile* through the generation of a new erythromycin sensitive strain 630, 630Δerm*, the aim was to further test this mutagenesis method through the creation of a *pyrE*$^{-}$ strain. The creation of a *pyrE* mutant auxotrophic for uracil is a fundamental step in the established roadmap for precise genome editing within *Clostridium* species, previously performed using Allele Coupled Exchange. The aim of this study was to generate a *pyrE* truncation, identical to that generated with ACE, in 630Δ*erm** using CRISPR/Cas9 genome editing. This will generate a strain compatible with the current pMTL-YN1/1C/1X series of allelic exchange vectors, allowing the repair of the *pyrE* truncation alongside the integration of cargo DNA at this chromosomal locus, and further validate the CRISPR/Cas9 system in *C. difficile*.

5.5.1 Construction of CRISPR/Cas9 vectors targeting *pyrE*

A 2.363 kb editing template was designed *in silico* to delete 234 bp from the 3' end of the *pyrE* gene (CD630_01870) in *C. difficile* 630 and was generated via SOE PCR from *C. difficile* 630. Restriction digests and ligation reactions were performed to clone this editing template between the *Asi*SI and *Asc*I sites of pPSI_cas_121, generating pPSI_cas_130.

Within the target 234 bp to be deleted from the 3' end of *pyrE*, sixteen 5'-N_{20}NGG-3' sequences were identified, only seven of which had on target scores >50. From these, the three sgRNA seed sequences listed in Table 5.4 were chosen and incorporated into pPSI_cas_130 via *Sal*I restriction digestion and subsequent HiFi assembly. Sanger sequencing confirmed the successful sgRNA seed replacement and the consequent generation of CRISPR/Cas9 vectors pPSI_cas_131 (*pyrE*1), pPSI_cas_132 (*pyrE*2) and pPSI_cas_133 (*pyrE*3) each targeting *pyrE*.

Name	Sequence	PAM	Position	Strand	On score	Off score
***pyrE*1**	AGAGTATTAGAAGCCTTAGG	TGG	382	+	60.8	99.0
***pyrE*2**	GAGTGTCCTTTATGTAAGGA	AGG	508	+	54.1	100
***pyrE*3**	ACCTACAACTTCTCCACCTA	AGG	417	-	53.6	94.0

Table 5-4 sgRNA seed sequences identified within the *pyrE* target sequence.

sgRNA seeds and the corresponding PAM sequences identified using the Benchling sgRNA design tool at the indicated positions within *C. difficile* 630 *pyrE* (CD01870). On- and Off-target scores refer to rules set by Doench *et al*., (2016).

5.5.2 Generation of *C. difficile* 630Δ*erm**Δ*pyrE*

To generate *C. difficile* 630Δ*erm**Δ*pyrE*, CRISPR/Cas9 vectors pPSI_cas_131, pPSI_cas_132 and pPSI_cas_133 were conjugated into *C. difficile* 630Δ*erm** from *E. coli* CA434 donor cells. Thiamphenicol-resistant, transconjugant colonies appeared 48-72 hours after plating of the conjugal mixture onto BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol, and were re-streaked onto the same media to confirm the presence of the respective plasmids. To confirm the desired truncation of *pyrE*, 36 subsequent thiamphenicol resistant colonies, 12 from conjugations with each of the three CRISPR/Cas9 vectors used, were screened via colony PCR using primers CD630_pyrD_sF1 and CD630_0189_sR3 (Figure 5.11A).

All three sgRNAs were able to generate the desired deletion, the efficiencies of which are shown in Table 5.5. To determine whether this mutagenesis efficiency would increase with further re-streaks onto selective agar, as observed by (Huang *et al*., 2016), all 36 colonies were passaged twice more onto BHIS agar supplemented with thiamphenicol and the colony PCR repeated after each passage (Figures 5.11B and 5.11C). With each passage, and for each sgRNA, the number of mutant alleles present and the proportion of pure mutants obtained, increased. However, during this passaging two colonies (numbers 19 and 21) exhibited an apparent reversion to wild-type, as evidenced by the single presence of a 2.056 kb band corresponding to wild-type sequence (Figure 5.11C), despite both colonies previously displaying the 1.822 kb sized band expected for the *pyrE* truncated mutant (Figure 5.11B).

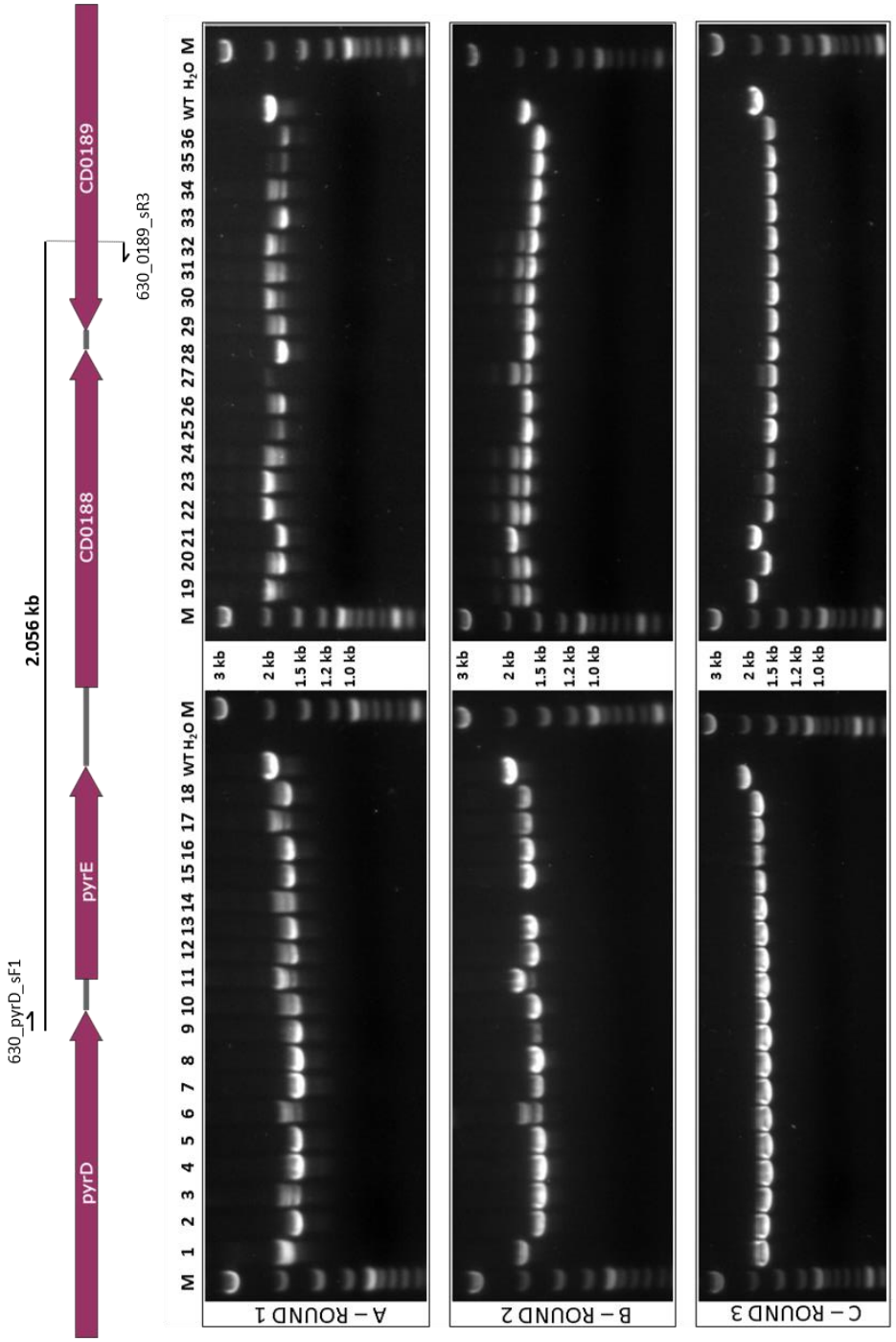


Figure 5-7 Colony PCR screening for *pyrE* truncated 630*Δerm strains following CRISPR/Cas9 mutagenesis.**

Top: Diagrammatic representation of PCR primer binding sites. Bottom: Agarose gel electrophoresis; M: DNA Marker (2-log ladder; NEB); Lanes 1-36: 36 thiamphenicol resistant colonies following (A) one passage, (B) two passages and (C) three passages on selective agar. WT: 630*Δerm* gDNA; H_2O: distilled water.

sgRNA	Round 1			Round 2			Round 3		
	M	**B**	**WT**	**M**	**B**	**WT**	**M**	**B**	**WT**
630*pyrE*1	**6/12**	**5/12**	**1/12**	**9/12**	**2/12**	**1/12**	**12/12**	**0/12**	**0/12**
630*pyrE*2	**4/12**	**6/12**	**2/12**	**5/11**	**5/11**	**1/11**	**9/12**	**1/12**	**2/12**
630*pyrE*3	**3/12**	**8/12**	**1/12**	**7/12**	**5/12**	**0/12**	**11/12**	**1/12**	**0/12**
Total	**13/36 (36%)**	**19/36 (53%)**	**4/36 (11%)**	**21/35 (60%)**	**12/35 (34%)**	**2/35 (6%)**	**32/36 (89%)**	**2/36 (6%)**	**2/36 (6%)**

Table 5-5 Efficiency of each gRNA in the generation of *C. difficile* 630*Δerm*Δ*pyrE* via CRISPR/Cas9 mutagenesis.**

Primary transconjugants following conjugation into 630*Δerm** of the three CRISPR/Cas9 vectors targeting *pyrE* were passaged for three rounds on BHIS agar supplemented with thiamphenicol. Values indicate the proportion of pure mutants (M), wild-type (WT) or a mixture of both alleles (B) present following colony PCR screening of the resulting thiamphenicol resistant colonies.

5.5.3 Sanger sequencing of pPSI_cas_132 from colonies 19 and 21

It was hypothesised that the reversion to wild-type exhibited by colonies 19 and 21 was caused by SNPs accumulated in the CRISPR/Cas9 vector pPSI_cas_132 present in these strains, following multiple passages. To investigate this, colony PCR using primers sgRNA2.6_R1 and Cas9_sR1 was performed on three colonies of each strain to generate 1.08 kb PCR products containing the sgRNA component, promoters P_{tcdB} and P_{thl} and the first 493 bp of the cas9 gene. Sanger sequencing of these PCR products was performed with the amplification primers and the resulting sequences aligned to the sequence of pPSI_cas_132. No mutations were observed in any of the triplicates for either colony 19 or 21. Hence, if mutations within pPSI_cas_132 are responsible for the reversion of these strains to wild-type, these must occur outside of the sequenced region.

5.6 Whole genome sequencing of *C. difficile* 630*Δerm**

To determine whether any genetic variants had been accumulated during CRISPR/Cas9 mutagenesis, each of the three independent *C. difficile* 630*Δerm** mutants plus the parental 630 strain were whole genome sequenced using the Illumina MiSeq platform (DeepSeq, Nottingham, UK). Alignments of the generated reads were performed using the CLC Genomics Workbench against a *C. difficile* 630 reference genome (AM108355.1) and variants were called according to the guidelines stated in Materials and Methods (Table 5.6).

Gene	Description	Position	630 reference	630 (CRG 856)			*Δerm** #1			*Δerm**#2			*Δerm** #3			AA
				S	F	C	S	F	C	S	F	C	S	F	C	
CD00730	*rplC*	103225	G	T	100	40	T	100	85	T	100	72	T	100	60	
CD02050	Transcription anti-terminator	268934	G				T	98	64	T	98	62	T	98	51	Gly165Cys
IG	-	690658	A	T	100	38	T	100	59	T	100	45	T	100	52	
CD11900	Acyl-coA N-acyltransferase	1391850	T	C	100	32	C	100	60	C	100	42	C	100	44	Phe133Leu
CD13880	pseudo	1607453	-	T	91.7	36	T	91	56	T	97	39	T	85	46	
CD17670	*gapB*	2044514	C	G	100	32	G	100	56	G	100	38	G	100	30	Pro33Ala
IG	-	2203033	A	T	97.2	36	T	100	48	T	100	35	T	100	30	
IG	-	2832892	G	T	100	21	T	100	50	T	100	57	T	100	22	
CD25320	Aminotransferase, alanine-glyoxylate transaminase	2924655	C	T	100	23	T	100	70	T	100	60	T	100	30	
CD26270	Conserved hypothetical protein	3034953	C	A	100	34	A	100	65	A	100	61	A	100	51	Gly68Cys
CD26670	*ptsG-BC*	3080703	C	T	100	28	T	100	65	T	100	60	T	100	53	Val228Ile
CD26850	*Putative sporulation stage protein E, spoIIE*	3105406	21 NT	-	100	13	-	100	29	-	100	30	-	100	18	ΔGlu115-Ala121
CD30890	PTS system, glucose-like IIBC	3590230	T	G	38	100	G	100	90	G	100	19	G	100	19	Glu258Asp
CD31561	Pseudo	3686535	-	A	52	100	A	98	49	A	100	22	A	100	20	Leu104fs
CD32450	*prdR*	3797112	C	T	63	98.4	T	100	82	T	98	57	T	100	46	Glu261Lys
IG	-	4007463	-	C	26	100	C	100	12	C	100	61	C	100	45	
IG	-	4007603	A	G	27	100	G	100	15	G	97	70	G	100	63	

Table 5-6 Resequencing analysis of *C. difficile* 630 (CRG856) and 630Δ*erm strains.**

SNPs and other variants identified following resequencing analysis of three 630Δ*erm** strains (#1, #2 and #3) and the parental strain 630 (CRG856) in the genes identified or in intergenic (IG) regions compared with the 630 reference genome (AM180355.1; Sebaihia *et al*., 2006). Position, genomic locus of variant; S, variant sequence at given locus; F, frequency of variant in generated reads; C, coverage of reads at given locus. Any resulting amino acid (AA) changes are also listed. Blank spaces denote no change from the 630 reference sequence.

Of the 17 genomic variants identified, 16 were common to all strains including the parental strain 630, indicating these discrepancies were present in the parental strain prior to CRISPR/Cas9 mutagenesis. The SNP at position 268934 was only called in the 630Δ*erm** triplicates and not the parental strain, which raised the possibility that this variant may have arisen during mutagenesis. However, a closer inspection of the WGS reads mapped to the reference genome at this locus revealed this SNP to be present in the parental strain. Furthermore, the coverage of reads in this area of the genome is low (≤10), which would explain why this SNP was not called initially. Hence, we can conclude that all 17 genomic variants identified between the 630 reference sequence and 630Δ*erm** strains were also present in the original 630 strain (CRG856) and that CRISPR/Cas9 mutagenesis was not responsible for the accumulation of any additional genomic variants.

A 21 bp deletion within CD630_26850, resulting in the in-frame deletion of seven amino acids, was detected in all four strains sequenced. To confirm this observation was genuine and not a mistake in the annotation of the 630 reference genome, a 623 bp region spanning the deletion site was PCR amplified using primers spoIIE_sF1 and spoIIE_sR1 from our 630 strain (CRG 856), 630*Δerm* and a new stock of *C. difficile* 630 (NCTC 13307) obtained from NCTC (National Collection of Type Cultures; Public Health England, UK). Sanger sequencing of these PCR products confirmed the 21 bp deletion in CD630_26850 occurred only in our 630 strain (CRG 856) and its progeny. This CD630_26850 gene is annotated as 'putative sporulation stage protein E, *spoIIE*', however, CD630_34900 is annotated as *spoIIE*. To confirm the identity of the *C. difficile spoIIE* gene and the CD630_26850 gene product, protein alignments for each gene product were performed against *B. subtilis* (Table 5.7). This confirmed that CD630_34900 encodes *spoIIE*, whilst the CD630_26850 protein sequence contains a C-terminal phosphatase domain belonging to the PP2C superfamily, a domain

found in sigma factor phosphatases including SpoIIE and RsbU. In summary, several genomic variations exist between our 630 strain (CRG 856) and the reference genome, but CRISPR/Cas9 mutagenesis was not responsible for the accumulation of any additional SNPs or other genetic variants.

Query	Query Length	Description	Coverage	E value	Identities
CD630_26850	589 aa	Phosphoserine phosphatase	28%	1e-16	49/170 (29%)
CD630_34900	790 aa	Stage II sporulation protein E	95%	6e-90	205/770 (27%)

Table 5-7 Protein alignments of CD630_26850 and CD630_34900 gene products against *B. subtilis*.

Amino acid sequences of both gene products were searched using the blastp algorithm against *B. subtilis* (taxid: 1423). Results of the best match in each case are shown.

5.7 CRISPR/Cas9 mutagenesis in *C. difficile* DH1916

Having used CRISPR/Cas9 to generate two separate deletion mutants in *C. difficile* 630, the system was to be further validated through exemplification in *C. difficile* strain DH1916. The *hsdR* gene, which encodes the restriction-subunit of the previously identified Type I restriction-modification system in DH1916, was chosen to be deleted with the intention of further increasing conjugation efficiency into this strain.

5.7.1 Construction of CRISPR/Cas9 vectors targeting DH1916_*hsdR*

A 2.045 kb editing template containing approximately 1 kb upstream and downstream of DH1916_*hsdR*, including the first and last three codons of *hsdR*, was generated via SOE PCR

as described in Materials and Methods. Restriction digestion and ligation reactions were performed to clone this editing template between the *Asi*SI and *Asc*I sites of similarly digested CRISPR/Cas9 vector pPSI_cas_111, used to generate a 3.648 kb deletion in Tn *5398* of strain 630. Sanger sequencing was used to confirm this insertion and the generated vector was named pPSI_cas_140.

As previously, candidate sgRNA seeds within the 3.336 kb *hsdR* gene of DH1916 were searched using the Benchling sgRNA design tool. Of the 139 identified 5'-N_{20}NGG-3' sequences, 3 were selected (Table 5.8) and incorporated into pPSI_cas_140 via *Sal*I restriction digest of the vector followed by separate HiFi assembly reactions. Successful replacement of the sgRNA seed region was confirmed in each case by Sanger sequencing with primer sgRNA_sF1, thus generating plasmids pPSI_cas_141 (*hsdR*1), pPSI_cas_142 (*hsdR*2) and pPSI_cas_143 (*hsdR*3) (Figure 5.12).

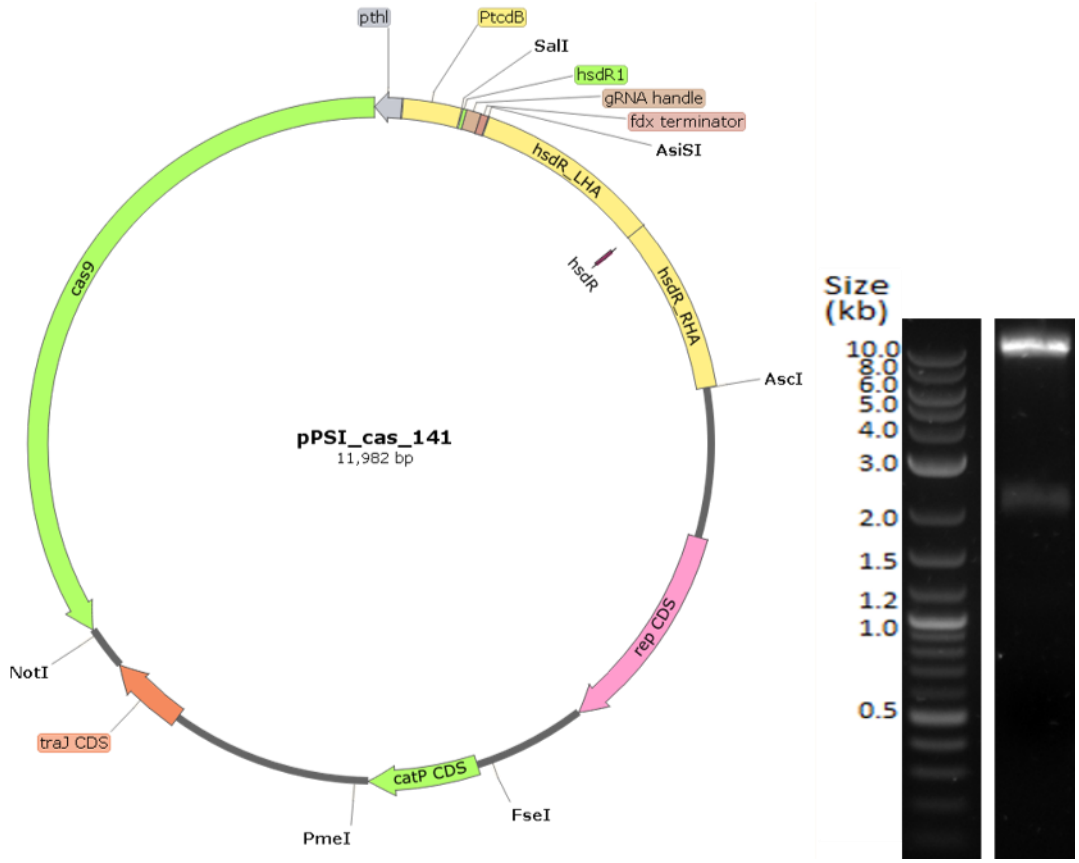


Figure 5-8 Diagrammatic representation of CRISPR/Cas9 vector pPSI_cas_141.

Left: Plasmid map for pPSI_cas_141, a CRISPR/Cas9 vector targeting *C. difficile* DH1916 *hsdR.* Features are as for pPSI_cas_111 except for the replacement of the *spo0A* editing template with one for generating a DH1916Δ*hsdR* strain, and the replacement of spo0A1 sgRNA seed with the hsdR1 sequence. Right: Agarose gel electrophoresis following restriction digestion with *Asi*SI and *Asc*I of pPSI_cas_141.

Name	Sequence	PAM	Position	Strand	On score	Off score
***hsdR*1**	TATTGAAAAAGCAATGACAG	AGG	2696	+	69.1	98.9
***hsdR*2**	ACTAGTGCGAACATAGAGAG	AGG	309	+	67.8	100
***hsdR*3**	TCACTATATGAAGTTATGAG	TGG	2503	-	66.7	99.7

Table 5-8 sgRNA seed sequences identified within the *hsdR* target sequence.

sgRNA seeds and the corresponding PAM sequences identified using the Benchling sgRNA design tool at the indicated positions within *C. difficile* DH1916 *hsdR* (DH1916_01870). On- and Off-target scores refer to rules set by Doench *et al*., (2016).

5.7.2 Generation of DH1916Δ*pyrE*Δ*hsdR* using CRISPR/Cas9 mutagenesis

CRISPR/Cas9 vectors pPSI_cas_141, pPSI_cas_142 and pPSI_cas_143 were transformed into *E. coli* NEB sExpress and the resulting strains used as conjugation donors for transfer of these plasmids into DH1916Δ*pyrE*. The efficiencies for conjugations of pPSI_cas_141-143 into DH1916Δ*pyrE* were determined, along with equivalent values for the conjugations of pPSI_cas_121-123 and pPSI_cas_131-133 into *C. difficile* 630, for comparison (Table 5.9). No thiamphenicol-resistant transconjugant colonies were obtained following conjugations of pPSI_cas_142 into DH1916Δ*pyrE*.

C. difficile strain	Vector (sgRNA)	Mean conjugation efficiency
DH1916Δ*pyrE*	pPSI_cas_141 (*hsdR*1)	2.77e-09 (± 1.86e-09)
	pPSI_cas_142 (*hsdR*2)	0
	pPSI_cas_143 (*hsdR*3)	1.01e-09 (± 8.25e-10)
630	pPSI_cas_121 (CD2008A)	5.71e-09*
	pPSI_cas_122 (CD2008B)	2.04e-09*
	pPSI_cas_123 (CD2008C)	2.08e-08*
630Δ*erm**	pPSI_cas_131 (*pyrE*1)	1.72e-05 (± 1.46e-05)
	pPSI_cas_132 (*pyrE*2)	1.70e-06 (± 1.12e-06)
	pPSI_cas_133 (*pyrE*3)	7.71e-07 (± 3.94e-07)

Table 5-9 Conjugation efficiencies using CRISPR/Cas9 vectors.

Conjugation efficiency calculated as thiamphenicol-resistant CFU/total *C. difficile* recipient CFU, resulting from conjugations of the stated plasmids into *C. difficile* DH1916Δ*pyrE* or 630 strains. Asterisked values denote results from a single experiment, all other values denote the means from three independent experiments. Numbers in ellipses indicate the standard error of the means.

Despite low efficiency, thiamphenicol-resistant transconjugant colonies were obtained following conjugations of pPSI_cas_141 and pPSI_cas_143 into DH1916. All ten such colonies were colony PCR screened for the in-frame deletion of *hsdR* using primers *hsdR*_sF1 and *hsdR*_sR1 (Figure 5.13). A 2.192 kb sized band corresponding to the expected deletion in *hsdR* was observed in all ten lanes, of which eight were pure mutants and two also contained the 5.510 kb sized band corresponding to the unmodified, parental strain. Colony PCR products from lanes 2, 5 and 8 were purified from the agarose gel and sanger sequenced with primers *hsdR*_sF1, *hsdR*_sF2 and *hsdR*_sR1. All three sequences were as expected for *hsdR* deletion mutants. These three DH1916Δ*pyrE*Δ*hsdR* strains, generated from independent conjugations and from which colony PCR products were obtained, were then cured of their respective plasmids. Patch plating onto BHIS agar with and without thiamphenicol was used to identify thiamphenicol-sensitive colonies which had undergone plasmid loss.

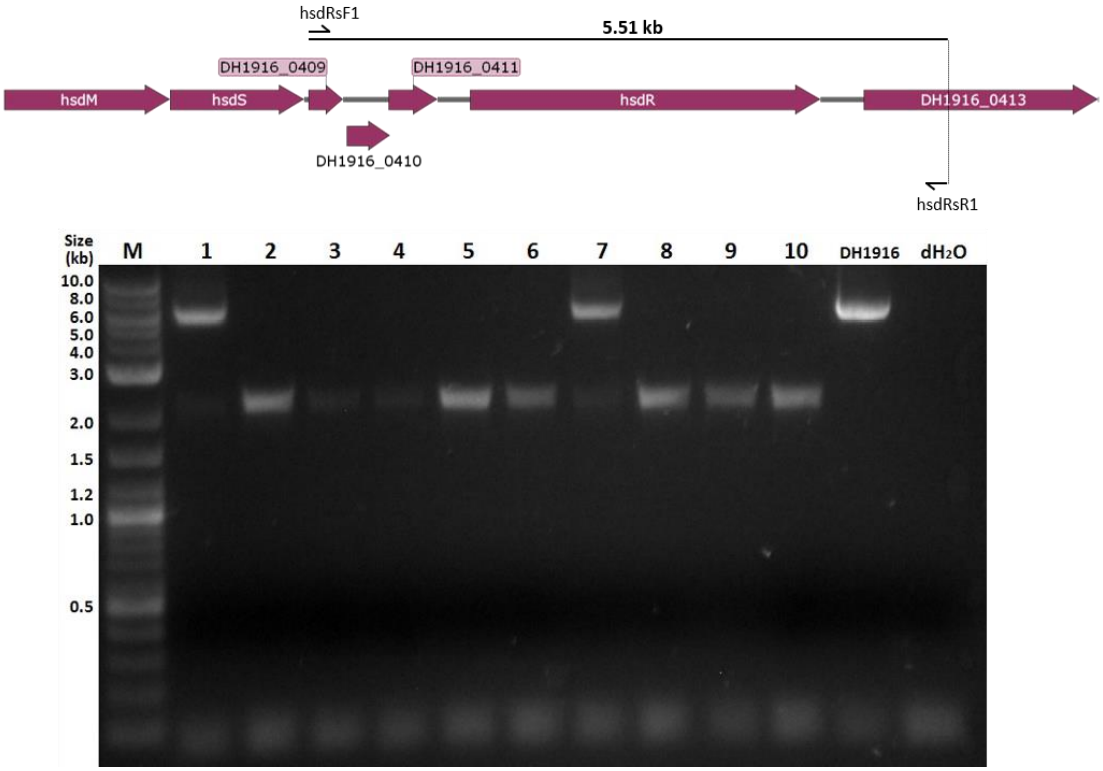


Figure 5-9 Colony PCR screening for *hsdR* deletion from DH1916Δ*pyrE* strains following CRISPR/Cas9 mutagenesis.

Top: Diagrammatic representation of PCR primer binding sites. Bottom: Agarose gel electrophoresis; M, DNA Marker (2-log ladder; NEB); Lanes 1-6, thiamphenicol resistant DH1916 colonies harbouring pPSI_cas_141; lanes 7-10, thiamphenicol-resistant DH1916 colonies harbouring pPSI_cas_143; DH1916: DH1916 gDNA; dH_2O: distilled water.

5.7.3 Assessment of conjugation efficiency in DH1916Δ*pyrE*Δ*hsdR*

To investigate the effects of the Type I restriction modification system on conjugation efficiency in *C. difficile*, conjugations from NEB sExpress donors harbouring the modular vectors pMTL81151-86151 were performed into DH1916Δ*pyrE* and the three newly generated DH1916Δ*pyrE*Δ*hsdR* strains (Figure 5.14). Plasmid pMTL81151, which lacks a Gram-positive replicon, was not successfully transferred into either DH1916 strain tested. Aside from pMTL81151, all other modular vectors tested showed increased conjugation efficiency of at least 100-fold into the *hsdR*$^-$ strain compared with DH1916Δ*pyrE*. Additionally, no thiamphenicol-resistant colonies were obtained following conjugations of cells harbouring pMTL85151 into DH1916Δ*pyrE*, whilst the mean conjugation efficiency of

pMTL85151 into DH1916Δ*pyrE*Δ*hsdR* was 1.5x10^{-6}. These results indicate that the HsdR Type-I restriction subunit is a clear barrier to conjugation-mediated DNA transfer, for all Gram-positive replicons tested, into *C. difficile* DH1916.

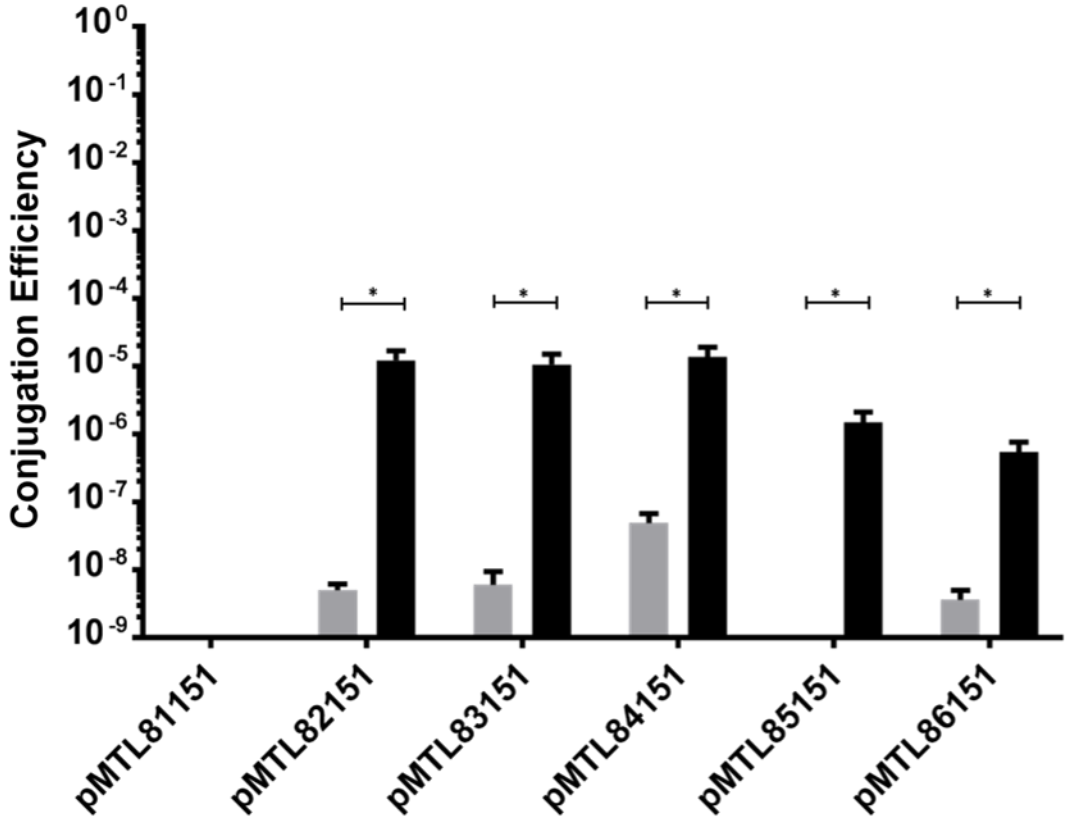


Figure 5-10 Effect of *hsdR* on conjugation efficiency into DH1916.

Conjugations into DH1916Δ*pyrE* (grey bars) or DH1916Δ*pyrE*Δ*hsdR* (black bars) overnight cultures from *E. coli* NEB sExpress cells harbouring various pMTL8000 modular vectors were performed as described in Materials and Methods. Conjugation efficiency calculated as the proportion of thiamphenicol resistant CFU obtained relative to the recipient *C. difficile* DH1916 CFU. Bars represent the means of three independent experiments. Error bars indicate the standard errors of the means. Statistical significance determined using multiple unpaired t-tests. Single asterisk denotes a *P*-value <0.05.

5.7.3.1 Determining the feasibility of suicide plasmid use in DH1916Δ*hsdR*

One of the most common methods to generate knockout mutants in bacteria is through the use of suicide vectors; plasmids which are unable to replicate in the target host. They can be used to carry out allelic exchange of wildtype genes by plasmid-encoded disrupted alleles, or may be exploited to deliver transposons to the host genome. Their deployment is, however, dependent on high rates of DNA transfer. To determine whether the increased conjugation efficiency observed following conjugations into DH1916 lacking *hsdR* was sufficient for the use of suicide plasmids, conjugations were performed into this strain and the parental DH1916Δ*pyrE* strain using *E. coli* cells harbouring pMTL-GL15, which carries a mariner transposon. The pMTL-GL15 suicide plasmid is identical to pMTL-SC1 (Figure 5.15), except that pMTL-GL15 does not contain the pBP1 Gram-positive replicon, thus making it replication deficient in clostridia. This *mariner* transposon has previously been shown to insert randomly into chromosomal TA target sequences via a cut-and-paste transposition mechanism, within several bacterial species, including *C. difficile* R20291 (Cartman & Minton, 2010).

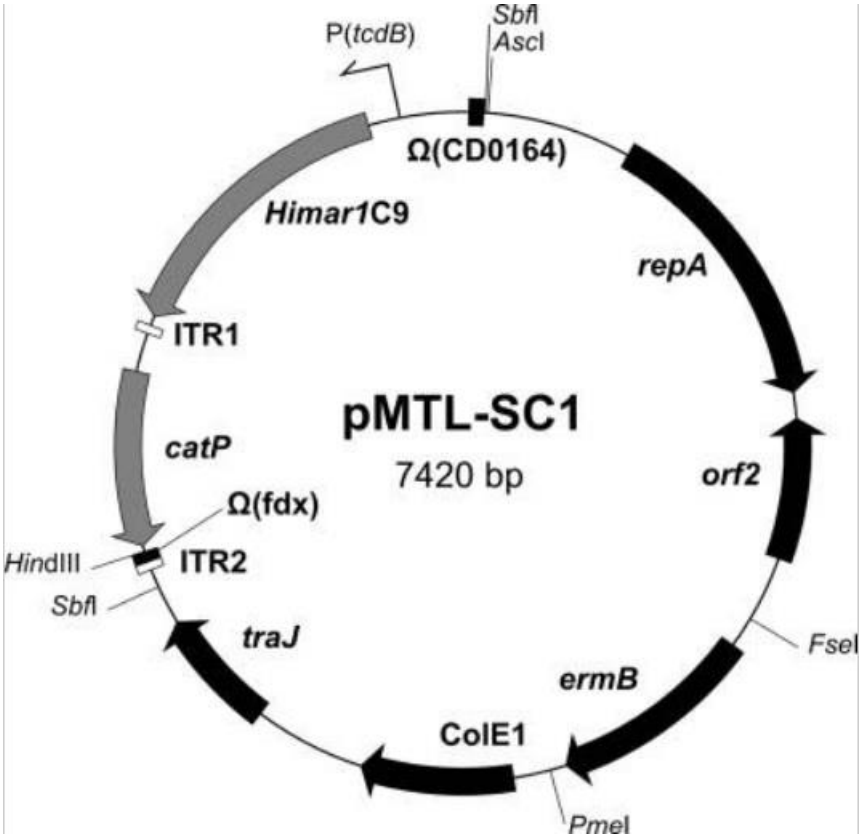


Figure 5-11 Diagrammatic representation of pMTL-SC1 constituents.

Vector map of pMTL-SC1 containing the hyperactive mariner transposase gene *Himar1* C9 under the control of the *C. difficile* toxin B promoter, P_{tcdB}, cloned into the pMTL82251 modular vector. Components *repA* and *orf2* correspond to the pBP1 Gram-positive replicon absent from pMTL-GL15. Image reproduced from (Cartman & Minton, 2010).

pMTL-GL15 was electroporated into competent *E. coli* NEB sExpress cells and the resulting chloramphenicol resistant strains used as conjugal donors for transfer of this vector into DH1916Δ*pyrE* and DH1916Δ*pyrE*Δ*hsdR* recipient cells. Transconjugant colonies in which the transposition of the *catP* gene into the DH1916 chromosome had occurred were subsequently selected on BHIS agar supplemented with D-cycloserine, cefoxitin and thiamphenicol. These thiamphenicol-resistant colonies were enumerated for each conjugation reaction performed and conjugation efficiency values calculated as previously (Table 5.10). No thiamphenicol-resistant colonies were obtained following conjugations of pMTL-GL15 into DH1916Δ*pyrE*. Whilst the mean conjugation efficiency was higher for this

vector into DH1916Δ*pyrE*Δ*hsdR* strains, the value of $7.55x10^{-9}$ is not sufficiently high to enable this system to be used to generate a random transposon library of mutants. Such systems can be used to generate libraries containing one million random transposon mutants. However, even with the improvements in conjugation efficiency observed for *hsdR*$^-$ strains, the level of conjugation efficiency observed for transfer of pMTL-GL15 into DH1916Δ*pyrE*Δ*hsdR* would require approximately 100,000 conjugation reactions to generate a library of one million transposon mutants.

Strain	Conjugation efficiency
DH1916Δ*pyrE*	0
DH1916Δ*pyrE*Δ*hsdR*	7.55e-09 (± 5.47e-10)

Table 5-10 Transposon mutagenesis efficiency from the suicide vector pMTL-GL15 into DH1916Δ*pyrE* and DH1916Δ*pyrE*Δ*hsdR*.

Conjugations from *E. coli* NEB sExpress cells harbouring pMTL-GL15 into DH1916 strains with or without functional *hsdR* were performed and the proportion of thiamphenicol-resistant CFU to total *C. difficile* conjugation recipient CFU for each strain was determined. Thiamphenicol resistant cells arise following successful conjugation and chromosomal integration of the *catP* gene from pMTL-GL15 via transposition. Values denote the means from three independent experiments. Values in parentheses indicate the standard errors of the means. The zero value for DH1916Δ*pyrE* indicates no thiamphenicol-resistant colonies were obtained following three independent conjugation reactions into this strain.

5.8 Discussion

This study has exemplified CRISPR/Cas9 genome editing for the first time in *C. difficile*. Furthermore, the editing efficiency of our CRISPR/Cas9 system compares favourably with similar systems used both in other clostridia and in more distantly-related organisms. Three genes were targeted for deletion in this study, CD630_20080, CD630_01870 (*pyrE*) and DH1916_00412 (*hsdR*) with overall editing efficiencies of 96%, 89% and 100% respectively.

Similar measures of efficiency for CRISPR/Cas9 systems used in other organisms are shown in Table 5.11.

Organism	Gene(s) targeted	Editing efficiency (%)	Reference
C. beijerinckii	*Pta*	100*	(Wang, *et al*., 2016)
C. cellulolyticum	*pyrF*	100	(Xu, *et al*., 2015)
C. ljungdahlii	*adhE1* *ctf* *pyrE*	75* 100 50	(Huang, *et al*. 2016)
C. autoethanogenum	*adh* *bdh*	50 64	(Nagaraju, *et al*. 2016)
B. subtilis	*amyE* *yvmC*	89 97	(Altenbuchner, 2016)
E. coli	*cadA*	86[#]	(Jiang, *et al*., 2015)

Table 5-11 Editing efficiencies reported using CRISPR/Cas9 genome editing systems in various organisms.

Editing efficiencies, from selected studies utilising the Type II CRISPR/Cas9 genome editing system with *S. pyogenes cas9*, were calculated as the proportion of strains containing the desired mutation relative to the total number of strains screened. Asterisked numbers denote studies in which <10 total colonies were screened. [#] denotes studies utilising a two-plasmid CRISPR/Cas system.

Thus, we now have a new method for precisely editing the *C. difficile* genome which overcomes many drawbacks encountered when using the current genetic tools on offer. Such advantages of CRISPR/Cas9 genome editing include; i) mutants are obtained faster than when using allelic exchange and as quick as with ClosTron mutagenesis, ii) fewer passages of transconjugants are required during mutagenesis, reducing the risks of SNPs and other genomic variants accumulating in the genome, iii) no antibiotic resistance marker is inserted into the genome which may result in downstream polar effects and limit further engineering of the strain, and, iv) both chromosomal crossover steps in homologous recombination with the editing template are simultaneously selected for. With this in mind, CRISPR/Cas9 genome editing should now be the preferred method for mutagenesis in *C. difficile*.

The aim of this study was to generate a new model erythromycin sensitive strain of *C. difficile* 630. Ultimately, this aim was not achieved due to the presence of pre-existing SNPs and genetic variations in the 630 strain (CRG856) used for mutagenesis. This study further highlights the importance of proper strain curation and routine genome sequencing. Without these measures, secondary genomic mutations may interfere with assayed phenotypes and obscure the true nature of the studied feature. One such variation was the 21 bp deletion within CD630_26850, present only in CRG856 and its progeny, resulting in a 7-amino acid deletion to the peptide sequence of this putative phosphoserine phosphatase. This mutation in CD630_26850 may cause conformational changes in the mature polypeptide which could reduce the activity of this enzyme. Given other proteins with this domain (SpoIIE and RsbU) are involved in regulating the sporulation process, such a change might explain the observed reduction in heat-resistant CFU development after 120 hours of 630 (CRG856; 40.8%) and 630Δ*erm** (21.2%) relative to values for 630Δ*erm*. This hypothesis is supported by a study in *B. subtilis* in which an N-terminal truncation of the SpoIIE protein abrogated localisation of this protein to the sporulation septum and resulted in the development of heat-resistant CFU at 30% of levels for wildtype cells (Arigoni *et al*., 1999).

Despite this, CRISPR/Cas9 genome editing was established in *C. difficile* and WGS revealed no additional genetic variants were accumulated during CRISPR/Cas9 genome editing. Hence, to generate a new model 630Δ*erm* strain as intended, the CRISPR/Cas9 vectors targeting CD630_20080 created in this study can be used to edit the recently obtained *C. difficile* 630 reference strain (NCTC 13307). Such a strain could then confirm the hypothesis that *erm1(B)* is responsible for the reversion to erythromycin resistance of 630Δ*erm*. The assay to investigate that hypothesis in this study was compromised in that only one triplicate of 630Δ*erm* exhibited reversion to erythromycin, which was then shown via colony PCR to

be contaminated with a strain possessing the wildtype *C. difficile* 630 Tn *5398* sequence. A future assay should take measures to minimise the risk of contamination and use a greater reaction volume which would allow for a greater number of cells to be present and increase the likelihood of observing a reversion event.

This study was unable to generate a *C. difficile* 630 *spo0A* deletion mutant, despite the very same CRISPR/Cas9 components later producing deletions of three other genomic targets. Only wild-type colonies arising from conjugations of pPSI_cas_111 into strain 630, suggests a lack of functioning of the CRISPR/Cas9 machinery in this instance. Possible explanations for these results include; i) the *spo0A* sgRNA seeds used were unable to direct Cas9 to the *spo0A* gene to perform its nucleolytic activity and kill cells not undergoing homologous recombination with the editing template, or ii) the *spo0A* editing template may contain elements which can interfere with the CRISPR/Cas9 machinery, abrogating either target recognition or nucleolytic activity. Future CRISPR/Cas9 genome editing studies will generate more data pertaining to the editing efficiency of particular sgRNAs which could then be used to aid the design of improved sgRNAs and/or identify motifs which inhibit Cas9 nucleolytic activity.

CRISPR/Cas9 genome editing was achieved with P_{tcdB} driving expression of the sgRNA component. However, since no *spo0A* mutant was produced when using *spo0A* sgRNAs under P_{tcdB} control, it cannot be concluded that the P_{araE} promoter is non-functional in *C. difficile*. A promoter assay in *C. difficile* would answer this question and also potentially identify a more suitable promoter, such as an inducible promoter. The use of inducible promoters to more tightly control the expression of one or more of the CRISPR/Cas9 components may allow for an increase in the number of mutants generated, since this

would give cells the opportunity to perform homologous recombination with the editing template prior to the induction of the nucleolytic activity of Cas9. This could also be achieved by using the promoter of a sporulation gene, as performed in a study by Wang *et al*., (2016) in which the native *spoIIE* promoter was used to drive *cas9* expression for the generation of CRISPR/Cas9 mutants in *C. beijerinckii*. Despite this, the use of P_{tcdB} to control expression of the sgRNA component in this study is advantageous in that it minimises the chance of the CRISPR/Cas9 machinery being active in *E. coli*. Furthermore, expression from P_{tcdB} is repressed in a glucose-dependent manner, therefore the use of BHIS agar in this study which contains 2 $g.L^{-1}$ glucose may impede expression of the sgRNA component until metabolism has sufficiently reduced the glucose concentration in the media. Future work could assess the editing efficiency of the CRISPR/Cas9 system generated in this study using different media with a range of glucose concentrations.

5.8 Summary

- CRISPR/Cas9 genome editing was successfully implemented in *C. difficile* strains 630 and DH1916
- Editing efficiencies for the three genes targeted in this study compared favourably with previous studies
- Deletion of the restriction-modification subunit, *hsdR*, from DH1916 increased conjugation efficiency by at least two orders of magnitude for each Gram-positive replicon tested

Chapter Six

General Discussion

6 General Discussion

6.1 Functions of the SpoVA proteins

In Chapter Three of this work, it was determined that SpoVA proteins are required for the accumulation of DPA and calcium into spores of *C. difficile* 630Δ*erm*. Whilst this result was not unexpected, its undertaking was necessitated by the unique nature of sporulation and germination mechanisms of *C. difficile* amongst the Firmicutes; examples of this include the lack of homologues to the GerA family of germinant receptors in *C. difficile* strains thus far, and the use of bile salts as positive and negative germination signals. Furthermore, our understanding of the mechanisms involved in *C. difficile* sporulation and germination is only just beginning to increase, thanks to the recent development of the appropriate reverse genetics tools. A greater understanding of the precise sporulation and germination mechanisms employed by *C. difficile* can ultimately lead to novel therapies. Since germination is a critical step in the *C. difficile* disease cycle, required to form the toxin-producing vegetative cells, this stage can be targeted for therapeutic intervention. Given the findings of this study, this could be achieved by targeting the SpoVA proteins to either prevent DPA release *in vivo* thereby potentially reducing the germination capability of the spores, or by triggering DPA release from endospores to potentially increase their susceptibility to routine disinfectants within healthcare settings. Following such a hypothesis, a recent study has treated *C. difficile* spores with iron oxide nanocrystals which significantly reduced DPA release from spores and prevented 64% of these spores from completing germination (Lee *et al*., 2017). These authors also tested the nanoparticles *in vivo* in a hamster infection model and observed decreased inflammation when iron oxide nanoparticles were present compared to without. Future work should identify where in the

spore these nanoparticles bind to investigate the partial, but not complete, reduction in DPA release from iron oxide nanoparticle treated spores.

This study has shown that *C. difficile* spores with reduced core DPA content are less able to withstand exposure to wet heat, UV radiation and ethanol, and return to vegetative cell growth than spores containing wildtype levels of DPA. Future work should investigate the specific role DPA plays in the resistance properties of *C. difficile* endospores, either directly, or via dehydration of the spore core. Firstly, it should be confirmed that the reduction in CFU obtained following the various stress treatments with spores lacking DPA is due to spore killing and not due to these spores being in a 'superdormant' state. Germination is an asynchronous event and there exists a sub-population of spores which exhibit a prolonged lag phase between activation and commitment to germination, hypothesised to be due to very low numbers of germinant receptors present within these spores' inner membranes (Ghosh *et al*., 2012). If these treatments are subsequently found to be sporicidal, the method of spore killing should then be identified. For example, to determine whether spores are killed via the accumulation of DNA damage, the resistance assays in this study could be repeated in *spoVA* mutants lacking *recA* and the resulting CFU.ml^{-1} levels compared to the results from this study. Furthermore, given the multi-factorial nature of spore resistance mechanisms, future studies should; i) assess the water content of *spoVA*$^{-}$ spores compared to wildtype, ii) compare the levels of SASPs in *spoVA*$^{-}$ spores relative to wildtype, and iii) use electron microscopy to observe any ultrastructure differences between wildtype spores and those lacking SpoVA proteins.

This study has shown that *C. difficile* spores lacking any, and all, of the three *spoVA* genes contain lower amounts of DPA relative to wildtype spores. A recent study in *C. sporogenes*

determined that the presence of a second, pentacistronic, *spoVA* operon (*spoVA2*) in strain PA3679 conferred greater heat resistance properties to its spores relative to strains lacking *spoVA2* (Butler *et al*., 2017). These results agree with another recent study in *B. subtilis* in which the additional presence of a second *spoVA* operon, present on a mobile genetic element, resulted in spores with increased levels of DPA and higher resistance properties (Berendsen *et al*., 2016). Hence, it appears that the presence of an additional copy of *spoVAC*, *spoVAD* and *spoVAEb* in *B. subtilis* and *C. sporogenes* leads to spores with increased DPA content, suggesting a dose-dependent link between DPA content and spore heat resistance. Consequently, future work in *C. difficile* could overexpress the single *spoVA* operon present, utilising the pMTL-YN1X overexpression vector designed for this purpose (Ng *et al*., 2013), to determine if this overexpression of the single *spoVA* operon results in similar increases in DPA content in, and heat-resistance properties of, *C. difficile* spores. If it does not, we may conclude that the observations in *C. sporogenes* and *B. subtilis* are due to specific features of the additional *spoVA2* operon. If the opposite is true, further studies could be performed to attempt to determine the maximum DPA content, and potentially also the maximum heat resistance, of *C. difficile* endospores.

6.2 Endospore Clumping

Following the purification of endospores of strains 630Δ*erm* and R20291 via repeated washing in distilled water, clumping was observed which prevented accurate determination of spore counts under phase contrast microscopy. However, such clumping was not observed for endospores of DH1916, therefore spore clumping appears to be a strain-specific trait which could be explained by variations in the outermost layer of the endospores of these strains. A recent study investigated the hydrophobicity of endospores from a variety of PCR-ribotype *C. difficile* strains and found an association between

hydrophobicity and presence of an exosporium (Joshi *et al*., 2012). An attractive hypothesis is that spore clumping is due to hydrophobic interactions between endospores and the washing buffer, and that endospores which do not clump when washed in distilled water are due to a diminished or absent exosporium in endospores of these strains. This could be investigated in future work by imaging endospores of DH1916 (non-clumping spores), 630Δ*erm* and R20291 (clumping spores) via electron microscopy and comparing the sizes of the exosporium layers. This hypothesis could also explain why the addition of Tween80 to the spore washing buffer, generating a more hypertonic solution, was able to abrogate the clumping of 630Δ*erm* and R20291 spores.

6.3 CRISPR/Cas9 genome editing

In chapter five of this work, CRISPR/Cas9 genome editing was established in *C. difficile* and used to generate three deletions of varying sizes across two different strains. This was performed using a functional variant of the *S. pyogenes cas9* gene, containing a frame-shift mutation in the RuvCI domain, which along with HNH is one of the two domains responsible for nucleolytic activity. Thus, it seems likely that this Cas9 variant generates single-stranded DNA cleavage (nicks), termed a Cas9 nickase (Cas9n), and is similar to that used to perform Cas9-mediated genome editing in *C. cellulolyticum* (Xu *et al*., 2015). Use of Cas9n is associated with improved efficiency of genome editing (Xu *et al*., 2015), which may explain the high levels of editing efficiency observed in this study. Future studies could investigate this by using the wildtype *S. pyogenes* Cas9 to re-generate the deletion mutants made in this study and compare the subsequent editing efficiencies of the two nucleases. Another Cas9 variant of potential interest for future studies is dCas9, in which both nucleolytic domains are non-functional resulting in a catalytically inactive protein lacking any nucleolytic activity. This dCas9 protein is then used in a process termed CRISPR-interference (CRISPRi), during

which dCas9 is guided to a target genomic locus by an appropriate sgRNA and the chromosomal binding of this dCas9-sgRNA complex efficiently blocks transcription elongation to silence the target gene. CRISPRi was successfully demonstrated in *E. coli* and resulted in up to 99.9% repression of target gene transcription (Larson *et al*., 2013). Future studies could establish CRISPRi gene silencing in *C. difficile* and subsequently use this reversible method to investigate the function of essential genes which we have previously been unable to inactivate.

In addition to the expanded use of heterologous CRISPR/Cas systems, future work should also investigate the native CRISPR/Cas system previously identified in *C. difficile* (Sebaihia *et al*., 2006). *C. difficile* possesses a type I-B CRISPR/Cas system and initial experiments confirmed the expression of CRISPR arrays in R20291 and by assaying the conjugation efficiency of various plasmids were able to putatively identify the protospacer-adjacent motif (PAM) of this system to be a 5'-CCW sequence (Boudry *et al*., 2015). As yet, there are no reports of using this native CRISPR/Cas system for genome editing in *C. difficile*, therefore such a study along with further characterisation of this understudied type I-B CRISPR system should be performed.

Many authors have referred to a 'CRISPR revolution' currently occurring, likening the numerate studies investigating the functioning and applications of CRISPR/Cas9 systems to the discovery and application of restriction enzymes in the 1970's. Aside from genome editing, CRISPR/Cas systems could also be used in molecular typing of *C. difficile* strains in the future. A recent study analysing CRISPR systems from within 217 *C. difficile* genomes identified predominantly type I-B CRISPR/Cas systems with an unusually high average number of CRISPR arrays (8.5) per genome (Andersen *et al*., 2016). Furthermore, these

authors suggest that there is sufficient variation within these CRISPR systems to permit CRISPR-based phylogenetic and typing analysis. Given that novel spacer sequences are inserted in a polarised manner at the leader end of the CRISPR locus, it is likely that analysis of these genomic regions will enhance future epidemiological studies of *C. difficile*. Future studies should also investigate the potential for CRISPR/Cas systems to be exploited for therapeutic interventions against *C. difficile*. For example, a recent study generated a phagemid containing the *S. pyogenes cas9* gene and a guide targeting the methicillin resistance gene, *mecA*, of *Staphylococcus aureus*, which selectively killed methicillin-resistant strains of this bacterium (Bikard *et al*., 2014). A similar approach could be taken to selectively kill toxigenic strains of *C. difficile*, for example by suppling a sgRNA targeting the PaLoc. In short, CRISPR/Cas9 genome editing represents as highly efficient method of mutagenesis which overcomes many of the shortfalls of previously available tools for reverse genetics studies in clostridia. Furthermore, the enormous potential applications of CRIPSR/Cas systems are hugely exciting but require further characterisation and proof of concept studies before these can be fully exploited.

Chapter Seven

Bibliography

7 Bibliography

Abt, M. C., McKenney, P. T., & Pamer, E. G. (2016). *Clostridium difficile* colitis: pathogenesis and host defence. *Nature Reviews Microbiology*, *14*(10), 609–620.

Ackermann, G., Tang, Y. J., Henderson, J. P., Rodloff, C., Silva, J., & Cohen, S. H. (2001). Electroporation of DNA sequences from the pathogenicity locus (PaLoc) of toxigenic *Clostridium difficile* into a non-toxigenic strain. *Molecular and Cellular Probes*, *15*(5), 301–306.

Akinosho, H., Yee, K., Close, D., & Ragauskas, A. (2014). The emergence of *Clostridium thermocellum* as a high utility candidate for consolidated bioprocessing applications. *Frontiers in Chemistry*, *2*(8), 1–18.

Altschul, S., Gish, W., Miller, W., Myers, E. & Lipman, D. (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215(3), 403-410.

Ammann, A. B., Kölle, L., & Brandl, H. (2011). Detection of bacterial endospores in soil by terbium fluorescence. *International Journal of Microbiology*, *2011*, 10–15.

Andersen, J. M., Shoup, M., Robinson, C., Britton, R., Olsen, K. E. P., & Barrangou, R. (2016). CRISPR Diversity and microevolution in *Clostridium difficile*. *Genome Biology and Evolution*, *8*(9), 2841–2855.

Anjuwon-Foster, B. R., & Tamayo, R. (2017). Phase Variation of *Clostridium difficile* Virulence Factors. *Gut Microbes*, *976*(9), 1-8.

Anne Greene, E., & Spiegelman, G. B. (1996). The Spo0A protein of *Bacillus subtilis* inhibits transcription of the *abrB* gene without preventing binding of the polymerase to the promoter. *Journal of Biological Chemistry*, *271*(19), 11455–11461.

Arigoni, F., Guérout-Fleury, a M., Barák, I., & Stragier, P. (1999). The SpoIIE phosphatase, the sporulation septum and the establishment of forespore-specific transcription in *Bacillus subtilis*: a reassessment. *Molecular Microbiology*, *31*(5), 1407–15.

Asheshov, E. A., & M Patricia, J. (1963). The Effect of Heat on the Ability of a Host Strain to Support the Growth of a Staphylococcus Phage. *Journal of General Microbiology*, *31*(1963), 97–107.

Baban, S. T., Kuehne, S. A., Barketi-Klai, A., Cartman, S. T., Kelly, M. L., Hardie, K. R., … Minton, N. P. (2013). The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. *PLoS ONE*, *8*(9), 1-12.

Bacci, S., Mølbak, K., Kjeldsen, M. K., & Olsen, K. E. P. (2011). Binary toxin and death after *Clostridium difficile* infection. *Emerging Infectious Diseases*, *17*(6), 976–982.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., … Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819), 1709-1712.

Barrangou, R., & Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. *Nature Microbiology*, *2*(6), e17092.

Berendsen, E. M., Boekhorst, J., Kuipers, O. P., & Wells-Bennik, M. H. J. (2016). A mobile genetic element profoundly increases heat resistance of bacterial spores. *International Society for Microbial Ecology*, *10*(11), 2633-2642.

Bignardi, G. E. (1998). Risk factors for *Clostridium difficile* infection. *Journal of Hospital Infection*, *40*(1), 1–15.

Bikard, D., Euler, C. W., Jiang, W., Nussenzweig, P. M., Goldberg, G. W., Duportet, X., … Marraffini, L. A. (2014). Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nature Biotechnology*, *32*(11), 1146–1150.

Birrer, G. A., Chesbro, W. R., & Zsigray, R. M. (1994). Electro-transformation of *Clostridium beijerinckii* NRRL B-592 with shuttle plasmid pHR106 and recombinant derivatives. *Applied Microbiology and Biotechnology*, *41*(1), 32–38.

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*, 2551–2561.

Boudry, P., Semenova, E., Monot, M., Datsenko, K., Lopatina, A., Sekulovic, O., … Soutourina, O. (2015). Function of the CRISPR-Cas System of the Human Pathogen *Clostridium difficil*e. *mBio*, *6*(5), 1–16.

Brazier, J. S., Raybould, R., Patel, B., Duckworth, G., Pearson, A., Charlett, A., & Duerden, B. I. (2008). Distribution and antimicrobial susceptibility patterns of *Clostridium difficile* PCR ribotypes in English hospitals, 2007-08. *Euro Surveillance: European Communicable Disease Bulletin*, *13*(41), 1–5.

Brouwer, M. S. M., Roberts, A. P., Hussain, H., Williams, R. J., Allan, E., & Mullany, P. (2013). Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. *Nature Communications*, *4*, 1–6.

Burns, D. A., Heap, J. T., & Minton, N. P. (2010). The diverse sporulation characteristics of *Clostridium difficile* clinical isolates are not associated with type. *Anaerobe*, *16*(6), 618–622.

Burns, D. (2011). Analysis of the spore germination mechanisms of *Clostridium difficile*. PhD thesis, University of Nottingham.

Butler, R. R., Schill, K. M., Wang, Y., & Pombert, J. F. (2017). Genetic characterization of the exceptionally high heat resistance of the non-toxic surrogate *Clostridium sporogenes* PA 3679. *Frontiers in Microbiology*, *8*(4), 1–11.

Carter, J. (2013). Deaths involving *Clostridium difficile*: England and Wales, 2008. *Health Statistics Quarterly*, (43), 43–47.

Cartman, S. T., Kelly, M. L., Heeg, D., Heap, J. T., & Minton, N. P. (2012). Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production, *Applied and Environmental* Microbiology, *78*(13), 4683–4690.

Cartman, S. T., & Minton, N. P. (2010). A mariner-Based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Applied and Environmental Microbiology*, *76*(4), 1103–1109.

Cassir, N., Benamar, S., & La Scola, B. (2016). *Clostridium butyricum*: From beneficial to a new emerging pathogen. *Clinical Microbiology and Infection*, *22*(1), 37–45.

Chang, T. W., Bartlett, J. G., Gorbach, S. L., & Onderdonk, A. B. (1978). Clindamycin-induced enterocolitis in hamsters as a model of pseudomembranous colitis in patients. *Infection and Immunity*, *20*(2), 526–529.

Ciampi, M. S., & Roth, J. R. (1988). Polarity effects in the *hisG* gene of salmonella require a site within the coding sequence. *Genetics*, *118*(2), 193–202.

Cole, S. A., & Stahl, T. J. (2015). Persistent and Recurrent *Clostridium difficile* Colitis. *Clinics in Colon and Rectal Surgery*, *1*(212), 65–69.

Coleman, J. P., & Hudson, L. L. (1995). Cloning and characterization of a conjugated bile acid hydrolase gene from *Clostridium perfringens*. *Applied and Environmental Microbiology*, *61*(7), 2514–2520.

Cooksley, C. M., Davis, I. J., Winzer, K., Chan, W. C., Peck, M. W., & Minton, N. P. (2010). Regulation of neurotoxin production and sporulation by a putative *agrBD* signaling system in proteolytic *Clostridium botulinum*. *Applied and Environmental Microbiology*,

76(13), 4448–4460.

Daniel, R. A., & Errington, J. (1993). Cloning, DNA sequence, functional analysis and transcriptional regulation of the genes encoding dipicolinic acid synthetase required for sporulation in *Bacillus subtilis*. *Journal of Molecular Biology*, *232*(2), 468-483.

Deakin, L., Clare, S., Fagan, R., Dawson, L., Pickard, D., West, M., Wren, B., Fairweather, N., Dougan, G. & Lawley, T. (2012). The *Clostridium difficile spo0A* gene is a persistence and transmission factor. *Infection and Immunity*, *80*(8), 2704-2711.

Dong, H., Zhang, Y., Dai, Z., & Li, Y. (2010). Engineering Clostridium strains to accept unmethylated DNA. *PLoS ONE*, *5*(2), 1–8.

Donnelly, M. L., Fimlaid, K. A. & Shen, A. (2016). Characterization of *Clostridium difficile* Spores Lacking Either SpoVAC or Dipicolinic Acid Synthetase, *Journal of Bacteriology*, *198*(11), 1694–1707.

Driks, A. (1999). *Bacillus subtilis* Spore Coat. *Microbiology and Molecular Biology Reviews*, *63*(1), 1–20.

Dubberke, E. R., & Olsen, M. A. (2012). Burden of *Clostridium difficile* on the healthcare system. *Clinical Infectious Diseases*, *55*(2), 2–6.

Errington, J. (2010). From spores to antibiotics via the cell cycle. *Microbiology*, *156*(1), 1–13.

Eyre, D. W., Walker, A. S., Wyllie, D., Dingle, K. E., Griffiths, D., Finney, J., … Peto, T. E. A. (2012). Predictors of first recurrence of *Clostridium difficile* infection: Implications for initial management. *Clinical Infectious Diseases*, *55*(2), 77–87.

Farrow, K. A., Lyras, D., & Rood, J. I. (2001). Genomic analysis of the erythromycin resistance element Tn 5398 from. *Microbiology*, *147*, 2717–2728.

Figueroa, I., Johnson, S., Sambol, S. P., Goldstein, E. J. C., Citron, D. M., & Gerding, D. N.

(2012). Relapse versus reinfection: Recurrent *Clostridium difficile* infection following treatment with fidaxomicin or vancomycin. *Clinical Infectious Diseases*, *55*(2), 104–109.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D., & Shen, A. (2013). Global Analysis of the Sporulation Pathway of *Clostridium difficile*. *PLoS Genetics*, *9*(8), 1-20.

Francis, M. B., Allen, C. A., Shrestha, R., & Sorg, J. A. (2013). Bile Acid Recognition by the *Clostridium difficile* Germinant Receptor, CspC, Is Important for Establishing Infection. *PLoS Pathogens*, *9*(5), 1-9.

Freeman, J., Bauer, M. P., Baines, S. D., Corver, J., Fawley, W. N., Goorhuis, B., … Wilcox, M. H. (2010). The changing epidemiology of *Clostridium difficile* infections. *Clinical Microbiology Reviews*, *23*(3), 529–549.

Galperin, M. Y., Mekhedov, S. L., Puigbo, P., Smirnov, S., Wolf, Y. I., & Rigden, D. J. (2012). Genomic determinants of sporulation in Bacilli and Clostridia: Towards the minimal set of sporulation-specific genes. *Environmental Microbiology*, *14*(11), 2870–2890.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9 – crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *PNAS*, *109*(39), 2579–2586.

Gautheret, D. & Lambert, A. (2001) Direct RNA Motif Definition and Identification from Multiple Sequence Alignments using Secondary Structure Profiles. *Journal of Molecular Biology, 313,* 1003–1011.

George, R. H., Symonds, J. M., Dimock, F., Brown, J. D., Arabi, Y., Shinagawa, N., … Burdon, D. W. (1978). Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. *British Medical Journal*, *1*(6114), 695.

Gerhard, R., Nottrott, S., Schoentaube, J., Tatge, H., Oiling, A., & Just, I. (2008). Glucosylation

of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells. *Journal of Medical Microbiology*, *57*(6), 765–770.

Ghose, C. (2013). *Clostridium difficile* infection in the twenty-first century. *Emerging Microbes & Infections*, *2*(9), e62.

Ghosh, S., Scotland, M., & Setlow, P. (2012). Levels of germination proteins in dormant and superdormant spores of *Bacillus subtilis*. *Journal of Bacteriology*, *194*(9), 2221–2227.

Giel, J. L., Sorg, J. A., Sonenshein, A. L., & Zhu, J. (2010). Metabolism of bile salts in mice influences spore germination in *Clostridium difficile*. *PLoS ONE*, *5*(1), e8740.

Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y., Cartman, S. T., Winzer, K., & Minton, N. P. (2012). Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker, *Nucleic Acids Research*, *40*(8), e59.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., & Minton, N. P. (2007). The ClosTron: A universal gene knock-out system for the genus Clostridium. *Journal of Microbiological Methods*, *70*(3), 452–464.

Heeg, D., Burns, D. A., Cartman, S. T., & Minton, N. P. (2012). Spores of *Clostridium difficile* clinical isolates display a diverse germination response to bile salts. *PLoS ONE*, *7*(2), e32381.

Henrich, T. J., Krakower, D., Bitton, A., & Yokoe, D. S. (2009). Clinical risk factors for severe *Clostridium difficile*-associated disease. *Emerging Infectious Diseases*, *15*(3), 415–422.

Higgins, D., & Dworkin, J. (2013). Recent progress in Bacillus subtilus sporulation. *Microbial Reviews*, *36*(1), 131–148.

Hilbert, D. W., & Piggot, P. J. (2004). Compartmentalization of gene expression during *Bacillus subtilis* spore formation. *Microbiology and Molecular Biology Reviews*, *68*(2),

234–262.

Hofmann, F. (1963). The Function of Bile Salts in Fat Absorption. the Solvent Properties of Dilute Micellar Solutions of Conjugated Bile Salts. *The Biochemical Journal*, *89*(1953), 57–68.

Huang, H., Chai, C., Li, N., Rowe, P., Minton, N. P., Yang, S., … Gu, Y. (2016). CRISPR/Cas9-Based Efficient Genome Editing in *Clostridium ljungdahli*i, an Autotrophic Gas-Fermenting Bacterium. *ACS Synthetic Biology*, *5*(12), 1355–1361.

Hussain, H. A., Roberts, A. P., & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δ*erm*) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. *Journal of Medical Microbiology*, *54*(2), 137–141.

Jamroskovic, J., Chromikova, Z., List, C., & Bartova, B. (2016). Variability in DPA and Calcium Content in the Spores of Clostridium Species. *Frontiers in Microbiology*, *7*(11), 1–11.

Jansen, R., Embden, J. D. A. Van, Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes, *Molecular Microbiology*, *43*(6), 1565–1575.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816-821.

Joshi, L. T., Phillips, D. S., Williams, C. F., Alyousef, A., & Baillie, L. (2012). Contribution of spores to the ability of clostridium difficile to adhere to surfaces. *Applied and Environmental Microbiology*, *78*(21), 7671–7679.

Kai Soo Tan, Boon Yu Wee, & Keang Peng Song. (2001). Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. *Journal of Medical Microbiology*,

50(7), 613–619.

Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J., & Lambowitz, A. M. (2001). Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nature Biotechnology*, *19*(12), 1162–1167.

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., & Cigan, M. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biology*, 1–13.

Kelly, C. P., & LaMont, J. T. (1998). *Clostridium difficile* Infection. *Annual Review of Medicine*, *49*(1), 375–390.

Khanna, S., Pardi, D. S., Aronson, S. L., Patricia, P., Orenstein, R., Sauver, J. L. S., … Zinsmeister, A. R. (2012). The Epidemiology of Community-acquired *Clostridium difficile* infection: A population-based study. *American Journal of Gastroenterology*, *107*(1), 89–95.

Kirby, J. M., Ahern, H., Roberts, A. K., Kumar, V., Freeman, Z., Acharya, K. R., & Shone, C. C. (2009). Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of *Clostridium difficile*. *Journal of Biological Chemistry*, *284*(50), 34666–34673.

Kirk, J. A., & Fagan, R. P. (2016). Heat shock increases conjugation efficiency in *Clostridium difficile*. *Anaerobe*, *42*, 1–5.

Klobutcher, L. a, Ragkousi, K., & Setlow, P. (2006). The *Bacillus subtilis* spore coat provides “eat resistance” during phagocytic predation by the protozoan Tetrahymena thermophila. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(1), 165–70.

Knudsen, S. M., Cermak, N., Delgado, F. F., Setlow, B., Setlow, P., & Manalis, S. R. (2016).

Water and small-molecule permeation of dormant *Bacillus subtilis* spores. *Journal of Bacteriology*, *198*(1), 168–177.

Kubiak, A. M., & Minton, N. P. (2015). The potential of clostridial spores as therapeutic delivery vehicles in tumour therapy. *Research in Microbiology*, *166*(4), 244–254.

Kuehne, S. A., Cartman, S. T., & Minton, N. P. (2011). Both, toxin A and toxin B, are important in *Clostridium difficile* infection. *Gut Microbes*, *2*(4), 711–713.

Kuehne, S. A., Collery, M. M., Kelly, M. L., Cartman, S. T., Cockayne, A., & Minton, N. P. (2014). Importance of toxin a, toxin b, and cdt in virulence of an epidemic *Clostridium difficile* strain. *Journal of Infectious Diseases*, *209*(1), 83–86.

Kuehne, S. A., Heap, J. T., Cooksley, C. M., Cartman, S. T., & Minton, N. P. (2011). ClosTron-mediated engineering of clostridium. *Methods in Molecular Biology*, *765*(8), 389–407.

Larson, M., Gilbert, L., Wang, X., Lim, W., Weissman, J., & Qi, L. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*, *8*(11), 2180–2196.

Lawson, P. A., Citron, D. M., Tyrrell, K. L., & Finegold, S. M. (2016). Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. *Anaerobe*, *40*, 95–99.

Lee, W.-T., Wu, Y.-N., Chen, Y.-H., Wu, S.-R., Shih, T.-M., Li, T.-J., … Shieh, D.-B. (2017). Octahedron Iron Oxide Nanocrystals Prohibited *Clostridium difficile* Spore Germination and Attenuated Local and Systemic Inflammation. *Scientific Reports*, *7*(1), 1–12.

Lessa, F., Mu, Y., Bamberg, W., Beldavs, Z., Dumyati, G., Dunn, J., Farley, M., Holzbauer, S., Meek, J., Phipps, E., Wilson, L., Winston, L., Cohen, J., Limbago, B., Fridkin, S., Gerding, D. & McDonald, L. (2015). Burden of *Clostridium difficile* infection in the United States. *New England Journal of Medicine*, *372*, 2369–2370.

Li, Y., Davis, A., Korza, G., Zhang, P., Li, Y. Q., Setlow, B., … Hao, B. (2012). Role of a SpoVA Protein in Dipicolinic Acid Uptake into Developing Spores of *Bacillus subtilis*. *Journal of Bacteriology*, *194*(8), 1875–1884.

Lopetuso, L. R., Scaldaferri, F., Petito, V., & Gasbarrini, A. (2013). Commensal Clostridia: leading players in the maintenance of gut homeostasis. *Gut Pathogens*, *5*(1), 23.

Loshon, C. A., Genest, P. C., Setlow, B., & Setlow, P. (1999). Formaldehyde kills spores of *Bacillus subtilis* by DNA damage and small, acid-soluble spore proteins of the α/β-type protect spores against this DNA damage. *Journal of Applied Microbiology*, *87*(1), 8–14.

Louie, T., Miller, M., Mullane, K., Weiss, K., Lentek, A., Golan, Y., Gorbach, S., Sears, P. & Shue, Y. (2011). Fidaxomicin versus vancomycin for *Clostridium difficile* infection. *New England Jornal of Medicine*, *364*, 422-431.

Lyras, D., O'Connor, J. R., Howarth, P. M., Sambol, S. P., Carter, G. P., Phumoonna, T., … Rood, J. I. (2009). Toxin B is essential for virulence of *Clostridium difficile*. *Nature*, *458*(7242), 1176–1179.

Macke, T., Ecker, D., Gutell, R., Gautheret, D., Case, DA. & Sampath, R. (2001) RNAMotif – A new RNA secondary structure definition and discovery algorithm. *Nucleic Acids Research. 29,* 4724–4735.

Makarova, K. S., Grishin, N. V, Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action, *Biology Direct*, *26*, 1–26.

Marsh, J. W., Arora, R., Schlackman, J. L., Shutt, K. A., Curry, S. R., & Harrison, L. H. (2012). Association of relapse of *Clostridium difficile* disease with BI/NAP1/027. *Journal of Clinical Microbiology*, *50*(12), 4078–4082.

Matamouros, S., England, P., & Dupuy, B. (2007). *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Molecular Microbiology*, *64*(5), 1274–1288.

Mohr, G., Smith, D., Belfort, M., & Lambowitz, A. M. (2000). Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes and Development*, *14*(5), 559–573.

Mojica, F., Diez-Villasenor, C., Garcia-Martinez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, (155), 733–740.

Mojica, F., Diez-Villasenor, C., Garcia-Martinez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, *60*, 174–182.

Na, X., Kim, H., Moyer, M. P., Pothoulakis, C., & LaMont, J. T. (2008). gp96 is a human colonocyte plasma membrane binding protein for *Clostridium difficile* toxin A. *Infection and Immunity*, *76*(7), 2862–2871.

Nawrocki, K. L., Edwards, A. N., Daou, N., Bouillaut, L., & McBride, S. M. (2016). CodY-dependent regulation of sporulation in *Clostridium difficile*. *Journal of Bacteriology*, *198*(15), 2113–2130.

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., … Minton, N. P. (2013). Expanding the Repertoire of Gene Tools for Precise Manipulation of the *Clostridium difficile* Genome: Allelic Exchange Using *pyrE* Alleles. *PLoS ONE*, *8*(2), e56051.

Nicholson, W. L., Munakata, N., Horneck, G., Melosh, H. J., & Setlow, P. (2000). Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiology and Molecular Biology Reviews*, *64*(3), 548–72.

O'Connor, J. R., Lyras, D., Farrow, K. A., Adams, V., Powell, D. R., Hinds, J., … Rood, J. I. (2006). Construction and analysis of chromosomal *Clostridium difficile* mutants. *Molecular Microbiology*, *61*(5), 1335–1351.

Oultram, J. D., Loughlin, M., Swinfield, T.-J. J., Brehm, J. K., Thompson, D. E., & Minton, N. P. (1988). Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. *FEMS Microbiology Letters*, *56*(1), 83–88.

Paidhungat, M., Setlow, B., & Driks, A. (2000). Characterization of Spores of *Bacillus subtilis* Which Lack Dipicolinic Acid Characterization of Spores of *Bacillus subtilis* Which Lack Dipicolinic Acid, *Journal of Bacteriology*, *182*(19), 5505–5512.

Papatheodorou, P., Carette, J., Bell, G., Schwan, C., Guttenberg, G., Brummelkamp, T. & Aktories, K. (2011). Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin *Clostridium difficile* transferase (CDT), *PNAS*, 108(39), 16422-16427.

Paredes-Sabja, D., Setlow, B., Setlow, P., & Sarker, M. R. (2008). Characterization of *Clostridium perfringens* spores that lack SpoVA proteins and dipicolinic acid. *Journal of Bacteriology*, *190*(13), 4648–4659.

Paredes-Sabja, D., Setlow, P., & Sarker, M. R. (2011). Germination of spores of Bacillales and Clostridiales species: Mechanisms and proteins involved. *Trends in Microbiology*, *19*(2), 85–94.

Paredes-Sabja, D. & Barker, M. (2012). Adherence of *Clostridium difficile* spores to Caco-2 cells in culture. *Journal of Medical Microbiology*, *61*, 1208-1218.

Permpoonpattana, P., Tolls, E. H., Nadem, R., Tan, S., Brisson, A., & Cutting, S. M. (2011). Surface layers of *Clostridium difficile* endospores. *Journal of Bacteriology*, *193*(23), 6461–6470.

Popoff, M., & Bouvet, P. (2009). Clostridial toxins. *Future Microbiology*, *4*(8), 1021–1064.

Public Health England. (2015). Voluntary surveillance of *Clostridium difficile*, England, Wales and Northern Ireland: 2014, *Health Protection* Report, *9*(21), 5–12.

Qureshi, N., & Blaschek, H. P. (2001). Recent advances in ABE fermentation: hyper-butanol producing *Clostridium beijerinckii* BA101. *Journal of Industrial Microbiology & Biotechnology*, *27*(5), 287–291.

Raju, D., Waters, M., Setlow, P., & Sarker, M. R. (2006). Investigating the role of small, acid-soluble spore proteins (SASPs) in the resistance of *Clostridium perfringens* spores to heat. *BMC Microbiology*, *6*(1), 50.

Redelings, M. D., Sorvillo, F., & Mascola, L. (2007). Increase in *Clostridium difficile*-related mortality rates, United States, 1999-2004. *Emerging Infectious Diseases*, *13*(9), 1417–1419.

Ridlon, J. M., Kang, D.-J., & Hylemon, P. B. (2006). Bile salt biotransformations by human intestinal bacteria. *Journal of Lipid Research*, *47*(2), 241–259.

Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. (2015). REBASE-a database for DNA restriction and modification: Enzymes, genes and genomes. *Nucleic Acids Research*, *43*(1), 298–299.

Rohlke, F., & Stollman, N. (2012). Fecal microbiota transplantation in relapsing *Clostridium difficile* infection. *Therapeutic Advances in Gastroenterology*, *5*(6), 403–420.

Ross, C., & Ernesto, A.-S. (2010). The Ger Receptor Family from Sporulating Bacteria. *Current Issues in Molecular Biology*, *12*(3), 147–158.

Rupnik, M., & Janezic, S. (2016). An Upate on *Clostridium difficile* Toxinotyping. *Journal of Clinical Microbiology*, *54*(1), 13–18.

Sayedy, L., Kothari, D., & Richards, R. (2010). Toxic megacolon associated *Clostridium difficile*

colitis. *World Journal of Gastrointestinal Endoscopy*, *2*(8), 293–297.

Schafer, A., Kalinowski, J., & Puhler, A. (1994). Increased fertility of *Corynebacterium glutamicum* recipients in intergeneric matings with *Escherichia coli* after stress exposure. *Applied and Environmental Microbiology*, *60*(2), 756–759.

Schantz, E. J., & Johnson, E. A. (1992). Properties and use of botulinum toxin and other microbial neurotoxins in medicine. *Microbiological Reviews*, *56*(1), 80–99.

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., … Parkhill, J. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature Genetics*, *38*(7), 779–786.

Setlow, B., Atluri, S., Kitchel, R., Koziol-Dube, K., & Setlow, P. (2006). Role of dipicolinic acid in resistance and stability of spores of *Bacillus subtilis* with or without DNA-protective α/β-type small acid-soluble proteins. *Journal of Bacteriology*, *188*(11), 3740–3747.

Setlow, B., Loshon, C. A., Genest, P. C., Cowan, A. E., Setlow, C., & Setlow, P. (2002). Mechanisms of killing spores of *Bacillus subtilis* by acid, alkali and ethanol. *Journal of Applied Microbiology*, *92*(2), 362–375.

Setlow, B., McGinnis, K. A., Ragkousi, K., & Setlow, P. (2000). Effects of major spore-specific DNA binding proteins on *Bacillus subtilis* sporulation and spore properties. *Journal of Bacteriology*, *182*(24), 6906–6912.

Setlow, P. (2003). Spore germination. *Current Opinion in Microbiology*, *6*(6), 550–556.

Setlow, P. (2006). Spores of *Bacillus subtilis*: Their resistance to and killing by radiation, heat and chemicals. *Journal of Applied Microbiology*, *101*(3), 514–525.

Smith, L. D., & King, E. O. (1962). Occurrence of *Clostridium difficile* in infections of man.

Journal of Bacteriology, *84*, 65–67.

Sorg, J. A., & Sonenshein, A. L. (2008). Bile salts and glycine as cogerminants for *Clostridium difficile* spores. *Journal of Bacteriology*, *190*(7), 2505–2512.

Sorg, J. A., & Sonenshein, A. L. (2010). Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. *Journal of Bacteriology*, *192*(19), 4983–4990.

Sreekumar, S., Baer, Z. C., Pazhamalai, A., Gunbas, G., Grippo, A., Blanch, H. W., … Toste, F. D. (2015). Production of an acetone-butanol-ethanol mixture from *Clostridium acetobutylicum* and its conversion to high-value biofuels. *Nature Protocols*, *10*(3), 528–537.

Stabler, R., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., … Wren, B. W. (2009). Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biology*, *10*(9), R102.

Stevens, D. L., & Bryant, A. E. (2002). The Role of Clostridial Toxins in the Pathogenesis of Gas Gangrene on JSTOR. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *35*(1), 93–100.

Sylvestre, P., Couture-Tosi, E., & Mock, M. (2002). A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. *Molecular Microbiology*, *45*(1), 169–178.

Tovar-rojo, F., Chander, M., Setlow, B., & Setlow, P. (2002). The Products of the *spoVA* Operon Are Involved in Dipicolinic Acid Uptake into Developing Spores of *Bacillus subtilis*, *Journal of Bacteriology*, *184*(2), 584–587.

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., …

Stephenson, K. (2009). Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *Journal of Bacteriology*, *191*(23), 7296–7305.

Velásquez, J., Schuurman-Wolters, G., Birkner, J. P., Abee, T., & Poolman, B. (2014). *Bacillus subtilis* spore protein SpoVAC functions as a mechanosensitive channel. *Molecular Microbiology*, *92*(4), 813–823.

Vepachedu, V. R., & Setlow, P. (2005). Localization of SpoVAD to the Inner Membrane of Spores of *Bacillus subtilis*. *Journal of Bacteriology*, *187*(16), 5677–5682.

Voth, D. E., & Ballard, J. D. (2005). *Clostridium difficile* Toxins: Mechanism of Action and Role in Disease. *Clinical Microbiology Reviews*, *18*(2), 247–263.

Wang, S., Xu, M., Wang, W., Cao, X., Piao, M., Khan, S., … Wang, B. (2016). Systematic review: Adverse events of fecal Microbiota transplantation. *PLoS ONE*, *11*(8), 1–24.

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., & Blaschek, H. P. (2016). Bacterial Genome Editing with CRISPR-Cas9: Deletion, Integration, Single Nucleotide Modification, and Desirable "clean" Mutant Selection in *Clostridium beijerinckii* as an Example. *ACS Synthetic Biology*, *5*(7), 721–732.

Warda, A. K., Tempelaars, M. H., Boekhorst, J., Abee, T., & Groot, M. N. N. (2016). Identification of CdnL, a putative transcriptional regulator involved in repair and outgrowth of heat-damaged bacillus cereus spores. *PLoS ONE*, *11*(2), 1–16.

Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., … McDonald, L. C. (2005). Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *Lancet*, *366*(9491), 1079–1084.

Wilson, K. H. (1983). Efficiency of various bile salt preparations for stimulation of *Clostridium*

difficile spore germination. *Journal of Clinical Microbiology*, *18*(4), 1017–1019.

Xu, T., Li, Y., Shi, Z., Hemme, C. L., Li, Y., Zhu, Y., … Zhou, J. (2015). Efficient genome editing in clostridium cellulolyticum via CRISPR-Cas9 nickase. *Applied and Environmental Microbiology*, *81*(13), 4423–4431.

Young, S. B., & Setlow, P. (2003). Mechanisms of killing of *Bacillus subtilis* spores by hypochlorite and chlorine dioxide. *Journal of Applied Microbiology*, *95*(1), 54–67.

Yutin, N., & Galperin, M. Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environmental Microbiology*, *15*(10), 2631–2641.

Zeng, X., Ardeshna, D., & Lin, J. (2015). Heat shock-enhanced conjugation efficiency in standard *Campylobacter jejuni* strains. *Applied and Environmental Microbiology*, *81*(13), 4546–4552.

Zhang, P., Kong, L., Setlow, P., & Li, Y. Q. (2010). Characterization of wet-heat inactivation of single spores of Bacillus species by dual-trap raman spectroscopy and elastic light scattering. *Applied and Environmental Microbiology*, *76*(6), 1796–1805.