Spore Germination and Recurrence of *Clostridioides difficile* Infection

by

Lorna Sophia Finch, BSc Honours



Thesis submitted to the University of Nottingham for the degree of

DOCTOR OF PHILOSOPHY

October 2020

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.

Lorna S. Finch

October 2020

Abstract

Clostridioides difficile-associated diarrhoea (CDAD) has emerged as a leading healthcare-associated infection of great clinical and economic significance. The infection is not only characterised by debilitating symptoms but is also complicated with a high incidence of recurrence and is increasingly found in the community. *C. difficile* spores are resistant to many chemical and physical agents, as a result spores shed in faecal matter are very difficult to eradicate, leading to infection, or re-infection, of cohabiting individuals through ingestion of spores from contaminated surfaces. Interventions that target sporulation could, therefore, have an enormous impact on disease transmission. Nonetheless, it is the toxin-producing vegetative cells that subsequently arise following germination that cause disease and so intervention at the stage of germination would be an effective means of disease prevention.

In this study various methods where employed towards the elucidation of ribotype prevalence and spore phenotypes of clinical isolates derived from 151 patients enrolled on a trial to assess the use of Rifaximin as a measure to control relapse of *Clostridioides difficile* infection (CDI). Genomic characterisation of a subset of isolates was performed to differentiate incidence of reinfection from relapse. Pangenome analysis together with genetic variation analysis of 23 key genes implicated in the colonisation, sporulation, germination and outgrowth of *C. difficile* was conducted to elucidate whether genotypic and phenotypic characteristics of isolates could be implicated as significant virulence or risk factors of recurrence in CDAD. A subset of 35 isolates were presented in this research. In this subset, 23 isolates are from 9 patients, who clinically presented with recurrent CDI, and 12 isolates from 6 patients, who clinically presented with non-recurrent CDI. Specific research questions included, whether variations of spore phenotypes and within the genomes among isolates over the infection period of CDI and recurrence could be observed.

To further support associations with data from clinical isolates, sporulation and germination phenotypes where screened by construction and characterisation of insertional ClosTron mutations in CD19300, CD05460, CD07730, CD07750, CD29630, CD08670 and CD24700, seven hypothetical genes, in the non-epidemic *C. difficile* 630Δ*erm* strain. Targets where proposed to be involved in sporulation and germination pathways, particularly CD24700,

encoding the germination protease (GPR). A recombinant GPR was produced to further asses its role along with the crude extraction of acid soluble spore proteins hypothesised to contain the GPR substrates, small acid soluble proteins (SASPs).

It was hypothesised that germination and outgrowth redundancies of strains relate to prolonged spore dormancy and persistence and thus these specific phenotypes would act as an explanation/virulence factor of relapse. Whereas reinfection would be predicted to be a result of competitive outgrowth of each co-infecting isolate, where antibiotic activity would disturb the faster outgrowing more so than the slower counterpart. The data collected in this study showed evidence to support these hypotheses. Interestingly, mutations leading to amino acid changes were detected in the germinant receptor, CspC, proline reductase, PrdB and exosporium glycoproteins, BclA1 and BclA3, for all isolates included in further genetic characterisation, whilst the initiation of spore germination was found to be significantly reduced in comparison to reference strain for all clinical isolates tested, indicating a higher level of control for Taurocholate-mediated germination. Additionally, levels of spore dormancy were detected in all clinical isolates tested, indicating the potential for relapse in all isolates. The inactivation of a hypothetical gene, CD24700, hypothesised to encode a germination protease (GPR), was revealed to have significant phenotypes supporting a clear role in spore viability, outgrowth and dormancy.

Together, this study has contributed and expanded our knowledge of germination and outgrowth of spores of the non-epidemic *C. difficile* 630Δ*erm* strain and of clinical isolates. It is hoped that sustained study will reveal further virulence factors of relapse and targets to prevent and tackle recurrent CDI.

Acknowledgments

To all of those that supported me in this ordeal, thank you! Special mention to my supervisors, Prof. Nigel Minton and Dr Sarah Kuehne and my honorary mentor and friend, who continued to push me to believe in myself even when I wanted to walk away from it all, Dr Sheyda Azimi.

I dedicate this thesis to myself.

Do not focus on how long it took for you to get here.

This is not the end of your life's achievements.

Do not compare yourself to those who did not have the same starting point in life.

You may have the same three letters after your name.

But your journey here was a different one.

Look back at how far you have come.

Now look forward…

Well done babe you did it!

Preface

I. Manuscript Structure

This thesis is divided into six chapters and are structured as follows:

a. The first and second chapters (Introduction, and Materials and Methods) presents an overview about the pathogen *Clostridioides difficile* and methods employed to carry out this research for this thesis, respectively.

b. The third chapter presents the research of 174 samples submitted from 93 in-patients during the clinical RAPID trial to assess the efficacy of Rifaximin to prevent recurrent *C. difficile* infection. In this study a team of five were involved in the bacterial isolation from stools collected at the Biomedical Research Unit at Nottingham Digestive Diseases Centre at Queens Medical Centre, Nottingham, UK. This chapter displays trial outcomes, ribotype prevalence, sequencing, assembly and annotation of isolates. The chapter continues with comparative genomic strategies including the genetic variation of 23 key genes involved in the colonisation, sporulation, germination and outgrowth of *C. difficile* along with the Pan-genomic study of 35 isolates for the elucidation of variables during recurrence. Genome sequence assembly was carried out by a team of three; Dr Emma Stevenson, Dr Patrick Budd and myself. Whilst all annotation and comparative genomic strategies of recurrence cohort isolates were solely my work.

c. The fourth chapter presents the further microbiological investigation of the 35 isolates in regard to sporulation and germination phenotypes. The data of eight strains from two recurrence patients and four strains from four non-recurrent patients are highlighted within this chapter, whilst the remaining strains are referred to in text and presented in the Appendix. Sporulation

assays were conducted on all isolates by a team of five, whilst all germination assays were conducted by myself only.

d. The fifth chapter presents the research of molecular microbiology techniques for gene inactivation of hypothetical proteins involved in germination. Targets were distributed among students and following initial screening, some targets swapped between students. All phenotype screening was solely my work. Focusing with the further complementation and sporulation and germination phenotypic characterisation of the Germination Protease gene mutant. The chapter continues with the characterisation of the generated recombinant germination protease protein as well as the extraction of acid-soluble spore proteins to elucidate the presence of small acid soluble proteins and their possible interaction with the recombinant germination protease protein.

The final chapters present the general discussion, conclusions and future perspectives of the work, followed by an appendix and references section.

List of contents

Declaration

Abstract

Acknowledgements

Preface

Table of contents

Table of figures

List of tables

List of abbreviations

Chapter 1: Introduction

Chapter 2: Materials and Methods

Chapter 3: Genomic Variation of *C. difficile* Isolates During Recurrent Infection

Chapter 4: Spore phenotypes of *C. difficile* Isolates During Recurrent Infection

Chapter 5: Characterisation of ClosTron Mutants and The Germination Protease

Chapter 6: General Discussion

Chapter 7: Appendix

Chapter 8: References

Table of Contents

Chapter 1: Introduction ***24***

1.1 *Clostridioides difficile* in the healthcare setting **24**

1.2 Treatment **24**

1.2.1 Management and therapeutic measures against *C. difficile* infection 24

1.2.2 Therapeutic measures against recurrent *C. difficile* infection 25

1.3 Epidemiology **27**

1.4 Virulence factors **29**

1.4.1 Toxins 31

1.5 Spores of *Clostridioides difficile* **34**

1.5.1 Bacterial spores and resistant properties 34

1.5.2 Sporulation 39

1.5.3 Initiation and regulation of sporulation in *C. difficile* 41

1.5.4 Bacterial spore germination 45

1.5.5 Germination pathways of *C. difficile* spores 45

1.5.6 Methods to detect sporulation, spore germination and outgrowth 52

1.6 Genetic Manipulation of *Clostridioides difficile* **53**

1.7 Molecular typing of *Clostridioides difficile* **54**

1.7.1 Typing methods of *C. difficile* 54

1.7.2 Emergence of 'hypervirulence' 55

1.7.2.1 Hypervirulence in sporulation/germination 57

1.7.2.2 Understanding the genetics of sporulation and germination of *C. difficile* and relapse 58

1.8 Project Aims **59**

Chapter 2: Materials and Methods ***61***

2.1 Selection of clinical isolates **61**

2.1.1 Study design and patient population 61

2.1.2 Randomization and follow-up 61

2.1.3 Efficacy assessments 62

2.1.4 Sampling schedule 62

2.1.5 Isolation of *C. difficile* from stool 63

2.2 Genotypic characterization of clinical isolates. **64**

2.2.1 PCR Ribotyping 64

2.2.1 List of 37 clinical *C. difficile* isolates selected for further characterisation 66

2.2.2 Whole genome sequencing and assembly 67

2.2.3 Genome annotation 67

2.2.4 Comparative genome analysis 67

2.2.5 Pan-genome analysis using the BPGA pipeline 69

2.4 Spore techniques **70**

2.4.1 Preparation of *C. difficile* spores in a nutrient-rich liquid medium 70

2.4.2 Preparation of spores on BHIS agar 70

2.4.3 Purification of spores 70

2.4.4 Measurement of *C. difficile* heat-resistant colony formation 71

2.4.5 Counting *C. difficile* spores by phase-contrast microscopy 71

2.4.6 Assaying the viability of *C. difficile* spores by decoating 71

2.4.7 Measurement of germination initiation 72
2.4.8 Measurement of spore outgrowth 72
2.4.9 Measurement of loss of spore heat resistance 73
2.4.10 Statistical analysis 73
2.4.11 List of other bacterial strains used in the study 74
2.4.12 List of plasmids 75
2.5 Bacterial strains, media and culture conditions 76
2.5.1 Bacterial strains 76
2.5.2 Culture media 76
2.5.3 Plasmids 76
2.5.4 Anaerobic and aerobic culture conditions 77
2.5.5 Preparation of electrocompetent *E. coli* cells 78
2.6 Molecular biology techniques 79
2.6.1 Oligo design 79
2.6.2 Plasmids 80
2.6.3 Sequence Checking 80
2.6.4 Restriction Endonuclease digestion 80
2.6.5 De-phosphorylation of DNA 80
2.6.6 Ligation 81
2.6.7 Membrane dialysis of DNA 81
2.6.8 Extraction of chromosomal DNA 81
2.6.9 Quantification of DNA 82
2.6.10 Amplification and assembly of DNA 82
2.6.11 Splicing by overlap extension PCR (SOE-PCR) 83
2.6.12 Agarose gel electrophoresis 84
2.6.13 Capillary electrophoresis 84
2.6.14 DNA purification 84
2.6.15 Extraction of plasmids from *E. coli* cultures 84
2.6.16 Plasmid transformation of *E. coli* 85
2.6.17 Conjugation 85
2.6.18 Southern blot 86
2.7 Integration of DNA into the genome of *C. difficile* 87
2.7.1 Generation of CD05460, CD07730, CD07750, CD08670, CD19300, CD24700 and CD29630 mutants using ClosTron 87
2.7.2 Isolation of single and double cross-over clones using Allelic exchange 88
2.8 Recombinant and native protein analysis 90
2.8.1 Inducing recombinant protein expression 90
2.8.2 Purification of 6xHis-tagged proteins under native conditions 91
2.8.3 Separation of proteins by SDS-PAGE 91
2.8.4 Coomassie staining of PAGE gels 92
2.8.5 Western Blot 92
2.8.6 Isolation of acid-soluble proteins from *C. difficile* endospores 93
2.8.7 Separation of native basic proteins by cathodic PAGE 93
2.8.8 Silver staining of PAGE gels 95
2.8.9 Protein quantification by BCA 96
2.8.10 Zymogen autoprocessing assay 96
2.8.11 Assay to determine the protection of DNA from DNAse treatment 96
2.8.12 Assay to determine GPR proteolysis of SASPs 97
Chapter 3: Genomic Variation of C. difficile Isolates During Recurrent Infection 99
3.1 Introduction 99

3.2 Results **102**
3.2.1 Bacterial isolates 102
3.2.2 Ribotype prevalence of isolates 103
3.2.3 Genomic diversity of isolates 108
3.2.4 Sequence variation of genes of *C. difficile* associated with spore germination and outgrowth 114
3.2.5 Pan-genome analysis 128
3.3 Discussion **140**
3.3.1 PCR Ribotype prevalence and its association to CDI recurrence 140
3.3.2 Sensitivity of PCR ribotype compared to SNP analysis to classify recurrence 140
3.3.3 Genomic diversity of isolates 142
3.3.4 Sequence variation of genes of *C. difficile* associated with spore germination and outgrowth and implications on recurrence 142
3.3.5 Pan-genomes of *C. difficile* and implications on recurrence 146
3.3.6 Rifaximin as a chaser therapy for 'relapse' 147
3.3.7 Limitations and benefits of this study 149
3.3.8 Conclusion and future implications 151
Chapter 4: Spore Phenotypes of C. difficile Isolates During Recurrent Infection ...153
4.1 Introduction **153**
4.2 Results **155**
4.2.1 Spore phenotypes of isolates from patients 01008, 11007 and 03008 treated with rifaximin or placebo for the prevention of CDI, presented as case studies 1, 2 and 3 155
4.2.2 Collated sporulation phenotypes of 37 isolates from 17 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence 169
4.2.3 Collated initiation of germination phenotypes of 20 isolates from 11 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence 173
4.2.4 Collated outgrowth phenotypes of 12 isolates from 6 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence 177
4.2.5 Collated spore dormancy phenotypes of 14 isolates from 7 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence 179
4.3 Discussion **181**
4.3.1 Sporulation in recurrent and non-recurrent CDI 181
4.3.1.1 Limitations and advancement of knowledge in the study of sporulation: Rates of sporulation cannot be assessed without spore titres 182
4.3.2 Rate of germination in recurrent and non-recurrent CDI 184
4.3.2.1 Limitations and advancement of knowledge in the study of spore germination: Significantly reduced rates of germination of clinical isolates compared to control strain CD630Δ*erm*, could be linked to mutations in regulators, *cspA* and csp*B,* of TA- dependant germination 187
4.3.3 Outgrowth of spores in recurrence 190
4.3.3.1 Limitations and advancement of knowledge in the study of spore outgrowth: Outgrowth from spores showed strain-to-strain variation with competitive outgrowth implied for co-infecting isolates 190
4.3.4 Loss of spore heat resistance in recurrence 191
4.3.4.1 Limitations and advancement of knowledge in the study of spore dormancy: Loss of heat resistance assays reveal a comparable level of spore dormancy for all isolates 192
4.3.5 Final Implications and conclusions 193
Chapter 5: Characterisation of ClosTron mutants reveal germination protease as a critical factor for spore viability 196
5.1 Introduction **196**

5.2 Results **200**
5.2.1 Mutant generation 200
5.2.2 Colony formation of heat-resistant spores of ClosTron mutant strains 203
5.2.3 Effect of ClosTron mutations on *C. difficile* sporulation 204
5.2.4 Effect of mutations on *C. difficile* spore viability 206
5.2.5 Complementation of *gpr* ClosTron mutant 208
5.2.6 Chromosome level complementation and overexpression studies in CT24700 mutant 215
5.2.7 Role of GPR in *C. difficile* spore germination and outgrowth 218
5.2.8 Recombinant GPR production 226
5.2.9 Identifying Small Acid Soluble Proteins of *C. difficile* 230
5.2.10 Assessing the proteolytic activity of recombinant GPR and the extent of DNA protection from DNAse acid soluble spore proteins 234
5.3 Discussion **237**
5.3.1 Key outcomes of ClosTron mutagenesis 237
5.3.2 Role of germination protease in *C. difficile* spore germination and outgrowth 238
5.3.3 Activity of recombinant GPR and SASPs in *Clostridioides difficile* 243
5.3.4 Concluding remarks 243
Chapter 6: General Discussion ***246***
6.1 Introduction **246**
6.2 Key findings, limitations and proposed follow-on work from this study **247**
6.3 Insights into the study of *C. difficile* spore germination, outgrowth and CDI recurrence: A final commentary on hypothesis and study methods **252**
Chapter 7: Appendix ***256***
Chapter 8: References ***272***

Table of Figures

Figure 1.1: Mortality rates for deaths mentioning *Clostridioides difficile* in England and Wales, registered between 2002 and 2012 from PHE (2013). 28

Figure 1.2: Rates in *C. difficile* infection in England from PHE mandatory surveillance 2017/2018 (PHE 2018). 28

Figure 1.3: Proposed model for establishment of *C. difficile*-mediated dysbiosis and successful bacteriotherapy from Lawley *et al.,* (2012). 30

Figure 1.4: Toxin loci of *C. difficile* from Rupnik, M *et al.,* (2009). 32

Figure 1.5: Structure of bacterial spores, from Kochan *et al.,* (2018). 35

Figure 1.6: Lifecycle of Endospore Formers from Shen *et al.,* (2019). 38

Figure 1.7: Stages of the sporulation process and structure of *C. difficile* spore from Zhu *et al.,* (2018). 39

Figure 1.8: Spore coat and exosporium structure in *C. difficile* from Shen *et al.,* (2019). 41

Figure 1.9: Regulatory pathway controlling Spo0A activation in *C. difficile* Shen *et al.,* (2019). 43

Figure 1.10: Proposed models of *Clostridioides difficile* germination from Kochan *et al.,* (2018). 50

Figure 1.11: Distribution of *Clostridioides difficile* ribotypes in England by quarter from April 2008 – March 2018 from PHE (2018). 55

Figure 1.12: Clades of *C. difficile* as determined by whole genome MLST from Ramirez-Vargas *et al.,* (2018). 56

Figure 2.1: Splicing by overlap extension PCR. 83

Figure 2.2: Screening for a ClosTron mutant. 88

Figure 3.1: Prevalence of PCR ribotypes in study. 104

Figure 3.2: Temporal graph of 23 recurrent *C. difficile* infection causing isolates from nine patients. 107

Figure 3.3: Phylogenetic tree of 15 *C. difficile* strains within the recurrent cohort. 112

Figure 3.4: A heatmap of average nucleotide identity (ANI) between genomes of the recurrent cohort. 113

Figure 3.5: The frequency of amino acid changes found in the strains of the recurrence cohort for 24 genes selected for their associated with growth, and spore germination in reference *C. difficile* strain CD630. 116

Figure 3.6: The distribution of amino acid changes found amongst 15 strains in the recurrence cohort for 16 genes of interest of reference *C. difficile* strain CD630.....117

Figure 3.7: MUSCLE alignment of CD630_03320 (*blcA1*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.. 122

Figure 3.8:MUSCLE alignment of CD630_08570 (*oppF*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.. 123

Figure 3.9: MUSCLE alignment of CD630_22470 (*cspBA*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.. 124

Figure 3.10: MUSCLE alignment of CD630_32410 (*prdB*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE. 125

Figure 3.11: MUSCLE alignment of CD630_33490 (*bclA3*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE. 126

Figure 3.12: MUSCLE alignment of CD630_22460 (*cspC*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE. 127

Figure 3.13: Core and Pan-genome analysis for 16 genomes from clinical isolates of *C. difficile*. 129

Figure 3.14: Number of new gene families revealed in the Pan-genome analysis of 16 isolates of *C. difficile* using the BPGA pipeline. 130

Figure 3.15: Major Cluster of Orthologus Groups (COG) categories of Pan-genome analysis of 16 isolates of *C. difficile*. 133

Figure 3.16: Detailed Cluster of Orthologus Groups (COG) distribution of Pan-genome analysis of 16 isolates of *C. difficile.* 134

Figure 4.1: Spore production after 24 hours is increased for clinal isolates of case studies 1-3 compared to laboratory control strains. 157

Figure 4.2: Spore production after 48 and 72 hours for clinal isolates of case studies 1-3. .158

Figure 4.3: Spore production after 96 and 120 hours for clinal isolates of case studies 1-3. 159

Figure 4.4: Initiation of spore germination of clinical isolates of case studies 1- 3 is significanly slower to that of laboratory control strains. 161

Figure 4.5: Spore outgrowth of isolates of case studies 1 and 2. 165

Figure 4.6: Loss of spore heat-resistance for isolates from case studies 1 and 2. 167

Figure 4.7: Sporulation rates at 0, 24 and 48 hrs for 37 isolates from patients enrolled of the trial of rifaximin chaser therapy for the prevention of CDI recurrence. 171

Figure 4.8: Sporulation rates after 72, 96 and 120 hrs for 37 isolates from patients enrolled of the trial of rifaximin chaser therapy for the prevention of CDI recurrence. 172

Figure 4.9: Initiation of spore germination rates in isolates from patients of the recurrence cohort. 174

Figure 4.10: Initiation of spore germination rates in isolates of four patients with non-recurrent CDI. 176

Figure 4.11: Spore outgrowth in in isolates from patients of the recurrence cohort. 178

Figure 4.12: Loss of heat resistance in isolates from patients of the recurrence cohort. 180

Figure 5.1: Colony PCR screening for putative *C. difficile* 630Δ*erm* ClosTron insertion mutants. 201

Figure 5.2: Southern blot analysis of ClosTron mutants using an intron specific probe. 202

Figure 5.3: Heat-resistant CFU development of *C. difficile* 630Δ*erm* mutants over a 5-day incubation period in BHIS broth. 203

Figure 5.4: Numbers of heat-resistant CFU and spore titres after five days incubation in BHIS broth. 205

Figure 5.5: Viability of *C. difficile* 630Δ*erm* and mutant spores after the five-day incubation period in BHIS broth.....207

Figure 5.6: Heat-resistant CFU development of plasmid based complemented CT24700 mutant over a five-day incubation period in BHIS broth.....209

Figure 5.7: Spore titres of the plasmid based complemented CT24700 mutant after five days incubation in BHIS broth. 210

Figure 5.8: Viability of *C. difficile* 630Δ*erm* and plasmid complemented CT24700 spores after five incubation period in BHIS broth.....211

Figure 5.9: PCR digest for the confirmation of vector construction for complementation and over expression in the chromosome of CT24700 genome. 213

Figure 5.10: PCR confirmation of chromosomal complementation and overexpression via ACE repair of CT24700. 214

Figure 5.11: Heat-resistant CFU development of chromosome based complementation and overexpression of CT24700 mutant over a 5-day incubation period in BHIS broth. 216

Figure 5.12: Spore viability of chromosome based complementation and overexpression of CT24700 mutant. 217

Figure 5.13: Spore germination of CT24700, chromosome based complementation and overexpression strains. 220

Figure 5.14: Spore outgrowth of CT24700, chromosome based complementation and overexpression strains. 221

Figure 5.15: Loss of heat resistance in spores of CT24700, chromosome based complementation and overexpression strains. 223

Figure 5.16: Percentage spore germination of CT24700, chromosome based complementation and overexpression strains. 224

Figure 5.17: Verification of pET28b+ plasmid carrying protein coding sequence of CD24700. 227

Figure 5.18: Recombinant GPR expression in pET28b(+) vectors and *E. coli* BL21(DE3*) host. 228

Figure 5.19: Ni-NTA column His-Tag purification of recombinant GPR protein. 229

Figure 5.20: Lysozyme controls for the validation of Cathodic PAGE gel to visualise basic proteins. 232

Figure 5.21: Crude acid-soluble spore protein extracts visualised on cathodic PAGE gel.. ..232

Figure 5.22: Acid-soluble spore protein extracts concentrated by YMCO-3000 filter centrifugation visulaised on cathodic PAGE gel.. 233

Figure 5.23: Auto-processing of recombinant GPR and its proteolytic activity on acid soluble spore protein extracts visualised by cathodic PAGE. 235

Figure 5.24: No protection of DNA by acid-soluble spore protein extracts from DNASE I treatment. 236

Figure 7.1: Flowchart of phenotype characterisation of *C. difficile* spores. 260

List of Tables

Table 1.1: Germination pathways of *Clostridioides difficile* adapted from Kochan *et al.,* (2018). 48

Table 2.1: A schematic of stool sample submission to the RAPID trial study 64

Table 2.2: Ribotyping master mix composition 66

Table 2.3: Ribotyping PCR cycle 66

Table 2.4. Clinical *C. difficile* isolates used in this study 67

Table 2.5: List of other bacterial strains used in this study 75

Table 2.6: List of plasmids used in this study.....76

Table 2.7. Antibiotic and growth supplements.....78

Table 2.8: List of oligonucleotides used in this study 80

Table 2.9: Thermocycling conditions for a routine PCR using NEB polymerase enzymes 83

Table 2.10: Composition of Cathodic PAGE gel, loading and running buffers 95

Table 2.11: Composition of Silver staining solutions.....96

Table 3.1: Summary of positive *Clostridioides difficile* isolation from stool samples. 103

Table 3.2: Summary of patient and genomic data of *C. difficile* isolates selected for further variant analysis in this study. 110

Table 3.3: Characteristics of 24 genes of reference *C. difficile* strain CD630 (GenBank Assembly ASM920v2) selected for their association with sporulation, germination, growth and antibiotic resistance used to compare the sequence variation of isolates 116

Table 3.4 : Amino acid changes for genes of interest from reference *C. difficile* strain CD630 exclusively observed in isolates 01008 wk-1 (E10) and 01008 wk 12 (G2) and 01016 wk-1 and 01016 wk 12. 122

Table 3.5: Core, accessory, unique and exclusively absent genes of 16 clinical isolates of *C. difficile* after pan-genome analysis using the BPGA pipeline. 132

Table 3.6: Results of a tblastn search querying 16 unique genes of clinical *C. difficile* isolate 01008_H4 against the NCBI database. 136

Table 3.7: Results of a tblastn search querying 11 unique genes of clinical *C. difficile* isolate 01013_wk12 against the NCBI database. 138

Table 3.8: Results of a tblastn search querying 13 unique genes of clinical *C. difficile* isolate 01018_wk-1 against the NCBI database. 139

Table 4.1: Case studies 1 – 3 of isolates from patient 01008, 11007 and 03008, respectively, treated with rifaximin or placebo chaser therapy for the prevention of recurrent CDI 157

Table 4.2: Statistics of the slope from 0 to 9 minutes of the initiation of germination for isolates of Case studies 1-3 163

Table 5.1: Statistics of the slope from 0 to 10 minutes of the initiation of germination for *gpr* inactivation, complementation and overexpression strains. 220

Table 5.2: SASPs of *C. difficile* 231

Table 7.1: Composition of *C. difficile* Minimal Medium 257

Table 7.2: Composition of cationic stacking and resolving polyacrylamide solutions 258

Table 7.3: Variant detection analysis provided by Dr Patrick Budd of 27 genomes of clinical *C. difficile* isolates, used to corroborate the findings in this study. 261

Table 7.4: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient. 262

Table 7.5: Percentage nucleotide sequence identity of 15 clinical isolates for 24 genes of interest from *C. difficile* reference strain CD630 267

Table 7.6: Number of amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates 268

Table 7.7: Amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates 269

Table 7.8: Statistics of the slope from 0 to 9 minutes of the initiation of germination for isolates of *C. difficile* 272

List of abbreviations

ACE - Allelic coupled exchange
ACG - American College of Gastroenterology
ATCC - American Type Culture Collection
ANI - Average nucleotide identity
BPGA - Bacterial Pan Genome Analysis
BCA - Bicinchoninic acid assay
BHI - Brain heart infusion
BHIS - Brain heart infusion supplemented
Blast - Basic alignment sequence tool
BLASTn – Nucleotide BLAST
BLASTp – Protein BLAST
CDCA- chenodeoxycholate
CDMM - *Clostridioides difficile* Minimal media
CDT - *C. difficile* binary toxin
CC – cefoxitin/cycloserine supplement
CCEY - Egg yolk cefoxitin cycloserine
CDAD -*Clostridioides difficile*-associated diarrhoea
CDI- *Clostridioides difficile* infection
CE -Capillary electrophoresis
CDRN - *Clostridioides difficile* Ribotyping Network
COG - Cluster of Orthologous groups
CROPs - Combined repetitive oligopeptides
CRG - Clostridium Research Group
Ca-DPA - Dipicolinic acid chelated with calcium
DDT - Dithiothreitol
dH_2O - Deionised water
DCO - Double cross-over
DNA- Deoxyribonucleic acid
dNTP - Deoxy nucleoside triphosphate
DPA - Dipicolinic acid
EBS – Exon-binding site
ESCMID - European Society of Clinical Microbiology and Infectious Diseases
FMT - Faecal microbiota transplant
FDA - Food and drug association
FOA - Fluoroorotic acid
GAM – Gifu anaerobic broth
GPR - Germination protease
H_2O - Hydrogen dioxide
HCl - Hydrogen chloride
HR-CFU - Heat-resistant colony-forming unit
HRP - Horseradish peroxidase (HRP)
HAI - Hospital acquired infection
ID - Identity
IDSA - Infectious Diseases Society of America
IBS - Irritable bowel syndrome
IPTG - Isopropyl-ß-D-thiogalactoside
LB – Luria/lysogeny broth
MALDI-TOF - Matrix-assisted laser desorption ionization time-of-flight mass spectrometry
$MgCl_2$ - Magnesium chloride
MW - Molecular weight
MWCO - Molecular weight cut off
MLEE - Multi-locus electrophoresis
MNV - Multiple nucleotide variant
N_2 - Nitrogen
NAP1 - North American pulsed field gel electrophoresis type 1
PAGE - Polyacrylamide gel electrophoresis
PBS - Phosphate-buffered saline
PCR - Polymerase chain reaction
PFGE- Pulse field gel electrophoresis
PG - Peptidoglycan
PHE - Public Health England
PMC - Pseudomembranous colitis

PPI - Proton-pump inhibitor
REA - Restriction endonuclease analysis
RL- Relapse
rRNA - Ribosomal ribonucleic acid
RT- Ribotype
SASPs - Small acid soluble proteins
SCFA - Short chain fatty acids
SDS - Sodium doceyl sulphate
SNP - Single nucleotide polymorphism
SNV - Single nucleotide variant
SSC - Saline-sodium-citrate
SOE-PCR - Splicing by overlap extension PCR
SOC - Super optimal broth with catabolite repression medium
NaOH - Sodium hydroxide
RAM - Retrotransposition-activated marker
TA - Taurocholate
TAE – Tris-acetate-EDTA
tblastn – Translated nucleotide BLAST
TBS - Tris-Buffered Saline
TCA - Taurocholate acid
TM - Thiamphenicol
TMB - Tetramethylbenzidine
tRNA - Transfer ribonucleic acid
UK- United Kingdom
VRE - Vancomycin-resistant *enterococci*
WGS - Whole-genome sequencing
WT - Wild type
Wk - week

Symbols and Units of measurement

α -Alpha
β-Beta
Δ-Delta
λ -Lambda
Ω - Ohm
A-Amps
bp-Basepair
CI- Confidence interval
CFU - colony forming unit
°C - degrees centigrade
g - Gram
g - gravitational force
h- Hour
kb-Kilobases
kDa - Kilodaltons
L-Litre
nm - Nanometre
mL-Millitre
µL Micro litre
µm- Micro metre
mM-Millimolar
µg-Micro gram
mins-Minutes
M - Molar
mA – Milli amp
M bp- Million basepairs
OD- Optical density
P - P value
% - Percentage
RCF – Relative centrifugal force
RPM- Rotations per minute
R^2 – R-squared
s - seconds
T0 - Time point 0 hours
wt/vol (w/v) – Weight/volume
V- Voltage
v/v- Volume/volume

Nucleic and Amino acids

Single Letter Code	Nucleobase
A	Adenine
C	Cytosine
G	Guanine
T	Thymine

Name	Three Letter Code	Single Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic Acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1

Chapter 1: Introduction

1.1 *Clostridioides difficile* in the healthcare setting

Clostridioides difficile, a spore forming Gram-positive obligate anaerobe, has been identified as a key mediator of antibiotic-associated diarrhoea (Poutanen and Simor, 2004). Briefly, the administration of high-risk antibiotics, that diminish the levels and diversity of "normal" gut microbiota within the colon, allows for the overgrowth of antibiotic resistant *C. difficile* that outgrow from ingested spores. These vegetative cells can release toxins (notably Toxin A and Toxin B) causing diarrhoea and other mild symptoms. Whilst carriage can be asymptomatic, especially in neonates, symptomatic *C. difficile* infection (CDI) can progress with serious complications ranging from toxic megacolon to fulminant pseudomembranous colitis (Poxton *et al.,* 2001). Recurrent CDI can result from later infection with the same or a different strain of *C. difficile* (Walters *et al.,* 1983; Young and McDonald, 1986; Wilcox *et al.,* 1998) and usually present within one to two months after discontinuing antibiotic therapy. Predisposing host factors and circumstances affecting the frequency and severity of disease include advanced age, underlying illness, recent hospitalisation or surgery, and recent administration of drugs that alter bowel motility, including antibiotics or antineoplastic agents with antibacterial activity (Eze *et al.,* 2017).

1.2 Treatment

1.2.1 Management and therapeutic measures against *C. difficile* infection

General principles of managing CDI, include the withdrawal of inducing antibiotics and other predisposing medications as soon as possible (Youngster *et al., 2*014; Goldstein *et al., 2*015) and supportive care such as fluid and electrolyte balance. Interventions led by infection control teams consist of isolation procedures and cleaning regimes to clear environmental spores in order to prevent further transmission (Surawicz *et al., 2*013). Clinical assessment of the severity of disease will dictate the therapeutic options offered to the patient. This can range from alternative antimicrobial usage, antibiotic therapy, to the addition of intravenous

immunoglobulin and gastro-intestinal surgical procedures (Hatton *et al., 2*011). For over 30 years, metronidazole or vancomycin have been the most commonly prescribed drugs for CDI, and until recently remained the first drugs of choice (Peng *et al.,* 2018; Daniels and Kufel, 2018).

1.2.2 Therapeutic measures against recurrent *C. difficile* infection

Recurrent CDI is defined by the complete resolution of presenting symptoms on appropriate therapy, with subsequent 'clinical' relapse and return of symptoms within eight weeks of the initial episode completion of treatment (Dubberke *et al., 2*007). Studies have shown that up to one-half of recurrent episodes are reinfections rather than relapses of infection with the original strain (Barbut *et al.,* 2000; Norén *et al.,* 2004). Recurrence rates of up to 36% have been reported among initially treated patients (Marsh *et al., 2*012) and can be caused by the original infecting strain (relapse) or by another strain (re-infection). This rate increases to 40-65% for subsequent episodes of recurrence (Kyne *et al., 2*002). Studies have estimated that 20%-50% of the reoccurring infections are caused by a new *C. difficile* organism, suggesting re-infection rather than relapse (Johnson *et al., 1*989; Avbersek *et al., 2*009). Recommendations by the Infectious Diseases Society of America (IDSA) and ESCMID are that the treatment of the first recurrence be the same as that at the initial episode and usage based on disease severity (Debast, Bauer, and Kuijper 2014). However, in most instance's recurrences are conventionally treated with prolonged courses of vancomycin (Rupnik *et al.,* 2009). Studies have shown that fidaxomicin has superior efficacy compared to vancomycin in reducing risk of disease recurrence and additional superiority in cure rates of high risk patients (Nathwani *et al., 2*014; Comely *et al., 2*012). Fidaxomicin is seen to be cost effective in such cases attributing to the reduced recurrence rates and beneficial implications for infection control; reducing further transmission of spores (Nathwani *et al., 2*014). Fidaxomicin is recommend by Public Health England (PHE) for patients thought to be at higher risk of a recurrence, such as those prescribed with concomitant antibiotics as a result of having multiple co-morbidities (Dutilh *et al.,* 2013).

Another option being explored is the synthetic antimicrobial, rifaximin, which has been used as a 'chaser therapy', following initial treatment of CDAD with vancomycin (Venugopal A. and Johnson S 2012) and has been shown to be successful in the management of multiple recurrences (Johnson *et al.,* 2007, 2009). The RNA synthesis inhibitor, derived from rifamycin (Lamanna and Orsi, 1984), has activity against both gram-positive and gram-negative aerobic and anaerobic bacteria and has been useful in a number of infections including traveller's diarrhoea, irritable bowel syndrome (IBS), Crohn's disease, ulcerative colitis and hepatic encephalopathy (Huang and DuPont, 2005). Whilst having a broad spectrum, studies in the use of rifaximin in colitis showed little overall impact on the faecal microbiota (Bajaj *et al.,* 2013; Brigidi *et al.,* 2002; Soldi *et al.,* 2015). The drug has been shown to be highly active against most strains of *C. difficile,* thought to be a result of minimal absorption occurring after ingestion, resulting in the drug being highly concentrated in the gut lumen (Major *et al., 2*018). Moreover, bile increases solubility of the drug, so concentrations should be maximal in the of *C. difficile* germination (Darkoh *et al.,* 2010), thus, rifaximin should suppress *C. difficile* proliferation while commensals recover to re-establish a competitive environment hostile to the pathogen (Major *et al., 2*018). Evenso, given all its theoretical advantages, rifaximin is still subject to the problems of other rifamycins, whereby a critical amino acid substitution in RpoB, the β-subunit of bacterial RNA polymerase, leads to resistance, as has been reported in clinical studies (O'Connor *et al.,* 2008).

Bacteriotherapy also commonly known as faecal microbiota transplant (FMT) is also currently considered in treatment of recurrent infections (PHE, 2013) and refractory recurrences in combination with oral vancomycin (Debast, Bauer, and Kuijper 2014). Less diversity in colonic microbiota in patients with recurrent CDI compared to healthy individuals has also been demonstrated in some studies (Khoruts *et al., 2*010). FMT aims to restore the gut microbiome and function by providing colonization resistance and was first described for use in pseudomembranous colitis (PMC) in 1958 (Eiseman *et al.,* 1958) but was not approved for use in England until 2014 (Massey *et al., 2*014; Kyne *et al., 2*002).

1.3 Epidemiology

Major outbreaks of CDI have been reported in healthcare settings worldwide, including North America (Loo *et al.,* 2005; McDonald *et al.,* 2005), the United Kingdom (O'Connor *et al.,* 2009) and continental Europe (Kuijper *et al.,* 2008) and have led to patient isolation, ward closures, and in some cases, hospital closure. It has been calculated that the annual cost of CDI to an average- sized district general hospital is approximately £400,000, which includes 2100 lost bed days (Wilcox, 1998). On average, a patient with a hospital acquired infection (HAI) will spend 2.5 times longer in hospital than an uninfected patient (Inweregbu *et al.,* 2005). It was reported that patients with a HAI stayed significantly longer than controls, averaging a total of 21.3 days in hospital, including an average of 14 days in a side room (Wilcox *et al.,* 1996). Rapid increases in the number of cases of positive *C. difficile* samples across England and Wales, since the introduction of voluntary reporting scheme by the Health Protection agency in 1990, resulted in this surveillance becoming mandatory in the UK in 2004. Reports continued to show dramatic increases in the incidence over the years, with mortality rates increasing from <2000 to >8000 between 2002 and 2007, respectively (see Figure 1.1). The drastic decrease in the mortality rates since 2008 has been attributed to the advances in the reporting system, infection control, and diagnosis and treatment turnaround. Despite a sustained decrease in incidence of reported cases (see Figure 1.2), from 18,022 to 14,687 between the financial year of 2012 and 2013 (Public Health England, 2013), in 2012 (at the start of this study) there were 1,646 deaths involving CDI in England and Wales (Public Health England, 2013). While 89.7% of deaths involving CDI during 2010 to 2012 occurred in NHS hospitals (Office for National Statistics, 2013). Such statistics highlighted that there was more that could be done to reduce the transmission, discomfort, and for some, death as a result of this infection.

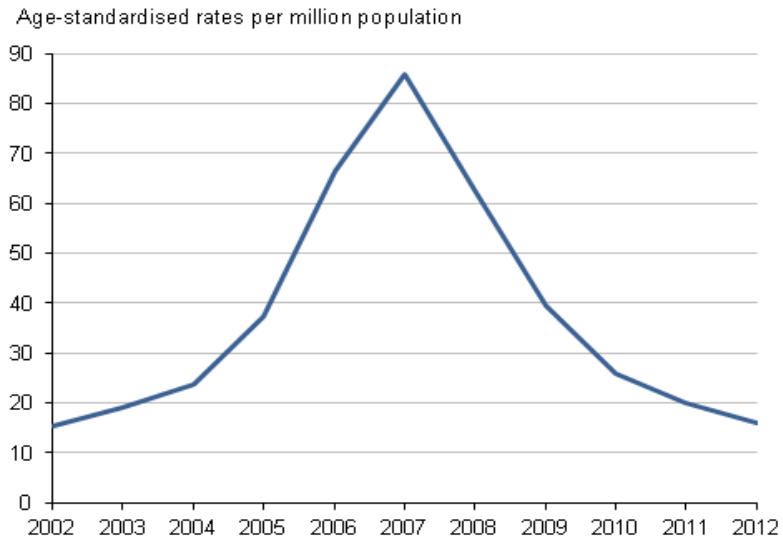


Figure 1.1: Mortality rates for deaths mentioning *Clostridioides difficile* in England and Wales, registered between 2002 and 2012 from PHE (2013).

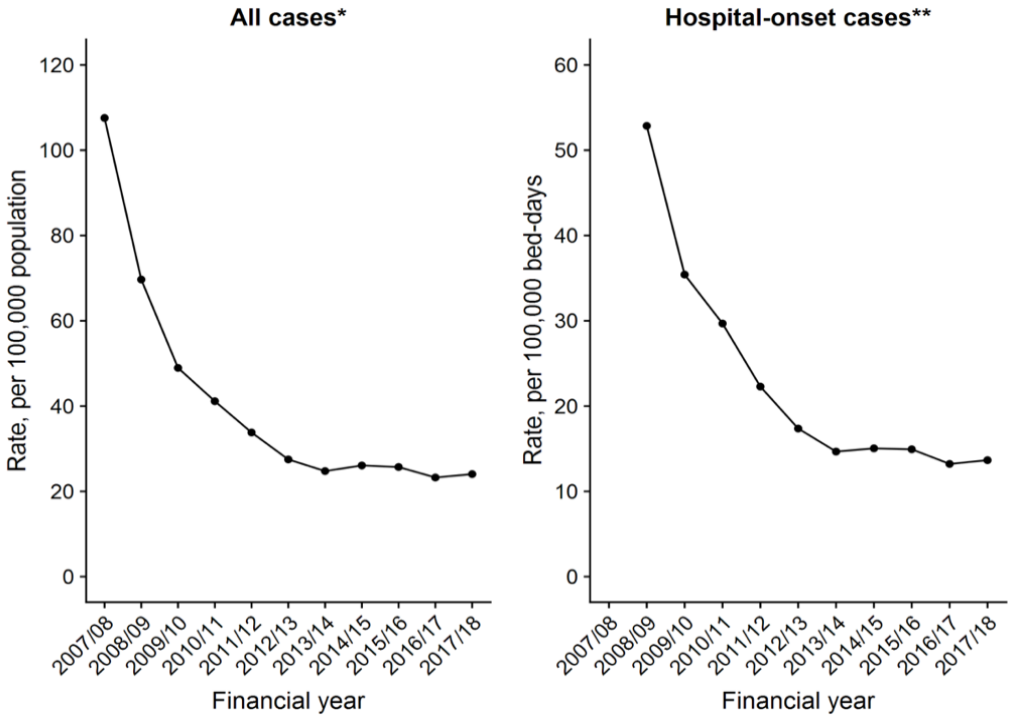


Figure 1.2: Rates in *C. difficile* infection in England from PHE mandatory surveillance 2017/2018 (PHE 2018).

1.4 Virulence factors

The pathogenesis of CDI begins with the disruption of colonic biota, also known as dysbiosis (Figure 1.3), by antibiotics allowing the overgrowth of *C. difficile*; a commensal bacterium of the human intestine present in ~4% of the population (Vernacchio *et al.,* 2006; Merida *et al.,* 1986). Colonization occurs by faecal-oral transmission, spread on the hands of healthcare staff and other people who come in to contact with colonized patients, or with environmental surfaces contaminated with spores. Virulence factors involved in colonisation and adherence include the surface layer proteins of vegetative cells (Drudy *et al.,* 2004; Ausiello *et al.,* 2006; Rupnik *et al.,* 2009) and the FliC and FliD components of *C. difficile* flagella (Tasteyre *et al.,* 2001). While the acidic environment of the stomach kills the ingested vegetative cells, acid resistant spores can pass through the bile acid and convert into vegetative forms in the anaerobic conditions of the bowel (Wilson, 1993; Wilson *et al.,* 1985).

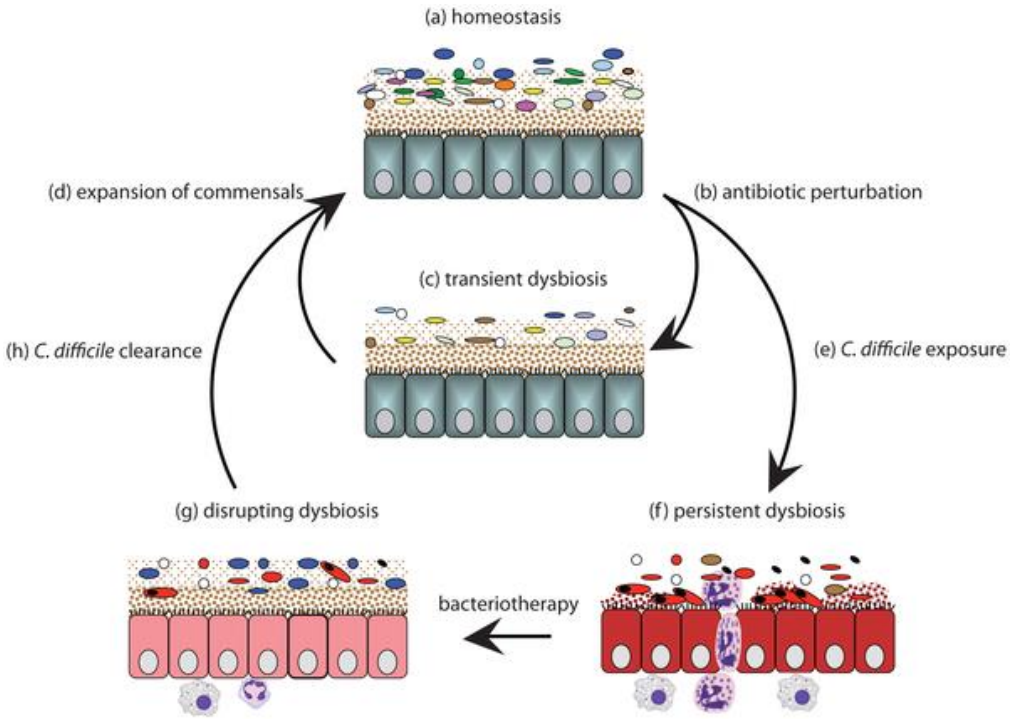


Figure 1.3: Proposed model for establishment of *C. difficile*-mediated dysbiosis and successful bacteriotherapy from Lawley *et al.,* (2012).

Intestinal homeostasis (a) is characterized by a diverse, stable microbiota that produces short chain fatty acids (SCFA); the end products of bacterial fermentation in the intestines that serve as important nutrients for the host (Macfarlane and Macfarlane, 2003). Antibiotic disturbance (b–c) kills susceptible bacteria resulting in a simplified community structure (and reduced SCFA production) and a loss of colonization resistance. In the absence of opportunistic infection, the microbiota is generally restored and diversity and SCFA production resolve (d) to re-establish homeostasis and colonization resistance (a). However, exposure to *C. difficile* (e) after antibiotic disruption (b) can lead to persistent dysbiosis (f) that is characterized by a pathogenic microbial community, reduced SCFA and pathology. Bacteriotherapy disrupts the microbial imbalances (g) leading to the clearance of *C. difficile* (h) and re-establishment of intestinal homeostasis (a). This model is supported by reports of recurrent *C. difficile* disease in humans being associated with a general reduction in intestinal bacterial diversity (Chang *et al.,* 2008).

1.4.1 Toxins

Pathogenic *C. difficile* releases large exotoxins; TcdA (toxin A) and TcdB (toxin B) (von Eichel-Streiber, Sauerborn, and Kuramitsu, 1992) and a third binary toxin known as *C. difficile* toxin (CDT) (Cartman *et al.,* 2010) (see Figure 1.4). Once in the cytoplasm of the lumen, these toxins inactivate Rho proteins, which are responsible for various signal transduction processes and the regulation of the actin cytoskeleton of cells (Hippenstiel *et al., 1*998). The ultimate disarrangement of the cytoskeleton leads to apoptosis. This allows for intestinal fluid secretion, due to diffuse lymphatic infiltrate in the lamina propia, which leads to mucosal injury including oedema and bulging, and the eventual separation of the basal portions and opening of epithelial cell junctions (Triadafilopoulos *et al., 1*987). This collectively leads to increased vascular permeability and acute inflammatory response including the recruitment of neutrophils and cytokines, often resulting in colitis and pseudomembrane formation. In severe cases of pseudomembranous colitis, this can cause intestinal perforation and even lead to the fatal complication of toxic megacolon.

Hamster models of CDI have been utilised to demonstrate the cause and effect relationship of both toxins A and B. Lyerly *et al.,* showed that when given multiple doses of toxin A, hamsters developed intestinal pathology, diarrhoea and death, while preparations of toxin B caused no symptoms of diarrhoea, intestinal pathology or death to the animals (Lyerly *et al.,* 1985). Implicating that toxin A had a significant role in pathogenesis over toxin B. Nevertheless, studies using Ussing's chamber model have shown that toxin B is 10 times more potent than toxin A in mediating damage in human colonic mucosa (Riegler *et al., 1*995) and has been found to be just as virulent in strains possessing toxin B only (Alfa *et al., 2*000). The importance of toxin B was further pushed after reports by Lyras *et al., (*2009) of avirulence in a A+B- mutant 630 strain of *C. difficile* compared to lethal disease in hamster for a A-B+ mutant (Lyras *et al.,* 2009). The enigma of the significance of the independent toxins in CDI was again presented after a study reestablished the importance of both toxins A and B. Kuehne *et al.,* created three isogenic toxin mutants and demonstrated that either toxin alone can produce fulminant disease in the hamster model *(*Kuehne *et al.,* 2010; 2014).

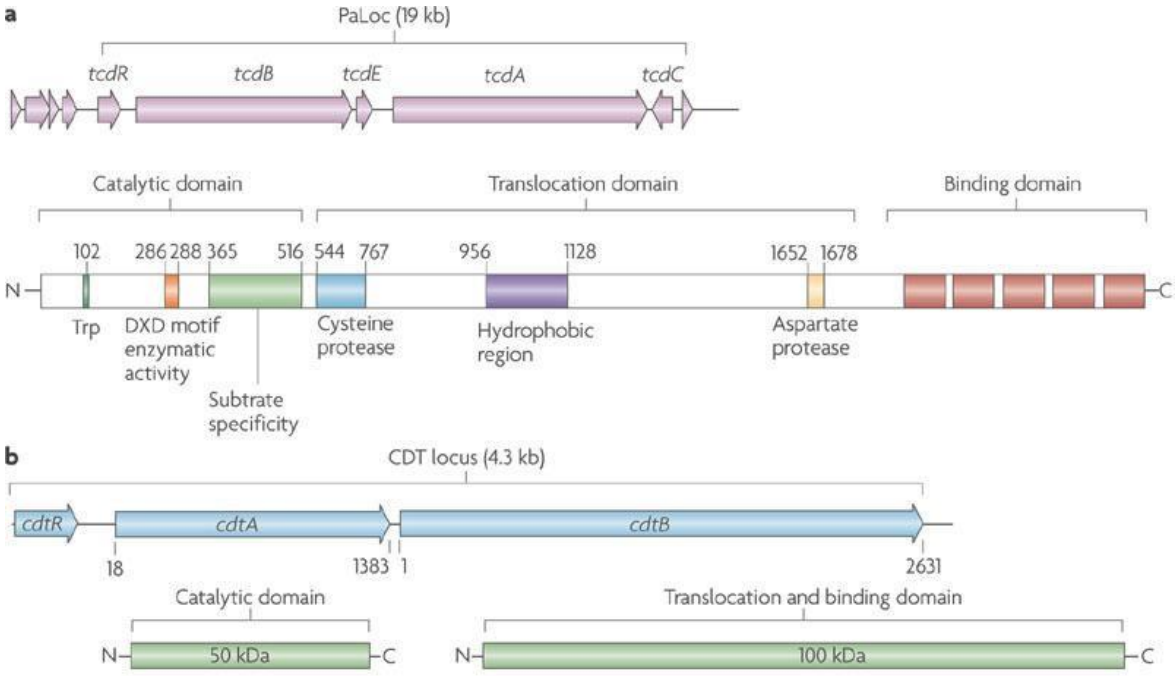


Figure 1.4: Toxin loci of *C. difficile* from Rupnik, M *et al.,* (2009).

a. *C. difficile* pathogenicity locus (PaLoc) and proposed protein domains organization of toxin B. Both *tcdA* and *tcdB* are encoded on the 19.6-kb pathogenicity locus, PaLoc. In addition to the two toxin genes *tcdA* and *tcdB*, that produce exotoxins A and B respectively, three additional regulatory genes, *tcdR*, *tcdE* and *tcdC*, are also located on this island. The genes *tcdR* and *tcdC*, have been found to encode proteins involved in the transcriptional regulation of the toxin genes (Mani, 2001; Matamouros, England, and Dupuy, 2007), whilst the product of *tcdE*, is required for the efficient secretion of TcdA and TcdB. The glycosyltransferase activity of *tcdB*, was found to be an inducer of the proinflammatory cytokines (Sun *et al.,* 2009), is located at the catalytic domain of the N terminus of the protein, and the C terminus, involved in receptor binding, is made up of combined repetitive oligopeptides (CROPs) (von Eichel-Streiber and Sauerborn, 1990; Ho *et al.,* 2005). The middle domain of the protein holds a putative transmembrane segment that is thought to be involved in membrane translocation of the catalytic domain into the cytosol (Jank and Aktories, 2008). The structure of TcdA is largely similar in organistation, however with a shorter CROPs domain than TcdB (Voth and Ballard, 2005). b. A third binary toxin named CD196 ADP-ribosyltransferase or *C. difficile* toxin (CDT) is encoded for by the CDT locus (Cartman *et al.,* 2010).

A third binary toxin was found in the CD196 strain by Poppof *et al.,* in 1988 (Popoff and Boquet 1988). Although this CDT showed no cytotoxic or lethal affect or relation to toxin A or B (Popoff and Boquet 1988), strains containing this binary toxin have amassed great interest in recent years. The view that CDT contributes to virulence was implicated in 2005, where McEllistrem *et al.,* detected an association between disease severity during a hospital outbreak and the presence of the binary toxin (McEllistrem *et al., 2*005). The epidemic strain responsible for the outbreaks across North America and Europe in 2005 was the ribotype 027, North American pulsed field gel electrophoresis type 1 (NAP1) (Warny *et al.,* 2005). It was later shown by Schwan *et al., (*2009) that this binary toxin had a role in virulence, inducing microtubule-based protrusions, increasing adherence of bacteria, in CDT producing types (NAP1/027) over non-CDT producing strains (VPI 10463). It was found that this ribotype uniformly carried the CDT binary toxin gene. In non-outbreak situations around 20 % of CDI causing strains have been found to carry CDT (Bauer *et al., 2*011). Moreover, infection with toxin A-negative, toxin B-negative, binary toxin-positive *C. difficile* isolate belonging to ribotype RT033, was isolated in a case of recurrent ulcerative colitis, however, this case was complicated by the isolation of a RT014/020 type strain from one of the patient samples (Androga *et al., 2*015).

1.5 Spores of *Clostridioides difficile*

1.5.1 Bacterial spores and resistant properties

In general, all bacterial spores are produced during a complex process known as sporulation, most often in response to nutrient deprivation. Spores are metabolically dormant and resistant to numerous environmental stresses, including oxygen, radiation, desiccation, ethanol, extremes of temperature, and low pH (Chen *et al., 2*008). Dormant bacterial spores can remain viable for hundreds of years (Kennedy, Reader, and Swierczynski, 1994), with studies suggesting that they may persist for millions of years (Cano and Borucki, 1995). Spores may then return to a metabolically active state, in a process known as germination. Within minutes upon sensing specific nutrients that signal that the external environment may be favourable for growth, spores outgrow into vegetative cells (Setlow, 2014). Responding to the challenge of surviving prolonged periods of dormancy in harsh conditions, spores have many mechanisms to protect themselves from damage which unfortunately also serve as defence from disinfection/sterilization procedures (Setlow, 2006). Many factors contribute to the resistant properties of bacterial spores, for many species this is largely because spore DNA is well protected by complex layers in their spore structure, of which, to a large extent, are remarkably similar between the various bacterial species (see Figure 1.5).

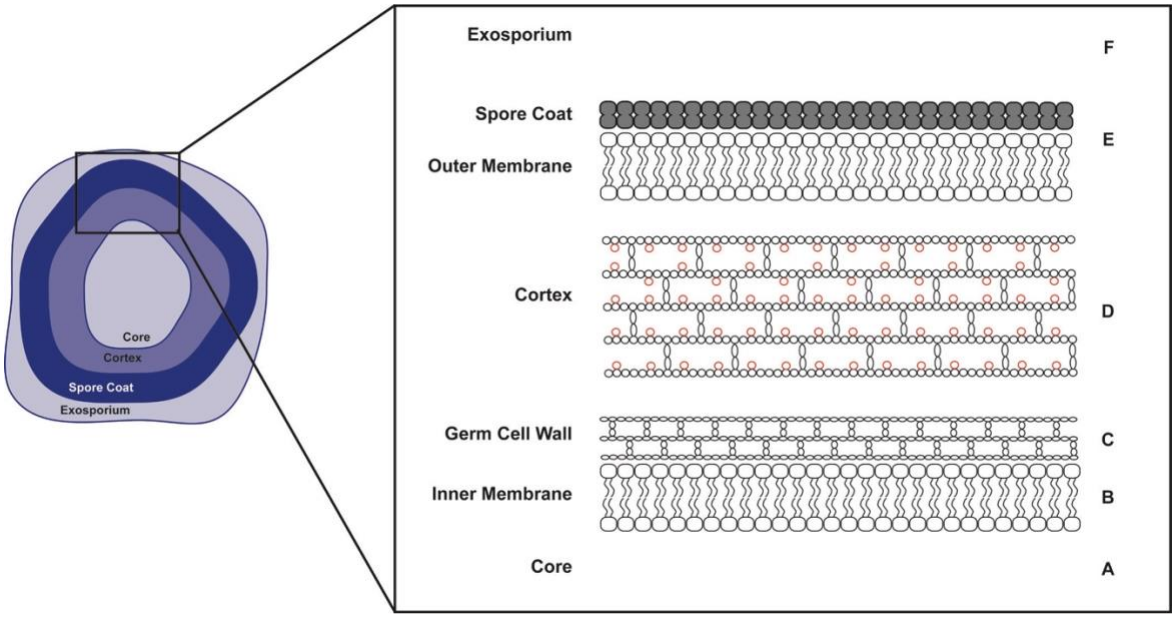


Figure 1.5: Structure of bacterial spores, from Kochan *et al.,* (2018).

Bacterial spores are composed of several layers, including a dehydrated spore core (A), an inner membrane (B), a germ cell wall (C), a spore cortex (D), a spore coat/outer membrane (E), and an exosporium (F). The spore cortex is a thick layer of modified peptidoglycan where up to 50% of N-acetylmuramic acid side chains are removed to produce muramic acid delta-lactam rings, which are depicted in red.

Located in the centre of the mature spore is the core. The spore core contains the genomic DNA, mRNA, ribosomes, protein, tRNA, and inactivated enzymes, and is very rich in pyridine-2,6-dicarboxylic acid, also known as dipicolinic acid (DPA), commonly as a calcium salt (Zhu *et al.,* 2018). These internal components are supported by the lower water content (Beaman *et al.,* 1982, Cowan *et al.,* 2003), water making up only 28–57% of spores' wet weight, a factor that is thought to contribute to both spores' enzymatic dormancy and their characteristic resistance to heat and some chemicals (Beaman and Gerhardt, 1986), (Paidhungat *et al., 2*000; Cowan *et al., 2*003) and (Sunde *et al., 2*009). This dehydration is effected by the transfer of dipicolinic acid chelated with calcium (CA-DPA) in to the core which replaces water, and is also known to be a major contributor to a spore's heat resistance properties (see Figure 1.5-A) (Paredes-Sabja *et al., 2*008; Paidhungat *et al., 2*000; Donnelly, Fimlaid, and Shen 2016). The DNA found in the core is supercoiled and bound with small, acid-soluble proteins (SASPs) that are believed to protect against DNA damage over long periods of time by providing resistance to heat and chemicals as well as by blocking transcription (Mason and Setlow 1986; Setlow 2007). Owing to the dormancy of the mature spore, any damage sustained to their DNA during this period needs to be rapidly repaired during outgrowth if the spore is to survive. Consequently, DNA repair is an important spore resistance mechanism. Spore DNA damage can be repaired by at least three mechanisms, including that mediated by spore photoproduct lyase (Yang *et al.,* 2015), recombination and excision repair (Moeller *et al., 2*008) and are RecA dependent and also can require some spore-specific repair proteins (Ramírez-Guadiana *et al., 2*016).

The spore core is surrounded by the inner membrane, which holds immobilised lipids and other membrane components that maintain the dehydration status of the core and limit permeability, (Popham *et al., 1*996; Nicholson *et al., 2*000; Cortezzo and Setlow, 2005; Setlow, 2006), thus serves as a permeability barrier against damaging chemicals (Swerdlow, Setlow, and Setlow, 1981; Westphal, *et al., 2*003; Setlow and Setlow, 1980) (see Figure 1.5-B). Protecting the inner membrane is the germ cell wall (see Figure 1.5-C) (Meador-Parton and Popham, 2000). Surrounding spore core is the spore cortex, a thick layer of modified peptidoglycan where muramic acid-delta lactam is present (see Figure 1.5-D) (Meador-Parton and Popham, 2000; Setlow, 2014). The presence of σ-lactam on about 50% of the muramic acid residues of cortex PG has been linked to the resistance to digestion by muramidases,

which play a crucial role in germination of spores by specific hydrolysis cortex lytic enzymes (CLEs) (Atrih, Bacher, Allmaier *et al., 1*999; Atrih, Bacher, Korner *et al., 1*999; Atrih *et al., 1*996; Popham 2002).This modification has been suggested to be a strategy that ensures that CLEs degrade the cortex and not the germ cell wall during germination (Kochan *et al., 2*018). The cortex layer is enclosed by the outer membrane, which contains the enzymes involved in cortex hydro- lysis (see Figure 1.5-E) (Chirakkal *et al., 2*002; Bagyan and Setlow, 2002; Ragkousi and Setlow, 2004) shown to have a role in maintaining spore dormancy, although the mechanism is not clear (Raju *et al., 2*006). Degradation of the spore cortex by chemicals and exogenous lytic enzymes is prevented by spore-specific proteins that are arranged in layers, making up the structure of the spore-coat following the outer-membrane (Donovan *et al.,* 1987). The outermost layer is a highly permeable layer of carbohydrates known as the exosporium (see Figure 1.5-F), found in *C. difficile* (Lawley *et al., 2*009; Permpoonpattana *et al., 2*011). This outermost layer has been suggested to have host immune activity and spore persistence during recurrent CDIs (Stewart, 2015; Diaz-Gonzalez *et al., 2*015) and is predicted to be involved in the pathogenesis (Redmond *et al., 2*004). Collectively, these structures provide resistance to environmental stresses that may be encountered during its spore life cycle (Gerding, Muto, and Owens, 2008).

Figure 1.6 displays the typical life cycle of endospore formers that is further discussed in the following subsections.

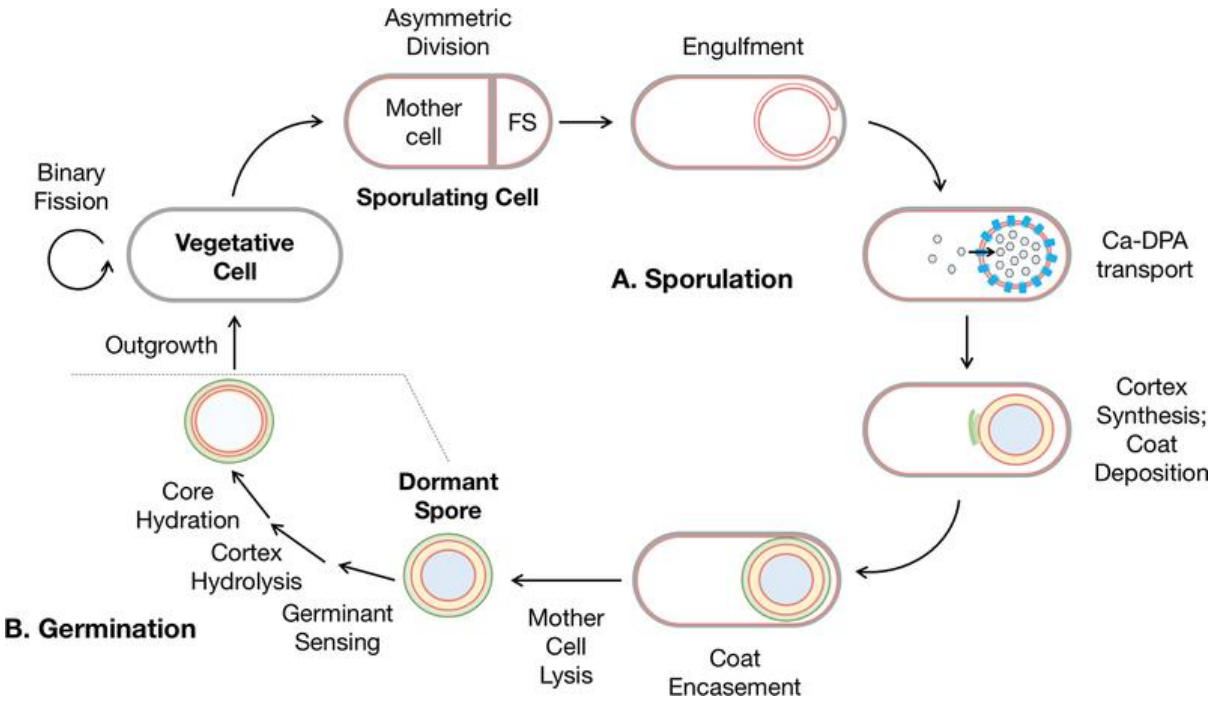


Figure 1.6: Lifecycle of Endospore Formers from Shen *et al.,* (2019).

(A) Sporulation: Upon sensing certain environmental conditions, endospore formers activate Spo0A and initiate sporulation. The first morphological event is the formation of a polar septum, which creates a larger mother cell and smaller forespore. The mother cell engulfs the forespore, and the two cells work together to assemble the dormant spore. Calcium dipicolinic acid (Ca-DPA) is synthesized in the mother cell and transported into the forespore in exchange for water. The cortex is formed between the two membranes, and coat proteins polymerize on the surface of the mother cell-derived membrane. Once the spore is mature, the mother cell lyses and releases the dorman spore into the environment. (B) Germination: Upon sensing the appropriate small molecule germinants, the spore initiates a signaling cascade that leads to activation of cortex hydrolases and core hydration, which is necessary for metabolism to resume in the germinating spore.

1.5.2 Sporulation

Clostridioides difficile, being an obligate anaerobe, is highly sensitive to oxygen, so the production of aerotolerant spores allows this organism to survive in the external environment until it infects a new host. When *C. difficile* cells meet certain environmental stimuli, they will initiate a sporulation pathway to produce sufficient dormant spores to survive in extreme situations (Setlow 2006; Rodriguez-Palacios and Lejeune, 2011; Deakin *et al., 2*012). The main process of *C. difficile* sporulation contains four stages of morphogenetic importance (Gil *et al., 2*017) (see Figure 1.7).

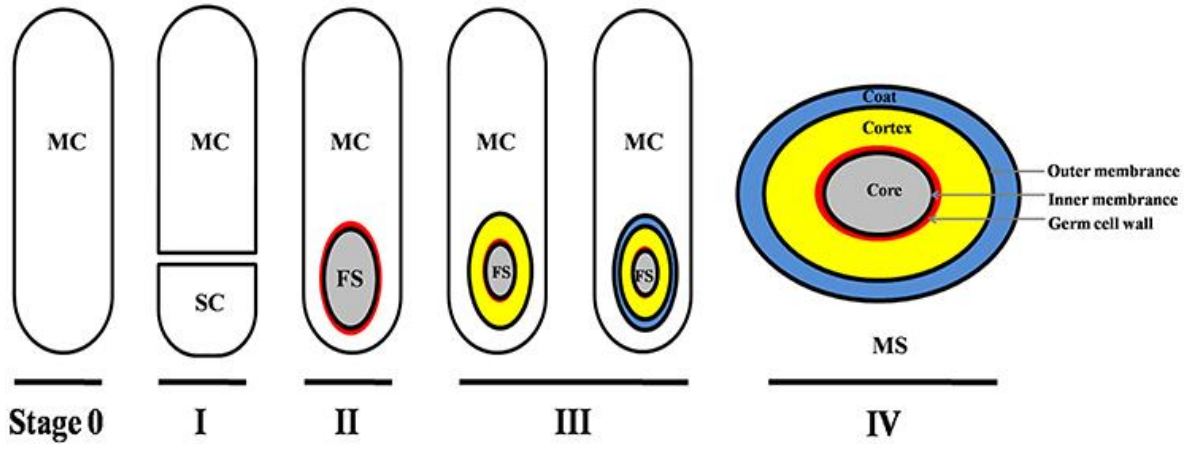


Figure 1.7: Stages of the sporulation process and structure of *C. difficile* spore from Zhu *et al.,* (2018).

MC, mother cell compartment; SC, smaller compartment; FS, forespore compartment; MS, mature spore.

Briefly, the mother cell (MC) (stage 0) undergoes asymmetric septation in stage I. There is an unequal cell division that separates two copies of the chromosome generating a large mother cell and a smaller forespore separated by the specific regulation of genes and production of a sporulation septum or smaller compartment within the larger MC. As the plasma membrane of the mother cell grows around and engulfs the forespore, it becomes surrounded by a double membrane (Parades-Sabja *et al.,* 2015). To ensure the gene expression in both bodies

remain coordinated, there is a complex cross talk between these compartments (Paredes-Sabja, Shen, and Sorg, 2014). This is achieved by the activity of specific sigma (σ) factors, that activate compartment-specific transcriptional regulation during sporulation (Fimlaid *et al., 2*015). These include σF and σG and σE and σK, which drive forespore and mother cell gene expression, respectively (Zhu, Sorg, and Sun, 2018). Further forespore and mothercell proteins, SpoIIAB, SpoIIAA, SpoIIE, SpoIIR, SpoIIGA, SpoIIIAH, SpoIIID and SpoIIQ have been found to be involved in the sporulation regulatory pathway that interact with the sigma (σ) factors (Fimlaid *et al., 2*013; Fimlaid *et al., 2*015; Fimlaid and Shen 2015; Serrano *et al., 2*016).

In stage II, the MC engulfs the SC resulting in a forespore contained within the cytoplasm of the MC. Important layers of the spore are formed in stage III. Synthesis of a spore cortex between the outer and inner forespore membranes is accompanied by a large decrease in the pH, volume and water content of the forespore into the cytoplasm (Paredes-Sabja, Shen, and Sorg, 2014). Large amounts of Ca-DPA, produced by the mother cell, is transferred into the forespore (approximately 25% of core dry weight), which decreases forespore protoplast water content even further (Paredes-Sabja, Shen, and Sorg, 2014). A sequence of gene expression results in the synthesis of specific proteins in the mother cell which are then used to layer the outer surface of the spore, forming the spore coat (Permpoonpattana *et al., 2*013). The clostridium-specific coat protein, SipL, is known to be essential for functional spore formation, working in corodination with spoIVA to promote encasement around the forespore (Touchette *et al.,* 2019; Benito de la Puebla *et al.,* 2020) (See Figure 1.8).

An outermost layer known as the exosporium is also found on some spores of *C. difficile* (Paredes-Sabja, Shen, and Sorg, 2014) (See Figure 1.8), with two exosporium morphotypes observed amongst clonal populations; thin and thick electron-dense layers, although both have hair-like extensions Pizarro-Guajardo *et al.,* 2016). Assembly of the outermost exosporium depends on the BclA collagen-like proteins, which likely create hair-like projections on the spore surface (Phetcharaburanin et al., 2014; Pizarro-Guajardo *et al.,* 2014), CdeC (Barra-Carrasco *et al.,* 2013) and CdeM (Shen *et al.,* 2019). Shen *et al.,* (2019) recently presented the hypothesis that the different morphotypes may have different roles during *C. difficile* infection such as spore persistence and immune evasion, as supported by mouse studies implicating exposproium proteins CdeM and BclA1 having a signification role

in colonization (Janoir *et al.,* 2013; Phetcharaburanin *et al.,* 2014). Finally, the MC lyses at stage IV and the mature spore is released into the surrounding environment.

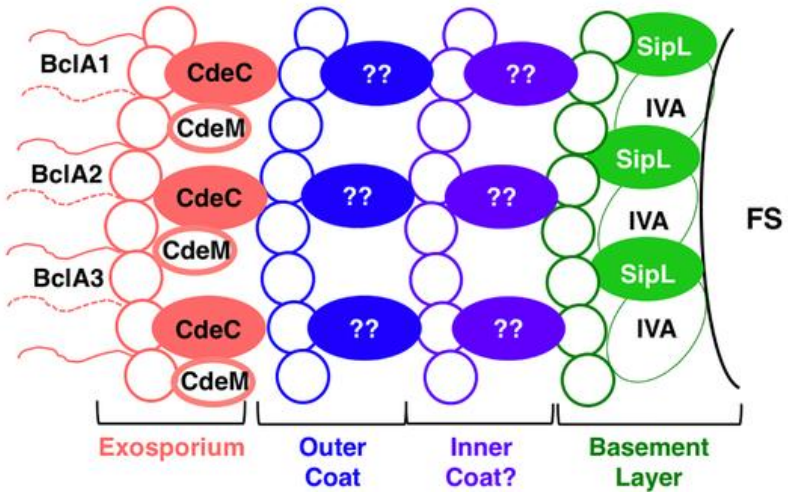


Figure 1.8: Spore coat and exosporium structure in *C. difficile* from Shen *et al.,* (2019).

The exopsorium layer made up of BclA 1-3, CdeC, and CdeM, whislt the proteins that make up the outer and inner coat layers are unknown, finally the basement layer is formed by SpoIVA (IVA) and SipL proteins recruiting coat proteins to the forespore and forming heat-resistant spores (Ribis *et al.,* 2017; Putnam *et al.,* 2013).

1.5.3 Initiation and regulation of sporulation in *C. difficile*

Progress has been made towards understanding of the mechanisms that regulate *C. difficile* sporulation. It is likely that environmental stimuli such as nutrient limitation, quorum sensing, and other unidentified stress factors are involved based on studies in other organisms (Higgins and Dworkin, 2012). *Bacillus* species, for example, produce spores as a response to starvation of carbon and/or nitrogen (Cowan *et al., 2*003; Setlow, 2003). This process is initiated by phosphorylation of the sporulation stage 0 protein A, Spo0A, encoded for by the gene *spo0A*, which plays an essential role in the initiation of sporulation; inactivation of *spo0A* results in complete deficiency in spore formation (Underwood *et al., 2*009). The protein Spo0A is a response regulation transcription factor which upon activation, can bind to specific target sequences, known as 0A boxes, near or in the promoters of genes under its direct regulation (Hatt and Youngman, 2000), including the sporulation-specific RNA polymerase

sigma factors, especially for σE, σF, σG, and σK (Fimlaid *et al., 2*015). Inactivation of *spo0A* results in complete deficiency in spore formation (Underwood *et al.,* 2009), and is described as the master transcription regulator (Fimlaid and Shen, 2015; Deakin *et al., 2*012). The protein Spo0A is a response regulation transcription factor which upon activation, can bind to specific target sequences, known as 0A boxes, near or in the promoters of genes under its direct regulation (Hatt and Youngman, 2000), including the sporulation-specific RNA polymerase sigma factors, especially for σE, σF, σG, and σK during regulatory feedback between the mothercell and the forespore (Fimlaid *et al., 2*015; Zhu, Sorg, and Sun, 2018). Inactivation of *spo0A* results in complete deficiency in spore formation (Underwood *et al.,* 2009), and is described as the master transcription regulator (Fimlaid and Shen, 2015; Deakin *et al., 2*012). The protein Spo0A is a response regulation transcription factor which upon activation, can bind to specific target sequences, known as 0A boxes, near or in the promoters of genes under its direct regulation (Hatt *et al.,* 2000). A review of other molecules involved in the regulation of sporulation in *C. difficile* was performed by Shen *et al.,* (2019) a schematic of the proposed pathways is shown in Figure 1.9.

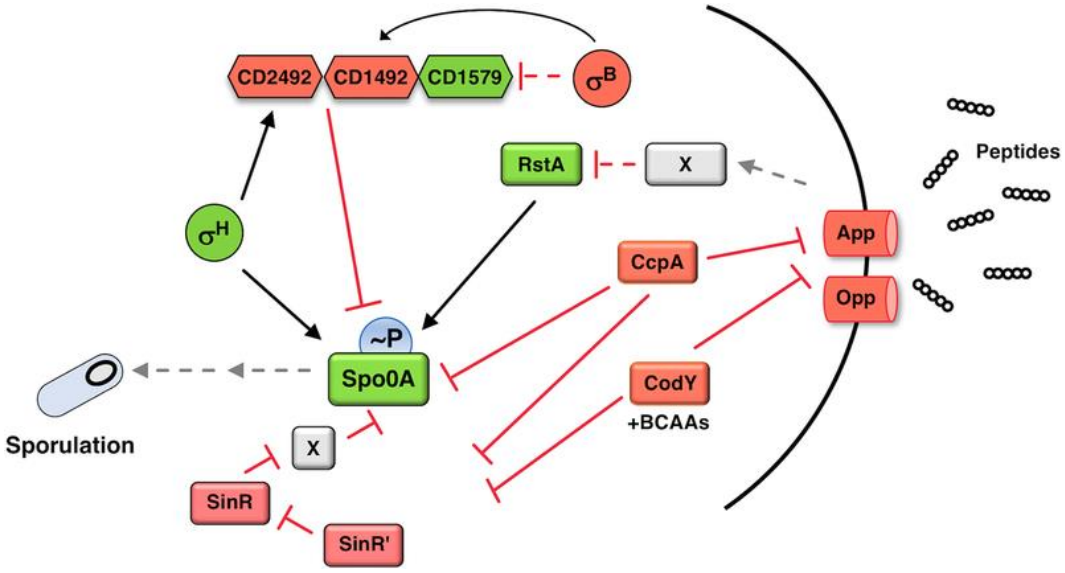


Figure 1.9: Regulatory pathway controlling Spo0A activation in *C. difficile* Shen *et al.,* (2019).

Positive regulators of Spo0A are highlighted in green, while those that inhibit Spo0A are highlighted in red , hexagons = histidine kinase/phosphatases; rounded rectangles = transcription factors; and circles = sigma factors. Red lines indicate negative regulation, and black lines indicate positive regulation. Solid lines indicate defined regulatory interactions, and dashed lines suggest proposed, and potentially indirect, regulatory effects. Grey rounded boxes idicate unkown intermediary. The reciprocal transcriptional regulation of early sporulation factors by Spo0A has also been omitted for simplicity.

The proteins CodY and CcpA established as nutritional sensor proteins in *Clostridioides* species have been shown to work as negative regulators of sporulation (Duncan *et al., 1*995; Hofmeister *et al., 1*995; Karow, Glaser, and Piggot 1995; Londonovallejo and Stragier 1995; Antunes *et al., 2*012; Nawrocki *et al., 2*016; Serrano *et al., 2*016). Among the genes CodY regulates are genes involved in spore formation, most notably spo0A (Zhu, Sorg, and Sun, 2018). Branched-chain amino acids (BCAAs) are a CodY cofactor (Dineen *et al.,* 2007) and their precursors are likely imported into vegetative cells primarily through the Opp and App oligopeptide transporters (Edwards *et al.,* 2014). It has been proposed that peptide transport coordinates the onset of sporulation in *C. difficile*, since the loss of Opp and App increases *C. difficile* sporulation ~20-fold (Edwards *et al.,* 2014). These results suggest that the peptides imported by *C. difficile* Opp and App serve as nutrients rather than quorum sensing molecules such that *C. difficile opp* and *app* mutants may initiate sporulation earlier due to decreased nutrient acquisition (Shen *et al.,* 2019).

Whilst CcpA, a LacI family DNA-binding transcriptional regulator, has been found to respond to the availability of carbohydrates (Deutscher, Francke and Postma, 2006). Interestingly, CcpA positively regulates sugar uptake, fermentation and amino acid metabolism, whilst repressing the use of alternative carbon sources (Fujita, 2009). In *C. difficile*, the PaLoc genes (*tcdR, tcdB, tcdA*, and *tcdC*) are directly regulated by CcpA to mediate glucose-dependent repression of toxin production and has been shown to indirectly regulate *C. difficile* sporulation (Antunes, Martin-Verstraete and Dupuy, 2011). A novel, positive regulator of sporulation initiation, RstA, was also found in *C. difficile* (Edwards, Tamayo and McBride, 2016) (see Figure 1.9). The peptide-interacting domain of RstA has been shown to positively affect the initiation of *C. difficile* sporulation. Interestingly, by affecting the flagellar-specific sigma factor (SigD) expression, RsTA was also shown to negatively regulate *C. difficile* toxin production and mobility. SinR and SinR', *C. difficile* orthologs for regulatory proteins were recently shown to promote sporulation in *C. difficile* (Girinathan *et al.,* 2018).

1.5.4 Bacterial spore germination

The transition from dormant bacterial spore to a vegetative cell has been described to involve three separate phases including activation, germination and outgrowth (Leggett *et al., 2*012). Dormancy is broken in the activation phase, which can be triggered by appropriate conditions of heat, pH or chemical exposure, rendering the dormant spore ready to enter the germination phase (Keynan and Evenchik, 1969). Whilst activation is a reversible process, by which spores are not necessarily committed to phases of germination and outgrowth (Keynan and Evenchik, 1969), once a spore is undergoes germination, the spore can no longer return to its dormant state (Gould, Stubbs, and King, 1970). Various environmental stimuli, known as germinants, initiate the germination process by interaction with germinant receptors, often varying depending on the species. These include, but are not limited to, metabolizable nutrient germinants; such as amino acids, nonmetabolizable germinants; such as some ionic species and chelates (for example Ca:DPA), and some physical treatments such as high pressures (Gould, Stubbs, and King, 1970; Setlow, 2003). The 'outgrowth' phase is defined as all developmental events taking place after germination that result in the return of the spore to vegetative cell growth, noted to include the initiation of metabolism, macromolecular synthesis, swelling of the spore, emergence from the outer spore layers and growth of the new cell (Leggert *et al.,* 2012).

1.5.5 Germination pathways of *C. difficile* spores

The lifestyle of spore germination in CDI starts when a patient ingests *C. difficile* spores. Once in the stomach spores are able to withstand the acidic environment and then pass through to the small intestine where they change into the metabolically active state through a process known as germination (Poutanen and Simor, 2004; Moir, 2006). The protective spore cell layers that provide much of their resistance properties are lost as they outgrow into vegetative cells, which are now able to colonize the large intestine and produce toxins and disease pathology (Dupuy and Sonenshein, 1998; Rupnik, Wilcox, and Gerding, 2009b; Warny *et al., 2*005; Carroll and Bartlett, 2011). As such spore germination plays a crucial role in the

development of CDI (Zhu, Sorg, and Sun, 2018; Setlow, 2001). Molecules known as germinants indicate to spores that environmental conditions are favourable cell survival.

In *C. difficile* spore germination is induced most effectively by the presence of cholate-derivatives, a component of bile (Lawley *et al., 2*009). Bile is produced to aid food digestion and protect against colonisation of the small intestine by bacterial pathogens (Ridlon *et al.,* 2006). The two main bile acids include cholic acid and chenodeoxycholic acid, which can be conjugated with taurine (giving taurocholate or taurochenodeoxycholate, respectively) or glycine (giving glycocholate and glycochenodeoxycholate, respectively). Taurocholate (TCA), in combination with glycine (Gly), can stimulate germination, whereas chenodeoxycholate (CDCA) inhibits germination in a *C. difficile* clinical isolate (Sorg and Sonenshein, 2009b; 2008a).

These acids can be further metabolised by the action of surface layer bile salt hydrolases of gut commensal *Bacteroides* spp. producing deconjugated primary bile acids cholate and CDCA (Ridlon, Kang, and Hylemon, 2006). As CDCA is actively reabsorbed 10-fold faster than TCA (Mekhjian *et al.,* 1979), germination of the spore would be favoured in the anaerobic environment of the small intestine (Sorg and Sonenshein, 2008; 2009). These deconjugated bile acids can be further metabolised by 7α-dehydroxylation into secondary bile salts by a few commensal microbiota of the small intestine such as *C. scindens* (Northfield and McColl, 1973; Mallonee and Hylemon, 1996; Ridlon *et al.,* 2006), thus reducing the availability of bile salt germinants resulting in an inhibitory effect on germination and spore outgrowth. However, use of broad-spectrum antibiotics can deplete such bacteria, allowing for spore germination and colonisation and subsequent disease by toxigenic *C. difficile* (Samuel *et al.,* 1973; Sorg and Sonenshein, 2008; 2009). It has also been reported that CDCA does not inhibit the germination, or outgrowth, of all *C. difficile* strains (Heeg *et al.,* 2012) and that these bile salts alone are not sufficient for efficient *C. difficile* germination, with additional molecules such as amino acids or divalent cations, referred to as co-germinants are essential for the initiation of germination signalling (Sorg and Sonenshein, 2008; Shrestha and Sorg, 2018; Kochan *et al., 2*017). These findings suggest that more than one germination pathway may be present in *C. difficile* and that these pathways are not found in other sporulating bacteria.

Germinants are known to interact with interaction with specific germination receptors (Pelczar *et al.,* 2007; Alturi *et al.,* 2006). In *C. diffcile*, a pseudoprotease known as CspC, located in the spore coat/outer membrane (Francis *et al.,* 2013), has been found which is to be the receptor for TA-mediated germination, proposed to stimulate cortex hydrolysis upon binding to bile acids (Francis *et al.,* 2013). More recently, CspA has also been suggested as the second *C. difficile* germinant receptor in regard to recognition of amino acid and divalent cation co-germinants (Shreatha *et al.,* 2019). GerG, a putative gel-forming protein, is found exclusively in *C. difficile* isolates and is required for *C. difficile* germination (Donnelly *et al.,* 2017). It is hypothesized that GerG functions by facilitating transport of Csp proteins across the outer forespore membrane into the cortex region during sporulation (Fimlaid *et al.,* 2015; Donnelly *et al.,* 2017).

In *Clostridioides difficile*, multiple distinct germination pathways have been identified, each requiring specific germinant-receptor interactions (see Table 1.1) (Kochan *et al.,* 2018; Shreatha *et al.,* 2019). The most efficient amino acid co-germinant for *C. difficile* spores is glycine, however it been found to be interchangeable with other amino acids, including L-alanine, L-histidine, and L-serine (Sorg and Sonenshien, 2008; Wheeldon *et al.,* 2011; Ramirez *et al.,* 2010; Shrestha *et al.,* 2017). TA and these amino acid co-germinants have been grouped into the Bile salt-amino acid germination pathway (Kochan *et al.,* 2018). Studies have also found that D-alanine and D-serine can facilitate germination together with taurocholate in *C. difficile* (Shrestha *et al.,* 2017). The enzyme alanine racemase Alr2 has also been shown to have role in this germination pathway, required for the conversion of L-alanine to D-serine. Divalent cations calcium and magnesium also function as cogerminants and have been found to play a major role in inducing germination *in vivo* (Kochan *et al.,* 2007). It has been hypothesised that this bile-salt-divalent cation germination pathway is the most physiologically relevant in the host, given that bile salts, calcium, and glycine are all present in the gastrointestinal tract (Kochan *et al.,* 2007).

Table 1.1: Germination pathways of *Clostridioides difficile* adapted from Kochan *et al.,* (2018) and Shreatha *et al.,* (2019).

Pathway	Germinant(s)	Receptors[1]
Bile salt-amino acid	Taurocholate and glycine, L-histidine, L-serine, or L-alanine	CspC, CspA
Bile salt-divalent cation	Taurocholate and calcium or magnesium	CspC, CspA
Bile salt-alanine racemase-dependent D-amino acid	Taurocholate and D-alanine or D-serine	CspC, CspA

[1] '?' indicate unknown co-germinant receptor(s)

While there are several environment-sensing pathways that can initiate spore germination, Each of the three germination pathways mentioned initiates a signalling cascade that leads to activation of SleC, the spore cortex lytic enzyme, found to be essential for *C. difficile* germination (Burns *et al.,* 2010). SleC activation initiates degradation of the cortex, resulting in full core rehydration and release of Ca-DPA from the spore core (Donnelly *et al.,* 2016; Kochan *et al.,* 2017; Shrestha *et al.,* 2017; Adams *et al.,* 2016). Subtilisin-like proteases of *C. difficile*, that are essential for germination, are encoded on the operon *cspBAC* carrying carries two genes, *cspBA* and *cspC,* that are expressed during sporulation (Adams *et al.,* 2013). The YabG protease processes the fusion protein into two proteins, CspB and CspA, which are incorporated into mature spores (see Figure 1.10 (B)) (Kevorkian *et al.,* 2016; Kevorkian and Shen, 2017). CspB cleaves the prodomain from pro-SleC, initiating cortex hydrolysis (Adams *et al.,* 2013), whilst, the incorporation of the germinant receptor, CspC, into the mature spore is affected by CspA (Kevorkian *et al.,* 2016; Kevorkian and Shen, 2017).

Rohlfing *et al*., (2019) bring into contention whether CspC acts a signalling molecular rather than a germinant receptor itself; identifing CspC mutations that confer either hypo- or hyper-sensitivity to bile acid germinant. Hyper-sensitive CspC variants exhibited bile acid-independent germination as well as increased sensitivity to amino acid and/or calcium co-germinants. Since mutations in specific residues altered CspC's responsiveness to these different signals (Francis *et al.,* 2013), CspC plays a critical role in regulating *C. difficile* spore germination in response to multiple environmental signals (Sorg and Sonenshein *et al.,* 2009).

Rohlfing *et al*., (2019) implicate CspC as being intimately involved in the detection of distinct classes of co-germinants in addition to bile acids and thus raises the possibility that CspC functions as a signalling node rather than a ligand-binding receptor. Shen *et al.,* (2019) point out that the binding of germinants to germinant receptors has yet to be biochemically demonstrated in any organism.

Significant progress has been made in regards to the identification molecules that regulate germination pathways in *C. difficile*, gaps remain in the knowledge relating to the permeability and mechanism by which germinants gain access to the spore's inner layers, as well as the interactions between bile salts and CspC, and the downstream signalling that occurs. Events that that lead to activation of CspB after bile salts interact with CspC, however, remain unknown (Kochan *et al.,* 2018). Upon sensing the appropriate small molecule germinants, the spore initiates a signalling cascade that leads to cortex hydrolysis which is initiated following the activation of cortex hydrolases and core hydration, which is necessary for metabolism to resume in the germinating spore (Shen *et al.,* 2019). Kochan *et al.,* (2018) provide two models of *C. difficile* germination (see Figure 1.10).

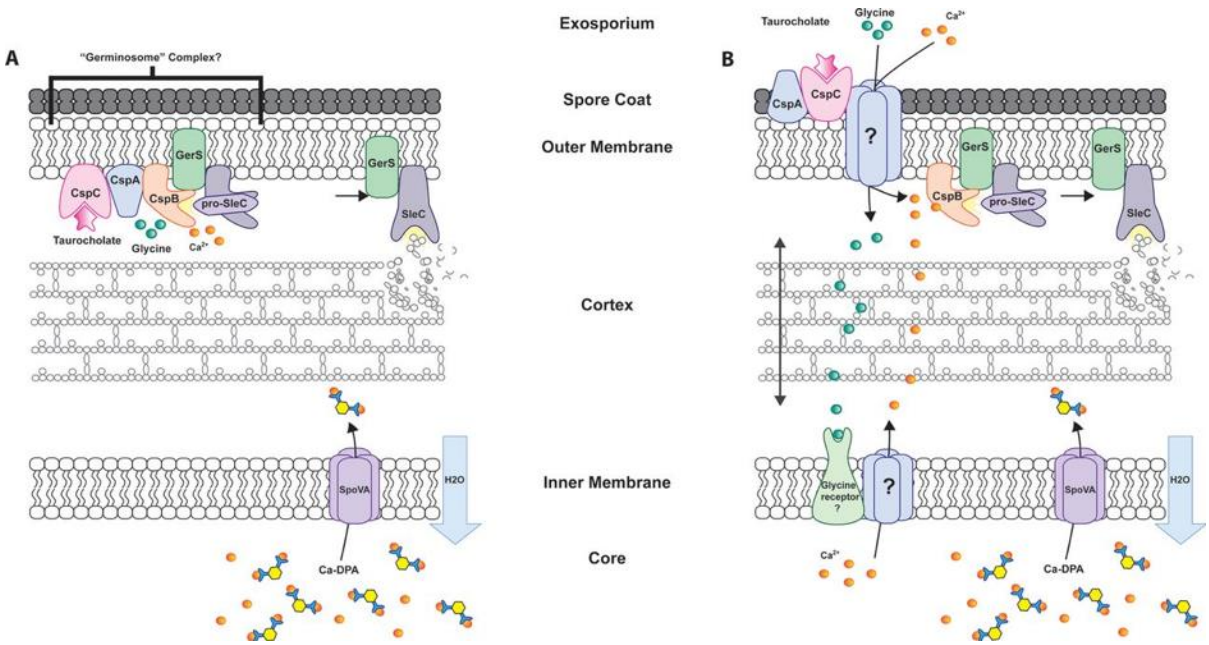


Figure 1.10: Proposed models of *Clostridioides difficile* germination from Kochan *et al.,* (2018).

(A) "Germinosome" model where proposed Csp protein-protein interactions lead to activation of SleC. (B) "Lock-and-key" model where CspC-taurocholate interactions facilitate cogerminant entry into the spore (Kochan *et al.,* 2018).

The 'Germinosome' model (see Figure 1.10 (A)), based on the work of Francis *et al.,* (2013), Adams *et al.,* (2013) and Fimlaid *et al.,* (2015), where proposed Csp protein-protein interactions lead to activation of SleC activating spore cortex hydrolysis (Francis *et al.,* 2013, Adams *et al.,* 2013; Fimlaid *et al.,* 2015). Whilst, the "lock-and-key" model (see Figure 1.10 (B)), based on the work of modified from that of Kochan *et al.,* (2017) hypothesizes that CspC faces outwards facilitating when taurocholate binds interactions facilitate cogerminant access and entry into the spore after interactions with their putative receptors. Kochan *et al.,* (2018) then go on to propose that calcium from the spore core (or the external environment) then binds to and activates CspB and subsequently SleC and cortex hydrolysis. In *C. difficile*, the signalling events that trigger cortex hydrolysis remain unknown. It is hypothesized that GerS regulates the activity of SleC by playing a role in modification of the spore cortex, allowing SleC to recognize it and initiate hydrolysis, similar to the role of CwlD in the cortex

modification of *Bacillus* species (Paidhungat *et al.,* 2002; Gilmore *et al.,* 2004; Fimlaid *et al.,* 2015; Diaz *et al.,* 2018). Ca-DPA is eventually released by a mechano-sensing mechanism whereby the core is able to swell after cortex hydrolysis (Francis *et al.,* 2015) and DPA is released to relieve hypo-osmotic stress (Francis and Sorg 2016).

Shen *et al*., (2019) made a comprehensive review of recent studies in germination amongst clostridial pathogens summate the critical insight they have provided into the early stages of germination namely germinant sensing, cortex hydrolysis, and Ca-DPA release, however they acknowledge that the little is known about subsequent events. Following the hydration of the spores' core, transcript levels increasing within 15 min of germination signifying that metabolism is allowed to resume (Dembek *et al.,* 2013). The Gpr germination protease of *B. subtilis*, which degrades SASPs coating the spore nucleoid, is activated following the release of Ca-DPA (Illades-Aguiar and Setlow, 1994). The proteolytic activity of Gpr on SASPs leads to the liberation of DNA, permitting transcription and translation to free up amino acids that are then provided as substrates for both catabolism and metabolism (Setlow, 2013). Yet it is currently unclear whether Ca-DPA activates Gpr homologs in clostridial pathogens (Shen *et al.,* 2019).

In addition to metabolism, the hydration of the core during spore germination concurrently leads to loss of resistance (Shen *et al.,* 2019). It has been found that mutant spores of *C. difficile* with higher spore water content are more susceptible to heat (Donnelly *et al.,* 2016; Paredes-Sabja *et al.,* 2008; Orsburn *et al.,* 2009). Whilst the susceptibility of spores to UV irradiation is speculated to be a result of reduced SASP binding affinity of SASPs to DNA due to increased hydration (Paredes-Sabja *et al.,* 2008; Orsburn *et al.,* 2009)). RNA-Seq analyses suggest that SspA and SspB are the predominant SASPs produced in *C. difficile* (Fimlaid *et al.,* 2013), which are the leading contibutors to resistance to UV and solar radiation in *B. subtilis* spores' (Moeller *et al.,* 2009). The role of SASPs in conferring resistance to *C. difficile* spores has not yet been examined (Shen *et al.* 2019).

1.5.6 Methods to detect sporulation, spore germination and outgrowth

A number of approaches have been developed to measure sporulation and germination phenotypes, however, considerations must be made to separate processes (Burns and Minton, 2011). Sporulation can be measured as expression on colony forming units after stress treatment (Buckley *et al.,* 2011; Goulding *et al.,* 2009; Kamiya *et al.,* 1992; Kuehne *et al.,* 2010; Lyras *et al.,* 2009; Poxton *et al.,* 2001; Sambol *et al.,* 2001). Whilst this assay provides information as to the production of spores over time, this phenotype would need to be separated by additional processes of germination and outgrowth as well as additional factors such as the death of non-sporulating cells and spore resistance. Spore titres can also be quantified using phase contrast microscopy, where phase-bright spores are easily distinguished from phase-dark vegetative cells, which can be enhanced using specific stains endospores stain such as Schaeffer-Fulton endospore stain (Mormak and Casida, 1985; Fawley *et al.,* 2007; Underwood *et al.,* 2009). The spore titre method allows the direct counting or estimation of the number of spores in a population, either of pure spores, to estimate purity, or in a background of vegetative cells differentiated by phase-brightness.

Measurements of germination can be expressed as the loss of phase-brightness (Leuschner and Lillford *et al.,* 2002) or drop in optical density (McCormick, 1964) over time indicating the in-take of water in the spore core. The assay provides information as the rate of germination initiation but can be complicated with outgrowth phenotypes, cell death, and spore resistance when alternative methods including the use of colony forming units is employed as used by Wheeldon *et al.,* (2011) and Carlson *et al.,* (2015). Additionally, the efficiency of germination can also be presented as a percentage of colony forming units in a mixed population of vegetative cells and spores; estimated using phase-bright microscopy (to enumerate the total population) and colony forming units (to indicate germinated spores). Similarly, the loss of heat-resistance of spores can indicate the process of germination as well as monitoring the release of DPA using terbium fluorescence (Francis *et al.,* 2015; Bhattercharjee *et al.,* 2015). DIC microscopy is an advanced technique that allows the germination of individual spores to be monitored (Zhang *et al.,* 2010). A spore decoating method can be used to assess whether phenotypes are attributed to molecules present in the spore cortex as described in Burns *et al., (*2010) as well as indicating the viability of spores to

outgrow into vegetative cells. Spore outgrowth and vegetative cell growth can be measured using rise in optical density (Akerlund *et al.,* 2008; Burns *et al.,* 2010a; Burns *et al.,* 2010b; Merrigan *et al.,* 2010; Underwood *et al.,* 2009; Vohra and Poxton, 2011), or the formation of colonies from spores.

Measures can be made to further validate phenotypes observed including the standardisation of growth mediums, and the use of controls to ensure no or minimal contamination of spores or vegetative cells, depending on the process being examined (Burns and Minton, 2011). In the starter culture or spore suspension at the start of the experiment, measurements at 0 h also known as T0 (or T_0) in order to provide a baseline (Vohra and Poxton, 2011; Burns and Minton, 2011). Burns and Minton (2011) review that differences in the experimental set up used to evaluate the same VPI 10463 strain produced conflicting sporulation proficiencies between studies (Akerlund *et al.,* 2008; Burns *et al.,* 2010a; Merrigan *et al.,* 2010; Vohra and Poxton, 2011). Even when using the same experimental methods, further measures of control are needed, including the use of appropriate control strains and sample sizes. When making conclusions on a particular group of strains (e.g. grouped by molecular type) to improve the confidence of data of potential strain-to-strain variation, it is import to sample as many representatives of that group as possible. A flowchart of the methods used in the Clostridium Research Group (CRG) to characterise *C. difficile* spore phenotype is presented in Figure 7.1 of the Appendix.

1.6 Genetic Manipulation of *Clostridioides difficile*

The strategy of genetic manipulation of *C. difficile* has been focused on generating mutants of genes involved in sporulation and/or germination. The genome of *C. diffcile* 630 was sequenced in 2006 (Sebaihia, M *et al.,* 2006). Following this, the proteomic analysis of *C. difficile* 630 spores (Lawley, TD *et al.,* 2009) identified a total of 336 spore-associated polypeptides, 39 of these being predicted to be secreted or surface associated. However, many of these are encoded by genes that have no predicted function in the genome annotation. Nine percent of the spore proteome consists of spore proteins encoded by genes

that are currently unique to *C. difficile*. Many of these are hypothetical proteins that can be preferentially targeted in mutant construction.

1.7 Molecular typing of *Clostridioides difficile*

1.7.1 Typing methods of *C. difficile*

Historically, the characterisation of *C. difficile* strains was based on phenotypic descriptions of features such as capsular and protein serotypes (Killgore, G *et al.,* 2008). Early bacterial population studies were aided by the use of multi-locus electrophoresis (MLEE), where isolates were discriminated based on the differential mobility of cellular enzymes during electrophoresis (Selander and Levin, 1980; Smith *et al.,* 1993; Maiden *et al.,* 1998). In an attempt to standardise the typing results across laboratories, multi-locus typing was developed (Maiden *et al.,* 1998). Advances in nucleic acid sequencing technologies prompted the development of ribotyping, which uses 16S ribosomal RNA gene sequence as a marker that is still in wide use today. The *Clostridioides difficile* Ribotyping Network (CDRN) provides ribotyping as a molecular epidemiology service across England and Northern Ireland to monitor the distribution of *C. difficile* types. However, ribotyping is not appropriate for identifying subtypes (Stackebrandt and Ebers, 2006; Medini *et al.,* 2008). Other typing schemes involve digesting DNA fragments with restriction enzymes and comparing the size of DNA fragments following electrophoresis such as pulse-field gel electrophoresis (PFGE) (using *smaI* digested DNA) and restriction endonuclease analysis (REA) (for HindII digested DNA) and are employed for epidemiological purposes in North America and the USA, respectively (Loo *et al., 2*005b; McDonald, 2005). Whilst whole-genome sequencing (WGS) and MALDI-TOF techniques are now available to type bacteria, the naming of strain types using historical typing methods are still used in reference to epidemiology of transmission of CDI.

1.7.2 Emergence of 'hypervirulence'

The emergence of *C. difficile* strains belonging to restriction endonuclease type BI, North American pulsed-field type 1 (NAP1) and PCR-ribotype 027 (BI/NAP1/027) was widely reported in 2005 after a series of outbreaks in North America, the USA and Europe. Strains of this type contributed to the increase in reported CDAD incidence (Pepin *et al., 2*004; Kuijper *et al., 2*007) and as such have been denoted as 'hypervirulent'.

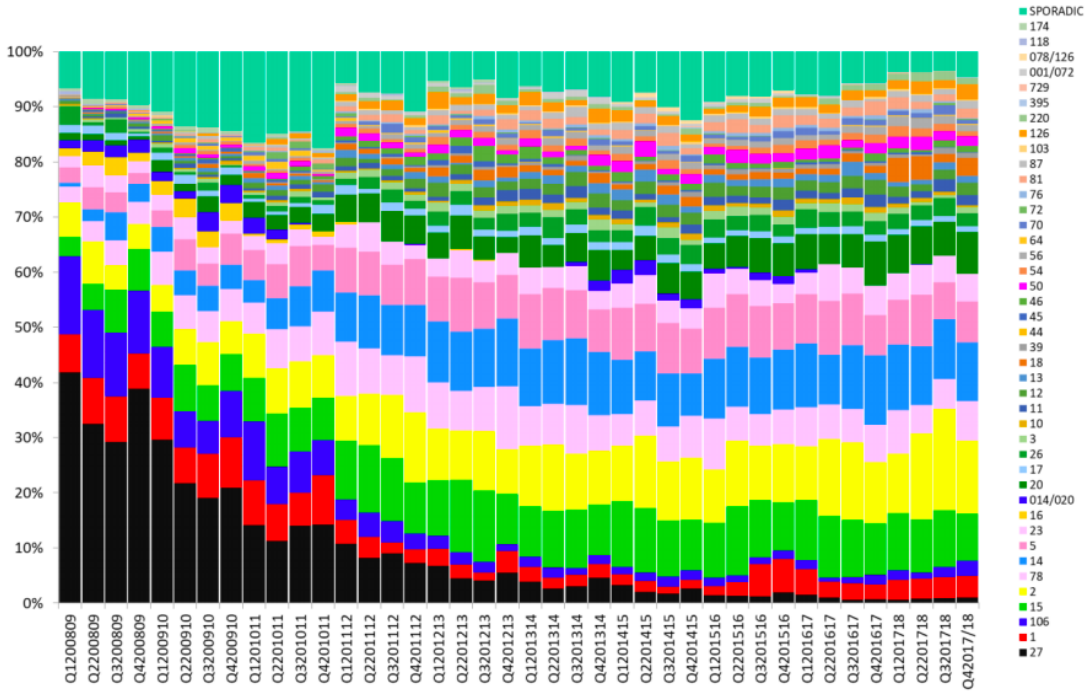


Figure 1.1: Distribution of *Clostridioides difficile* ribotypes in England by quarter from April 2008 – March 2018 from PHE (2018).

The BI/NAP1/027 type was also identified to be responsible for the increase in severe complications of the disease, linked to higher relapse rates and increased mortality, as well as a greater resistance to fluoroquinolones (Kuijper, E *et al.,* 2006). In 2007, the RT078 strain emerged in the Netherlands, it was first isolated in animals and thus is proposed to be involved in zoonotic CDI (Knight and Riley 2019). This strain is shown in humans with similar disease severity and mortality rates as strains of the 027/NAP1 type but was shown to also infect patients younger and more frequently in the community (Knight and Riley 2019). These so called hypervirulent strains both contain the binary toxin gene and possess variants of the

tcdC gene allowing toxin production at all growth phases as proposed by Warny *et al., (*2005). Since 2008 there has been a decrease in the incidence of RT027 strains across England (see Figure 1.2).

To date, WGS and MLST have revealed 6 clades and the additional clades C-II, C-II and C-III within the clonal structure of *C. difficile* lineage (Ramirez-Vargas *et al., 2*018) (see Figure 1.3). Further phylogenetic analysis has indicated that PCR RT's are clade specific (Dingle *et al.,* 2011). The hypervirulent strains RT027 type are found in clade 2. Interest surrounds strains of RT017 found within clade 4 to have a variant toxin profile (A^- B^+ CDT^-), noted to cause widespread disease associated with global outbreaks (Knight *et al.,* 2015). Whilst clade 5 has also been of particular interest, containing RT078 (ST11) strains associated with livestock (Jhung *et al.,* 2008), and RT237 strains found to have atypical rearrangement of toxins A and recovered from pigs and humans with strains (Knight *et al.,* 2013).

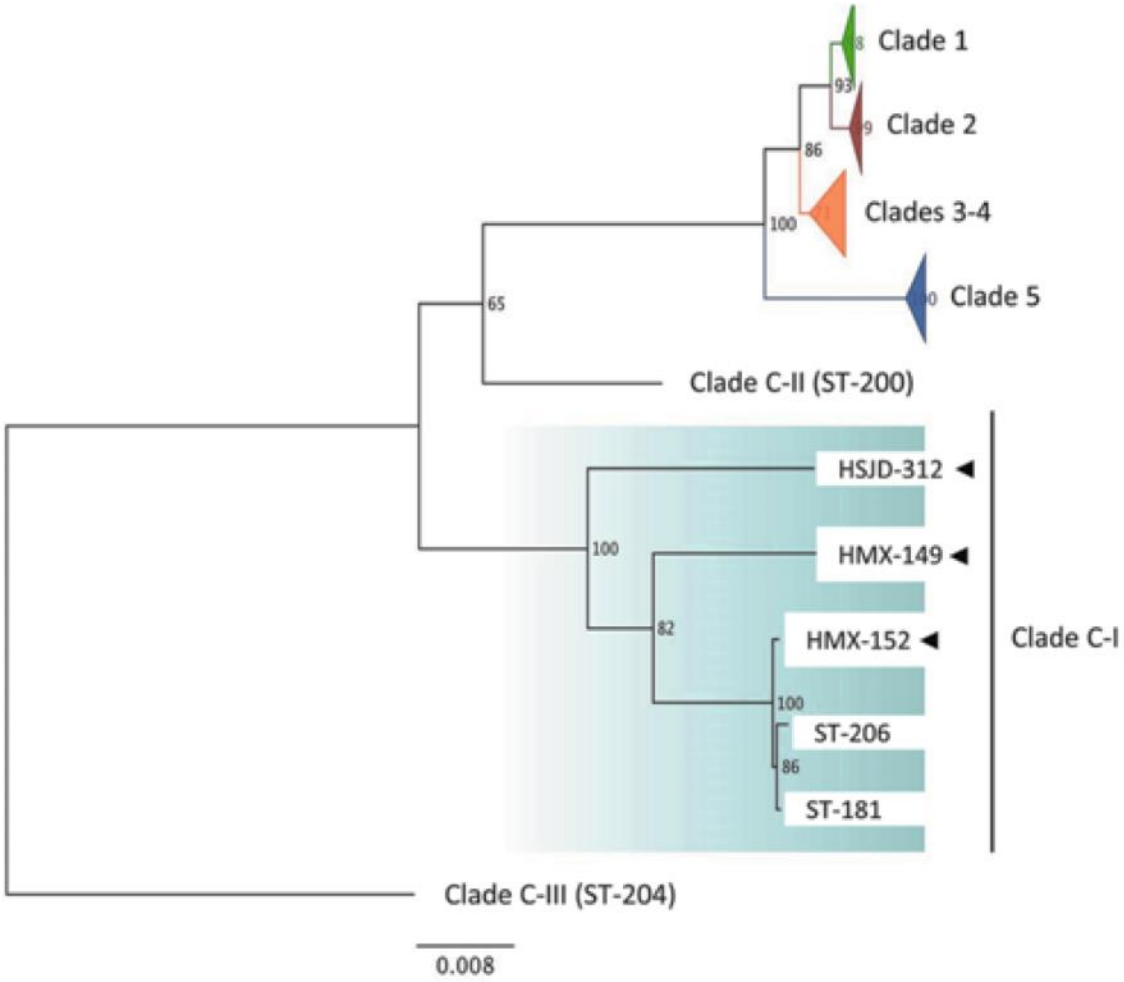


Figure 1.3: Clades of *C. difficile* as determined by whole genome MLST from Ramirez-Vargas *et al.,* (2018).

1.7.2.1 Hypervirulence in sporulation/germination

The increase in disease incidence and severity is likely to be linked to the transmission propensity of these hypervirulent strains (Merrigan *et al.,* 2010). It has been observed that hypervirulent strains sporulate earlier and with greater efficiency than non-hypervirulent strains (Wilcox and Fawley 2000; Fawley *et al.,* 2007; Åkerlund *et al.,* 2008). Enhanced sporulation rates of hypervirulent *C. difficile* would increase the propensity for dissemination of the pathogen into the environment, significantly impacting disease transmission. Spores also contribute to the survival and persistence of *C. difficile* within hosts during antibiotic treatment (Walters *et al.,* 1983; Underwood *et al.,* 2009), epidemic strains possessing altered sporulation efficiencies could therefore be the reason for the reports of higher rates of relapse seen in this group. However, recent studies have shown that an epidemic strain isolate of BI/NAP1/027, R20291, formed less spores more slowly than the non-epidemic strain 012 isolate of CD630Δ*erm* with no association between sporulation rates and ribotypes was observed (Burns *et al.,* 2010; Burns and Minton, 2011).

There has so far been a disparity in the literature towards the germination properties of strains within the so-called 'hypervirulent' BI/NAP1/027 type. Saxton *et al., (*2009) suggest that following antibiotic administration in patients, the germination properties of this type may be altered (Saxton *et al.,* 2009). They found that, in a human gut model, *C. difficile* PCR ribotype 027 germinates more readily, remains in a vegetative form and produces the cytotoxin substantially longer than *C. difficile* PCR ribotype 001. However, the opposite was found later in *in vitro* studies, where the non-epidemic strain CD630, germinated more than that of a R20291 isolate (Burns *et al.,* 2010). It was further highlighted that there was a level of variation between strains within the BI/NAP1/027 group (Heeg *et al., 2*012; Oka *et al., 2*012), with no ribotype specific germination response to sodium taurocholate. Even so, isolates within the 027 ribotype had been later shown to have significantly higher germination efficiencies in response to sodium taurocholate than isolates of types 106 and 078 (Moore *et al., 2*013).

Divergent germination characteristics within hypervirulent strains may presumably allow for increased susceptibility of disease in patients who ingest spores of these types, significantly inducing disease onset and the likelihood of relapse. One study (Oka *et al., 2*012) found no

correlation between the DNA typing group of isolates and infection status (i.e. single, relapse, or reinfection) (Oka *et al.,* 2012). However, this same study showed significantly increased germination rates without sodium taurocholate for isolates recovered from relapse cases. Spores of strains with greater germination efficiencies that would remain in the gut after levels of antibiotics dropped would be able to affect the germination and rapid growth of residual *C. difficile* subsequently causing a relapse of viable disease. Oka *et al.,* (2012) went on to speculate that the isolates recovered from relapse strains may be more sensitive to co-germinants other than taurocholate in the test medium or may have an atypical spore coat structure which facilitates germination. It is plausible to assume that in such cases recurrence may occur more frequently with strains that have phenotypic characteristics advantageous for germination and intestinal survival.

1.7.2.2 Understanding the genetics of sporulation and germination of *C. difficile* and relapse.

C. difficile spores represent the principal transmission route of CDAD (Burns and Minton, 2011). With emergence of so-called 'hypervirulent' strains, linked to increased disease severity, and higher relapse rates, the need to understand how *C. difficile* produces infection and disease remains a valid concern. As these spores must germinate to allow for vegetative cell growth and toxin production, a detailed understanding of both the sporulation and the germination process may have direct applications for disease prevention in healthcare environments.

Strain-to-strain variation of sporulation and germination characteristics currently shows that these features do not appear to be ribotype associated. It can be criticised that, to date, studies have varied extensively in the methods employed and have been limited to as few as one representative strain per PCR-Ribotype. Standardisation of procedures needs to be implemented if meaningful comparisons between studies are to be made. More studies including larger sample sizes across more strains are needed to confirm whether such a relationship for relapse and germination exists, together with extensive molecular studies to

identify what components 'hypervirulent' strains may have that allow them to cause more severe disease and relapse.

1.8 Project Aims

The aim of this project is to undertake molecular and genetic studies to further define and understand the processes that control the initiation of sporulation and germination.

To perform functional assignments to hypothetical proteins in sporulation and germination pathways of *C. diffcile* to complete the regulation cascade.

To establish a link between germination/sporulation phenotype and relapse of clinical isolates.

Chapter 2

Chapter 2: Materials and Methods

2.1 Selection of clinical isolates

2.1.1 Study design and patient population

Clinical isolates of *C. difficile* for this study were obtained from a multi-centre, two-arm parallel group double-blind, randomized, placebo-controlled trial to assess the use of rifaximin as a measure to control relapse of CDI. Eligible patients were recruited from 2012 to 2014, were at least 18 years old, successful in the treatment of clinically diagnosed *C. difficile*-associated diarrhoea (CDAD) using standard therapy (metronidazole or vancomycin) given according to standard local hospital guidelines, able to swallow tablets, able to stop chronic antibiotic use, women of child bearing potential willing and able to use at least one highly effective contraceptive method throughout the study. Patients were excluded if they were pregnant or breast feeding, hypersensitive to the active substance, to any rifamycin (e.g. rifampicin or rifabutin) or to any of its excipients (Tablet core: Sodium starch glycolate type A, glycerol distearate, colloidal anhydrous, silica, talc and microcrystalline cellulose. Tablet coating: hypromellose, titanium dioxide (E171), disodium edentate, propylene glycol and red iron oxide E172), or >5 days post standard therapy (metronidazole or vancomycin) for clinically diagnosed CDAD. The study was approved by The National Research Ethics Service Committee East Midlands and Leicester and gained multi-centre study approval by the Multi-Research Ethics Committee (study reference 12/EM/0292) and all patients signed informed consent.

2.1.2 Randomization and follow-up

Enrolled patients were randomly assigned in a 1:1 ratio to receive oral Rifaximin (200 mg tablets) or a placebo at a dosage of 2 x 200 mg tablets three times a day for the first 2 weeks, then 1 x 200 mg tablet three times a day for the final 2 weeks, immediately after finishing a 10–14 day course of standard anti-CDI antibiotics. Patients were randomized after a positive toxin test was reported and before the end of standard anti-CDI therapy. Randomization was

performed by the investigational drug pharmacist at the hospital who was not involved in the conduct of the study. Study medication and matching placebo were dispensed with a specific study number to ensure blinding of investigators and patients. All patients were inpatients at the time of randomization.

2.1.3 Efficacy assessments

The primary endpoint was a difference in the percentage of relapse between rifaximin and placebo at week 12. Secondary outcome measures assessed outside of this thesis included bowel symptoms, length of stay, safety/adverse events and the difference in relapse of CDAD within 6 months of start of therapy.

2.1.4 Sampling schedule

Stools were submitted from each patient, where possible, at 4 time-points during treatment and follow-up periods (see Table 2.1). Up to 20 colonies from each sample was investigated for the presence of *C. difficile* to allow maximal screening of recurrence by reinfection or relapse causing strains. The work involved in the isolation of *C. difficile* from stool was carried out by a team of five; N. Kinsmore, Dr E. Stevenson, Dr P. Budd, Dr C. Meaney and myself at the Biomedical Research Unit of Queens Medical Centre, University of Nottingham Teaching Hospital Trust.

Table 2.1: A schematic of stool sample submission to the RAPID trial study

Trial time point/Sample name	Sample submission	Trial stage
Week - 1	CDI Diagnosis [a]	Treatment period
Week 0	Chaser therapy start [b]	
Week 4	Chaser therapy end [c]	Follow-up period
Week 12	3-month follow up [d]	
Relapse (RL)	Relapse sample submission [e]	

a = Diagnostic sampling of patient whilst on standard therapy of metronidazole or vancomycin for the identification and confirmation of *C. difficile* infection
b = Sampling at start of 4 -week trial chaser therapy of placebo or rifaximin for the identification of *C. difficile* carriage
c = End of treatment period sampling for the identification of post chaser therapy carriage of *C. difficile*
d = Sampling three months post standard therapy for the identification of asymptomatic and persistent carriage of *C. difficile*
e = Relapse sampling of any symptomatic infection, 3-months post standard therapy

2.1.5 Isolation of *C. difficile* from stool

Frozen stool samples were thawed and aliquots of 300 µL are mixed with PBS 1:1 by vortexing. Samples were heat shocked at 80 °C for 15 minutes before being centrifuged at 1,500 x *g* for 5 minutes. The supernatant was then transferred to the anaerobic workstation (model MG1000 Don Whitley scientific Limited, United Kingdom) and 50 µL spread onto three separate pre-reduced Brazier (Oxiod), egg yolk cefoxitin cycloserine (CCEY) agar plates. Following incubation for 48 hours, up to 20 colonies for each stool specimen where suspended in 200 µL BHIS broth overnight in a 96-well plate. A 10-fold dilution in PCR-grade water of this culture was stored at -20 °C as crude DNA templates for colony PCR, whilst the remaining culture was stored at -80 °C following the addition of ~10 µL of 100% glycerol.

2.2 Genotypic characterization of clinical isolates.

2.2.1 PCR Ribotyping

Every colony that was isolated from each stool sample was subjected to in-house ribotyping. PCR amplification of the 16S rRNA intergenic spacer region was carried out according to a modified protocol obtained from the former Anaerobe Reference Laboratory in Cardiff. Briefly, the 96-well plates containing 10-fold dilutions of culture in PCR-grade water were defrosted and heated to 95 °C for 20 mins. Into a clean 96-well plate, a 10 µL aliquot of heated sample or control was transferred into corresponding wells, with the addition of 40 µL master mix. Plates were sealed with sealing tape and transferred into a thermocycler for PCR. Master mix components and ribotyping PCR program can be found in Table 2.2 and Table 2.3, respectively. Following the PCR, the seal film was removed, and samples heated to 75 °C until ~20 µL reaction remained. PCR-ribotype profiles were analysed by capillary electrophoresis as described in section 2.6.13. Individual profiles were assessed, and then one isolate from each distinct typing profile that had been obtained from each stool sample was re-cultured onto BHIS-CC agar and stored as a glycerol stock. These stock strains were sent for confirmation via the *C. difficile* ribotyping network (CDRN) service in Leeds and those confirmed as *C. difficile* where used for downstream characterization.

Table 2.2: Ribotyping master mix composition

Component	Compound	Volume (µL)
Ribotyping Master mix (1 reaction)	P3 primer (100µM)	0.25
	P5 primer (100µM)	0.25
	dNTPs in 5X reaction buffer	10
	1.5 mM $MgCl_2$	1.5
	OneTaq	0.25
	H_2O	37.75
dNTPs in 5X reaction buffer	PCR-grade H_2O	300
	Polymerase Buffer (supplied with OneTaq)	500
	dNTPs	200

Table 2.3: Ribotyping PCR cycle

Step	Temperature °C	Time (M:S[1])	Cycles
1	95	02:00	-
2	92	01:00	30 cycles of step 2 - 4
3	55	01:00	
4	72	01:30	
5	95	01:00	-
6	55	00:04	-
7	72	05:00	-
8	14	Hold	-

[1]M:S = Minutes:Seconds

2.2.1 List of 37 clinical *C. difficile* isolates selected for further characterisation

Table 2.4. Clinical *C. difficile* isolates used in this study

Isolate	Source	Enrolment site
01008 WK-1 (E2)	Stools of In-patient, BRU-ID 01008	Queens Medical Center, Nottingham, UK
01008 WK-1 (E10)		
01008 WK12 (G2)		
01008 RL (H4)		
01013 WK-1 (C17)	Stools of In-patient, BRU-ID 01013	Queens Medical Center, Nottingham, UK
01013 WK-1 (C19)		
01013 WK12		
01016 WK-1	Stools of In-patient, BRU-ID 01016	Queens Medical Center, Nottingham, UK
01016 WK12		
01018 WK-1 (C17)	Stool of In-patient, BRU-ID 01018	Queens Medical Center, Nottingham, UK
01018 WK-1 (C20)		
11001 WK4	Stools of In-patient, BRU-ID 11001	Kettering General Hospital, Kettering, UK
11001 RL		
11007 WK-1 (C2)	Stools of In-patient, BRU-ID 11007	Kettering General Hospital, Kettering, UK
11007 WK-1 (C5)		
11007 WK0		
11007 RL		
12001 WK 0	Stools of In-patient, BRU-ID 12001	Dorset County Hospital, Dorset, UK
12001RL		
12004 WK0	Stools of In-patient, BRU-ID 12004	Dorset County Hospital, UK
12004 RL		
18001 WK0	Stools of In-patient, BRU-ID 18001	New Cross Hospital, Wolverhampton, UK
18001 RL		
01006 WK-1	Stools of In-patient, BRU-ID 01006	Queens Medical Center, Nottingham, UK
01006 WK0		
01006 WK12		
01015 WK-1	Stools of In-patient, BRU-ID 01015	Queens Medical Center, Nottingham, UK
01015 WK4		
01017 WK4	Stool of In-patient, BRU-ID 01017	Queens Medical Center, Nottingham, UK
01023 WK-1	Stools of In-patient, BRU-ID 01023	Queens Medical Center, Nottingham, UK
01023 WK0		
01028 WK-1	Stools of In-patient, BRU-ID 01028	Queens Medical Center, Nottingham, UK
01028 WK0		
03008 WK0	Stools of In-patient, BRU-ID 03008	Kings Mill Hospital, Sutton in Ashfield, UK
03008 WK4		

2.2.2 Whole genome sequencing and assembly

Genomic DNA of *C. difficile* was extracted from strains using an extraction method with phenol-chloroform-isoamyl alcohol (25:24:1). Purity and quantity were measured using the NanoDrop spectrophotometer (ThermoFischer Scientific) and Qubit 3.0 Fluorometer (ThermoFischer Scientific) according to manufacturer's instructions. Genomic DNA of all isolates were sent for Illumina sequencing using MisSeq 250-PE technology (MicrobeNG, University of Birmingham). Where available, genomic DNA from diagnostic isolates were sent for Pacific Bioscience (PacBio) sequencing (McGill University and Genome Québec Innovation Centre), by Dr E. Stevenson, University of Nottingham. Paired-end reads from the MiSeq runs of treatment and follow-up isolates belonging to the same PCR ribotype of their corresponding diagnostic sample were mapped to PacBio contigs using CLC Genomics Workbench, version 8.0.2 (Qiagen) as the reference. *De novo* mapping of trimmed Illumina reads was carried out using the *de novo* mapping and assembly workflow of CLC genomics workbench software. Consensus contigs were exported in fasta format for downstream applications.

2.2.3 Genome annotation

Genomes were annotated using the automated PROKKA (Seemann, 2014) pipeline, with the following settings "-- compliant -- genus Clostridioides use genus".

2.2.4 Comparative genome analysis

The basic variant detection workflow (CLC Genomics Workbench, version 8.0.2 [Qiagen]) was used to identify single-nucleotide variations (SNVs), insertions, and deletions of sequences between isolates of the same PCR ribotype from the same patient. Base pair calls were identified across genomes using CLC's fixed ploidy variant detection with default noise filters. A consensus of ≥75% was required to support a call, of which were required to be homozygous under a diploid model. Only calls supported by >5 reads, including one in each

direction were accepted. Variant calls that were identified after default quality filtering were used to identify single and multiple nucleotide variants (SNV and MNV) between sequenced isolates from the same patient infected with the same ribotype. Isolates with reinfections from different ribotypes were not subject to SNP analysis. Default parameters and settings were used for all tools unless stated otherwise.

A phylogenetic tree of genomes was built using the program bcgTree (version 1.1.0) which implements a pipeline for tree building based on essential single-copy core bacterial genes, based on a Hidden Markov model (Ankenbrand and Keller 2016), with 100 tree iterations were performed for RAxML bootstrapping using ggtree in R package version 0.9.4 (Guangchuang *et al.,* 2017; 2018). This program requires coding regions translated to protein sequences as input. The identification of coding regions and translation of nucleotide to protein sequences was automatically performed using Prodigal (version 2.6.3) during the PROKKA annotation of genomes outputted as gff format files, which were used as the input for phylogenic tree construction. The reference genome *C. difficile* 630 was downloaded from Genbank (Assembly ASM920v2) (Sebaihi *et al.,* 2006) and included to provide an indication of similarity to the type strain, while genome of *C. perfringens* ATCC 13124 was downloaded from Genbank (Assembly ASM1328v1) (Myers *et al.,* 2006) to act as an outgroup.

To further assess relatedness of strains, the average nucleotide identity (ANI) was calculated. Using the ANIcalculator (version 1) (Varghese *et al.,* 2015), a program to bi-directionally align two genomes, the fraction aligned, and the matching percent identity of the aligned regions was reported as a single ANI value. The max ANI possible is 100%, while the max alignment fraction is 1. The R suite with additional packages ggplot2 and cowplot was used to visualize this data as a heatmap of ANI, revealing clustering patterns in the genomes.

To identify sequence variation in genes of interest from the reference *C. difficile* strain CD630 (GenBank Assembly ASM920v2), nucleotide and amino acid sequence variations in selected genomes of isolates were further investigated. From a list of 24 genes of interest, chosen for their association to *C. difficile* virulence, nucleotide fasta sequences were downloaded from Genbank using their CD630 gene locus tags as the search term. Using BLASTn (version 2.2.28+) to run each reference gene against each genome, the best hit was selected using the default BLAST output format to compare sequence alignment and identity. Protein fasta sequences

of the 24 genes of interest were also downloaded from Genbank using their locus tags as the search term. BLASTp (version 2.2.28+) was used to compare these sequences to the previously translated protein sequences for each genome, outputted from prokka genome annotation, to reveal amino acid identity. For sequences that were not all identical, amino acid sequence alignments were made against the reference sequence and the matching isolate sequences using MUSCLE (version 3.8.31) and visualized alignments in the UGENE (version 1.9.8) graphical interface.

2.2.5 Pan-genome analysis using the BPGA pipeline

The Pan-genome analysis was carried out using the Bacterial Pan Genome Analysis (BPGA) pipeline (Chaudhari, Gupta, and Dutta, 2016). BPGA programme was run through Windows command line, performed with default clustering tool, USEARCH. Neighbour joining trees of concatenated core genes generated through BPGA pipeline were then visualised in CLC Genomics Workbench (version 8.0.2 [Qiagen]).

2.4 Spore techniques

2.4.1 Preparation of *C. difficile* spores in a nutrient-rich liquid medium

Sporulation of *C. difficile* was achieved by incubating cultures anaerobically in BHIS broth for 5 days at 37°C. To ensure that no spores were present when the sporulation medium was inoculated, a starter culture was prepared in BHIS broth using a 1% inoculum of a *C. difficile* overnight culture and incubated until the optical density at 600 nm (OD_{600}) was between 0.2 and 0.5. The sporulation medium was then inoculated with 0.01 volume of this exponential starter culture.

2.4.2 Preparation of spores on BHIS agar

To achieve efficient sporulation of *C. difficile*, 100 µL of overnight culture were spread onto 10-15 BHIS agar plates and incubated anaerobically for a minimum of 5 days. Plates were then incubated at 4°C for 3-4 hours, to loosen the spore cell mixture, which was harvested in 1 mL sterile dH_2O. Spore-cell mixtures were incubated overnight at 4°C to encourage the release of mature spores from the mother cell (Sorg and Sonenshein 2010) before being stored at -20 °C.

2.4.3 Purification of spores

Spore-cell mixtures were thawed and centrifuged at 4°C for 4 minutes at 16, 200 x *g.* The pellet was re-suspended in 1 mL of ice-cold sterile dH_2O and washed for a minimum 10 wash steps, removing cell debris each time. The culture was then estimated to be ≥98 % pure under phase-contrast microscopy and stored at -20 °C in dH_2O.

2.4.4 Measurement of *C. difficile* heat-resistant colony formation

Sporulation cultures were set up as described in section 2.4.2. During the 5-day culture, samples (500 µL) were removed from the anaerobic cabinet at different time points (0h (T_0), 24h, 48h, 72h, 96h and 120h) and heated at 65 °C for 30 min to kill the vegetative cells but not the spores. To ensure no contamination with vegetative cells for heat-treated samples, aliquots were centrifuged for 3s, to allow all the suspension was evenly heated and that none remained in the lid of the microcentrifuge tube. To control for any effects of oxygen exposure during heat treatment, a non-heated sample was also removed from the anaerobic cabinet for 30 min and left on the bench. Samples were then serially diluted in phosphate-buffered saline (PBS) and returned to the anaerobic cabinet, where they were plated onto BHIS agar supplemented with 0.1% taurocholate (Sigma, United Kingdom). *C. difficile spo0A* ClosTron mutants of 630Δ*erm* were used as sporulation-negative controls (Heap *et al.,* 2009). Plates were incubated for 24 h before colony forming units (CFUs) were enumerated.

2.4.5 Counting *C. difficile* spores by phase-contrast microscopy

Sporulation cultures were set up as described in sections 2.4.2 and 2.4.3 Following incubation in BHIS broth for five days, spores were enumerated by using a Bright-Line haemocytometer (Sigma, United Kingdom) in a Nikon Eclipse Ci phase contrast microscope under the 100x objective and Ph3 annular diaphragm. To load the haemocytometer, the coverslip was secured and placed in the appropriate position, and 7-10 µL of *C. difficile* spore culture was loaded onto the haemocytometer. *C. difficile* spores were then visualised by phase-contrast microscopy. Spores were counted in 10, 200 µm haemocytometer squares, averaged, and expressed as spores/mL using the appropriate dilution factors.

2.4.6 Assaying the viability of *C. difficile* spores by decoating

Sporulation cultures were set up and purified as described in sections 2.4.2 and 2.4.3. Following incubation in BHIS broth for five days, cultures were repeatedly washed with ice-

cold deionised H_2O (dH_2O) until they contained no cell debris or vegetative cells as observed by phase-contrast microscopy, as described in section 2.4.5. Spores were stored at -20 °C in dH_2O. Purified spores were decoated by resuspending a spore pellet in 1 mL of 50 mM Tris-HCl (pH 8.0)-8 M urea-1% (wt/vol) sodium dodecyl-sulfate-50 mM dithiothreitol (DDT) and incubating the preparation at 37°C for 90 min (Popham, Sengupta, and Setlow, 1995). Following incubation, decoated spores were washed three times in PBS. Samples were then heat treated as described above, serially diluted and plated onto BHIS medium supplemented with lysozyme (1 µg/mL). *C. difficile sleC* ClosTron mutant of 630Δ*erm* were used as germination affected control.

2.4.7 Measurement of germination initiation

Pure spore cultures were prepared as described in sections 2.4.2 and 2.4.3. Cultures were adjusted to have an OD_{600} of 1.0, split into 500 µL, heat treated at 65 °C for 30 minutes and centrifuged for 4 min at 16,200 k x g (the remaining pellet of 500 µL pure spore aliquot was stored at -20 °C for subsequent outgrowth assays that same day). Spore pellets were re-suspended in 1mL pre-reduced BHIS broth and 180 µL of spore suspension and a negative control of 180 µL BHIS broth only added to wells of a microtitre plate in triplicate. The wells were then inoculated with 20 µL BHIS broth + 1% taurocholate (TC), to give a final TC concentration of 0.1%. The plate was sealed with a multiwall plate seal (Greiner Bioone, UK) immediately transferred to a preheated GloMax-Multi microplate reader (Promega, UK) and the initiation of germination measured as a loss of OD_{600} over a 120 min period at 37 °C.

2.4.8 Measurement of spore outgrowth

A pelleted aliquot of pure spores were prepared as described in sections 2.4.2 and 2.4.3. Spore pellets were re-suspended in 1 mL pre-reduced BHIS broth and 180 µL of spore suspension and a negative control of 180 µL BHIS broth only added to wells of a microtitre plate in triplicate. The wells were then inoculated with 20 µL BHIS broth + 1% taurocholate, to give a final TC concentration of 0.1%. The plate was sealed with a multiwall plate seal

(Greiner Bioone, UK) immediately transferred to a preheated GloMax-Multi microplate reader (Promega, UK) and outgrowth measured as a rise in OD_{600} over a 16 h period at 37 °C.

2.4.9 Measurement of loss of spore heat resistance

Pure spore cultures were prepared as described in sections 2.4.2 and 2.4.3. Cultures were adjusted to have an OD_{600} of 1.0 in 1 mL H_2O and heat treated at 65 °C for 30 minutes. Ten aliquots of 100 µL were centrifuged for 4 min at 16,200 x *g*. One aliquot was resuspended in 100 µL PBS, labelled as time point zero (T_0) and incubated on ice. The remaining spore pellets were re-suspended in 500 µL 0.1% TC-BHIS broth and incubated for 5, 10, 20, 30, 60, 90, 120, 180 or 240 mins to allow germination at room temperature. At the end of each time point, the aliquoted spore/germinant mixture was immediately heat treated for 30 min at 65 to kill germinated-vegetative cells, leaving only non-germinating spores behind. Once heated, the aliquots are placed on ice until the final time point is heated and all are serially diluted in PBS including T_0. Dilutions were then transferred to the anaerobic cabinet, where they were plated onto BHIS agar supplemented with 0.1% taurocholate (Sigma, United Kingdom). Plates were incubated for 24 h before CFU were enumerated. Heat resistance, represented as CFU/mL, was calculated as CFU/mL at T_0 subtracted by CFU/mL at time point. An experimental flowchart of this assay is presented in Figure 7.2 of the appendix.

2.4.10 Statistical analysis

Statistical analysis was performed using GraphPad Prism, version 8.4.3 (GraphPad Software, San Diego California USA) and Microsoft Excel, version 16.42 (Microsoft Office 365, Redmond Washington, USA). A statistical significance level of $P < 0.05$ was adopted, were P values <0.05 considered significant, <0.01 highly significant and <0.001 and <0.0001 very highly significant.

2.4.11 List of other bacterial strains used in the study

Table 2.5: List of other bacterial strains used in this study

Strain ID	Description	Source/Reference
CD630Δ*erm*	*C. difficile* 630Δ*erm*	Hussain *et al.,* 2005
EK28	Clinical *C. difficile* isolate PCR RT 078	Ed Kujiper
CD630 Δ*erm* Δ*spo0A*	*C. difficile* 630Δ*erm*Δ*spo0A*::intron *ermB*	Heap*, et al.,* 2007
CD630 Δ*erm* Δ*sleC*	*C. difficile* 630Δ*erm*Δ*sleC*::intron *ermB*	Burns *et al.,* 2010
TOP10	*Escherichia coli* TOP10	Invitrogen
CA434	*Escherichia coli* CA434	Purdy *et al.,* 2002
CT19300	*C. difficile* 630Δ*erm*_CD19300::intron *ermB*	This study
CT24700	*C. difficile* 630Δ*erm*_CD24700::intron *ermB*	This study
CT05460	*C. difficile* 630Δ*erm* _CD05460::CDintron *ermB*	This study
CT07730	*C. difficile* 630Δ*erm* _CD07730::intron *ermB*	This study
CT07750	*C. difficile* 630Δ*erm* _CD07750::intron *ermB*	This study
CT08670	*C. difficile* 630Δ*erm* _CD08670::intron *ermB*	This study
CT29630	*C. difficile* 630Δ*erm* _CD29630::intron *ermB*	This study
CRG000	CD630 Δ*erm*Δ*pyrE*	Nigel Minton
CT24700 Δ*pyrE*	CD630 Δ*erm* Δ*pyrE:CT24700*	This study
CT24700 + $_{chr}$Comp	CD630 Δ*erm* Δ*pyrE:CT24700:pyrE:CD24700:fdxpromoter*	This study
CT24700 + $_{chr}$Over	CD630 Δ*erm* Δ*pyrE:CT24700:pyrE:CD24700:nativepromoter*	This study
BL21(DE3*)	*Escherichia coli* BL21(DE3*)	New England BioLabs (Catalogue # C2527)

2.4.12 List of plasmids

Table 2.6: List of plasmids used in this study

Plasmid	Description	Source
pMTL007C-E2	ClosTron plasmid for retargeting	DNA 2.0, Heap *et al.,* 2010
pMTL007C-E2::CD0546-164/165a	ClosTron plasmid retargeted for intron insertion at position 164/165a of the *C. difficile* CD0546 gene	DNA 2.0
pMTL007C-E2::CD0773-155/156a	ClosTron plasmid retargeted for intron insertion at 155/156a of the *C. difficile* CD0773 gene	DNA 2.0
pMTL007C-E2::CD0775-167/168a	ClosTron plasmid retargeted for intron insertion at 167/168a of the *C. difficile* CD0775 gene	DNA 2.0
pMTL-RN1	ACE vector with pCD6 replicon and native *pyrE* plasmid to correct *pyrE* truncation	Minton
pMTL-RN1c::CD24700	ACE vector used to insert GPR with native promoter into chromosome with simultaneous *pyrE* correction	This study
pMTL-RN1x::CD24700	ACE vector used to insert GPR with *fdx* promoter into chromosome with simultaneous pyrE correction	This study
pMTL-ME2	Allelic exchange plasmid containing pCD6 replicon and truncated *pyrE* gene from *C. acetobutylicum*	Minton
pMTL-ME2-CD12900KOC	pMTL-ME2 plasmid carrying a recombinant allele for the deletion of CD12900	This study
pMTL-ME2-CD26880KOC	pMTL-ME2 plasmid carrying a recombinant allele for the deletion of CD26880	This study
pMTL-ME2-CD32490KOC	pMTL-ME2 plasmid carrying a recombinant allele for the deletion of CD32490	This study
pET28b+	pET system expression vector for IPTG inducible expression of recombinant N-, C- terminally 6XHis-tagged proteins in *E. coli*,	Novagen
pET28b+::CD24700	pET system expression vector for IPTG inducible expression of recombinant N-, C- terminally 6XHis-tagged CD24700 protein in *E. coli*.	This study

2.5 Bacterial strains, media and culture conditions

2.5.1 Bacterial strains

The lab derived bacterial strains used in this study include *Clostridioides difficile* strain 630 and *C. difficile* EK28. *Esherica coli* TOP10 (Invitrogen) was used as the cloning host for pMTL-based plasmids and *E. coli* CA434(151) for conjugation donor strain for transferring plasmid DNA into *C. diffcile. E. coli* BL21(DE3) was used for hosting constructed pET vector plasmids. A list of *C. difficile* strains can be found in Table 2.5.

2.5.2 Culture media

C. difficile strains were grown pre-reduced brain heart infusion (BHI) supplemented with 0.5% w/v yeast extract (Oxoid) and 0.1% w/v L-cysteine (Sigma, UK), or *Clostridioides difficile* Minimal Medium (CDMM) (see Table 7.1 of the appendix) (Karlsson *et al.,* 1999). *C. difficile* was grown in an anaerobic workstation (Don Whitley Scientific) containing an environment composed of 80% N_2, 10% CO_2 and 10 % H_2 at 37°C. *E. coli* strains were routinely grown on Lysogeny Broth (Sigma) or in LB broth (Fisher Chemicals). Expression for protein purification was carried out using *E. coli* strain BL21(DE3*) (Novagen) with cultures were grown with shaking at 37°C. Liquid and agar BHIS, CDMM, LB and SOC medium (Invitrogen) were supplemented where appropriate with additives and antibiotics as listed in Table 2.7. Bacterial strains were stored in 20% glycerol at -80°C.

2.5.3 Plasmids

All plasmids used in this study are listed in Table 2.6. Plasmid maps were designed and analysed using Snapgene or Benchling software. The re-targeted vectors carrying target genes were insertionally inactivated with the ClosTron system (Heap *et al., 2*007) using the modular ClosTron plasmid pMTL007C-E2 (Heap *et al., 2*009) and were pre-designed by the Clostridium Research Group (CRG) and delivered, sequence-verified, ready for transfer into *C. difficile* from DNA 2.0.

2.5.4 Anaerobic and aerobic culture conditions

All *C. difficile* strains were grown in an anaerobic workstation (model MG1000 Don Whitley Scientific Limited, United Kingdom) with a gas mixture of 80% nitrogen, 10% hydrogen and 10% carbon dioxide set at a growth temperature of 37 °C. All *C. difficile* liquid media were allowed to reach anaerobicity with overnight reduction in the anaerobic cabinet, and all solid growth media were allowed to reach anaerobicity by reducing for at least 4 h prior to inoculation. All *Escherichia coli* strains were grown using Lysogeny-Broth (LB) broth or agar as appropriate at 37 °C, with 200 rpm shaking for liquid cultures. Media was supplemented with antibiotics as appropriate, listed in Table 2.7. To select against non-clostridial organisms, growth media was supplemented with Cefoxitin 8 µg /mL and Cycloserine 250 µg /mL. When selecting for *E. coli* harbouring pMTL or pET vector construction plasmids, growth media was supplemented with chloramphenicol (25 µg/mL in agar or 12.5 µg /mL in broth), or Kanamycin (50 µg /mL), respectively. All strains were stored in their respective culture media with 10% glycerol (filter sterilized) in a -80 °C freezer.

Table 2.7. Antibiotic and growth supplements

Supplement	**Working concentration *C. difficile* (µg/mL)**	**Working concentration *E. coli* (µg/mL)**
5-FOA	4 mg/mL	-
Cefoxitin	8	-
Cycloserine	250	-
Chloramphenicol	-	25 (in agar) 12.5 (in liquid medium)
Erythromycin	10	100
IPTG	4 mM	-
Kanamycin	-	50
Lysozyme	1	-
Sodium taurocholate	1% (w/v)	-
Thiamphenicol	15 (in agar) 7.5 (in liquid medium)	-

2.5.5 Preparation of electrocompetent *E. coli* cells

A 1 mL overnight culture of *E. coli* was used to inoculate 200 mL of LB broth (pre-warmed to 37 °C) in 500 mL conical flasks. *E. coli* cells were cultured at 37 °C with 200 rpm shaking until an optical density at 600 nm (OD_{600}) of between 0.5 and 1 was reached, indicative of the exponential growth phase. Cells were harvested by pooling into separate 50 mL falcon tubes, cooling on ice for 30 min, and then centrifugation for 15 min at 4000 × *g* in a rotor pre-chilled to 4 °C. Pellets were then gently re-suspended in 40 mL of chilled, sterile deionised H_2O (dH_2O) before centrifugation. Following centrifugation, the pellets were re-suspended in 500 µL of sterile 10% glycerol, pooled, and stored at -80 °C in 50 µL aliquots.

2.6 Molecular biology techniques

2.6.1 Oligo design

Oligonucleotide primers for the amplification of DNA by polymerase chain reaction, Gibson assembly or NEB HiFi DNA assembly were designed manually using FASTA format sequences derived from genome sequences and analysed using Snapgene software. A list of oligonucleotides used in this study are shown in Table 2.8.

Table 2.8: List of oligonucleotides used in this study

Primer name	Sequence	Source
CD0546-F1	GGGATATGTGCAATTTCTATGGC	This work
CD0546-R1	TGTAAATCCTGCTGAAACGCA	This work
CD0773-F1	AGACCAAATAAGCCCAAAACCT	This work
CD0773-R1	AGGTCGTAAACTCCTAATCCAGT	This work
CD0775-F1	GGTAACGTATGTAACAACTGGTG	This work
CD0775-R1	CTGTAGCAGTTGTACCTCCTAAA	This work
EBS Universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Heap *et al.,* 2007
EBS 2	TGAACGCAATTTCTAAATTTCGATTTAACTTCGAT AGAGGAAAGT	Kuehne *et al.,* 2011
Intron *Sal*I-R1	ATTACTGTGACTGGTTTGCACCACCCTCTTCG	Kuehne *et al.,* 2011
pMTL-YN1 -F	CATGATATAGGTATGCC	Minton
pMTL-YN1-R	GTGCTGCAAGGCGATTAAG	Minton
T7 Promoter	TAATACGACTCACTATAGGG	Eurofins
T7 Terminal	CTAGTTATTGCTCAGCGGT	Eurofins

2.6.2 Plasmids

All plasmids used in this study are listed in Table 2.6. Plasmid maps were designed and analysed using Snapgene, Benchling or NEBuilder software. The re-targeted vectors carrying target genes were insertionally inactivated with the ClosTron system (Heap *et al.,* 2009) and were pre-designed by the Clostridium Research Group and delivered, sequence-verified, ready for transfer into *C. difficile*.

2.6.3 Sequence Checking

DNA sequencing was performed by Source Bioscience, UK. Sequence data analysis was performed on online and software platforms Benchling and SnapGene (version 5.0).

2.6.4 Restriction Endonuclease digestion

Restriction enzymes and buffers were supplied by New England Biolabs (NEB) and reactions were performed according to manufacturer's recommend temperatures and conditions. Unless otherwise stated, all single and double digest reactions were carried out in total volumes of 20 µL with an incubation period of 2 h at 37 °C.

2.6.5 De-phosphorylation of DNA

Antarctic phosphatase (NEB) was used to remove 5' phosphate groups on linearized DNA to prevent self-ligation, according to manufacturer's instructions.

2.6.6 Ligation

Digested DNA products were ligated together using T4 ligase and ligation buffer (Promega, UK) according to manufacturer's recommendations. A control ligation of a vector without insert was also carried out to assess any re-ligation events.

2.6.7 Membrane dialysis of DNA

Ligated products were dialysed to remove excess salts from the ligation buffer through a Millipore 0.25 µm nitrocellulose membrane. The full volume of the ligation reaction was placed on the membrane on dH_2O for 30 min at room temperature prior to transformation.

2.6.8 Extraction of chromosomal DNA

Genomic DNA was extracted by phenol:chloroform:isoamyl alcohol (25:24:1) extraction. An overnight culture of cells was centrifuged at 12,000 x *g* for 2 min. The pellet was resuspended in 180 µL lysozyme and incubated at 37 °C for 30 min with gentle agitation throughout. The cell suspension was additionally treated with 4 µL RNase and incubated at room temperature for a further 30 min. To complete protein degradation 25 µL proteinase K, 85 µL dH_2O and 110 µL 10% w/v sodium doceyl sulphate (SDS) solution was added to the cell suspension and mixed by inversion and incubated at 65 °C for 30 min. This mixture was then transferred to phase lock gel heavy tubes (5Prime, Germany) with an equal volume (400 µL) phenol:chloroform:isoamyl alcohol mix and mixed thoroughly by inversion. This mixture was then centrifuged for 3 min at 16,200 x *g* and the top layer was transferred into a fresh phase lock gel heavy tube. This extraction was repeated twice, before the top layer was transferred into a fresh 1.5 mL microcentrifuge tube containing 0.1 volume 3M sodium acetate and 2.0 volumes ice-cold ethanol. The content was then mixed by inversion and immediately incubated at -80°C for 30 min to facilitate precipitation. The tube was then centrifuged at 16,200 x *g* for 15 minutes and the precipitate was immediately washed in 1mL 70% ethanol

and centrifuged again at 16,200 x *g* for 3 min. The pellet was then air dried for 45 min at room temperature and re-dissolved in 50 µL dH_2O.

2.6.9 Quantification of DNA

Plasmid and chromosomal DNA concentrations were quantified by absorbance at 260/280 nm using the Nano-Drop ND-1000 spectrophotometer (Labtech) according to manufacturer's instructions.

2.6.10 Amplification and assembly of DNA

Amplification of DNA by polymerase chain reaction (PCR) as performed in a thermal cycler block (Biometra Professional Trio Thermocycler) with cycling conditions as shown in Table 2.9. Reactions by PCR were carried out using Phusion, OneTaq and Q5 Hot start high fidelity DNA polymerases and master mix buffers (New England Biolabs) according to manufacturer's instructions. Reactions by NEBuilder HiFi DNA assembly master mix (New England Biolabs) were carried out according to manufacturer's instructions.

Table 2.9: Thermocycling conditions for a routine PCR using NEB polymerase enzymes

Step	Temperature (°C)	Time (M:S[2])
Initial Denaturation	98	0:30
25–35 Cycles	98	0:10
	50–77[1]	0:30
	72	20–30 s/kb
Final Extension	72	2:00
Hold	4–10	-

[1]Annealing temperature calculated using NEBTm Calculator, [2]M:S = Minutes:Seconds

2.6.11 Splicing by overlap extension PCR (SOE-PCR)

Portions of non- flanking DNA can be fused together by a method called overlap extension PCR. A schematic representation of SOE-PCR is shown in Figure 2.1. Firstly, two non-flanking sequence regions are initially amplified using their respective flanking primer pairs. These generated PCR products are purified and pooled to form the template of the final PCR reaction, where they are fused together using hybrid primers containing complementary overlapping sequences of the non-flanking sequence regions. Hybrid primer pairs span splice point regions which indicate where the intermediate DNA sequence is to be 'spliced out', creating an alternate recombinant allele. This technique is used to create recombinant alleles or knock-out cassettes, where restriction endonuclease sites can be incorporated onto flanking primers to allow conventional cloning into knock-out vectors.

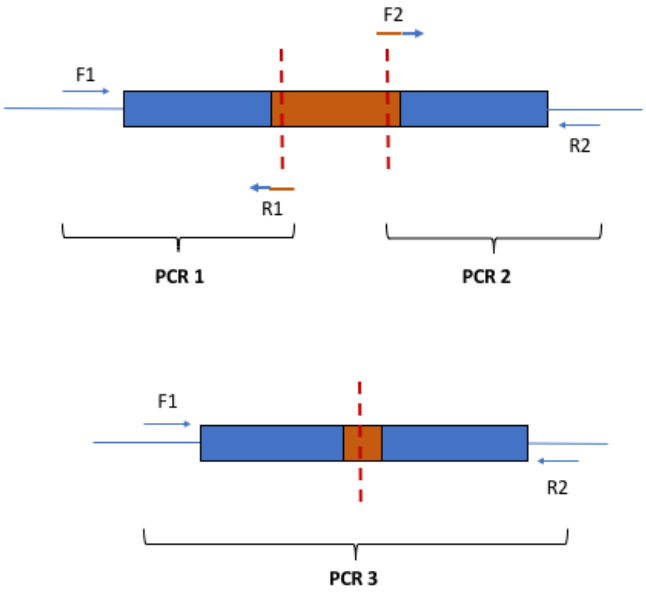


Figure 2.1: Splicing by overlap extension PCR.

This method consists of three rounds of PCR, where flanking (F1 and F2, blue arrows) and hybrid (R1 and F2, blue and orange arrows) primer pairs are used to generate recombinant alleles, splicing out the intermediate sequence (orange) region at splicing points (red dashed line).

2.6.12 Agarose gel electrophoresis

Analysis of PCR products and restriction fragments was performed by agarose gel electrophoresis. Loading dye (Promega) was added to samples according to manufacturer's instructions. A 2-Log DNA ladder (NEB) or λ DNA/HindIII marker (NEB), were used as appropriate. A gel of 0.8-1.5% agarose in Tris-acetate-EDTA (TAE) buffer with 0.001% SYBR Safe green DNA stain was allowed to polymerise and samples run at 90-120 V for 45-110 mins in 1 x TAE buffer. The gel was visualised on a Bio-Rad Gel Doc XR imaging system.

2.6.13 Capillary electrophoresis

PCR-ribotype profiles were analysed with a QIAxcel Advanced capillary electrophoresis system (Qiagen) and QIAxcel DNA High Resolution Kit (Qiagen) according to manufacturer's instructions, using the OL400 program with the QX 15-bp to 1-kb alignment marker and the QX 50-bp to 800-bp size marker.

2.6.14 DNA purification

PCR and digestion fragments were extracted from agarose gel using a scalpel and purified using a Monarch DNA Gel Extraction kit (NEB) according to manufacturer's instructions. PCR products were purified using a Monarch PCR DNA Cleanup Kit according to manufacturer's instructions.

2.6.15 Extraction of plasmids from *E. coli* cultures

Plasmids were isolated using an NEB Monarch Plasmid Miniprep Kit according to manufacturer's instructions.

2.6.16 Plasmid transformation of *E. coli*

A 50 µL aliquot of electro-competent *E. coli* cells was thawed on ice. Plasmid DNA was extracted as described in section 2.6.15. An aliquot of 2 µL DNA was then added to the *E. coli* cells aliquot and gently mixed with a pipette tip. The mixture was then gently pipetted into a cold 0.2 cm gap electroporation cuvette and a pulse applied across the cuvette using an electroporator with 2.5kV voltage, 25 µF capacitance and 200 Ω resistance. A 400 µL supplement of super optimal broth with catabolite repression (SOC) medium (Invitrogen) was then immediately added to the mixture and the entire contents of the electroporation cuvette were then transferred to a microcentrifuge (Eppendorf) tube and immediately placed on ice until ready for incubation at 37 °C, at 200 rpm, for 1 h. This incubation step is necessary to allow expression of antibiotic resistance genes. The mixture was then plated onto LB agar at suitable dilutions, supplemented with appropriate antibiotics for selection (see Table 2.7).

2.6.17 Conjugation

Plasmid DNA was transferred into *C. difficile* by conjugation from an *E. coli* donor CA434, as previously described (Purdy *et al., 2*002). Briefly, both donor and recipient cells were grown to stationary phase in overnight cultures and washed with phosphate-buffered saline (PBS) as necessary to remove traces of antibiotics. A pellet of donor cells was prepared by centrifugation for 2 min at 4000 × *g*, before being moved to the anaerobic cabinet and re-suspended in a suspension of recipient cells. The donor-recipient suspension was then inoculated onto BHIS agar in discrete spots and incubated anaerobically at 37 °C. For conjugation into *C. difficile* 630Δ*erm* (Hussain *et al.,* 2005), suspensions were incubated for 4-8 h. Cells were then re-suspended in 750 µL of PBS by gently harvesting and emulsifying colonies, and then plated onto media containing cycloserine (250 µg/mL), cefoxitin (8 µg/mL) or thiamphenicol (15 µg/mL) as appropriate to counter select against growth of *E. coli* donor strains and to positively select for *C. difficile* transconjugants. Plates were incubated anaerobically at 37 °C for 24-72 h before colonies were picked and purified by subculture for further analysis.

2.6.18 Southern blot

2 µL genomic DNA (gDNA) was digested O/N with 5 µL restriction buffer (New England Bioscience (NEB)), 2 µL restriction enzyme (HindIII or EcoRV) (NEB) and 16 µL dH_2O and then separated O/N on a 0.8 % agarose gel at 110 V with an untreated λ DNA/HindIII marker (NEB). Nucleic acids were then transferred to a positively change nylon membrane. Briefly, the agarose gel was incubated on wick paper in 0.4M NaOH and overlaid with the gel sized membrane sheet pre-soaked in 0.4M NaOH. A stack of dry blotting sheets was placed on top of the filter paper and weighed down. The stack assembly was left to transfer O/N. The transfer stack was then disassembled, and the membrane UV fixed for 2 min. The membrane was then washed in 2 X saline-sodium-citrate (SSC) in a hybridisation tube. After washing, the membrane was immersed in 10 mL DIG Easy Hyb buffer (Roche) with 20 µL of complementary DIG- labelled intron- specific probe, generated by PCR amplification of intron template using EBS2 and *Sal*I-R1 primers and dioxigenin (DIG) high prime reaction. Hybridization was carried out O/N at in a hybridization oven at 42°C. with rotation. The following day, the membrane was washed twice; with a 2X low stringency wash solution (2X SSC and 0.1% SDS) at room temperature and with a 2X high stringency was solution (0.5X SSC and 0.1% SDS) at 68°C for 15 min. The membrane was then equilibrated for 1 min in maleic acid-T buffer (1X maleic acid buffer and 0.3% tween20) at room temperature and incubated in a 1X blocking buffer solution (1x blocking buffer in 1X maleic acid-T buffer) at room temperature for 20 min. The membrane was then incubated with 3 µL anti-DIG Ab Fab fragments coupled with alkaline phosphatase (Roche) diluted in 30 mL 1X blocking solution at room temperature for 30 min. The membrane was washed twice maleic acid-T buffer for 15 min and allowed to equilibrate in a 1X detection buffer (100 mM Tris-HCL [pH 9.5] and 100 mM NaCl) for 2 min at room temperature. The membrane was incubated with 1.5 mL chemiluminescent alkaline phosphatase substrate, CSPD, and sealed in acetate sheets at 37°C for 10 min. The light signal produced during the enzyme reaction was detected by 17 min exposure to X-ray film. The film was the washed in developer (Sigma Aldrich) and fixer (Sigma Aldrich) under red light.

2.7 Integration of DNA into the genome of *C. difficile*

2.7.1 Generation of CD05460, CD07730, CD07750, CD08670, CD19300, CD24700 and CD29630 mutants using ClosTron

Target genes were insertionally inactivated with the ClosTron system as previously described (Heap *et al.,* 2009). Target sites were identified and chosen by the CRG group and the intron retargeted to this site by DNA synthesis company, DNA 2.0. Re-targeted pMTL007C-E2 plasmids carrying this retargeted intron were sequence verified and delivered ready for transformation into *E. coli* donor strain and then into *C. difficile* by conjugation.

Seven genes encoding hypothetical proteins thought to be involved in the germination of *C. difficile* spores were specifically targeted for inactivation by insertional mutagenesis. These target genes included CD05460, CD07730, CD07750, CD08670, CD19300, CD24700 and CD29630, with their target sites for intron insertion between nucleotides 164 and 165 on the antisense strand of CD05460, 155 and 156 on the anti-sense strand of CD07730 and 167 and 168 on the anti-sense strand of CD07750, 285 and 286 on the sense strand of CD19300, and 234 and 235 on the sense strand of CD24700. ClosTrons of CD08670, CD29630 were previously generated in-house (CRG). During this study genes CD07730 and CD07750 were later revealed to be *spoVAC* (Shen *et al.,* 2016) and *spoVAE* (Sorg *et al.,* 2016).

The retargeted plasmids for each gene can be found in Table 2.6 and were separately transferred into *C. difficile* by conjugation from the donor *E. coli* CA434 strain as described previously. Transconjugants containing the ClosTron vector were selected for using thiamphenicol, and then serially sub-cultured to obtain pure clones. Insertion of the Group II:RAM element of the ClosTron into the chromosome was selected for by resistance to erythromycin. Genomic DNA was extracted from erythromycin- resistant clones for PCR confirmation of intron insertion to the target site using primer pairs flanking the target gene, the intron-target gene junction and the retrotransposition-activated marker (RAM) element of the ClosTron, as illustrated in Figure 2.2. The oligonucleotides used for construction of insertional mutants used in this study are listed in Table 2.8.

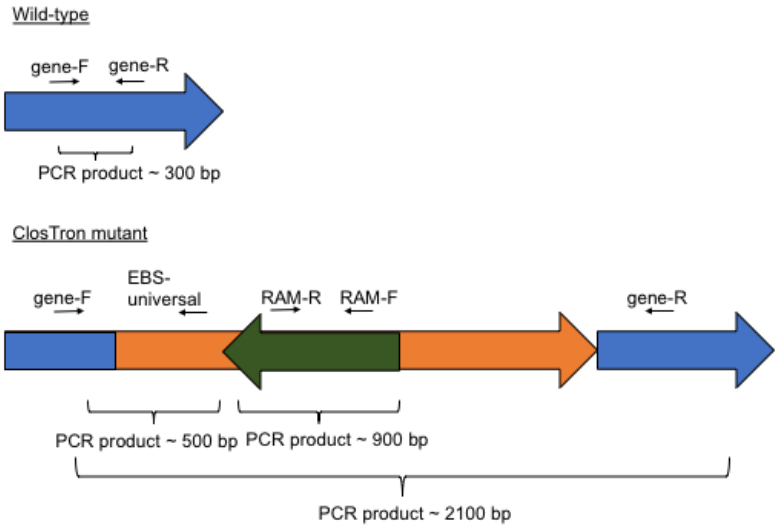


Figure 2.2: Screening for a ClosTron mutant.

A schematic representation of the targeted wild-type gene showing an appropriate position for the design of screening primers. The schematic representation of a ClosTron mutant showing annealing sites of the RAM-primers and possible PCR products that would be generated, including the exon–intron junction.

2.7.2 Isolation of single and double cross-over clones using Allelic exchange

The pMTL-ME2 vectors, carrying recombinant alleles of target genes, were used for in-frame deletion as described by Ng *et al., (*2013). These recombinant alleles were designed to have 1kb flanking homology regions upstream and downstream of the target and were created using splicing by overlap PCR (SOE-PCR) as described above, and amplified to carry flanking restriction endonuclease sites, *AscI* and *SbfI*. The resultant ~ 2 kb knockout cassette was purified by gel extraction, sequence verified and cloned into the pMTL-ME2-RN3 vector by restriction digest and ligation as described above. Plasmids generated carrying the knockout cassettes for the deletion of three *C. difficile* SASP genes, CD1290, CD2688 and CD3249. The knockout cassette for the fourth SASP gene, CD32490, was not generated due to unsuccessful attempts of SOE-PCR.

Complementation of the *pyrE* in-frame deletion was achieved by conjugating the pMTL-YN1 plasmid carrying the WT *pyrE* allele into the isolated double-crossover mutants. The pMTL-ME2 vectors also allow for large sequences of DNA can be integrated into the chromosome of *C. difficile* at the *pyrE* locus, for the coupled repair of the *pyrE* gene and target mutation. In order to complement the mutants, Wild type alleles together with their native promoter sequence were cloned into the pMTL-RN1 repair vector carrying the corrected version of *pyrE*. Native promoter sequences including were identified by using BPROM (SoftBerry). Ribosome Binding Site (RBS) and a 200 bp 5' non-coding region containing the putative promoter signals as identified by BPROM (SoftBerry) were cloned into pMTL-YN1C using SacI and BamHI restriction sites.

The resultant complementation plasmid, RN1c was then transformed into *E. coli* and subsequently conjugated into the mutant strain with a Δ*pyrE* background, concomitantly repairing *pyrE* and with the insertion of a functional copy of the gene inactivated by ClosTron deleted or allelic exchange, for complementation of the mutation at an appropriate gene dosage. Similarly, the overexpression of target genes originally knocked out can be achieved by re-introducing the corresponding gene into the pMTL-RN1x plasmid carrying the strong *fdx* promoter. Two plasmids, pMTL-RN1c::CD24700 and pMTL-RN1x::CD24700 were cloned for the complementation and overexpression of *C. difficile* 630Δ*erm*_CD24700::intron *ermB* strain, respectively. Conjugated strains were incubated for 96 h then restreaked onto non-supplemented CDMM to select for uracil prototrophy. Restoration of the *pyrE* phenotype was observed by the colony formation indicating successful allele exchange and confirmed by colony PCR using primers flanking the *pyrE* locus followed by Sanger sequencing of the amplified product.

2.8 Recombinant and native protein analysis

2.8.1 Inducing recombinant protein expression

For the expression of recombinant proteins from an *E. coli* host, the expression vector carrying the proteins coding sequence must be induced. Electrocompetent *E. coli* host, BL21(DE3), was transformed with the pET-28b(+) plasmid carrying the coding sequence of the GPR protein (CD24700), generated using the NEBuilder online tool at a recommended DNA molar ratio of vector to insert = 1:2. Briefly, CD2700 was amplified using gene specific primers and 50 ng of purified amplicon was mixed with 100 ng of pET-28b(+) plasmid and 10 µL NEBuilder HiFi DNA assembly master mix to a final reaction volume of 20 uL was incubated for 15 minutes at 50 °C. Once sanger sequence confirmed using T7 primer pairs (see Table 2.8), 2 µL of the assembled product was then electroporated into TOP10 and then transformed into BL21(DE3*) cells.

A single colony of BL21(DE3) carrying plasmid pET28b(+)::CD24700 was grown in 10 mL culture of LB supplemented with 50 µg/mL kanamycin LB broth with shaking overnight. This 10 mL starter culture was used to inoculate a pre-warmed flask of 250 mL LB broth supplemented with 50 µg/mL kanamycin and grown at 37 °C with shaking until the OD_{600} reached 0.6. Two 500 µL aliquots were taken as non-induced controls. pET vectors utilise the *lac* repressor system for transcription of sequences down stream of this promoter, leading to almost exclusive expression of the recombinant protein this sequence encodes when induced with isopropyl-ß-D-thiogalactoside (IPTG). Expression was induced by the addition of IPTG to a final concentration of 1mM and incubated overnight at 15 °C. Two additional 500µL aliquots were taken, pelleted and stored at -20 °C. All four aliquots were pelleted and stored at -20 °C. Cultures of BL21(DE3) host *E. coli* with and without empty pET28b(+) vectors were also grown and aliquots taken, pelleted and store at -20 °C, as negative expression controls. The remaining induced culture was harvested by centrifugation, pelleted and stored at -20 °C, ready for purification.

2.8.2 Purification of 6xHis-tagged proteins under native conditions

Use of pET expression vectors utilises the affinity of His-tagged recombinant proteins to Nickel based purification systems. The pET28b(+) vector used allowed for 6xHis- tagging of the recombinant GPR protein at both the N- and C-terminus, promoting the accessibility of these tags in the native tertiary structure. Prior to purification, pellets were treated using Bug Buster master mix (Novagen) according to manufacturer's instructions. Purification under native conditions was carried out using QIAexpress Ni-NTA Fast Start kit (Qiagen) according to manufacturer's instructions. Briefly, the cell pellet was resuspended in lysis buffer and treated with lysozyme on ice for 30 mins and sonicated on ice with six 10 s burst at 200 W with 10 s cooling periods in between. The lysate was centrifuged at 10, 000 x g for 30 min at 4 °C. 4 mL supernatant was incubated with 1 mL Ni-TA 50% slurry/lysis buffer at 4 °C for 60 min. The mixture was then loaded onto the Ni-NTA column which was then washed twice in 4 mL wash buffer and eluted twice in 0.5 mL elution buffer. The cell lysate, supernatant, flow-through, wash and eluate fractions were all analysed by SDS-PAGE analysis.

2.8.3 Separation of proteins by SDS-PAGE

All pelleted aliquots of induced, non-induced and negative recombinant protein expression controls were resuspended with 1X NuPAGE LDS Sample Buffer (4X) (Thermo Fischer Scientific) and normalised based on the OD_{600} at harvest for comparative loading on Novex NuPAGE 10% Bis-Tris Protein Gels (Thermo Fischer Scientific). Prior to loading, samples were heated to 70 °C for 10 min, pelleted and 13 µL supernatant loaded into the wells of the gel and run at 90 V for 120 min in 1X SDS running buffer (5% Novex NuPAGE MES Running Buffer (20X) in dH_2O, (Thermo Fischer Scientific) by electrophoresis using an XCell SureLock Mini-Cell (Thermo Fischer Scientific). For gels going onto western blotting, samples were diluted 1:10 with LDS before loading and a positive western blot control was also run. A Kaleidoscope Precision Plus Protein Standard (BioRad) was loaded as a marker.

2.8.4 Coomassie staining of PAGE gels

Proteins in a polyacrylamide gel were visualised by Coomassie staining by incubating in a Coomassie stain solution (0.1% Coomassie Brilliant Blue R-250, 50% Methanol and 10% glacial acetic acid) for 1 hr with gentle shaking, followed by the incubation in a destain solution (40% methanol and 10% glacial acetic acid) for 1 hr with gentle shaking. After destaining gels are left if H_2O overnight with gentle shaking, to reveal blue protein bands against a clear background.

2.8.5 Western Blot

Protein resolved by SDS-PAGE as mentioned above can be used for Western blot analysis. Briefly, SDS-PAGE gels are electronically transferred using the Trans-Blot Turbo Mini Nitrocellulose Transfer Pack (Bio-Rad) and run in the cassette of the Transfer System Mini Transfer System (BioRad) according to the pre-set Mini Gel Format run at 1.3 A constant; up to 25 V for 7 min. Following protein transfer to nitrocellulose membrane, protein-free sites that remain on the membrane most be blocked to minimise nonspecific signalling due binding of primary or secondary antibodies, or interactions with detection reagents. The western blot membrane was incubated in 30mL blocking solution (1.5% skimmed milk powder in Tris-Buffered Saline (TBS)) for 1 h, gently shaking at room temperature. The 6x His-tagged recombinant proteins can be detected using Anti-His antibodies. The membrane was then placed in 30 mL fresh blocking buffer containing 1/3000 mouse monoclonal antibody to 6X His tag horseradish peroxidase (HRP) conjugate (GeneTex) at 4 °C overnight with gentle shaking. The membrane was washed 6 times in TBS with 0.1% Tween-20 buffer at room temperature. Protein bands developed as dark blue when the 3, 3', 5, 5'-tetramethylbenzidine (TMB) liquid substrate (Sigma) is added, as an insoluble reaction product in the presence of HRP.

2.8.6 Isolation of acid-soluble proteins from *C. difficile* endospores

To extract acid-soluble spores from *C. difficile* spores a method adapted from Raju *et al.,* (2006) and Setlow (1974). A suspension of pure spores was prepared as described in sections 2.4.2 and 2.4.3 and repeated until 100 mg of dry weight spores was achieved. Dry spores were transferred into a pre-autoclaved microcentrifuge tube containing 0.1 g of 0.1 mm glass beads and metal beads (Precellys). Following the addition of 500 µL extraction buffer (3% acetic acid-30mM HCL), spores were ruptured in a Precellys 24 set at 6500 rpm 2x23 sec, with 1 min incubations on ice in between settings. The sample was transferred into an Eppendorf and incubated on ice for 30 mins with intermittent vortexing. Following centrifugation at 16,200 RCF for 10 min at 4 °C, the supernatant was collected into a fresh Eppendorf. The remaining pellet was suspended in 500 µL extraction buffer and centrifuged again. This extraction was repeated until a total of 2 mL supernatant is collected in the fresh Eppendorf tube, transferred into a 3 kDa molecular weight cut off (MWCO) dialysis membrane (Spectrum) and dialysed against 1L of 1% acetic acid at 4 °C for 24 h on a stirrer, with the dialysis buffer changed 3 times. After 24 h the contents are transferred into a flacon tube, frozen in liquid N_2 and lyophilised in a Thermo MicroModulyo Freeze Dryer (Thermo Scientific). The lyophilised acetic acid extract was dissolved in 1% acetic acid and stored at -80 °C.

2.8.7 Separation of native basic proteins by cathodic PAGE

Since SASPs are positively charged a separation method to produce high resolution for small molecular weight basic proteins was employed. Native basic proteins can be separated in low pH polyacrylamide gels, in which they migrate from anode to cathode during electrophoresis. This method was referred to as cathodic PAGE. Cationic stacking and resolving polyacrylamide gels were run in low pH buffers and under inverted electrode polarity. Lists of the composition of the Cathodic PAGE stacking and resolving gels, running buffers and loading dye used to separate and visualise acid-soluble proteins extracted from *C. difficile* spores can be found in Table 2.10 and Table 7.2 of the appendix for the composition of solutions A-F. 1 mg /mL lysosome was used as a positive control. Briefly, 10 µL of loading dye was added to 20 µL sample or control and the gel run at 20 mA in 1X running buffer for 45 min in a mini-Elpho

(Biophoretics) with electrodes reversed for current to run anode to cathode. Gels were Coomassie stained as described in section 2.8.4, except destained overnight.

Table 2.10: Composition of Cathodic PAGE gel, loading and running buffers

Gel composition	Solution [a]	Relative amount
Lower resolving gel	A	1
	C	2
	F	4
	H_2O	1
Upper stacking gel	B	1
	D	2
	E	4
	H_2O	1

Solution	Compound	Volume
Running buffer (diluted 1/8 before use)	*B*-alanine	31.2 g
	Glacial acetic acid	8.0 ml
	H_2O	to 1 L
Loading dye	Glycerol	1.0 mL
	Solution B	2.0 mL
	0.25% methyl green	0.25 ml
Gel Stain	Methanol	50 mL
	Glacial acetic acid	10 mL
	H_2O	50 mL
	Coomassie Brillant Blue R	0.275 g
Destain	Methanol	75 mL
	Glacial acetic acid	50 mL
	H2O	875 mL

[a] See Table 7.2 of the appendix for the composition of solutions A-F.

2.8.8 Silver staining of PAGE gels

Following staining of PAGE gels, a further silver staining method with higher sensitivity was used. A list of the composition of solutions used during silver staining are listed in Table 2.11. Coomassie stained PAGE gels were incubated in dH_2O overnight, with shaking. The is followed by incubation with 100mL of 5mg/L dithiotreitol (DTT) for 1 h and washed in 150 mL of dH_2O for 1 h. The gel was then stained in 100 mL silver stain for 30 min. The stained gel was then washed 3 times in dH_2O and developed in 150 mL developing solution for up to 15 min. As soon as bands were visible, the reaction was stopped with the addition of 20% methanol, 7.5% acetic acid in H_2O and finally placed in water.

Table 2.11: Composition of Silver staining solutions

Component	Compound	Volume
Silver strain	0.36% NaOH	21 mL
	880 ammonium solution	1.2 mL
	$AgNO_3$ solution	4 mL
	H_2O	To 100 mL
Developing solution	1% citric acid	2.5 mL
	Formaldehyde	200 µL
	H_2O	To 500 mL

$AgNO_3$ solution	$AgNO_3$	0.8 g
	H_2O	4 mL

2.8.9 Protein quantification by BCA

Total protein concentrations were quantified using absorbance at 562nm and compared to a calibration of albumin protein standards using the Pierce BCA Protein Assay Kit (ThermoFischer Scientific), according to manufacturer's instructions. Acid-soluble spore extracts were concentrated using a YM-3 Centricon Centrifugal Filter Device (Millipore) according to manufacturer's instructions.

2.8.10 Zymogen autoprocessing assay

The purified recombinant GPR protein exists as two forms, firstly, the inactive zymogen and then through an auto-process of self-cleavage into the active form. To assess this activity, an assay based on Carroll and Setlow (2005) and modified by Wetzel and Fischer (2015) was employed. Briefly, 3 µg of purified recombinant GPR were incubated with 100 µL of autoprocessing buffer (100 mM 2-[N-morpholino] ethane-sulfonic acid (pH 6.2), 5 mM $CaCl_2$, 10% glycerol (v/v) 50% DMSO (v/v)) or a negative control of 100 µL H_2O at 30 °C overnight. The reaction was stopped by the addition of 5 µL 1x NuPAGE LDS Sample Buffer (4X) (Thermo Fischer Scientific) and analysed by 10% gradient SDS-PAGE followed by Coomassie staining, western blot analysis and silver staining.

2.8.11 Assay to determine the protection of DNA from DNAse treatment

The protection of chromosomal *C. difficile* DNA from DNase degradation by SASPs in acid-soluble protein spore extracts was assessed based on assays developed by Raju *et al.,* (2006). Briefly, 1 µg chromosomal DNA was incubated with either 0 µg, 4 µg, 10 µg, 20 µg acid-soluble protein extract in 25 µL TAE buffer with and without 3 µL 10X Turbo DNase I buffer and 2 µL Turbo DNase for 2 h at 37 °C. Mixtures without Turbo DNase treatment acted as negative controls and that 1 µg chromosomal (chr) DNA and Turbo treatment as a positive control. Reactions were stopped with the addition of 100 µL SDS- solution (1.25% SDS, 25 mM EDTA, 13 µL 5M NaCl and 900 µL 96% ethanol) and immediately incubated at - 20 °C overnight to

allow precipitation of DNA. Samples were centrifuged at 13,000 RCF for 30 min at 4 °C and pellets washed in 70% ethanol, this process was repeated twice before pellet was air dried and dissolved in 10 µL H_2O. Samples were run in 1.5% agarose TAE alongside a 2 -Log DNA ladder (NEB) and visualised after gel electrophoresis in a Bio-Rad Gel Doc XR imaging system.

2.8.12 Assay to determine GPR proteolysis of SASPs

The ability of purified recombinant GPR to cleave small acid soluble proteins (SASPs) in the acid-soluble spore extracts from *C. difficile* spores was assessed by incubations at 37 °C in 100 µL of autoprocessing buffer. Briefly, mixtures of purified zymogen recombinant GPR with extracted spore proteins at a mass ratio of 1:5, 1:10 and YM-3 centrifugal filter concentrated acid-soluble spore extracts. Aliquots of 20 µL were removed after 12 h and 24 hr incubations and reactions stopped with the addition of 5 µL 1X NuPAGE LDS Sample Buffer (4X) (Thermo Fischer Scientific). Samples were loaded on 10% gradient SDS-PAGE followed by Coomassie staining, western blot analysis and silver staining.

Chapter 3

Chapter 3: Genomic Variation of *C. difficile* Isolates During Recurrent Infection

3.1 Introduction

Recurrence accounts for 5% -35% of CDI among patients (Bakken, 2009; Johnson, 2009) and it is not uncommon for those patients to have multiple relapses. Relapse in the clinical definition covers both recurrence by relapse with the same strain as that of the initial episode and/or reinfection with a different strain and may occur months or years after resolution of the initial infection (Johnson and Gerding, 2009). Since classifications include an external reinfection by the original strain, this presents a therapeutic dilemma as to whether patients should be treated with metronidazole or vancomycin for apparent treatment failures. Misclassifications of 'new cases' prove costly in terms of re-treatment, length of hospital stay and increases in surveillance reporting.

Whilst treatment of recurrent CDI remains limited, ongoing studies showing a reduction of recurrence include the use of antibiotics fidaxomicin and rifaximin, and monoclonal antibodies against toxins A and B**.** A pilot study of rifaximin as a chaser therapy to vancomycin treatment (n=79), showed a significant decrease in the incidence of recurrence compared to placebo (31% and 15% respectively, p=0.11) (Garey *et al., 2*011). Further studies in clinical settings are required to confirm its efficacy in the treatment of CDIs, as well as inferences of virulence factors correlating with recurrence.

A common limitation of clinical studies is the ability to differentiate relapse from re-infection due to the different typing methods used. One study demonstrated the ability to discriminate between PCR ribotype 027 isolates using whole genome sequencing (Stabler *et al., 2*009), with further research highlighting changes in genomes over short and long-time scales (He *et al., 2*010). Comparative whole genome analysis of multiple samples from the same patient could shed light on the genetic changes occurring over the course of infection and may have implications towards features of relapsing strains. Similarly, this would allow the genetic discrimination between that of reinfection strains, and if these can translate into phenotypic advantages during the course of infection.

As well as variation across genomes, observations of sequence variation within specific genes will also allow for observations for key genotypes that may be present amongst strains. Based on associations in literature for roles in *C. difficile* sporulation, germination, growth and antibiotic resistance, a subset of 24 genes were selected for further examination.

Clinical isolates presented in this study may harbour the same amino acid mutations in the DNA-directed RNA polymerase subunit beta (encoded by *rpoB*), known to lead to resistance to rifamycin. The gene is therefore of interest due to the trial of rifaximin (a rifamycin derivative) as a chaser therapy for recurrence presented in this study. An oligopeptide permease, also known as spo0K in spore formers, is encoded by five genes in the *opp* operon, (*oppB, oppA, oppC, oppD* and *oppF*) (Edwards *et al., 2*016), which have been shown to have a role in the initiation of sporulation *C. difficile* and was hence selected for further investigation. A major component of the *C. difficile* exosporium is the *Bacillus* collagen-like protein of *anthracis* (BclA) glycoprotein, with encoding genes *bclA1* and *bclA3* shown to have roles in spore persistence in mice studies of recurrent CDI (Phetcharaburanin *et al., 2*014), as such these genes were also selected to observe any changes in strains, particularly for those persistently carried in this study.

SpoVA proteins are known to be involved in the uptake of water in place of Ca-DPA, hydrating the spore during germination (Vepachedu and Setlow 2007) and so the entire operon, *spoVAE*, was selected for investigation, as it may pertain to differences in germination and outgrowth amongst strains. The roles of YabG, CspA, CspB and CspC along with SleC in *C. difficile* germination are well noted in the literature (Adams *et al.,* 2013; Kevorkian *et al.,* 2016; Kevorkian and Shen, 2017). Whilst GerS, a putative lipoprotein, has a hypothesised role in regulating SleC (Fimlaid *et al.,* 2015; Kochan *et al.,* 2018) and so any sequence variations in encoding genes of these proteins would help to link further phenotypic study of germination and outgrowth presented later in this work. The genes encoding the germination protease (GPR), putative sporulation protein YyaC and small acid soluble proteins (SASPs) were also selected due to their demonstrated role in spore resistance in *B. subtilis, C. perfringens* and *C. botulinum* (Loshon *et al., 1*999; Tennen *et al., 2*000; Setlow *et al., 2*000; Raju *et al.*2006; Leyva-Illades *et al., 2*007; Paredes-Sabja *et al., 2*008a; Meaney *et al.,* 2016; Wang *et al., 2*006), and their role in *C. difficile* germination and outgrowth presented later in this work.

In this chapter, results of PCR ribotyping, whole genome sequencing methods and sequence analysis of a subset of 24 genes are presented in an effort towards the elucidation of ribotype prevalence and the association of any sequence variants as significant virulence or risk factors of recurrence in CDAD.

3.2 Results

3.2.1 Bacterial isolates

A total of 178 in-patients were recruited between 2012 to 2014 from 23 hospital sites across the UK for enrolment on the trial to assess the use of rifaximin as a measure to control relapse of CDI. At the end of the enrolment period, 151 patients met the inclusion criteria. Of the enrolled patients, 77 (51%) were given a course of rifaximin and 74 (49%) were given the placebo (table 1.4.1.1). However, a total of 258 samples were submitted from only 92 (60.9%) patients, with 46 (50%) receiving a course of rifaximin and 46 (50%) given the placebo. These 92 patients and their submitted samples therefore form the total cohort investigated and discussed in this chapter. The positive isolation of *C. difficile* from these samples are displayed in Table 3.1.

Table 3.1: Summary of positive *Clostridioides difficile* isolation from stool samples.

	No. positive cultures/Total no. samples (%)		
	Placebo	Rifaximin[a]	Total
Treatment period[b]			
Week -1[c]	19/28 (67.8%)	14/24 (58.3%)	33/52 (63.4%)
Week 0	23/36 (63.8%)	21/34 (61.7%)	44/70 (62.8%)
Follow-up period[b]			
Week 4	20/28 (71.4%)	19/30 (63.3%)	39/58 (67.2%)
Week 12	19/25 (76.0%)	17/23 (73.9%)	26/48 (54.1%)
Relapse[d]	10/13 (76.9%)	9/13 (69.2%)	19/26 (73.0%)

[a] Rifaximin dosage: 1st 2 weeks- 2x 200 mg tablets 3 times daily (total = 1.2 g/day) 2nd 2 weeks- 1x 200 mg tablet 3 times daily (total = 0.6 g/day). [b] The first two weeks were the rifaximin or placebo treatment period and weeks 4 to 12 where the follow up period. [c] This sample is the trial enrolment sample taken on diagnosis of CDI, with standard therapy (metronidazole or vancomycin). [d] Relapse samples submitted as being 'clinically relapsed'. Four samples could not be designated into the treatment or follow-up period due to transcription discrepancies.

Of all 258 samples submitted only 173 (67%) were successfully cultured as *C. difficile* positive, in 48 (52%) patients. A total of 70 (76%) patients submitted 132 follow-up stool samples, collected more than one month after their initial specimen date, of these only 94 (71%) submitted from 47 patients were positive for *C. difficile* by culture. A total of 15 (16.3%) patients submitted 'clinically relapsed' samples, with six given a course of Rifaximin and nine given the placebo. Resulting in a recurrence rate of 6.5% for the rifaximin treated cohort and the 9.78% placebo treated cohort, respectively. A total of 26 (10%) 'clinically relapsed' samples were submitted, of this only 19 (73%) samples were successfully cultured as *C. difficile* positive. A drop of 8.1% and 2.1% in *C. difficile* culture-positive samples submitted at weeks 4 and week 12, respectively, where patients received Rifaximin over the placebo. There was an observed reduction in *C. difficile* culture-positive 'clinical relapse' samples by 7.7% for patients treated with Rifaximin compared to the placebo.

3.2.2 Ribotype prevalence of isolates

To assess the ribotype prevalence of patients and to assess those who were carrying multiple ribotypes, up to 20 single colonies per patient sample were stored for ribotype analysis, as described in the materials and methods chapter. Of the 173 samples successfully cultured as *C. difficile* positive, only 55 (31.79%) were successfully ribotyped from thirty-one patients, with eighteen different PCR ribotypes identified in this study. Ribotypes 002 and 014 were the most prevalent during our collection period across all enrolment sites, accounting for 21.8% and 20% of ribotype-able isolates, respectively (Figure 3.1).

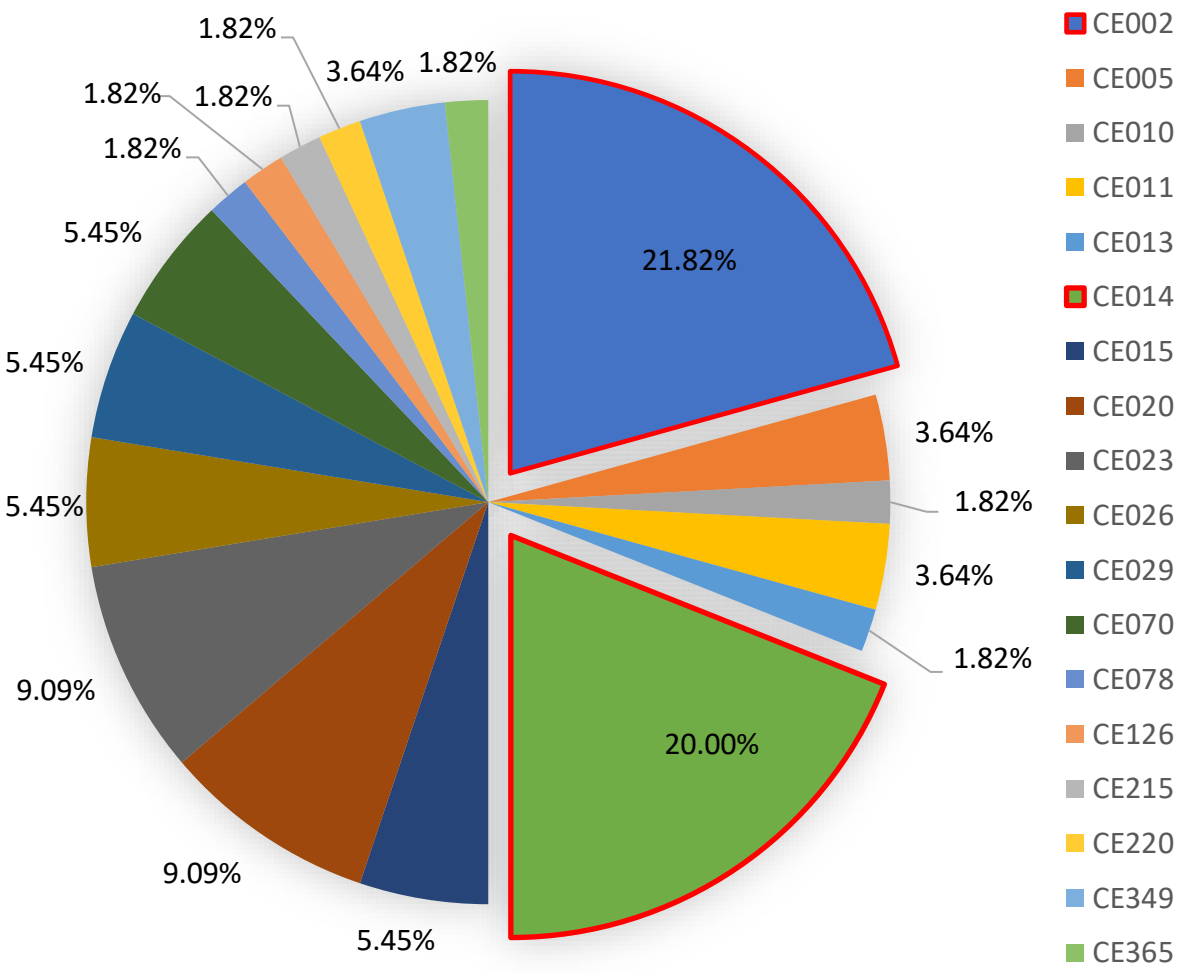


Figure 3.1: Prevalence of PCR ribotypes in study.

Every colony of *C. dificile* that was isolated from each stool sample was subjected to in-house PCR ribotyping. The most prevalent PCR ribotypes 002 (21.82%) and 014 (20%) during our collection period are bordered in red. The percentage frequency of ribotype-able *C. difficile* cultures and the distribution amongst PCR ribotype resolved by Capillary electrophoresis (CE).

A subset of reinfection events was made evident by different PCR ribotypes being present for the isolates from the same sample. By selecting up to 20 colonies per sample, we were able to screen for up to 20 isolates. These isolates may indeed be different strains with the same ribotype or conversely, different strains with different ribotypes. The latter was observed for five (16%) patients through the extensive colony screening method employed in this study. A total of twelve isolates were involved in reinfections observable by PCR ribotyping and belonged to types 002, 010, 014, 015, 026 and 029. The most common occurring ribotypes of this reinfection by ribotype observation were 002 and 014 occurring in seven out of twelve isolates (015 occurring twice and 010, 026 and 029 occurring once, respectively). A subset of 23 isolates from nine patients, now referred to as the recurrence cohort, were further characterised for phylogenic and genotypic diversity due to their recurrence status implied by PCR ribotyping and indication of clinical relapse (RL submitted sample).

A temporal graph illustrating the ribotype carriage over time for patients in this recurrent cohort are shown in Figure 3.2. One recurrence case was observed prior to trial enrolment (wk-1 sample) from patient 01008, and so this isolate, 01008 wk-1, also represents this patient's 1st relapse. There were 5 clinically indicated relapse events for patients 01008, 11007, 12001, 12004 and 18001 observed during the trial period, with patient 01008 submitting the 2nd relapse sample, 01008 RL, after week 12 of enrolment. 'Clinical relapse' episodes were observed up to three months after the trial therapy (wk 0 samples) started for four patients (11007, 12001, 12004 and 18001). With patient 01008 having a 2nd clinical relapse 6 months after the diagnostic sample was submitted (RL sample). Reinfection and coinfection cases were identified by the detection of different PCR ribotypes in the same sample from patients 01008 (coinfection and reinfection), 01013 (reinfection), 01018 (coinfection) and 11007 (coinfection and reinfection). Reinfection events were observed between 3 and 6 months after the presumed original isolate responsible for symptomatic disease was detected.

Whole genomic DNA was extracted from 20 isolates of the recurrence cohort; 01008 wk-1 (E2), 01008 wk-1 (E10), 01008 wk12 (G2), 01008 RL (H4), 01013 WK-1 (C17), 01013 WK12, 01016 WK-1, 01016 WK12, 01018 WK-1 (C17), 01018 WK-1 (C20), 11007 WK-1 (C2), 11007 WK-1 (C5), 11007 WK0, 11007 RL, 1200 1 WK 0, 12001 RL, 12004 WK0, 12004 RL, 18001 WK0 and 18001 RL. Isolates from patient 11001 were omitted from further investigation as there was no sample available that represented diagnostic status at the start of the trial or during the treatment period. Whole genome sequencing and variant analysis was preformed on these isolates to investigate phylogeny and sequence similarity.

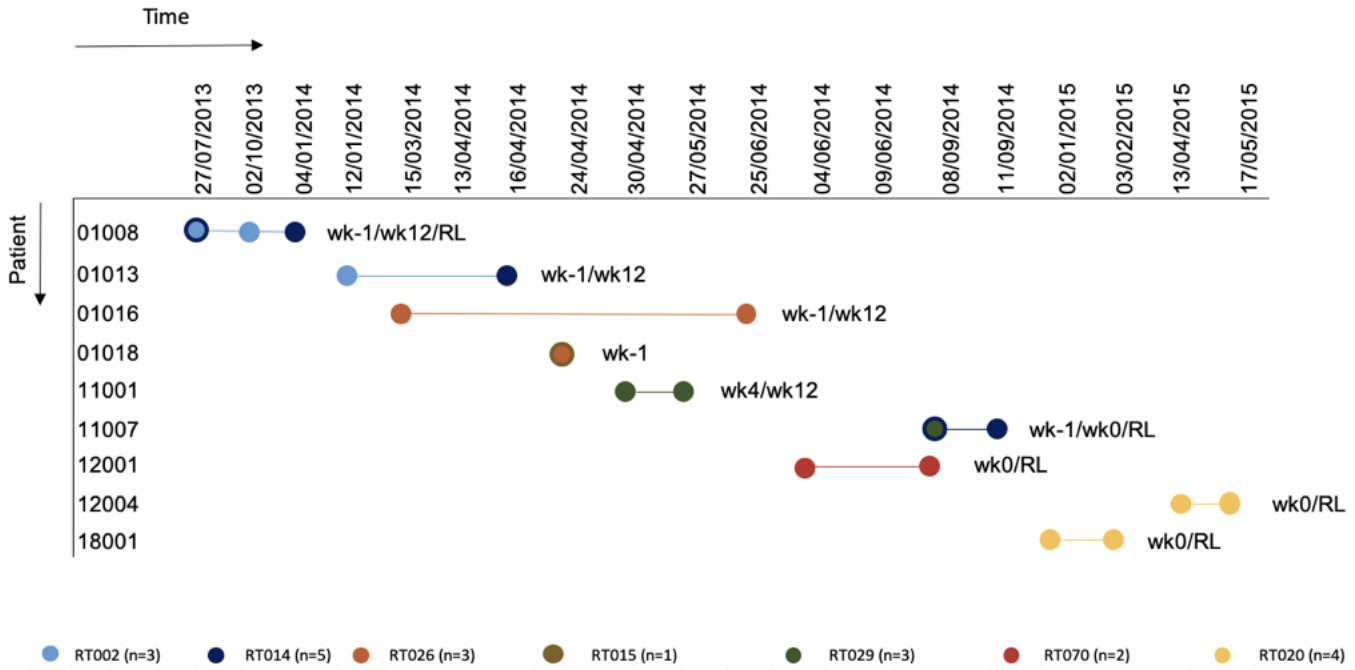


Figure 3.2: Temporal graph of 23 recurrent *C. difficile* infection causing isolates from nine patients.

Each row depicts two to three samples of a patient (Y axis) at the time points at which they were submitted (X axis). Isolates selected for further characterisation obtained from these samples are represented by circles, each coloured according to the PCR ribotype(s) (RT) present; · RT002 (n=2), · RT014 (n=5), · RT026 (n=3), · RT015 (n=1), · RT029 (n=3), · RT070 (n=2), · RT020. Circles with alternate colours bordering represent the additional PCR ribotype isolated from that sample. Sample identifiers are labelled next to each isolate. Wk = Trial week of sample submssion, RL = Relapse sample submitted.

3.2.3 Genomic diversity of isolates

Genomic variation of strains between ribotypes has been previously highlighted in literature as well as strain diversity noted within ribotypes (Kurka *et al., 2*014). This study aimed to illustrate variants between isolates over time to properly asses their infection status as reinfection or recurrence. Variant calling using Qiagen Genomic Gx CLC suite (CLC) pipeline between isolates of the same patient within the 'recurrence cohort' are presented in a summary with their infective status for the host in Table 3.2. Data outputs of specific variants for isolates can be found in Table 7.4 of the appendix.

Out of the 12 patients presented in this chapter as the recurrence cohort, five were prescribed Rifaximin, whilst of three patients were prescribed placebo, and a total of 20 positive stool cultures were isolated. A total of 19 isolates were sequenced, 17 producing full genomes (see Table 3.2). The average genome was 4.21 M bp, with strain wk-1 (E2) from patient 01008 having the largest genome (4.65 M bp) and the smallest genome (4.13 M bp) belonging to strain wk-1 (C17) from patient 01018.

Clinical recurrence (referred to by healthcare workers as 'relapse') was described for 5 patients; 01008, 11007, 12001, 12004 and 18001, displaying symptomatic infection within 3 months of treatment. With a clinical recurrence rate of 25% (n=3) and 16.7 % (n=2) was observed for those patients administered rifaximin and placebo, respectively (see Table 3.2). Of these cases, only three relapse causing strains were confirmed by SNP analysis, where <100 SNPs were found between the initial infecting strain and the relapse isolate in patients 01008, 12001 and 12004. Whilst strains isolated from 'relapse' samples of patients 11007 and 18001 were found to have >100 SNPs between them and the initial infecting strain, indicating that they were genetically distinct strains and so reclassified as reinfection cases. Further inspection revealed that a total of two relapse isolates were recovered from patients on Rifaximin compared to two on placebo, representing an equal relapse rate of 50% for both treatments.

Table 3.2: Summary of patient and genomic data of *C. difficile* isolates selected for further variant analysis in this study.

Patient ID	Pre-Trial Treatment[a]	Trial therapy	Strain ID [b]	PCR RT [c]	Size of genome, bp	SNPs[d]	Microbiological infection status [e]
01008	Met	Rifaximin	wk-1 (E2)	(CE) 002	Contig 1 67,300		Co-infection
			wk-1 (E10)	(CE)014	Contig 1 4,465,937	4	1st Relapse (Co-infection)
			wk12 (G2)	(CE) 002	Contig 1 4,207,893	6	Persistent Carriage
			RL (H4)	(CE)014	Contig 1 4,207,498	1	2nd Relapse
01013	U	Rifaximin	WK-1 (C17)	(CE)002	4,161,787	-	Symptomatic carriage
			WK12	(CE)014	4,161,784	-	Asymptomatic Reinfection
01016	U	Rifaximin	WK-1	(CE)026	4,172,179	-	Symptomatic carriage
			WK12	(CE)026	4,172,179	11	Persistent Carriage
01018	U	Placebo	WK-1 (C17)	(CE)026	4,130,634	-	Co-infection
			WK-1 (C20)	(CE)015	N/A	-	Co-infection
11007	Met	Placebo	WK-1 (C2)	(CE)014	4,275,823	-	Co-infection
			WK-1 (C5)	(CE)029	Contig 1 1,24,061	-	Co-infection
			WK0	(CE)014	4,275,823	>100	Reinfection
			RL	(CE)014	4,275,823	>100	Relapse
12001	Met	Rifaximin	WK 0	(CE)070	4,187,972	-	Carriage
			RL	(CE)070	4,187,972	6	Relapse
12004	Van	Placebo	WK0	(CE)020	4,179,842	-	Carriage
			RL	(CE)020	4,179,842	9	Relapse
18001	Van	Rifaximin	WK0	(CE)020	4,179,842	-	Carriage
			RL	(CE)020	4,179,842	>100	Reinfection

[a] Met: metronidazole, Van = vancomycin, U = unknown, [b] Wk = week, RL= relapse, [c] PCR RT = PCR ribotype visualised by capillary electrophoresis (CE), [d] SNPs = Number of single nucleotide polymorphisms using the earliest isolate of the same PCR ribotype as the reference genome, [e] Co-infection = symptomatic carriage of an additional strain isolated from the same diagnostic sample, Reinfection = carriage of a different strain to that isolated from prior samples, Relapse = symptomatic carriage of the same strain isolated from prior samples, Carriage = carriage of the same strain as that isolated from prior samples during the course of treatment, Persistent carriage = asymptomatic carriage of the initial infecting strain beyond treatment.

Table 3.2 also presents the incidence of reinfection and persistent carriage in this recurrence cohort. One asymptomatic reinfection event was observed in patient 01013 at week 12, this isolate was found to have a different PCR ribotype, CE 014, indicating a new strain and was detected within 3 months of treatment. Additional reinfection events were observed for patient 11007 at week 0 and patient 18001 (Relapse sample), in both cases, the same PCR ribotype was detected, CE 014 and 020, respectively as the initial infecting isolate. However, over 100 SNPs were identified indicating a new strain had been acquired, for patient 11007 at the start of the placebo trial treatment period and for patient 18001, within the post Rifaximin chaser therapy period. For patient 11007 this same reinfection strain was later isolated from the relapse sample (11007 RL), again having >100 SNPs from the original infecting isolate, indicating that this strain was carried asymptomatically through the placebo chaser therapy period and later led to symptomatic disease. Interestingly, SNP analysis indicates that 3 genetically distinct strains were isolated from patient 11007, two during a co-infection and one as a reinfection that later led to relapse. Persistent carriage was observed in two patients, 01008 and 01016, where asymptomatic carriage of the same initial infecting strain beyond treatment was detected 3 months later (week 12 samples).

Whole genome comparisons allowed the identification of variants across strains within the same patient to identify recurrence status. More information can be drawn as to whether strains have any sequence similarity between all strains isolated in this recurrence cohort to shed light on possible events of transmission between patients. A phylogenetic tree using *Clostridium perfringens* as the outgroup to anchor the tree is presented in Figure 3.3. The strain 01013 wk-1 (C1) was not included in this analysis, as downloadable contigs were corrupt, whilst strain 01008 wk-1 (E2) was excluded from further analysis as the contig is too short to reliably contain enough core genes for tree building. Strains included in this investigation include 01008 wk-1 (E1), 01008 wk12 (G2), 01008 RL (H4), 01013 wk-1 (C17), 01016 wk12, 01016 wk-1, 01018 wk-1 (C17), 11007 RL, 11007 wk0, 11007_wk-1 (C2), 12001 RL, 12001 wk0, 12004 RL, 12004 wk0, 18001 RL and 18001 wk0.

Strains appear to be clustering by patient, suggesting that there was no transmission between patients, however, as genomes are so similar, this is not conclusive. The bootstrap values indicate the degree of confidence in each branching point, with 100 being the max. Whilst a higher bootstrap value means stronger support for that branch structure, in this study as all genomes are so closely related, lower values on the tree most likely represent the difficulty the program had deciding which groups of strains were more closely related. Looking at the group of 01016 and 01018 strains, the program is confident that in the branching structure. But towards the top of the tree, there are much lower bootstrap values, indicating lower confidence. In this case, the 01008 strains are split between two different clades, but the bootstrap value of those clades is only 25, which is very low. This means that the program did not have the highest confidence that those two groups should be considered different. Likely, the genomes are all so closely related that the program can't tell the difference between them, this is supported by the low-level SNP difference between strains isolated from this patient that were included in this tree.

Another approach to assess relatedness between strains, includes the average nucleotide identity (ANI) calculated using ANIcalculator (version 1) presented in Figure 3.4. All analysed genomes were > 98% similar, with most > 99% similar, indicating that they are very closely related (approximately strain level). As expected, the results of the ANI analysis indicate that genomes are closely related. They are nearly all >99% similar, indicating that they are very likely the same species, if not the same strain. The heatmap of ANI also indicates that 01008 wk 12 (H4) clusters separately from 01008 wk 12 (G2) and 01008 wk-1 (E10), supporting the branch structure for these strains. The genomes of the same PCR ribotypes are shown to have the highest relatedness, with the highest ANI observed between strains. This analysis also confirms that the branching in the phylogenetic tree is picking up small differences between genomes. No strain had an ANI of 100 against any other indicating that no transmission occurred within this recurrence cohort.

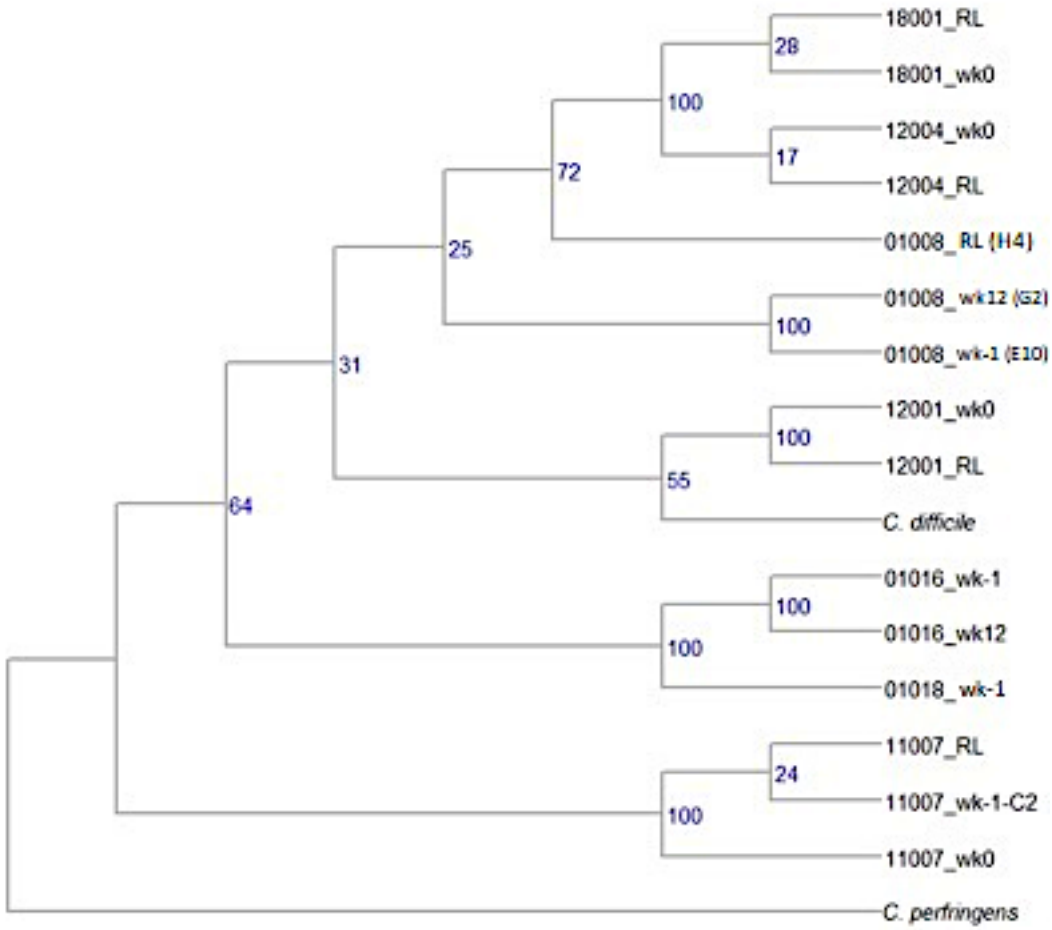


Figure 3.3: Phylogenetic tree of 15 *C. difficile* strains within the recurrent cohort.

A phylogenetic tree illustrating the relatedness between the full genomes obtained for strains of *C. difficile* within the recurrence cohort, constructed based on alignments of single-copy core bacterial genes (Ankenbrand & Keller, 2016). The genome of *C. difficile* 630 (GenBank Assembly ASM920v2) was included in this tree as a reference, and the genome of *C. perfringens* ATCC 13124 (GenBank Assembly ASM1328v1) was included as an outgroup (Sebaihia et al., 2006; Myers et al., 2006). Bootstrap values are reported at each branch node in blue.

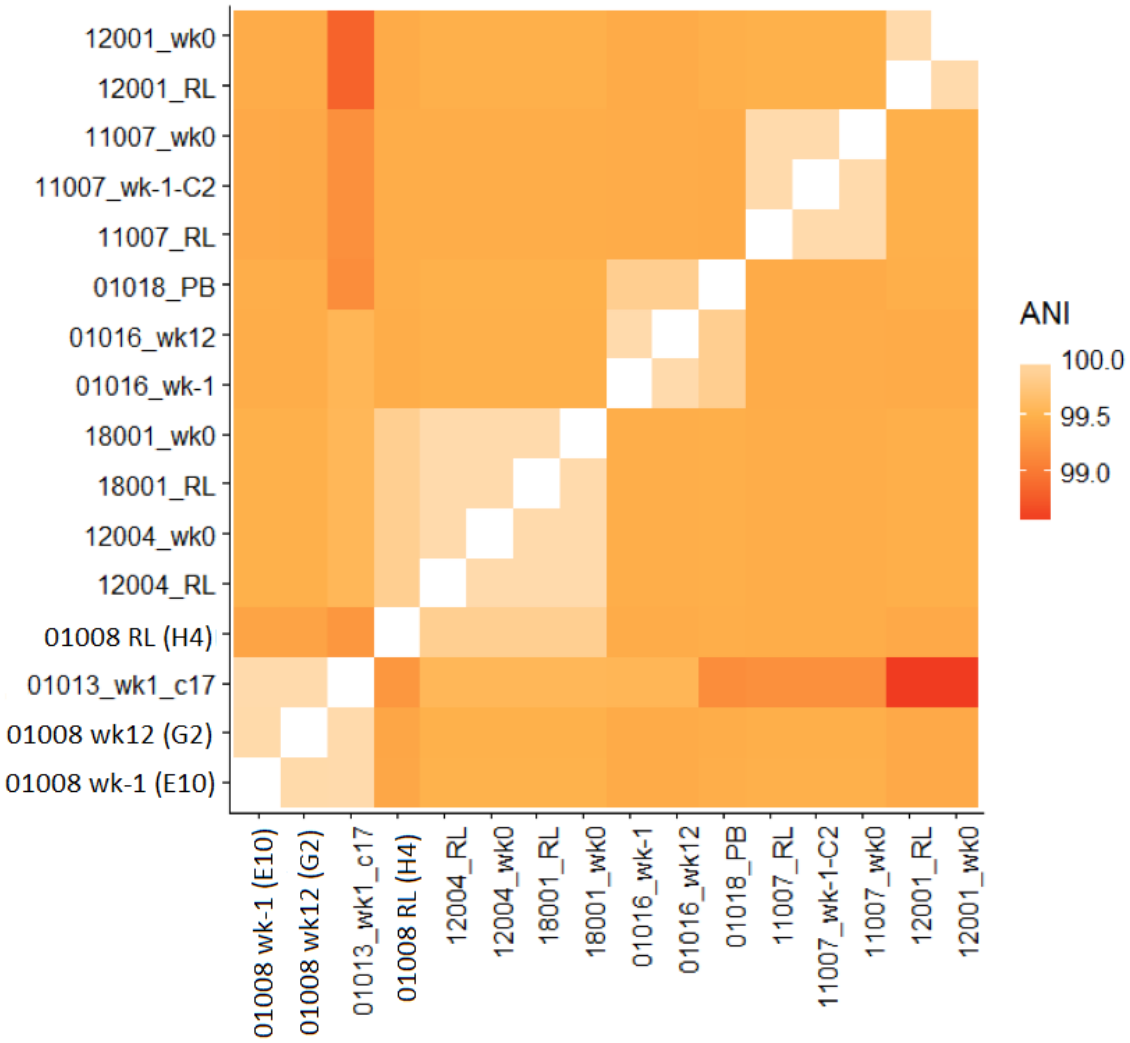


Figure 3.1: A heatmap of average nucleotide identity (ANI) between genomes of the recurrent cohort.

Heatmap to illustrate the relatedness between strains, using the average nucleotide identity (ANI). Representing the matching percent identity of the aligned regions of each genome pairs. The max ANI possible is 100%. Lighter colours represent higher ANI, so more closely related.

3.2.4 Sequence variation of genes of *C. difficile* associated with spore germination and outgrowth

One research question in this study was how sequences in key spore phenotype and growth genes differ between genomes. Using CD630 (GenBank Assembly ASM920v2) as the reference strain, to which most gene phenotype studies are based in the literature, nucleotide and amino acid sequence variations in the genomes of the recurrent cohort were further investigated. Strains included in this investigation include 01008 wk -1 (E1), 01008 wk12 (G2), 01008 RL (H4), 01013 wk-1 (C17), 01016 wk12, 01016 wk-1, 01018 wk-1 (C17), 11007 RL, 11007 wk0, 11007_wk-1 (C2), 12001 RL, 12001 wk0, 12004 RL, 12004 wk0, 18001 RL and 18001 wk0.

From a list of 24 genes of interest, chosen for their association to *C. difficile* sporulation, germination, growth and antibiotic resistance (see Table 3.3), nucleotide fasta sequences were downloaded from Genbank using their CD630 gene locus tags as the search term. A table of percentage identity for the best hit in each genome can be found in Table 7.8 of the appendix. A list of all amino acid changes found is provided in appendix Table 7.10. The frequency of amino acid changes in genes of interest and the distribution of these changes amongst strains can be found in Figure 3.5 and Figure 3.6, respectively. UGENE visualised alignments for selected genes presented in Figure 3.7 - Figure 3.12.

Table 3.3: Characteristics of 24 genes of reference *C. difficile* strain CD630 (GenBank Assembly ASM920v2) selected for their association with sporulation, germination, growth and antibiotic resistance used to compare the sequence variation of isolates

Gene Name	**Gene Locus**	**Length**	**Phenotypic Association** [a]
DNA-directed RNA polymerase subunit beta (r*poB* [1])	CD630_00660	1238	Antibiotic resistance
putative sporulation protein yyac (*yyaC* [2])	CD630_01290	175	Germination
putative exosporium glycoprotein (*bclA*1[3])	CD630_03320	2082	Colonization
Spore cortex-lytic enzyme pre-pro-form (*sleC*[4])	CD630_05510	423	Germination
Stage V sporulation protein AC (*spoVAC*[5])	CD630_07730	145	Germination
Stage V sporulation protein AD (*spoVAD*[5])	CD630_07740	337	Germination
Stage V sporulation protein AE (*spoVAE*[5])	CD630_07750	120	Germination
ABC-type transport system, oligopeptide-family permease (*opp*B[6])	CD630_08530	328	Sporulation
ABC-type transport system, oligopeptide-family permease (*opp*C[6])	CD630_08540	303	Sporulation
ABC-type transport system, oligopeptide-family extracellular solute-binding protein (*oppA*[6])	CD630_08550	522	Sporulation
ABC-type transport system, ATP-binding component (*oppD*[6])	CD630_08560	337	Sporulation
Fragment of ABC-type transport system, oligopeptide-family ATP-binding protein (*oppF*[6])	CD630_08570	314	Sporulation
putative small acid-soluble spore protein (putative *ssp* [2])	CD630_12900	66	Germination
Subtilisin-like serine germination related protease (*cspC*[4])	CD630_22460	557	Germination
Subtilisin-like serine germination related protease (*cspBA*[4])	CD630_22470	1116	Germination
Glycine reductase complex selenoprotein A (*grdA*[7])	CD630_23520	107	Cell growth
Germination protease (*gpr* [2])	CD630_24700	317	Germination
Small, acid-soluble spore protein alpha (*sspA*[2])	CD630_26880	73	Germination
Small acid-soluble spore protein (*ssp*[2])	CD630_32201	81	Germination
Proline reductase (*prdB*[7])	CD630_32410	150	Cell growth
Small, acid-soluble spore protein beta (*sspB*[2])	CD630_32490	67	Germination
Exosporium glycoprotein BclA3 (*bclA3*[3])	CD630_33490	656	Colonization
putative lipoprotein (*gerS*[8])	CD630_34640	196	Germination
Sporulation-specific protease (*yabG*[4])	CD360_35690	286	Germination

[a]Phenotpic association taken from literature: [1] *rpoB* taken from O'Connor *et al., (*2008), [2]*yyaC, gpr, ssp, sspA* and *sspB* taken from Wetzel and Fischer (2015), [3] *bclA1* and *bclA3* taken from Phetcharaburanin *et al., (*2014), [4]*sleC, cspC* and *cspBA* and *yabG* taken from Kevorkian *et al., (*2016), [5] *spoVA* operon taken from Vepachedu *et al., (*2006), [6] *opp* operon taken from Edwards *et al., (*2014), [7]*grdA* and *prdB* taken from Bouillaut *et al., (*2012), [8]*gerS* taken from Fimlai *et al., (*2015) and Diaz *et al., (*2018).

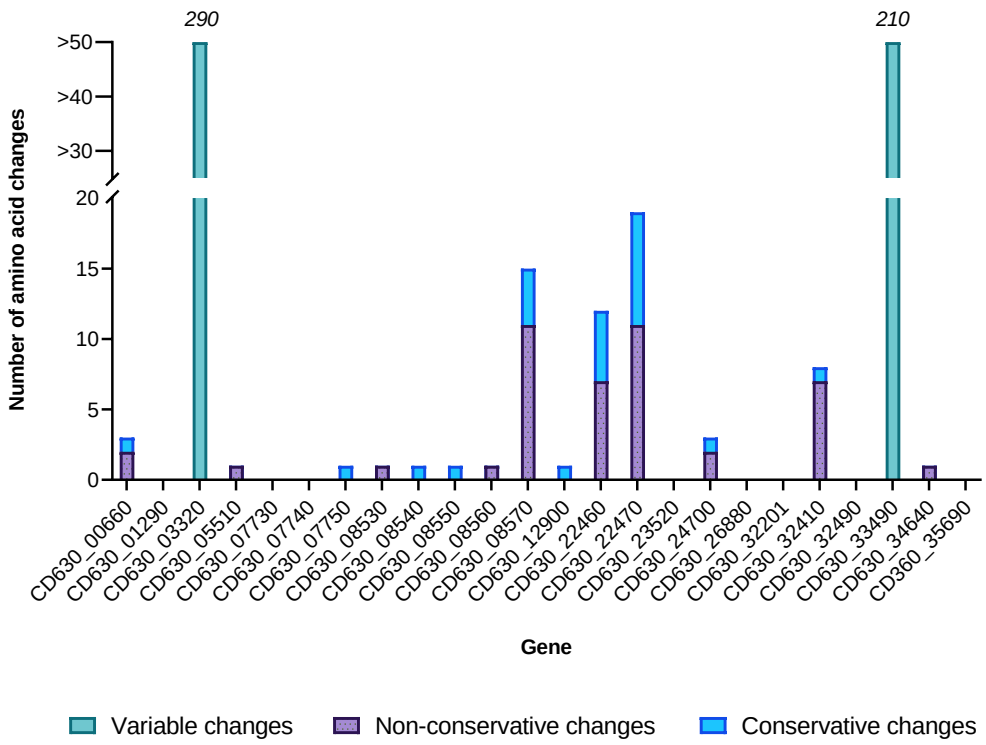


Figure 3.5: The frequency of amino acid changes found in the strains of the recurrence cohort for 24 genes selected for their associated with growth, and spore germination in reference *C. difficile* strain CD630.

For genes CD630_03320 and CD630_33490 variable regions with an excess of 290 and 210 amino acid changes were obsevered, respectively. As such these changes are epresented in the key as variable changes.

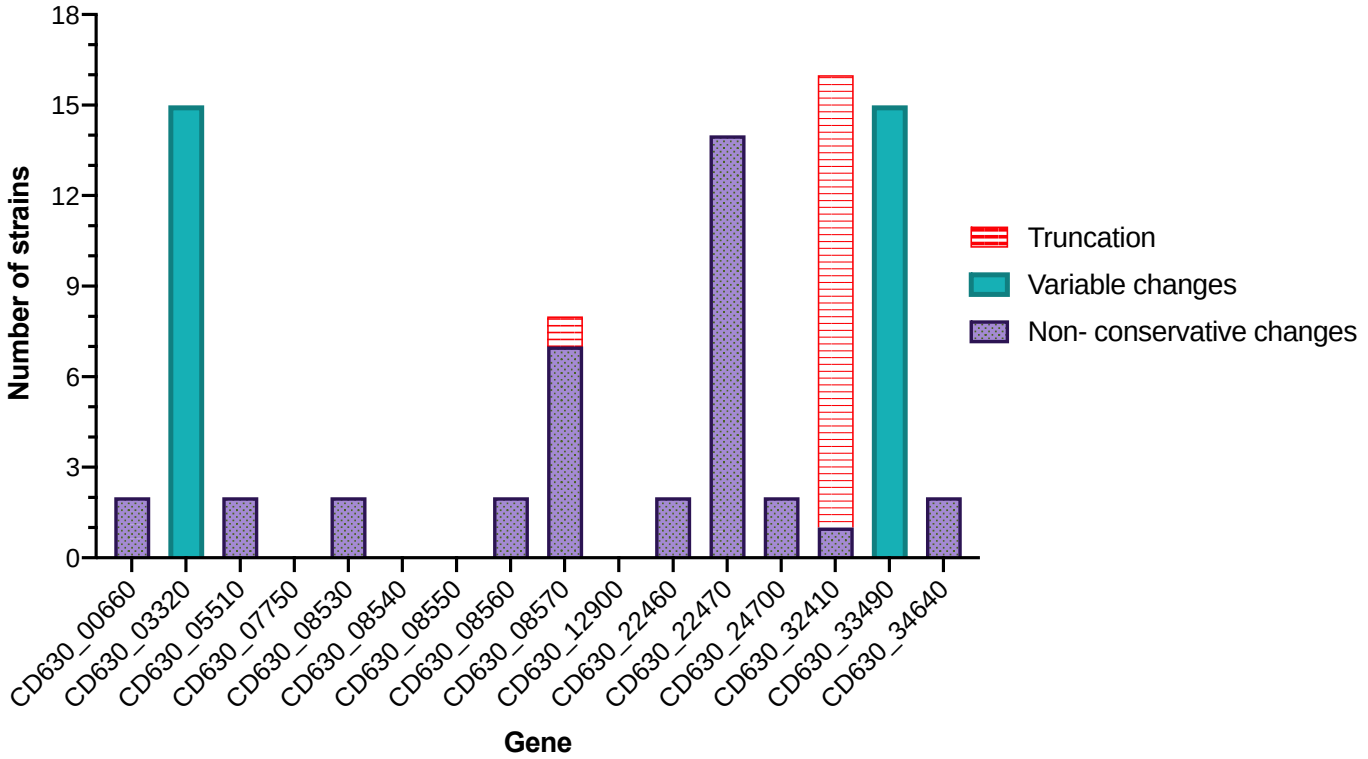


Figure 3.6: The distribution of amino acid changes found amongst 15 strains in the recurrence cohort for 16 genes of interest of reference *C. difficile* strain CD630

Several genes were mis-aligned to the reference at the start of the gene in all isolates during this alignment. As all the isolates matched each other and because this region was at the very start of the alignment, it likely reflects disagreement between the reference and the alignments as to where the gene starts, rather than an actual change in sequence. Genes CD630_07730 (*spoVAC*), CD630_07740 (*spoVAD*), CD630_23520 (*grdA*), CD630_26880 (*sspA*), CD630_32201 (*ssp*), CD630_32490 (*sspB*), and CD360_35690 (*yabG*), showed 100% nucleotide identity to all strains, with mismatches ranging from 99-80% for the remaining genes of interest across strains (see Table 7.8 of the appendix). Downloadable contigs for strain 01013 wk-1 (C1) were later found to be corrupt, with alignments showing less than 55% nucleotide identity to all genes of interest and 0% nucleotide identity to gene CD630_33490 (*bclA3*). Assessment of the exported genome revealed that it was too short for meaningful analysis and so not included in downstream genomic analysis. No missense mutations were found as a result of sequence mismatches for any isolate in gene CD630_01290 (*yyaC*).

Missense mutations found in 16 genes, led to a total of 70 non-synonymous amino acid changes, including single amino acid insertions, whilst more than 50 total amino acid changes, including gaps, were observed in two genes containing variable repeat regions (see Figure 3.5, and Table 7.9 and Table 7.10 of the appendix for detailed amino acid changes). Of these changes, 66% consisted of non-conservative amino acid changes, by which an amino acid of differing physiochemical properties is substituted (n=46, excluding variable repeat regions) in genes CD630_00660 (*rpoB*), CD630_03320 (*bclA1*), CD630_05510 (*sleC*), CD630_07750 (*spoVAE*), CD630_08530 (*oppB*), CD630_08540 (*oppC*), CD630_08550 (*oppA*), CD630_08560 (*oppD*), CD630_08570 (*oppF*), CD630_12900 (*ssp*), CD630_22460 (*cspC*), CD630_22470 (*cspBA*), CD630_24700 (*gpr*), CD630_32410 (*prdB*), CD630_33490 (*bclA3*) and CD630_34640 (*gerS*) (see Figure 3.6, and Table 7.9 and Table 7.10 of the appendix for detailed amino acid changes). The largest level of variation, with more than 5 amino acid changes across isolated strains was observed in genes CD630_03320 (*bcla1*), CD630_08570 (*oppF*), CD630_22470 (*cspBA*), CD630_32410 (*prdB*) and CD630_33490 (*blcA3*) (see Table 7.10 of the appendix for

detailed amino acid changes). The remaining genes had small amino acid changes, usually between 1-3 changed amino acids in only a few strains.

MUSCLE alignments of amino acid sequences for CD630_03320 (*bcla1*), show variable repeats and some changes in conserved regions in all isolates (see Figure 3.7), this gene consists of two conserved regions with a variable region of glycine repeats in the middle. All strains differed from the reference sequence in this region. Alignments for amino acid sequences for gene CD630_33490 (*blcA3*) showed a similar pattern, with all strains showing differences in the variable region, additionally all strains had sequences longer than reference gene and some variants were also observed in 2nd conserved region (see Figure 3.11) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE. A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red box highlights the variable region of this gene across sequences.). Additionally, all strains showed some variation (1-6 amino acid changes) in the amino acid sequence of gene CD630_08570 (*oppF*) (see Table 7.10 of the appendix for detailed amino acid changes). All strains except 01018 wk-1 (C17) showed variation in amino acid sequences of gene CD630_22470 (*cspBA*). MUSCLE alignments also indicated some truncated genes; CD630_32410 (*prdB*) was truncated in all strains, and genes CD630_08570 (*oppF*) and CD630_03320 (*bclA1*) were truncated in strains 01008 wk-1 (E10) and 01008 RL (H4) showed respectively (see Figure 3.10, Figure 3.8 and Figure 3.7, respectively).

A total of eight non-conservative amino acid changes were found across co-infecting strains in all genes. Whilst a total of 11 non-conservative amino acid changes were found in across relapse strains in all genes. No amino acid changes in the genes of interest were exclusively found in coinfecting, persistently carried or relapsing strain groups. However, isolates 01008 wk-1 (E10) and 01008 wk 12 (G2) and 01016 wk-1 and 01016 wk 12, carried in patients who had persistent CDI, had the highest number of amino acid changes across the selected genes of interest, sharing 46 of the total 70 non-synonymous changes found across all strains, whilst

also containing 24 exclusive amino acid changes. Interestingly, all non-synonymous amino acid changes observed for the gene CD630_22460 (*cspC*), were only found in persistently carried strains. In all a total of 37 non-conservative amino acid changes were found in this group. All non-conservative amino acid changes observed for genes CD630_34640 (*gerS*), CD630_32410 (*prdB*), CD630_00660 (*rpoB*), CD630_05510 (*sleC*) and CD630_22460 (*cspC*) were exclusive to persistently carried strains (see Table 3.4).

Table 3.4: Amino acid changes for genes of interest from reference *C. difficile* strain CD630 exclusively observed in isolates 01008 wk-1 (E10) and 01008 wk 12 (G2) and 01016 wk-1 and 01016 wk 12.

Gene Locus	Gene name	Associated Phenotype	Strain	Amino acid change
CD630_00660	*rpoB*	Rifamycin antibiotic resistance	01008 wk-1 (E10) 01008 wk 12 (G2)	489 Q-> K 492 D-> V
CD630_05510	*sleC*	Spore germination	01016 wk-1 01016 wk 12	240 A->S
CD630_22460	*cspC*	Spore germination	01008 wk-1 (E10) 01008 wk 12 (G2)	72 N-> D 415 D-> E
			01016 wk -1 01016 wk 12	39 L-> I 157 N-> S 158 R-> T 162 S-> G 184 I-> V 267 A-> T 331 V-> A 374 R-> G 394 N-> G 530 E-> D
CD630_32410	*prdB*[1]	Vegetative cell growth	01008 wk 12 (G2)	123 H-> Q 124 V-> K 126 T-> K 132 A-> E 135 A-> R 136 K-> I 146 V-> L 147 M-> L
CD630_34640	*gerS*	Germinant receptor homologue	01008 wk-1 (E10) 01008 wk 12 (G2)	94 G-> D

[1]truncation in all strains, amino acid changes found in strain 01008 wk 12 (G2) only

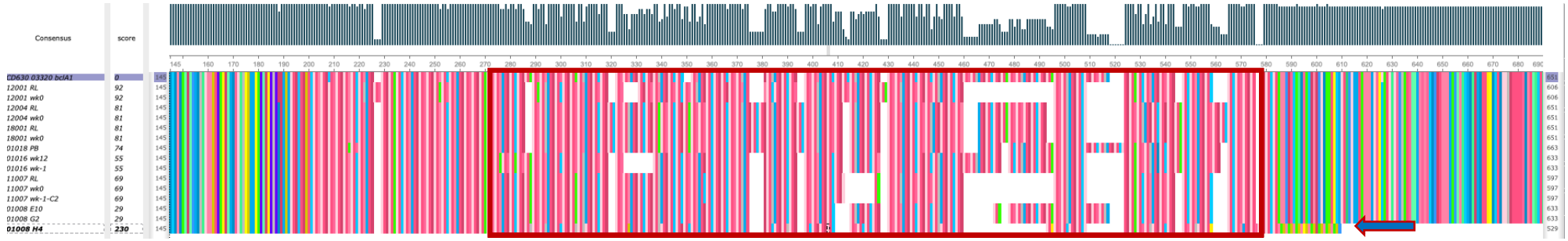


Figure 3.7: MUSCLE alignment of CD630_03320 (*blcA1*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red box to illustrates the variable region in between the two conserved regions of this gene and blue arrow indicates truncated sequence observed amongst strains.

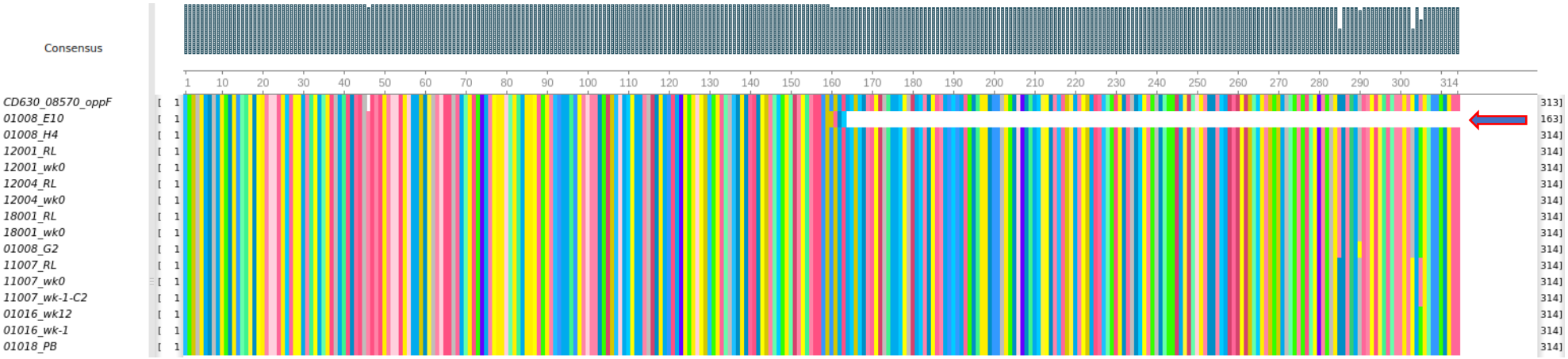


Figure 3.8: MUSCLE alignment of CD630_08570 (*oppF*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The blue arrow highlights the truncated sequence observed amongst strains.

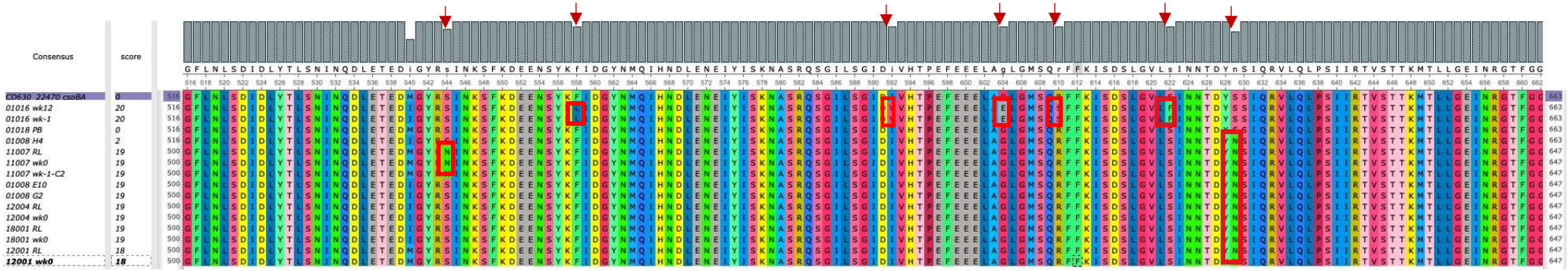


Figure 3.9: MUSCLE alignment of CD630_22470 (*cspBA*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red boxes and red arrows highlight example amino acid changes observed across sequences.

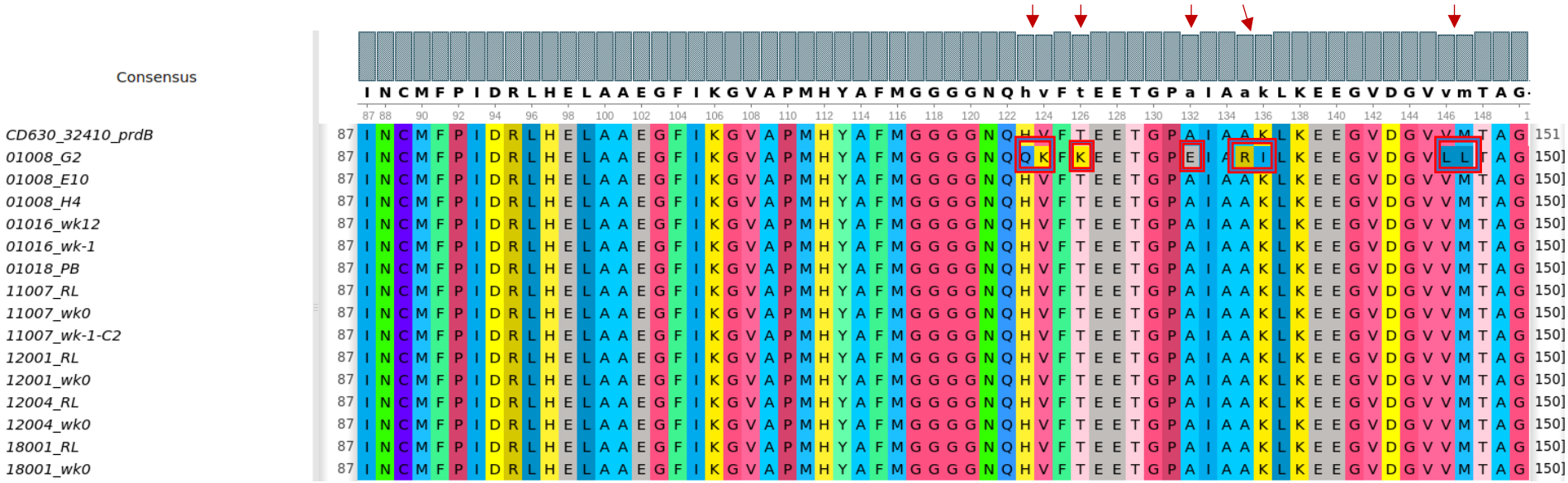


Figure 3.10: MUSCLE alignment of CD630_32410 (*prdB*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red boxes and arrows highlight example amino acid changes observed across sequences.

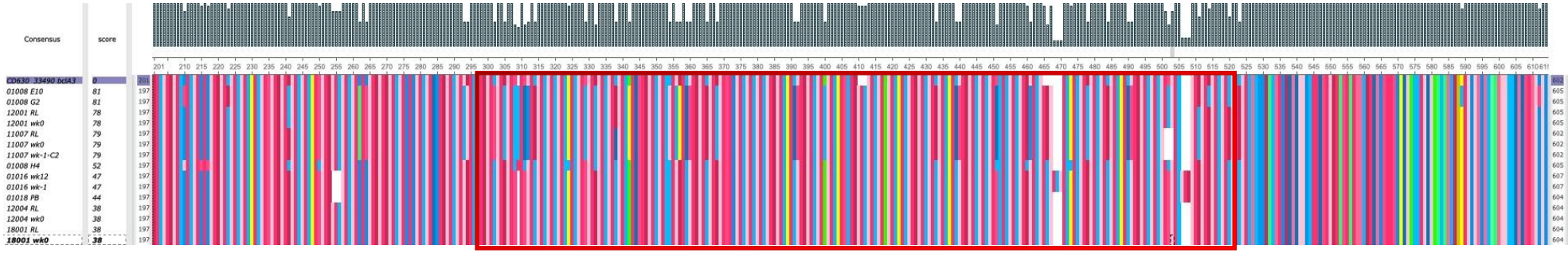


Figure 3.11: MUSCLE alignment of CD630_33490 (*bclA3*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red box highlights the variable region of this gene across sequences.

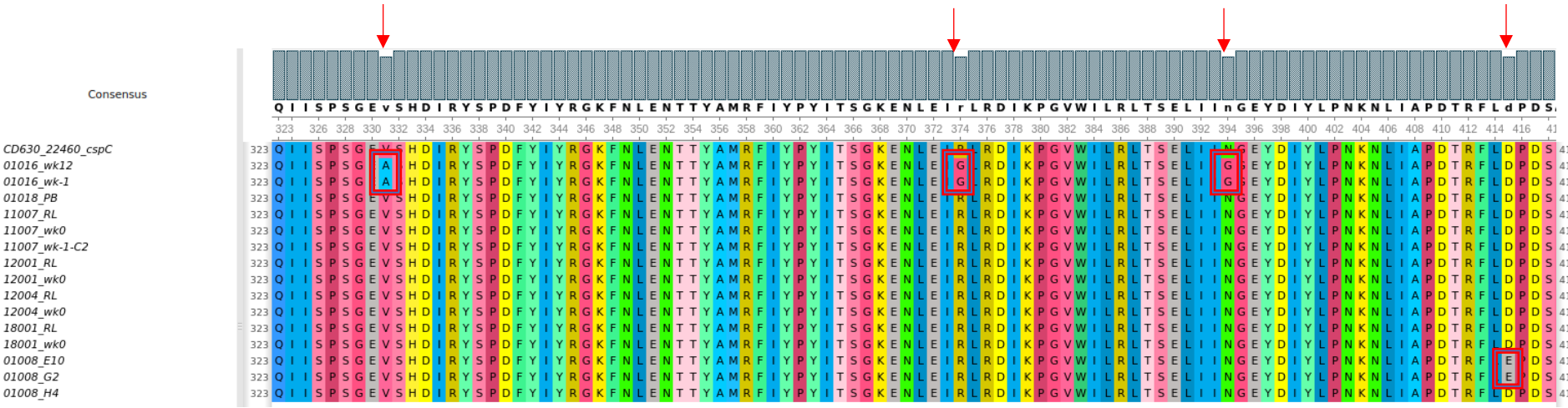


Figure 3.12: MUSCLE alignment of CD630_22460 (*cspC*) against translated protein sequences for each genome of the recurrence cohort as visualised using UGENE.

A ClustalW colour scheme was selected for output in which each residue in the alignment is assigned a colour corresponding to the amino acid profile. The red boxes and red arrows highlight the amino acid changes across sequences.

3.2.5 Pan-genome analysis

Pan genome analysis was performed to highlight possible functional classes of genes in the accessory genomes of strains that may potentially cluster with recurrence and/or persistence. The predicted protein sequences of the 16 genomes from strains 01008 wk-1 (E1), 01008 wk12 (G2), 01008 RL (H4), 01013 wk12, 01016 wk12, 01016 wk-1, 01018 wk-1 (C17), 11007 RL, 11007 wk0, 11007 wk-1 (C2), 12001 RL, 12001 wk0, 12004 RL, 12004 wk0, 18001 RL and 18001 wk0, were used as input for the pan-genome analysis using the BPGA pipeline, with the USEARCH clustering tool. The Pan-genome size was calculated at 3101 genes at n=16, whilst the trajectory of the pan-genome continues to increase, the curve appears to be entering a plateau (see Figure 3.13), characteristic of it being nearly closed in this recurrence cohort, as described by (Knight *et al., 2*015). The number of core genes converges at 3101 at n = 16, accounting for 64.3% of the total gene pool. Whilst the accessory genome varied from 246 to 612 genes across strains (see Table 3.5).

New gene families, defined as several genes containing the highest number of homologous sequences without any associated function (Ufarté *et al.,* 2015), were identified mostly amongst persistent carriage and relapsing causing strains, 01008 wk12 (G2), 01008 RL (H4), 01013 wk12, 01016 wk12, 11007, 12001 RL and 12004 RL, with the exception of 01008 wk-1 (E1) and 01018 wk-1, a symptomatic carriage causing strain (Figure 3.14). The largest number of gene families were observed in strains 01008 wk-1 (E1) (n=458), 01016 wk12 (n=66), 11007 RL (n=170) and 12001 RL (n=120). Whilst the content of the accessory genomes specific for each strain will of course need further data mining, the size of the accessory genome for the relapse causing strain 01008 RL (H4) is >40% smaller than the other strains isolated from the same patient in this study (see Table 3.5). This strain also had the highest number of unique genes (n=16), followed by strains 01013 wk12 (n=13) and 01018 wk-1 (C17) (n=11) (see Table 3.5). The Pan-genome analysis revealed 17 exclusively absent genes in four strains (01008 RL

(H4), n=4), 01013 wk12, n=6, 01018 wk-1 (C17), n=6 and 18001 wk0, n=1) that were present within all other genomes analysed (see Table 3.5).

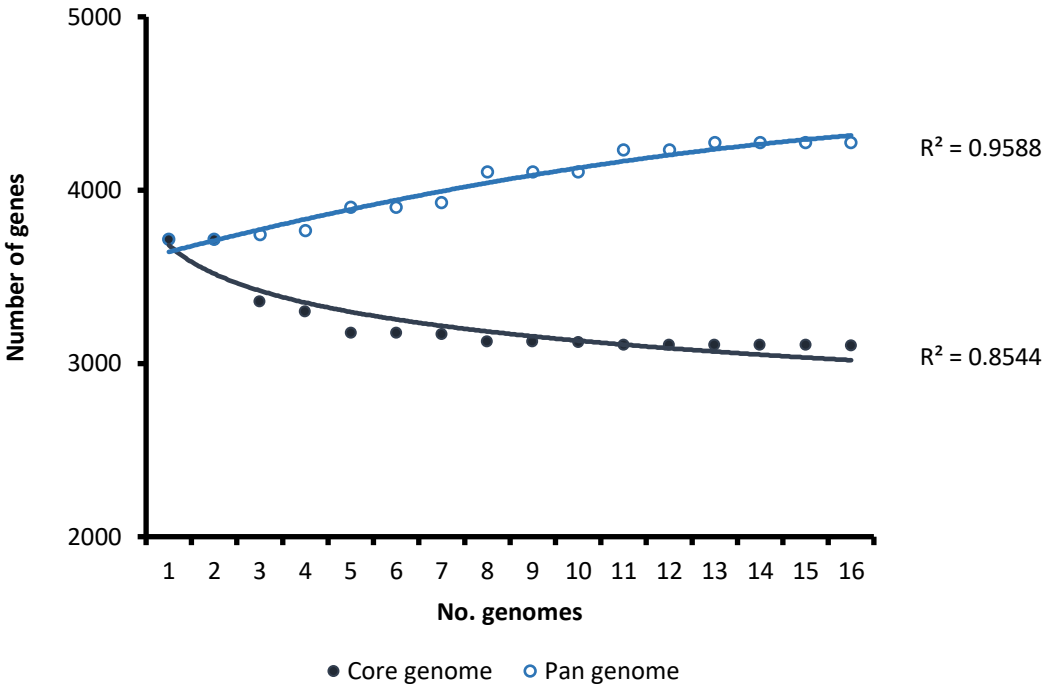


Figure 3.13: Core and Pan-genome analysis for 16 genomes from clinical isolates of *C. difficile*.

A total of 16 *C. difficile* genomes from clinical isolates have been analysed using the default parameters of the BPGA pipeline. The total number of genes in the pan (open circles) and core (closed circles) genomes are plotted as a function of the number of genomes sequentially added (n=16). For the Pan-genome the continuous curve shows the non-linear fit (R2 = 0.9588 and 0.8544, respectively).

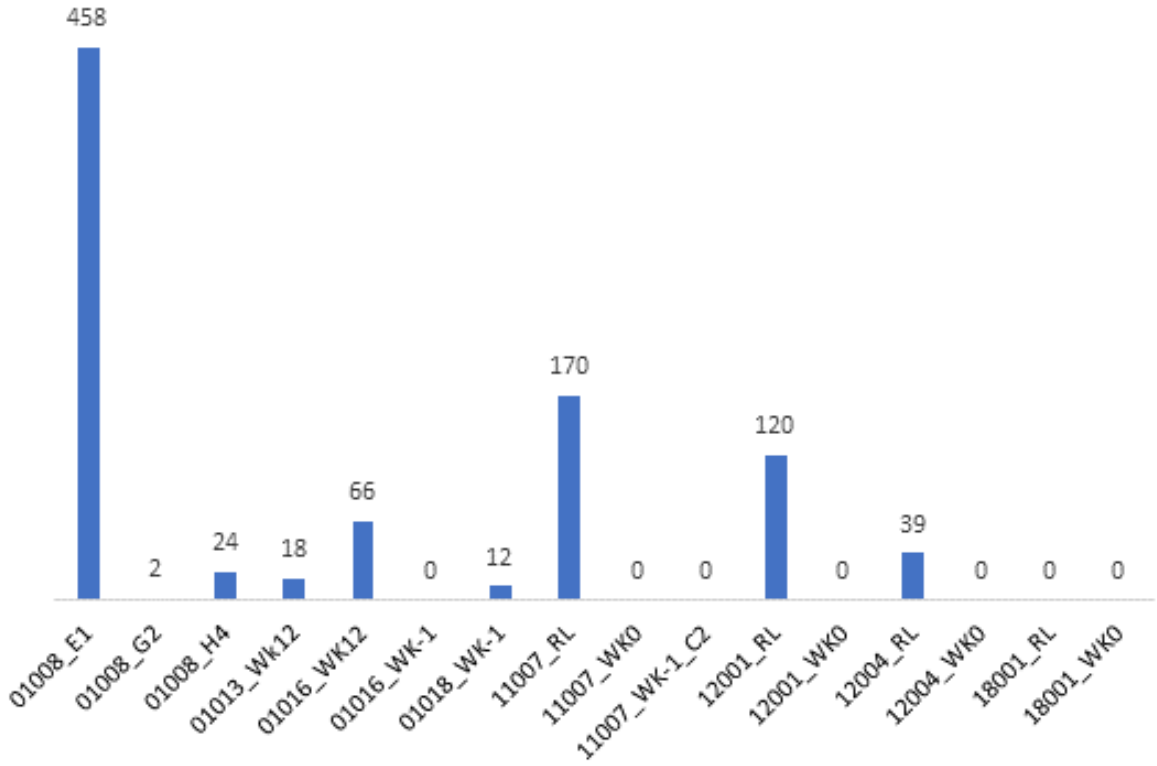


Figure 3.14: Number of new gene families revealed in the Pan-genome analysis of 16 isolates of *C. difficile* using the BPGA pipeline.

Table 3.5: Core, accessory, unique and exclusively absent genes of 16 clinical isolates of *C. difficile* after pan-genome analysis using the BPGA pipeline.

Genome no.	Organism name	No. of core genes[a]	No. of accessory genes[b]	No. of unique genes[c]	No. of exclusively absent genes[d]
1	01008_E1	3101	612	0	0
2	01008_G2	3101	611	0	0
3	01008_H4	3101	268	16	**4**
4	01013_Wk12	3101	329	11	**6**
5	01016_WK12	3101	247	0	0
6	01016_WK-1	3101	246	0	0
7	01018_WK-1	3101	288	13	**6**
8	11007_RL	3101	411	0	0
9	11007_WK0	3101	412	0	0
10	11007_WK-1_C2	3101	406	0	0
11	12001_RL	3101	370	0	0
12	12001_WK0	3101	370	0	0
13	12004_RL	3101	314	0	0
14	12004_WK0	3101	314	0	0
15	18001_RL	3101	315	0	0
16	18001_WK0	3101	306	0	**1**

Pan genome as defined by Chaudhari et al., (2016): [a] Core genes = conserved or shared by all strains in a group of isolates , [b] Accessory genes = dispensable or specific to a single strain, [c] Unique genes = stain-specific or shared by two or more but not all strains, [d] Exclusively absent genes = orthologous families that contain genes from all genomes except one specific genome.

The predicted core, accessory and unique genes in all 16 strains were further classified according to Cluster of Orthologous groups (COG) categories pathways using default parameters within the BPGA pipeline. Interestingly 20% of the major COG pathway categories remained poorly characterised for the core genomes (see Figure 3.15). Analysis showed the core (conserved) gene pool was found to be most abundant in all metabolism classes in comparison to the accessory genome (see Figure 3.15). Whilst further functional classification according to COG category found the accessory (dispensable) genome was most abundant in defence mechanisms, transcription, replication, and recombination and repair pathways (see Figure 3.16). The distribution of unique (strain-specific) genes according to COG categories was among transcription, replication, recombination and repair, energy production and conversion, carbohydrate transport and metabolism pathways. Results of a tblastn search against the NCBI database querying 40 unique genes of from 3 clinical *C. difficile* strains are presented in Table 3.6 - Table 3.8.

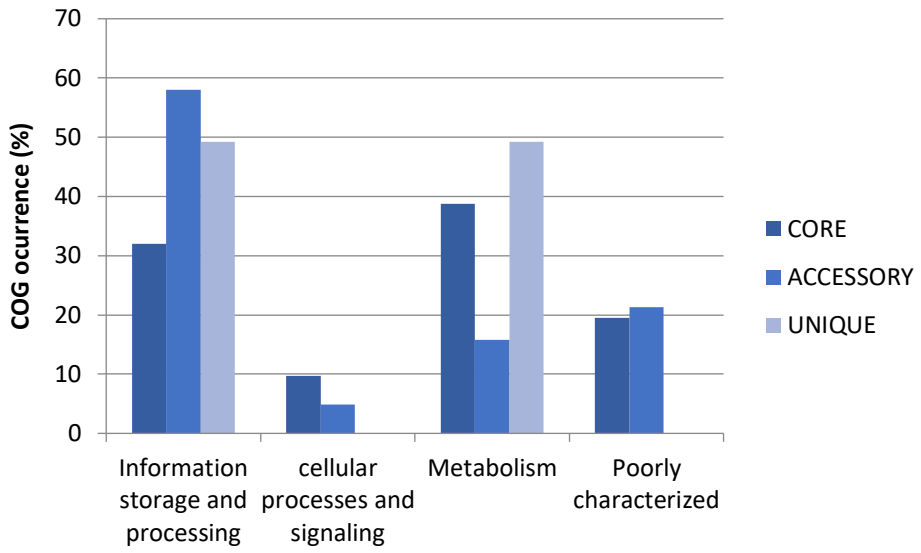


Figure 3.15: Major Cluster of Orthologus Groups (COG) categories of Pan-genome analysis of 16 isolates of *C. difficile*.

Comparison of the major COG distribution between core, accessory and unique genes of the 16 *C. difficile* strains analysed using the default parameters of the BPGA pipeline.

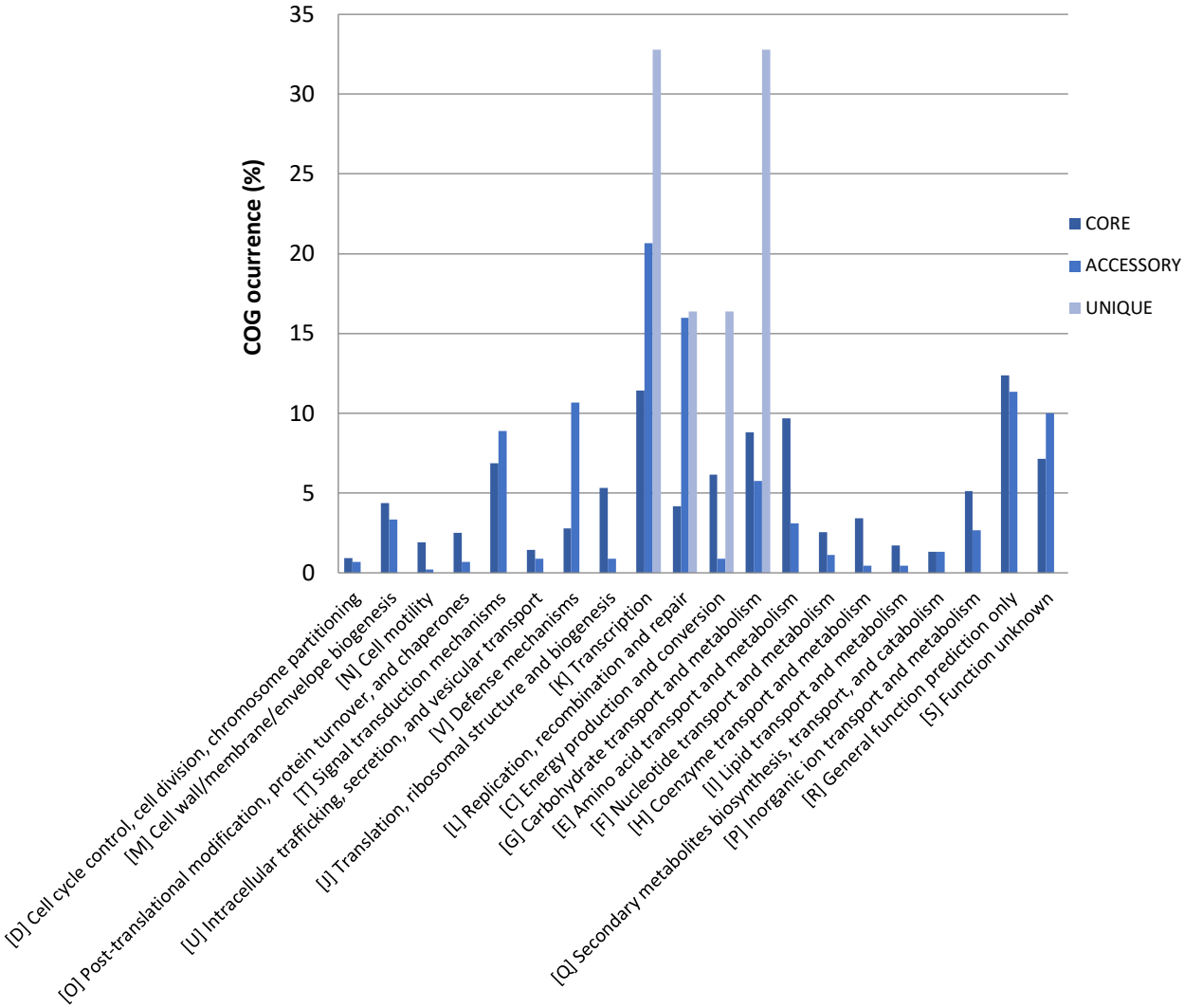


Figure 3.16: Detailed Cluster of Orthologus Groups (COG) distribution of Pan-genome analysis of 16 isolates of *C. difficile.*

Detailed functional classification of specific cellular systems and biochemical pathways of COG distribution profiles of core, accessory and unique genes amongst the 16 *C. difficile* strains analysed using the default parameters of the BPGA pipeline. Bracketed letters correspond to the functional catergory code.

Table 3.6: Results of a tblastn search querying 16 unique genes of clinical *C. difficile* isolate 01008_H4 against the NCBI database.

01008_H4.faa query ID	tblastn match description	Query cover	E value	Identity	Accession	Homologous protein annotation
unique/1776/3/Org3_Gene2048	*Clostridioides difficile* strain F II 3 Cwp29 gene, complete cds	100%	0	100%	JQ389413.1	CWP29, Cell wall protein
unique/3889/3/Org3_Gene1287	*Clostridioides difficile* strain DH/NAP11/106/ST-42, complete genome	100%	3.00E-43	100%	CP022524.1	CGC51_08825, AraC family transcriptional regulator - glycerol oxidative pathway *C. butyricum* VPI 1718
unique/3891/3/Org3_Gene1217	*Clostridioides difficile* strain FDAARGOS_267 chromosome, complete genome	100%	2.00E-43	100%	CP020424.2	A6J95_01875, integrase
unique/4005/3/Org3_Gene1528	*Clostridioides difficile* CD196 complete genome, strain CD196	100%	1.00E-41	100%	FN538970.1	CD196_1881, putative uncharacterized protein
unique/4006/3/Org3_Gene1808	*Clostridioides difficile* strain DH/NAP11/106/ST-42, complete genome	100%	6.00E-35	99%	CP022524.1	CGC51_11850, DNA-binding response regulator
unique/4087/3/Org3_Gene3474	*Clostridioides difficile* strain DH/NAP11/106/ST-42, complete genome	100%	3.00E-26	84%	CP022524.1	CGC51_01940, flagellar protein
unique/4095/3/Org3_Gene3608	*Clostridioides difficile* strain LEM1, complete genome	100%	2.00E-33	98%	CP019469.1	conserved hypothetical protein
unique/4143/3/Org3_Gene460	*Clostridioides difficile* strain DH/NAP11/106/ST-42, complete genome	100%	7.00E-29	100%	CP022524.1	CGC51_01940, flagellar protein
unique/4156/3/Org3_Gene3600	*Clostridioides difficile* strain AK plasmid pAK1, complete sequence	100%	2.00E-25	98%	CP027015.1	not known
unique/4197/3/Org3_Gene330	*Clostridioides difficile* strain LEM1, complete genome	100%	9.00E-26	100%	CP019469.1	not known
unique/4206/3/Org3_Gene30	*Clostridioides difficile* strain Z31 chromosome, complete genome	100%	3.00E-24	100%	CP013196.1	not known
unique/4211/3/Org3_Gene2736	*Clostridioides difficile* 630 complete genome	100%	5.00E-11	69%	FN668375.1	CD630_32330, putative sigma-54 interacting protein

Table 3.6 Continued - Results of a tblastn search querying 16 unique genes of clinical *C. difficile* isolate 01008_H4 against the NCBI database.

01008_H4.faa query ID	tblastn match description	Query cover	E value	Identity	Accession	Homologous protein annotation
unique/4220/3/Org3_Gene913	*Clostridioides difficile* BI9 chromosome	100%	3.00E-21	100%	FN668944.1	not known
unique/4229/3/Org3_Gene1525	*Clostridioides difficile* strain Z31 chromosome, complete genome	100%	2.00E-19	100%	CP013196.1	not known
unique/4242/3/Org3_Gene1988	*Clostridioides difficile* 630 complete genome	100%	8.00E-19	97%	AM180355.1	conserved hypothetical protein
unique/4244/3/Org3_Gene3606	*Clostridioides difficile* strain CD161 plasmid unnamed1, complete sequence	100%	9.00E-19	97%	CP029155.1	not known

Table 3.7: Results of a tblastn search querying 11 unique genes of clinical *C. difficile* isolate 01013_wk12 against the NCBI database.

01013_wk12.faa query ID	tblastn match description	Query cover	E value	Identity	Accession	Homologous protein annotation
unique/2492/4/Org4_Gene1500	*Clostridioides difficile* strain C192 Cwp26 gene, complete cds	100%	2.00E-161	99%	JQ389338.1	Cwp26, cell wall protein
>unique/2662/4/Org4_Gene541	no significant similiarity to *Clostridioides difficile* taxid:1496 or Clostridioides group firmicutes taxid:1239 or NCBI database from tblastn search					
>unique/3734/4/Org4_Gene533	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_105	100%	3.00E-44	88%	LK932309.1	conserved hypothetical protein
>unique/3800/4/Org4_Gene3193	*Clostridioides difficile* strain W0003a chromosome	100%	4.00E-50	100%	CP025047.1	phospho-sugar mutase
unique/4034/4/Org4_Gene548	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_44	100%	5.00E-09	45%	LK932381.1	conserved hypothetical protein
>unique/4149/4/Org4_Gene546	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	73%	4.00E-17	100%	CP011847.1	hypothetical protein
unique/4186/4/Org4_Gene991	*Clostridioides difficile* strain 630 delta erm, complete genome	97%	5.00E-14	81%	CP016318.1	putative peptidase, S9 family
>unique/4214/4/Org4_Gene3662	*Clostridioides difficile* BI9 chromosome	100%	6.00E-20	98%	FN668944.1	not known
>unique/4238/4/Org4_Gene562	no significant similiarity to *Clostridioides difficile* taxid:1496 or Clostridioides group firmicutes taxid:1239 or NCBI database from tblastn search					
>unique/4255/4/Org4_Gene3718	*Clostridioides difficile* strain Z31 chromosome, complete genome	100%	4.00E-16	100%	CP013196.1	Periplasmic pH-dependent serine endoprotease DegQ precursor
>unique/4264/4/Org4_Gene1662	Pepto*Clostridioides difficile* genome assembly 7032985, scaffold BN1096_Contig_56	100%	1.00E-14	100%	LK932509.1	BN1096_560302, Glycerol kinase 1

Table 3.8: Results of a tblastn search querying 13 unique genes of clinical *C. difficile* isolate 01018_wk-1 against the NCBI database.

01018_wk-1.faa query ID	tblastn match description	Query cover	E value	Identity	Accession	Homologous protein annotation
unique/757/7/Org7_Gene3162	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	0	93%	CP011847.1	CDIF27639_03151, ABC transporter ATP-binding protein
>unique/962/7/Org7_Gene3104	*Clostridioides difficile* strain W0022a chromosome	100%	0	94%	CP025046.1	hypothetical protein
unique/2317/7/Org7_Gene3095	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	2.00E-146	94%	CP011847.1	CDIF27639_03068, ABC transporter ATP-binding protein
unique/3123/7/Org7_Gene3161	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	2.00E-100	100%	CP011847.1	CDIF27639_03149, phage protein
>unique/3185/7/Org7_Gene3160	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	1.00E-96	99%	CP011847.1	CDIF27639_03148, phage repressor, annotated as Transcriptional regulator, HTH-type in CD630 AM180355.1 (CD630_29500)
>unique/3484/7/Org7_Gene3144	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_129	100%	2.00E-79	100%	LK932335.1	conserved hypothetical protein
>unique/3618/7/Org7_Gene3145	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_129	85%	3.00E-58	100%	LK932335.1	conserved hypothetical protein
>unique/3666/7/Org7_Gene3149	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_37	100%	4.00E-49	82%	LK932373.1	conserved hypothetical protein
>unique/3714/7/Org7_Gene3103	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_90	100%	1.00E-59	100%	LK932432.1	conserved hypothetical protein

Table 3.8 Continued - Results of a tblastn search querying 13 unique genes of clinical *C. difficile* isolate 01018_wk-1 against the NCBI database.

01018_wk-1.faa query ID	tblastn match description	Query cover	E value	Identity	Accession	Homologous protein annotation
>unique/3954/7/Org7_Gene3159	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	5.00E-42	100%	CP011847.1	CDIF27639_03147, phage repressor (transcriptional regulator, phage-type in CD630 AM180355.1 CD630_29490 - 100% ID)
>unique/3983/7/Org7_Gene3119	*Clostridioides difficile* strain 08ACD0030 chromosome, complete genome	100%	1.00E-38	97%	CP010888.1	TW87_07910, XRE family transcriptional regulator
>unique/4075/7/Org7_Gene3157	*Clostridioides difficile* strain DSM 27639 chromosome, complete genome	100%	2.00E-34	100%	CP011847.1	hypothetical protein
>unique/4097/7/Org7_Gene3143	Pepto*Clostridioides difficile* genome assembly 7032994, scaffold BN1097_Contig_129	100%	3.00E-32	100%	LK932335.1	conserved hypothetical protein

3.3 Discussion

3.3.1 PCR Ribotype prevalence and its association to CDI recurrence

Major outbreaks worldwide caused by the 'hypervirulent' PCR ribotype 027, have led to thorough studies linking it to severe diarrhoea, higher mortality and more recurrences than patients infected with other ribotypes (Hubert *et al., 2*007; Loo *et al., 2*005a; Redelings, 2007). The decline in prevalence globally, demonstrates the impact of surveillance and research on infection control. Nevertheless, there are other ribotypes that are now emerging and becoming more predominant in UK hospitals. It is still unclear what genetic characteristics differentiate these ribotypes from others and which make them prevalent in recurrence. In this study, there was no correlation between a single ribotype for all cases of recurrence. Ribotypes 002 and 014 were the most prevalent in this study and were the most represented in reinfection cases.

3.3.2 Sensitivity of PCR ribotype compared to SNP analysis to classify recurrence

Within the 'recurrence cohort', clinical recurrence (referred to by healthcare workers as 'relapse') was described for 5 patients, 01008, 11007, 12001, 12004 and 18001, displaying symptomatic infection within 3 months of treatment. The criteria of fellow researchers working on this trial, towards broader genomic studies, was that a difference of >100 SNPs found between strains indicated reinfection of a new strain, whist <100 SNPs found indicated a relapse of the same strain (see appendix Table 7.4). This local classification of recurrence, based on WGS, was also used in this study to maintain continuity and so all but one case of relapse was confirmed by SNP analysis. Whereas SNV difference, based on SNVs observed in the core genome, to classify transmission is known to be more stringent in the literature. Of the 5 relapse samples submitted in this recurrence cohort, only 1 would have been confirmed by SNV analysis, according to the criteria of Eyre *et al.,* (2013), whereby <2 SNVs were found

between the initial infecting strain and the relapse isolate and genetically distinct strains classified as >10 SNVs (Eyre *et al.,* 2013). The remaining clinical 'relapse' strains would have been classed as indeterminate, with 3-10 SNVs between the initial infecting strain and the clinical 'relapse' isolate, except strain 01016 wk 12 containing 11 SNPs from the index strain. However, it has been commented that these thresholds may be underestimating the evolutionary distance between strains due to the genetically quiescent nature of *C. difficile* spores (Khanna *et al., 2*012).

Whole genome SNP analysis of isolates was able to differentiate two cases of reinfection from carriage for patients 18001 (Relapse sample) and 11007 (week 0 sample), which would not be possible using PCR ribotyping alone. PCR ribotyping detected all cases of recurrence, as defined by the return of symptomatic disease post treatment. However, WGS was able to discriminate between a case of reinfection by SNP analysis, that would have been recognised as asymptomatic carriage by PCR ribotyping. Despite the higher discriminatory power of WGS, in this cohort there was a 100% sensitivity for both methods in detecting recurrence. But for the purposes of this study, WGS did not outperform the current method of PCR ribotyping, used in routine reference laboratories.

WGS analysis was able to confirm the clinical recurrence rate of 62.5 % (n=5 patients) in this recurrence cohort. It was also shown to be much more sensitive for the discrimination of recurrence cases revealing a relapse rate of 80% (n=4 patients), correctly classifying relapse and reinfection. Whilst PCR ribotyping, identified the same relapse rate of 80% (n=4 patients), two misclassifications were found (strain 01008 RL (H4) incorrectly as reinfection and 18001 RL incorrectly as relapse based on PCR ribotype detected). Whilst clinical diagnosis of CDI does not aim to differentiate between reinfection or relapse, it should also be highlighted that it does not monitor for persistence either. Isolation of positive *C. difficile* from two patients at week 12, one representing a case of asymptomatic reinfection (01013 wk 12) and one of persistence carriage (01016 wk 12), reveals that these events would be missed entirely by routine diagnostic assays employed. Despite both cases being presented as asymptomatic,

there is the potential for these carriers to shed *C. difficile* into the environment and present as transmission reservoirs, in addition to potentially becoming symptomatic again at a later timepoint.

Whilst the depth of sampling and characterisation of isolates for routine CDI does not currently stretch to those carried out in this study, questions should be raised as to whether further surveillance measures to discriminate between relapse and re-infection or to identify cases of asymptomatic carriers should be asked in the healthcare setting to reduce the level of incidence and transmission and inform infection control procedures.

3.3.3 Genomic diversity of isolates

The average genome was 4.21 M bp, with strain 01008 wk-1 (E2) having the largest genome (4.65 M bp) and the smallest genome (4.13 M bp) belonging to strain 01018 wk-1 (C17). Phylogenetic analysis showed that all strains were very closely related, but generally clustered by patient ID. Clustering of strains within the same patient was also confirmed by ANI. No transmission of strains between patients was indicated by phylogenic analysis.

3.3.4 Sequence variation of genes of *C. difficile* associated with spore germination and outgrowth and implications on recurrence

Although several genes had variations in amino acid sequence between the isolates, the most dramatic changes were in CD630_03320 (*bclA1*). This gene consists of two conserved regions with a variable region of glycine repeats in the middle. All strains differed from the reference in this region. CD630_33490 (*bclA3*) showed a similar pattern, although with fewer changes in CD630_03320 (*bclA1*). The *Bacillus* collagen-like protein of *anthracis* (BclA) glycoprotein is a major component of the exosporium that carries collagen-like repeats along regions of the

peptide used for attachment to oligosaccharides in some spore formers (Sylvestre *et al*., 2003; Steichen *et al*., 2003). All three *C. difficile* BclA proteins (BclA1, BclA2 and BclA3) contain a similarly extensive, collagen-like region with multiple glycine repeats central to the peptide, flanked by N- and C-terminal domains of variable length (Phetcharaburanin *et al., 2*014), as observed in this sequence analysis. The BclA1 protein is located on the spore surface of *C. diffcile* spore (Pizarro-Guajardo *et al*., 2014) and appears to be a critical factor in the initial stages of colonization (Phetcharaburanin *et al., 2*014). Small amino acid changes, usually between 1-5 changed amino acids were predicted from the translated protein sequences for additional genes but in only a few strains. A couple of alignments seemed to indicate truncated genes; CD630_32410 (*prdB*) was truncated in all strains, whilst CD630_08570 (*oppF*) and CD630_03320 (*bclA1*) were truncated in strains 01008 wk-1 (E10) and 01008 RL (H4) showed respectively.

For *C. difficile*, both glycine and L-proline are required for optimal growth in defined media (Osgood *et al., 1*993; Halsam *et al., 1*986; Karasawa *et al., 1*995) and their metabolism is achieved with enzymes gylcine reductase and proline reductase, encoded by genes *grdA* and *prdB*, respectively. One study has shown a more explicit role for *prdB* over *grdA*, where the lack of *grdA* did not affect growth of *C. difficile* in rich medium. Whilst the proline-dependent expression of *prdB* was still required for optimal growth of *C. difficile,* even in rich complex medium where bacteria can find and utilize a multitude of carbon and nitrogen sources (Bouillaut *et al., 2*012). It would be interesting to see if the truncation observed in *prdB* for all clinical isolates, in addition to the amino acid changes exclusively detected in the persistently carried strain 01008 wk12 (G2) would have any effect on the cell growth or spore outgrowth.

An oligopeptide permease, also known as Spo0K in spore formers, is encoded by five genes in the *opp* operon, (*oppB, oppA, oppC, oppD* and *oppF*) (Edwards *et al., 2*016). These groups of molecules are known to be involved in other organisms, mainly as quorum sensing molecules, in processes contributing to competence in *Bacillus* and *Streptococcus* species, plasmid transfer in *Enterococcus faecalis*, and the expression of virulence factors in *Bacillus*

thuringiensis (Edwards *et al., 2*014; Moraes *et al., 2*014). The role of *opp* molecules in initiating sporulation in *C. difficile* has been attributed instead to responses to nutrient availability and uptake, rather than quorum sensing (Edwards *et al., 2*014). Their study showed that mutants lacking oligopeptide permease genes, resulted in increased sporulation, likely because of the inability of the mutant to import small peptides as a nutrient source (Edwards *et al., 2*014). Therefore, the truncation of *oppF* in strain 01008 wk-1 (E10) may play a potential role in *C. difficile* sporulation and or virulence. The truncation of *bclA1* was also observed in this patient and was acquired between initial infection and the relapse (01008 RL (H4)) episode for which this truncation was detected. It has been reported that strain R20291, a 'hypervirulent' 027 strain, that carries a truncated BclA1 protein had a higher ID_{50} when used to infect mice (Phetcharaburanin *et al., 2*014). Whilst mice in this study were less susceptible to infection when challenged with *bclA1*$^{-}$ spores, these spores germinated faster than the wild-type strain. It would be interesting therefore to see if this strain had any similar phenotypes observed in this study.

Persistent carriage of *C. difficile* strains was detected four weeks after treatment for two patients (01008 and 01016) on rifaximin chaser therapy. Whilst only one case went onto relapse, this presents as failure of rifaximin to clear *C. difficile* from both patients. This presents the possibility of genetic fitness for these strains over others in this recurrence cohort. Interestingly, all missense mutations that were exclusively for genes CD630_34640 (*gerS*), CD630_32410 (*prdB*), CD630_00660 (*rpoB*), CD630_05510, (*sleC*) and CD630_22460 (*cspC)* were found in strains that were persistently carried. Interestingly, both mutations in *rpoB*, are known to confer rifamycin antibiotic resistance in other bacterial species. With the glutamine-to-lysine amino acid change at position 489 (489 Q-> K), corresponding to the same position at which the glutamine residue of RpoB in *Thermus aquaticus* binds to rifampicin (O'Connor *et al., 2*008). Additionally, the aspartic acid-to-valine amino acid change at position 492 (492 D-> V), corresponds to the same position at which aspartic acid residue alterations confer rifamycin resistance in *Staphylococcus aureus* (O'Connor *et al., 2*008; Murphy *et al., 2*006). It is therefore likely that these amino acid changes also confer resistance to rifamycin,

with the potential for this to extend resistance to rifaximin (a rifamycin derivative). Whilst none of the genes of interest contained missense changes that were exclusively found in coinfecting strains or relapsing strains, they were found for a non-persistently carried isolate 01008 wk-1 (E10). It will be interesting to see if any of the genetic changes observed for the genes will be translated into phenotypic changes for the strains of which they where detected.

This study, however, does not consider whether silent nucleotide or conservative amino acids are found in the promoter regions of genes, that could potentially affect levels of expression, of the genes themselves or other genes for which they may be involved in regulation. Additionally, all nucleotide changes would need to be confirmed by sanger sequencing to eliminate possible WGS sequencing or mapping and annotation errors. Whilst interesting, the genes of interest selected are not exhaustive to the associated phenotypes, more genes would need to be included to more confidently attribute genetic fitness of strains to mechanisms of growth, antibiotic resistance, germination, sporulation and persistence. Furthermore, other virulence factors exist outside of this focus of sporulation and germination, such as biofilm production and adherence for example.

Nevertheless, this small sample set does highlight some interesting genetic differences between strains involved in recurrence and for those that are persistently carried. The vast sequence variation observed in this recurrence cohort implicates genes CD630_03320 (*blcA1*), CD630_33490 (*bclA3*), CD630_08570 (*oppF*), CD630_32410 (*prd*B) and CD630_22470 (*cspBA*) in clinical infection when compared to the non-clinical circulating reference strain CD630. Additionally, the findings of exclusive amino acid changes in genes CD630_34640 (*gerS*) CD630_32410 (*prdB*), CD630_00660 (*rpoB*), CD630_05510, (*sleC*) and CD630_22460 (*cspC*) implicate a role in the advantageous fitness of persistently carried strains over others isolated from the same host, specifically in mechanisms of germination and growth.

3.3.5 Pan-genomes of *C. difficile* and implications on recurrence

As mentioned, genes of interest used in the sequence variation analysis are not exhaustive. Other genes within phenotypes of sporulation, spore dormancy, germination and outgrowth, and indeed other functional classes of unknown importance may play key roles in the genetic fitness of strains, which could potentially be associated with recurrence and persistence.

Pan and core genome analysis revealed that the genomes of isolates have not undergone massive gene rearrangements indicated by the low number of unique and exclusively absent genes predicted in the genomes. The distribution of unique genes according to COG categories was among transcription, replication, recombination and repair, energy production and conversion, carbohydrate transport and metabolism pathways. This then indicates that for the isolates that have such changes, further characterisation is needed to evaluate whether these novel genes contribute to the fitness or virulence within recurrent CDI. Functional variation can be seen in both the core and accessory genomes of the strains isolated in this recurrence cohort. It is not yet clear whether functional variation of isolates can be discriminated down to the specific accessory genes they carry. Selective advantages such as ecological adaptation, antibiotic resistance, virulence mechanisms, or colonization of a new host is described to be conferred by the supplementary biochemical pathways and functions of the accessory genome as well as unique genes of strains (Chaudhari *et al., 2*016). Further data mining of both core and accessory gene outputs is needed to evaluate this. New gene families were consistently observed for the genomes of recurrence causing isolates and warrant further data mining of the contents of these and whether they can be associated with mobile genetic elements, antibiotic resistance or other virulence factors. Larger pangenome analysis including genomes of non-recurrent isolates is needed to confirm this trend.

3.3.6 Rifaximin as a chaser therapy for 'relapse'

Overall from this clinical trial data shows a reduction in the isolation *C. difficile* strains at all time points for patients in the rifaximin group compared to those on placebo. 'Clinical relapse' episodes were observed up to three months after the trial therapy (wk 0) started for five patients (01008, 11007, 12001, 12004 and 18001) in this cohort, indicating a clear need for chaser therapy to prevent recurrence. Data for a subset of *C. difficile* isolates sampled from patients treated with rifaximin (n=5) and placebo (n=3) were presented in this chapter. The use of rifaximin being assessed for its ability to prevent recurrent CDI was the aim for the clinical trial, however, recurrence was observed in both study groups. For patients treated with rifaximin (01008, 01013, 01016, 12001 and 18001) three experienced a relapse (patients 01008, 12001 and 12004) whilst four patients experienced a reinfection (patients 01008, 01013, 11007 and 18001) within 12 weeks post diagnosis.

Interestingly, patient 01008, prescribed metronidazole during clinical treatment, had non-conservative amino acid changes in *rpoB* detected in two strains (01008 wk-1 (E10) and 01008 wk12 (G2) that were known to confer resistance to rifampicin (in other bacterial species. Neither of these strains were isolated from the relapse submitted sample; yet strain 01008 RL (H4) containing no missense mutations in the *rpoB* gene was also carried during initial infected (01008 wk1 (E2)) and isolated beyond Rifaximin chaser therapy. Many speculations can be drawn for this case:

- Mutations in *rpoB* for strains 01008 wk-1 (E10) and 01008 wk12 (G2) may had conferred resistance to rifaximin, thus evaded clearance during chaser therapy.

- Strains (01008 wk-1 (E10) and 01008 wk-1 (E2)) exposed to sub-inhibitory concentrations of rifaximin evaded clearance due to undergoing processes of sporulation or existed in environmental niches such as biofilms and emerged beyond rifaximin treatment at week 12 of the trial (01008 wk12 (G2)) and later in the relapse sample (01008 RL (H4)).

- Spores of all strains in this patient existed at some level throughout the course of infection and evaded clearance due to non-sporicidal nature of rifaximin. The sampling schedule of the trial led to the bacterial isolation by culture. Partial clearance of spores and vegetative cells for strains from the host resulted in detection of strains that were able to persist.

All speculations would need experimental testing, and all could indeed prove inadequate, with additional unknown factors contributing to the complexity of this case. It would be interesting to see if spore phenotypes could support or indeed disprove hypotheses made.

For patients treated with placebo in the subset presented in this thesis (01018, 11007 and 12004) two experienced reinfection and one a relapse. Whilst it could be argued that these recurrent events were identified by further microbiological and genomic study outside the current scope of routine diagnostics and so not necessarily indicative of symptomatic recurrence, clinical recurrence (described only as relapse) was also reported in 5 patients. To then assess the ability of rifaximin to prevent 'clinical recurrence' (relapse), this chaser therapy performed quite poorly in this small subset, with three patients (01008, 12001 and 12004) clinically described as relapsing with CDI despite being administered Rifaximin, compared to only one on placebo (patient 18001). Evenso, isolate 18001 RL, was submitted as a relapse sample but a new strain with a different PCR ribotype was detected, indicating a symptomatic reinfection event. As such this does not indicate a failure of rifaximin as a chaser therapy, but this may be due to new resistant strain being acquired after the chaser treatment period in this patient. Even so, with this limited sample set, conclusions as to efficacy of Rifaximin as a chaser therapy to prevent recurrence for all patients enrolled on this clinical trial may have had a different outcome.

The use of a chaser therapy would still not target the dormant spores and so would not likely to be effective with spores for remain dormant for longer periods of time. Additionally, it is unknown whether chaser therapy would be effective on slow outgrowing spores during relapse or in cases of co-infecting strains. Questions remain as to whether genomic or

environmental changes affect the regulation of germination or outgrowth of spores for relapse or reinfection causing spores to stay undetected and untreated in the host and environment leading to the cycle of CDI and complication of recurrence to return.

3.3.7 Limitations and benefits of this study

Following the pilot study of rifaximin as a chaser therapy (Garey *et al., 2*011), researchers commented that further studies in clinical settings were required to confirm its efficacy in the treatment of CDIs. This study aimed to obtained samples from 151 patients that met the inclusion criteria. Whilst samples were submitted from only 92 patients, this study had a larger sample set then that of the pilot study, that found a decrease in the incidence of recurrence for patients on rifaximin compared to placebo chaser therapy. A total of 15 patients submitted 'clinically relapsed' samples, revealing a total 16.3% recurrence rate. A small reduction in recurrence rate for the rifaximin treated cohort (6.5%) compared to the placebo treated cohort (9.78%) was observed supporting the findings of the pilot study. Additionally, a larger reduction in *C. difficile* culture-positive 'clinical relapse' samples, as well as a larger reduction in the number of *C. difficile* culture-positive samples submitted at week 12 for patients treated with Rifaximin compared to the placebo. Indicating a large improvement in the persistent carriage and relapse episodes for patients treated with Rifaximin.

Published work has also commented on the need for inferences of virulence factors correlating with recurrence (Garey *et al.,* 2008; Garey *et al., 2*011; Eyre *et al.,* 2012). In this study comparative whole genome analysis of multiple samples from the same patient revealed genetic changes occurring over the course of infection. Further genetic variation analysis for genes of phenotypic interest indicated amino acid variations amongst isolates of the recurrence cohort, that may have implications towards features of relapsing strains. Whilst this analysis can be critiqued in that it was limited to only 24 genes and only to isolates

involved in recurrent infection, it is hoped that phenotypic studies reveal whether these genetic changes can translate into measurable phenotypic advantages. It could also be critiqued that mutations need to be confirmed to exclude errors of sequencing, mapping, annotation or SNP calling methods. Nevertheless, one significant advantage of this study is the use of a genetically related strain as the reference strain not only for genome assembly but also SNP calling. Other studies use the same reference, either CD630 or a ribotype 027 strain (Wroblewski *et al.,* 2009; Didelot *et al.,* 2012; Eyre *et al.,* 2012; 2013; He *et al*., 2013; Kumar *et al.,* 2016), limiting the genomic variation called by SNP algorithms to infer recurrence status of clinical isolates, but also inferences of variation to genes of known phenotypes for which the reference strain CD630 is commonly based. Reasonably, any amino acid changes observed for genes of phenotypic interest will be of interest, as differences in phenotypes may be associated with genomic variation observed.

A major limitation of this study however is the lack of transcriptomic data for genes of interest as well as the nominal genes of interest investigated. Genes with key phenotypes in toxin production and antibiotic resistance were excluded from this study. There is already an over representation of genetic variation attributed to these virulence factors, with no definite association with germination phenotypes due to difference in experimental methods, of which this study aims to interrogate. Additionally, these phenotypes have been included in the work of other researchers on this trial (Kinsmore, Natasha Louise (2019) *Understanding certain Clostridium difficile virulence and antibiotic resistance factors and how these relate to patient clinical outcome*. PhD thesis, University of Nottingham). Evenso, data on toxin titres would have supported assumptions of symptomatic disease or asymptomatic carriage, as no clinical information was available to confirm symptomatic disease at each sampling time point, particularly for asymptomatic persistent carriage assumed at week 12 sampling points. Unfortunately, the genomic and phenotypic analyses of the sample set in the research presented here was not performed on each isolate due the workload and time constrains of this study. A sister study including genomic analysis for a larger sample set including non-recurrent isolates obtained for this trial was also underway at the time of this research.

Nevertheless, it is anticipated that genomic variation highlighted for clinical isolates in this chapter may be associated with measurable phenotypic differences by the experimental methods employed in Chapter 4, which would be of significant benefit for this study.

3.3.8 Conclusion and future implications

Alternative management of CDAD, such with use of other antibiotics anti-toxin antibodies, the use of probiotics, have been debated in the literature (Aslam *et al., 2*005; Dendukuri *et al., 2*005). At a glance these options may seem favourable in circumventing *C. difficile's* large number of mobile genetic elements carrying antibiotic resistant genes (Sebaihia *et al., 2*006b), but are yet to be implemented in the clinical setting. Timely development of narrow-spectrum antibiotics such as rifaximin and FDA approved fidaxomicin (Shah *et al., 2*010), appears promising. However, transmission of the disease, caused by heat resistant spores which are difficult to eradicate using routine disinfectants, has remained problematic in health care settings. And so, whilst chaser therapy may provide a benefit for preventing continued infection of the same susceptible strain, reinfection events of strains that are resistant to these therapies can only be prevented by enforcing stringent sporicidal cleaning strategies. Questions as to whether sampling techniques in this study should be adopted routinely to enhance surveillance strategies, as well as recurrence classification should be considered. As mentioned, detailed sequence variation in the genes of interest were observed in this recurrence cohort. Whilst not exhaustive, they do shed light on potential genetic mechanisms of advantageous fitness for strains isolated from the same host. It would be interesting to see whether these have any measurable implication on the phenotypes of clinical isolates, particularly in sporulation, spore dormancy, germination and outgrowth.

Chapter 4

Chapter 4: Spore Phenotypes of *C. difficile* Isolates During Recurrent Infection

4.1 Introduction

Correlations between relapse and virulence factors including toxinotype, ribotype, cytotoxicity and even sporulation has not been investigated. The only organism-based risk factor demonstrated to date is that of germination (Oka *et al., 2*011; Moore *et al., 2*013). Whilst spores notably contribute to the survival and persistence of *C. difficile* within hosts during antibiotic treatment (Underwood *et al., 2*009; Walters *et al., 1*983), it is spore germination followed by efficient outgrowth of strains that allow for increased susceptibility of disease in patients inducing disease onset and the likelihood of relapse.

It is now known that strain- to- strain variation in sporulation rates is not type associated (Burns *et al., 2*011). Burns *et al.,* (2011) stipulate many critiques over small sample size, growth mediums, limited time points and confounding factors surrounding the enumeration of colony forming units in the methods of other studies. Oka *et al.,* (2012) express sporulation rate (as a percentage) as the ratio of spores to vegetative cells within a population. However, this relative measure can be affected by the growth differences among strains and by the fate of non-sporulating vegetative cells. So, whilst their study indicates that there is no significant correlation between sporulation ability and PCR or PFGE type or recurrence, the limitations of their method cannot rule out the opposite findings.

A disparity between germination rates has been observed when studies use dissimilar methods. The Moore *et al.,* (2013) study employed several techniques for the quantification of germination efficiencies and found that they were significantly higher in recurrent CDI or when metronidazole therapy failed than when therapy was successful [(49.00±10.49 %, 95 % CI 46.25–51.75 %) versus (41.42±9.43 %, 95 % CI 37.93–44.91 %); $P \leq 0.01$]. Both Moore *et al.,* (2013) and Oka *et al.,* (2012) employed methods of converting spore counts and colony

forming units of outgrowth on Gifu Anaerobic Medium (GAM) or BHI or BHIS supplemented with sodium taurocholate media into percentage (Nakamura *et al., 1*985; Kamiya *et al.,* 1989). Whilst other studies went further to include measurement of spore germination and outgrowth by loss in optical densities at 600 nm and rise in optical density at 600 nm respectively, (Moore *et al., 2*013). Moore *et al.,* (2013), using drop in optical density, observed decreased germination rates correlating to PCR ribotype 027, while Wheeldon *et al.,* (2011) and Carlson *et al.,* (2015) used loss of heat resistance and inability to outgrow in the absence of germinant to identify germination percentages. Here Wheeldon observed an increase in germination rates correlating to PCR ribotype 027. Whilst Carlson *et al.,* (2015) found that isolates that caused severe disease exhibited significantly lower germination efficiency, and no significant difference within the 027 type. Interestingly though, isolates of the most predominate type in hospitals during the collection period, type 014-020, exhibited higher levels of germination compared to those tested including ribotype 027. They highlighted that further study will be needed to identify differences inherent to these isolates that lead to this increased control over germination and to elucidate the role of this control in the pathogenesis of *C. difficile*.

In this chapter, the afore mentioned methods as well as Sorg and Sonenshein's (Sorg and Sonenshein, 2008b; Sorg and Sonenshein, 2010; 2009a) methods were employed to detect the effect of bile salts on the spores of *C. difficile* to elucidate the rate, efficiency and percentage of germination. The aim was to elucidate germination as a significant virulence and risk factor of recurrence in CDAD. Specific research questions include whether environmental pressures affecting spore phenotypes can be observed, supported by WGS data among isolates over the infection period of CDI and recurrence. A focus on three patient case studies will be presented individually, whilst the remaining data for 37 isolates of 17 patients will be presented as collated data sets.

4.2 Results

4.2.1 Spore phenotypes of isolates from patients 01008, 11007 and 03008 treated with rifaximin or placebo for the prevention of CDI, presented as case studies 1, 2 and 3

Case study 1, includes isolates from patient 01008, selected for data presentation as a recurrent case study of CDI, where rifaximin was administered as the chaser therapy. This patient was already noted to have an ongoing CDI infection before trial enrolment, with strains belonging to PCR ribotypes 002 and 004. Only four isolates obtained after completion of routine antibiotic treatment marking the start of trial enrolment will be presented in this study (see Table 4.1). During the period of enrolment, both relapse and reinfection events were revealed during microbiological analysis. Two isolates, wk-1 (E2) and Wk-1 (E10), belonging to PCR ribotype 002 and 014, respectively, were obtained from a diagnostic sample submitted at the beginning of enrolment. Whilst two isolates, Wk12 (G2) and RL (H4), belonging to PCR ribotype 002 and 014, respectively, were obtained from week 12 and relapse submitted samples, respectively.

Case study 2, includes isolates from patient 11007, selected for data presentation as a recurrent case study of CDI, where placebo was administered as the chaser therapy. This patient was already noted to have a previous CDI infection before trial enrolment, with a strain belonging to ribotype 029. (see Table 4.1). Two isolates, wk-1 C2 and Wk-1 C5, belonging to ribotype 014 and 029, respectively, were obtained during a reinfection event at wk-1. Whilst two isolates, Wk0 and RL, belonging to ribotype 014 were obtained from week 0 and relapse submitted samples, respectively. And so, for this patient relapse and reinfection events were revealed during the period of enrolment.

Case study 3, includes isolates from patient 03008, selected for data presentation as a non-recurrent case study of CDI. During the period of enrolment, rifaximin was administered as the chaser therapy. Despite the detection of *C. difficile* from stool post treatment at week 4, no carriage of *C. difficile* was observed beyond this sampling point, additionally, no recurrent

events were revealed during the trial period. The spore phenotype analysis of two isolates, Wk0 and Wk4, belonging to ribotype 002 are presented in this chapter (see Table 4.1).

Table 4.1: Case studies 1 – 3 of isolates from patient 01008, 11007 and 03008, respectively, treated with rifaximin or placebo chaser therapy for the prevention of recurrent CDI

Case study	Patient	Treatment	Strain ID	Sample Date [a]	PCR Ribotype	Microbiological infection status
1	01008	Rifaximin	WK-1 (E2)	21/06/2013 (27/07/2013)	(CE) 002	Co-infection
			WK-1 (E10)	21/06/2013 (27/07/2013)	(CE)014	1st Relapse (Co-infection)
			WK12 (G2)	02/10/2013	(CE) 002	Persistent carriage
			RL (H4)	04/01/2014	(CE)014	2nd Relapse
2	11007	Placebo	WK-1 (C2)	08/09/2014	(CE)014	Co-infection
			WK-1 (C5)	08/09/2014	(CE)029	Co-infection
			WK0	08/09/2014	(CE)014	Asymptomatic Reinfection
			RL	11/09/2014	(CE)014	Relapse
3	03008	Rifaximin	WK0	18/11/2013	(CE) 002	Symptomatic carriage
			WK4	13/07/2013	(CE) 002	Carriage

[a] Dates in brackets indicate the date of specimen arrival to the laboratory

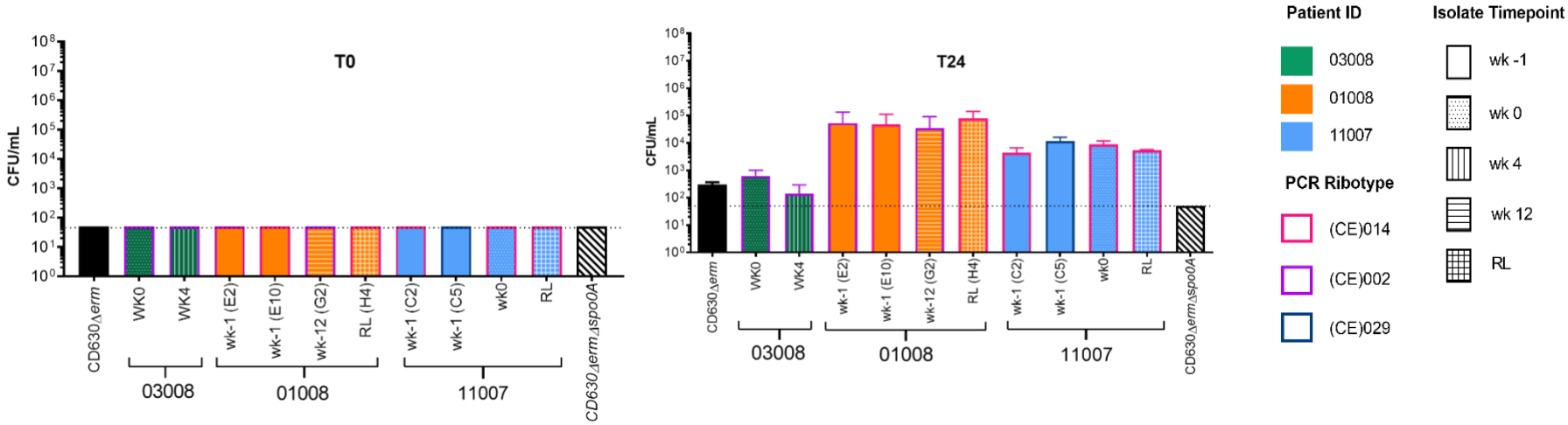


Figure 4.1: Spore production after 24 hours is increased for clinal isolates of case studies 1-3 compared to laboratory control strains.

Sporulation rates of isolates from patients 01008, 11007 and 03008, as measured by CFU formation on BHIS agar supplemented with 0.1% Tc after 0 hours (T0) and 24 hrs (T24) incubation of 0.1% starter culture of *C. difficile* cells in BHIS and 30 min heat treatment at 65 °C. Positive and negative control strains for this assay include lab strain CD630Δ*erm* and CD630Δ*erm*Δspo0A, respectively. The detection limit was 50 CFU/mL, as illustrated by the black dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the mean. See right of graph a key for patient ID (fill), PCR ribotype (border) and timepoint (pattern).

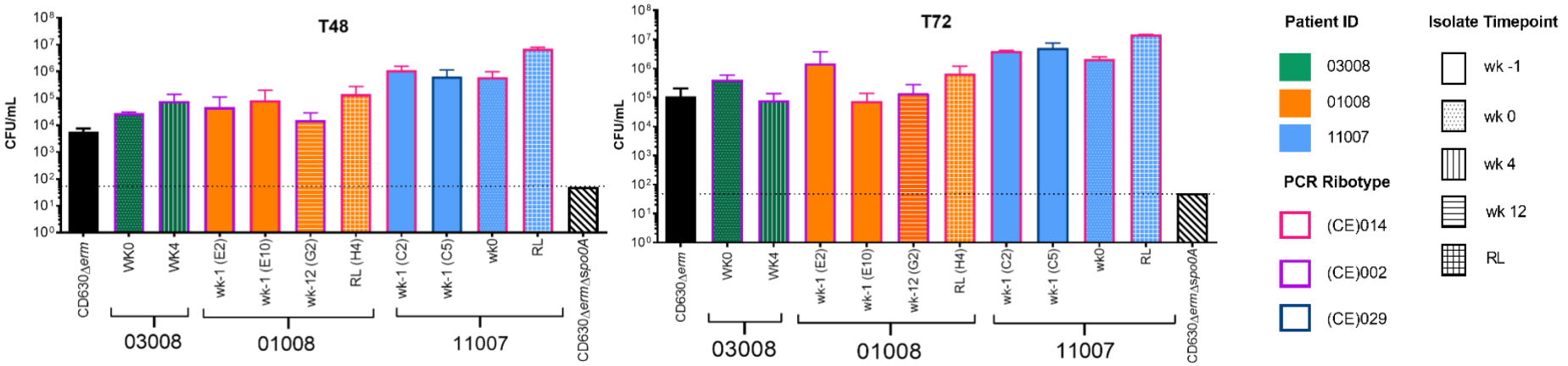


Figure 4.2: Spore production after 48 and 72 hours for clinal isolates of case studies 1-3.

Sporulation rates of isolates from patients 01008, 11007 and 03008, as measured by CFU formation on BHIS agar supplemented with 0.1% Tc after 48 hours (T48) and 72 hrs (T72) incubation of 0.1% starter culture of *C. difficile* cells in BHIS and 30 min heat treatment at 65 °C. Positive and negative control strains for this assay include lab strain CD630Δ*erm* and CD630Δ*erm*Δ*spo0A*, respectively. The detection limit was 50 CFU/mL, as illustrated by the black dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the mean. See right of graph a key for patient ID (fill), PCR ribotype (border) and timepoint (pattern).

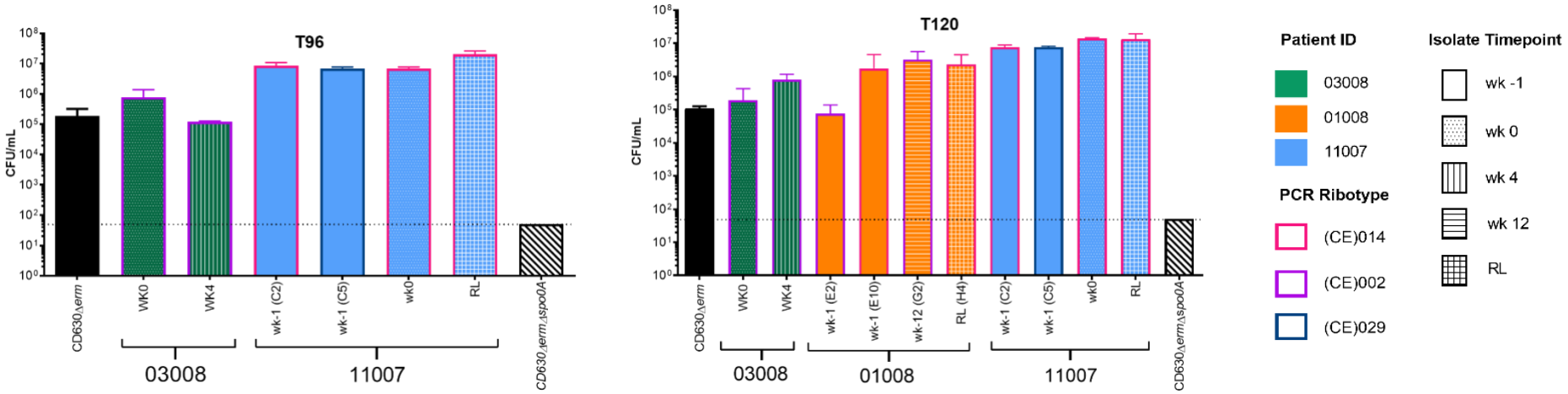


Figure 4.3: Spore production after 96 and 120 hours for clinal isolates of case studies 1-3.

Sporulation rates of isolates from patients 01008, 11007 and 03008, as measured by CFU formation on BHIS agar supplemented with 0.1% Tc after 96 hours (T96) and 120 hrs (T120) incubation of 0.1% starter culture of *C. difficile* cells in BHIS and 30 min heat treatment at 65 °C. Positive and negative control strains for this assay include lab strain CD630Δ*erm* and CD630Δ*erm*Δ*spo0A*, respectively. The detection limit was 50 CFU/mL, as illustrated by the black dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the mean. See right of graph a key for patient ID (fill), PCR ribotype (border) and timepoint (pattern).

The isolates were tested to determine their ability to form heat-resistant colonies over a 5-day period on BHIS agar supplemented with the bile salt taurocholate (see Figure 4.1- Figure 4.3). A heat-resistant colony-forming unit (HR-CFU) represents a spore that has germinated and returned to vegetative cell growth. The HR-CFU development of all recurrence patient isolate's form 2 logs more of CFU/mL compared to the control strains and non-recurrence isolates by 24 hours. However, no statically significant difference was observed using a one-way ANOVA with a Dunnett multiple comparison test. All isolates continue develop an increased level CFU/mL through all time points, apart from time points at 48 and 72 hours were isolate 01008 wk-1 (E10) and week 12 (G2), respectively, that diverge from this trend. The initiation of spore germination for isolates for case studies 1-3, was also investigated (see Figure 4.4), as measured by a drop in OD_{600} over 20min as described in section 2.4.7.

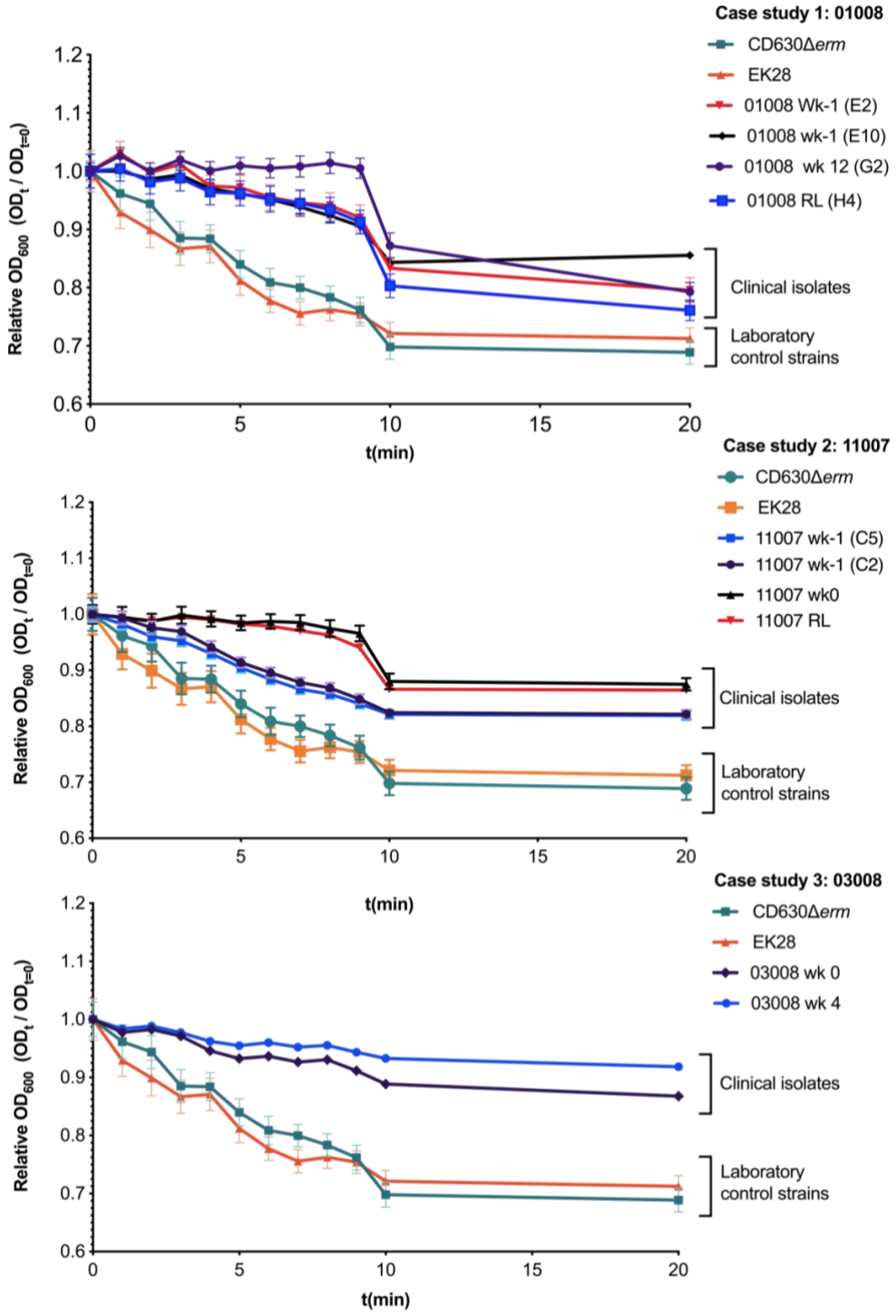


Figure 4.4: Initiation of spore germination of clinical isolates of case studies 1- 3 is significanly slower to that of laboratory control strains.

Initiation of germination rates of isolates from patients 01008, 11007 and 03008 as measured by a drop in OD600 over 20 minute incubation of pure spores with BHIS and 0.1% Tc. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Positive control strains for this assay include lab strain CD630Δ*erm* and clinical isolate EK28.

Table 4.2: Statistics of the slope from 0 to 9 minutes of the initiation of germination for isolates of Case studies 1-3

Isolate	Slope	Standard error of slope
CD630Δ*erm*	-0.02655	0.00155
EK28	-0.02683	0.00262
01008 Wk-1 (E2)	-0.01076	0.00148
01008 wk-1 (E10)	-0.01068	0.00086
01008 wk 12 (G2)	-0.00021	0.00103
01008 RL (H4)	-0.00952	0.00074
11007 wk-1 (C2)	-0.01804	0.00078
11007 wk-1 (C5)	-0.01826	0.00056
11007 wk 0	-0.00302	0.00059
11007 RL	-0.00544	0.00086
03008 wk 0	-0.00923	0.00098
03008 wk 4	-0.83418	0.00071

The initiation of spore germination of all clinical isolates from case studies 1-3 (Figure 4.4) were highly significant compared to control strain CD630Δ*erm* according to an uncorrected Dunn's test; 01008 wk-1 (E2) P = 0.0013, 01008 wk-1 (E10) $P = 0.0001$, 01008 wk12 (G2) $P \leq 0.0001$, 01008 RL (H4) P= 0.0159, 11007 wk-1 (C2) $P = 0.0072$, 11007 wk-1 (C5) $P = 0.0344$, 11007 wk0 $P < 0.0001$ and 11007 RL $P < 0.0001$, 03008 wk0 $P = 0.0003$ and 03008 wk 4 $P \leq 0.0001$. Whilst control strains, CD630Δ*erm* and EK28, showed no statistical significance difference to each other. Linear regression analysis of the slope for 0 to 9 mins was used to formulate an ordinary one-way ANOVA using Turkey's multiple comparison test to observe significant difference between rate of germination for isolates. Isolates of all patients have a slower rate of drop in OD_{600} shown by the smaller slope from 0 to 9 minutes, with a highly

significant difference of $P \leq 0.0001$ compared to control strains CD630Δ*erm* and EK28 indicating a slower rate of germination initiation (see Table 4.2). During the 9 to 10 minute time point there is a steeper drop in OD for all control strains and patient isolates compared to rest of the assay, indicating a collective germination event of a larger proportion of spores. Additionally, all isolates of patients 01008, 11007, 03008 have a higher endpoint than control strains indicating that a proportion of spores have not initiated germination.

Isolates 01008 wk-1 (E2) and 01008 wk 12 (G2) were indicated to be the same strain, with only a 6 SNP difference from WGS analysis in Chapter 3. Interestingly, initiation of germination for the week 12 isolate proceeds much slower compared to its counterpart at diagnosis, as illustrated by a highly significant and smaller slope (see Table 4.2), $P = 0.0001$, and drop in OD_{600} for this isolate not appearing until after 9 minutes initiation in this assay. This could relate to the persistence of this isolate over the course of infection, with spores persisting during the routine course of treatment and rifaximin chaser therapy**.** The relapsing isolates wk-1 (E10) and RL (H4) share a similar profile for initiation of germination where drops in OD_{600} are near identical, only diverging in the end-point reading, with no statistical difference shown between slopes. Here the 2nd relapse isolate RL (H4) continues to germinate at the 20min time point as indicated by the continued drop in OD, much like the RT002 isolates. However, the first relapse strain, 01008 wk-1 (E10) shows the highest endpoint for this assay, suggestive of incomplete germination for a higher proportion of spores than the other isolates of this patient. With ahighly significant difference shown for isolates 01008 wk-1 (E10) vs. 01008 wk 12 (G2) ($P = 0.0002$), and 01008 wk 12 (G2) vs. 01008 RL (H4) ($P = 0.001$), respectively.

The diagnostic isolates of patient 11007, wk-1 C2 and C5, despite being identified as two completely different strains with different ribotypes, shared a similar profile for initiation of germination where drops in OD_{600} are near identical, with no statistically significant difference shown between them. Rates here proceed at a relatively quicker decline, shown by larger slopes compared to their later counterpart isolates from week 0 and RL samples (see Table 4.2), where a highly significant difference is observed by a multi comparison test, $P < 0.0001$. Even though isolates 11007 wk-1 (C2) and 11007 wk 0 were indicated to be the same strain from the SNP analysis in Chapter 3, germination profiles indicated from this assay are not shared. Interestingly, initiation of germination for the post treatment isolates (wk 0 and RL)

share near identical profiles and are observed to proceed much more slowly compared to the counterpart at diagnosis, with drops in OD_{600} not appearing until after 9 minutes initiation in this assay. Similarly, post diagnostic isolates show the highest endpoint for this assay suggestive of incomplete germination for a higher proportion of spores than the other isolates of this patient.

Isolates 03008 wk0 and 03008 wk4 were indicated to be the same strain from the SNP analysis by Dr Partick Budd (see Table 7.3 of the appendix). Interestingly, initiation of germination for the wk 4 isolate proceeds more slowly compared to its earlier corresponding isolate, as shown by the significantly smaller slope (see Table 4.2). This could relate to the carriage of this isolate over the course of infection, with spores persisting during the routine course of treatment and rifaximin chaser therapy. The week 4 shows the highest endpoint for this assay suggestive of incomplete germination for a higher proportion of spores than the other isolate of this patient. This could reason why the isolate is carried throughout anti-CDI therapy for this patient.

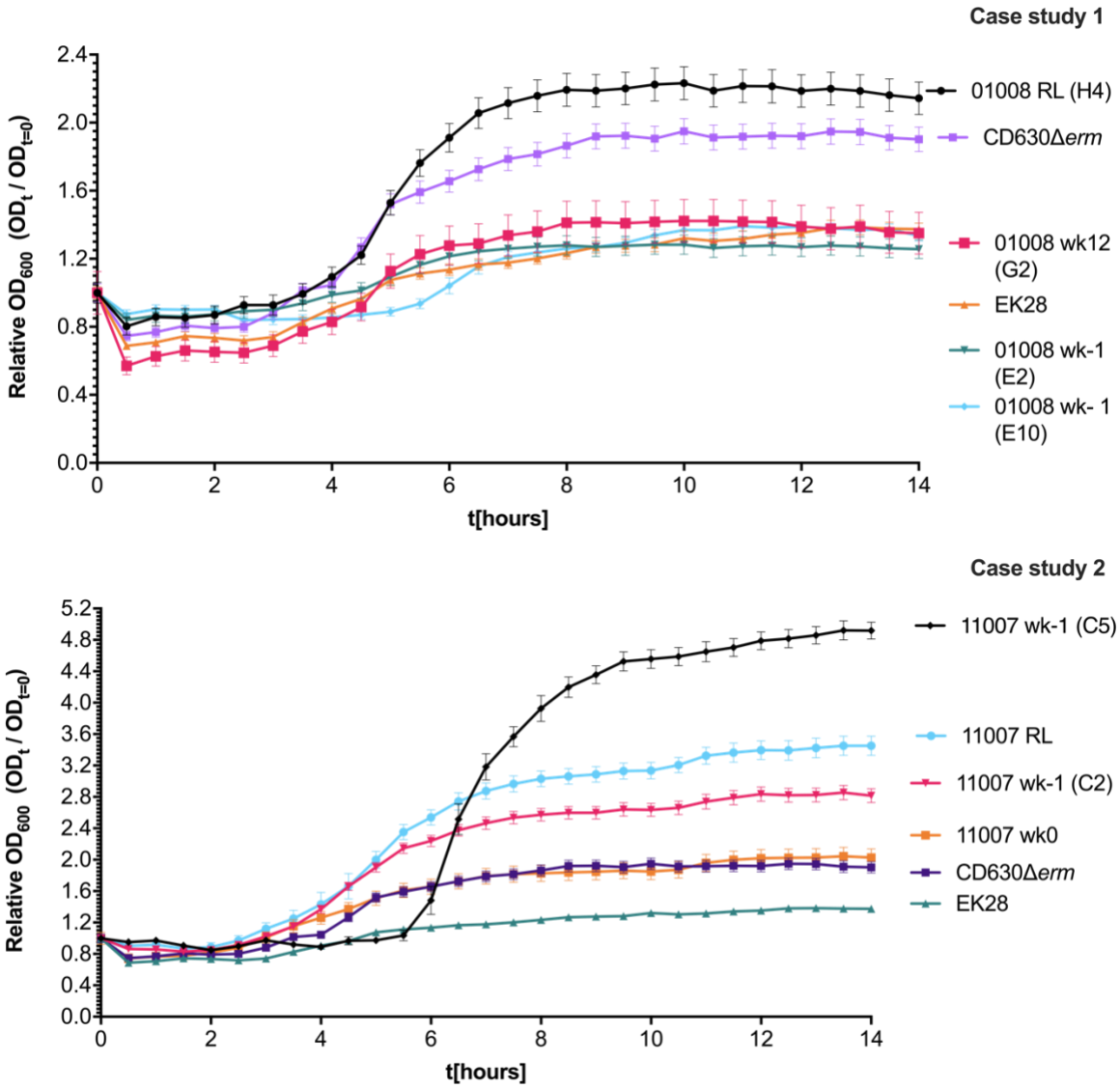


Figure 4.5: Spore outgrowth of isolates of case studies 1 and 2.

Spore outgrowth of isolates from patient 01008 and 11007, as measured by a rise in relative OD600 over 14 hrs incubation with BHIS and 0.1% Tc. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Positive control strains for this assay include lab strain CD630Δ*erm* and clinical isolate EK28.

The outgrowth from spores for isolates of patient 01008 and 11007 is presented in Figure 4.5. Isolate wk-1 (E10), from patient 01008, has the longest lag phase compared to other isolates and controls in this experiment indicating the spore outgrowth is much slower than control strains and other isolates from this patient, which enter into exponential phase approximately 3 hours earlier. The co-infecting isolate at this diagnostic enrolment period, wk-1 (E2), had a smaller slope during exponential phase compared to other later isolates, followed by isolate wk-1 (E10), m = 0.0882 and m = 0.0109, respectively. With a highly significant difference found for diagnostic isolates, wk-1 (E10) and wk -1 (E2) (P = 0.0002), and for that of wk 12 (G2), m = 0.1753 (P = 0.001), respectively. This is indicative of a smaller proportion of spores being able to outgrow compared to the later isolates. Isolate 01008 RL (H4) had the largest slope of the exponential phase in this experiment, m = 0.3833, with a highly significant difference shown to isolate week 12 (G2), m = 0.1753, P = 0.0010, as well as the highest end point of exponential phase compared to other isolates from this patient.

Isolate wk-1 (C5), from patient 11007, has the longest lag phase in this experiment indicating the spore outgrowth is significantly slower than control strains and other isolates from this patient, which enter into exponential phase 3-4 hours earlier. However, this isolate also has the largest slope of the exponential phase, m = 0.9873, indicating a faster rate of growth of vegetative cells than all case study isolates tested. The co-infecting isolate at this diagnostic enrolment period, wk-1 (C2) had a highly significant difference and smaller slope during exponential phase, m = 0.3498, $P < 0.001$, which is indicative of a smaller proportion of spores being able to outgrow and a competitive level of fitness. Isolates 11007 wk-1 (C2) and 11007 RL share similar lag phase, and comparable slopes for the exponential phase, m = 0.3498 and m = 0.4265, respectively, with no statistically significant difference found between the two, indicating similar rates and proportion of total spore outgrowth. However, C2 enters log phase before C5 with a lower end point. Isolate 11007 wk-1 (C5) has the highest end point of exponential phase, followed by RL, wk-1 (C2) and wk0 isolates.

To assess whether this hypothesised persistence and so dormancy of spores for case studies 1 and 2 could be observed by other methods, a loss of heat resistance assay was employed, as described in section 2.4.9, to measure the number of spores that persist despite incubating in a germinant environment over 4 hours (see Figure 4.6).

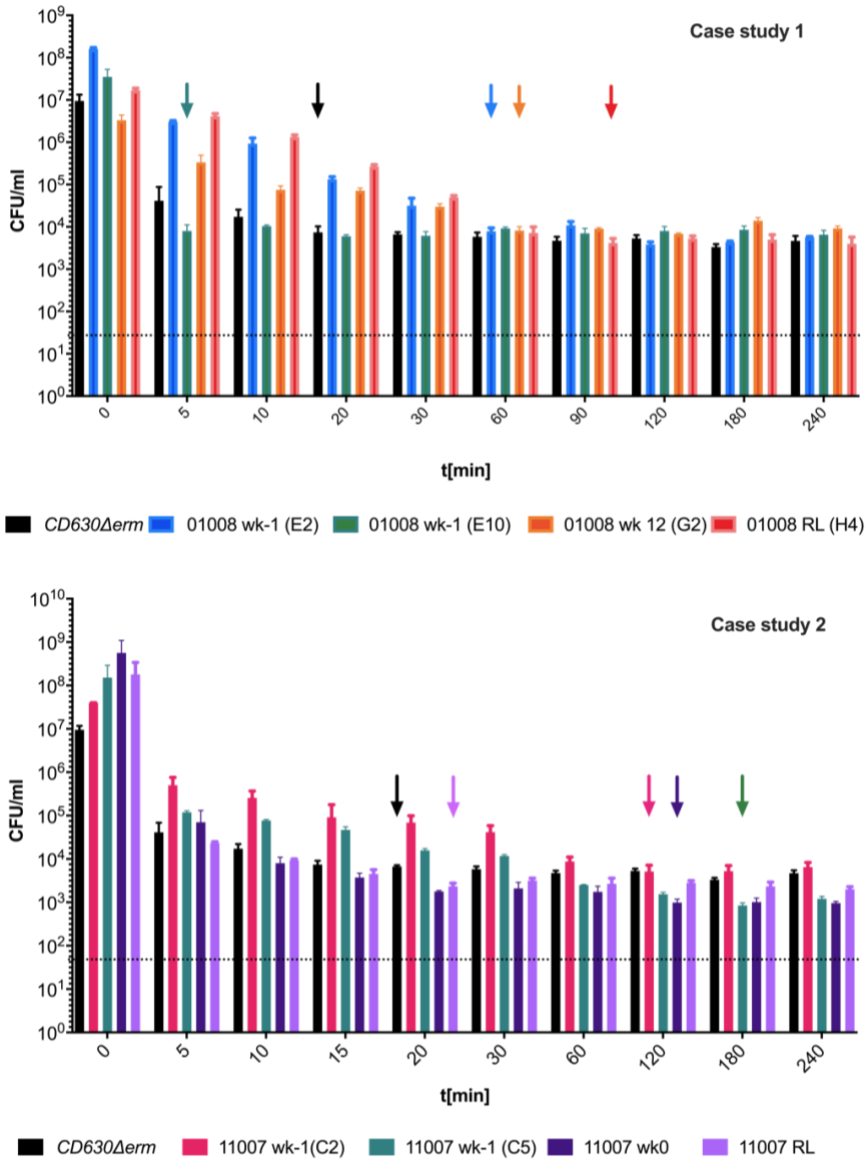


Figure 4.6: Loss of spore heat-resistance for isolates from case studies 1 and 2.

Loss of spore heat resistance of isolates from patient 01008 and 11007. Spore suspensions were incubated with germinant (0.1% Tc in BHIS broth) for permitted time points, where they were heat treated to kill any germinated or outgrown cells. Any HR spores remaining at this time point were allowed to germinate and out growth on BHIS agar supplement with 0.1% Tc with 24 hrs where CFU/mL was determined. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Positive control strain used for this assay includes lab strain CD630Δ*erm*. The detection limit was 50 CFU/mL, as illustrated by the dotted line. Red arrows indicate the estimated end of spore germination events observed for isolates during this assay as compared to CD630Δ*erm* as indicated by the black arrow.

The continued measurement of detectable heat resistance CFU in Figure 4.6 - case study 1 throughout the duration of this assay illustrates a level of spore dormancy across all isolates of case studies as well as the control strain, CD630Δ*erm*. For case study 1, it was observed that isolate 01008 wk-1 (E10) enters spore dormancy earlier than all isolates of this patient, as well as those of case study 2 and the control strain, CD630Δ*erm*, with no substantial losses of heat resistance observed beyond 5 minutes. This indicates that all the spores that have germinated, and of those that can efficiently outgrow, have done so within 5 mins, supporting the observation in outgrowth phenotype in Figure 4.5, where it had the longest lag phase and smallest slope during exponential phase. This collectively indicates that a smaller number of spores outgrew compared to other isolates, conclusively attributing this phenotype to spore outgrowth and not vegetative cell growth. The remaining isolates of case study 1 entered substantial spore dormancy by 60 minutes in this assay, with a detectable endpoint of ~10^4 CFU/mL. For patient 11007, the RL isolate enter spore dormancy by 20 minutes, which is earlier than that of diagnostic and treatment period isolates, wk-1 (C2), wk-1 (C5) and wk 0, where no substantial losses of heat resistance observed beyond 180 minutes, (see Figure 4.6 - case study 2). At the end of the assay all isolates of cases studies 1 and 2 and the control strain had a detectable endpoint of <10^4 CFU/mL, indicating a non-germination fate of spores for a significant proportion of spores in this assay.

4.2.2 Collated sporulation phenotypes of 37 isolates from 17 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence

All isolates obtained in this trial were tested to determine their ability to form heat-resistant colonies over a 5-day period on BHIS agar supplemented with the bile salt taurocholate. Data from 37 isolates are presented in this section (see Figure 4.7 and Figure 4.8). At 0 hours incubation all isolates begin at the expected baseline of 50 CFU/mL at the limit of detection, as indicated by the blue dashed line on figure, except for isolates 12004 wk0, 12004 RL and 18001 wk0 (see Figure 4.7 0h), with wk0 isolates (P = 0.0293 and P = 0.0218) shown to be significant and the relapse strain, 12004 RL, shown to have a highly significant difference (P < 0.0001) compared to control strain CD630Δ*erm*Δ*spo0A* at T0, according to an uncorrected Dunn's test. For these isolates, belonging to PCR ribotype CE020, assays were again repeated in triplicate with the same observation at T0. Such findings indicate that these isolates are able to enter into sporulation much earlier than all others tested in this assay.

By 24 hours, all isolates produce detectable CFU, as with the wild-type control strain CD630Δ*erm* (see Figure 4.7 24h). Isolates 01017 wk4, 01016 wk12 produce comparable CFU's to wildtype control, whilst the remaining isolates produce over 1e+4 CFU/mL by this time point. The average CFU/mL for non-recurrence isolates is 9.6e+4, whilst it is 6e+4 for the recurrent cohort as indicated by the solid and dashed red lines respectively. PCR ribotypes CE023, CE011 and CE020 highest level of CFU/mL on average across all isolates by 24 hours. By 48 hours, all isolates produce detectable CFU above 1e+3 CFU/mL, as with the wild-type control strain CD630Δ*erm* (see Figure 4.7 48h). The average CFU/mL for non-recurrence isolates is 2.54e+5, whilst it is 5.58e+5 for the recurrent cohort as indicated by the solid and dashed red lines respectively. This indicates that the recurrent isolates produced on average 3.04e+5 CFU/mL more than non-recurrent isolates by 48 hours. PCR ribotypes CE023, CE011, CE015 and CE020 highest level of CFU/mL on average across all isolates by 48 hours.

By 72 hours, all isolates produce detectable CFU above 1e+4 CFU/mL, as with the wild-type control strain CD630Δ*erm* (see Figure 4.8 72h). The average CFU/mL for non-recurrence isolates is 1.39e+6, whilst it is 1.76e+6 for the recurrent cohort as indicated by the solid and dashed red lines respectively. This indicates that the recurrent isolates produced on average 3.7e+5 CFU/mL more than non-recurrent isolates by 72 hours. PCR ribotypes CE005, CE011

and CE015 highest level of CFU/mL on average across all isolates by 72 hours. By 96 hours, almost all isolates produce detectable CFU above 1e+5 CFU/mL, as with the wild-type control strain CD630Δ*erm* (see Figure 4.8 96h). The average CFU/mL for non-recurrence isolates is 3.6e+6, whilst it is 4.7e+6 for the recurrent cohort, as indicated by the solid and dashed red lines, respectively, indicating that the recurrent isolates produced on average 1.1e+6 CFU/mL more than non-recurrent isolates. PCR ribotypes CE005, CE011 and CE015 highest level of CFU/mL on average across all isolates by 96 hours. By the end of the assay, at 120 hours, all isolates produce detectable CFU above 1e+5 CFU/mL, as with the wild-type control strain CD630Δ*erm* (see Figure 4.8 120h). The average CFU/mL for non-recurrence isolates is 3.67e+6, whilst it is 5.2e+6 for the recurrent cohort, as indicated by the solid and dashed red lines, respectively, indicating that the recurrent isolates produced on average 1.53e+6 CFU/mL more than non-recurrent isolates by 120 hours. PCR ribotypes CE005, CE011, CE070, CE015 and CE020 highest level of CFU/mL on average across all isolates by 120 hours. Overall, a statistically significant difference was observed for the average CFU/mL production by 120 hours when comparing the control, CD630Δ*erm*, to the recurrence isolates, $P = 0.0292$, but no statistically significant difference was observed between the average CFU/mL production of non-recurrence or recurrence isolates.

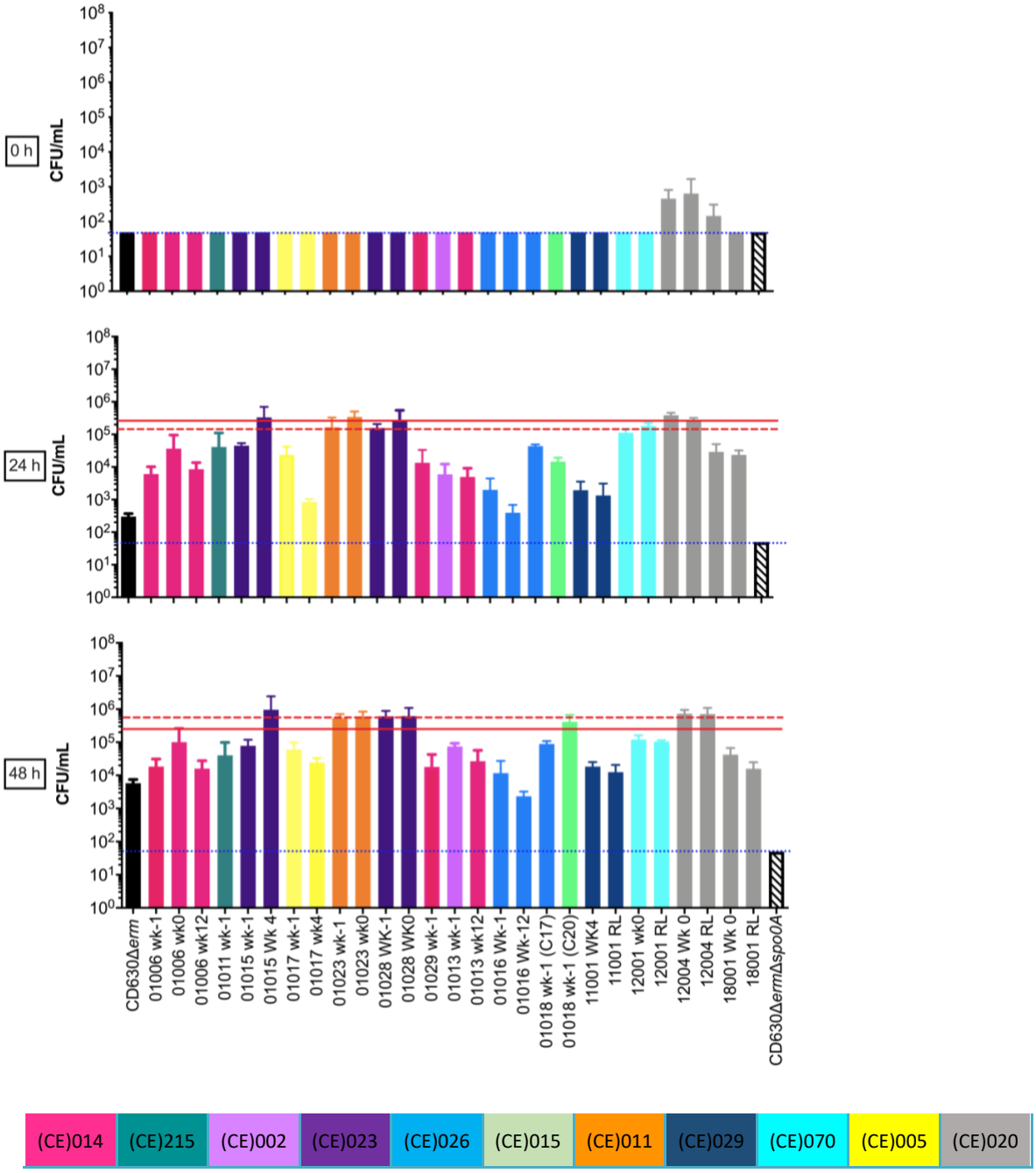


Figure 4.7: Sporulation rates at 0, 24 and 48 hrs for 37 isolates from patients enrolled of the trial of rifaximin chaser therapy for the prevention of CDI recurrence.

Sporulation rates as measured by CFU formation on BHIS agar supplemented with 0.1% Tc at 0, 24 and 48 hrs incubation of 0.1% starter culture of *C. difficile* cells in BHIS with 30 min at 65 °C. Positive and negative controls include CD630Δ*erm* and CD630Δ*erm*Δ*spo0A*, respectively. A 50 CFU/mL detection limit is illustrated by the blue dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means, with coloured bars indicating PCR ribotype. The average CFU/mL of the non-recurrent and recurrent isolates are indicated by the solid and dashed red line. See below graph a colour key of PCR ribotype.

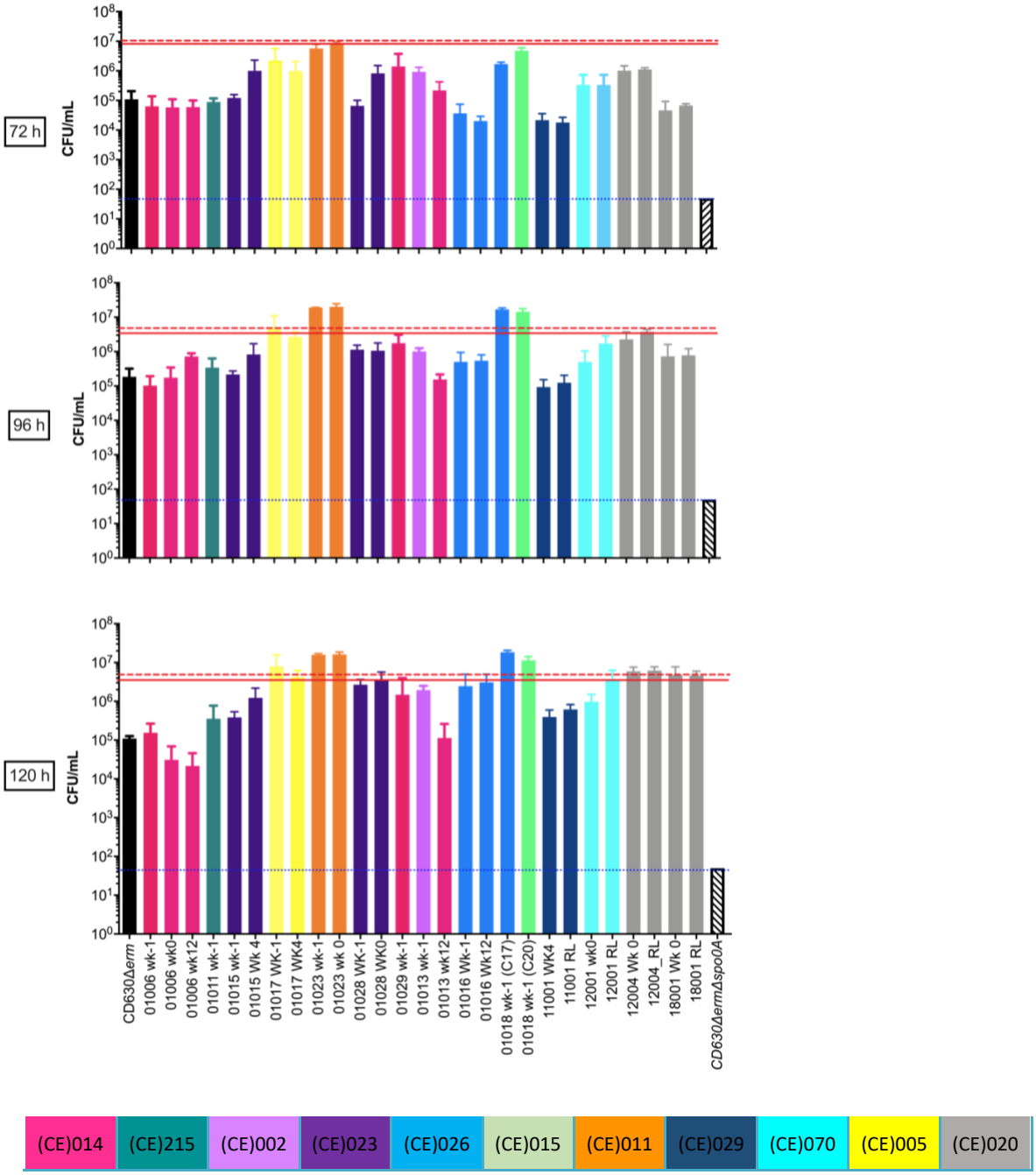


Figure 4.8: Sporulation rates after 72, 96 and 120 hrs for 37 isolates from patients enrolled of the trial of rifaximin chaser therapy for the prevention of CDI recurrence.

Sporulation rates as measured by CFU formation on BHIS agar supplemented with 0.1% Tc at 72, 96 and 120 hrs incubation of 0.1% starter culture of *C. difficile* cells in BHIS with 30 min at 65 °C. Positive and negative controls include CD630Δ*erm* and CD630Δ*erm*Δ*spo0A*, respectively. A 50 CFU/mL detection limit is illustrated by the blue dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means, with coloured bars indicating PCR ribotype. The average CFU/mL of the non-recurrent and recurrent isolates are indicated by the solid and dashed red line. See below graph a colour key of PCR ribotype.

4.2.3 Collated initiation of germination phenotypes of 20 isolates from 11 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence

A total of 20 additional isolates were analysed for rates of germination initiation (see Figure 4.9 and Figure 4.10, and of the appendix for slope statistics). During the 9 to 10 minute time point there is a larger drop in OD for all control strains and patient isolates compared to rest of the assay, indicating a collective germination event of a larger proportion of spores. Isolates 01013 wk-1 and 11001 RL were shown to have no statistically significant difference compared to CD630Δ*erm* using the uncorrected Dunn's test and have faster rates of germination compared to all other isolates, with a very high significant difference ($P < 0.0001$). All remaining isolates of patients tested have a slower rate of drop in OD_{600} shown by the smaller slope from 0 to 9 minutes, compared to control strains CD630Δ*erm* and EK28 indicating a slower rate of germination initiation, with very highly ($P < 0.001$) significant difference found according to an ordinary one-way ANOVA using Turkey's multiple comparison test. Additionally, all patient isolates tested have a higher endpoint than control strains indicating that a higher proportion of spores have not initiated germination.

This phenotype was observed for 3 out of 7 recurrence patients (see Figure 4.9), whilst this was not seen for the remaining patients studied in the non-recurrent cohort (see Figure 4.10). Admittedly, only two patients of the non-recurrent cohort, 01015 and 01028 (see Figure 4.10), had more than one isolate during this analysis, however both of these isolates shared near identical slopes for initiation of germination for those patients.

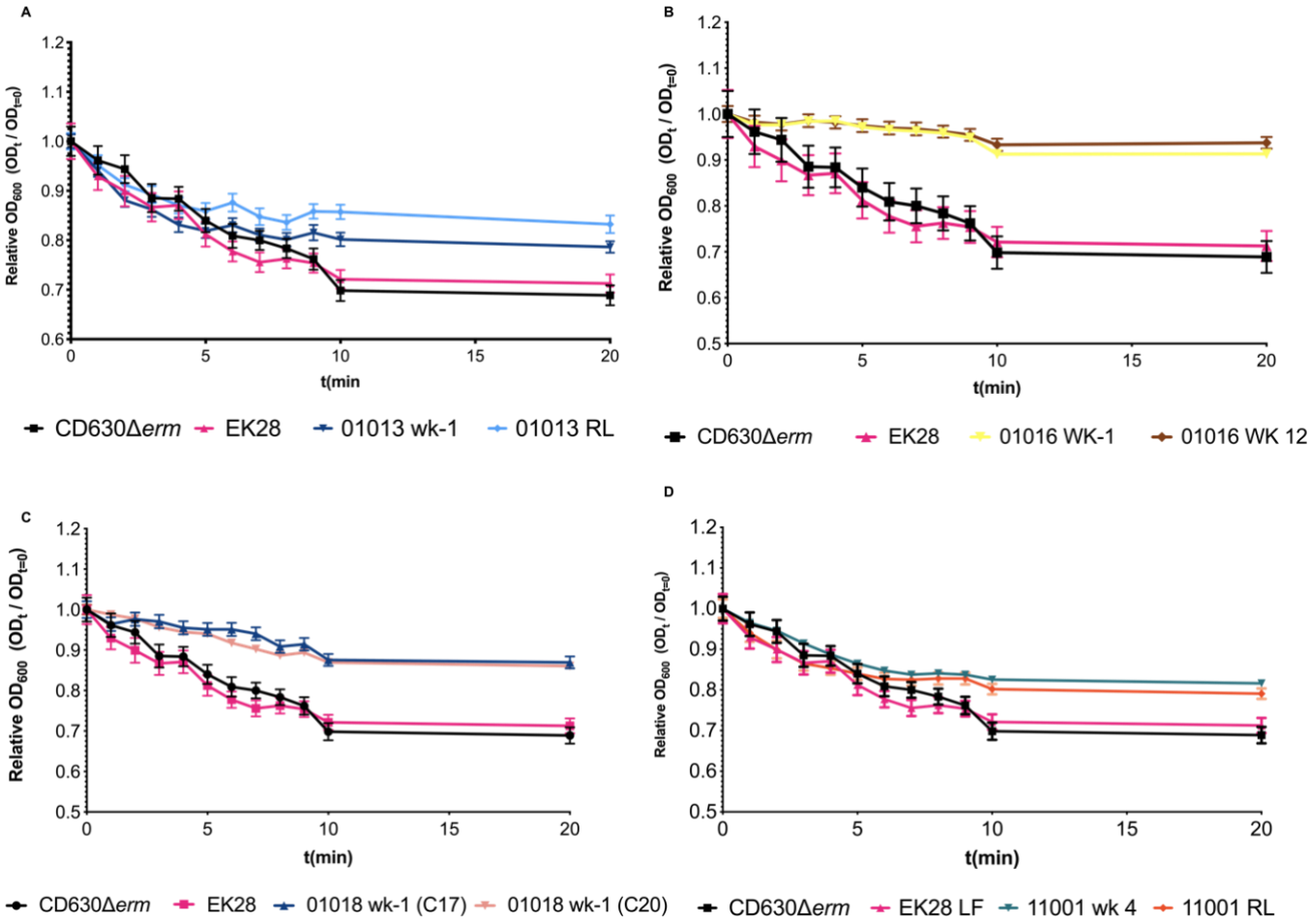


Figure 4.9: Initiation of spore germination rates in isolates from patients of the recurrence cohort.

Panels A-D: Initiation of germination rates of isolates four patients with recurrent or persistent CDI enrolled on the trial of rifaximin chaser therapy for the prevention of CDI recurrence. Initiation of germination rates were measured by a drop in OD_{600} over 20 minute incubation with BHIS and 0.1% Tc. Positive control strains for this assay include lab strain CD630Δ*erm* and clinical isolate EK28. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Isolates that were statistically significant compared to control strain CD630Δ*erm* according to an uncorrected Dunn's test include; 01013 RL and 01018 wk-1 (C20) $P = 0.0002$, 01016 wk-1, 01016 wk12, 01018 wk-1 (C17), $P < 0.0001$ and 11001 wk4 $P = 0.0049$. Control strains showed no statistical significance difference to each other.

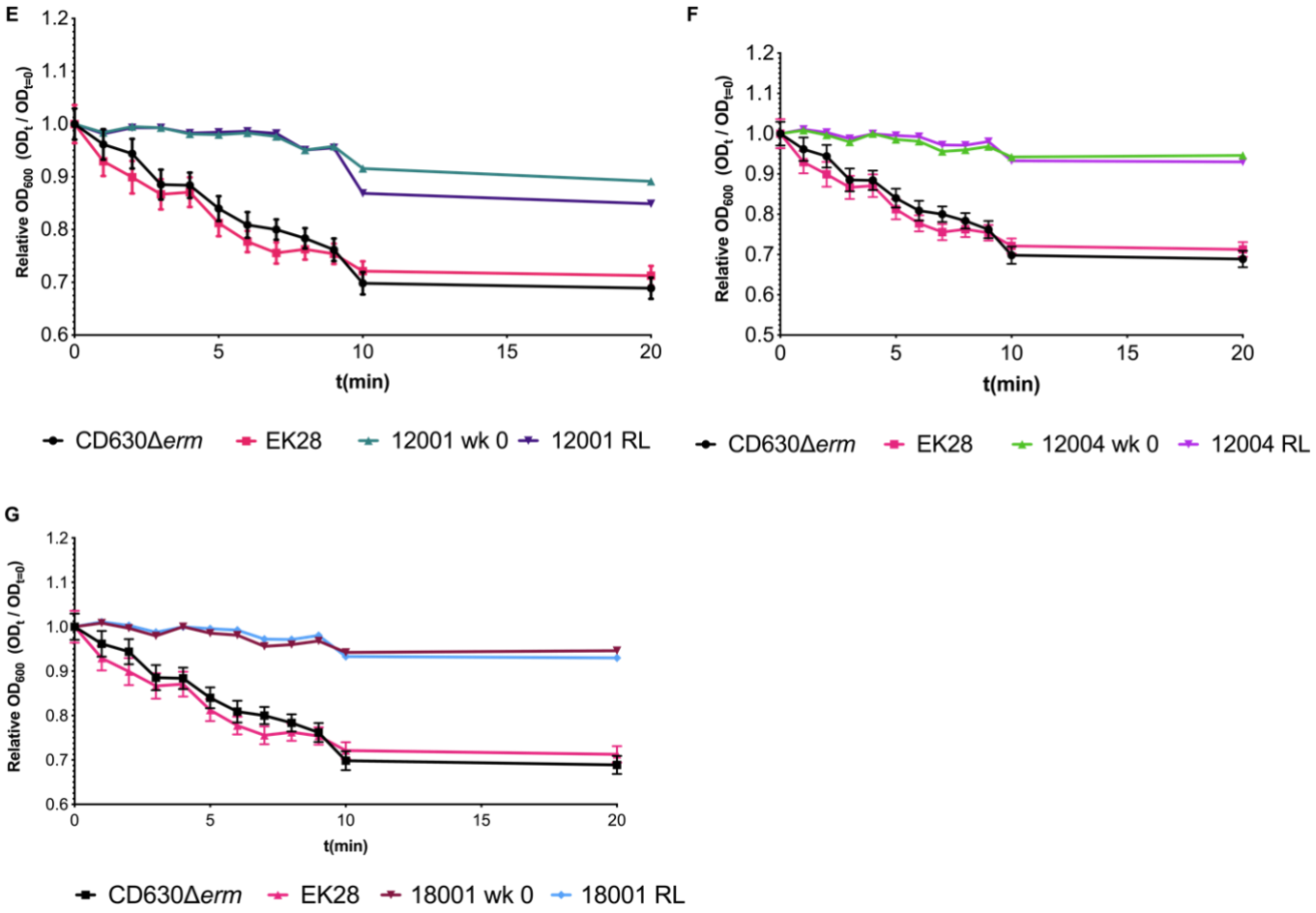


Figure 4.9 Continued: Initiation of spore germination rates in isolates from patients of the recurrence cohort.

Panels E-G: Initiation of germination rates of isolates three patients with recurrent CDI enrolled on the trial of rifaximin chaser therapy for the prevention of CDI recurrence. Initiation of germination rates were measured by a drop in OD600 over 20 minute incubation of pure spores with BHIS and 0.1% Tc. Positive control strains for this assay include lab strain CD630Δ*erm* and clinical isolate EK28. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Isolates that were statistically significant compared to control strain CD630Δ*erm* according to an uncorrected Dunn's test include; 12001 wk0, 12001 RL, 12004 wk0, 12004 RL, 18001 wk0 and 18001 RL $P < 0.0001$. Control strains showed no statistical significance difference to each other.

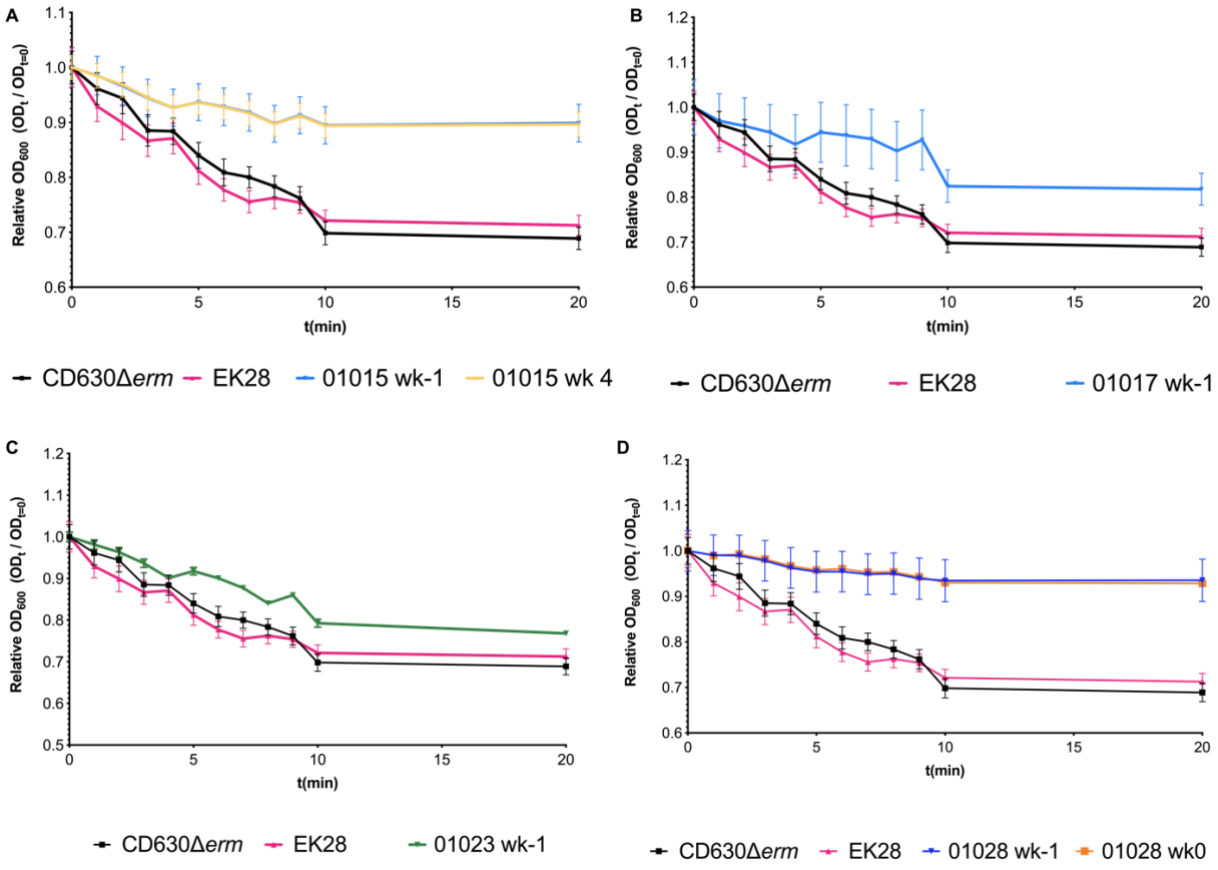


Figure 4.10: Initiation of spore germination rates in isolates of four patients with non-recurrent CDI.

A-D: Initiation of spore germination rates of isolates from 4 patients with non-recurrent CDI enrolled on the trial of rifaximin chaser therapy for the prevention of CDI recurrence. Initiation of germination rates were measured by a drop in OD_{600} over 20 minute incubation of pure spores with BHIS and 0.1% Tc. Positive control strains for this assay include lab strain CD630Δ*erm* and clinical isolate EK28. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Isolates that were statistically significant compared to control strain CD630Δerm according to an uncorrected Dunn's test include; 01015 wk-1 $P = 0.0001$, 01015 wk4 $P = 0.0002$, 01017 wk-1 $P = 0.0011$, 01023 $P = 0.0006$, 01028 wk-1 and 01028 wk0 $P < 0.0001$. Control strains showed no statistical significance difference to each other

4.2.4 Collated outgrowth phenotypes of 12 isolates from 6 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence

Figure 4.11 shows the rate of spore outgrowth from 12 isolates of 6 patients with recurrent CDI. All isolates, excluding 01018 wk-1 (C20) had lag phases longer than the control strains and did not enter the exponential phase until after 4-5 hours in this experiment, indicating that these spores outgrow much slower in comparison to the control strains. Isolate 01018 wk-1 (C20) had the shortest lag phase of around 2 hours, indicative of relatively rapid spore outgrowth compared to all isolates and control strains tested (see Figure 4.10 - B).

All other isolate pairs from the same patient had comparable outgrowth curves, only diverging at the end point OD. Here, the common trend being the 'relapse' (RL) or later isolate reaching a lower OD compared to their corresponding earlier isolates. This was however, contradicted by the two case studies of patients 01008 and 11007 (see Figure 4.5), where the opposite was true.

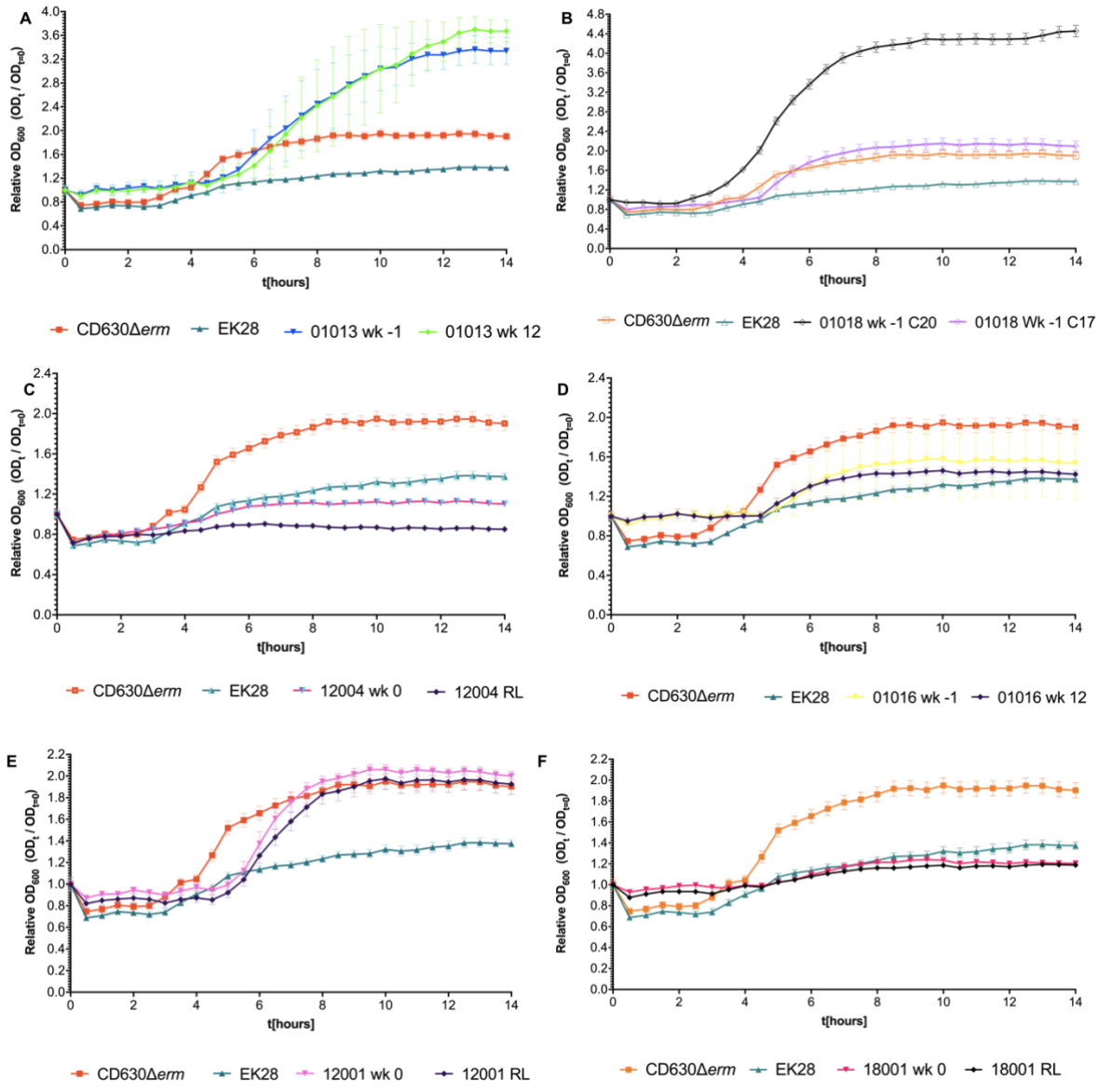


Figure 4.11: Spore outgrowth in in isolates from patients of the recurrence cohort.

A-F: Spore outgrowth for isolates from 6 patients with recurrent CDI enrolled on the trial of rifaximin chaser therapy for the prevention of CDI recurrence as measured by a drop in OD_{600} over 14 hrs incubation with BHIS and 0.1% Tc. as measured by a rise in OD_{600} over 14 hrs incubation with BHIS and 0.1% Tc. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. Positive control strains for this assay include lab strain CD630Δerm and clinical isolate EK28. Control strain EK28 had a significant difference to CD630Δ*erm*, $P = 0.0336$.

4.2.5 Collated spore dormancy phenotypes of 14 isolates from 7 patients enrolled on the trial of rifaximin therapy for the prevention of CDI recurrence

To assess whether this hypothesised dormancy of spores could be observed for the remaining isolates of the recurrence cohort, as seen in case studies 1 and 2 (see Figure 4.6), loss of heat resistance was assessed (see Figure 4.12). As with the case studies there is a continued measurement of detectable heat resistance CFU throughout the duration of this assay illustrating a level of spore dormancy across all isolates. For isolate 01013 wk-1, loss of heat resistance seems to end by the 5 minute time point, followed by the corresponding week 12 isolate ending by 10 minutes (see Figure 4.12 - A). This observation suggests that these isolates enter into spore dormancy earlier than the control strain and of all clinical isolates tested in this assay.

The remaining isolates of this data set entered substantial spore dormancy by 60 minutes in this assay, with a detectable endpoint of ~10^4 CFU/mL. Interestingly, no complete loss of heat resistance was observed for the period allowed in this assay for any isolates tested.

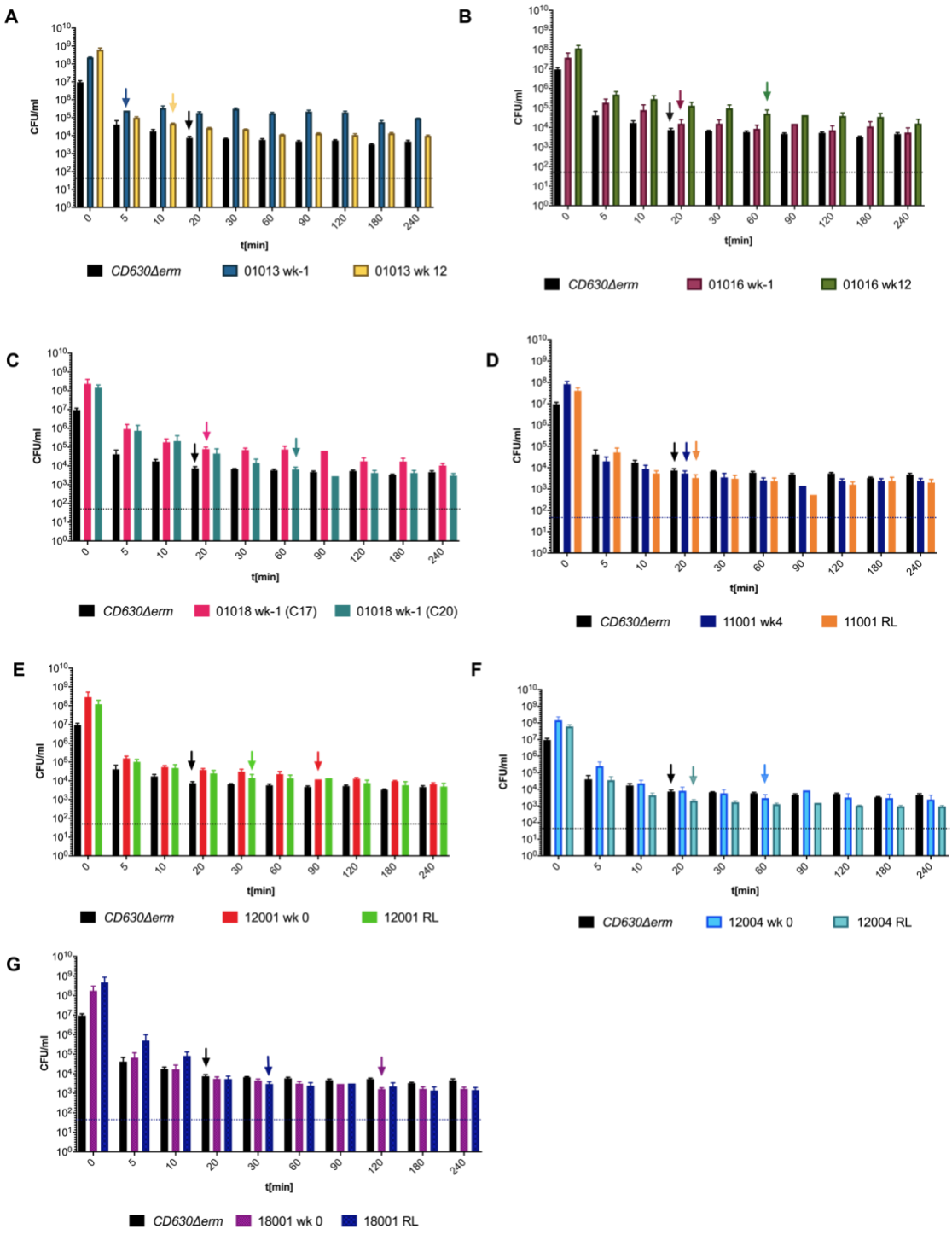


Figure 4.12: Loss of heat resistance in isolates from patients of the recurrence cohort.

A-G: Loss of spore heat-resistance in isolates from 7 patients enrolled on the trial of rifaximin chaser therapy for the prevention of CDI recurrence. Loss of heat resistance was measured by HR-CFU formation of spores remaining after incubations with BHIS + 0.1% Tc for up to 240 minutes, coloured arrows indicate the estimated end of spore germination events observed for isolates during this assay as compared to CD630Δ*erm* as indicated by the black arrow.

4.3 Discussion

4.3.1 Sporulation in recurrent and non-recurrent CDI

At 0 hours incubation isolates 12004 wk0, 12004 RL and 18001 wk0, belonging to PCR ribotype CE020, were able to produce detectable HR-CFU much earlier than all others tested in this assay, which remained at the limit of detection. Interestingly, isolates belonging to PCR ribotype CE011 produced, on average, one of the highest levels of CFU/mL across the 24- 120 hr time points of the assay.

The collated rate of sporulation data showed all isolates having a minimum serial increase of 1-log in CFU/mL at each time point over the 120hr assay (see Figure 4.7 and Figure 4.8). Apart from the 72hr time point (Figure 4.8 – 72h), the average CFU/mL showed a consistent increased average CFU/mL development for the recurrent isolates compared to the non-recurrent isolates tested over the course of this assay. A finding that could indicate a faster rate of spore production. However, the increased ability of isolates to develop heat- resistant CFU, could be a consequence of several possible traits in either sporulation, germination or outgrowth. An increase in the observed colony formation following heat shock would suggest that the patient isolates are able to out-compete their counterparts in one or more of these processes. A spore titre would be needed to confirm whether this phenotype was a result of an increased production of spores for the recurrent isolates compared to non-recurrent isolates tested.

All isolates from the patients 01008, 11007 and 03008, presented as case studies, produced higher CFU/mL than the positive control strain in the rate of sporulation assays (see Figure 4.1 and Figure 4.2). This finding that could, again, imply an increased rate of sporulation. The ability to produce many spores would evidence the ability of the isolate to be transmitted from host to host, to the environment and indeed re-ingestion into the shedding host for re-infection, a phenotype that would be better determined by spore titres at assay endpoint. Whether vegetative cells can continue to enter the sporulation stage within the host as a

mechanism of evading antibiotic targeting has yet to be established. Such a phenomenon would exacerbate the risk of relapse by means of producing dormant spores that would result in persistence and possible recurrence following germination and outgrowth into toxin producing cells. Such a feature seems unlikely given the colonisation status and availability of nutrients during CDI dysbiosis, in addition to the anaerobic environment maintained in the gut promotive of vegetative cell survival.

Whilst there remains a significant importance of assessing sporulation phenotypes for infection control, this feature pertains to transmission potential. Even the effect of increased risk of ingestion and so reinfection via increased spore production, really depends on whether these spores go on to germinate to cause infection. Relapsing infection, however, relies on the ability of these spores to remain dormant and persist beyond treatment and then go on to cause disease which relies on germination and outgrowth into toxin producing vegetative cells. To conclude, whilst there seems to be a preliminary correlation between the rate of sporulation and recurrence, the assay itself was confounded by the fact that it encompasses processes of heat resistance, germination and outgrowth also. To confirm such an assumption a spore titre would be needed as well as a larger sample set of isolates.

4.3.1.1 Limitations and advancement of knowledge in the study of sporulation: Rates of sporulation cannot be assessed without spore titres

The assay employed here overcomes limitations of other methods such as the Oka *et al.,* (2012) study, in regards to extended time points, by quantifying HR-CFUs at 24 hour intervals, instead of using a ratio of vegetative cells to spores by microscopy after 5-days. Importantly, observations of the initiation of sporulation can be observed for isolates much earlier than 120 hours, as was the case at time point 0 of this assay for the isolates belonging to PCR ribotype CE020. PCR type associations can loosely be made for CE011, producing on average one of the highest levels of CFU/mL, however, only two isolates of this type where obtained

in this study and no statistical significance difference observed for type associations, much like that of Burns *et al.,* (2011) and Oka *et al.,* (2012). Unfortunately, this experiment was limited in separating the rate of sporulation phenotype from that of germination, outgrowth and viable and detectable colony formation, as spore titres for isolates was not examined. And so, this larger production of CFU/mL cannot be exclusively attributed to the sporulation phenotype.

All 16 recurrent cohort strains tested for genomic variation showed amino acid changes for the sporulation gene CD630_08570, *oppF,* known to support the initiation of *C. difficile* sporulation (Edwards *et al., 2*014). Higher average rates of sporulation on was observed for recurrent isolates from 48-120hrs compared to non-recurrent isolates. Whilst this was not seen as significantly different by statistical tests, low level production could amass to a potential reservoir of spores that would be able to evade non-sporicidal anti-CDI therapy. Unfortunately, genomic variation analysis was not performed for isolates of a non-recurrent cohort and so this observation cannot be conclusively associated to this gene. Further amino acid changes were observed for genes CD630_08530 (*oppB*), CD630_08560 (*oppD*) in isolates 12001 wk 0 and 12001 RL, whilst *oppF* was found to be truncated in strain 01008 wk-1 (E10). Studies have observed that mutants lacking oligopeptide permease genes, resulted in increased sporulation (Edwards *et al.,* 2014). Whilst the 01008 wk-1 (E10) isolate did produce a high level of HR-CFU by the end of rate of sporulation assays, this was not higher than the clinical isolate average. However, it was higher than the control strain, CD630Δ*erm*, the WT comparator of the published study and so similar conclusions can be drawn.

At 0 hours incubation isolates 12004 wk0, 12004 RL and 18001 wk0 were able to produce detectable HR-CFU, indicating that they were able to enter into sporulation much earlier than all others tested in this assay, which remained at the limit of detection However, no amino acid changes, truncations of deletions in oligopeptide permease genes were observed for these isolates, and so cannot support the same findings of (Edwards *et al., 2*014). As mentioned previously, the assay used in this current study was not able to separate sporulation phenotypes from that germination or viable spore outgrowth. Edwards *et al.,*

(2014) also used the same method of CFU enumeration after 0, 2, 24 and 48 h incubation, to investigate sporulation efficiency. However, went further to separate the phenotype by enumerating spores by phase-contrast and fluorescent microscopy for *In vitro* and *In vivo* studies as well as growth from spores, observing the same phenotype of higher spore production in the mutant, but no differences in growth. Whilst their method of growth from spores can be critiqued by against using enumeration of CFU/mL, thus not separating the growth phenotype from germination, outgrowth or spore dormancy, this advantageously separates sporulation from these latter phenotypes. Therefore, illustrating the need for spore titres of these clinical isolates to exclusively reveal any associations with this family of genes to a sporulation phenotype.

4.3.2 Rate of germination in recurrent and non-recurrent CDI

The initiation of spore germination of all clinical isolates from case studies 1-3 had a statistically significant difference compared to control strain CD630Δ*erm,* where smaller slopes were observed throughout the experiment. This was trend was supported by collated data, except for two isolates. Two isolates, 01013 wk-1 and 11001 RL, belonging to PCR ribotypes CE002 and CE029, respectively, had faster rates of germination than all clinical isolates tested. During the 9 to 10 minute time point there is a larger drop in OD_{600}, for control strains and clinical isolates, compared to rest of the assay. The difference observed for all isolates, excluding 01013 wk-1 and 11001 RL (see Figure 4.4, Figure 4.9 and Figure 4.10) has been shown to be extremely or very highly significant in the rates of germination initiation compared to CD630Δ*erm* using the uncorrected Dunn's test. Interestingly, clinical isolates had a slower rate of drop in OD_{600} shown by the smaller slope from 0 to 9 minutes, compared to control strains CD630Δ*erm* and EK28 indicating a slower rate of germination initiation. During the 9 to 10 minute time point there is a steeper drop in OD_{600} for a majority of patient isolates compared to the rest of the assay, indicating a collective germination event of a larger

proportion of spores. Additionally, all patient isolates have a higher endpoint than control strains indicating that a larger proportion of spores had not initiated germination.

Interestingly, the persistent and relapse causing isolates of the recurrent case studies are shown in this study to have higher end points than their earlier counterparts indicating that a larger proportion of spores did not initiate germination. This could relate to the persistence of isolates over the course of infection, with spores persisting beyond the routine course of treatment and chaser therapy. For patient 03008, although indicated to have a non-recurrent infection by SNP analysis and clinical diagnosis, samples were submitted at wk 0 and 4 indicating carriage during anti-CDI therapy. Here, the same phenotype of the later isolate having higher end points to the earlier counterpart, again indicating a larger proportion of spores had not initiated germination (Figure 4.4 – case study 3). This phenotype was observed for 3 out of 6 relapse patients of the recurrence cohort (Figure 4.9) (01013 was a reinfection) but was not seen for the remaining patients studied in the non-recurrent cohort. Admittedly, only two patients of the non-recurrence cohort, 01015 and 01028 (Figure 4.10), had more than one strain during this analysis, however both strains shared near identical rates of germination for those patients. Additionally, such a limit on strains submitted would be expected for patients with non-recurrent infections where there should be no symptomatic carriage of *C. difficile* beyond treatment at the time intervals of this trial, as implicated as an antonym of the definition of recurrence. Even so, the major limitations of this study included the low sample size, incomplete submission of requested samples, and continued trial enrolment. So, no conclusion on the rate of initiation of germination correlating with relapse or persistence can be confidently made.

The study of Moore *et al.,* (2013) using the same drop in OD_{600} method to observe germination found that germination efficiency was significantly higher in recurrent CDI or when metronidazole therapy failed than when therapy was successful. Conversely, several post treatment and recurrent isolates had higher end points to their earlier counterparts for the initiation of germination study, indicating that a larger proportion of spores had not initiated germination. Moore *et al.,* (2013) also correlated a decrease germination rates to

PCR ribotype 027, however, only 12 isolates representing three ribotypes, including ribotypes 106 and 078, where included in that study. Interestingly, all clinical isolates had reduced rates of germination in comparison to the control strain. As no PCR ribotype 027 isolates were detected in this study, no comment can be made as to whether this phenotype is associated with so called hypervirulent strains of this type.

Whilst germination percentages were not obtained for clinical isolates in this chapter, a level of dormancy was observed for all isolates tested using the loss of heat resistance assay. Coupled with the non-germination fate for post treatment and recurrent isolates observed by the rate of germination initiation, similarities can be drawn with that of Carlson *et al.,* (2015), where isolates that caused severe disease exhibited significantly lower germination efficiency. Carlson *et al.,* (2015) also observed that isolates of the most predominate type in hospitals during the collection period, type 014-020, exhibited higher levels of germination compared to those tested including ribotype 027. Two isolates, 01013 wk-1 and 11001 RL, belonging to PCR ribotypes CE002 and CE029, respectively, had faster rates of germination than all clinical isolates tested, with these ribotypes accounting for 21.82% and 5.45%, of those detected in this trial respectively. Whilst there is some overlap between findings, this was not observed for the next most frequently detected PCR ribotype CE014, nor for the germination rates of isolates belonging to this type. Again, as no PCR ribotype 027 isolates were detected in this study, no comment can be made as to whether this phenotype is associated with this so called hypervirulent strain type.

Carlson *et al.,* (2015) went on to highlight that further study was needed to identify differences inherent to these isolates that allowed control over germination rates. This is a key novelty of this study in that WGS identified SNP differences not only to key genes of the reference strain, but also between isolate genomes to which they were mapped. An additional limitation of phenotypic associations to PCR ribotypes is that classifications of recurrence are based on typing methods. Furthermore, when SNP differences are made in published work, clinical isolates are mapped to the reference strain, CD630, rather than to that of a suspected strain relative. As mentioned in Chapter 3, sequence variation in 23 genes

of interest, with published association to spore phenotypes, were observed in this recurrence cohort. It was hoped that these genetic differences would shed light on potential mechanisms of advantageous fitness for strains isolated from the same host.

4.3.2.1 Limitations and advancement of knowledge in the study of spore germination: Significantly reduced rates of germination of clinical isolates compared to control strain CD630Δ*erm*, could be linked to mutations in regulators, *cspA* and csp*B,* of TA- dependant germination

One of the novel findings in this study includes the phenotypes observed for initiation of TA-dependant germination of clinical isolates, where incomplete germination or non-germination fate for a large number of spores compared to control strains was concluded. Large levels of genomic variation were found between all 16 strains of the recurrent cohort strains in Chapter 3 for genes CD630_03320 (*bclA1*) and CD630_33490 (*blcA3*) and CD630_22470 (*cspBA*) with published phenotypes associated with spore gemination. As non-recurrent isolates where not investigated for genomic variation these amino acid changes cannot be exclusively associated for slow and low germination rates observed for all clinical isolates. Nevertheless, this measurable phenotype remains an important and novel finding of the study.

It is possible that these genes do play an important role in germination for all clinical isolates. As mentioned in Chapter 3, the BclA1 protein has been acknowledged as a critical factor in the initial stages of *C. diffcile* colonization (Phetcharaburanin *et al.,* 2014). The heterogeneity of the exosporium ultrastructure and of BclA1 proteins of *C. difficile* spores in various isolates of clinical relevance compared to the reference strain has been noted (Pizarro-Guajardo *et al.,* 2016). *In vivo* studies have shown that *bclA1*$^{-}$ spores germinated faster than the wild-type strain, CD630Δ*erm*, whilst no difference was observed for the *bclA3*$^{-}$ mutant (Phetcharaburanin *et al.,* 2014). Whilst the opposite was true of relapse isolate 01008 RL (H4)

for which a truncation of *bclA1* was observed, despite using the same assay methods, germination was far reduced compared to the control strain. The truncation observed for the clinical isolate in this study was acquired somewhere after the initial infection and relapse episode of this strain. It could be hypothesised that any genomic variation observed over time would be advantageous to the overall fitness of this strain. For example, in a response to overcome environmental pressures of antibiotic exposure, as well as competition with a co-infecting isolate found during the same trial period. However, it is unknown whether this truncation contributed to the virulence of this relapse causing strain. It has been reported that strain R20291, a 'hypervirulent' 027 strain, also carries a truncated BclA1 protein. The involvement of BclA1 in virulence was experimentally assessed and a higher ID_{50} was observed when mice were challenged with spores of the hypervirulent strain (Phetcharaburanin *et al.,* 2014).

Similarly, it is possible that the variation observed for the subtilisin-like serine germination related protease gene CD630_22470, *cspBA*, contributed to the difference in germination initiation observed amongst clinical isolates. In order to maintain spore dormancy, cortex lytic enzymes remain inactive until they receive a specific signal downstream of germinant-receptor interactions (Kochan *et al.,* 2018). CspB is hypothesized to be important for activating the SCLE, pro-SleC, to its active, cortex-degrading form (Kevorkian *et al.,* 2016), whilst *cspA* is essential for spore germination by controlling the levels of CspC, the bile acid germinant receptor, into the developing spore (Francis *et al.,* 2013a; Kevorkian *et al.,* 2016). Therefore, it could be hypothesised that any amino acid change observed for the fusion gene would affect either the separation of CspB and CspA domains, and/ or the peptide structure of these individual proteins. Interestingly, mutations in this fusion gene have been observed to produce germination null phenotypes (Francis *et al.,* 2013). Random mutant strains with germination null phenotypes were also found to have SNPs in genes *cspBA* and *cspC,* compared to parental and reference strain UK1 PCR ribotype 027. This study performed further functional studies of these genes to confidently attribute germination null phenotypes to these genes.

Furthermore, amino acid changes in *cspC* and *gerS* genes were also observed for isolates involved with the most genetic variation of those studied, as well as for both cases of persistent carriage, 01008 wk-1 (E10), 01008 wk 12 (G2), 01016 wk -1 and wk 12. Whilst additional amino acid changes were observed in CD630_05510 (*sleC*) for isolates 01016 wk -1 and wk 12. Whilst isolates 01008 wk-1 (E2) and 01008 wk 12 (G2) were indicated to be the same strain, with only a 6 SNP difference from WGS analysis in Chapter 3, the initiation of germination for the week 12 isolate proceeds slower compared to its counterpart with an extremely high significant difference, $P = 0.0001$. The genetic variation acquired for the later isolate in *gerS* and *cspC*, may have a role in persistence. Whilst isolate 01008 wk-1 (E10), had a similar rate of germination, but ended with a higher end point OD compared to the relapsing counterpart, isolate 0108 RL (H4). The same was observed for isolates 01016 wk -1 and wk 12, however, did not have a large difference in end point OD. Collectively, these genetic variations in key germination genes could relate to the persistence of isolates over the course of infection, with spores persisting during the routine course of treatment and rifaximin chaser therapy. Further studies for specific effects on CA-DPA release, regarding *sleC*, and germination in co-germinant medium, regarding *gerS*, would be required to elucidate any additional effect of these genomic differences.

Additional amino acid changes in gene CD630_07750, s*poVAE*, were observed for isolates 11007 wk-1 (C2), 11007 wk 0 and 11007 RL. The key finding for this patient was that post treatment isolates had slower rates of germination and higher OD at endpoint compared their diagnostic counterparts, additionally, the outgrowth from spores was significantly different for strain 11007 wk-1 (C5). However, the observation of genetic variation in *spoVAE* for all isolates tested cannot be attributed to phenotypic differences observed as isolate 11007 wk-1 (C5) was not included in further genomic analysis.

Unfortunately, the fastest rates of germination initiation observed, for isolates 01013 wk-1 and 11001 RL, compared to all clinical isolates, could not be associated with any genetic variation using genomic analysis, particularly as isolate 11007 RL was not included in the analysis, which itself was already limited to the 24 genes of interest.

4.3.3 Outgrowth of spores in recurrence

A competitive fitness was inferred for co-infecting isolates; where one had a shorter lag phase compared to the other. This phenotype was hypothesised as a model of delayed, yet, advantageous colonisation potential for the slow outgrowing isolates. Whilst the faster outgrowing isolates would be subjected to antibiotic activity and so clearance from the patient, as evidenced by the former isolate strain persisting throughout the sampling period for two of the case studies (patients 01008 and 11007, see Figure 4.5). A contradictory theme of the collated data however (see

Figure 4.11), was the post treatment isolates having lower end points than their earlier counterpart isolates, where the opposite was true for the two recurrence patient case studies. It was concluded that these features (the exponential and log phases of the curve) was a vegetative cell growth phenotype that would have significance on symptomatic infection potential. However, not an assumed recurrence associated phenotype where spore persistence or dormancy would lead to infection post treatment. A further assay to illustrate this was performed based on the loss of spore heat resistance.

4.3.3.1 Limitations and advancement of knowledge in the study of spore outgrowth: Outgrowth from spores showed strain-to-strain variation with competitive outgrowth implied for co-infecting isolates

Differences in lag phase of outgrowth from spore experiments illustrated a competitive fitness for co-infecting isolates; where one had a shorter lag phase compared to the other. Interestingly, all clinical isolates showed truncations in the gene CD630_32410 (*prdB*) in Chapter 3, with additional amino acid changes observed for isolate 01008 wk 12 (G2). However, no significant difference was observed for outgrowth from spore assays. One critique could be that whilst *prdB* has been shown to be optimal for *C. difficile* growth, the assay employed in this study did not correctly assess vegetative cell growth or outgrowth

from spores independently from germination as a suspension of spores was used rather than a normalized started culture of vegetative cells. Additionally, to further assess the effect of genomic variation in clinical isolates compared to the controls strain, experiments would need to include varying growth mediums, to see if optimal growth is affected as a result of the truncation or amino acid changes found in *prdB* for clinical isolates.

4.3.4 Loss of spore heat resistance in recurrence

The continued measurement of detectable heat resistance CFU (see Figure 4.6 and Figure 4.12) throughout the duration of this assay illustrates a level of spore dormancy across all isolates. Isolate 01013 wk-1 enters spore dormancy earliest than all isolates and the control strain in this assay, with no substantial losses of heat resistance observed beyond 5 minutes. This indicates that all the spores that have germinated, and of those that can efficiently outgrow, have done so within 5 mins. Supporting the observation in outgrowth phenotype (see Figure 4.11), where it had the longest lag phase of nearly 6 hours, and the smallest slope during exponential phase. Collectively indicating that a smaller number of spores outgrew compared to other isolates. Similar findings are observed for the case study of patient 01008 wk-1 (E10), in both outgrowth and loss of heat resistance phenotypes. However, it may also be that these strains produce more spores, however, this speculative hypothesis would need to be confirmed by a spore titre assay.

The remaining isolates of case study 1 entered substantial spore dormancy by 90 minutes in this assay, with a detectable endpoint of ~10^4 CFU/mL. Even so, minute losses in heat resistance were detected for all isolates and whilst not observed as logs in decline, the infectious dose of spores is not easily assessed when it comes to spore dormancy and recurrence, minute developments of CFU could translate to the beginning of a recurrent episode in the host. For example, whilst the major loss in heat resistance was observed to end at 60 mins, for isolate 01008 RL (H4), 3.166 x 10^3 CFU/mL was produced from 'dormant' spores

that germinated between 60 and 120 mins (the end of the assay). Such germination could amass to high levels of dividing vegetative cells a potential indicated by the high growth rates during the exponential phase of this isolate in Figure 4.5.

Low levels of decline in heat-resistance CFU was detected across different time points by methods employed in this study, up until the end of the assay for all isolates. This indicates that germination isn't entirely completed at a single time point and not for the total population of spores of a strain. It may be presumed that germination could continue after a period of longer dormancy such as time intervals observed in clinical recurrence. Nevertheless, this speculation cannot be confirmed without continued time points.

4.3.4.1 Limitations and advancement of knowledge in the study of spore dormancy: Loss of heat resistance assays reveal a comparable level of spore dormancy for all isolates

One key finding of this study was that no complete loss of heat resistance was observed for any isolate tested illustrating a level of spore dormancy for all strains, irrespective of PCR ribotype or infection status. Taken alone, this level of dormancy would imply a potential for all isolates to be persistently carried. Host factors clearly play a role as isolates were not persistently detected by microbiological investigation, however, controlled rates of germination and strain-to-strain variation in outgrowth may also contribute, amongst other virulence factors. It would be interesting to see if remaining spores do eventually germinate beyond the time constrains of this study and if any associations could be made to recurrent status or ribotype if such events occurred.

Interestingly, differences were observed for time points at which isolates enter into spore dormancy. Particularly, isolates 01008 wk-1 (E10) and 01013 wk-1, belonging to PCR ribotype CE014 and CE002, respectively, that enter spore dormancy earlier than all clinical isolates tested as well as the control strain, CD630Δ*erm*, with no substantial losses of heat resistance

observed beyond 5 minutes, followed by isolate 01013 wk 12 of PCR ribotype CE014. Whilst analysis of genetic variation of genes in Chapter 3, did not highlight these strains, analysis was limited to the genes selected. Amino acid variations were found for the gene CD630_12900, encoding SASPs for isolates 11007 wk-1 (C2), 11007 wk 0 and 11007 RL. These proteins have been found to have a role in the resistance of spore DNA damage. The effect of resistance phenotypes on spores or that on levels of dormancy was not investigated in this study and so no association ca be made for this gene using the methods employed by this study.

4.3.5 Final Implications and conclusions

It was hypothesized that germination and outgrowth redundancies of strains relate to prolonged spore dormancy and persistence and thus these specific phenotypes would act as an explanation/virulence factor of relapse. Whereas reinfection would be predicted to be a result of competitive outgrowth of each co-infecting isolate, where antibiotic activity would affect the faster outgrowing more so than the slower counterpart. The data collected shows some support to these hypotheses, with a number of early infecting isolates having a lower proportion of spores initiating germination compared to their post treatment and relapse counterparts. Whilst this trend was only observed in the recurrence cohort, the limited data set restricts such a conclusion. Low-level germination was observed exclusively to clinical isolates and could be related to mutations in regulators and receptors of germination. Spore dormancy was detected in all clinical isolates tested indicating the potential for persistent carriage in all isolates. Limitations of this study include minimal non-recurrence isolates being testing for dormancy and genomic variation, so no definite inferences as to a recurrence vs non-recurrence phenotype can be confidently made.

With the definition of germination varying in depth between research groups; some including outgrowth, others not, it may be that this collective process has an implication on recurrence and the differentiation into specific phenotypes was not ascertained in this study. Future

work is needed to aid the separation of phenotypes in the rate of sporulation including a spore titre to ascertain sporulation potential. By employing methods to independently assess sporulation, germination and spore outgrowth, conclusions of the findings observed will have had more weight towards specific processes, a comment to made on all published studies and well as this one. Genetic variation observed in sporulation, germination and outgrowth genes in genomes of recurrent isolations, did have distinct implication on the phenotypes of clinical isolates of the recurrent, particularly in germination. Further work would be needed to show this genetic variation in non-recurrent isolates. One of the main novelties of this studies is the significantly reduced rates of germination observed for isolates compared to the control strain. Additionally, by not mapping strains to a single reference strain, individual SNP differences between isolates of the presumed same family were resolved and allowed more robust and supporting distinctions between clinical and laboratory strains. Such mutations highlight the mechanisms that lead to this increased control over germination and support a role of this control in the pathogenesis of *C. difficile*.

Chapter 5

Chapter 5: Characterisation of ClosTron mutants reveal germination protease as a critical factor for spore viability

5.1 Introduction

CDAD has emerged as a leading healthcare-associated infection of great clinical and economic significance. The infection is not only characterised by debilitating symptoms but is also complicated with a high incidence of recurrence and is increasingly found in the community. *C. difficile* spores are resistant to all manner of chemical and physical agents, and as a result spores shed in faecal matter are very difficult to eradicate, leading to infection, or re-infection, of cohabiting individuals through inadvertent ingestion of spores from contaminated surfaces (Riggs *et al., 2*007; Gerding, Muto, and Owens, 2008). Interventions that target sporulation could, therefore, have an enormous impact on disease transmission. Nonetheless, it is the toxin-producing vegetative cells that subsequently arise following germination that cause disease and so intervention at the stage of germination would be an effective means of disease prevention.

Much of the resistant features of *C. difficile* spores are a consequence of to its' uniquely layered structure surrounding the spore core. A key protective feature within the spore core is the presence of 5–7 kDa α/β-type small acid soluble proteins (SASP) that make up 10-20% of total core protein and provide much of the protection to spore DNA (Fairhead *et al.,* 1993). SASPs, a group of very abundant small proteins that are found exclusively in spores, which are synthesized late in sporulation only in the developing spore (Setlow, 1988; 1995). SASPs are degrade early during germination and are hypothesised to providing a vital source of free amino acids for the outgrowing spore (Setlow, 1988; 1995). The two main types include the α/β-type and the γ-type SASP, however the γ-type SASPs, not known play no known role in spore resistance, are not found in all Clostridia and some members of the order *Bacillales* (Setlow and Waites, 1976; Granum *et al., 1*987; Raju *et al., 2*006; Vyas *et al., 2*011). By binding

directly to and saturating spore DNA, α/β-type SASPs provide an important component of spore resistance against chemical and other treatments that target spore DNA (Setlow, 1988; 1995). It has been demonstrated *in vitro* that α/β-type SASPs protect *B. subtilis* spores and DNA, effecting resistance to DNase and several restriction enzymes (Setlow *et al., 1*992), hydrogen peroxide (Setlow and Setlow, 1993a; Imlay and Linn, 1988 and Setlow *et al., 2*000), alkylating agents and UV (Setlow *et al., 1*998; Loshon *et al., 1*999; Setlow and Setlow, 1993a; Setlow, 1995; 2001; Lee *et al., 2*008). Whilst α/β-type SASP mediated DNA protection from wet heat (Setlow, 1995; Setlow *et al., 2*000; Raju *et al.,* 2006; Leyva-Illades *et al., 2*007), nitrous acid and formaldehyde (Loshon *et al., 1*999; Tennen *et al., 2*000; Paredes-Sabja *et al., 2*008a; Meaney *et al.,* 2016) have been demonstrated in spores and DNA of *B. subtilis, C. perfringens* and *C. botulinum*.

Although these α/β-type SASPs amass to high levels in the late stages of sporulation and remain stable in the dormant spore of *B. subtilis*, they are rapidly degraded into amino acids at the onset of germination (Setlow, 1995). This degradation is facilitated by a sequence specific protease termed germination protease (GPR) which cleaves SASPs at a specific site (Setlow 1988). GPR is produced during early stages of sporulation within the developing forespore and undergoes auto-processing for catalytic activity. In short, the larger P_{46} zymogen form of GPR is converted into a smaller P_{41} form (Setlow, 1994; Nessi *et al.,* 1998; Sanchez-Salas and Setlow, 1993; Illades-Aguiar and Setlow, 1994). It has been predicted that owing to the dehydration status of the dormant spore, the P_{41} form does not degrade α/β-type SASPs during spore development or dormancy (Sanchez-Salas *et al.,* 1992; Setlow, 1994; Sanchez-Salas and Setlow, 1993). Upon initiation of germination, the increased water content would permit activity of pre-embedded GPR and thus cleavage of α/β-type SASPs, would un-coat DNA, freeing it for transcription (Setlow, 1994; Sanchez-Salas, *et al.,* 1992). The cleaved SASPs are further degraded into amino acids that are fed into protein synthesis and energy metabolism in early spore outgrowth and subsequently the return to vegetative cell growth (Setlow, 1994; Popham, Sengupta, and Setlow, 1995; Sonenshein, Losick, and Hoch, 2002). Studies have found that in *Bacillus megaterium* and *B. subtilis* mutants with an inactivated

gpr gene, growth, sporulation, artificial spore germination remained unaffected, whilst SASP degradation and outgrowth were significantly delayed (Sanchez-Salas *et al., 1*992), thus evidencing a role of GPR in efficient spore outgrowth.

Knowledge of sporulation and germination pathways in *Clostridia* remains largely based on the equivalent processes in *B. subtilis*, with speculations based on the presence of putative orthologues in *C. difficile* and close relatives. The strategy of genetic manipulation of *C. difficile* has so far been focused on generating mutants of homologous genes involved in sporulation and/or germination. The genome of *C. difficile* 630 was sequenced in 2006, following this, the proteomic analysis of *C. difficile* 630 spores (Lawley *et al., 2*009) identified a total of 336 spore-associated polypeptides, 39 of these being predicted to be secreted or surface associated. However, many of these genes have no predicted function. Nine percent of the spore proteome consists of spore genes that are currently unique to *C. difficile*. Many of these genes encode hypothetical proteins that can be preferentially targeted in mutant construction. A putative germination protease gene (CD24700) has been detected in the genome of Clostridia, however at the time of this study, it had not yet been characterised in *Clostridioides difficile* (Sebaihia *et al.,* 2006b). Protein studies in *C. perfringens* and *C. acetobutylicum* have illustrated proteolytic activity of recombinant GPR on recombinant α/β-type SASPs and the protection of DNA by recombinant α/β-type SASPs and crude acid-soluble spore extracts (Wetzel and Fischer, 2015; Li and McClane, 2008). The number of α/β-type SASP in Clostridial species ranges from three to eleven genes, which are scattered throughout the genome, however, their role has not yet been demonstrated in *C. difficile*.

Of the 3981 predicted genes in the *C. difficile* strain 630 genome, an estimated 24% (968) are hypothetical proteins with unknown function (Kent, 2002). Lawley *et al.,* (*2*009) identified a total of 336 spore-associated proteins, with 29 spore genes being unique to *C. difficile* in comparison to 11 Clostridial genomes. The majority of these genes encode proteins with no known function and so were preferentially targeted for directed mutagenesis in our research group. Sporulation and germination phenotypes were screened by construction and characterisation of insertional ClosTron mutations in targeted genes of in the non-epidemic

C. difficile 630Δ*erm* strain. Targeted genes included CD19300 (hypothetical protein, with discounted MBL fold hydrolase annotation on NCBI), CD05460 (membrane protein), CD07730 (*spoVACDE*), CD07750 (*spoVACDE*), CD29630 (hypothetical protein), CD08670 (putative protein) and CD24700 (*gpr*). Mutant strains generated were respectively named CT19300, CT05460, CT07730, CD07750, CT29630, CT08670 and CT24700. In this study, the role of CD24700 (*gpr*) in *C. difficile* was further characterised with phenotype complementation and protein studies of recombinant GPR and its activity on SASPs determined by the crude extraction of acid-soluble spore proteins.

5.2 Results

5.2.1 Mutant generation

To assess the importance of CD19300, CD05460, CD29630 and CD08670 in *C. difficile* germination, the ClosTron system (Heap *et al., 2*007; Heap *et al., 2*010) was used to create independent insertional mutants of these genes in the *C. difficile* 630Δ*erm* background (see Section 2.7). Confirmation of insertions in CD630_05460, CD630_07730 and CD630_07750 were positive, showing PCR products at ~2.0 kb corresponding to insertion of the intron into the wild type (WT) gene. Junction PCR products at 0.3 kb confirm insertion at the target site of genes (see Figure 5.1). One erythromycin-resistant clone of each mutant was picked and the corresponding PCR products (Lanes 1, 10 and 23) as verified by Sanger sequencing at the intron insertion site.

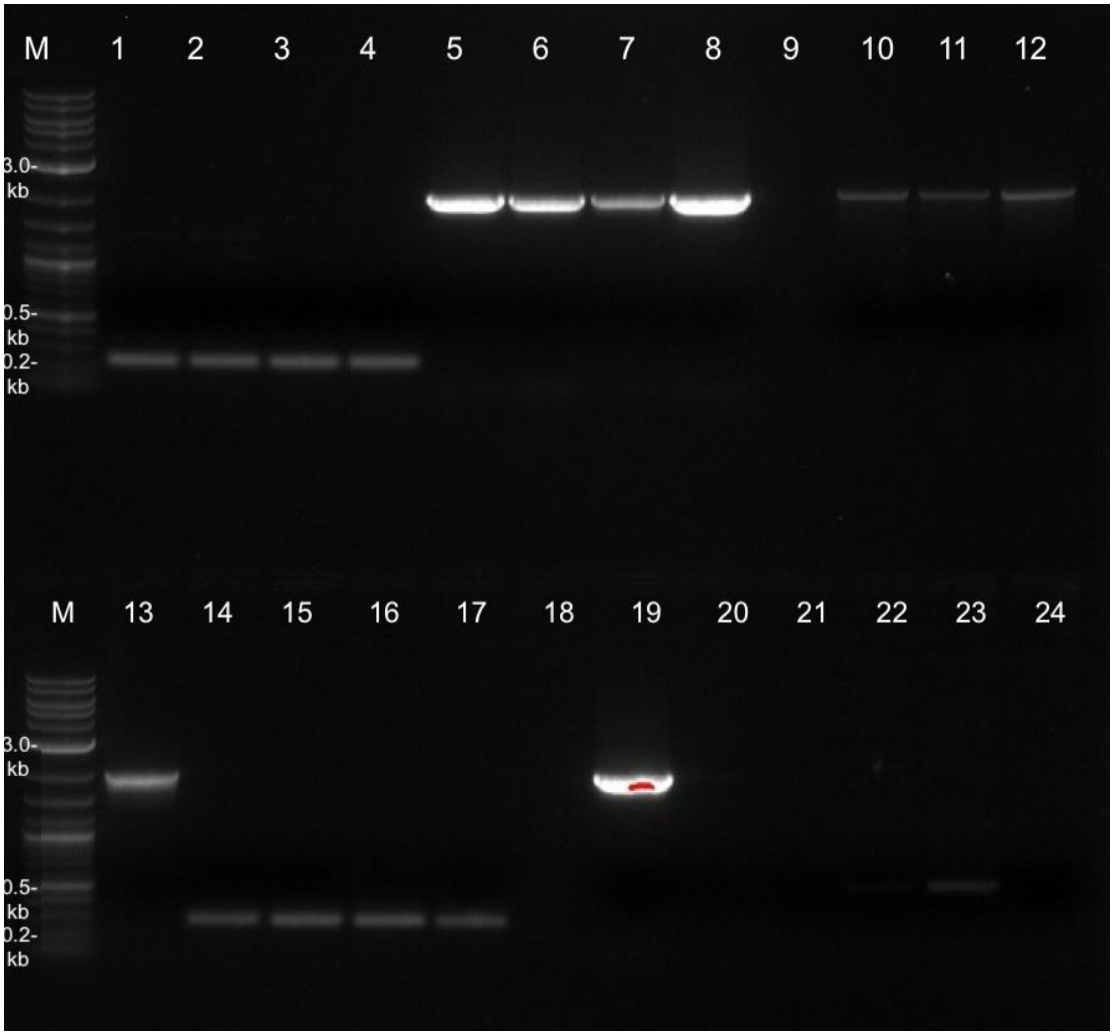


Figure 5.1: Colony PCR screening for putative *C. difficile* 630Δ*erm* ClosTron insertion mutants.

PCR was used initially to screen for intron insertion using the gene-specific primer pairs F1 and R1 for each mutant (expected band size of ~2.1 kb) and intron-specific EBS Universal primer and the gene-specific primer F1 (630Δ*erm*) used to screen the junction across the intron insertion at the target site of the WT gene (expected band size of 0.3 kb). Lanes M; 2-log DNA ladder (NEB, United Kingdom), 1-4; 4 colonies of CT05460 amplified with EBS universal and F1, 5-8; 4 colonies of CT05460 amplified with F1 and F2, 9, 18 and 24; blank, 10-12 and 13; 4 colonies of CT07730 amplified with F1 and R1, 14-17; 4 colonies of CT07730 amplied with EBS universal and F1, 19; 1 colony of CT07750 amplified with with F1 and F2 and 20-23; 4 colonies of CT07750 amplified with EBS universal and F1. Expected band sizes are indicated to the left of the gel.

To establish that the constructed mutants contained a single intron insertion a Southern blot was performed. Single bands <6557bp - >2027bp observed for CT05460, CT07730, CT07750, CT08670 and CT29630 indicate single intron insertions confirming ClosTron mutagenesis (see Figure 5.2). Mutant strains CT19300 and CT24700 were produced and confirmed by PCR and Southern blot by Patrick Ingle.

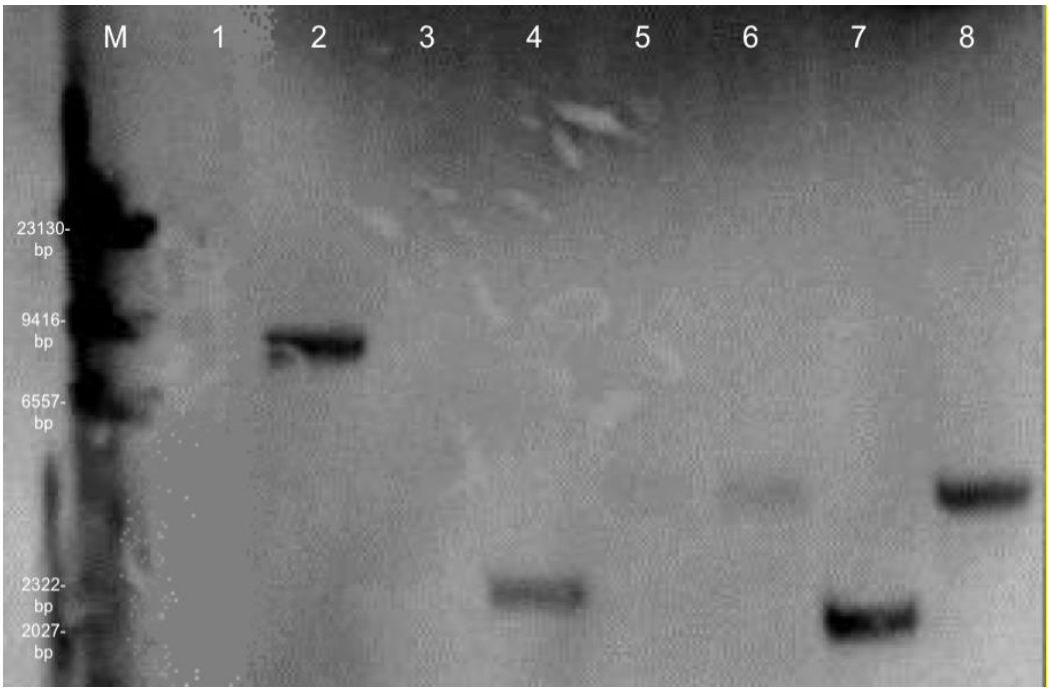


Figure 5.2: Southern blot analysis of ClosTron mutants using an intron specific probe.

In all cases genomic DNA has been cleaved with HindIII. M: λ DNA/HindIII marker, lanes 1 and 3: blank, lane 2: pMTL007C-E2, lane 4: CT05460, lane 5: CT07730 (*spoVAC*), lane 6: CT07750 (s*poVAE*), lane 7: CT08670 and lane 8: CT29630. Expected band sizes are indicated to the left of the gel.

5.2.2 Colony formation of heat-resistant spores of ClosTron mutant strains

The mutant strains were tested to determine their ability to form heat-resistant colonies over a 5-day period on BHIS agar supplemented with the bile salt taurocholate (see Figure 5.3). The HR-CFU development of strain CT24700 showed a significant difference ($P = 0.0169$), only starting after 96h and having a substantially (5-fold) lower HR-CFU at endpoint compared to WT. While CT19300, CT08670 and CT05460 show decreased development of HR-CFU over the five-day period and at endpoint, they show no statistical significance using the uncorrected Dunn's test. Interestingly, the initial development of HR-CFU for CT29630 follows closely to WT but, however, seems to plateau at 48 h, with 3-fold lower HR-CFU at end point compared to WT.

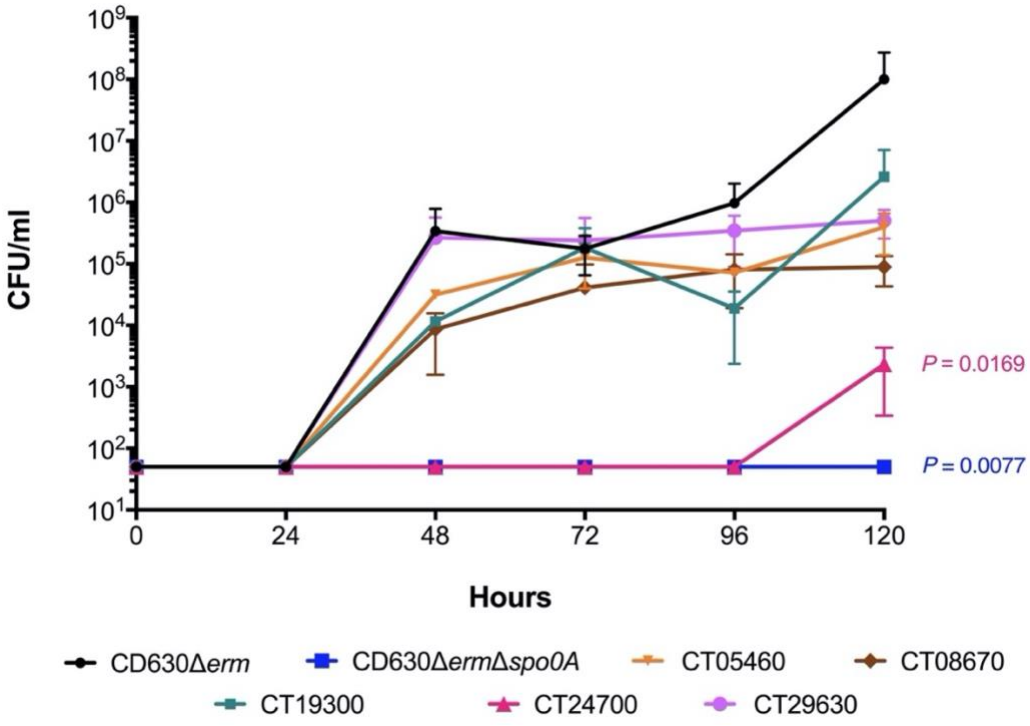


Figure 5.3: Heat-resistant CFU development of *C. difficile* 630Δ*erm* mutants over a 5-day incubation period in BHIS broth.

The detection limit was 50 CFU/mL, as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*spo0A*. Symbols represent the averages of three independent experiments and error bars of the standard errors of the mean.

5.2.3 Effect of ClosTron mutations on *C. difficile* sporulation

The decreased/delayed ability of ClosTron mutants to develop heat- resistant CFU could be a consequence of a number of possible defects in sporulation, germination and outgrowth. A reduction in the observed colony formation following heat shock would suggest that the mutant strains are unable to complete one or more of these processes. To examine the sporulation efficiencies of the mutant strains, spores were counted after five days by phase-contrast microscopy (see Figure 5.4) and the results were compared to the observed development of heat-resistant CFU (see Figure 5.3). Except for a sporulation negative control, CD630Δ*erm*Δ*spo0A*, in which the master regulator of sporulation, *spo0A*, was insertionally inactivated (Heap, JT *et al.,* 2007b), spore titres amongst the mutants showed no significant difference in comparison to CD630Δ*erm*. This suggests that inactivating these genes does not affect *C. difficile* spore formation, but that in these mutants, spores are unable to form colonies after heat treatment and stimulation by sodium taurocholate. Following the flowchart of initial mutant phenotype characterisation (see Figure 7.1 of the appendix), results indicate that the targeted genes are involved in the germination, outgrowth or spore viability of *C. difficile*.

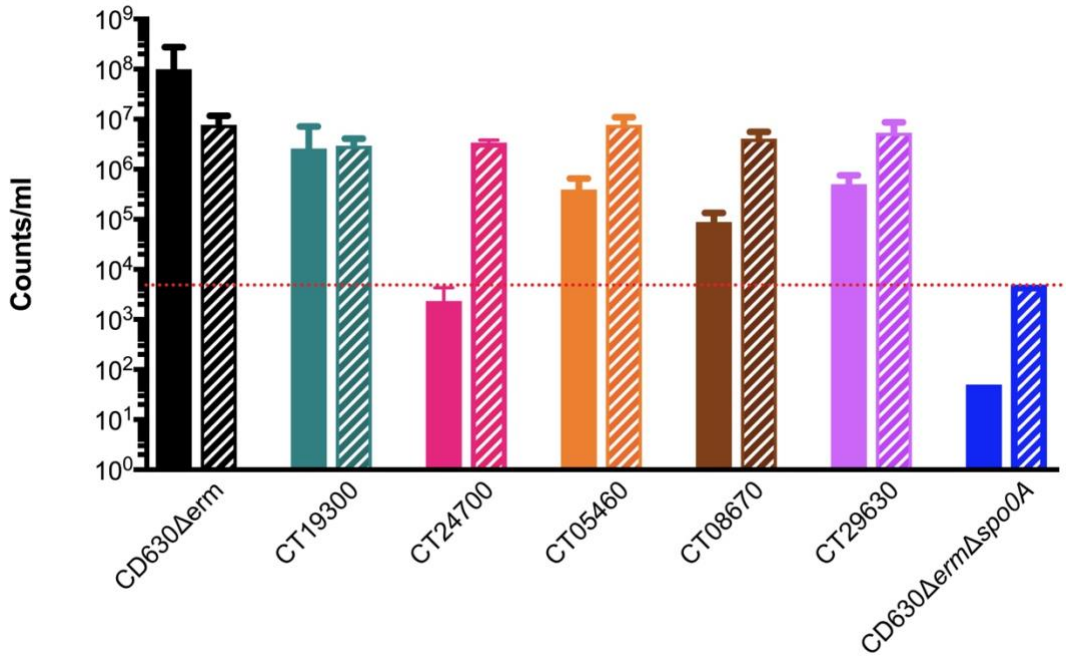


Figure 5.4: Numbers of heat-resistant CFU and spore titres after five days incubation in BHIS broth.

Spore counts (bars with diagonal lines) and heat-resistant CFU (filled bars) were determined for *C. difficile* 630Δ*erm*, and mutants after 5-days. Spore titres and heat-resistant CFU were enumerated as described in 2.4.5. The bars represent the averages of three independent experiments, and the error bars indicate the standard errors of the means. The detection limit for colony counts was 50 CFU/mL, and the detection limit for spore counts was 5 $\times 10^3$ spores as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*spo0A*.

5.2.4 Effect of mutations on *C. difficile* spore viability

It was important to demonstrate that the observed phenotypes from the sporulation rates are not a result of altered spore heat resistance. Decoating of spores has previously been used to distinguish between spore viability and germination phenotypes (Popham *et al.,* 1995; Paredes-Sabja *et al.,* 2009) as decoating and plating with supplemental lysozyme allows a germination-deficient spore to form HR-CFU. Recovery of decoated spores on lysozyme plates shows CT19300, CT05460, CT08670 and CT29630 are able to produce HT-CFU similar to WT levels (see Figure 5.5) proving that these spores are viable and that the reduction of heat-resistant CFU development is due to these mutations affecting germination or vegetative cell outgrowth (see Figure 7.1 of the appendix). Whilst CT24700 showed no detectable spore recovery for four out of six replicates, 2 replicates developed CFU's very close the limit of detection, with a very highly significant difference to the WT (P = 0.0047) (see Figure 5.5) suggesting that this gene is involved in the viability of *C. difficile* spores (see Figure 7.1 of the appendix).

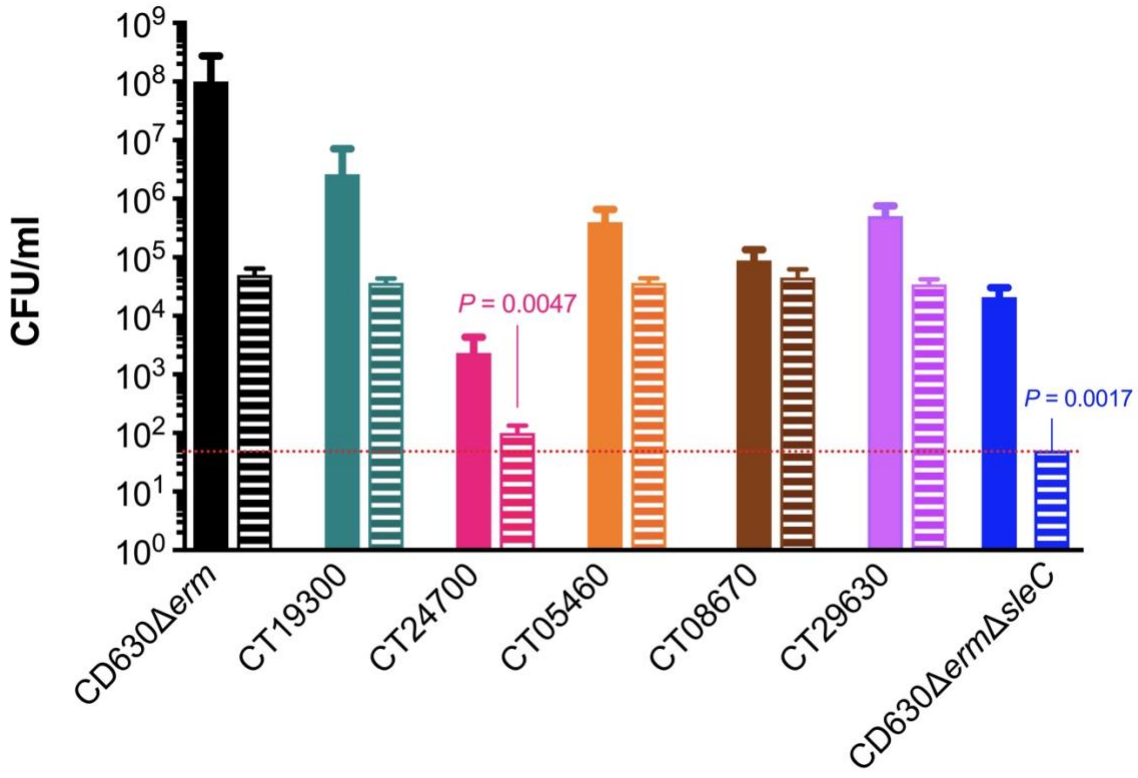


Figure 5.3: Viability of *C. difficile* 630Δ*erm* and mutant spores after the five-day incubation period in BHIS broth.

Heat-resistant CFU were determined as described in Section 2.4. Dormant spore suspensions were heated before being plated onto BHIS agar supplemented with 0.1% taurocholate (solid filled bars). Spore suspensions were decoated in decoating solution and serially diluted before they were plated onto BHIS agar supplemented with 1 µg/mL lysozyme (bars with horizontal lines). The bars indicate the averages of three independent experiments, and the error bars indicate standard errors of the means. The detection limit for the assay was 50 CFU/mL as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*sleC.*

5.2.5 Complementation of *gpr* ClosTron mutant

In order to show the phenotype observed for the ClosTron generated *gpr* mutant (CT24700) was a consequence of gene inactivation, complementation studies were performed (see section 2.7.3). Plasmid based complementation of the CT24700 mutant revealed some limitations. The pMTL-based complementation (CT24700+vComp) and overexpression (CT24700+vOver) plasmids were conjugated into the mutant CT24700 strain, along with an empty plasmid control (CT24700+Empty) in CD630Δ*erm* strain, and transformants were subject to spore phenotype assays. It was intended that the phenotype of CT24700 would be restored to WT levels. However, the CT24700 + vComp strain failed to demonstrate phenotypes comparable to WT as illustrated by the dashed red arrows on Figure 5.6. This trend was also seen for the WT strain carrying the empty pMTL8151 plasmid (CD630Δ*erm*+vEmpty). This is most probably due to lack of antibiotics used during assays to maintain the plasmid, coupled with the duration of the assay contributing to possible plasmid loss and limiting effects on growth. Plasmid based complementation did, however, confirm that GPR had no role in spore formation (see Figure 5.7). Initial overexpression studies showed that, comparable to the complementation strain, development of heat-resistant CFU's was delayed by 24 - 48 hrs in Figure 5.6, with reduced end-point levels, indicated by the dashed and solid red arrows, respectively. The limited CFU development of the overexpression strain observed in the spore viability assay (see Figure 5.8) cannot be meaningfully discussed as the assay overall failed as indicated by the HR-CFU development of negative control strain.

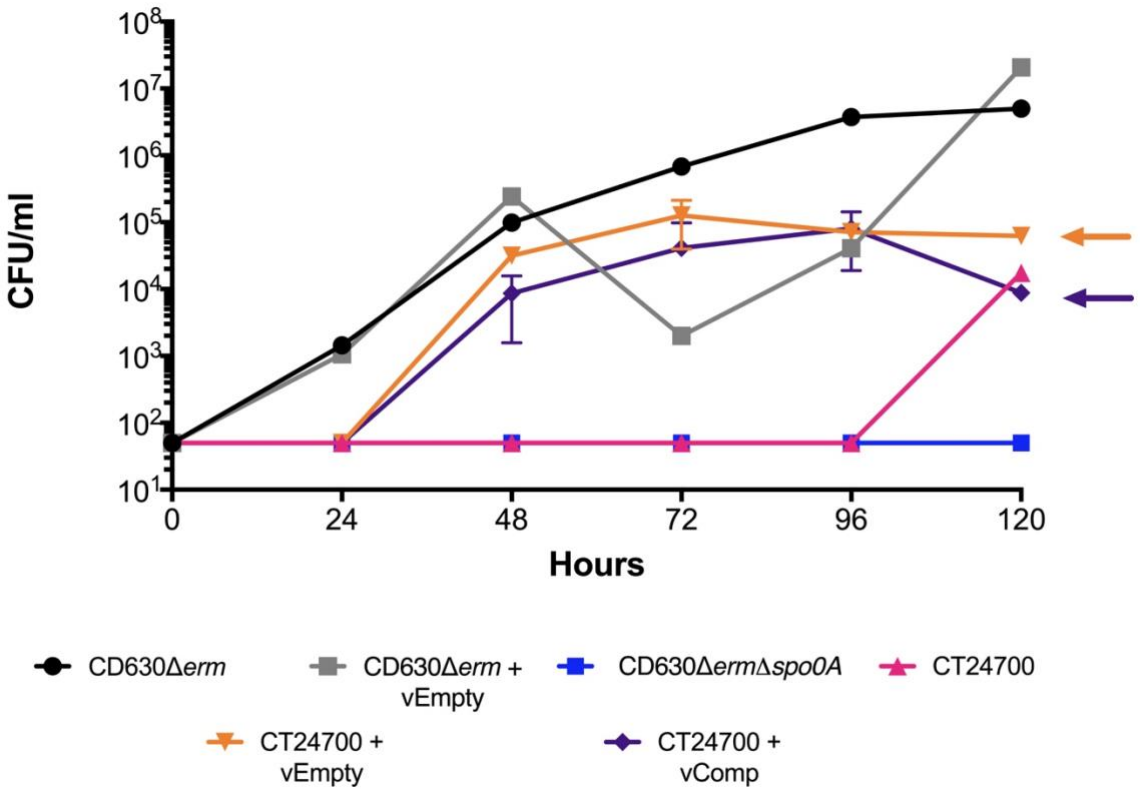


Figure 5.6: Heat-resistant CFU development of plasmid based complemented CT24700 mutant over a five-day incubation period in BHIS broth.

The detection limit was 50 CFU/mL as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*spo0A*. Symbols represent the averages of three independent experiments and error bars of the standard errors of the mean. CT24700 *P* = 0.0077 and *P* = 0.0021 for CD630Δ*erm*Δ*spo0A*.

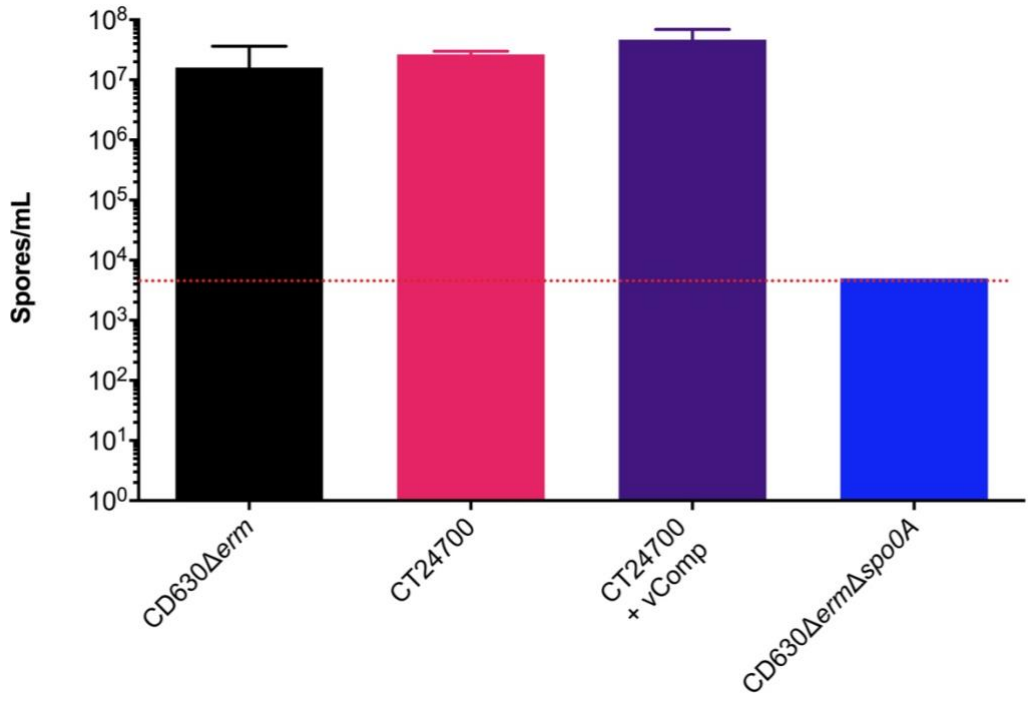


Figure 5.7: Spore titres of the plasmid based complemented CT24700 mutant after five days incubation in BHIS broth.

Spore counts determined for CD630Δ*erm*, CT24700 and CT24700 carrying a complementation vector (CT24700 + vComp). Spore titres and heat-resistant CFU were enumerated as described in Section 2.4. The bars represent the averages of three independent experiments, and the error bars indicate the standard errors of the means. The detection limit for colony counts was 50 CFU/mL, and the detection limit for spore counts was 5 $\times 10^3$ spores as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*spo0A*.

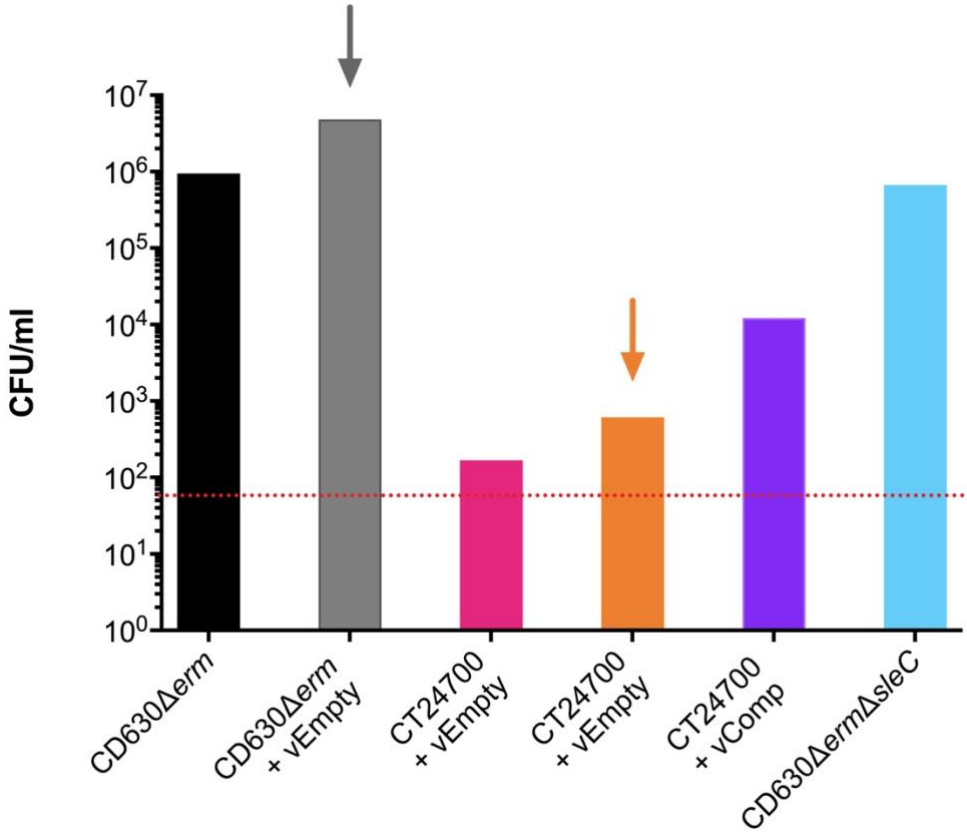


Figure 5.4: Viability of *C. difficile* 630Δ*erm* and plasmid complemented CT24700 spores after five incubation period in BHIS broth.

Spore suspensions were decoated and serially diluted before they were plated onto BHIS agar supplemented with 1 µg/mL lysozyme. The bars indicate the averages of three independent experiments, and the error bars indicate standard errors of the means. The detection limit for the assay was 50 CFU/mL as illustrated by the dotted line, however the negative control strain CD630Δ*erm*Δ*sleC* failed this experiment. A significant diference was shown for clostron mutant with and without empty vector compared to control strain CD630Δ*erm*, $P = 0.0059$ and $P = 0.0389$ for CT24700 and CT24700+vEmpty, respectively.

It was decided that chromosome level complementation by allelic coupled exchange would circumvent the limitations of the plasmid-based repair alleviating concerns of plasmid loss or appropriate gene dosage (see section 2.7.3.1). The complementation construct was generated by inserting the CD24700 gene with its native promoter with amplified restriction sites NotI and XhoI (~1,325 bp) into pMTL-RN1c. The overexpression construct was generated by inserting the CD24700 gene with amplified restriction sites *SacI* and *XhoI* (~ 980 bp) into pMTL-RN1x. Figure 5.9 shows the construct confirmation of ACE complementation (lanes 1-5) and ACE overexpression (lanes 6-10) vectors by plasmid DNA digests with the appropriate restriction enzymes. Following ACE repair of CT24700 at the *pyrE* locus in a CD630Δ*erm*Δ*pyrE* background, the complement strain (CT24700 + $_{chr}$Comp) and overexpression strain (CT24700 + $_{chr}$Over) were subject to PCR confirmation (see Figure 5.10) using primers flanking 500 bp both up and downstream of the *pyrE* (CD01870) gene (585 bp). The *pyrE* locus of CT24700 in a CD630Δ*erm* background revealed a ~ 2.129 kb product (lane 1), whilst the *pyrE* locus in of CT24700 in a CD630Δ*erm*Δ*pyrE* background (CT24700 Δ*pyrE*) revealed a >10 Kb product (lane 2). The CD630Δ*erm*Δ*pyrE* strain (CRG000) was generated by Yen Ng using the defective vector pMTL-JH18::lambda6.5 with a 6.5 Kb Lambda DNA insert within the allelic exchange cassette (Heap *et al., 2*012; Ng *et al., 2*013). This insert supported the selection of single and double crossovers on thiamphenicol supplemented BHIS agar, as the vector on its own had a replication defect. The *pyrE* locus of CT24700 in a CD630Δ*erm*Δ*pyrE* background complemented with CD24700 and its native promoter revealed a 3.23 kb product (lane 3), whilst the *pyrE* locus of CT24700 in a CD630Δ*erm*Δ*pyrE* background complemented with CD2470 and strong *fdx* promoter revealed a ~ 3.129 kb product (lane 4).

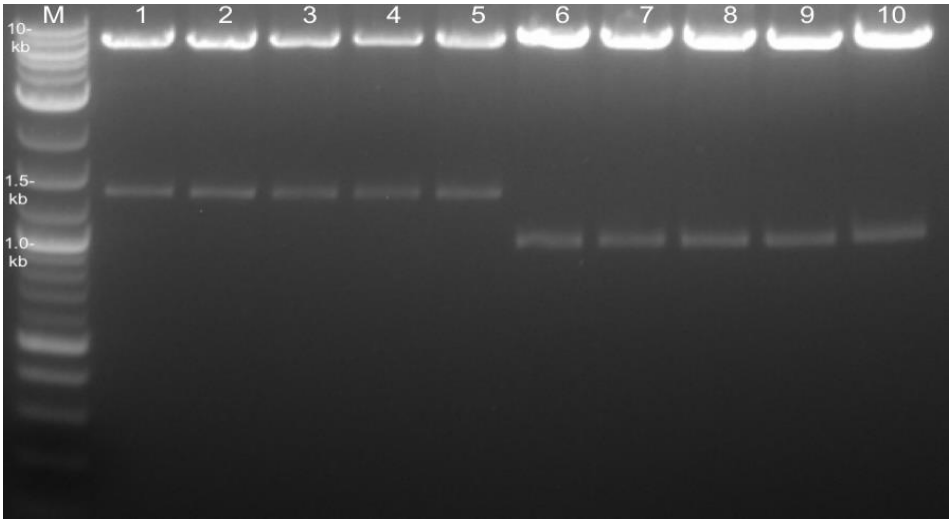


Figure 5.9: PCR digest for the confirmation of vector construction for complementation and over expression in the chromosome of CT24700 genome.

M: 2-Log DNA ladder, lanes 1-5 NotI and XhoI double digest of plasmid carrying CD24700 with native promoter, lanes 6-10: SacI and XhoI double digest of plasmid carrying CD24700 under the control of strong fdx promoter. Expected band sizes are indicated to the left of the gel.

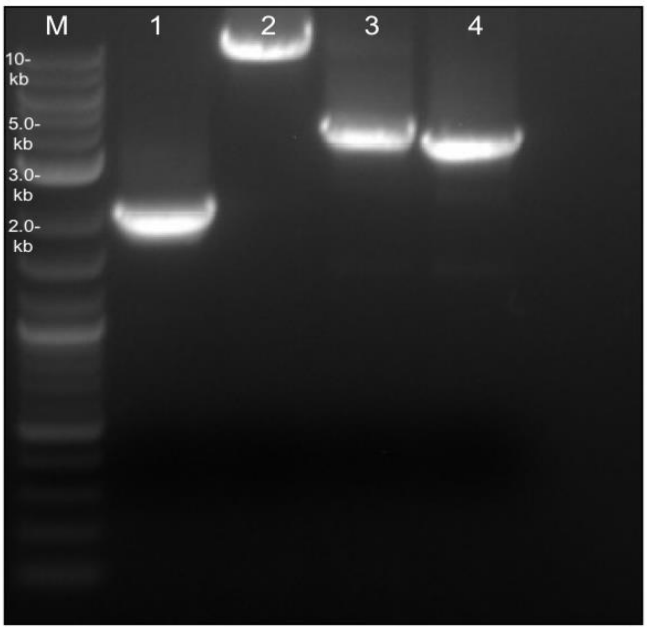


Figure 5.10: PCR confirmation of chromosomal complementation and overexpression via ACE repair of CT24700.

Primers flanking 500 bp up and downstream of the *pyrE* locus were used to screen the complementation and overexpression of CT24700 strain by ACE. L1: 2-Log DNA ladder, lane 1: WT *pyrE* gene (CD01870) locus in a CD630Δerm background, lane 2: pyrE locus of CT24700 in a CD630Δ*erm*Δ*pyrE* background (CT24700 Δ*pyrE*), lane 3: p*yrE* locus of CT24700 in a CD630Δ*erm*Δ*pyrE* background complemented with CD24700 and its native promoter at chromosome level (CT24700 + chrComp), and lane 4: pyrE locus of CT24700 in a CD630Δ*erm*Δ*pyrE* background complemented with CD2470 and strong fdx promoter at chromosome level (CT24700 + chrOver). Expected band sizes are indicated to the left of the gel.

5.2.6 Chromosome level complementation and overexpression studies in CT24700 mutant

Strains of the chromosomal complementation and overexpression were tested to determine their ability to form heat-resistant colonies over a 5-day period on BHIS agar supplemented with the bile salt taurocholate (see Figure 5.11). The HR-CFU development of CT24700 only starts after 96h and has a substantially (5-fold) lower HR-CFU at end-point compared to WT. Complementation (CT24700 + $_{chr}$Comp) and overexpression (CT24700 + $_{chr}$Over) strains have restored HR-CFU phenotypes comparable to WT, as indicated by green and purple arrows, respectively (see Figure 5.11). Spore viability was restored to comparable WT levels for CT24700 + $_{chr}$Comp strain, whilst restoration was achieved to a slightly reduced efficacy for overexpression strain CT24700 +$_{chr}$Over (see green and purple arrows, respectively, in Figure 5.12).

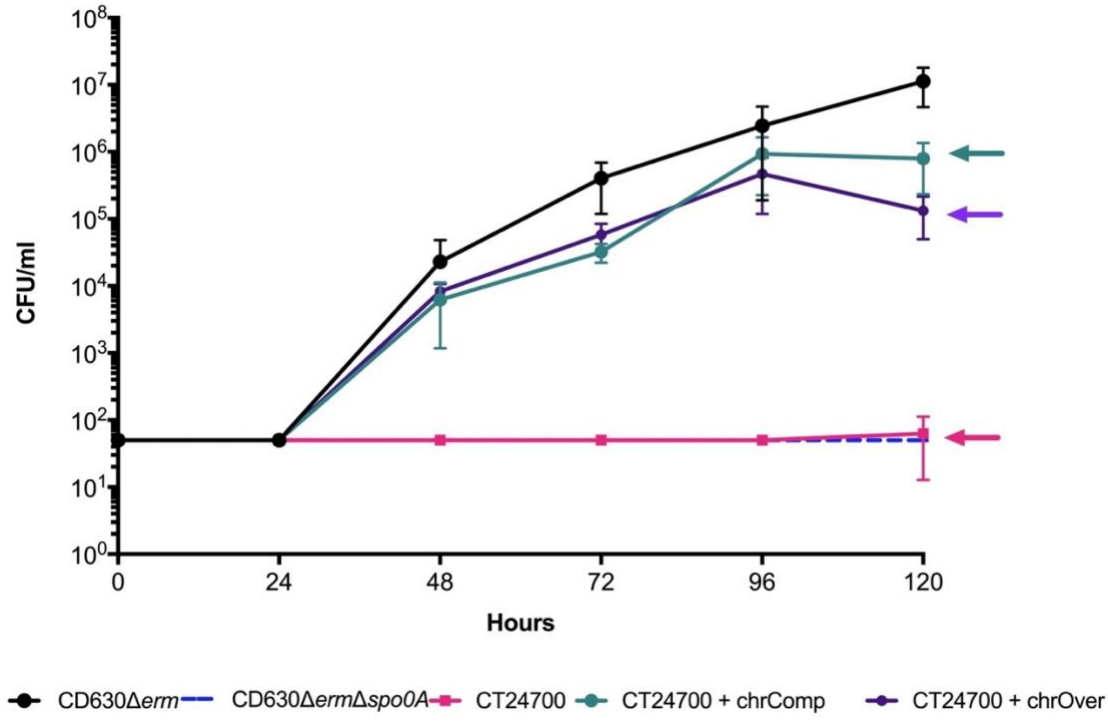


Figure 5.5: Heat-resistant CFU development of chromosome based complementation and overexpression of CT24700 mutant over a 5-day incubation period in BHIS broth.

Heat-resistant CFU development of CD630Δ*erm* mutants over a five-day incubation period in BHIS broth, as described in section 2.4. The detection limit was 50 CFU/mL, as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*spo0A*. Symbols represent the averages of three independent experiments and error bars the standard errors of the mean.

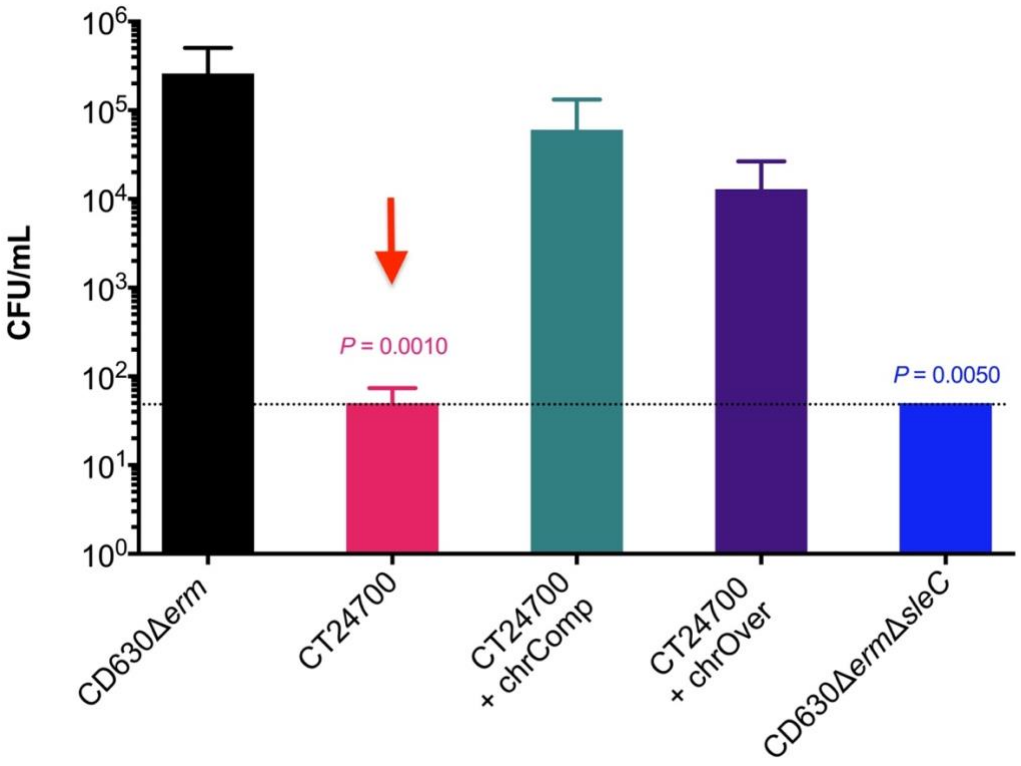


Figure 5.6: Spore viability of chromosome based complementation and overexpression of CT24700 mutant.

Spore suspensions were decoated and serially diluted before they were plated onto BHIS agar supplemented with 1 µg/mL lysozyme (bars with horizontal lines). The bars indicate the averages of three independent experiments, and the error bars indicate standard errors of the means. The detection limit for the assay was 50 CFU/mL, as illustrated by the dotted line and negative control strain CD630Δ*erm*Δ*sleC*.

5.2.7 Role of GPR in *C. difficile* spore germination and outgrowth

Further characterisation studies were carried out to elucidate whether spore viability phenotypes could be attributed to a role of *gpr* specifically in germination or outgrowth. The rate of germination was measured as a drop in OD_{600} of spores incubated with 0.1% Tc over time. Figure 5.13 shows that rates of germination for positive control strains CD630Δ*erm* and EK28 are comparable, with a slope of -0.0275 and -0.0256, respectively from 0 to 10 min. Whereas, a statistically significant difference was observed for all mutant (highly significant difference, P <0.0001), complement (significant difference, P <0.024) and over expression (highly significant difference, P <0.0011) strains in comparison to CD630Δ*erm* (see Table 5.1). Spores of CT2700 show minimal drop in OD, as supported by the small, yet present slope over 10 min (see Table 5.1), indicating minimal spore germination. Unfortunately, the complemented strain showed unexpected results during the initiation of germination experiment, with no return to WT initiation of germination phenotype despite repeated attempts. The overexpression strain CT24700+$_{chr}$Over has the fastest decline in OD, as shown by the largest slope from 0 to 10 min (see Table 5.1), indicative of spores completing germination much quicker when *gpr* expression is under the control of a stronger promoter. Moreover, the drop in OD, is considerably more reduced (see purple arrow) than that observed for spores of the *gpr* mutant and complementation strains (see pink and green arrows, respectively).

Table 5.1: Statistics of the slope from 0 to 10 minutes of the initiation of germination for *gpr* inactivation, complementation and overexpression strains.

Isolate	**Slope**	**Standard error of slope**	***P* value** [a]
CD630Δ*erm*	-0.027580383	0.001399909	-
EK28	-0.0256436	0.00224272	ns
CT24700	-0.0049241	0.00178521	< 0.0001
CT24700+$_{chr}$Comp	-0.0016241	0.00173711	0.024
CT24700+$_{chr}$Over	-0.0121314	0.00689398	0.0011

[a] P value for the statistical significance when compared to CD630Δerm, ns = no significant difference found

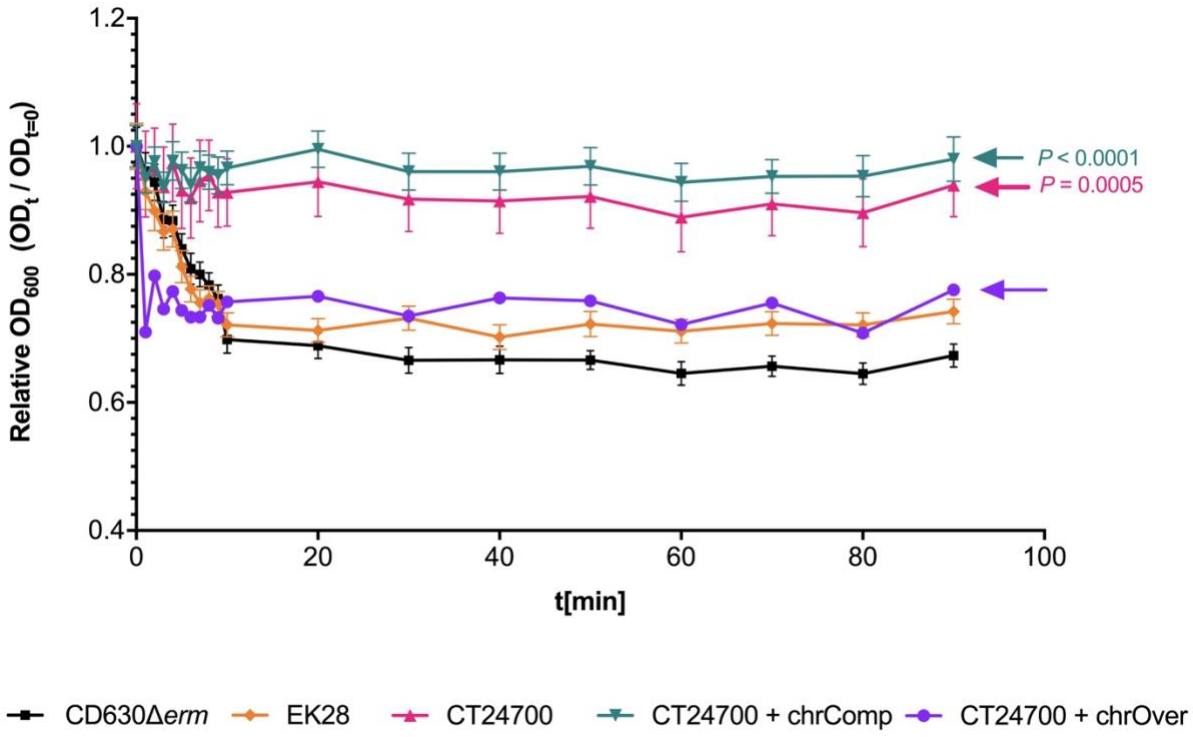


Figure 5.13: Spore germination of CT24700, chromosome based complementation and overexpression strains.

All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. A statistically significant difference in the drop in OD over 90 min was found for strains CT24700 and CT24700+ chrComp $P \leq 0.0001$ and $P = 0.0005$, respectively, in comparision to control strain, CD630Δ*erm*.

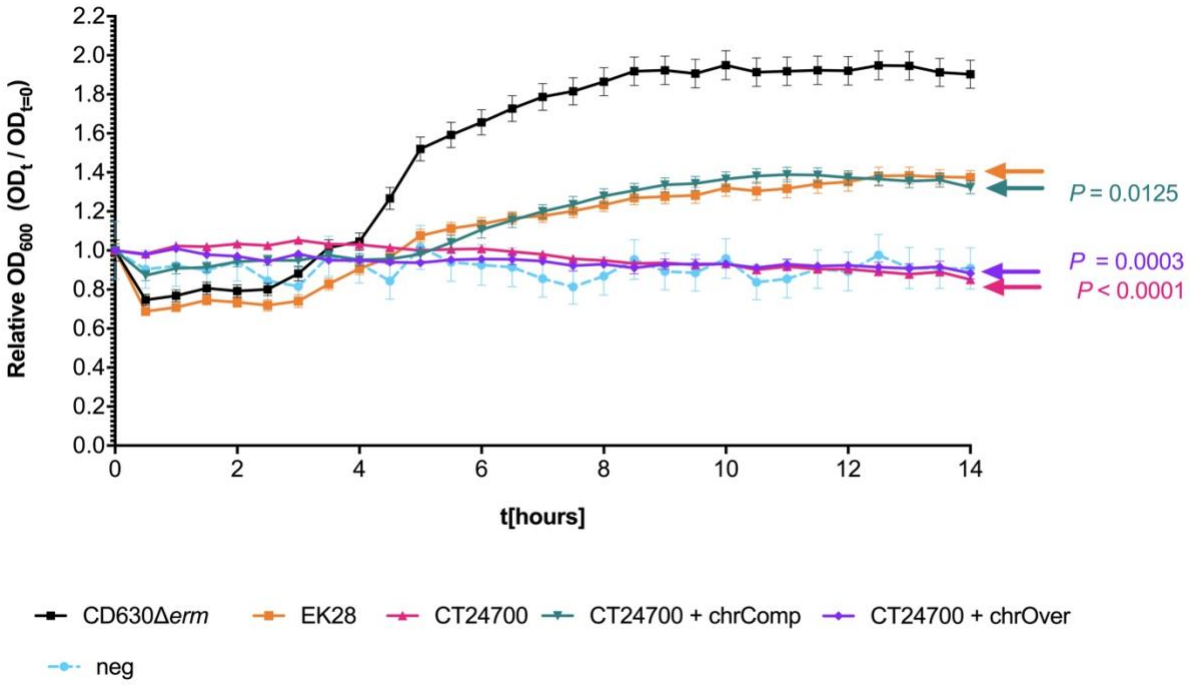


Figure 5.14: Spore outgrowth of CT24700, chromosome based complementation and overexpression strains.

Cell growth from spores of CT24700 during the incubation of BHIS supplemented with 0. 1% (w/v) taurocholate. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. The baseline adjustment (Relative OD600 (ODt/ODt=0)) was used to normalise the data so that the initial reading for each isolate was equivalent to OD600 of 1.0. A statistically significant difference was found for strains CT24700, CT24700+ chrComp and CT24700+ chrOver , $P \leq 0.0001$ and $P = 0.0125$ and $P = 0.0003$, respectively, in comparision to control strain, CD630Δ*erm*.

Both control strains, CD630Δ*erm* and EK28, have the shortest lag phases of all strains that show growth profiles, indicating a higher proportion of spores initiate outgrowth. The CD630Δ*erm* control strain reaches exponential growth phase quicker than the manipulated strains and EK28, and so is able to produce significantly higher levels of bacteria as expressed as Relative $OD_{600.}$ EK28 (see purple arrow in Figure 5.14) reaches exponential growth phase only slightly slower than that of CD630 Δ*erm* yet reaches stationary phase with significantly lower levels of bacteria, indicating that a smaller proportion of spores outgrow into viable vegetative cells of the EK28 strain compared to CD630Δ*erm*. The outgrowth of the complemented strain (see green arrow) has a much longer lag phase than that of the control strains but reaches comparable levels of vegetative cells as EK28. This suggests that the complemented strain takes longer to outgrowth than control strains, suggestive of a smaller population of spores being able to outgrow into viable vegetative cells. Interestingly, spores of the inactivated *gpr* and over expressed *gpr* strain did not outgrow within the time frame of this study (see pink and purple arrows, respectively). Indicating that the levels of GPR do indeed affect the ability of spores to outgrow following germination.

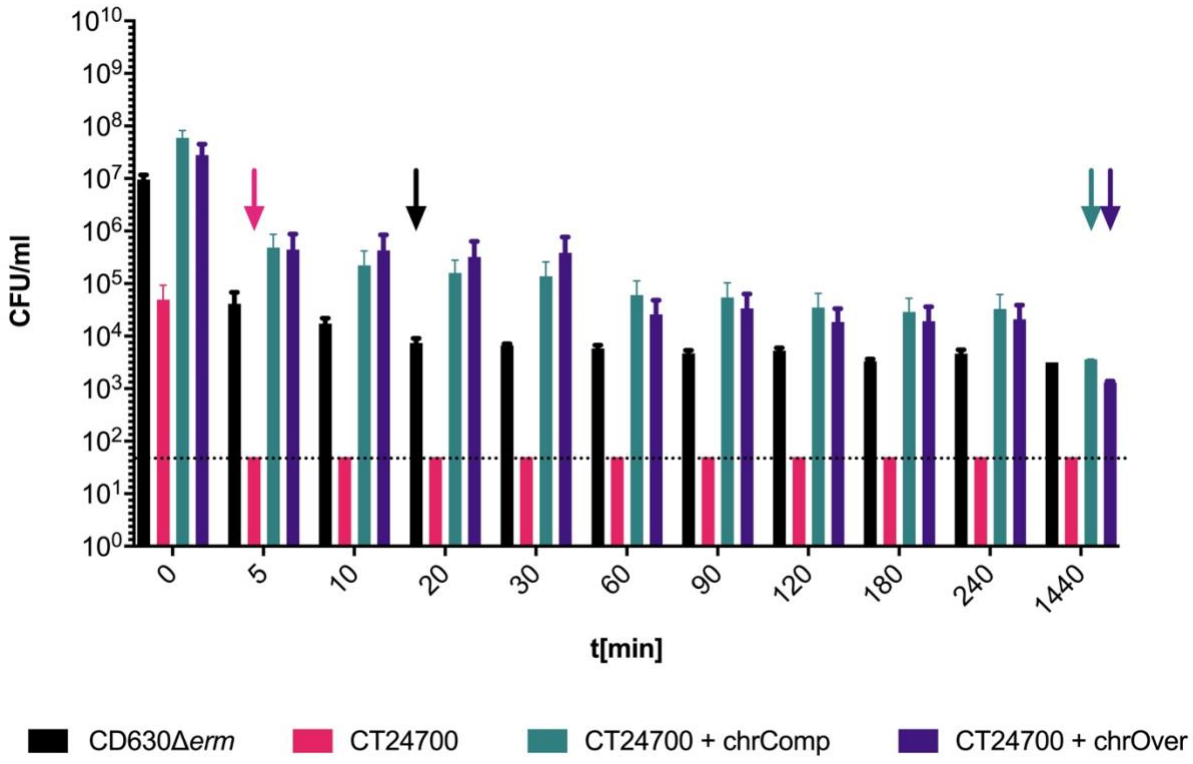


Figure 5.15: Loss of heat resistance in spores of CT24700, chromosome based complementation and overexpression strains.

Spore suspensions were incubated with germinant for permitted time point, where they were heat treated to kill any germinated or out grown cells. Any HR spores remaining at this time point continue to be allowed to germinate and out growth on BHIS agar supplement with 0.1% Tc with 24 hrs where CFU/ml was determined. The detection limit was 50 CFU/mL, as illustrated by the dotted line. All data represent the average of three independent experiments of three independent spore stocks and error bars indicate the standard error of the means. CT24700 and CT24700+ chrComp $P = 0.0222$ and $P = 0.0326$, respectively.

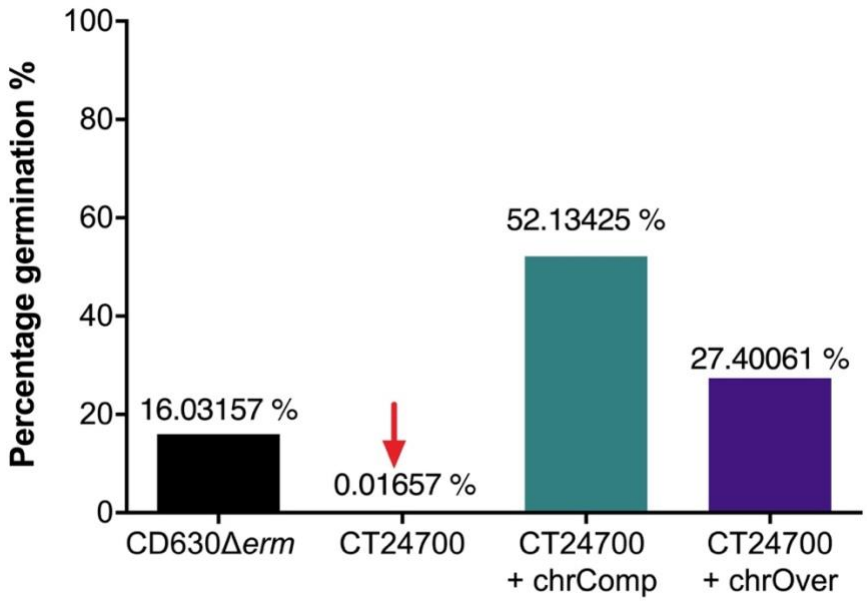


Figure 5.16: Percentage spore germination of CT24700, chromosome based complementation and overexpression strains.

The percentage germination of spores determined by spore counts/ml of a single OD 1.0 spore suspension in ratio of CFU/ml determined at T0 of the loss of heat resistance assay in Figure 5.15.

Figure 5.15 suggests that of all CT24700 spores that remain at 5 mins (and those that can efficiently outgrow) have germinated and done so within 10 mins. However, on quantifying the CFU/mL, growth of vegetative cells did not become apparent until between 24-48 hrs incubation on 0.1% Tc plates. This indicates that outgrowth of CT24700 is significantly delayed and reduced in the measure of HR-CFU development ($P = 0.0222$) and similarly supported by lack of outgrowth (highly significant difference, $P < 0.0001$) within 14 hrs, as measured by a drop in OD_{600} (see Figure 5.14) showing assay reliability. The hypothesis that germination was also delayed was confirmed by two methods of detection, contradicting the implication of no HR-CFU being observed after 5 mins within the loss of HR assay. Firstly, the measure of germination efficacy revealed minute proportion of spores germinating (0.01657%), as indicated by the red arrow in Figure 5.16, poor germination efficacy was also supported with the delayed (limited detection of) initiation of germination as measured by the drop in OD method (see Figure 5.13). The fact that over $1x10^3$ spores/mL remain at 24 hrs (see of Figure 5.15) for all strains except CT24700 (see pink arrow), proves a level of spore dormancy and that this dormancy is affected by the presence of an active GPR to allow efficient outgrowth.

There is a consistent proportion of HR WT spores of control strain CD630Δ*erm* remaining (~$1x10^4$) at each of the 8 time points across 20 mins to 1440 mins (see Figure 5.15), indicative of the total number of spores that cannot germinate despite further incubation with germinant. Whilst there is actually minimal decline in HR WT spores remaining, the consistency supports that there is still a significant proportion of spores that do not go on to germinate. This consistent proportion of non-germinating spores (HR-CFU) is somewhat similarly shown for the complement strain. However, the gradual yet apparent decline supports the hypothesis of lower expression of GPR reducing the rate of germination compared to the control strain CD630Δ*erm*. Interestingly, for the overexpression strain, staggered germination events are seen at T_5, T_{60} and T_{24}.

5.2.8 Recombinant GPR production

It was decided that protein studies would support phenotypic characterising of the germination protease and so production of the recombinant protein was carried out. The expression vector pET-28b(+) was chosen for the IPTG induction of recombinant GPR expression and subsequent purification via nickel affinity to 6xHis-tags. Confirmation of this vector construct carrying the correct GPR protein coding sequence and His-tag position was confirmed by diagnostic digest with Blp1 and Xho1 (see Figure 5.17 lane 2) and products confirmed by Sanger sequencing with primers spanning the T7 promoter sequence. Constructs were transformed into *E. coli* expression hosts BL21 (DE3*) and expression of recombinant GPR induced with 1mM IPTG. Levels of expression were confirmed by SDS-PAGE and Western blot (see red arrows in Figure 5.18) and finally purified (see red arrows in Figure 5.19), respectively, to a concentration of 510 µg/ml as determined by BCA assay.

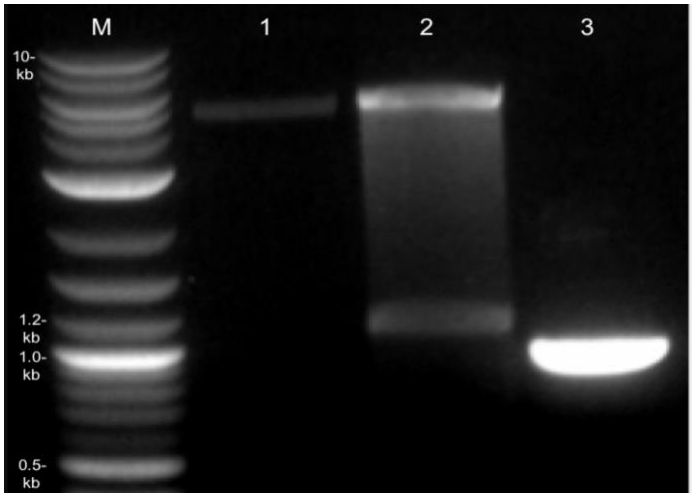


Figure 5.17: Verification of pET28b+ plasmid carrying protein coding sequence of CD24700.

For verification, double diagnostic digest using enzymes BlpI and xbaI were used. M: 2-log ladder, Lane 1: pET28b(+); lane 2: pET28b(+)::CD24700 (954 bp); lane 3: CD24700 assembled with NEB HiFi primers overlapping MCS region of pET28b(+). Expected band sizes are indicated to the left of the gel.

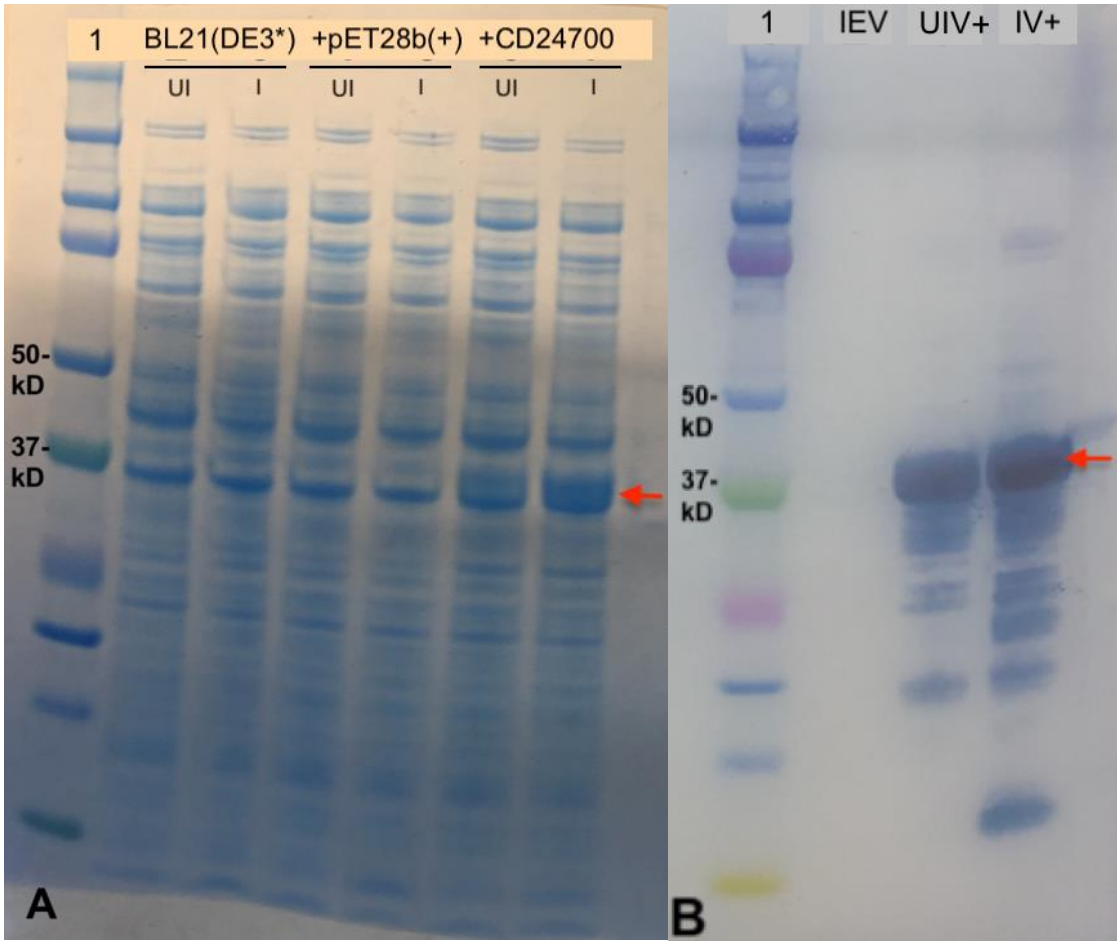


Figure 5.18: Recombinant GPR expression in pET28b(+) vectors and *E. coli* BL21(DE3*) host.

(A) SDS-PAGE of induced and un-induced GPR expression and empty host and vector controls: lane 1= Protein Standards ladder, lane 2= un-induced E. coli BL21(cells) cells, lane 3 = IPTG induced E. coli BL21(DE3) cells, lane 4= un-induced pET28b(+) E. coli BL21 cells, lane 5 = IPTG induced pET28b(+) E. coli BL21(DE3) cells, lane 6= un-induced CD24700:pET28b(+) E. coli BL21 cells, lane 7 = IPTG induced CD24700:pET28b(+) E. coli BL21(DE3) cells. (B) Western blot of induced and un-induced GPR expression and empty vector control, using a Penta-His antibody (Qiagen): lane 1 = Protein standards ladder, lane 2 Induced empty vector control (IEV) = IPTG induced pET28b(+) E. coli BL21(DE3) cells, lane 3 Uninduced vector with CD24700 (UIV+) = un-induced CD24700:pET28b(+) E. coli BL21 cells, lane 4 Induced vector with CD24700 (IV+) = IPTG induced CD24700:pET28b(+) E. coli BL21(DE3) cells. Western blot positive control shown in Figure 5.19 - B. Expected band sizes are indicated to the left of the gel.

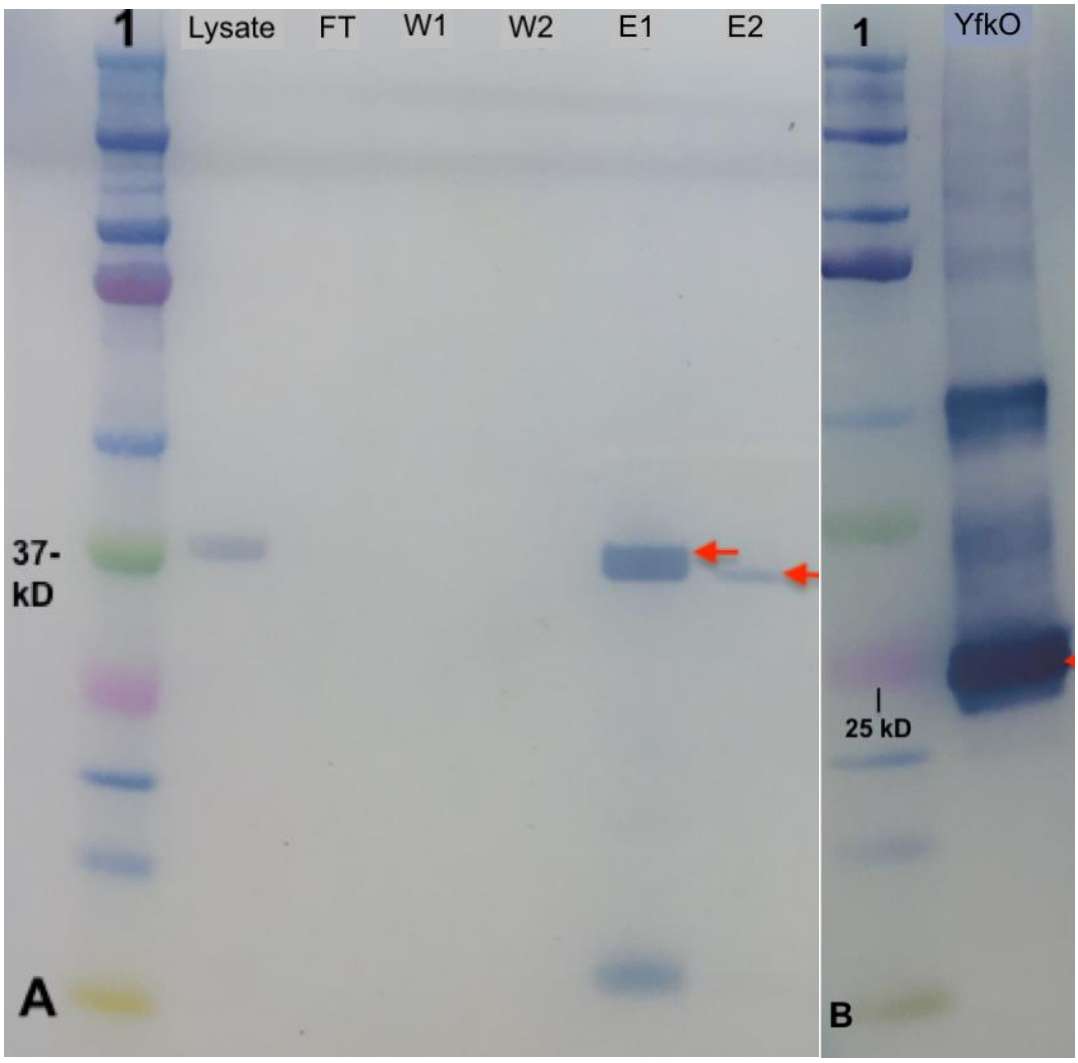


Figure 5.19: Ni-NTA column His-Tag purification of recombinant GPR protein.

Western blot of purification fractions and YfkO positive control, using Penta-His antibody (Qiagen). (A) Ni-NTA column purification, under denaturing conditions; Lanes 1 Protein standards ladder, Lysate: cell lysate, FT: flow through, W1: wash 1, W2: Wash 2, E1: Elute 1, E2: Elute 2. (B) Positive western blot control; Lane 1: Protein Standards ladder and YfkO: YfkO protein (~25 kDa). Expected band sizes are indicated on the gel.

5.2.9 Identifying Small Acid Soluble Proteins of *C. difficile*

Table 5.2: SASPs of *C. difficile*

ORF	Annotation (KEGG)	Gene (bp)	Protein (aa)	MW (kDa)	pI (pH)	Most homolgous gene* (annotation, ORF, genome, % ID)
CD12900	Putative small acid-soluble spore protein SASP	201	66	7.522	6.57	Hypothetical protein, CPE1424, CPE, 87%
CD26880	sspA	222	73	8.051	9.10	*sspA*, CA_C2365, CAC, 72%
CD32201	small acid-soluble spore protein	255	84	9.580	9.77	Small acid soluble, protein beta, CPE1903, CPE, 77%
CD32490	sspB	204	67	7.392	8.01	*ssp1*, CPE2161, CPE, 72%
CD24700	*gpr*, Germination protease	954	317	34, 957	4.76	*gpr,* CPE2041, CPE, 68% (Query coverage = 77%) *gpr*, CA_C1275, CAC, 67% (Query coverage = 65%)

*homologous genes revealed in blastn alignments to organism set search of *B. subtilis* 168 (BSU), *C. acetobutylicum* ATCC 824 (CAC) and *C. perfringens* strain 13 (CPE) genomes.

To investigate an enzyme-substrate interaction between recombinant GPR and hypothesised small acid soluble proteins in *C. difficile* listed in Table 5.2, a crude extraction of acid soluble spore proteins was performed. Proteins were extracted from 1 mg of pure *C. difficile* spores using an acetic acid-based extraction protocol adapted from Li *et al.,* (2009) and to maintain their activity (basic properties) were then analysed by cathodic PAGE. To establish the cathodic PAGE method and a suitable basic protein positive control, a gradient of lysozyme concentrations was loaded and run by cathodic PAGE (see red arrow in Figure 5.20). A 4mg/ml concentration of lysozyme was then deemed suitable and subsequently loaded as a control alongside crude acid-soluble spore protein extracts to confer presence of SASPs in *C. difficile* spores (see Figure 5.21 - A). As a faint smear was observed for the loading of the extract by coomassie and silver staining techniques (see red arrow in Figure 5.21 - B), centrifugal

concentration of the sample was employed to visualise discrete banding. This method concentrated the sample from 60 µg/ml to 1.265 mg/ml as determined by BCA assay. Folowing the silver staining of this concentrated sample, four additional discrete bands were revealed (see Figure 5.22). Unfortunately, a basic protein ladder was not accessible at the time of experimentation and so the sizes and molecular weight of acid soluble proteins within spore extracts were not estabilished. However, on-going work to confirm the prescence of SASPs by Mass spectrometry on the 5 indicated bands excised from the silver stained gel (see red and blue arrows in Figure 5.22 - B), was to be completed as future work outside of this research.

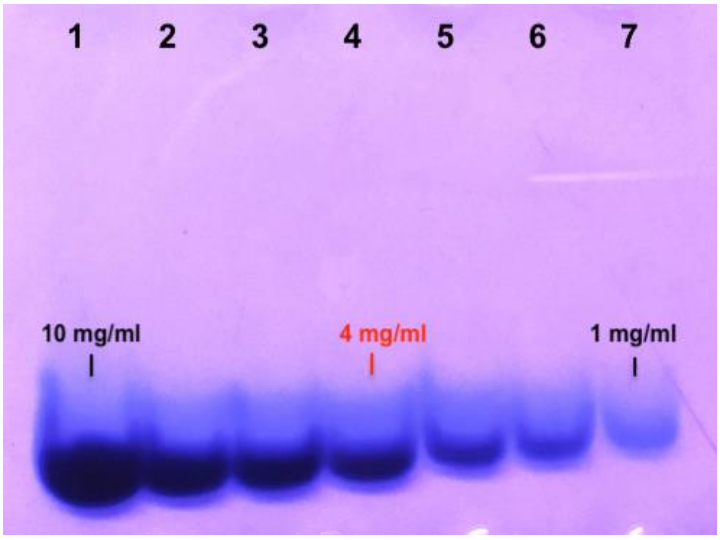


Figure 5.20: Lysozyme controls for the validation of Cathodic PAGE gel to visualise basic proteins.

A gradient of 10 mg/mL – 1 mg/mL Lysome (Lanes 1-7) was analysed by cathodic PAGE.

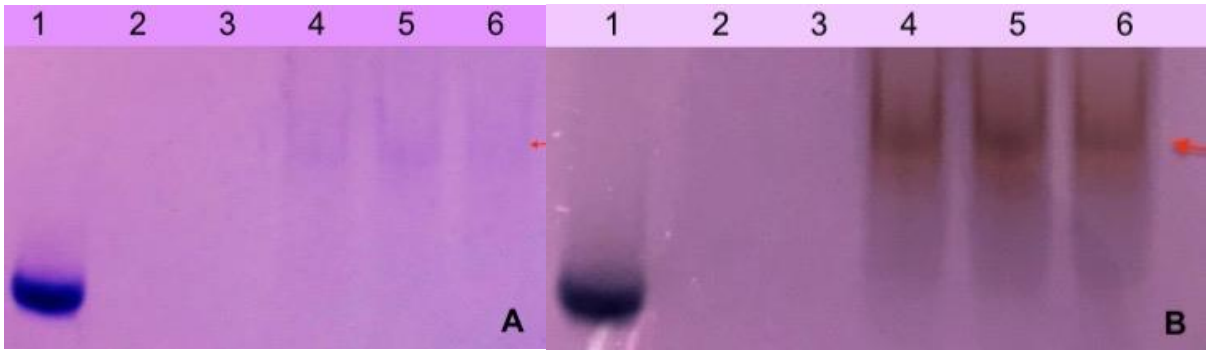


Figure 5.21: Crude acid-soluble spore protein extracts visualised on cathodic PAGE gel.

(A) Lane 1 = 4 mg/ml lysozyme, Lanes 2-3 = blank, Lanes 4 -5 = Crude acid-soluble spore protein extracts. (B) Coomassie and silver stain of gel (A).

Coomassie staining of crude acid-soluble protein spore extracts run on Cathodic PAGE reaveled faint yet apparent smears and were more obvious with the further processing for silver staining with higher senstivity than coomassie technique. The total protein concentration of crude acid-soluble protein spore extract was determined to 60 µg/mL by BCA assay. As no discrete bands were observed for extracts when loading the gel at this concentration by either staining methods, the sample was concentrated by YMCO-3000 filter centrifugation (see Figure 5.22). A more intense band was observed by coomassie staining however faint banding smears remained. Following silver staining several faint bands became more discrete. The five bands indicated on silver stained gel were excised and processed for mass spectrometry analysis for future work within the research group, data not shown.

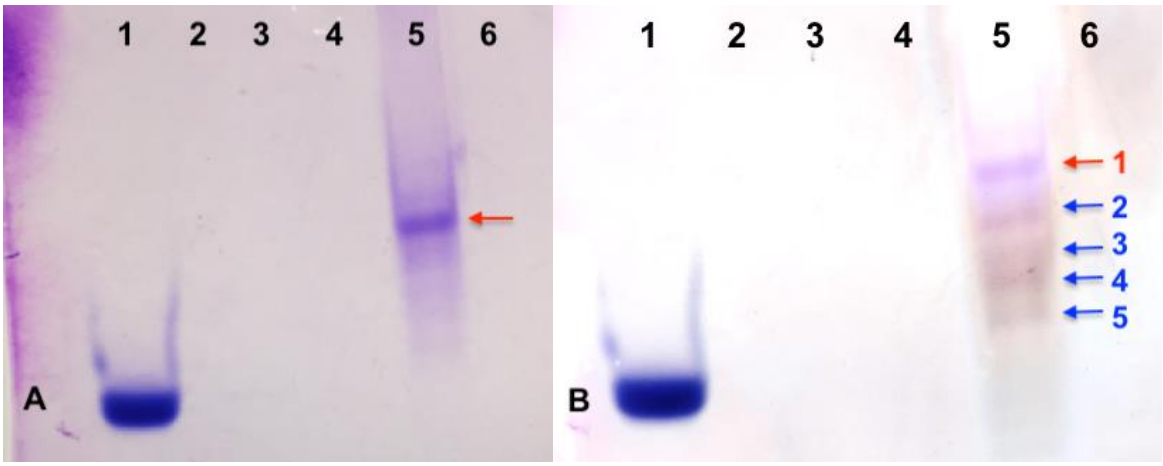


Figure 5.22: Acid-soluble spore protein extracts concentrated by YMCO-3000 filter centrifugation visulaised on cathodic PAGE gel.

(A) Coomassie destained for 48 hrs to permit clarity of banding by removal of background stain lanes 1: 4 mg/mL lysozyme, 2 - 4 and 6: blank and 5: crude acid soluble spore protein extract concentrated by centrifugation. (B) silver stain of gel (A).

5.2.10 Assessing the proteolytic activity of recombinant GPR and the extent of DNA protection from DNAse acid soluble spore proteins

The autoprocessing of GPR from the inactive to the active form can be inferred by incubation in autoprocessing buffer at 30 °C as described in section 2.8.10. Analysis by SDS-PAGE ought to reveal an additional smaller band, such as that seen in *C. acetobutylicum* of 35.5 kDa along with the larger at ~ 36.5 kDa (Wetzel and Fischer, 2015). However, Figure 5.23 show only faint single bands (Lanes 1-2: blue arrows), when assaying the autoprocessing of 3µg GPR in a 100 µl reaction. A ratio of mass concentrations of 10:1 and 20:1 GPR:SASP (4 ug) was previously proven to show a difference in the protection of DNA from DNAse by SASPs of *C. acetobutylicum* (Wetzel and Fischer, 2015) and so a range in ratio was assayed during this study. Due to the total volume constraints of acid soluble spore protein extracts, ratios of 5:1 (20 µg GPR: 4 µg crude extract), 50:10 (20 µg GPR: 4 µg crude extract in a 10 µL reaction) and 1:6 (20 µg recombinant GPR: 126.5 µg concentrated extract) were assayed. The proteolytic activity of GPR against SASPs can be inferred by when incubations with GPR and autoprocessing buffer at 37 °C over 12 -24 hours are stopped with the addition of SDS and visualised by SDS-PAGE followed by silver staining as described in section 2.8.12.

Interestingly, Figure 5.23 shows with the addition of SASPs, when the concentration of recombinant GPR is increased to 20µg, banding at ~37 kDa appears during the 100 µL reaction after 12 h (Lane 6: blue arrow) and after 24hr incubation a thicker band is also revealed (Lane 7: red arrow). The concentration of GPR in the 100 µL reaction was increased 10-fold, when a ratio of 50:10 was used, resulting in more intense banding (~ 37 kDa) produced after the 12h time point compared to the prior reaction (Lane 8 blue arrow) and again after the 24h incubation with SASPs (red and blue arrows in Lane 9). Two discrete bands are revealed at ~ 37 kDa when a 1:6 ratio of GPR: SASPs is incubated after 12h (Lane 10: red and blue arrows) and 24h (Lane11: red and blue arrows) indicating autoprocessing activity of recombinant GPR. Smears of banding at < 25 kDa appear during the addition of small acid soluble protein extracts (lanes 7-11), specifically < 10 kDa (lanes 10-11: red arrows) however no inference can

be made as to whether these are SASPs as no MW marker was used during the Cathodic-PAGE gels and Mass spectrometry data was not yet available.

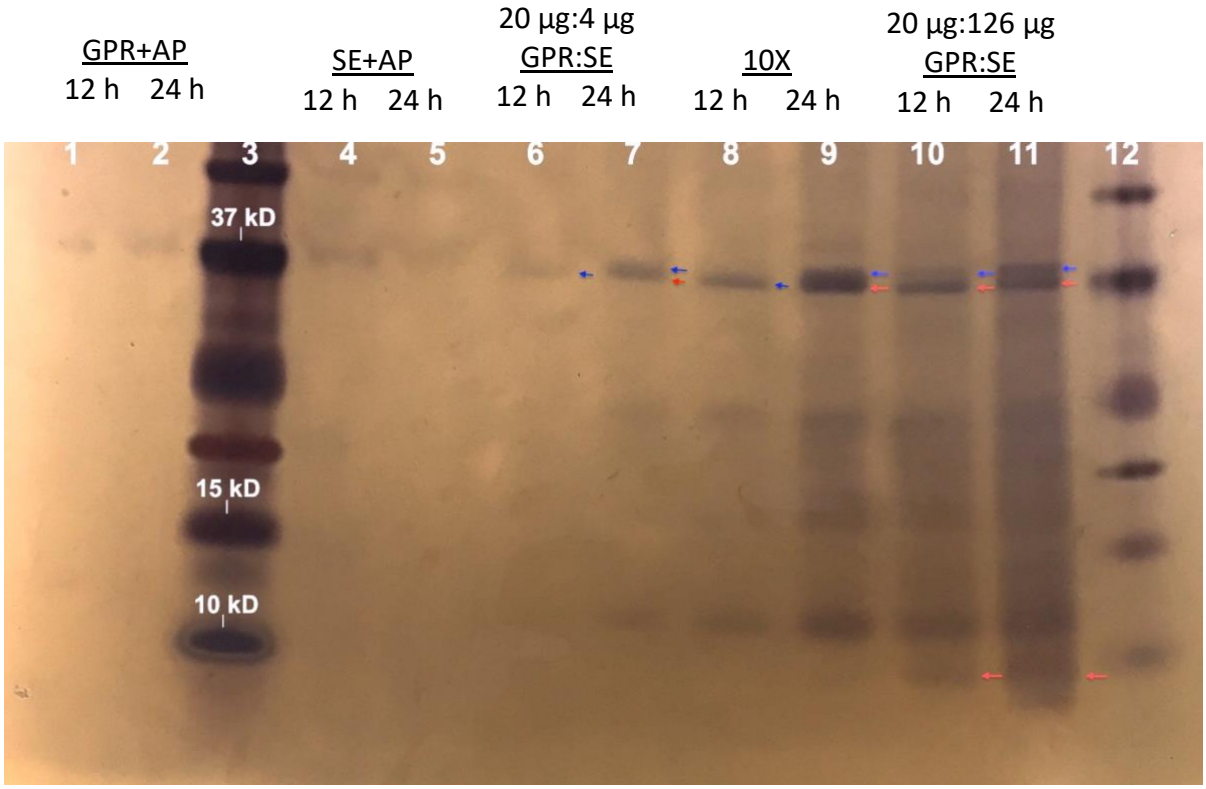


Figure 5.23: Auto-processing of recombinant GPR and its proteolytic activity on acid soluble spore protein extracts visualised by cathodic PAGE.

Recombinant GPR (3 µg) was incubated with water or auto-processing (AP) buffer for 30°C overnight lanes 1 and 2 respectively. Lanes 3 and 12: protein standards ladder. Aliquots of 4 µg crude acid-soluble protein spore extracts (SE) were incubated with autoprocessing buffer at 37°C for 12 and 24 hours (lanes 4 and 5 respectively). Recombinant GPR was added at a mass ratio of 5:1 (20 µg: 4 µg) and incubated for 37°C in autoprocessing for 12 and 24 hrs (lanes 6-7 respectively). The concentration of GPR:spore protein extract in the 100 µl reaction was increased 10-fold (50:10) to promote detection of bands and again incubated at 37°C in autoprocessing buffer for 12 and 24 hrs (lanes 8-9, respectively). Aliquots of 126.5 µg concentrated extract were incubated with 20 µg recombinant GPR to give a ratio of 1:6 and incubated at 37°C in autoprocessing buffer for 12 and 24 hrs (lanes 10-11, respectively). A breif guide to each reaction is indicated above the gel.

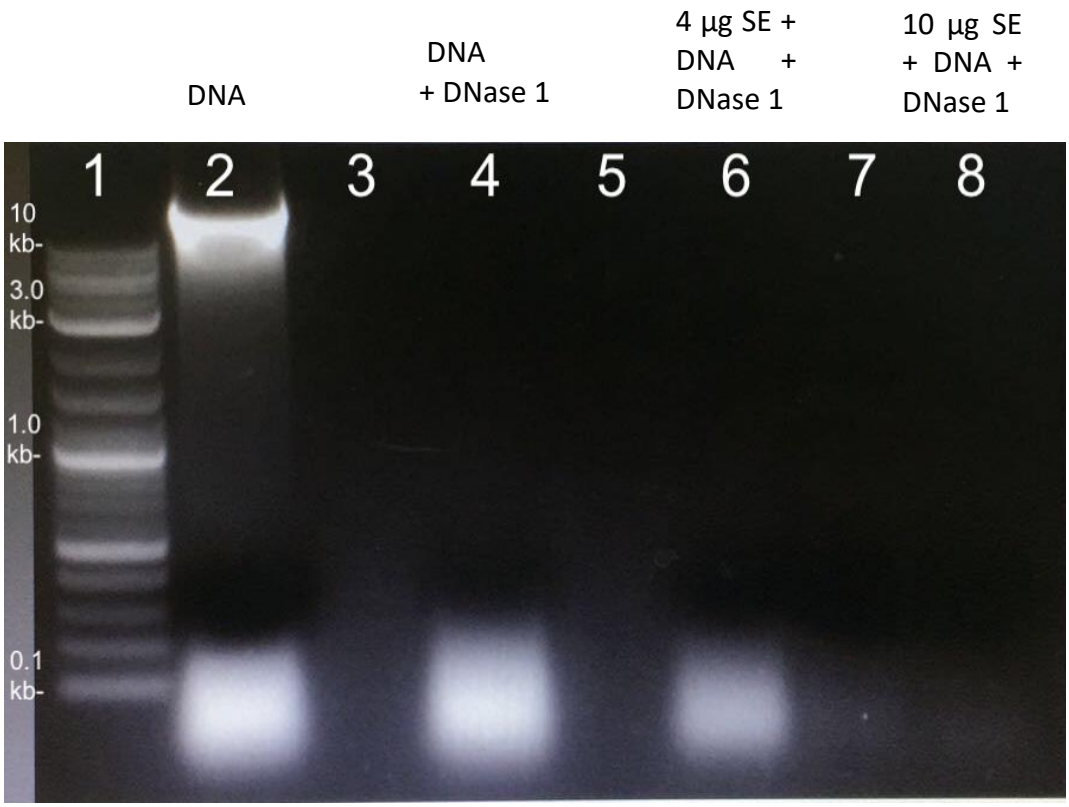


Figure 5.24: No protection of DNA by acid-soluble spore protein extracts from DNASE I treatment.

Lane 1: 2-log ladder, lane 2: 1 µg/ml chromosomal *C. difficile* 630 DNA in water, lane 4: 1 µg/ml chromosomal *C. difficile* 630 DNA + DNase I, lane 3 and 5: blank, lane 6: 1 µg/ml chromosomal *C. difficile* 630 DNA + DNase I + 4 µg crude acid soluble protein extracts, lane 7: 1 µg/ml chromosomal *C. difficile* 630 DNA + DNase I + 10 µg crude acid soluble protein extracts and lane 8: 1 µg/ml chromosomal *C. difficile* 630 DNA + DNase I + 20 µg crude acid soluble protein extracts. A breif guide to each reaction is indicated above the gel.

The protection of chromosomal *C. difficile* DNA from DNase degradation by SASPs in acid-soluble protein spore extracts was assessed as described in section 2.8.11. No protection from DNase I by extracts at either 4 µg, 10 µg or 20 µg was observed in the 100 µL reaction (see Figure 5.24).

5.3 Discussion

5.3.1 Key outcomes of ClosTron mutagenesis

In this study, we describe the initial characterisation of putative genes (CD19300, CD24700, CD05460, CD08670 and CD29630) in germination and sporulation pathways of *C. difficile* 630Δ*erm* through use of ClosTron mutagenesis. Heat-treated spores of the *gpr* mutant (CT24700) displayed a significantly delayed and reduced development of colonies when germination was induced with taurocholate and showed no detectable recovery of decoated spores with lysozyme. Inactivation of hypothetical protein genes (CD19300, CD05460, CD08670 and CD29630) showed a reduction in heat-resistant colony formation, but to a level of no significant difference as determined by the uncorrected Dunn's test and formed parental levels of colonies after spore decoating. Initial characterisation of these ClosTron mutants suggests that the germination protease (*gpr*) plays a role in the development of viable spores, whilst the hypothetical proteins, CD19300, CD05460, CD08670 and CD29630, may have a role in germination and/or vegetative cell outgrowth in *C. difficile* 630Δ*erm*. The level of statistical significance for CT24700 in initial spore phenotype assays when compared to the CD630Δ*erm* control strain led to further characterisation studies.

Complementation of *gpr* inactivation revealed comparable trends to WT in the rate of sporulation, spore counts and spore viability, indicating that the phenotypes observed were in fact due to the *gpr* inactivation. Spores of CT2700 showed an extremely limited rate of spore germination initiation. However, the complemented strain showed unexpected results during the initiation of germination experiment. The overexpression of *gpr* strain has the fastest decline in OD_{600}, indicative of spores completing germination much quicker when *gpr* expression is under the control of a stronger promoter. However, until a return to WT initiation of germination phenotype can be demonstrated with complementation strain, such conclusions of its involvement in germination remain to be definitive. Interestingly, spores of the inactivated *gpr* and over expressed *gpr* strain did not outgrow within the time frame of

this study (see pink and purple arrows, respectively). Indicating that the levels of GPR do indeed affect the ability of spores to outgrow following germination.

5.3.2 Role of germination protease in *C. difficile* spore germination and outgrowth

Following difficulties trusting such disparities between lab strain CD630Δ*erm* and clinical isolates observed in Chapter 4:, a second positive germination control was used to illustrate that differences are observed between lab strains also. Rates of germination for control strains CD630Δ*erm* and EK28 are comparable. Spores of control strains presented the largest drop in OD_{600} observed using this assay indicating that a higher proportion of spores complete germination than that of CT24700, complemented and overexpression strains. Rates of germination initiation for control strains follow trends observed by Bhattacharjee *et al.,* (2016), where the initiation of spore germination as measured by a drop in Relative OD_{600} from 1.0 to ~ 0.7 within 10 mins. Likewise, this rate and comparable maximal drop in OD_{600} was observed for the GPR overexpression strain, whilst it reached a maximal drop more closely related to EK28, compared to CD630Δ*erm*. This leads to the inference that spores complete germination much quicker when GPR expression is under the control of a stronger promoter. This hypothesis could be extended to that of control strains, where the expression of GPR may have been upregulated in comparison to the GPR mutant and in clinical isolates, with EK28 (also a clinical isolate) expressing more GPR than CD630Δ*erm*. Conversely, this would suggest that spores with inactivated GPR and likewise for spores of clinical isolates could be assumed to have down regulated levels of GPR. In this instance, the hypothesis being that this would permit spores to remain dormant for a longer time, that may have been limited to the time constrains employed by this assay. This assumption could be further speculated by the results of the loss of heat resistance assay within this study. Evenso, to prove or disprove this hypothesis studies of *gpr* expression levels such as by RNAseq, would need to be employed.

However, this hypothesis is complicated by the fact that the CT24700+*chr*Comp strain doesn't complement the initiation germination phenotype to levels observed of WT spores, despite numerous replicates being tested. The native promoter sequence of *gpr* was identified using the Softberry BPROM tool (Solovyev, 2011). It may be that the promoter sequence used didn't include all and maybe errors were present in the sequences of transcription factor binding sites within the promoter as that of the complete version of the lab reference strains. Indeed, the automated prediction tool is limited to an 80% accuracy and specificity. Therefore, errors in binding sites of activators or repressors of GPR transcription within the promoter sequence may have affected the rate of transcription, thus affecting expression. Additionally, the chromosomal location of genes can influence levels of expression. The position of CD24700 is flanked by CD24690 (spore stage II protein, SpoIIP), CD24710 (conserved hypothetical protein) and CD24720 (conserved hypothetical protein (ArsR transcription regulator)), which have all been found to be upregulated during spore germination in transcriptome studies (Dembek *et al., 2*013). The complementation at the *pyrE* locus could have had an influence on the levels of *gpr* expression. It could therefore be deduced that the usual transcription of *gpr* in WT spores was to a higher level than that of the promoter sequence and chromosome locus used in the complement strain CT24700+*chr*Comp. As it is established that *gpr* doesn't play a role in sporulation detectable by assays employed in this study, complementation of initial spore phenotype screens showed comparable but reduced levels of CFU/mL to WT in the rate of sporulation and in viability of spores determined by ability to outgrowth after lysozyme treatment. When using CFU formation as the measure of a sporulation phenotype, the assays employed are immediately presented with contradictions, as it is actually the measure of the ability of vegetative cells to sporulate, withstand heating, then go on to germinate and outgrow on media supplemented with Tc and not the measure of a sporulation phenotype alone. Therefore, it may be argued that whilst the CT24700+*chr*Comp strain complemented the initial sporulation screening phenotype it may not however, have been complemented to produce a gene dosage as comparable to WT as expected when using the promoter sequence selected BROM analysis.

The CD630Δ*erm* reference strain reaches exponential growth phase quicker than that of manipulated CT24700 strains and EK28 (see Figure 5.14), and so represents a higher proportion of spores able to outgrow and produce significantly higher levels of bacteria as expressed as Relative $OD_{600.}$ EK28 (purple arrow) reaches exponential growth phase only slightly slower than that of CD630Δ*erm* yet reaches stationary phase with significantly lower levels of bacteria, indicating that a smaller proportion of spores outgrow into viable vegetative cells compared to CD630Δ*erm*. The outgrowth of the complemented strain (solid red arrow) has a much longer lag phase than that of the control strains but reaches levels of vegetative cells comparable to EK28. This suggests that the complemented strain takes longer to outgrow than control strains, this is supported by a significantly shorter exponential growth phase suggestive of a smaller population of spores being able to outgrow into viable vegetative cells. Continuing the presumption of lower expression of *gpr* in the CT24700+chrComp strain this result could speculatively infer a role of GPR expression in the outgrowth.

Interestingly, spores of the mutant GPR and over expressed GPR strain did not outgrow within the time frame of this study (green and red dashed arrow respectively), despite extensive replicates being assayed. This indicates that the levels of GPR do indeed affect the ability of spores to outgrow following germination. This observation is somewhat expected for the GPR inactivated mutant, where even with reduced rates of germination, a small proportion of spores would be able to allow the intake of water and DPA promoting the eventual disassociation of SASPs to allow a small population of outgrowing vegetative cells. It may be the case that the time to observe this phenomenon via this assay exceeded those employed in this study. However, in the case of the overexpression strain, where levels of GPR ought to be restored so that they produce an outgrowth curve, or even a continuation of the hypothesis where levels would be expected to match that of CD630Δ*erm,* the reverse was true. The failure of spore outgrowth to be measured by this assay for the overexpression of GPR and inactivated GPR, that are obviously detectable using measures of CFU's in other assays suggests an experimental error. Spores of CD630 are notoriously difficult to purify and

possible clumping could have hindered changes in optical density from being measured. The measure of loss of heat resistance by CFU development was therefore assayed to reveal a germination and outgrowth phenotype.

Figure 5.15 suggests that all CT24700 (significant difference, $P = 0.022$) spores that remain at 5 mins and of those that can efficiently outgrow, have done so within 10 mins. However, on quantifying the CFU/ml, growth of vegetative cells did not become apparent until between 24-48 h incubation on 0.1% Tc plates. The fact that over $1x10^3$ spores/mL remain at 24 h (red asterisk) for all strains except CT24700, indicates a level of spore dormancy that is affected by the presence of an active GPR. It is also possible that with impaired ability of GPR to dissociate SASPs from spores, the maximum efficacy of impaired GPR or alternate SASP degrading pathways is reached within 10 mins, in a much shorter time compared to WT. this means a higher proportion of SASPs remain bound to DNA preventing the return to vegetative cell growth of the rest of the partially germinated spore population. That is to suggest that whilst spores may detect germinants and carryout function of early stages of germination this is retarded to such a point that complete/efficient outgrowth is delayed past the limits of this assay. This is supported by the determination of CFU/mL for CT24700 aliquot at T_5, did not become completely detectable until between 24-48 hrs incubation on 0.1% Tc plates, whilst all control strains and clinical isolates in this study were detectable by 24 hrs.

This argument then presents a 'model' of spore dormancy, which however can only be confirmed by the eventual outgrowth and detectable HR-CFU or by spore counts of suspensions at times continuing past 24 hrs. If outgrowth did not occur by a 48 hrs time point, how long would this assay be required to go on for? Recurrent infection is defined as the return to symptomatic disease within 8 weeks of initial treatment completion and indeed recurrence are known to occur longer than that, take for example patients 01008, 01013 and 12001 were recurrences were reported 3 months after the resolution of first infection episode in Chapter 3. Experimental complications for this, however, include the possible crystallisation of salts in the growth media over time (PBS and germinant alone would not support complete outgrowth). Such a barrier could limit the confidence of reliable spore

counts possibly presenting as microscope artefacts, in combination to the turbidity of BHIS solution to potentially mask phase-bright analysis of spores. Moreover, if and how often the media would need to be replaced or agitated as a result, to evenly distribute the availability of germinant to spores. This assay would need further study and modifications to account for such considerations.

Much like the studies of *Bacillus* (Sanchez-Salas *et al.,* 1992), *gpr* inactivation leads to a markedly delayed spore outgrowth in *C. difficile*. The inferred persistence of α/β SASPs beyond stages of normal degradation permits them to remain bound to DNA within the spore core and subsequently interrupts normal DNA transcription depriving the spore of exit from the phase of dormancy. Additionally, the inhibited production of amino acids by GPR cleavage and α/β SASP degradation obstructs protein synthesis during early outgrowth. Again, in parallel to *Bacillus* studies (Sanchez-Salas *et al.,* 1992), in this research germination and outgrowth of spores eventually ensues and the SASPs are somehow disassociated. This phenomenon suggests that an alternative means of SASP degradation is present. Importantly, the early studies of *gpr* mutants in Bacillus implicated that a portion of DNA is SASP-free in fully germinated spores that lack *gpr*, even though there is no SASP degradation (Sanchez-Salas *et al.,* 1992). Allowing then for the transcription of GPR-like protease(s) during germination that ensure the partial degradation of SASP, such as observed in mutagenesis studies of the Clp-like protein of *B. subtills*, *ymfB* and its adaptor *ylzJ* (Traag *et al., 2*013), or *yyaC* (Wang *et al., 2*006). Alternatively, the slow but eventual outgrowth of spore observed both in this study and historically, may owe to the moderate affinity of α/β SASP to DNA (Sanchez-Salas *et al.,* 1992). During early germination, increases in spore core water content may promote eventual dissociation over time, whereby the spore chromosome eventually becomes SASP free, allowing germination and outgrowth to proceed.

5.3.3 Activity of recombinant GPR and SASPs in *Clostridioides difficile*

The auto-processing activity of GPR was only inferred during the proteolysis activity assays on crude acid soluble spore protein extracts. No proteolytic activity of GPR could be inferred by this study as well as no protection of DNA from DNase I treatment by crude acid soluble spore proteins. Unfortunately, assays were carried out only once and with minimal stocks of spore protein extracts. The selection of concentrations and inappropriate dilutions in 100µl buffer reactions means the failure to detect possible activity of recombinant GPR or crude spore protein extracts during this study cannot be ruled out. Use of the YM-3 centrifugal concentrator, whilst revealing more discrete banding of spore protein extracts run on cathodic-PAGE, may have promoted the loss of proteins at the target MW of SASPs. Additionally, work by Fimlaid *et al.,* (2013), where expression of *gpr* is used as a reporter for sigma factor F in *C. difficile* used a specific antibody to GPR, whereas in this study an anti-His antibody was used during Western blot analysis. Use of a specific antibody to the desired protein would had given a more reliable indication towards the level of purity in eluates, as well as an additional size exclusion column of purified eluates to increase purity. As during recombinant protein purification, the contamination of mixed species, due to potentially leaky expression of T7 promoter, could had also reduced concentrations of recombinant GPR protein in downstream experiments. Future work would need to include, confirmation of recombinant protein purity as well as repeat activity assays with modified concentrations and reaction volumes, including use of a MW marker for basic proteins.

5.3.4 Concluding remarks

Finally, this study shows that the germination protease is implicated in having a role in spore viability, outgrowth and the ability of spores to remain dormant, as shown by inactivated spores showing no recovery after decoating and no measurable outgrowth of spores, as well as the complete loss of heat-resistance compared to the control strains using experimental

methods in this study. However, the hypothesis made relating to role of *gpr* in the germination of *C. difficile* spores would need to be further explored as complementation of the inactivated mutant strain did not return the phenotype to WT levels. Unfortunately, the role of recombinant GPR for the degradation of SASPs cannot be clearly concluded from this study, due to experimental limitations, repeat assays with modifications in concentrations of protein in activity assays should be made. Moreover, the conclusion of small acid spore proteins having no protective effect on DNA protection from DNase cannot be inferred as experimental limitations of protein concentration were also encountered, repeat assays would need to be carried out to address this. Attempts to produce recombinant CD12900, CD26880, CD32201, CD32490 and YyaC were aborted due to Sanger sequencing revealing errors in cloned pET vector constructs that would allow the incorrect amino acid coding sequence to be driven in-frame to the T7 promoter. Additionally, attempts to create in-frame deletions of SASPs genes were aborted when double crossover clones for the in-frame deletion of CD12900, CD26880 and CD32201 all returned as wild-type by Sanger sequencing.

Research targeting the association and degradation of SASPs remains important in understanding both the resistant properties and germination characteristics of *C. diffcile* spores. Such research could see the development of translatable prophylactic CDI interventions; whereby spores remain dormant indefinitely by ensuring SASP remain permanently bound to DNA or modifications of GPR that induce early degradation and germination permitting clearance of vegetative cells by cleaning and antibiotic agents. However, with varying germination and sporulation characteristics amongst strains of *C. difficile*, there may be other unknown factors that will interfere with a uniform prophylactic measure against all *C. difficile* types. Justifiably then the need for detailed molecular understanding of the sporulation and germination process remains a valid concern.

Chapter 6

Chapter 6: General Discussion

6.1 Introduction

The genome of *C. difficlie* 630, the model strain of *C. difficile*, was sequenced in 2006 (Sebaihia *et al.,* 2006). Following this, the proteomic analysis of *C. difficile* 630 spores (Lawley *et al.,* 2009) identified a total of 336 spore-associated polypeptides, 39 of these being predicted to be secreted or surface associated. However, many of these are encoded by genes that have no predicted function in the genome annotation. Nine percent of the spore proteome consists of spore proteins encoded by genes that are currently unique to *C. difficile*. Many of these are hypothetical proteins that can be preferentially targeted in mutant construction, as carried out for a small set of hypothetical targets in this study. There have since been significant advances in the understanding of sporulation and germination pathways of *C. difficile* (Kochan *et al.,* 2018).

To establish a link between germination/sporulation phenotypes and CDAD, a study of recurrence in clinical isolates, using genotypic and phenotypic analysis of strains responsible for CDI in the Nottingham area, and those involved in relapse was conducted. It was hoped that together these studies would provide insights into the genomes of *C. difficile* involved in recurrence and the genetic and phenotypic variation related spore phenotypes. The overreaching goal was the identification of genetic variants that correlated with phenotypes that would have implications on infection behaviour. Specific research aims included

- To undertake molecular and genetic studies to define and understand the processes that control the initiation of sporulation and germination.
- To perform functional assignments to hypothetical proteins in sporulation and germination pathways of *C. difficile* in an effort to complete the regulation cascade.
- To establish a link between germination/sporulation phenotype and relapse of clinical isolates.

6.2 Key findings, limitations and proposed follow-on work from this study

1. Rifaximin as a chaser therapy protects patients from CDI recurrence

Within this thesis a clinical trial study for use of rifaximin as a chaser therapy for CDI recurrence was assessed. Statistical analysis was limited for this clinical trial study. With variables including patient age, sex, duration of symptoms prior to treatment initiation, CDI severity and presence of prescribed concomitant antibiotics not collected. The lack of interrogation means there was no available analysis of these variables to identify clinical factors associated with recurrence in this cohort and whether the use of rifaximin alleviated these potential risk groups. Nevertheless, data from this study indicates that chaser therapy with rifaximin is preventative of recurrence of CDI. It is hypothesised that any residual bacterial loads or uncleared spores or biofilms would lead to a risk of clinical recurrence, in that the presence of toxins and pro-inflammatory response would give rise to CDI pathology (Kuehne *et al*., 2010). However no information was available for symptom duration following the time points between diagnosis and week 12 for these patients to link the inference of prevention of microbiological reinfection or relapse, that is to say the strains isolated in the study that appear to have persisted giving rise to symptomatic CDI. Additionally, no details on disease severity or comorbidity were collected and so the effect of compromised immune response on the susceptibility of patients to recurrence could not be inferred. That is to question whether patients who appear to have been protected by rifaximin could have also been by less immune challenged compared to the placebo group, or whether the groups differed in mild or severe CDI, or symptomatic relapses being a confounded result of chronic diseases such as Irritable bowel syndrome. With a burden of over-antibiotic use worldwide, such information would be useful in guiding clinicians of chaser therapy depending on the severity of CDI, or immune status of the patient.

2. Spore germination of clinical *C. difficile* isolates is more tightly controlled in comparison to the reference strain CD630Δ*erm.*

Using the loss of optical density method to measure the initiation of spore germination in Chapter 4, it was found that all clinical isolates involved in both recurrent and non-recurrent

CDI were significantly different to control strains. Findings include significantly smaller slopes at the beginning of the experiment and higher end point OD at the end of the experiment for all clinical isolate strains. Together these results indicated that a lower proportion of spores initiate germination and of those that do, do so significantly slower than the control strains.

Linking phenotypic findings to that of genotypic data was limited mainly due the unequal sample size and lack of representative strains involved in non-recurrent CDI during genomic analysis of Chapter 3. Nevertheless, of the 23 genes of interest selected for analysis of genetic variation in strains of recurrent CDI, a higher level of sequence variation was found in genes encoding, exosporium proteins BclA1 and BclA3, the fusion gene *cspBA*, encoding both regulators of the Bile acid germination receptor CspC, CspB and CspA and in the gene encoding the proline reductase enzyme, PrdB. Of significant interest was *cspBA*, notably for the reported germination null phenotypes described in studies of random mutants, found to have mutations in all three *cspBAC* genes of *C. difficile* (Francis *et al.,* 2013). Whilst mutations found in clinical isolates were not shared by germination null mutants in the published study, it was found to have the highest level of non-conserved amino acid changes for all isolates investigated.

Given the key role *cspBA* has in spore germination it is hypothesised that these mutations will be related to the reduced initiation of germination in TA mediated germination observed in this study. Kochan *et al.,* (2018), summarise three models of germination in *C. difficile*. For the bile salt-divalent cation pathway, they hypothesise that calcium from the spore core (or the external environment) binds to and activates CspB and subsequently SleC to activate cortex hydrolysis. Further studies to assess the initiation, rate and efficiency of germination including addition methods mentioned section 1.6.6, such also include the addition of Calcium as well as amino acid co-germinants to further elucidate the mechanism and germination lifestyle of these clinical isolates. Further transcription studies to assess the associations of cspBAC should also be considered, as well as the inclusion of *sigE*, known to regulate *cspBAC* expression (Kochan *et al.,* 2018).

3. Loss of heat resistance assays reveal a comparable level of spore dormancy for all isolates.

Using the loss of heat resistance assay to measure the number of non-germinating/non-outgrowing spores for strains of *C. difficile* in chapter 4, the detection of HR-CFU throughout all time points for all isolates and controls tested indicated a level of dormant spores. As discussed, alone, this would imply a potential for all isolates to be persistently carried. Whilst this is true, host factors would play an important role, for example varying levels of clearance as well as optimum host conditions that would promote germination and outgrowth, particularly is susceptible individuals with reduced diversity of nutrient competing microbiota. Additional strain factors could contribute to the level of persistence or carriage or dormant spores could also include biofilm formation, a process not presented in this study. It could be hypothesised that spores and vegetative cells could persist in the environment as well as the host in the form of biofilms that would reinforce the resistance mechanisms, such as evasion of antibiotic activity or physical clearance due the existing in protective layers, providing 'safety in numbers'. Experimental assays to asses biofilm formation would help measure the persistence potential of isolates, that, in this study remains hypothetical.

When considering the findings using the current study, implied dormancy (within the time constrains of experimental methods) can further be separated from non-germinating/non-outgrowing spore fates. A non-germinating fate for a significant proportion of spores (within the time constrains of experimental methods) is clearly demonstrated for all isolates tested, as discussed above. Whereas, spore outgrowth has been shown to vary between isolates tested. Whilst outgrowth of spores is measurably detected using method employed, the fate of spores could also be supported by a spore titre assay. Using methods described in section 1.6.6, the proportion of phase bright or Schaeffer-Fulton endospore stained could indicate the level of intact dormant spores (that would remain phase bright for example) compared to spores arrested in germination (that would by swollen indicating water uptake and cortex hydrolysis). Using both parameters in comparison to the level of outgrown vegetative cells over time (including points beyond those used in this study) it may be possible to make further associations of spore germination, outgrowth, viability and dormancy to the recurrent status of isolates tested.

4. The germination protease gene, *gpr*, of *C. difficile* plays a role in spore viability, outgrowth and the ability of spores to remain dormant.

In Chapter 5, Spores of the inactivated *gpr* strain showed no recovery after decoating and together with the overexpression strain showed no measurable spore outgrowth using the rise in OD method. Interestingly, a complete loss of heat-resistance significantly earlier than WT, complemented and *gpr* overexpression strains was implied by the CFU development below the limit of detection by 5 minutes of the 24-hour assay. Whilst complemented and overexpression strains continued to show detectable HR-CFU, indicating that spores were able to germinate and outgrow into viable vegetative cells. Unfortunately, the role of GPR in the germination of spores could not be established, due to conflicting complementation results in the initiation of germination assay. The germination protease is known to cleave small acid soluble proteins (SASPs) that provide protection of spore DNA in *C. acetobutylicum* (Wetzel and Fischer, 2015). However, experimental limitations of this study meant that this protein interaction was not demonstrated for *C. difficile*. Evenso, the phenotypes displayed by inactivation studies of *gpr* noticeably indicate a role the outgrowth and viability of spores.

Further genomic characterisation of recurrent *C. difficile* isolates revealed amino acid changes in CD630_24700 encoding the *gpr* for isolates 12001_wk0, 12001_RL, 01016 wk-1 and 01016 wk12. However, no significantly different germination, outgrowth or loss of heat resistance phenotypes were detected for these strains in comparison to all isolates test using methods employed in this study. It may be that mutations found, did not affect the catalytic activity of GPR in these strains. Additionally, mutations in promoter regions for investigated genes were not established in this study, and so effects regarding the expression levels of the GPR in clinical isolates cannot be linked to the phenotypes presented in inactivation studies. Interestingly, for a set of isolates from patient 11007 presented in a case study of Chapter 3, amino acid mutations of the SASP encoding gene CD630_12900 were detected in isolates 11007 wk-1 (C2), 11007 wk 0 and 11007 RL, as well as the gene CD630_07750 encoding SpoVAE. SpoVA proteins are known to be involved in the uptake of water in place of Ca-DPA, hydrating the spore during germination (Vepachedu and Setlow, 2007). It has also been found that SASPs have moderate affinity to DNA are able to dissociate in the absence of *gpr* (Sanchez-Salas *et al.,* 1992), hypothesised here to be a result of increased in water content during germination that eventually leads to the dissociation of SASPs over time. Together

then the function of SpoVAE could be hypothesised to influence the dissociation of SASPs. However, this hypothesis would need to be experimentally determined. The key finding for this patient was that post treatment isolates had slower rates of germination and higher OD at endpoint compared their diagnostic counterparts, additionally, the outgrowth from spores was significantly different for strain 11007 wk-1 (C5). Unfortunately, as isolate 11007 wk-1 (C5) was not included in further genomic analysis of Chapter 3, phenotypes presented here cannot be exclusively attributed the observation of genetic variation in *spoVAE*.

Interestingly, differences were observed for time points at which clinical isolates enter into spore dormancy in Chapter 4. Particularly, isolates 01008 wk-1 (E10) and 01013 wk-1, belonging to PCR ribotype CE014 and CE002, respectively, that enter spore dormancy earlier than all clinical isolates tested as well as the control strain, CD630Δ*erm*. Methods used showed no substantial losses of heat resistance beyond 5 minutes, much like the *gpr* inactivation mutant, followed by isolate 01013 wk 12 of PCR ribotype CE014 at 10 mins. there is actually minimal decline in HR WT spores remaining in Chapter 4 and 5, the consistency supports that there is still a significant proportion of spores that do not go on to germinate. This consistent proportion of non-germinating spores (HR-CFU) is somewhat similarly shown for the complement strain. Interestingly, for the *gpr* overexpression strain, staggered germination events are seen at T_5, T_{60} and T_{24}. This indicates that germination doesn't occur at a single time point for all spores and that of the spores that remain relatively significant proportion of them do continue to germinate at further time points. It may be presumed that this pattern could continue after a period of longer dormancy such as those observed in recurrence. It could be concluded that in environments were spores recognise to be suitable for vegetative cell survival, expression of GPR can promote sporadic germination events during dormancy.

Whilst analysis of genetic variation of genes did not show amino acid changes in CD630_24700 for these strains, further analysis as to the expression of *gpr* will be needed to exclude or implicate GPR in such loss of heat resistant phenotypes. Additionally, due to the limited set of genes of interest in this study, associated functions in spore germination and outgrowth of *C. difficile* should also be considered for follow up work.

6.3 Insights into the study of *C. difficile* spore germination, outgrowth and CDI recurrence: A final commentary on hypothesis and study methods

The definition of recurrence is the re-appearance of symptoms following resolution of an initially treated episode and can occur as a result of repeat infection of the same or new strain. Therefore, if any treatment failures do occur, they do so at two points. Firstly, with the failure of spores being eradicated from the host. The large intestine can have several microenvironments where spores can reside and lay dormant for months to years and following completion of treatment periods can germinate and colonise all areas again resulting in disease onset. The second, includes failures in infection control whereby spores are not effectively eradicated in the external environment. From this it can be concluded that spore germination and spore resistance are key microbial risk/virulence factors for recurrent infection.

With the majority of relapses occurring in the first 14 days after successful treatment (Wilcox *et al*., 1998), and the majority of relapses occurring within 8 weeks after initial infection (Kamboj *et al*., 2011), this study highlights the need for longitudinal follow-up, with patients identified as suffering recurrent disease up to three months later. A 10 % clinical relapse rate was identified in Chapter 3 of this study, but with only 73% of the clinically relapsed patients having recoverable *C. difficile* isolated from stool this raises questions as to whether this is an indication of mild or severe CDI, reduced bacterial or spore load or could it be a critique of recovery medium or conditions used to study clinical isolates. Later chapters show a reduced rate of germination for clinical isolates compared to the control strain. With alternative germination pathways being discovered it could be that the classic model of propagating *C. difficile* spores in taurocholate supplemented medium needs to be reviewed. There is a case now for the addition of co-factors, such as amino acids and calcium (Kochan *et al*., 2018) being required to provide optimal recovery of strains from clinical samples, where there is a dependency of shed spores being able to outgrow for further propagation to enable downstream study.

It was hypothesized that germination and outgrowth redundancies of strains relate to prolonged spore dormancy and persistence and thus these specific phenotypes would act as an explanation/virulence factor of relapse. Whereas reinfection would be predicted to be a

result of competitive outgrowth of each co-infecting isolate, where antibiotic activity would affect the faster outgrowing more so than the slower counterpart. The data collected in Chapter 4 shows some support to these hypotheses, with a number of early infecting isolates having a lower proportion of spores initiating germination compared to their post treatment and relapse counterparts. Whilst this trend was only observed in the recurrence cohort, the limited data set restricts such a conclusion. Low-level germination was observed exclusively to clinical isolates and could be related to mutations in regulators and receptors of germination found in Chapter 3. Spore dormancy was detected in all clinical isolates tested indicating the potential for persistent carriage in all isolates and the molecular mechanism for this pathway touched upon in Chapter 5. However, limitations of this study include minimal non-recurrence isolates being included in the analysis of genomic variation and spore dormancy and so no definite inferences as to a recurrence vs non-recurrence phenotype can be confidently made.

With the definition of germination varying in sensitivity between research groups; some including outgrowth and viability, others not, it may be that their implications on recurrence are restricted to the collective process. By employing methods to independently assess sporulation, germination and spore outgrowth, conclusions of the findings observed will have had more weight towards specific processes. Whilst the differentiation into specific phenotypes was not completely ascertained in this study, this was a novel attempt to shed light on spore fate for clinical isolates that future work can progress from.

Genetic variation observed in sporulation, germination and outgrowth genes in genomes of recurrent isolations, could have distinct implication on the phenotypes of clinical isolates of the recurrent. It was aimed that this would be an outcome of this study. With such a large undertaking of work, it is to be credited that genomic analysis was considered, by way of using the earliest strain isolated as the scaffold of genome assembly, to provide a more accurate assessment of variant calling. By not mapping strains to a single reference strain, individual SNP differences between isolates allow more robust and supporting distinctions between clinical and laboratory strains. Many studies have made inferences of genetic differences between strains, linking the findings to PCR ribotype with minimal comparator groups and use of a single laboratory or serially passaged clinical control strain as the backbone of

genome assembly and variant calling (Wroblewski *et al.,* 2009; Didelot *et al.,* 2012; Eyre *et al.,* 2012; 2013; He *et al*., 2013; Kumar *et al.,* 2016). Studies have found varying levels of germination and the existence of super dormant spores amongst isolates with little investigation as to what links these phenotypic differences. Identification of key mutations and knockout studies of germinant receptors, exosporium proteins and sporulation genes has been extensively studied in control strains, whilst the identification of these variants in clinical isolates together with phenotypic observation has been limited.

One of the main novelties of this studies is the significantly reduced rates of germination observed for isolates compared to the control strain in Chapter 4. Identification of mutations for the isolates studied point towards the mechanisms that lead to this increased control over the germination process and support a role of this control in the pathogenesis of *C. difficile*. However, with genetic variation analysis in Chapter 3 only being performed on the recurrent cohort sub-study, findings associating key genes with persistence, relapse or re-infection are noticeably weak without a non-recurrence comparator sub-group. This limitation is also reflected in the restricted number of sporulation, germination, growth and antibiotic resistance genes selected for study. Nevertheless, the observed genetic variation of genes selected within the clinical isolates available still provides relevant and novel findings. Potential expansion of the sub-group to included non-recurrence strains, would supplement associations as well as support the mode of action in the model strain CD630 within the literature. Developments in the context of germination and outgrowth are welcomed, being that these pathways are not as well discerned in *C. difficile* as others, such as the regulation of sporulation or toxin production for example.

Within clostridial strains the latter pathways of the spores' lifecycle are becoming more elucidated, particularly with reference to spore coat formation and the role of SASPs in spore resistance. Whilst this study does elucidate a role for *gpr* in germination and outgrowth in *C. difficile*, a finding not yet demonstrated in other clostridial pathogens, the proteolytic activity of GPR on SASPs or the SASP mediated protection of DNA was not found with the methods employed. Further investigation into SASPs and role of resistance of *C. difficile* spores, as well as the mobilisation of specific spore coat and exosporium proteins remain. With spores ubiquitously known as the main transmission vehicles in transmission, the existence of super-

dormancy in the environment and biofilm produced spores in the host, answers to these questions remain a concern.

Proposed research questions for future consideration:

- Do vegetative cells sense the presence of antimicrobials as a trigger for sporulation in the colon, as to evade cell death?
- How quickly can vegetative cells sporulate in the host?
- Do alternative germination pathways, utilising calcium and amino acid co-germinants show differences to the tight control of TA-mediated germination in clinical isolates?
- What effect does biofilm formation have on the ability of spores to persist and can they be experimentally determined?
- Can spore dormancy or cell viability be further experimentally separated? If so, can associations be made for the levels non-germinating and non-outgrowing spores in recurrent or persistently carried isolates?
- What is the role of GPR in the germination cascade of *C. difficile* spores?
- Can the hypothetical role for SpoVAE mediated hydrolysis and the dissociation of SASPs be experimentally determined?

Chapter 7: Appendix

Table 7.1: Composition of *C. difficile* Minimal Medium

Component	Stock concentration (mg/mL)	Final concentration (mg/mL)
Amino acids (5X)		
Casamino acids	50	10
L-Tryptophan	2.5	0.5
L-cysteine	2.5	0.5
Salts (10X)		
Na_2HPO_4	50	5
$NaHCO_3$	50	5
KH_2PO_4	9.0	0.9
NaCl	9.0	0.9
Trace salts (50X)		
$(NH_4)2SO_4$	2.0	0.04
$CaCl_2 \cdot 2H_2O$	1.3	0.026
$MgCl_2 \cdot 6H_2O$	1.0	0.02
$MnCl_2 \cdot 4H_2O$	0.5	0.01
$CoCl_2 \cdot 6H_2O$	0.05	0.001
Glucose (20X)		
D-glucose	200	10
Iron (100X)		
$FeSO_4 \cdot 7H_2O$	0.4	0.004
Vitamins (1000X)		
D-biotin	1.0	0.001
Calcium-D-panthothenate	1.0	0.001
Pyridoxine	1.0	0.001
5-fluorocytosine	10	0.05

Table 7.2: Composition of cationic stacking and resolving polyacrylamide solutions

Solution	Component	Volume
A	1M KOH	24.0 mL
	Glacial acetic acid	11.2 mL
	Tetramethylenediamine (TEMED)	2.3 mL
	H_2O	62.5 mL
B	1M KOH	24.0 mL
	Glacial acetic acid	1.44 mL
	TEMED	0.23 mL
	H_2O	74.33 mL
C	Acrylamide (30%)	28 g
	Bis-acrylamide (0.8%)	0.74 g
	H_2O	100 mL
D (Filtered through Whatman #1 paper)	Acrylamide	10 g
	Bis-acrylamide	2.5 g
	H_2O	100 mL
E	Riboflavin	0.4 mg
	H_2O	100 mL
F	Ammonium persulfate	0.4 g
	H_2O	100 mL

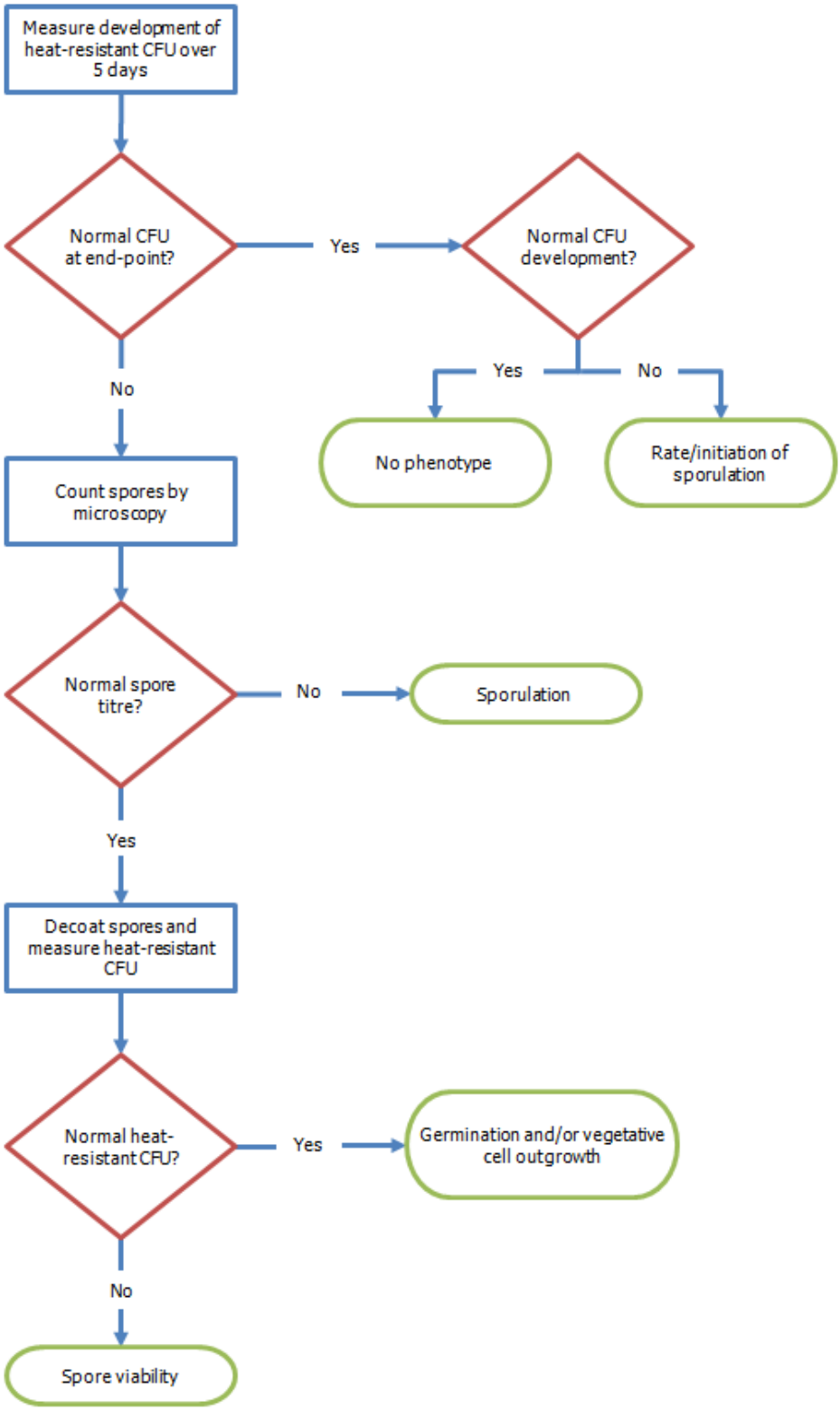


Figure 7.1: Flowchart of phenotype characterisation of *C. difficile* spores.

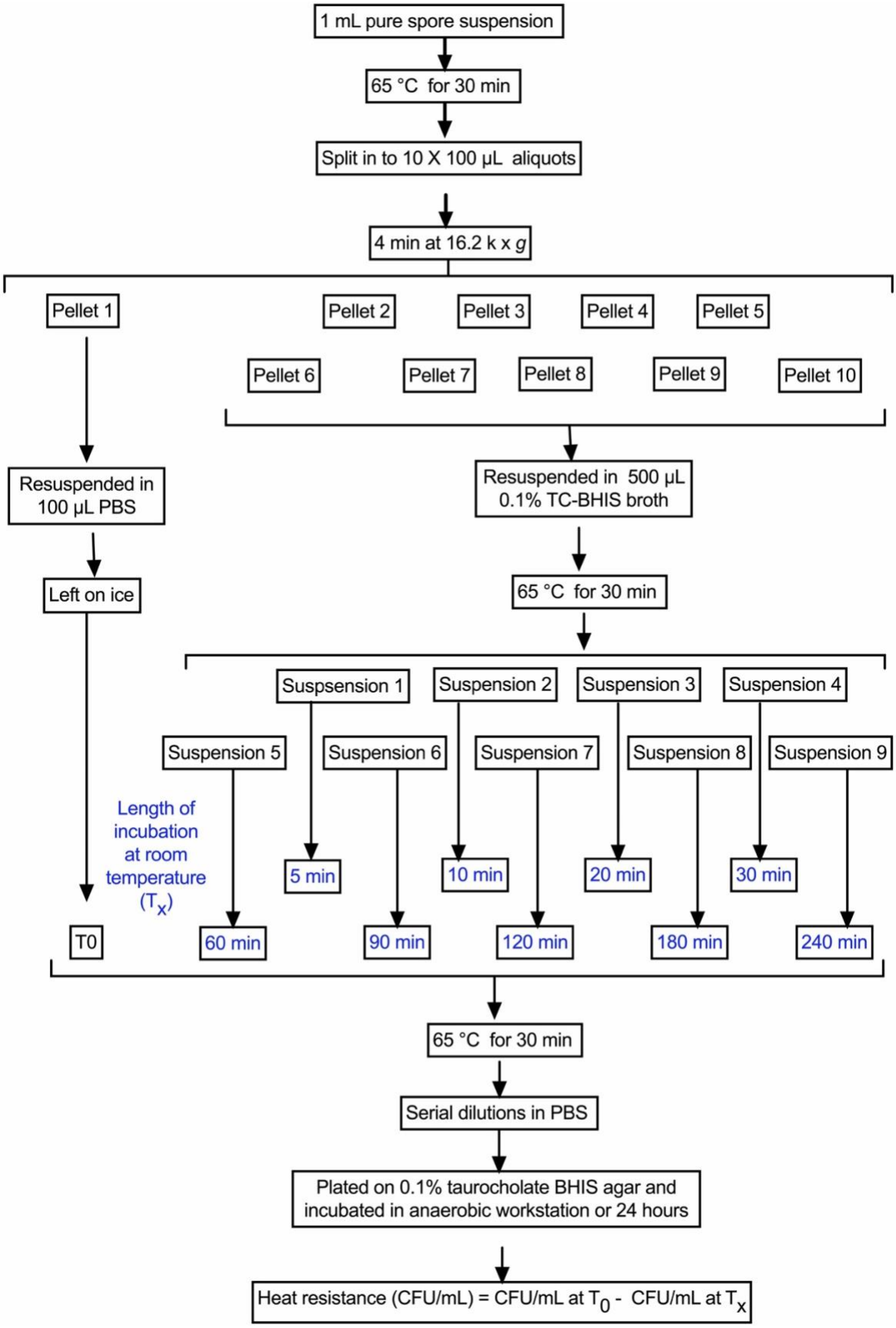


Figure 7.2: Experimental flowchart of the loss of heat resistance assay

Table 7.3: Variant detection analysis provided by Dr Patrick Budd of 27 genomes of clinical *C. difficile* isolates, used to corroborate the findings in this study.

Patient	Sample	Mutations	Reinfection Mutations	Relapse/ Reinfection	Ribotype
1016	Wk-1	-	-	-	26
	Wk 12	21	-	Relapse	26
11007	Wk-1	-	-	-	14
	Wk 0	17607	-	Reinfection	14
	RL	17587	39	Reinfection	14
12001	Wk 0	-	-	-	70
	RL	4	-	Relapse	70
12004	Wk 0	-	-	-	20
	RL	19	-	Relapse	20
18001	Wk-1	-	-	-	20
	Wk 0	561	-	Reinfection	20
	RL	546	0	Reinfection	20
01006	Wk-1	-	-		
	Wk0	65	-	Persistant	
	Wk12	1378	1243	Reinfection	
01013	Wk-1	-	-		
	Wk12	9322	-	Reinfection	
01015	Wk-1	-	-		
	Wk4	9322	-	Reinfection	
01017	Wk-1	-	-		
	Wk4	78	-	Persistant	
01023	Wk-1	-	-		
	Wk0	152	-	Persistant	
01028	Wk-1	-	-		
	Wk0	29691	-	Reinfection	
03008	Wk0	-	-		
	Wk4	2737	-	Reinfection	

Table 7.4: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient.

Isolate	Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Forward /reverse balance	Average quality	Coding region	Coding region change	Amino acid change
01008 Wk-1 (E10)	147783	SNV	A	T	195	198	98.48	0.49	34.89	Aminobenzoyl -glutamate transport protein	c.210T>A	
	556339	Deletion	A	-	146	149	97.99	0.46	32.74	Oligopeptide transport ATP-binding protein OppF	c.479delA	p.Gln160fs
	2974073	SNV	G	A	151	211	71.56	0.44	36.64			
	3762298	SNV	C	A	161	161	100	0.47	36.15	DNA-directed RNA polymerase beta subunit	c.1465C>A	p.Gln489Lys
01008 wk12 (G2)	2587297	SNV	T	C	128	171	74.85	0.45	36.85	-	-	-
	2587339	SNV	C	G	134	179	74.86	0.45	35.57	-	-	-
	2587344	MNV	AC	GT	131	175	74.86	0.44	35.6	-	-	-
	2587347	SNV	T	C	132	175	75.43	0.44	36.03	-	-	-
	2587359	MNV	CA	TC	138	182	75.82	0.35	36.87	-	-	-
	2587362	SNV	A	G	138	181	76.24	0.35	35.8	-	-	-

MNV: multiple nucleotide variant, SNV: single nucleotide variant

Table 7.4 Continued: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient.

Isolate	Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Forward /reverse balance	Average quality	Coding region	Coding region change	Amino acid change
01008 wk12 (G2)	3216110	SNV	T	A	155	309	50.16	0.48	37.37	D-proline reductase, 45 kDa subunit / D-proline reductase, 23 kDa subunit	c.1290A>T	
	3216125	SNV	T	A	159	310	51.29	0.49	36.54		c.1275A>T	
	3216152	SNV	G	T	161	314	51.27	0.47	37.65		c.1248C>A	
	3216157	SNV	C	A	158	309	51.13	0.46	37.15		c.1243G>T	p.Gly415Cys
	3216164	SNV	T	C	157	308	50.97	0.46	36.73		c.1236A>G	
	3216184	MNV	AT	CC	153	302	50.66	0.47	37.09		c.1215_1216delATinsGG	p.Leu406Val
	3216187	SNV	C	T	154	301	51.16	0.47	37.56		c.1213G>A	p.Val405Ile
	3216194	SNV	C	T	153	299	51.17	0.48	37.02		c.1206G>A	p.Met402Ile
	3216196	SNV	T	C	154	299	51.51	0.48	36.54		c.1204A>G	p.Met402Val
	3216218	SNV	T	C	151	296	51.01	0.46	36.81		c.1182A>G	
	3216220	SNV	G	T	150	296	50.68	0.45	36.88		c.1180C>A	p.Gln394Lys
	3216224	SNV	T	A	156	302	51.66	0.44	36.18		c.1176A>T	
	3216231	MNV	GTA	ACT	150	296	50.68	0.44	36.33		c.1167_1169delTACinsAGT	p.Thr390Val
	3216242	SNV	C	T	155	297	52.19	0.45	36.71		c.1158G>A	
	3216245	SNV	A	C	154	296	52.03	0.43	37.29		c.1155T>G	
	3216247	SNV	C	T	155	296	52.36	0.44	37.43		c.1153G>A	p.Val385Ile
	3216259	SNV	G	A	158	302	52.32	0.42	36.77		c.1141C>T	
	3216266	SNV	A	G	157	302	51.99	0.41	37.47		c.1134T>C	
	3216281	SNV	T	A	157	308	50.97	0.41	36.86		c.1119A>T	
	3216298	MNV	TA	CT	149	295	50.51	0.42	37.27		c.1101_1102delTAinsAG	p.Ile368Val
	3216301	SNV	C	T	148	296	50	0.42	37.8		c.1099G>A	p.Val367Ile

MNV: multiple nucleotide variant, SNV: single nucleotide variant

Table 7.4: Continued: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient.

Isolate	Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Forward /reverse balance	Average quality	Coding region	Coding region change	Amino acid change
01008 wk12 (G2)	3762308	SNV	A	T	143	174	82.18	0.43	36.96	DNA-directed RNA polymerase beta subunit	c.1475A>T	p.Asp492Val
	3834256	SNV	T	A	43	76	56.58	0.48	33.05	Glucosamine--fructose-6-phosphate aminotransferase	c.210T>A	
01008 RL (H4)	2562162	SNV	T	C	84	85	98.82	0.48	34.56	-	-	-

MNV: multiple nucleotide variant, SNV: single nucleotide variant

Table 7.4: Continued: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient.

Isolate	Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Forward /reverse balance	Average quality	Coding region change	Amino acid change
01016 wk 12	289980	Insertion	-	A	98	102	96.08	0.47	35.97	-	-
	608501	Insertion	-	A	100	114	87.72	0.49	35.98	-	-
	671867	SNV	C	T	82	84	97.62	0.46	36.39	-	-
	901149	Insertion	-	T	83	88	94.32	0.46	34.75	-	-
	1117336	Insertion	-	A	78	80	97.5	0.47	35.46	-	-
	1259267	Insertion	-	A	76	79	96.2	0.46	36.58	-	-
	2433726	Insertion	-	T	83	88	94.32	0.49	35.89	-	-
	2812852	Insertion	-	A	71	79	89.87	0.49	35.77	-	-
	3487082	Insertion	-	T	88	96	91.67	0.49	35.67	-	-
	3786343	SNV	C	T	108	111	97.3	0.49	34.93	-	-
	4156110	Insertion	-	A	92	98	93.88	0.46	36.62	-	-

MNV: multiple nucleotide variant, SNV: single nucleotide variant

Table 7.4: Continued: Variant detection analysis of CLC assembled genomes for 'recurrent cohort' isolates in *C. difficile* isolates with <100 sequence variants. Reference position and reference were taken from the genome data of the earliest isolate of said patient.

Isolate	Reference Position	Type	Reference	Allele	Count	Coverage	Frequency	Forward /reverse balance	Average quality	Coding region change	Amino acid change
12001 RL	228694	SNV	G	A	205	207	99.03	0.46	33.47	-	-
	703282	SNV	C	A	45	77	58.44	0.47	33.56	-	-
	703468	SNV	T	G	45	89	50.56	0.47	33	-	-
	791981	SNV	C	A	88	90	97.78	0.47	35.19	-	-
	2667306	SNV	A	G	111	112	99.11	0.5	33.82	-	-
	3900258	Insertion	-	T	81	87	93.1	0.49	35.53	-	-
12004 RL	249674	Insertion	-	A	75	85	88.24	0.39	36.47	-	-
	594810	Insertion	-	T	73	83	87.95	0.49	36.29	-	-
	1268970	Insertion	-	T	74	81	91.36	0.5	36.43	-	-
	1525844	Insertion	-	T	77	84	91.67	0.47	36.53	-	-
	1715090	Insertion	-	T	85	94	90.43	0.47	36.68	-	-
	1751026	Insertion	-	T	70	84	83.33	0.5	34.3	-	-
	2521560	Insertion	-	T	102	119	85.71	0.44	34.6	-	-
	2567634	Insertion	-	T	105	116	90.52	0.48	36.06	-	-
	4054837	Insertion	-	A	107	110	97.27	0.48	35.93	-	-

MNV: multiple nucleotide variant, SNV: single nucleotide variant

Table 7.5: Percentage nucleotide sequence identity of 15 clinical isolates for 24 genes of interest from *C. difficile* reference strain CD630

	Percentage nucleotide sequence identity to *C. difficile* CD630 Reference Gene (%)																							
Clinical Isolate Strain ID	CD630_00660	CD630_01290	CD630_03320	CD630_05510	CD630_07730	CD630_07740	CD630_07750	CD630_08530	CD630_08540	CD630_08550	CD630_08560	CD630_08570	CD630_12900	CD630_22460	CD630_22470	CD630_23520	CD630_24700	CD630_26880	CD630_32201	CD630_32410	CD630_32490	CD630_33490	CD630_34640	CD360_35690
18001_RL	99	99	88	100	100	100	100	100	100	99	99	99	100	100	99	100	100	100	99	100	100	99	100	100
18001_wk0	99	99	88	100	100	100	100	100	100	99	99	99	100	100	99	100	100	100	99	100	100	99	100	100
12004_wk0	99	99	88	100	100	100	100	100	100	99	99	99	100	100	99	100	100	100	99	100	100	99	100	100
12004_RL	99	99	88	100	100	100	100	100	100	99	99	99	100	100	99	100	100	100	99	100	100	99	100	100
01008_RL (H4)	100	99	90	100	100	100	100	100	100	99	99	99	100	100	99	100	100	100	99	100	100	98	100	100
01008_wk 12 (G2)	99	99	94	100	100	100	100	100	100	100	99	98	100	99	99	100	100	100	99	95	100	80	99	100
01008_ wk-1(E10)	99	99	94	100	100	100	100	100	100	100	99	99	100	99	99	100	100	100	99	100	100	80	99	100
12001_wk0	100	99	88	100	100	100	100	99	100	100	99	99	100	100	99	100	99	100	99	100	100	81	100	100
12001_RL	100	99	88	100	100	100	100	99	100	100	99	99	100	100	99	100	99	100	99	100	100	81	100	100
01016_wk-1	100	99	93	99	100	100	100	100	99	100	99	99	100	98	98	100	99	100	99	100	100	97	100	100
01016_wk12	100	99	93	99	100	100	100	100	99	100	99	99	100	98	98	100	99	100	99	100	100	97	100	100
01018 *wk-1 (C17)*	99	99	90	100	100	100	100	100	99	100	99	99	100	100	100	100	100	100	99	100	100	98	100	100
11007_RL	100	99	89	100	100	100	99	100	100	100	99	99	98	100	99	100	100	100	99	100	100	82	100	100
11007_wk-1 (C1)	100	99	89	100	100	100	99	100	100	100	99	99	98	100	99	100	100	100	99	100	100	82	100	100
11007_wk0	100	99	89	100	100	100	99	100	100	100	99	99	98	100	99	100	100	100	99	100	100	82	100	100

Table 7.6: Number of amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates

Gene	Conservative changes	Non-conservative changes [a]	Total no. amino acid changes
CD630_00660	1	2	3
CD630_01290	0	0	0
CD630_03320	>20		
CD630_05510	0	1	1
CD630_07730	0	0	0
CD630_07740	0	0	0
CD630_07750	1	0	1
CD630_08530	0	1	1
CD630_08540	1	0	1
CD630_08550	1	0	1
CD630_08560	0	1	1
CD630_08570	4	11	15
CD630_12900	1	0	1
CD630_22460	5	7	12
CD630_22470	8	11	19
CD630_23520	0	0	0
CD630_24700	1	2	3
CD630_26880	0	0	0
CD630_32201	0	0	0
CD630_32410	1	7	8
CD630_32490	0	0	0
CD630_33490	>20		
CD630_34640	0	1	1
CD360_35690	0	0	0
Total	24	44	68 (70 [b])

[a] Including amino acid insertions, [b] Including total amino acid changes in genes CD630_03320 and CD630_33490 as n=2.

Table 7.7: Amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates

Gene	Isolate(s)	Reference coding position	Reference residue	Alternative residue	Amino acid change
CD630_00660	01018 *wk-1 (C17)*	750	I	V	Conservative
	18001_RL				
	18001_wk0				
	12004_wk0				
	12004_RL				
	01008_G2	492	D	V	Non-conservative
	01008_E10	489	Q	K	Non-conservative
CD630_03320	All	Variable repeats, some changes in conserved regions			
CD630_05510	01016_wk-1	240	A	S	Non-conservative
	01016_wk12				
CD630_07750	11007_RL	113	V	I	Conservative
	11007_wk-1-C2				
	11007_wk0				
CD630_08530	12001_wk0	327	T	A	Non-conservative
	12001_RL				
CD630_08540	01016_wk-1	189	K	R	Conservative
	01016_wk12				
	01018 *wk-1 (C17)*				
CD630_08550	18001_RL	512	I	V	Conservative
	18001_wk0				
	12004_wk0				
	12004_RL				
	01008_H4				
CD630_08560	12001_wk0	263	S	L	Non-conservative
	12001_RL				
CD630_08570	18001_RL	46	-	+S	Insertion
	18001_wk0				
	12004_wk0				
	12004_RL				
	01008_H4	285	L	S	Non-conservative
	12001_wk0	301	D	E	Conservative
	12001_RL	305	S	N	Non-conservative
	01008_G2	263	S	L	Non-conservative
		46	-	+ S	Insertion
		285	L	S	Non-conservative
		290	E	D	Conservative
		301	D	E	Conservative
		305	S	N	Non-conservative
	01008_E10	46	-	+S	Insertion
		Truncated			
	01016_wk-1	46	-	+S	Insertion
	01016_wk12				
	01018 *wk-1 (C17)*	305	S	N	Non-conservative
	11007_RL	46	-	+S	Insertion
	11007_wk-1-C2				
	11007_wk0				

Table 7.7 Continued: Amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates

Gene	Isolate(s)	Reference coding position	Reference residue	Alternative residue	Amino acid change
CD630_12900	11007_RL	4	V	I	Conservative
	11007_wk-1-C2				
	11007_wk0				
CD630_22460	01008_G2	72	N	D	Conservative
	01008_E10	415	D	E	Conservative
	01016 wk-1	39	L	I	Conservative
	01016 wk12	157	N	S	Non-conservative
		158	R	T	Non-conservative
		162	S	G	Non-conservative
		184	I	V	Non-conservative
		267	A	T	Non-conservative
		331	V	A	Conservative
		374	R	G	Non-conservative
		394	N	G	Non-conservative
		530	E	D	Conservative
CD630_24700	12001_wk0	48	R	E	Non-conservative
	12001_RL	52	A	S	Non-conservative
	01016 wk-1	62	I	V	Conservative
	01016 wk12				
CD630_32410	All	Truncated			
	01008_G2	123	H	Q	Non-conservative
		124	V	K	Non-conservative
		126	T	K	Non-conservative
		132	A	E	Non-conservative
		135	A	R	Non-conservative
		136	K	I	Non-conservative
		146	V	L	Conservative
		147	M	L	Non-conservative
CD630_33490	All	Variable region, some variants in 2nd conserved region, longer than reference gene			
CD630_34640	01008_G2	94	G	D	Non-conservative
	01008_E10				

Table 7.7 Continued: Amino acid changes amongst 24 genes of interest from *C. difficile* reference strain CD630 found amongst 15 clinical isolates

Gene	Isolate(s)	Reference coding position	Reference residue	Alternative residue	Amino acid change
CD630_22470	18001_RL	540	M	I	Non-conservative
	18001_wk0				
	12004_wk0				
	12004_RL				
	01008_H4				
	01008_G2				
	01008_E10				
	01016_wk-1				
	01016_wk12				
	11007_RL	544	S	N	Non-conservative
	11007_wk-1-C2				
	11007_wk0				
	18001_RL	629	S	N	Non-conservative
	18001_wk0				
	12004_wk0				
	12004_RL				
	01008_H4				
	01008_G2				
	01008_E10				
	12001_wk0				
	12001_RL				
	11007_RL				
	11007_wk-1-C2				
	11007_wk0				
	01016 wk-1 01016 wk 12	471	I	T	Non-conservative
		141	D	N	Conservative
		278	V	I	Conservative
		283	R	H	Conservative
		319	I	V	Conservative
		515	F	L	Non-conservative
		558	F	L	Non-conservative
		592	I	V	Conservative
		604	G	E	Non-conservative
		610	R	S	Non-conservative
		622	S	F	Non-conservative
		794	D	G	Non-conservative
		827	T	S	Conservative
		884	D	E	Conservative
		911	E	K	Non-conservative
		914	S	R	Non-conservative
		1021	G	S	Non-conservative
		1035	D	N	Conservative
		1171	S	A	Non-conservative

Table 7.8: Statistics of the slope from 0 to 9 minutes of the initiation of germination for isolates of *C. difficile*

Isolate	Slope	Standard error of slope
CD630Δ*erm*	-0.0265546	0.00155353
EK28	-0.0268272	0.00261616
12001 wk 0	-0.0044652	0.00091138
12001 RL	-0.004195	0.00112082
11001 wk 4	-0.0187595	0.00199138
11001 RL	-0.0171961	0.00325155
01018 wk-1(C17)	-0.008477	0.00113957
01018 wk-1 (C20)	-0.0130754	0.00070902
12004 wk 0	-0.0051134	0.0010924
12004 RL	-0.0036285	0.00087597
18001 wk 0	-0.0051134	0.0010924
18001 RL	-0.0036285	0.00087597
01016 WK-1	-0.0043872	0.000766
01016 WK 12	-0.0039693	0.00056289
01013 wk-1	-0.0187686	0.00364684
01013 RL	-0.0149667	0.00280072
01015 wk 4	-0.0103381	0.00127439
01015 wk-1	-0.0101631	0.00131826
01017 wk-1	-0.0076371	0.00179625
01023 wk-1	-0.0167983	0.0015257
01028 wk-1	-0.0067617	0.00061325
01028 wk0	-0.0063198	0.00057158

Chapter 8: References

Adams, C. M., Eckenroth, B. E., Putnam, E. E., Doublie, S., & Shen, A. (2013). Structural and Functional Analysis of the CspB Protease Required for *Clostridium* Spore Germination. PLOS Pathogens, 9(2). doi:10.1371/journal.ppat.1003165

Akerlund, T., Persson, I., Unemo, M., Noren, T., Svenungsson, B., Wullt, M., & Burman, L. (2008). Increased sporulation rate of epidemic *Clostridium difficile* Type 027/NAP1. Journal of Clinical Microbiology, 46, 1530 - 1533

Alfa, M. J., Kabani, A., Lyerly, D., Moncrief, S., Neville, L. M., Al-Barrak, A., Harding, G. K., Dyck, B., Olekson, K., & Embil, J. M. (2000). Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile* responsible for a nosocomial outbreak of *Clostridium difficile*-associated diarrhea. Journal of Clinical Microbiology, 38(7), 2706-271

Androga, G. O., Hart, J., Foster, N. F., Charles, A., Forbes, D., & Riley, T. V. (2015). Infection with Toxin A-Negative, Toxin B-Negative, Binary Toxin-Positive *Clostridium difficile*in a Young Patient with Ulcerative Colitis. Journal of Clinical Microbiology, 53(11), 3702-3704. doi:10.1128/JCM.01810-15

Angelotti, R., Maryanski, J. H., Butler, T. F., Peeler, J. T., Campbell, J. E (1968). Influence of Spore Moisture Content on the Dry-Heat Resistance of *Bacillus subtilis var. niger*. Applied Microbiology, 16 (5) 735-745

Ankenbrand, M. J., & Keller, A. (2016). bcgTree: automatized phylogenetic tree building from bacterial core genomes. Genome, 59(10), 783-791. doi:10.1139/gen-2015-0175

Antunes, A., Camiade, E., Monot, M., Courtois, E., Barbut, F., Sernova, N. V., Rodionov, D. A., Martin-Verstraete, I., & Dupuy, B. (2012). Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. Nucleic Acids Research, 40(21), 10701-10718. doi:10.1093/nar/gks864

Antunes, A., Martin-Verstraete, I., & Dupuy, B. (2011). CcpA-mediated repression of *Clostridium difficile*toxin gene expression. Molecular Microbiology, 79(4), 882-899. doi:10.1111/j.1365-2958.2010.07495.x

Aronson, A. I., & Fitzjames, P. (1976). Structure and morphogenesis of bacterial spore coat. Bacteriological Reviews, 40(2), 360-40

Aronson, A. I., Ekanayake, L., & Fitzjames, P. C. (1992). Protein Filaments May Initiate The Assembly Of The *Bacillus subtilis* Spore Coat. Biochimie, 74(7-8), 661-667. doi:10.1016/0300-9084(92)90138-5

Aslam, S., Hamill, R. J. & Musher, D. M. 2005. Treatment of *Clostridium difficile* associated disease: old therapies and new strategies. The Lancet Infectious Diseases, 5, 549-557

Atrih, A., & Foster, S. J. (2001). Analysis of the role of bacterial endospore cortex structure in resistance properties and demonstration of its conservation amongst species. Journal of Applied Microbiology, 91(2), 364-372. doi:10.1046/j.1365-2672.2001.01394.x

Atrih, A., Bacher, G., Allmaier, G., Williamson, M. P., & Foster, S. J. (1999). Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. Journal of Bacteriology, 181(13), 3956-3966

Atrih, A., Bacher, G., Korner, R., Allmaier, G., & Foster, S. J. (1999). Structural analysis of *Bacillus megaterium* KM spore peptidoglycan and its dynamics during germination. Microbiology-UK, 145, 1033-1041. doi:10.1099/13500872-145-5-1033

Atrih, A., Zollner, P., Allmaier, G., & Foster, S. J. (1996). Structural analysis of *Bacillus subtilis* 168 endospore peptidoglycan and its role during differentiation. Journal of Bacteriology, 178(21), 6173-6183. doi:10.1128/jb.178.21.6173-6183.1996

Atrih, A., Zollner, P., Allmaier, G., Williamson, M. P., & Foster, S. J. (1998). Peptidoglycan structural dynamics during germination of *Bacillus subtilis* 168 endospores. Journal of Bacteriology, 180(17), 4603-4612

Avbersek, J., Janezic, S., Pate, M., Rupnik, M., Zidaric, V., Logar, K., Vengust, M., Zemljic, M., Pirs, T., & Ocepek, M. (2009). Diversity of *Clostridium difficile* in pigs and other animals in Slovenia. Anaerobe, 15(6), 252-255

Awad, M. M., Ellemor, D. M., Boyd, R. L., Emmins, J. J., & Rood, J. I. (2001). Synergistic effects of alpha-toxin and perfringolysin O in *Clostridium perfringens* -mediated gas gangrene. Infection and Immunity, 69(12), 7904-7910. doi:10.1128/iai.69.12.7904-7910.2001

Babakhani, F., Bouillaut, L., Gomez, A., Sears, P., Nguyen, L., & Sonenshein, A. L. (2012). Fidaxomicin inhibits spore production in *Clostridium difficile*. Clinical Infectious Diseases, 55 Suppl 2, S162-169. doi:10.1093/cid/cis453

Bagyan, I., & Setlow, P. (2002). Localization of the cortex lytic enzyme CwlJ in spores of *Bacillus subtilis*. Journal of Bacteriology, 184(4), 1219-1224. doi:10.1128/jb.184.4.1219-1224.2002

Bajaj, J. S., Heuman, D. M., Sanyal, A. J., Hylemon, P. B., Sterling, R. K., Stravitz, R. T., Fuchs, M., Ridlon, J. M., Daita, K., Monteith, P., Noble, N. A., White, M. B., Fisher, A., Sikaroodi, M., Rangwala, H., & Gillevet, P. M. (2013). Modulation of the Metabiome by Rifaximin in Patients with Cirrhosis and Minimal Hepatic Encephalopathy. PLOS ONE, 8(4), e60042. doi:10.1371/journal.pone.0060042

Bakken, J. S. (2009). Fecal bacteriotherapy for recurrent *Clostridium difficile* infection. Anaerobe, 15(6), 285-289. doi:10.1016/j.anaerobe.2009.09.007

Bakken, J. S., Borody, T., Brandt, L. J., Brill, J. V., Demarco, D. C., Franzos, M. A., Kelly, C., Khoruts, A., Louie, T., Martinelli, L. P., Moore, T. A., Russell, G., & Surawicz, C. (2011). Treating *Clostridium difficile* infection with fecal microbiota transplantation. Clinical Gastroenterology and Hepatology, 9(12), 1044-1049. doi:10.1016/j.cgh.2011.08.014

Banawas, S., Paredes-Sabja, D., Korza, G., Li, Y., Hao, B., Setlow, P., & Sarker, M. R. (2013). The *Clostridium perfringens* Germinant Receptor Protein GerKC Is Located in the Spore Inner Membrane and Is Crucial for Spore Germination. Journal of Bacteriology, 195(22), 5084-5091. doi:10.1128/jb.00901-13

Barbut, F., Richard, A., Hamadi, K., Chomette, V., Burghoffer, B., & Petit, J.-C. (2000). Epidemiology of Recurrences or Reinfections of *Clostridium difficile*-Associated Diarrhea. Journal of Clinical Microbiology, 38(6), 2386-2388

Barra-Carrasco, J., Olguín-Araneda, V., Plaza-Garrido, A., Miranda-Cárdenas, C., Cofré-Araneda, G., Pizarro-Guajardo, M., Sarker, M. R., & Paredes-Sabja, D. (2013). The *Clostridium difficile* exosporium cysteine (CdeC)-rich protein is required for exosporium morphogenesis and coat assembly. Journal of Bacteriology, 195(17), 3863–3875. https://doi.org/10.1128/JB.00369-13

Bartlett, J. G. (2010). *Clostridium difficile*: progress and challenges. Annals of the New York Academy of Sciences, 1213(1), 62-69. doi:10.1111/j.1749-6632.2010.05863.x

Bauer, M. P., Notermans, D. W., van Benthem, B. H., Brazier, J. S., Wilcox, M. H., Rupnik, M., Monnet, D. L., van Dissel, J. T., & Kuijper, E. J. (2011). *Clostridium difficile* infection in Europe: a hospital-based survey. Lancet, 377(9759), 63-73. doi:10.1016/s0140-6736(10)61266-4

Beaman, T. C., & Gerhardt, P. (1986). Heat-Resistance of Bacterial-Spores Correlated with Protoplast Dehydration, Mineralization, and Thermal Adaptation. Applied and Environmental Microbiology, 52(6), 1242-1246

Beaman, T. C., Pankratz, H. S., & Gerhardt, P. (1971). Paracrystalline Sheets Reaggregated From Solubilized Exosporium Of *Bacillus cereus*. Journal of Bacteriology, 107(1), 320.

Bender, G. R., & Marquis, R. E. (1985). Spore Heat-Resistance and Specific Mineralization. Applied and Environmental Microbiology, 50(6), 1414-1421

Bertsch, L. L., Bonsen, P. P. M., & Kornberg, A. (1969). Biochemical Studies of Bacterial Sporulation and Germination. XIV. Phospholipids in *Bacillus megaterium*. Journal of Bacteriology, 98(1), 75-81

Bhattacharjee, D., McAllister, K. N., & Sorg, J. A. (2016). Germinants and Their Receptors in *Clostridia*. Journal of Bacteriology, 198(20), 2767-2775. doi:10.1128/jb.00405-16

Bottone, E. J. (2010). *Bacillus cereus*, a Volatile Human Pathogen. Clinical Microbiology Reviews, 23(2), 382. doi:10.1128/cmr.00073-09

Bowen, W. R., Fenton, A. S., Lovitt, R. W., & Wright, C. J. (2002). The measurement of *Bacillus mycoides* spore adhesion using atomic force microscopy, simple counting methods, and a spinning disk technique. Biotechnology and Bioengineering, 79(2), 170-179. doi:10.1002/bit.10321

Brigidi, P., Swennen, E., Rizzello, F., Bozzolasco, M., & Matteuzzi, D. (2002). Effects of Rifaximin Administration on the Intestinal Microbiota in Patients with Ulcerative Colitis. Journal of Chemotherapy, 14(3), 290-295. doi:10.1179/joc.2002.14.3.290

Brown, C. J., Boutros, M., Morris, A., & Divino, C. M. (2014). CAGS and ACS evidence based reviews in surgery. Is a diverting loop ileostomy and colonic lavage an alternative to colectomy for the treatment of severe *Clostridium difficile*-associated disease? Canadian Journal of Surgery, 57(3), 214-216. doi:10.1503/cjs.005014

Buckley, A. M., Spencer, J., Candlish, D., Irvine, J. J., & Douce, G. R. (2011). Infection of hamsters with the UK *Clostridium difficile* ribotype 027 outbreak strain R20291. Journal of Medical Microbiology, 60(8), 1174-1180. doi:10.1099/jmm.0.028514-0

Burns, D. A., & Minton, N. P. (2011). Sporulation studies in *Clostridium difficile*. Journal of Microbiological Methods, 87(2), 133-138. doi:10.1016/j.mimet.2011.07.017

Burns, D. A., Heap, J. T., & Minton, N. P. (2010). *Clostridium difficile* spore germination: an update. Research in Microbiology, 161(9), 730-734. doi: 10.1016/j.resmic.2010.09.007

Burns, D. A., Heap, J. T., & Minton, N. P. (2010). SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. Journal of Bacteriology, 192(3), 657-664. doi:10.1128/jb.01209-09

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. doi:10.1016/j.anaerobe.2010.10.001

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. doi:http://dx.doi.org/10.1016/j.anaerobe.2010.10.001

Burns, D. A., Heeg, D., Cartman, S. T., & Minton, N. P. (2011). Reconsidering the Sporulation Characteristics of Hypervirulent *Clostridium difficile* BI/NAP1/027. PLOS ONE, 6(9), e24894

Burns, D. A., Heeg, D., Cartman, S. T., & Minton, N. P. (2011). Reconsidering the Sporulation Characteristics of Hypervirulent *Clostridium difficile* BI/NAP1/027. PLOS ONE, 6(9), e24894

Cano, R. J., & Borucki, M. K. (1995). Revival and Identification of Bacterial-Spores In 25-Million-Year-Old To 40-Million-Year-Old Dominican Amber. Science, 268(5213), 1060-1064. doi:10.1126/science.7538699

Carlson, P. E., Jr., Kaiser, A. M., McColm, S. A., Bauer, J. M., Young, V. B., Aronoff, D. M., & Hanna, P. C. (2015). Variation in germination of *Clostridium difficile* clinical isolates correlates to disease severity. Anaerobe, 33, 64-70. doi:10.1016/j.anaerobe.2015.02.003

Carroll, K. C., & Bartlett, J. G. (2011). Biology of *Clostridium difficile*: Implications for Epidemiology and Diagnosis. In S. Gottesman & C. S. Harwood (Eds.), Annual Review of Microbiology, Vol 65 (Vol. 65, pp. 501-521)

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. doi:10.1128/aem.02525-09

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. doi:10.1128/aem.00249-12

Cartman, S., Heap, J., Kuehne, S., Cockayne, A., & Minton, N. (2010) The emergence of 'hypervirulence' in *Clostridium difficile*. International Journal of Medical Microbiology, 300, 387-395

CasillasMartinez, L., & Setlow, P. (1997). Alkyl hydroperoxide reductase, catalase, MrgA, and superoxide dismutase are not involved in resistance of *Bacillus subtilis* spores to heat or oxidizing agents. Journal of Bacteriology, 179(23), 7420-7425. doi:10.1128/jb.179.23.7420-7425.1997

Chang, C. (2008). M1175 Mice with Gut Flora Reduced by Antibiotics Suffer Increased Intestinal Injury After Gamma-Irradiation. Gastroenterology, 134(4, Supplement 1), A-354

Chaudhari, N. M., Gupta, V. K., & Dutta, C. (2016). BPGA- an ultra-fast pan-genome analysis pipeline. Scientific Reports, 6, 24373. doi:10.1038/srep24373

Chen, X., Katchar, K., Goldsmith, J. D., Nanthakumar, N., Cheknis, A., Gerding, D. N., & Kelly, C. P. (2008). A Mouse Model of *Clostridium difficile*-Associated Disease. Gastroenterology, 135(6), 1984-1992

Childress, K. O., Edwards, A. N., Nawrocki, K. L., Anderson, S. E., Woods, E. C., & McBride, S. M. (2016). The Phosphotransfer Protein CD1492 Represses Sporulation Initiation in *Clostridium difficile*. Infection and Immunity, 84(12), 3434-3444. doi:10.1128/iai.00735-16

Chirakkal, H., O'Rourke, M., Atrih, A., Foster, S. J., & Moir, A. (2002). Analysis of spore cortex lytic enzymes and related proteins in *Bacillus subtilis* endospore germination. Microbiology-SGM, 148, 2383-2392. doi:10.1099/00221287-148-8-2383

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., & McDonald, L. C. (2010). Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). Infection Control & Hospital Epidemiology , 31. doi:10.1086/651706

Coleman, W. H., De, C., Li, Y.-q., Cowan, A. E., & Setlow, P. (2007). How moist heat kills spores of *Bacillus subtilis*. Journal of Bacteriology, 189(23), 8458-8466. doi:10.1128/jb.01242-07

Collie, R. E., & McClane, B. A. (1998). Evidence that the enterotoxin gene can be episomal in *Clostridium perfringens* isolates associated with non-food-borne human gastrointestinal diseases. Journal of Clinical Microbiology, 36(1), 30-36

Comely, O. A., Miller, M. A., Louie, T. J., Crook, D. W., & Gorbach, S. L. (2012). Treatment of First Recurrence of *Clostridium difficile* Infection: Fidaxomicin Versus Vancomycin. Clinical Infectious Diseases, 55, S154-S161. doi:10.1093/cid/cis462

Cortezzo, D. E., & Setlow, P. (2005). Analysis of factors that influence the sensitivity of spores of *Bacillus subtilis* to DNA damaging chemicals. Journal of Applied Microbiology, 98(3), 606-617. doi:10.1111/j.1365-2672.2004.02495.x

Cortezzo, D. E., Koziol-Dube, K., Setlow, B., & Setlow, P. (2004). Treatment with oxidizing agents damages the inner membrane of spores of *Bacillus subtilis* and sensitizes spores to subsequent stress. Journal of Applied Microbiology, 97(4), 838-852. doi:10.1111/j.1365-2672.2004.02370.x

Cowan, A. E., Koppel, D. E., Setlow, B., & Setlow, P. (2003). A soluble protein is immobile in dormant spores of *Bacillus subtilis* but is mobile in germinated spores: Implications for spore dormancy. Proceedings of the National Academy of Sciences of the United States of America, 100(7), 4209-4214. doi:10.1073/pnas.0636762100

Cowan, A. E., Olivastro, E. M., Koppel, D. E., Loshon, C. A., Setlow, B., & Setlow, P. (2004). Lipids in the inner membrane of dormant spores of Bacillus species are largely immobile. Proceedings of the National Academy of Sciences of the United States of America, 101(20), 7733-7738. doi:10.1073/pnas.0306859101

Craftslighty, A., & Ellar, D. J. (1980). Structure and function of the spore outer-membrane in dormant and germinating spores of *Bacillus-megaterium*. Journal of Applied Bacteriology, 48(1), 135-145. doi:10.1111/j.1365-2672.1980.tb05215.x

Cybulski, R. J., Jr., Sanz, P., Alem, F., Stibitz, S., Bull, R. L., & O'Brien, A. D. (2009). Four Superoxide Dismutases Contribute to Bacillus anthracis Virulence and Provide Spores with Redundant Protection from Oxidative Stress. Infection and Immunity, 77(1), 274-285. doi:10.1128/iai.00515-08

Daniel, R. A., Drake, S., Buchanan, C. E., Scholle, R., & Errington, J. (1994). The *Bacillus-subtilis spoVD* gene encodes a mother-cell-specific penicillin-binding protein required for spore

morphogenesis. Journal of Molecular Biology, 235(1), 209-220. doi:10.1016/s0022-2836(05)80027-0

Daniels, L. M., & Kufel, W. D. (2018). Clinical review of *Clostridium difficile* infection: an update on treatment and prevention. Expert Opinion on Pharmacotherapy, 19(16), 1759-1769. doi:10.1080/14656566.2018.1524872

Darkoh, C., Lichtenberger, L. M., Ajami, N., Dial, E. J., Jiang, Z.-D., & DuPont, H. L. (2010). Bile Acids Improve the Antimicrobial Effect of Rifaximin. Antimicrobial Agents and Chemotherapy, 54(9), 3618-3624. doi:10.1128/aac.00161-10

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., Wren, B. W., Fairweather, N. F., Dougan, G., & Lawley, T. D. (2012). The *Clostridium difficile spo0A* gene is a persistence and transmission factor. Infection and Immunity, 80(8), 2704-2711. doi:10.1128/iai.00147-12

Debast, S. B., Bauer, M. P., & Kuijper, E. J. (2014). European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for *Clostridium difficile* infection. Clinical Microbiology and Infection, 20 Suppl 2, 1-26. doi:10.1111/1469-0691.12418

de la Puebla, H. B., Giacalone, D., Cooper, A., & Shen, A. (2020). Role of SpoIVA ATPase Motifs During *Clostridioides difficile* Sporulation. Journal of Bacteriology, JB.00387-20. Advance online publication. https://doi.org/10.1128/JB.00387-20

Dembek, M., Stabler, R. A., Witney, A. A., Wren, B. W., & Fairweather, N. F. (2013). Transcriptional analysis of temporal gene expression in germinating *Clostridium difficile* 630 endospores. PLOS ONE, 8(5), e64011. doi:10.1371/journal.pone.0064011

Dembek, M., Stabler, R. A., Witney, A. A., Wren, B. W., & Fairweather, N. F. (2013). Transcriptional analysis of temporal gene expression in germinating *Clostridium difficile* 630 endospores. PLOS ONE, 8(5), e64011. doi:10.1371/journal.pone.0064011

Dendukuri, N., Costa, V., Mcgregor, M. & Brophy, J. M. (2005). Probiotic therapy for the prevention and treatment of *Clostridium difficile*-associated diarrhea: a systematic review. Canadian Medical Association Journal, 173, 167-170

Deutscher, J., Francke, C., & Postma, P. W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiology and Molecular Biology Reviews, 70(4), 939-+. doi:10.1128/mmbr.00024-06

Didelot, X., Eyre, D. W., Cule, M., Ip, C. L., Ansari, M. A., Griffiths, D., Vaughan, A., O'Connor, L., Golubchik, T., Batty, E. M., Piazza, P., Wilson, D. J., Bowden, R., Donnelly, P. J., Dingle, K. E., Wilcox, M., Walker, A. S., Crook, D. W., Peto, T. E., & Harding, R. M. (2012). Microevolutionary analysis of *Clostridium difficile* genomes to investigate transmission. Genome biology, 13(12), R118. https://doi.org/10.1186/gb-2012-13-12-r118

Diaz, O. R., Sayer, C. V., Popham, D. L., & Shen, A. (2018). *Clostridium difficile* Lipoprotein GerS Is Required for Cortex Modification and Thus Spore Germination. mSphere, 3(3), e00205-18. https://doi.org/10.1128/mSphere.00205-18

Diaz-Gonzalez, F., Milano, M., Olguin-Araneda, V., Pizarro-Cerda, J., Castro-Cordova, P., Tzeng, S.-C., Maier, C. S., Sarker, M. R., & Paredes-Sabja, D. (2015). Protein composition of the outermost exosporium-like layer of *Clostridium difficile* 630 spores. Journal of Proteomics, 123, 1-13. doi:10.1016/j.jprot.2015.03.035

Dineen, S. S., Villapakkam, A. C., Nordman, J. T., & Sonenshein, A. L. (2007). Repression of *Clostridium difficile* toxin gene expression by CodY. Molecular Microbiology, 66:206–219.

Dingle, T. C., Mulvey, G. L., & Armstrong, G. D. (2011). Mutagenic Analysis of the *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence in Hamsters. Infection and Immunity, 79(10), 4061-4067. doi:10.1128/iai.05305-11

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. doi:10.1128/jb.00986-15

Donovan, W., Zheng, L., Sandman, K., & Losick, R. (1987). Genes encoding spore coat polypeptides from bacillus-subtilis. Journal of Molecular Biology, 196(1), 1-10. doi:10.1016/0022-2836(87)90506-7

Driks, A. (1999). *Bacillus subtilis* spore coat. Microbiology and Molecular Biology Reviews, 63(1), 1

Dubberke, E. R., Reske, K. A., Noble-Wang, J., Thompson, A., Killgore, G., Mayfield, J., Camins, B., Woeltje, K., McDonald, J. R., McDonald, L. C., & Fraser, V. J. (2007). Prevalence of *Clostridium difficile* environmental contamination and strain variability in multiple health care facilities. American Journal of Infection Control, 35(5), 315-318

Duggan, S. T. (2011). Fidaxomicin: in *Clostridium difficile* infection. Drugs, 71(18), 2445-2456. doi:10.2165/11208220-000000000-00000

Duncan, L., Alper, S., Arigoni, F., Losick, R., & Stragier, P. (1995). Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. Science, 270(5236), 641-644. doi:10.1126/science.270.5236.641

Dupuy, B., & Sonenshein, A. L. (1998). Regulated transcription of *Clostridium difficile* toxin genes. Molecular Microbiology, 27(1), 107-120. doi:10.1046/j.1365-2958.1998.00663.x

Dutilh BE, Backus L, Edwards RA, Wels M, Bayjanov JR, van Hijum SA. (2013). Explaining microbial phenotypes on a genomic scale: GWAS for microbes, Briefings in Functional Genomics, 12(4):366–80

Edwards, A. N., & McBride, S. M. (2014). Initiation of sporulation in *Clostridium difficile*: a twist on the classic model. Federation of European Microbiological Societies Microbiology Letters, 358(2), 110-118. doi:10.1111/1574-6968.12499

Edwards, A. N., Nawrocki, K. L., & McBride, S. M. (2014). Conserved Oligopeptide Permeases Modulate Sporulation Initiation in *Clostridium difficile*. Infection and Immunity, 82(10), 4276-4291. doi:10.1128/iai.02323-14

Edwards, A. N., Tamayo, R., & McBride, S. M. (2016). A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. Molecular Microbiology, 100(6), 954-971. doi:10.1111/mmi.13361

Edwards, D. I. (1993). Nitroimidazole drugs--action and resistance mechanisms. II. Mechanisms of resistance. Journal of Antimicrobial Chemotherapy, 31(2), 201-210. doi:10.1093/jac/31.2.201

Eiseman, B., Silen, W., Bascom, G. S., & Kauvar, A. J. (1958). Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. Surgery, 44(5), 854-859

Errington, J. (1993). *Bacillus-subtilis* sporulation - regulation of gene-expression and control of morphogenesis. Microbiological Reviews, 57(1), 1-33.

Eyre, D. W., Babakhani, F., Griffiths, D., Seddon, J., Del Ojo Elias, C., Gorbach, S. L., Peto, T. E. A., Crook, D. W., & Walker, A. S. (2013). Whole Genome Sequencing Demonstrates Fidaxomicin is Superior to Vancomycin for Prevention of *Clostridium difficile* Relapse and Reinfection. Journal of Infectious Diseases. doi:10.1093/infdis/jit598

Eyre, D. W., Cule, M. L., Griffiths, D., Crook, D. W., Peto, T. E. A., Walker, A. S., & Wilson, D. J. (2013). Detection of Mixed Infection from Bacterial Whole Genome Sequence Data Allows Assessment of Its Role in *Clostridium difficile* Transmission. PLOS Computational Biology, 9(5), e1003059. doi:10.1371/journal.pcbi.1003059

Eyre, D. W., Cule, M. L., Wilson, D. J., Griffiths, D., Vaughan, A., O'Connor, L., Ip, C. L. C., Golubchik, T., Batty, E. M., Finney, J. M., Wyllie, D. H., Didelot, X., Piazza, P., Bowden, R., Dingle, K. E., Harding, R. M., Crook, D. W., Wilcox, M. H., Peto, T. E. A., & Walker, A. S. (2013). Diverse Sources of *C. difficile* Infection Identified on Whole-Genome Sequencing. New England Journal of Medicine, 369(13), 1195-1205. doi:10.1056/NEJMoa1216064

Eyre, D. W., Fawley, W. N., Best, E. L., Griffiths, D., Stoesser, N. E., Crook, D. W., Peto, T. E., Walker, A. S., & Wilcox, M. H. (2013). Comparison of multilocus variable-number tandem-repeat analysis and whole-genome sequencing for investigation of *Clostridium difficile*

transmission. Journal of Clinical Microbiology, *51*(12), 4141–4149. https://doi.org/10.1128/JCM.01095-13

Eyre, D. W., Walker, A. S., Freeman, J., Baines, S. D., Fawley, W. N., Chilton, C. H., Griffiths, D., Vaughan, A., Crook, D. W., Peto, T. E. A., & Wilcox, M. H. (2013). Short-Term Genome Stability of Serial *Clostridium difficile* Ribotype 027 Isolates in an Experimental Gut Model and Recurrent Human Disease. PLOS ONE, 8(5), e63540. doi:10.1371/journal.pone.0063540

Eze, P., Balsells, E., Kyaw, M. H., & Nair, H. (2017). Risk factors for *Clostridium difficile* infections - an overview of the evidence base and challenges in data synthesis. Journal of Global Health, 7(1), 010417-010417. doi:10.7189/jogh.07.010417

Fawley, W. N., Underwood, S., Freeman, J., Baines, S. D., Saxton, K., Stephenson, K., Owens, R. C., & Wilcox, M. H. (2007). Efficacy of Hospital Cleaning Agents and Germicides Against Epidemic *Clostridium difficile* Strains. Infection Control & Hospital Epidemiology, 28(8), 920-925. doi:10.1086/519201

Fazzini, M. M., Schuch, R., & Fischetti, V. A. (2010). A Novel Spore Protein, ExsM, Regulates Formation of the Exosporium in *Bacillus cereus* and *Bacillus anthracis* and Affects Spore Size and Shape. Journal of Bacteriology, 192(15), 4012-4021. doi:10.1128/jb.00197-10

Fimlaid, K. A., & Shen, A. (2015). Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. Current Opinion in Microbiology, 24, 88-95. doi:http://dx.doi.org/10.1016/j.mib.2015.01.006

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). doi:10.1371/journal.pgen.1003660

Fimlaid, K. A., Jensen, O., Donnelly, M. L., Francis, M. B., Sorg, J. A., & Shen, A. (2015). Identification of a Novel Lipoprotein Regulator of *Clostridium difficile* Spore Germination. PLOS Pathogens, 11(10). doi:10.1371/journal.ppat.1005239

Finegold, S. M. (2008). Therapy and epidemiology of autism-clostridial spores as key elements. Medical Hypotheses, 70(3), 508-511

Finnegan, M., Linley, E., Denyer, S. P., McDonnell, G., Simons, C., & Maillard, J.-Y. (2010). Mode of action of hydrogen peroxide and other oxidizing agents: differences between liquid and gas forms. Journal of Antimicrobial Chemotherapy, 65(10), 2108-2115. doi:10.1093/jac/dkq308

Fitz-James, P. (1971). Formation of protoplasts from resting spores. Journal of Bacteriology, 105(3), 1119-1136

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). doi:10.1371/journal.ppat.1003356

Francis, M. B., Allen, C. A., & Sorg, J. A. (2015). Spore Cortex Hydrolysis Precedes Dipicolinic Acid Release during *Clostridium difficile* Spore Germination. Journal of Bacteriology, 197(14), 2276-2283. doi:10.1128/jb.02575-14

Freeman, C. D., Klutman, N. E., & Lamp, K. C. (1997). Metronidazole. A therapeutic review and update. Drugs, 54(5), 679-708. doi:10.2165/00003495-199754050-00003

Fujita, Y. (2009). Carbon Catabolite Control of the Metabolic Network in *Bacillus subtilis*. Bioscience Biotechnology and Biochemistry, 73(2), 245-259. doi:10.1271/bbb.80479

Fujitani, S., George, W. L., Morgan, M. A., Nichols, S., & Murthy, A. R. (2011). Implications for vancomycin-resistant Enterococcus colonization associated with *Clostridium difficile* infections. American Journal of Infection Control, 39(3), 188-193. doi:10.1016/j.ajic.2010.10.024

Garey, K. W., Sethi, S., Yadav, Y., & DuPont, H. L. (2008). Meta-analysis to assess risk factors for recurrent Clostridium difficile infection. Journal of Hospital Infection. 70(4):298-304

Garey, K. W., Ghantoji, S. S., Shah, D. N., Habib, M., Arora, V., Jiang, Z.-D., & DuPont, H. L. (2011). A randomized, double-blind, placebo-controlled pilot study to assess the ability of rifaximin to prevent recurrent diarrhoea in patients with *Clostridium difficile* infection. Journal of Antimicrobial Chemotherapy, 66(12), 2850-2855. doi:10.1093/jac/dkr377

Geoghegan, O., Eades, C., Moore, L. S. & Gilchrist, M. (2017). *Clostridium difficile*: diagnosis and treatment update. Clinical Pharmacist, 9. doi:10.1211/CP.2017.20202242

Gerding, D. N., Muto, C. A., & Owens, R. C., Jr. (2008). Measures to control and prevent *Clostridium difficile* infection. Clinical Infectious Diseases, 46 Suppl 1, S43-49. doi:10.1086/521861

Gil, F., Lagos-Moraga, S., Calderon-Romero, P., Pizarro-Guajardo, M., & Paredes-Sabja, D. (2017). Updates on *Clostridium difficile* spore biology. Anaerobe, 45, 3-9. doi:10.1016/j.anaerobe.2017.02.018

Gilmore, M. E., Dandyopadhyay, D., Dean, A. M., Linnstaedt, S. D., & Popham, D. L. (2004). Production of muramic beta-lactam in *Bacillus subtilis* spore peptidoglycan. Journal of Bacteriology, 186(1), 80-89. doi:10.1128/jb.186.1.80-89.2004

Girinathan, B. P., Ou, J., Dupuy, B., & Govind, R. (2018). Pleiotropic roles of *Clostridium difficile sin* locus. PLoS Pathogens, 14:e1006940.

Goldstein, E. J., Johnson, S., Maziade, P. J., McFarland, L. V., Trick, W., Dresser, L., Millette, M., Mazloum, H., & Low, D. E. (2015). Pathway to Prevention of Nosocomial *Clostridium difficile* Infection. Clinical Infectious Disease, 60 Suppl 2, S148-158. doi:10.1093/cid/civ142

Gould, G. W., Stubbs, J. M., & King, W. L. (1970). Structure and Composition Of Resistant Layers In Bacterial Spore Coats. Journal of General Microbiology, 60, 347-&. doi:10.1099/00221287-60-3-347

Goulding, D., Thompson, H., Emerson, J., Fairweather, N. F., Dougan, G., & Douce, G. R. (2009). Distinctive Profiles of Infection and Pathology in Hamsters Infected with *Clostridium difficile* Strains 630 and B1. Infection and Immunity, 77(12), 5478-5485. doi:10.1128/iai.00551-09

Granum, P. E., Richardson, M., & Blom, H. (1987). Isolation and amino-acid-sequence of an acid-soluble protein from *Clostridium-perfringens* spores. Federation of European Microbiological Societies Microbiology Letters, 42(2-3), 225-230

Haslam, S. C., Ketley, J. M., Mitchell, T. J., Stephen, J., Burdon, D. W., & Candy, D. C. (1986). Growth of *Clostridium difficile* and production of toxins A and B in complex and defined media. Journal of Medical Microbiology, 21(4), 293-297. doi:10.1099/00222615-21-4-293

Hatt, J. K., & Youngman, P. (2000). Mutational analysis of conserved residues in the putative DNA-binding domain of the response regulator Spo0A of *Bacillus subtilis*. Journal of Bacteriology, 182(24), 6975-6982. doi:10.1128/jb.182.24.6975-6982.2000

Hatton, K., Cooley, N., Demertzi, E., & Azadian, B. (2011). The impact of a multidisciplinary clinical review of inpatients with *Clostridium difficile* infection: Category: Lesson in Microbiology & Infection Control. Journal of Infection, 63(6), e82-e83. doi:10.1016/j.jinf.2011.04.143

Hatton, K., Cooley, N., Demertzi, E., & Azadian, B. (2011). The impact of a multidisciplinary clinical review of inpatients with *Clostridium difficile* infection: Category: Lesson in Microbiology & Infection Control. Journal of Infection, 63(6), e82-e83. doi:10.1016/j.jinf.2011.04.143

He, M., Sebaihia, M., Lawley, T. D., Stabler, R. A., Dawson, L. F., Martin, M. J., Holt, K. E., Seth-Smith, H. M. B., Quail, M. A., Rance, R., Brooks, K., Churcher, C., Harris, D., Bentley, S. D., Burrows, C., Clark, L., Corton, C., Murray, V., Rose, G., Thurston, S., van Tonder, A., Walker, D., Wren, B. W., Dougan, G., & Parkhill, J. (2010). Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proceedings of the National Academy of Sciences, 107(16), 7527-7532. doi:10.1073/pnas.0914322107

He, M., Miyajima, F., Roberts, P., Ellison, L., Pickard, D. J., Martin, M. J., Connor, T. R., Harris, S. R., Fairley, D., Bamford, K. B., D'Arc, S., Brazier, J., Brown, D., Coia, J. E., Douce, G., Gerding, D., Kim, H. J., Koh, T. H., Kato, H., Senoh, M., … Lawley, T. D. (2013). Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*. Nature Genetics, 45(1), 109–113. https://doi.org/10.1038/ng.2478

Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y.-K., 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. doi:10.1093/nar/gkr1321

Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C., & Minton, N. P. (2010). The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods, 80(1), 49-55. doi:10.1016/j.mimet.2009.10.018

Heap, J. T., Pennington, O. J., Cartman, S. T., & Minton, N. P. (2009). A modular system for *Clostridium* shuttle plasmids. Journal of Microbiological Methods, 78(1), 79-85. doi:10.1016/j.mimet.2009.05.004

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. doi:10.1016/j.mimet.2007.05.021

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. doi:10.1371/journal.pone.0032381

Higgins, D., & Dworkin, J. (2012). Recent progress in *Bacillus subtilis* sporulation. Federation of European Microbiological Societies Microbiology Reviews, 36(1), 131-148. doi:10.1111/j.1574-6976.2011.00310.x

Hippenstiel, S., Kratz, T., Krüll, M., Seybold, J., Eichel-Streiber, C. v., & Suttorp, N. (1998). Rho Protein Inhibition Blocks Protein Kinase C Translocation and Activation. Biochemical and Biophysical Research Communications, 245(3), 830-834

Ho, J. G., Greco, A., Rupnik, M., & Ng, K. K. (2005). Crystal structure of receptor-binding C-terminal repeats from *Clostridium difficile* toxin A. Proceedings of the National Academy of Sciences of the USA, 102(51), 18373-18378. doi:10.1073/pnas.0506391102

Hofmeister, A. E. M., Londonovallejo, A., Harry, E., Stragier, P., & Losick, R. (1995). Extracellular signal protein triggering the proteolytic activation of a developmental transcription factor in *Bacillus-subtilis*. Cell, 83(2), 219-226. doi:10.1016/0092-8674(95)90163-9

Huang, D. B., & DuPont, H. L. (2005). Rifaximin--a novel antimicrobial for enteric infections. Journal of Infection, 50(2), 97-106. doi:10.1016/j.jinf.2004.05.019

Hubert, B., Loo, V. G., Bourgault, A. M., Poirier, L., Dascal, A., Fortin, E., Dionne, M., & Lorange, M. (2007). A portrait of the geographic dissemination of the *Clostridium difficile* North American pulsed-field type 1 strain and the epidemiology of *C. difficile*-associated disease in Quebec. Clinical Infectious Disease, 44(2), 238-244. doi:10.1086/510391

Hussain, H. A., Roberts, A. P., & Mullany, P. (2005). Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630 Delta erm) and demonstration that the conjugative transposon Tn916 Delta E enters the genome of this strain at multiple sites. Journal of Medical Microbiology, 54(2), 137-141. doi:10.1099/jmm.0.45790-0

Imlay, J. A., & Linn, S. (1988). DNA damage and oxygen radical toxicity. Science, 240(4857), 1302-1309. doi:10.1126/science.3287616

Janoir, C., Denève, C., Bouttier, S., Barbut, F., Hoys, S., Caleechum, L., Chapetón-Montes, D., Pereira, F. C., Henriques, A. O., Collignon, A., Monot, M., & Dupuy, B. (2013). Adaptive strategies and pathogenesis of *Clostridium difficile* from in vivo transcriptomics. Infection and Immunity, 81(10), 3757–3769. https://doi.org/10.1128/IAI.00515-13

Jhung, M. A., Thompson, A. D., Killgore, G. E., Zukowski, W. E., Songer, G., & Warny, M. (2008). Toxinotype V *Clostridium difficile* in humans and food animals. Emerging Infectious Diseases, 14. doi:10.3201/eid1407.071641

Johnson, S. (2009). Recurrent *Clostridium difficile* infection: causality and therapeutic approaches. Int J Antimicrob Agents, 33 Suppl 1, S33-36. doi:10.1016/s0924-8579(09)70014-7

Johnson, S., & Gerding, D. N. (2009). Introduction to the special issue on *Clostridium difficile*. Anaerobe, 15(6), 225-226

Johnson, S., Adelmann, A., Clabots, C. R., Peterson, L. R., & Gerding, D. N. (1989). Recurrences of *Clostridium difficile* diarrhea not caused by the original infecting organism. Journal of Infectious Diseases, 159(2), 340-343. doi:10.1093/infdis/159.2.340

Johnson, S., Schriever, C., Galang, M., Kelly, C. P., & Gerding, D. N. (2007). Interruption of recurrent *Clostridium difficile*-associated diarrhea episodes by serial therapy with vancomycin and rifaximin. Clinical Infectious Disease, 44(6), 846-848. doi:10.1086/511870

Kamboj, M., Khosa, P., Kaltsas, A., Babady, N.E., Son, C. and Sepkowitz, K.A. (2011). Relapse versus reinfection: surveillance of *Clostridium difficile* infection. Clinical Infectious Diseases, 53(10), 1003- 1006

Kamiya, S., Ogura, H., Meng, X. Q., & Nakamura, S. (1992). Correlation between cytotoxin production and sporulation in *Clostridium difficile*. Journal of Medical Microbiology, 37(3), 206-210. doi:10.1099/00222615-37-3-206

Kamiya, S., Yamakawa, K., Ogura, H., & Nakamura, S. (1989). Recovery of spores of *Clostridium difficile* altered by heat or alkali. Journal of Medical Microbiology, 28(3), 217-221. doi:10.1099/00222615-28-3-217

Karasawa, T., Ikoma, S., Yamakawa, K., & Nakamura, S. (1995). A defined growth medium for *Clostridium difficile*. Microbiology, 141(2), 371-375. doi:10.1099/13500872-141-2-371

Karlsson, S., Burman, L. G., & Ã…kerlund, T. (1999). Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. Microbiology, 145(7), 1683-1693. doi:10.1099/13500872-145-7-1683

Karow, M. L., Glaser, P., & Piggot, P. J. (1995). Identification of a gene, spoiir that links the activation of sigma(e) to the transcriptional activity of sigma(f) during sporulation in *Bacillus-subtilis*. Proceedings of the National Academy of Sciences of the United States of America, 92(6), 2012-2016. doi:10.1073/pnas.92.6.2012

Kennedy, M. J., Reader, S. L., & Swierczynski, L. M. (1994). Preservation records of microorganisms - evidence of the tenacity of life. Microbiology-SGM, 140, 2513-2529. doi:10.1099/00221287-140-10-2513

Kent, W. (2002). BLAT--the BLAST-like alignment tool. Genome Research, 12, 656 - 664

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. Genome Research, 12(6), 996-1006. doi:10.1101/gr.229102

Kevorkian, Y., & Shen, A. (2017). Revisiting the Role of Csp Family Proteins in Regulating *Clostridium difficile* Spore Germination. Journal of Bacteriology, 199(22). doi:10.1128/jb.00266-17

Kevorkian, Y., Shirley, D. J., & Shen, A. (2016). Regulation of *Clostridium difficile* spore germination by the CspA pseudoprotease domain. Biochimie, 122, 243-254. doi:10.1016/j.biochi.2015.07.023

Keynan A, Evenchick Z. 1969. Activation, p 359–396 In Gould GW, Hurst A, editors. (ed), The Bacterial Spore. Academic Press, London, United Kingdom

Khanna, S., Pardi, D. S., Aronson, S. L., Kammer, P. P., & Baddour, L. M. (2012). Outcomes in community-acquired *Clostridium difficile* infection. Alimentary Pharmacology & Therapeutics, n/a-n/a. doi:10.1111/j.1365-2036.2011.04984.x

Khoruts, A., Dicksved, J., Jansson, J. K., & Sadowsky, M. J. (2010). Changes in the Composition of the Human Fecal Microbiome After Bacteriotherapy for Recurrent *Clostridium difficile*-associated Diarrhea. Journal of Clinical Gastroenterology, 44(5), 354-360 310.1097/MCG.1090b1013e3181c1087e1002

Killgore, G., Thompson, A., Johnson, S., Brazier, J., Kuijper, E., Pepin, J., Frost, E. H., Savelkoul, P., Nicholson, B., van den Berg, R. J., Kato, H., Sambol, S. P., Zukowski, W., Woods, C., Limbago, B., Gerding, D. N., & McDonald, L. C. (2008). Comparison of Seven Techniques for Typing International Epidemic Strains of *Clostridium difficile*: Restriction Endonuclease Analysis, Pulsed-Field Gel Electrophoresis, PCR-Ribotyping, Multilocus Sequence Typing, Multilocus Variable-Number Tandem-Repeat Analysis, Amplified Fragment Length Polymorphism, and Surface Layer Protein A Gene Sequence Typing. Journal of Clinical. Microbiology., 46(2), 431-437. doi:10.1128/jcm.01484-07

Kling, P. A., & Burman, L. G. (1989). Serum and tissue pharmacokinetics of intravenous metronidazole in surgical patients. Acta Chirurgica Scandinavica, 155(6-7), 347-350

Knight, D. R., & Riley, T. V. (2019). Genomic Delineation of Zoonotic Origins of *Clostridium difficile*. Frontiers in Public Health, 7(164). doi:10.3389/fpubh.2019.00164

Knight, D. R., Elliott, B., Chang, B. J., Perkins, T. T., & Riley, T. V. (2015). Diversity and Evolution in the Genome of *Clostridium difficile*. Clinical Microbiology Reviews, 28(3), 721-741. doi:10.1128/cmr.00127-14

Knight, D. R., Thean, S., Putsathit, P., Fenwick, S., & Riley, T. V. (2013). Cross-Sectional Study Reveals High Prevalence of Non-PCR Ribotype 078 Strains in Australian Veal Calves at Slaughter. Applied and Environmental Microbiology, 79(8), 2630-2635. doi:10.1128/aem.03951-12

Kochan, T. J., Foley, M. H., Shoshiev, M. S., Somers, M. J., Carlson, P. E., & Hanna, P. C. (2018). Updates to *Clostridium difficile* Spore Germination. Journal of Bacteriology, 200(16), e00218-00218. doi:10.1128/jb.00218-18

Kochan, T. J., Somers, M. J., Kaiser, A. M., Shoshiev, M. S., Hagan, A. K., Hastie, J. L., Giordano, N. P., Smith, A. D., Schubert, A. M., Carlson, P. E., Jr., & Hanna, P. C. (2017). Intestinal calcium and bile salts facilitate germination of *Clostridium difficile* spores. PLOS Pathogens, 13(7). doi:10.1371/journal.ppat.1006443

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A., & Minton, N. P. (2010). The role of toxin A and toxin B in *Clostridium difficile* infection. Nature, 467(7316), 711-U797. doi:10.1038/nature09397

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. doi:10.1093/infdis/jit426

Kuijper, E. J., & Monnet, D. L. (2008). The spread of hypervirulent *Clostridium difficile* PCR Ribotype 027 in Europe. International Journal of Infectious Diseases, 12(Supplement 1), e22-e22

Kuijper, E. J., Coignard, B., Brazier, J. S., Suetens, C., Drudy, D., Wiuff, C., Pituch, H., Reichert, P., Schneider, F., Widmer, A. F., Olsen, K. E., Allerberger, F., Notermans, D. W., Barbut, F., Delmee, M., Wilcox, M., Pearson, A., Patel, B. C., Brown, D. J., Frei, R., Akerlund, T., Poxton, I. R., & Tull, P. (2007). Update of *Clostridium difficile*-associated disease due to PCR ribotype 027 in Europe. European Surveillance, 12, E1 - 2

Kuijper, E. J., Coignard, B., & Tull, P. (2006). Emergence of *Clostridium difficile*-associated disease in North America and Europe. Clinical Microbiology and Infection, 12 Suppl 6, 2-18. doi:10.1111/j.1469-0691.2006.01580.x

Kumar, N., Miyajima, F., He, M., Roberts, P., Swale, A., Ellison, L., Pickard, D., Smith, G., Molyneux, R., Dougan, G., Parkhill, J., Wren, B. W., Parry, C. M., Pirmohamed, M., & Lawley, T. D. (2016). Genome-Based Infection Tracking Reveals Dynamics of *Clostridium difficile* Transmission and Disease Recurrence. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, 62(6), 746–752. https://doi.org/10.1093/cid/civ1031

Kyne, L., Sougioultzis, S., McFarland, L. V., & Kelly, C. n. P. (2002). Underlying Disease Severity as a Major Risk Factor for Nosocomial *Clostridium difficile* Diarrhea Infection Control and Hospital Epidemiology, 23(11), 653-659. doi:10.1086/501989

Lamanna, A., & Orsi, A. (1984). In vitro activity of rifaximin and rifampicin against some anaerobic bacteria. Chemioterapia, 3(6), 365-367

Lawley, T. D., Clare, S., Walker, A. W., Goulding, D., Stabler, R. A., Croucher, N., Mastroeni, P., Scott, P., Raisen, C., Mottram, L., Fairweather, N. F., Wren, B. W., Parkhill, J., & Dougan, G. (2009). Antibiotic Treatment of *Clostridium difficile* Carrier Mice Triggers a Supershedder State, Spore-Mediated Transmission, and Severe Disease in Immunocompromised Hosts. Infection and Immunity, 77(9), 3661-3669. doi:10.1128/iai.00558-09

Lawley, T. D., Clare, S., Walker, A. W., Stares, M. D., Connor, T. R., Raisen, C., Goulding, D., Rad, R., Schreiber, F., Brandt, C., Deakin, L. J., Pickard, D. J., Duncan, S. H., Flint, H. J., Clark, T. G., Parkhill, J., & Dougan, G. (2012). Targeted Restoration of the Intestinal Microbiota with a

Simple, Defined Bacteriotherapy Resolves Relapsing *Clostridium difficile* Disease in Mice. PLoS Pathogens, 8(10), e1002995. doi:10.1371/journal.ppat.1002995

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D. J., Parkhill, J., Choudhary, J., & Dougan, G. (2009). Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. Journal of Bacteriology, 191(17), 5377-5386. doi:10.1128/jb.00597-09

Leal, J., Gregson, D. B., Ross, T., Church, D. L., & Laupland, K. B. (2008). Epidemiology of Clostridium species bacteremia in Calgary, Canada, 2000-2006. Journal of Infection, 57(3), 198-203

Leeds, J. A., Sachdeva, M., Mullin, S., Barnes, S. W., & Ruzin, A. (2013). In vitro selection, via serial passage, of *Clostridium difficile* mutants with reduced susceptibility to fidaxomicin or vancomycin. Journal of Antimicrobial Chemotherapy, 69(1), 41-44. doi:10.1093/jac/dkt302

Leggett, M. J., McDonnell, G., Denyer, S. P., Setlow, P., & Maillard, J.-Y. (2012). Bacterial spore structures and their protective role in biocide resistance. Journal of Applied Microbiology, 113(3), 485-498. doi:10.1111/j.1365-2672.2012.05336.x

Leyva-Illades, J. F., Setlow, B., Sarker, M. R., & Setlow, P. (2007). Effect of a small, acid-soluble spore protein from *Clostridium perfringens* on the resistance properties of *Bacillus subtilis* spores. Journal of Bacteriology, 189(21), 7927-7931. doi:10.1128/jb.01179-07

Li, H., Ruan, J., & Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research, 18, 1851 - 1858

Li, J., & McClane, B. A. (2008). A Novel Small Acid Soluble Protein Variant Is Important for Spore Resistance of Most *Clostridium perfringens* Food Poisoning Isolates. PLOS Pathogens, 4(5), e1000056. doi:10.1371/journal.ppat.1000056

Löfmark, S., Edlund, C., & Nord, C. E. (2010). Metronidazole Is Still the Drug of Choice for Treatment of Anaerobic Infections. Clinical Infectious Diseases, 50(Supplement 1), S16-S23. doi:10.1086/647939

Loll, P. J., & Axelsen, P. H. (2000). The structural biology of molecular recognition by vancomycin. Annual Review Biophysics and Biomolecular Structure, 29, 265-289. doi:10.1146/annurev.biophys.29.1.265

Londonovallejo, J. A., & Stragier, P. (1995). Cell-cell signaling pathway activating a developmental transcription factor in *Bacillus-subtilis*. Genes & Development, 9(4), 503-508. doi:10.1101/gad.9.4.503

Loo, V. G., Poirier, L., Miller, M. A., Oughton, M., Libman, M. D., Michaud, S., Bourgault, A. M., Nguyen, T., Frenette, C., Kelly, M., Vibien, A., Brassard, P., Fenn, S., Dewar, K., Hudson, T. J., Horn, R., Rene, P., Monczak, Y., & Dascal, A. (2005a). A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. New England Journal of Medicine, 353, 2442 - 2449

Loo, V. G., Poirier, L., Miller, M. A., Oughton, M., Libman, M. D., Michaud, S., Bourgault, A. M., Nguyen, T., Frenette, C., Kelly, M., Vibien, A., Brassard, P., Fenn, S., Dewar, K., Hudson, T. J., Horn, R., Rene, P., Monczak, Y., & Dascal, A. (2005b). A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. New England Journal of Medicine, 353(23), 2442-2449. doi:10.1056/NEJMoa051639

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 alpha/beta-type protect spores against this DNA damage. Journal of Applied Microbiology, 87(1), 8-14. doi:10.1046/j.1365-2672.1999.00783.x

Lyerly, D. M., Saum, K. E., MacDonald, D. K., & Wilkins, T. D. (1985). Effects of *Clostridium difficile* toxins given intragastrically to animals. Infection and Immunity, 47(2), 349-352

Lyras, D., O'Connor, J. R., Howarth, P. M., Sambol, S. P., Carter, G. P., Phumoonna, T., Poon, R., Adams, V., Vedantam, G., Johnson, S., Gerding, D. N., & Rood, J. I. (2009). Toxin B is essential for virulence of *Clostridium difficile*. Nature, 458(7242), 1176-1181. doi:10.1038/nature07822

Maiden, M., Bygraves, J., Feil, E., Morelli, G., Russell, J., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., & Caugant, D. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America, 95, 3140 - 3145

Major, G., Bradshaw, L., Boota, N., Sprange, K., Diggle, M., Montgomery, A., Jawhari, A., & Spiller, R. C. (2019). Follow-on Rifaximin for the Prevention of recurrence following standard treatment of Infection with *Clostridium difficile* (RAPID): a randomised placebo controlled trial. Gut, 68(7), 1224-1231. doi:10.1136/gutjnl-2018-316794

Mani, N., and Dupuy, B. (2001). Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proceedings of the National Academy of Sciences of the United States of America, 98, 5844-5849

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. doi:10.1128/jcm.02291-12

Mason, J. M., & Setlow, P. (1986). Essential role of small, acid-soluble spore proteins in resistance of *Bacillus subtilis* spores to UV light. Journal of Bacteriology, 167(1), 174-178. doi:10.1128/jb.167.1.174-178.1986

Massey, S., Yeager, L. A., Blumentritt, C. A., Vijayakumar, S., Sbrana, E., Peterson, J. W., Brasel, T., LeDuc, J. W., Endsley, J. J., & Torres, A. G. (2014). Comparative *Burkholderia* pseudomallei natural history virulence studies using an aerosol murine model of infection. Scientific Reports, 4. doi:10.1038/srep04305

Matamouros, S., England, P., & Dupuy, B. (2007). *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Molecular Microbiology, 64(5), 1274-1288. doi:10.1111/j.1365-2958.2007.05739.x

McBride, S. M., & Sonenshein, A. L. (2011). Identification of a Genetic Locus Responsible for Antimicrobial Peptide Resistance in *Clostridium difficile*. Infection and Immunity, 79(1), 167-176. doi:10.1128/iai.00731-10

McDonald, C. (2005). *Clostridium difficile*: responding to a new threat from an old enemy. Infection Control and Epidemiology, 26, 672 - 675

McEllistrem, M., Carman, R., Gerding, D., Genheimer, C., & Zheng, L. (2005). A hospital outbreak of *Clostridium difficile* disease associated with isolates carrying binary toxin genes. Clinical Infectious Diseases, 40, 265 - 272.

Meador-Parton, J., & Popham, D. L. (2000). Structural analysis of *Bacillus subtilis* spore peptidoglycan during sporulation. Journal of Bacteriology, 182(16), 4491-4499. doi:10.1128/jb.182.16.4491-4499.2000

Merrigan, M., Venugopal, A., Mallozzi, M., Roxas, B., Viswanathan, V. K., Johnson, S., Gerding, D. N., & Vedantam, G. (2010). Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. Journal of Bacteriology, 192(19), 4904-4911. doi:JB.00445-10 [pii]

Miller, M. A. (2007). Clinical Management of *Clostridium difficile*-Associated Disease. Clinical Infectious Diseases, 45(Supplement_2), S122-S128. doi:10.1086/519257

Minton, N. P., Mauchline, M. L., Lemmon, M. J., Brehm, J. K., Fox, M., Michael, N. P., Giaccia, A., & Brown, J. M. (1995). Chemotherapeutic tumour targeting using clostridial spores. Federation of European Microbiological Societies Microbiology Reviews, 17(3), 357-364. doi:10.1111/j.1574-6976.1995.tb00219.x

Moeller, R., Setlow, P., Horneck, G., Berger, T., Reitz, G., Rettberg, P., Doherty, A. J., Okayasu, R., & Nicholson, W. L. (2008). Roles of the major, small, acid-soluble spore proteins and spore-specific and universal DNA repair mechanisms in resistance of *Bacillus subtilis* spores to ionizing radiation from x rays and high-energy charged-particle bombardment. Journal of Bacteriology, 190(3), 1134-1140. doi:10.1128/jb.01644-07

Moir, A. (2006). How do spores germinate? Journal of Applied Microbiology, 101(3), 526-530. doi:10.1111/j.1365-2672.2006.02885.x

Moore, P., Kyne, L., Martin, A., & Solomon, K. (2013). Germination efficiency of clinical *Clostridium difficile* spores and correlation with ribotype, disease severity and therapy failure. Journal of Medical Microbiology, 62(Pt 9), 1405-1413. doi:10.1099/jmm.0.056614-0

Moraes, P. M., Seyffert, N., Silva, W. M., Castro, T. L., Silva, R. F., Lima, D. D., Hirata, R., Jr., Silva, A., Miyoshi, A., & Azevedo, V. (2014). Characterization of the Opp peptide transporter of *Corynebacterium pseudotuberculosis* and its role in virulence and pathogenicity. Biomed Research International, 2014, 489782. doi:10.1155/2014/489782

Mormak, D. A., & Casida, L. E. (1985). Study of *Bacillus-subtilis* endospores in soil by use of a modified endospore stain. Applied and Environmental Microbiology, 49(6), 1356-1360

Muller, M. (1983). Mode of action of metronidazole on anaerobic bacteria and protozoa. Surgery, 93(1 Pt 2), 165-171

Mullish, B. H., & Williams, H. R. (2015). Obstacles to establishing an NHS faecal transplant programme. British Medical Journal, 351, h6043. doi:10.1136/bmj.h6043

Munasinghe, P. C., & Khanal, S. K. (2010). Biomass-derived syngas fermentation into biofuels: Opportunities and challenges. Bioresource Technology, 101(13), 5013-5022. doi:10.1016/j.biortech.2009.12.098

Murphy, C. K., Mullin, S., Osburne, M. S., van Duzer, J., Siedlecki, J., Yu, X., Kerstein, K., Cynamon, M., & Rothstein, D. M. (2006). In Vitro Activity of Novel Rifamycins against Rifamycin-Resistant *Staphylococcus aureus*. Antimicrobial Agents and Chemotherapy, 50(3), 827-834. doi:10.1128/aac.50.3.827-834.2006

Myers, G. S., Rasko, D. A., Cheung, J. K., Ravel, J., Seshadri, R., DeBoy, R. T., Ren, Q., Varga, J., Awad, M. M., Brinkac, L. M., Daugherty, S. C., Haft, D. H., Dodson, R. J., Madupu, R., Nelson, W. C., Rosovitz, M. J., Sullivan, S. A., Khouri, H., Dimitrov, G. I., Watkins, K. L., Mulligan, S., Benton, J., Radune, D., Fisher, D. J., Atkins, H. S., Hiscox, T., Jost, B. H., Billington, S. J., Songer, J. G., McClane, B. A., Titball, R. W., Rood, J. I., Melville, S. B., & Paulsen, I. T. (2006). Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. Genome Research, 16(8), 1031-1040. doi:10.1101/gr.5238106

Nakamura, S., Yamakawa, K., Izumi, J., Nakashio, S., & Nishida, S. (1985). Germinability and heat resistance of spores of *Clostridium difficile* strains. Microbiology and Immunology, 29(2), 113-118

Nathwani, D., Cornely, O. A., Van Engen, A. K., Odufowora-Sita, O., Retsa, P., & Odeyemi, I. A. (2014). Cost-effectiveness analysis of fidaxomicin versus vancomycin in *Clostridium difficile* infection. Journal of Antimicrobial Chemotherapy, 69(11), 2901-2912. doi:10.1093/jac/dku257

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. doi:10.1128/jb.00220-16

Neal, M. D., Alverdy, J. C., Hall, D. E., Simmons, R. L., & Zuckerbraun, B. S. (2011). Diverting loop ileostomy and colonic lavage: an alternative to total abdominal colectomy for the treatment of severe, complicated *Clostridium difficile* associated disease. Annals of Surgery, 254(3), 423-427; discussion 427-429. doi:10.1097/SLA.0b013e31822ade48

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., Cartman, S. T., & 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. doi:10.1371/journal.pone.0056051

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-+. doi:10.1128/mmbr.64.3.548-572.2000

Noren, T., Akerlund, T., Back, E., Sjoberg, L., Persson, I., Alriksson, I., & Burman, L. G. (2004). Molecular epidemiology of hospital-associated and community-acquired *Clostridium difficile* infection in a Swedish county. Journal of Clinical Microbiology, 42(8), 3635-3643. doi:10.1128/jcm.42.8.3635-3643.2004

O'Connor, J. R., Galang, M. A., Sambol, S. P., Hecht, D. W., Vedantam, G., Gerding, D. N., & Johnson, S. (2008). Rifampin and Rifaximin Resistance in Clinical Isolates of *Clostridium*

difficile. Antimicrobial Agents and Chemotherapy, 52(8), 2813-2817. doi:10.1128/aac.00342-08

Office for National Statistics. (2006). Deaths involving *Clostridium difficile*: England and Wales, 2006 to 2010. Retrieved from http://www.ons.gov.uk/ons/rel/subnational-health2/deaths-involving-clostridium-difficile/2006-to-2010/statistical-bulletin.html

Oka, K., Osaki, T., Hanawa, T., Kurata, S., Okazaki, M., Manzoku, T., Takahashi, M., Tanaka, M., Taguchi, H., Watanabe, T., Inamatsu, T., & Kamiya, S. (2011). Molecular and microbiological characterization of *Clostridium difficile* isolates from single, relapse, and re-infection cases. Journal of Clinical Microbiology. doi:10.1128/jcm.05588-11

Oka, K., Osaki, T., Hanawa, T., Kurata, S., Okazaki, M., Manzoku, T., Takahashi, M., Tanaka, M., Taguchi, H., Watanabe, T., Inamatsu, T., & Kamiya, S. (2012). Molecular and microbiological characterization of *Clostridium difficile* isolates from single, relapse, and reinfection cases. Journal of Clinical Microbiology, 50(3), 915-921. doi:10.1128/JCM.05588-11

Osgood, D. P., Wood, N. P., & Sperry, J. F. (1993). Nutritional aspects of cytotoxin production by *Clostridium difficile*. Applied and Environmental Microbiology, 59(12), 3985-3988

Paidhungat, M., Setlow, B., Daniels, W. B., Hoover, D., Papafragkou, E., & Setlow, P. (2002). Mechanisms of induction of germination of *Bacillus subtilis* spores by high pressure. Applied and Environmental Microbiology, 68(6), 3172-3175. doi:10.1128/aem.68.6.3172-3175.2002

Paidhungat, M., Setlow, B., Driks, A., & Setlow, P. (2000). Characterization of spores of *Bacillus subtilis* which lack dipicolinic acid. Journal of Bacteriology, 182(19), 5505-5512. doi:10.1128/jb.182.19.5505-5512.2000

Paredes-Sabja, D., Bond, C., Carman, R. J., Setlow, P., & Sarker, M. R. (2008). Germination of spores of *Clostridium difficile* strains, including isolates from a hospital outbreak of *Clostridium difficile*-associated disease (CDAD). Microbiology, 154(8), 2241-2250. doi:10.1099/mic.0.2008/016592-0

Paredes-Sabja, D., Raju, D., Torres, J. A., & Sarker, M. R. (2008). Role of small, acid-soluble spore proteins in the resistance of *Clostridium perfringens* spores to chemicals. International Journal of Food Microbiology, 122(3), 333-335. doi:10.1016/j.ijfoodmicro.2007.12.006

Paredes-Sabja, D., Sarker, N., Setlow, B., Setlow, P., & Sarker, M. R. (2008). Roles of *dacB* and *spm* proteins in *Clostridium perfringens* spore resistance to moist heat, chemicals, and UV radiation. Applied and Environmental Microbiology, 74(12), 3730-3738. doi:10.1128/aem.00169-08

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. doi:10.1128/jb.00325-08

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends in Microbiology, 22(7), 406-416. doi:http://dx.doi.org/10.1016/j.tim.2014.04.003

Paredes-Sabja, D., Torres, J. A., Setlow, P., & Sarker, M. R. (2008). *Clostridium perfringens* spore germination: Characterization of germinants and their receptors. Journal of Bacteriology, 190(4), 1190-1201. doi:10.1128/jb.01748-07

Pelczar, P. L., Igarashi, T., Setlow, B., & Setlow, P. (2007). Role of GerD in germination of *Bacillus subtilis* spores. Journal of Bacteriology, 189(3), 1090-1098. doi:10.1128/jb.01606-06

Peng, Z., Addisu, A., Alrabaa, S., & Sun, X. (2017). Antibiotic Resistance and Toxin Production of *Clostridium difficile* Isolates from the Hospitalized Patients in a Large Hospital in Florida. Frontiers in Microbiology, 8, 2584. doi:10.3389/fmicb.2017.02584

Peng, Z., Ling, L., Stratton, C. W., Li, C., Polage, C. R., Wu, B., & Tang, Y. W. (2018). Advances in the diagnosis and treatment of *Clostridium difficile* infections. Emerging Microbes & Infections, 7(1), 15. doi:10.1038/s41426-017-0019-4

Pepin, J., Valiquette, L., Alary, M. E., Villemure, P., Pelletier, A., Forget, K., Pepin, K., & Chouinard, D. (2004). *Clostridium difficile*-associated diarrhea in a region of Quebec from

1991 to 2003: a changing pattern of disease severity. Canadian Medical Association Journal, 171(5), 466-472. doi:10.1503/cmaj.1041104

Permpoonpattana, P., Phetcharaburanin, J., Mikelsone, A., Dembek, M., Tan, S., Brisson, M.-C., La Ragione, R., Brisson, A. R., Fairweather, N., Hong, H. A., & Cutting, S. M. (2013). Functional Characterization of *Clostridium difficile* Spore Coat Proteins. Journal of Bacteriology, 195(7), 1492-1503. doi:10.1128/jb.02104-12

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. doi:10.1128/jb.05182-11

Phetcharaburanin, J., Hong, H. A., Colenutt, C., Bianconi, I., Sempere, L., Permpoonpattana, P., Smith, K., Dembek, M., Tan, S., Brisson, M. C., Brisson, A. R., Fairweather, N. F., & Cutting, S. M. (2014). The spore-associated protein BclA1 affects the susceptibility of animals to colonization and infection by *Clostridium difficile*. Molecular Microbiology, 92(5), 1025–1038. https://doi.org/10.1111/mmi.12611

Pizarro-Guajardo, M., Olguín-Araneda, V., Barra-Carrasco, J., Brito-Silva, C., Sarker, M. R., & Paredes-Sabja, D. (2014). Characterization of the collagen-like exosporium protein, BclA1, of *Clostridium difficile* spores. Anaerobe, 25, 18–30. https://doi.org/10.1016/j.anaerobe.2013.11.003

Pizarro-Guajardo, M., Calderón-Romero, P., Castro-Córdova, P., Mora-Uribe, P., & Paredes-Sabja, D. (2016). Ultrastructural Variability of the Exosporium Layer of *Clostridium difficile* Spores. Applied and Environmental Microbiology, 82(7), 2202–2209. https://doi.org/10.1128/AEM.03410-15

Pizarro-Guajardo, M., Calderón-Romero, P., & Paredes-Sabja, D. (2016). Ultrastructure Variability of the Exosporium Layer of *Clostridium difficile* Spores from Sporulating Cultures and Biofilms. Applied and Environmental Microbiology, 82(19), 5892–5898. https://doi.org/10.1128/AEM.01463-16

Popham, D. L. (2002). Specialized peptidoglycan of the bacterial endospore: the inner wall of the lockbox. Cellular and Molecular Life Sciences, 59(3), 426-433. doi:10.1007/s00018-002-8435-5

Popham, D. L., Helin, J., Costello, C. E., & Setlow, P. (1996). Muramic lactam in peptidoglycan of *Bacillus subtilis* spores is required for spore outgrowth but not for spore dehydration or heat resistance. Proceedings of the National Academy of Sciences of the United States of America, 93(26), 15405-15410. doi:10.1073/pnas.93.26.15405

Popham, D. L., Sengupta, S., & Setlow, P. (1995). Heat, hydrogen peroxide, and UV resistance of *Bacillus subtilis* spores with increased core water content and with or without major DNA-binding proteins. Applied and Environmental Microbiology, 61(10), 3633

Popoff, M. R., & Boquet, P. (1988). *Clostridium spiroforme* toxin is a binary toxin which ADP-ribosylates cellular actin. Biochemical and Biophysical Reseatch Communications, 152(3), 1361-1368. doi:10.1016/s0006-291x(88)80435-2

Poutanen, S. M., & Simor, A. E. (2004). *Clostridium difficile*-associated diarrhea in adults. Canadian Medical Association Journal, 171(1), 51-58. doi:10.1503/cmaj.1031189

Poutanen, S. M., & Simor, A. E. (2004). *Clostridium difficile*-associated diarrhea in adults. Canadian Medical Association Journal, 171(1), 51-58

Poxton, I. R., McCoubrey, J., & Blair, G. (2001). The pathogenicity of *Clostridium difficile*. Clinical Microbiology and Infection, 7(8), 421-427. doi:10.1046/j.1198-743x.2001.00287.x

Public Health England (2013). Updated guidance on the management and treatment of *Clostridium difficile* infection. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/321891/Clostridium_difficile_management_and_treatment.pdf%20

Public Health England (2018). MRSA, MSSA, Gram-negative bacteraemia and CDI: quarterly report. URL: https://www.gov.uk/government/statistics/mrsa-mssa-and-e-coli-bacteraemia-and-c-difficile-infection-quarterly-epidemiological-commentary

Purdy, D., O'Keeffe, T. A., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., Ostrowski, A., & Minton, N. P. (2002). Conjugative transfer of Clostridial shuttle vectors from Escherichia coli to *Clostridium difficile* through circumvention of the restriction barrier. Molecular Microbiology, 46(2), 439-452. doi:10.1046/j.1365-2958.2002.03134.x

Ragkousi, K., & Setlow, P. (2004). Transglutaminase-mediated cross-linking of GerQ in the coats of *Bacillus subtilis* spores. Journal of Bacteriology, 186(17), 5567-5575. doi:10.1128/JB.186.17.5567-5575.2004

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. doi:10.1186/1471-2180-6-50

Ramírez-Guadiana, F. H., Barajas-Ornelas, R. d. C., Corona-Bautista, S. U., Setlow, P., & Pedraza-Reyes, M. (2016). The RecA-Dependent SOS Response Is Active and Required for Processing of DNA Damage during *Bacillus subtilis* Sporulation. PLOS ONE, 11(3), e0150348. doi:10.1371/journal.pone.0150348

Ramírez-Vargas, G., López-Ureña, D., Badilla, A., Orozco-Aguilar, J., Murillo, T., Rojas, P., Riedel, T., Overmann, J., González, G., Chaves-Olarte, E., Quesada-Gómez, C., & Rodríguez, C. (2018). Novel Clade C-I *Clostridium difficile* strains escape diagnostic tests, differ in pathogenicity potential and carry toxins on extrachromosomal elements. Scientific Reports, 8(1), 13951. doi:10.1038/s41598-018-32390-6

Rao, K., Erb-Downward, J. R., Walk, S. T., Micic, D., Falkowski, N., Santhosh, K., Mogle, J. A., Ring, C., Young, V. B., Huffnagle, G. B., & Aronoff, D. M. (2014). The systemic inflammatory response to *Clostridium difficile* infection. PLOS ONE, 9(3), e92578. doi:10.1371/journal.pone.0092578

Rechner PM, Agger WA, Mruz K, Cogbill TH. (2001). Clinical features of Clostridial bacteremia: a review from a rural area. Clinical Infectious Diseases 2001; 33:349–53

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. doi:10.3201/eid1309.061116

Redmond, C., Baillie, L. W. J., Hibbs, S., Moir, A. J. G., & Moir, A. (2004). Identification of proteins in the exosporium of *Bacillus anthracis*. Microbiology-SGM, 150, 355-363. doi:10.1099/mic.0.26681-0

Reysset, G. (1996). Genetics of 5-nitroimidazole resistance in Bacteroides species. Anaerobe, 2(2), 59-69. doi:10.1006/anae.1996.0008

Ribis, J. W., Ravichandran, P., Putnam, E. E., Pishdadian, K., & Shen, A. (2017). The Conserved Spore Coat Protein SpoVM Is Largely Dispensable in *Clostridium difficile* Spore Formation. mSphere, 2(5), e00315-17. https://doi.org/10.1128/mSphere.00315-17

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. doi:10.1194/jlr.R500013-JLR200

Riegler, M., Sedivy, R., Pothoulakis, C., Hamilton, G., Zacherl, J., Bischof, G., Cosentini, E., Feil, W., Schiessel, R., LaMont, J. T., & et al. (1995). *Clostridium difficile* toxin B is more potent than toxin A in damaging human colonic epithelium in vitro. Journal of Clinical Investigation, 95(5), 2004-2011. doi:10.1172/jci117885

Riggs, M. M., Sethi, A. K., Zabarsky, T. F., Eckstein, E. C., Jump, R. L., & Donskey, C. J. (2007). Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic *Clostridium difficile* strains among long-term care facility residents. Clinical Infectious Diseases, 45(8), 992-998. doi:10.1086/521854

Rodriguez-Palacios, A., & Lejeune, J. T. (2011). Moist-heat resistance, spore aging, and superdormancy in *Clostridium difficile*. Applied Environmental Microbiology, 77(9), 3085-3091. doi:10.1128/aem.01589-10

Rohlfing, A. E., Eckenroth, B. E., Forster, E. R., Kevorkian, Y., Donnelly, M. L., Benito de la Puebla, H., Doublié, S., & Shen, A. (2019). The CspC pseudoprotease regulates germination of

Clostridioides difficile spores in response to multiple environmental signals. PLoS Genetics, 15(7), e1008224. https://doi.org/10.1371/journal.pgen.1008224

Rupnik, M., Wilcox, M. H., & Gerding, D. N. (2009). *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nature Reviews Microbiology, 7(7), 526-536. doi:10.1038/nrmicro2164

Sambol, S. P., Tang, J. K., Merrigan, M. M., Johnson, S., & Gerding, D. N. (2001). Infection of Hamsters with Epidemiologically Important Strains of *Clostridium difficile*. Journal of Infectious Diseases, 183(12), 1760-1766. doi:10.1086/320736

Samuel, P., Holtzman, C. M., Meilman, E., & Sekowski, I. (1973). Effect of Neomycin and Other Antibiotics on Serum Cholesterol Levels and on 7{alpha}-Dehydroxylation of Bile Acids by the Fecal Bacterial Flora in Man. Circulation Research, 33(4), 393-402

Sanchez-Salas, J. L., Santiago-Lara, M. L., Setlow, B., Sussman, M. D., & Setlow, P. (1992). Properties of Bacillus megaterium and *Bacillus subtilis* mutants which lack the protease that degrades small, acid-soluble proteins during spore germination. Journal of Bacteriology, 174(3), 807. doi:10.1128/jb.174.3.807-814.1992

Saujet, L., Pereira, F. C., Serrano, M., Soutourina, O., Monot, M., Shelyakin, P. V., Gelfand, M. S., Dupuy, B., Henriques, A. O., & Martin-Verstraete, I. (2013). Genome-Wide Analysis of Cell Type-Specific Gene Transcription during Spore Formation in *Clostridium difficile*. Plos Genetics, 9(10). doi:10.1371/journal.pgen.1003756

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., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B., & Parkhill, J. (2006). The multidrug-resistant human pathogen

Clostridium difficile has a highly mobile, mosaic genome. Nature Genetics, 38(7), 779-786. doi:10.1038/ng1830

Serrano, M., Crawshaw, A. D., Dembek, M., Monteiro, J. M., Pereira, F. C., Pinho, M. G., Fairweather, N. F., Salgado, P. S., & Henriques, A. O. (2016). The SpoIIQ-SpoIIIAH complex of *Clostridium difficile* controls forespore engulfment and late stages of gene expression and spore morphogenesis. Molecular Microbiology, 100(1), 204-228. doi:10.1111/mmi.13311

Setlow, B., & Setlow, P. (1980). Measurements of the pH within dormant and germinated bacterial spores. Proceedings of the National Academy of Sciences of the United States of America, 77(5), 2474-2476. doi:10.1073/pnas.77.5.2474

Setlow, B., & Setlow, P. (1993). Binding of small, acid-soluble spore proteins to DNA plays a significant role in the resistance of *Bacillus-subtilis* spores to hydrogen-peroxide. Applied and Environmental Microbiology, 59(10), 3418-3423

Setlow, B., & Setlow, P. (1993). Dipicolinic acid greatly enhances production of spore photoproduct in bacterial-spores upon UV irradiation. Applied and Environmental Microbiology, 59(2), 640-643.

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. doi:10.1128/jb.182.24.6906-6912.2000

Setlow, B., Sun, D. X., & Setlow, P. (1992). Interaction between DNA and alpha/beta-type small, acid-soluble spore proteins - a new class of DNA-binding protein. Journal of Bacteriology, 174(7), 2312-2322

Setlow, P. (1988). Small, acid-soluble spore proteins of *Bacillus* species: structure, synthesis, genetics, function, and degradation. Annual Review of Microbiology, 42, 319-338. doi:10.1146/annurev.mi.42.100188.001535

Setlow, P. (1994). Mechanisms which contribute to the long-term survival of spores of *Bacillus* species. Society for Applied Bacteriology Symposium Series, 23, 49s-60s

Setlow, P. (1995). Mechanisms for the prevention of damage to DNA in spores of Bacillus species. Annual Review of Microbiology, 49(1), 29-54. doi:10.1146/annurev.mi.49.100195.000333

Setlow, P. (2001). Resistance of spores of Bacillus species to ultraviolet light. Environmental and Molecular Mutagenesis, 38(2-3), 97-104. doi:10.1002/em.1058

Setlow, P. (2003). Spore germination. Current Opinion in Microbiology, 6(6), 550-556. doi:10.1016/j.mib.2003.10.001

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. doi:10.1111/j.1365-2672.2005.02736.x

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. Trends in Microbiology, 15(4), 172-180. doi:10.1016/j.tim.2007.02.004

Setlow, P. (2014). Germination of Spores of Bacillus Species: What We Know and Do Not Know. Journal of Bacteriology, 196(7), 1297-1305. doi:10.1128/jb.01455-13

Setlow, P., & Waites, W. M. (1976). Identification of several unique, low-molecular-weight basic-proteins in dormant spores of *Clostridium-bifermentans* and their degradation during spore germination. Journal of Bacteriology, 127(2), 1015-1017

Shah, D., Dang, M. D., Hasbun, R., Koo, H. L., Jiang, Z. D., DuPont, H. L., & Garey, K. W. (2010). *Clostridium difficile* infection: update on emerging antibiotic treatment options and antibiotic resistance. Expert Review of Anti-Infective Therapy, 8(5), 555-564. doi:10.1586/eri.10.28

Shapiro, R. L., Hatheway, C., & Swerdlow, D. L. (1998). Botulism in the United States: a clinical and epidemiologic review. Annuals of Internal Medicine, 129(3), 221-228. doi:10.7326/0003-4819-129-3-199808010-00011

Shen, A., Fimlaid, K. A., & Pishdadian, K. (2016). Inducing and Quantifying *Clostridium difficile* Spore Formation. Methods in Molecular Biology, 1476, 129-142. doi:10.1007/978-1-4939-6361-4_10

Shen, A., Edwards, A. N., Sarker, M. R., & Paredes-Sabja, D. (2019). Sporulation and Germination in Clostridial Pathogens. Microbiology spectrum, 7(6), 10.1128/microbiolspec. GPP3-0017-2018. https://doi.org/10.1128/microbiolspec.GPP3-0017-2018

Shrestha, R., Lockless, S. W., & Sorg, J. A. (2017). A *Clostridium difficile* alanine racemase affects spore germination and accommodates serine as a substrate. Journal of Biological Chemistry, 292(25), 10735-10742. doi:10.1074/jbc.M117.791749

Shrestha, R., & Sorg, J. A. (2018). Hierarchical recognition of amino acid co-germinants during *Clostridioides difficile* spore germination. Anaerobe, 49, 41-47. doi:10.1016/j.anaerobe.2017.12.001

Shrestha, R., Cochran, A. M., & Sorg, J. A. (2019). The requirement for co-germinants during Clostridium difficile spore germination is influenced by mutations in yabG and cspA. PLoS Pathogens, 15(4), e1007681. https://doi.org/10.1371/journal.ppat.1007681

Siezen, R. J., & Leunissen, J. A. M. (1997). Subtilases: The superfamily of subtilisin-like serine proteases. Protein Science, 6(3), 501-523.

Soldi, S., Vasileiadis, S., Uggeri, F., Campanale, M., Morelli, L., Fogli, M. V., Calanni, F., Grimaldi, M., & Gasbarrini, A. (2015). Modulation of the gut microbiota composition by rifaximin in non-constipated irritable bowel syndrome patients: a molecular approach. Clinical and Experimental Gastroenterology, 8, 309-325. doi:10.2147/ceg.S89999

Solovyev, V. (2011). V. Solovyev, A Salamov (2011) Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed. R.W. Li), Nova Science Publishers, p.61-78. In (pp. 61-78)

Sonenshein, A. L., Losick, R., & Hoch, J. A. (2002). *Bacillus subtilis* and Its Closest Relatives: American Society of Microbiology

Sorg, J. A., & Sonenshein, A. L. (2008a). Bile Salts and Glycine as Cogerminants for *Clostridium difficile* Spores. J. Bacteriol., 190(7), 2505-2512. doi:10.1128/jb.01765-07

Sorg, J. A., & Sonenshein, A. L. (2009). Chenodeoxycholate Is an Inhibitor of *Clostridium difficile* Spore Germination. Journal of Bacteriology, 191(3), 1115-1117. doi:10.1128/jb.01260-08

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. doi:10.1128/jb.00610-10

Stabler, R., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., Lawley, T., Sebaihia, M., Quail, M., Rose, G., Gerding, D., Gibert, M., Popoff, M., Parkhill, J., Dougan, G., & Wren, B. (2009). Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biology, 10(9), R

Steichen, C., Chen, P., Kearney, J. F., & Turnbough, C. L., Jr. (2003). Identification of the immunodominant protein and other proteins of the Bacillus anthracis exosporium. Journal of Bacteriology, 185(6), 1903-1910. doi:10.1128/jb.185.6.1903-1910.2003

Stevens, D. L., Aldape, M. J., & Bryant, A. E. (2012). Life-threatening clostridial infections. Anaerobe, 18(2), 254-259. doi:10.1016/j.anaerobe.2011.11.001

Stewart, G. C. (2015). The Exosporium Layer of Bacterial Spores: A Connection to the Environment and the Infected Host. Microbiol Mol Biol Rev, 79(4), 437-457. doi:10.1128/mmbr.00050-15

Sun, X., He, X., Tzipori, S., Gerhard, R., & Feng, H. (2009). Essential role of the glucosyltransferase activity in *Clostridium difficile* toxin-induced secretion of TNF-alpha by macrophages. Microbial Pathogenesis, 46(6), 298-305. doi:10.1016/j.micpath.2009.03.002

Sunde, E. P., Setlow, P., Hederstedt, L., & Halle, B. (2009). The physical state of water in bacterial spores. Proceedings of the National Academy of Sciences, 106(46), 19334-19339. doi:10.1073/pnas.0908712106

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., McFarland, L. V., Mellow, M., & Zuckerbraun, B. S. (2013). Guidelines for diagnosis, treatment,

and prevention of *Clostridium difficile* infections. American Journal of Gastroenterology, 108(4), 478-498; quiz 499. doi:10.1038/ajg.2013.4

Swerdlow, B. M., Setlow, B., & Setlow, P. (1981). Levels of H+ and other monovalent cations in dormant and germinating spores of *Bacillus megaterium*. Journal of Bacteriology, 148(1), 20-29

Sylvestre, P., Couture-Tosi, E., & Mock, M. (2003). Polymorphism in the collagen-like region of the *Bacillus anthracis* BclA protein leads to variation in exosporium filament length. Journal of Bacteriology, 185(5), 1555-1563. doi:10.1128/jb.185.5.1555-1563.2003

Tasteyre, A., Barc, M.-C., Collignon, A., Boureau, H., & Karjalainen, T. (2001). Role of FliC and FliD Flagellar Proteins of *Clostridium difficile* in Adherence and Gut Colonization. Infection and Immunity, 69(12), 7937-7940. doi:10.1128/iai.69.12.7937-7940.2001

Tennen, R., Setlow, B., Davis, K. L., Loshon, C. A., & Setlow, P. (2000). Mechanisms of killing of spores of *Bacillus subtilis* by iodine, glutaraldehyde and nitrous acid. Journal of Applied Microbiology, 89(2), 330-338. doi:10.1046/j.1365-2672.2000.01114.x

Theys, J., Pennington, O., Dubois, L., Anlezark, G., Vaughan, T., Mengesha, A., Lambin, P. (2006). Repeated cycles of Clostridium-directed enzyme prodrug therapy result in sustained antitumour effects *in vivo*. British Journal of Cancer, 95(9), 1212–1219. doi:10.1038/sj.bjc.6603367

Todd, S. J., Moir, A. J. G., Johnson, M. J., & Moir, A. (2003). Genes of Bacillus cereus and *Bacillus anthracis* encoding proteins of the exosporium. Journal of Bacteriology, 185(11), 3373-3378. doi:10.1128/jb.185.11.3373-3378.2003

Touchette, M. H., Benito de la Puebla, H., Ravichandran, P., & Shen, A. (2019). SpoIVA-SipL Complex Formation Is Essential for *Clostridioides difficile* Spore Assembly. Journal of Bacteriology, 201(8), e00042-19. https://doi.org/10.1128/JB.00042-19

Traag, B. A., Pugliese, A., Setlow, B., Setlow, P., & Losick, R. (2013). A conserved ClpP-like protease involved in spore outgrowth in *Bacillus subtilis*. Molecular Microbiology, 90(1), 160-166. doi:10.1111/mmi.12355

Triadafilopoulos, G., Pothoulakis, C., O'Brien, M. J., & LaMont, J. T. (1987). Differential effects of *Clostridium difficile* toxins A and B on rabbit ileum. Gastroenterology, 93(2), 273-279. doi:10.1016/0016-5085(87)91014-6

Ufarté, L., Potocki-Veronese, G., & Laville, É. (2015). Discovery of new protein families and functions: new challenges in functional metagenomics for biotechnologies and microbial ecology. Frontiers in Microbiology, 6, 563-563. doi:10.3389/fmicb.2015.00563

Uhr, G. T., Dohnalová, L., & Thaiss, C. A. (2019). The Dimension of Time in Host-Microbiome Interactions. mSystems, 4(1), e00216-00218. doi:10.1128/mSystems.00216-18

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., Wilcox, M. H., & 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. doi:10.1128/JB.00882-09

van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., Visser, C. E., Kuijper, E. J., Bartelsman, J. F., Tijssen, J. G., Speelman, P., Dijkgraaf, M. G., & Keller, J. J. (2013). Duodenal infusion of donor feces for recurrent *Clostridium difficile*. New England Journal of Medicine, 368(5), 407-415. doi:10.1056/NEJMoa1205037

Venugopal, A. A., & Johnson, S. (2012). Current state of *Clostridium difficile* treatment options. Clinical Infectious Diseases, 55 Suppl 2, S71-76. doi:10.1093/cid/cis355

Venugopal, A. A., & Johnson, S. (2012). Fidaxomicin: a novel macrocyclic antibiotic approved for treatment of *Clostridium difficile* infection. Clinical Infectious Diseases, 54(4), 568-574. doi:10.1093/cid/cir830

Vepachedu, V. R., & Setlow, P. (2007). Role of SpoVA Proteins in Release of Dipicolinic Acid during Germination of *Bacillus subtilis* Spores Triggered by Dodecylamine or Lysozyme. Journal of Bacteriology, 189(5), 1565-1572. doi:10.1128/jb.01613-06

Vohra, P., & Poxton, I. R. (2011). Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. Microbiology, 157(5), 1343-1353. doi:10.1099/mic.0.046243-0

von Eichel-Streiber, C., & Sauerborn, M. (1990). *Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. Gene, 96(1), 107-113

von Eichel-Streiber, C., Sauerborn, M., & Kuramitsu, H. K. (1992). Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and Streptococcus mutans glucosyltransferases. Journal of Bacteriology, 174(20), 6707-6710. doi:10.1128/jb.174.20.6707-6710.1992

Voth, D. E., & Ballard, J. D. (2005). *Clostridium difficile* Toxins: Mechanism of Action and Role in Disease. Clinical Microbiology Reviews, 18(2), 247-263. doi:10.1128/cmr.18.2.247-263.2005

Vyas, J., Cox, J., Setlow, B., Coleman, W. H., & Setlow, P. (2011). Extremely Variable Conservation of gamma-Type Small, Acid-Soluble Proteins from Spores of Some Species in the Bacterial Order Bacillales. Journal of Bacteriology, 193(8), 1884-1892. doi:10.1128/jb.00018-11

Walters, B. A., Roberts, R., Stafford, R., & Seneviratne, E. (1983). Relapse of antibiotic associated colitis: endogenous persistence of *Clostridium difficile* during vancomycin therapy. Gut, 24(3), 206-212. doi:10.1136/gut.24.3.206

Wang, S. T., Setlow, B., Conlon, E. M., Lyon, J. L., Imamura, D., Sato, T., Setlow, P., Losick, R., & Eichenberger, P. (2006). The forespore line of gene expression in *Bacillus subtilis*. Journal of Molecular Biology, 358(1), 16-37. doi:10.1016/j.jmb.2006.01.059

Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., Frost, E., & 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. doi:10.1016/s0140-6736(05)67420-x

Westphal, A. J., Price, P. B., Leighton, T. J., & Wheeler, K. E. (2003). Kinetics of size changes of individual *Bacillus thuringiensis* spores in response to changes in relative humidity.

Proceedings of the National Academy of Sciences of the United States of America, 100(6), 3461-3466. doi:10.1073/pnas.232710999

Wetzel, D., & Fischer, R. J. (2015). Small acid-soluble spore proteins of *Clostridium acetobutylicum*are able to protect DNA in vitro and are specifically cleaved by germination protease GPR and spore protease YyaC. Microbiology, 161(11), 2098-2109. doi:10.1099/mic.0.000162

Wheeldon, L. J., Worthington, T., & Lambert, P. A. (2011). Histidine acts as a co-germinant with glycine and taurocholate for *Clostridium difficile* spores. Journal of Applied Microbiology, 110(4), 987-994. doi:10.1111/j.1365-2672.2011.04953.x

Wilcox, M. H., Cunniffe, J. G., Trundle, C., & Redpath, C. (1996). Financial burden of hospital-acquired *Clostridium difficile* infection. Journal of Hospital Infection, 34(1), 23-30. doi:http://dx.doi.org/10.1016/S0195-6701(96)90122-X

Wilcox, M. H., Fawley, W. N., Settle, C. D., & Davidson, A. (1998). Recurrence of symptoms in *Clostridium difficile* infection-relapse or reinfection? Journal of Hospital Infection, 38(2), 93-100

Wilcox, M. H., & Fawley, W. N. (2000). Hospital disinfectants and spore formation by *Clostridium difficile*. Lancet, 356(9238), 1324-1324. doi:10.1016/s0140-6736(00)02819-1

Wroblewski, D., Hannett, G. E., Bopp, D. J., Dumyati, G. K., Halse, T. A., Dumas, N. B., & Musser, K. A. (2009). Rapid molecular characterization of Clostridium difficile and assessment of populations of *C. difficile* in stool specimens. Journal of Clinical Microbiology, 47(7), 2142–2148. https://doi.org/10.1128/JCM.02498-08.

Yang, L., & Li, L. (2015). Spore photoproduct lyase: the known, the controversial, and the unknown. The Journal of Biological Chemistry, 290(7), 4003-4009. doi:10.1074/jbc.R114.573675

Yoon, S., & Brandt, L. J. (2008). W1281 Fecal Transplantation Via Colonoscopy for *C. difficile*-Associated Diarrhea; a Case Series of 6 Patients. Gastroenterology, 134(4), A-671. doi:10.1016/S0016-5085(08)63132-7

Youngster, I., Russell, G. H., Pindar, C., Ziv-Baran, T., Sauk, J., & Hohmann, E. L. (2014). Oral, capsulized, frozen fecal microbiota transplantation for relapsing *Clostridium difficile* infection. Journal of the American Medical Association, 312(17), 1772-1778. doi:10.1001/jama.2014.13875

Zhang, Y., & Klaassen, C. D. (2010). Effects of feeding bile acids and a bile acid sequestrant on hepatic bile acid composition in mice. Journal of Lipid Research, 51(11), 3230-3242. doi:10.1194/jlr.M007641

Zhu, D., Sorg, J. A., & Sun, X. (2018). *Clostridioides difficile* Biology: Sporulation, Germination, and Corresponding Therapies for *C. difficile* Infection. Frontiers in Cellular and Infection Microbiology, 8(29). doi:10.3389/fcimb.2018.00029