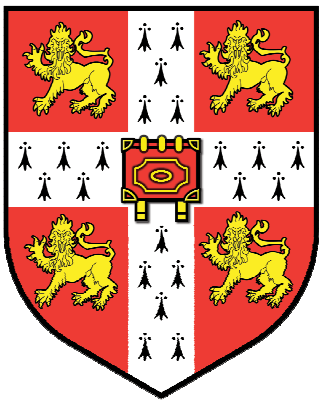
Defining the *Clostridium difficile spo0A* regulon and its role in disease and transmission

By

Laura Jane Pettit

2012



This dissertation is submitted for the degree of
Doctor of Philosophy

Darwin College
University of Cambridge

Abstract

Clostridium difficile is an intestinal dwelling bacterium that is a major cause of antibiotic-associated diarrhoea, representing a major healthcare-associated problem and significant economic burden. Unlike most other healthcare pathogens, *C. difficile* produces highly infective and resistant spores that are excreted by infected patients, producing an environmental transmission reservoir that confounds standard disinfection regimens. Thus, infection with *C. difficile* is now endemic in many hospitals.

The aim of this study was to assess the role of the *C. difficile spo0A* gene in disease and transmission, and to identify the genes under the control of Spo0A, that is, its regulon. Here, we use a murine model of infection to examine the role of the *C. difficile spo0A* gene during infection and transmission. We demonstrate that *C. difficile spo0A* mutant derivatives cause exaggerated virulence in mice, linked to an increase in the production of toxins A and B, but that these mutant derivatives are unable to persist within and effectively transmit between mice. Thus, the *C. difficile* Spo0A protein plays a key role in persistent infection, including recurrence and host-to-host transmission in this model. This data has potential clinical implications related to the management of hospital patients.

The *C. difficile spo0A* gene encodes for a highly conserved transcriptional regulator of sporulation. Here, we define the *C. difficile spo0A* regulon using transcriptomic and proteomic analyses. We validate Spo0A as a regulator of a number of sporulation genes and confirm that Spo0A negatively regulates toxin production. Spo0A also negatively regulates key components of the *C. difficile* flagellar assembly apparatus and modulates several metabolic pathways, including the fermentation of carbohydrates leading to the production of butyrate.

Thus, the *C. difficile spo0A* gene is a global transcriptional regulator that coordinates multiple virulence, sporulation and metabolic phenotypes during *C. difficile* disease and transmission.

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Gordon Dougan for giving me the opportunity to complete this Ph.D in his laboratory, as well as for his ongoing support and expert supervision. I also extend my gratitude to my mentor, Dr Trevor Lawley for his direction and guidance throughout this work. Additionally, I wish to acknowledge the Wellcome Trust, whose generous funding has underwritten this work.

I would also like to make a special mention of those who helped me during parts of my Ph.D, and without whom this work could not have been completed: Dr Simon Clare for his expert ideas, insights and knowledge of mouse infections, David Goulding for his incredible EM images, Dr Derek Pickard for his interesting approach to molecular microbiology, and Lars Barquist and Melissa Martin for teaching me the ways of ssRNA-Seq.

To all the other people who encouraged, supported and advised me during my Ph.D, I cannot list you all, but thank you for everything.

A massive thanks to everyone in Team 15 is also deserved, for their willingness to give their time, expertise and support. I look forward to seeing you all again!

Finally, I would like to thank my family for their unconditional and ongoing support. I dedicate this thesis to my parents, Andrew and Susan, for never doubting that I could do it, and to my wonderful husband Charles, for his endless reassurance, support and encouragement. I truly could not have done it without you.

Declaration

I hereby declare that this thesis is the result of my own work and includes no material written by any other person or material which is the outcome of work done in collaboration except where specifically indicated in here or in the Materials and methods section. I was fully involved in all aspects of the design and experimental work associated with this thesis.

Dr Simon Clare (Wellcome Trust Sanger Institute, Cambridge), assisted the author in performing tissue collection from experimental animals and performed mouse immunisations. Mr David Goulding (Wellcome Trust Sanger Institute, Cambridge) performed Transmission Electron Microscopy, in assistance with the author. Mass spectrometry was performed by the core proteomics facility (Wellcome Trust Sanger Institute, Cambridge) in assistance with the author. The analysis of the results was performed by the author. Transcript mapping was performed by the core pathogen informatics team as part of the ssRNA-Seq mapping pipeline (Wellcome Trust Sanger Institute, Cambridge). The analysis of the results was performed by the author.

C. difficile 630 Δ*spo0A* and R20291Δ*spo0A* mutants were generated by Dr Lisa Dawson (London School of Hygiene and Tropical Medicine, London) and Dr Robert Fagan (Imperial College, London), respectively. Butyrate quantification was performed in collaboration with Dr Sylvia Duncan (Rowett Institute of Nutrition and Health, Aberdeen). Spo0A purification was performed by Dr Wiep Klass Smits (Leiden University, Leiden).

None of the material presented herein has been submitted previously for the purpose of obtaining another degree. This thesis does not exceed 60,000 words, as required by the School of Biological Sciences.

Laura J. Pettit
October, 2012

Table of contents

Abstract.....i
Acknowledgements.....ii
Declaration.....iii
Table of contents.....iv
List of tables.....xii
List of figures.....xiii
List of abbreviations.....xvi

Chapter 1: Introduction

1.1 *Clostridium difficile*.....2
1.1.1 Discovery and natural history.....2
1.1.1.1 "Clindamycin colitis".....3
1.1.2 Genomics and phylogeny.....3
1.1.2.1 Strains of *C. difficile* and ribotype distribution.....4
1.1.2.2 Ribotype 027.....4
1.2 *C. difficile* pathogenicity.....6
1.2.1 Risk factors and susceptibility to *C. difficile* infection.....7
1.2.1.1 Antibiotics and the role of the host microbiota.....7
1.2.1.2 Age-related susceptibility.....9
1.2.1.3 Immunosupression and other risk factors.....10
1.2.1.4 Adaptive immunity and protective factors.....11
1.2.2 Acquisition and colonisation.....13
1.2.3 Toxins A and B and the expression of disease.....14
1.2.3.1 Genomic organisation and structure.....16
1.2.3.2 Mode of action.....19
1.2.3.3 Immune activation.....20
1.2.4 Spectrum of disease.....20
1.2.4.1 Asymptomatic carriage and newborn resistance to disease.....21
1.2.4.2 Pseudomembranous and fulminant colitis.....22

1.2.5 Diagnosis....23
1.2.6 Treatments....24
1.2.6.1 Non-antimicrobial therapies for CDAI....25
1.2.6.1.1 Faecal replacement therapy....25
1.2.6.1.2 Probiotics....26
1.2.6.2 Standard antibiotic therapy....27
1.2.7 Recurrent and relapsing disease: the antibiotic paradox....28
1.3 Epidemiology of *C. difficile*....29
1.3.1 Surveillance in the UK....29
1.3.2 Outbreaks and epidemics: the rise of the 027s in the UK....32
1.3.2.1 Hospital-acquired infection and transmission....33
1.3.2.2 Changing epidemiology: community-acquired *C. difficile*....35
1.3.2.3 Economic burden of *C. difficile* disease....36
1.3.2.4 Infection control....37
1.4 *C. difficile* life cycle....38
1.4.1 Biology of spores....38
1.4.1.1 Role of *C. difficile* spores....39
1.4.2 Structure of the spore....41
1.4.3 General process of spore formation, sporulation and germination....43
1.4.3.1 Stages of sporulation....43
1.4.3.2 Stages of germination....45
1.4.4 *Bacillus*: the paradigm spore former....46
1.5 Stage 0 sporulation protein A....48
1.5.1 Pathway to Spo0A activation....50
1.5.2 Structure and role of Spo0A as a transcription factor....50
1.5.2.1 Mutual regulation of gene expression by Spo0A and σ^H....51
1.5.2.2 The *Bacillus* Spo0A regulon....52
1.6 Genomics and genetic tools....53
1.6.1 Mutagenesis....53
1.6.1.1 ClosTron technology....53

1.6.2 Transcriptomics and proteomics....54
1.6.2.1 ssRNA-Seq....54
1.7 Aims of thesis....55

Chapter 2: Materials and methods

2.1 Materials....57
2.1.1 Bacteria and plasmids....57
2.1.2 Oligonucleotides....59
2.1.3 Mice....60
2.2 Methods....61
2.2.1 Bacterial culture....61
2.2.2 *In vivo* methods....61
2.2.2.1 Mouse infections....62
2.2.2.1.1 Oral gavage with *C. difficile*....62
2.2.2.1.2 Enumeration of viable *C. difficile* from mouse faeces....63
2.2.2.1.3 Competitive index infections....63
2.2.2.1.4 Transmission experiments....64
2.2.2.2 Antibody generation and serum extraction....67
2.2.3 Tissue methods....67
2.2.3.1 Paraffin embedding and sectioning of caecum tissue....67
2.2.3.1.1 Histology....68
2.2.4 Microscopy....68
2.2.4.1 Indirect immunofluorescence....68
2.2.4.2 Preservation for transmission electron microscopy....69
2.2.4.3 Flagellar negative staining....70
2.2.4.4 Tissue processing for ImmunoGold electron microscopy....70
2.2.5 Molecular methods....71
2.2.5.1 Polymerase chain reaction (PCR)....71
2.2.5.2 Cloning PCR amplicons....72
2.2.5.3 Plasmid DNA extraction....73
2.2.5.4 Restriction digestion....73

2.2.5.5 DNA ligation....74
2.2.5.6 Preparation of electrocompetent cells....74
2.2.5.7 Transformation via electroporation....74
2.2.5.8 Conjugation....75
2.2.5.9 ClosTron mutagenesis....75
2.2.5.9.1 Screening and verification of mutants....79
2.2.5.10 Genetic complementation of the *spo0A* mutation....80
2.2.5.11 TcdA and TcdB quantification....80
2.2.5.12 Butyrate quantification....82
2.2.6 RNA methods....82
2.2.6.1 RNA stabilisation....82
2.2.6.2 RNA extraction....83
2.2.6.3 DNA removal....83
2.2.6.4 Reverse transcription of isolated RNA....84
2.2.7 DNA sequencing....84
2.2.7.1 Library construction and Illumina HiSeq cDNA sequencing....84
2.2.7.2 Transcript mapping....85
2.2.7.3 ssRNA-Seq transcript normalisation....86
2.2.8 Protein methods....87
2.2.8.1 Spo0A purification and generation of anti-Spo0A antibodies....87
2.2.8.2 Protein preparation and cell lysis....88
2.2.8.3 One-dimensional gel electrophoresis of proteins....88
2.2.8.4 Western blotting....89
2.2.8.5 Coomassie staining....89
2.2.8.6 Protein digestion and peptide extraction....90
2.2.8.7 Isotopomeric dimethyl labeling....90
2.2.8.8 LC-MS/MS analysis....92
2.3 Statistical analysis....94
2.3.1 R....94

Chapter 3: The *Clostridium difficile spo0A* gene is a persistence and transmission factor

4.1 Publications arising from this chapter....96

4.2 Introduction....96

4.2.1 Human-virulent *C. difficile*....96

4.2.2 *C. difficile* transmission....97

4.2.3 Antibiotic treatment of *C. difficile* infection: disease recurrence....97

4.2.4 Animal models of *C. difficile* infection....98

4.2.4.1 Murine model of *C. difficile* disease....98

4.3 Aims of the work described in this chapter....99

4.4 Results....99

4.4.1 *C. difficile spo0A* is essential for spore formation....99

4.4.2 Antibiotic-induced model of *C. difficile* disease....102

4.4.3 Role of *spo0A* in *C. difficile* pathogenesis....103

4.4.3.1 *C. difficile spo0A* mutant derivatives cause acute disease in mice....104

4.4.3.2 *C. difficile* Spo0A mediates pathological responses in mice....106

4.4.3.2.1 Pathological response to *C. difficile* R20291 infection....106

4.4.3.2.2 Pathological response to *C. difficile* 630Δ*erm* infection....109

4.4.3.3 Association between Spo0A and toxin synthesis....110

4.4.3.3.1 *C. difficile* R20291 produces more TcdA and TcdB than 630Δ*erm* strains....111

4.4.3.3.2 *spo0A* is a negative regulator of toxin production in *C. difficile* R20291....111

4.4.3.3.3 Role of *C. difficile* 630Δ*erm* Spo0A in toxin production....113

4.4.3.3.4 Attempt to quantify toxin from faecal samples....114

4.4.4 Role of *spo0A* in persistence and recurrence....114

4.4.4.1 The *C. difficile spo0A* gene is required for persistent infection....115

4.4.4.2 *C. difficile spo0A* mediates disease recurrence....118

4.4.4.3 Disease recurrence is likely the result of relapse.....120
4.4.5 Role of *spo0A* in *C. difficile* transmission.....121
4.4.5.1 *C. difficile spo0A* is required for efficient host-to-host transmission.....122
4.5 Discussion.....125

Chapter 4: *Clostridium difficile spo0A* is a global regulator of virulence and transmission genes

5.1 Introduction.....129
5.1.1 Transcriptomics.....129
5.1.2 Proteomics.....130
5.1.3 An integrated approach to studying spore formation.....130
5.2 Aims of the work described in this chapter.....131
5.3 Results.....131
5.3.1 *C. difficile spo0A* is expressed during exponential growth.....131
5.3.2 *C. difficile spo0A* is a global regulator of gene expression.....133
5.3.2.1 Genome-wide identification of genes significantly regulated by Spo0A.....135
5.3.2.2 Confirmation of toxin production and sporulation regulated phenotypes by ssRNA-Seq.....141
5.3.2.2.1 Spo0A is an indirect negative regulator of TcdA production.....141
5.3.2.2.2 Spo0A is required for the expression of later stage sporulation genes.....142
5.3.2.2.3 Spo0A may be auto-regulated in *C. difficile* via a feedback loop.....144
5.3.2.3 Spo0A is a regulator of *C. difficile* metabolism.....146
5.3.2.3.1 Spo0A regulation of energy metabolism.....148
5.3.2.4 Spo0A regulates the expression of surface-associated transcripts.....149
5.3.2.4.1 Surface-associated transcripts positively regulated by Spo0A.....149

5.3.2.4.2 Surface-associated transcripts negatively regulated by Spo0A..... 150
5.3.3 Global relationship between the transcriptome and proteome of exponentially growing *C. difficile*.....150
5.3.3.1 Analysis of proteins from exponentially growing *C. difficile*.....151
5.3.3.1.1 SDS-PAGE.....151
5.3.3.1.2 Mass-spectrometry.....151
5.3.3.2 Correlation between transcriptome and proteome expression.....153
5.3.3.2.1 Proteins up-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....156
5.3.3.2.2 Proteins down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....157
5.3.3.2.3 Identification of "core" Spo0A-regulated genes.....158
5.3.4 Spo0A positively regulates butyrate production in *C. difficile* 630Δ*erm*…...161
5.3.4.1 Spo0A increases butyrate production in *C. difficile*.....164
5.3.5 Spo0A negatively regulates components of the *C. difficile* flagellar assembly apparatus.....166
5.3.5.1 Spo0A represses *C. difficile* flagellum assembly.....169
5.4 Discussion.....174

Chapter 5: Final discussion

6.1 A murine model of *C. difficile* infection: insights into the role of Spo0A in persistence and transmission.....180
6.2 Insights from an "omics" approach to studying the *C. difficile* Spo0A regulon.....182
6.3 Final conclusions and future directions……..... 183

References.....185

Appendices

Appendix 1: Publication.....211
Appendix 2: Transcripts significantly up-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....219
Appendix 3: Transcripts significantly down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....222

Appendix 4: Proteins significantly up-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....225
Appendix 5: Proteins significantly down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*.....227

List of tables

Chapter 2: Materials and methods

Table 2.1. Strains and plasmids used in this study.....57

Table 2.2. Oligonucleotides used in this study.....59

Table 2.3. Summary of transmission routes.....65

Table 2.4. Mass shifts and dimethyl labels resulting from isotope-labelled formaldehyde and cyanoborohydride.....91

Chapter 4: *Clostridium difficile spo0A* is a global regulator of virulence and transmission genes

Table 4.1. **"**Core" *C. difficile* 630Δ*erm* genes regulated by Spo0A.....159

Table 4.2. Summary of flagellar composition in *C. difficile* ribotypes 012 (630Δ*erm*) and 027 (R20291) 171

List of figures

Chapter 1: Introduction

Figure 1.1. Deep branching phylogenetic tree of human virulent *C. difficile*.....5
Figure 1.2. Key determinants in the acquisition and expression of CDAI.....8
Figure 1.3. *C. difficile* colonisation and pathogenesis.....15
Figure 1.4. Organisation of the *C. difficile* PaLoc region and toxin structure.....17
Figure 1.5. Psuedomembranous colitis due to *C. difficile* infection.....23
Figure 1.6. Number of death certificates citing *C. difficile*.....31
Figure 1.7. Primary reservoirs of healthcare-associated *C. difficile* infection.....34
Figure 1.8. Schematic representation of the *Clostridium difficile* life cycle.....40
Figure 1.9. Ultrastructure of the bacterial endospore.....42
Figure 1.10. Schematic overview of spore formation, sporulation and germination.....44
Figure 1.11. Canonical overview of clostridial spore formation.....47
Figure 1.12. Amino acid sequence alignment of Spo0A from common *C. difficile* ribotypes.....49

Chapter 2: Materials and methods

Figure 2.1. Schematic diagram demonstrating the experimental models used to investigate the role of the *C. difficile spo0A* gene in host transmission.....66
Figure 2.2. Schematic representation of ClosTron mutagenesis plamids.....77
Figure 2.3. Schematic representation of the ClosTron mutagenesis system.....78

Chapter 3: The *Clostridium difficile spo0A* gene is a persistence and transmission factor

Figure 3.1. Genetic and phenotypic characterisation of *C. difficile* 630Δ*erm* and R20291 *spo0A* mutants.....100
Figure 3.2. *C. difficile* R20291 can cause a long-term, persistent infection in mice.....103
Figure 3.3. *C. difficile* R20291 *spo0A* mutants demonstrate increased virulence in mice.....105

Figure 3.4. *C. difficile* R20291 *spo0A* mutants cause increased mucosal damage in mice.....107
Figure 3.5. *C. difficile* R20291 Spo0A mediates pathological responses in mice.....108
Figure 3.6. *C. difficile* 630Δ*erm*Δ*spo0A* mutants cause increased mucosal damage in mice.....110
Figure 3.7. *C. difficile spo0A* mutants produce elevated levels of TcdA and TcdB.....112
Figure 3.8. *C. difficile* Spo0A mediates intestinal persistence in mice.....116
Figure 3.9. *C. difficile spo0A* gene is required for persistent colonisation in mice.....117
Figure 3.10. *C. difficile spo0A* gene is required for relapsing disease in mice.....119
Figure 3.11. Host-to-host transmission of *C. difficile* is mediated by Spo0A.....123

Chapter 4: *Clostridium difficile spo0A* is a global regulator of virulence and transmission genes

Figure 4.1. Growth and sporulation properties of *C. difficile* 630Δ*erm* derivatives.....132
Figure 4.2. Western blot analysis of Spo0A expression in *C. difficile* 630Δ*erm* derivatives.....134
Figure 4.3. Identification of differentially expressed genes in *C. difficile* 630Δ*erm*Δ*spo0A* by transcriptional profiling.....136
Figure 4.4. Global overview of the *C. difficile* 630Δ*erm* Spo0A-regulated transcriptome.....137
Figure 4.5. Heat map illustrating the expression data of the top 40 differentially expressed genes.....139
Figure 4.6. Distribution and relative abundance of *C. difficile* 630Δ*erm* Spo0A-regulated transcripts, according to functional classification.....140
Figure 4.7. Linear representation of the *C. difficile* 630 genome with transcriptome reads mapping to the PaLoc region.....142
Figure 4.8. Linear representation of the *C. difficile* 630 genome with transcriptome reads mapping to the *spoIIIA* operon.....144
Figure 4.9. Spo0A may regulate *sigH* expression via a direct interaction.....145
Figure 4.10. SDS-PAGE of cellular proteins of *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*.....152

Figure 4.11. Total proteins identified by LC-MS/MS in *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*.....154
Figure 4.12. Venn diagram demonstrating the correlation between the Spo0A-regulated transcriptome and proteome of *C. difficile* 630Δ*erm*…....155
Figure 4.13. Linear representation the *C. difficile* 630 genome with transcriptome reads mapping to a region involved in butyrate production.....162
Figure 4.14. Metabolic pathway of butyrate biosynthesis.....163
Figure 4.15. Butyrate quantification formed by *C. difficile* 630Δ*erm* derivatives.....165
Figure 4.16. Schematic model of the bacterial flagellar assembly apparatus.....167
Figure 4.17. Schematic representation of the protein-protein interactions among components of the flagellar apparatus.....168
Figure 4.18. Spo0A is a negative regulator of flagellar synthesis.....170
Figure 4.19. Western Blot analysis of flagellar gene expression in *C. difficile* 630Δ*erm* derivatives.....173
Figure 4.20. Schematic representation of the *C. difficile*-specific sporulation cascade, indicating genes belonging to the Spo0A regulon at the transcriptional level.....175

Chapter 5: Final discussion

Figure 5.1. Summary of the phenotypes associated with the *C. difficile* Spo0A regulon.....181

List of abbreviations

BYA	Billion years ago
CDAI	*Clostridium difficile*-associated infection
cDNA	Complementary DNA
CDS	Coding sequence
CI	Competitive index
CFU	Colony forming units
Cy3	Cyanine-3
DAPI	4', 6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
DOH	Department of Health
DPA	Dipicolinic acid
ELISA	Enzyme-linked immunosorbent assay
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
FRT	Faecal replacement therapy
HPA	Health Protection Agency
IAA	Idoacetamide
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IGEM	ImmunoGold electron microscopy
IL	Interleukin
IBD	Inflammatory bowel disease
IPTG	Isopropyl-β-d-thiogalactopyranoside
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-MS/MS	Liquid chromatography with tandem mass spectrometry
LCT	Large clostridial toxin
LTCF	Long-term care facility
NAP1	North American Pulsotype 1
NHS	National Health Service

ONS	Office for National Statistics
PaLoc	Pathogenicity locus
PCR	Polymerase chain reaction
PMC	Pseudomembranous colitis
PMN	Polymorphonuclear neutrophil
PPI	Proton pump inhibitor
PTS	Phosphotransferase system
RAM	Retrotransposition-activated marker
REA	Restriction endonuclease analysis
RNA	Ribonucleic acid
SASP	Small acid-soluble protein
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SCFA	Short-chain fatty acid
SOE PCR	Splicing by overlapping extension polymerase chain reaction
ssRNA-Seq	Strand-specific cDNA sequencing
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TCEP	Tris(2-carboxyethyl)phosphine
TEAB	Triethylammonium bicarbonate
TEM	Tranmission electron microscopy
TLR	Toll-like receptor
TMB	3, 3', 5, 5'-tetramethylbenzidine
TNF-α	Tumour necrosis factor-alpha
WTSI	Wellcome Trust Sanger Institute