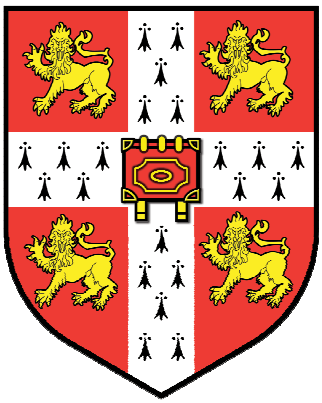
Defining the *Clostridium difficile spo0A* regulon and its role in disease and transmission

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

Chapter 2: Materials and methods

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

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

List of tables

List of figures

Chapter 1: Introduction

Chapter 2: Materials and methods

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

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

Chapter 5: Final discussion

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

1 Introduction

1.1 *Clostridium difficile*

Clostridium difficile is a Gram-positive, sporogenic anaerobe that has rapidly emerged in the past two decades from relative obscurity to become a dominant healthcare-associated pathogen (1-3). At present, *C. difficile* is arguably the leading cause of antibiotic-associated diarrhoea in developed countries, and is a significant cause of morbidity and mortality in hospitalised patients (4, 5). As such, *C. difficile* represents a major healthcare-associated problem and significant drain on resources.

In this chapter, I discuss the rapid emergence of *C. difficile* as an important healthcare-associated pathogen, as well as the clinical implications of *C. difficile* colonisation, transmission and infection. Novel approaches for the management of *C. difficile* are also reviewed.

1.1.1 Discovery and natural history

Originally identified by Hall and O'Toole in 1935 from the stools of healthy neonates (6), the bacterium was initially named *Bacillus difficillis*, "the difficult bacterium" due to the difficulties encountered in its isolation and culture, but it was subsequently reassigned to the genus *Clostridium*. However, given that the organism was considered a commensal of the gut microbiota rather than a threat to human health, its discovery generated little attention and few studies resulted from its initial discovery.

1.1.1.1 "Clindamycin colitis"

At that time and in the subsequent decades, *C. difficile*-associated infections (CDAI) were generally infrequent and only caused sporadic cases of diarrhoea in the hospitalised elderly. This picture of *C. difficile* as a nuisance to the healthcare system rather than a significant human pathogen has changed dramatically over recent years. Today, *C. difficile* is regarded as a primary healthcare-associated pathogen capable of causing major epidemics that are becoming increasingly frequent and severe (7-9). Infection with *C. difficile* is now endemic in many hospitals and has become an infection control emergency. These factors make *C. difficile* an important economic challenge as well as a medical one.

1.1.2 Genomics and phylogeny

C. difficile strain 630 (PCR ribotype 012, isolated in 1985) comprises a 4,290,252 bp genome with a low G+C content of 29.06%, and encodes a total of 3,776 coding sequences (CDSs) (10). The genome of *C. difficile* R20291 (PCR-ribotype 027, isolated in 2004-2005) shares a "core" of 3,247 CDS with *C. difficile* 630, including genes important for pathogenesis, spore formation and antimicrobial resistance (11). Interestingly, *C. difficile* has a strong coding bias, with > 80% of CDSs encoded from the forward strand (10).

1.1.2.2 Strains of *C. difficile* and ribotype distribution

C. difficile is a genetically diverse species with a highly dynamic and mosaic genome primed for genetic exchange (12-14). There are currently in excess of 280 different *C. difficile* PCR ribotypes, suggesting considerable genetic diversity (Dr T. Lawley, personal communication). A phylogenetic tree based on whole genome sequences which illustrates the relationship between key *C. difficile* ribotypes is shown in Figure 1.1. Whilst certain ribotypes can dominate at given points in time and space (such as during an outbreak), ribotype prevalence tends to fluctuate both temporally and geographically (13).

1.1.2.3 Ribotype 027

In North America from around 2000, CDAIs appeared to be more severe and frequent with an observed increased resistance to fluoroquinolones and recalcitrance to standard antibacterial therapies (8, 15, 16). This resulted in the identification of a new ribotype of *C. difficile*, commonly referred to as PCR ribotype 027, North American Pulsotype 1 (NAP1) or restriction endonuclease analysis (REA) type BI. These designations are based on different strain typing methods but refer to the same strain or clade of organisms harbouring this particular genomic signature, herein known as ribotype 027. The emergence of the 027s strongly correlates with fluoroquinolone use, which were at the time one of the most commonly prescribed antibiotic classes (17).

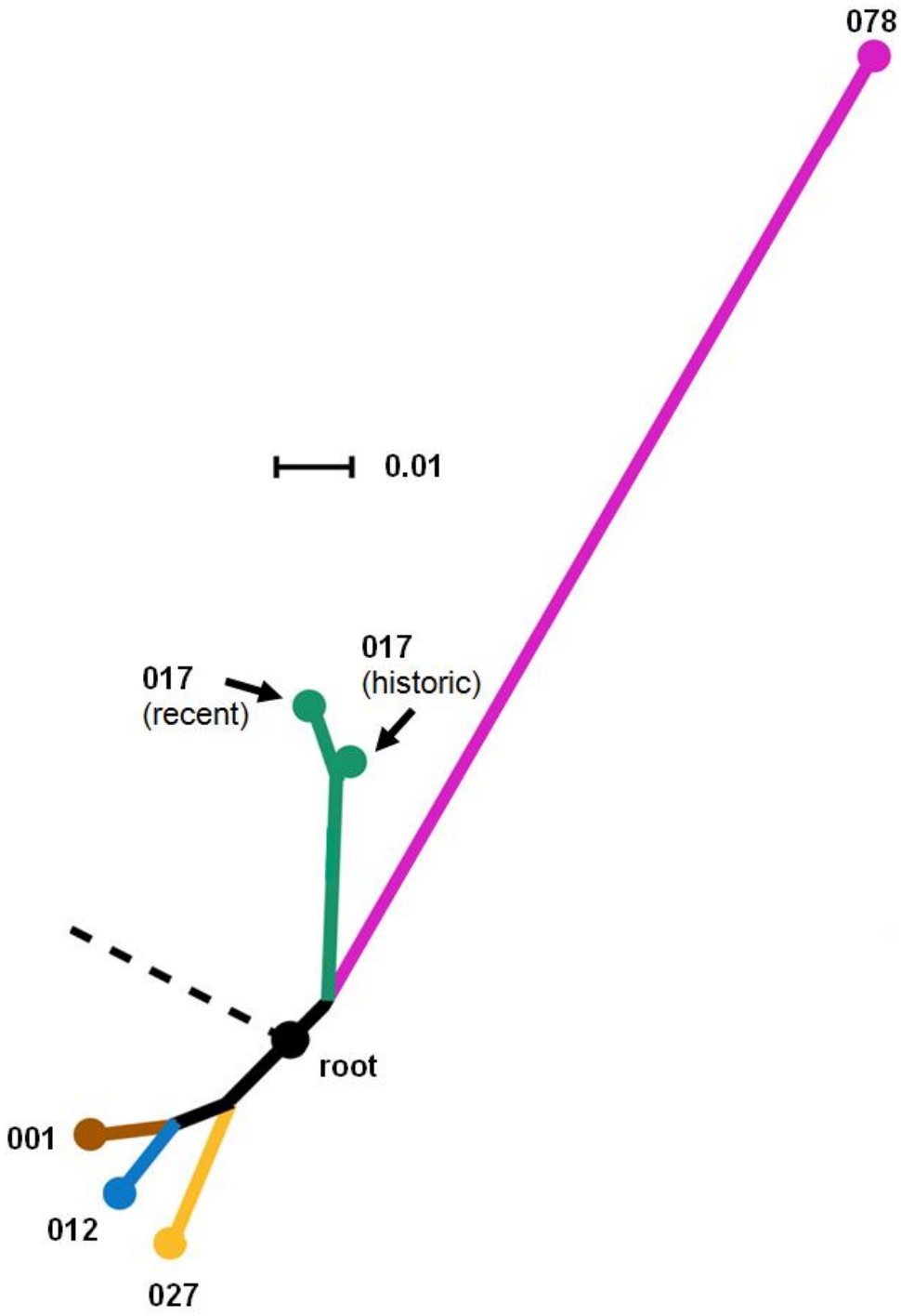


Figure 1.1. Deep branching phylogenetic tree of human virulent *C. difficile*. Phylogeny of *C. difficile* based on whole-genome sequences, describing the association between common human virulent *C. difficile* ribotypes. The root connects to *Clostridium bartlettii* and *Clostridium hiranonis*. Pathogenic isolates have been described in all of the ribotypes shown. Scale bar represents the number of substitutions per site. Adapted from: (13).

It is now thought that two distinct epidemic lineages of ribotype 027 originating in North America independently acquired identical mutations (at Thr82Ile) in the *gyrA* gene, which confers fluoroquinolone resistance (18). Recent work indicates that these two epidemic lineages were responsible for the rapid trans-continental and trans-Atlantic dissemination of ribotype 027 from North America to the healthcare systems in the UK (from 2003), continental Europe (from 2005) and Australia (from 2010), as well as South Korea and Switzerland (Miao He, in review). These 027 strains continue to disseminate globally (19, 20).

It has been postulated that the so-called "hypervirulence" associated with ribotype 027 strains is the result of increased toxin production, a property which is described in greater detail in section 1.2.3.1. The role of ribotype 027 in highly publicised UK hospital outbreaks and epidemics is discussed in section 1.3.2.

1.2 *C. difficile* pathogenicity

Most *C. difficile* infections are acquired nosocomially, and indeed the reported rates of *C. difficile* disease are greatest within the healthcare setting (21, 22). There are likely multiple reasons for this, such as the presence of a high density of immune compromised patients and heavy antibiotic usage. The healthcare system also provides an environment where high densities of highly infectious spores are encountered. Although *C. difficile* has the capacity to cause disease, the majority of colonised patients are likely to be asymptomatic (22), indicating

that exposure alone cannot precipitate disease. It is thus more likely that several factors dictate the risk of developing symptomatic *C. difficile* disease, including a combination of those in the host and the environment.

It has been postulated that a triad of events are required to incite symptomatic disease: (i) a heightened susceptibility to infection (typically by distorting the enteric microbiota via widespread antibiotic administration, or through immunosenescence/debilitation typical of an ageing population), (ii) exposure to and acquisition of a virulent strain of the pathogen, and (iii) successful colonisation and toxin production (23). The key determinants in the expression of disease are discussed in this section, and a general overview of the factors that influence it are given in Figure 1.2.

1.2.1 Risk factors and susceptibility to *C. difficile* infection

1.2.1.1 Antibiotics and the role of the host microbiota

The enteric microbiota, which comprises ~10^{12} bacteria/g faeces (24), is a major protective barrier against *C. difficile* colonisation. Perturbation of this protective commensal microbiota is perhaps the foremost predisposing factor for symptomatic *C. difficile* infection, and > 90% of symptomatic cases transpire during or shortly after antimicrobial therapy (21). Indeed, elderly patients with a history of antibiotic treatment within the previous three months represent the majority of CDAI cases (25). Following antimicrobial treatment, there is a loss

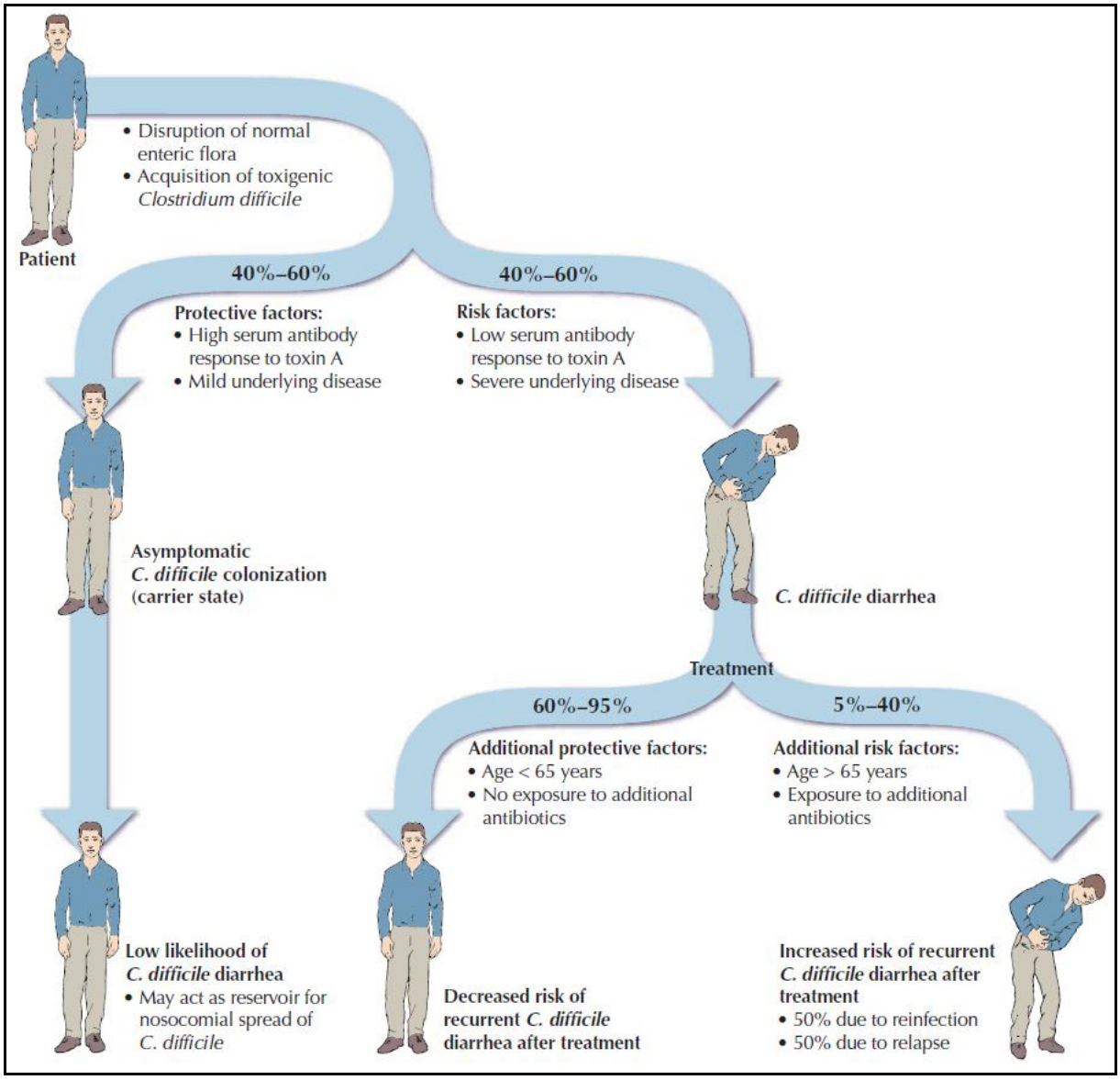


Figure 1.2. Key determinants in the acquisition and expression of CDAI. Flow diagram describing the events leading to *C. difficile* colonisation, and the development of systematic disease. A triad of events are typically associated with the development of *C. difficile* disease, and include (i) perturbation of the enteric microbiota, (ii) acquisition of a virulent strain, and (iii) toxin production. Host factors such as age, health status, and ability to mount an anamnestic IgG response to toxin A may also dictate the risk of developing symptomatic disease. Taken from: (23).

of “colonisation resistance” and reduced microbial competence, resulting in an ecological niche in which *C. difficile*- acquired either exogenously or endogenously- can proliferate (26, 27). The resulting infections are no longer considered just unpleasant complications of antibiotic usage, but rather as serious, life-threatening, economic and medical burdens.

Virtually all classes of antibiotics have the potential to trigger *C. difficile* disease, and as such the antibiotic era (circa 1970s) unsurprisingly brought with it an increase in the number of CDAIs (28). However, the propensity to facilitate disease is to a great extent dependent on the class of antibiotics used for treatment. The greatest risk is often attributed to antibiotics with a substantial anaerobic spectrum of activity (29). Historically, these have included broad-spectrum antibiotics such as penicillins, cephalosporins and most famously clindamycin, which become widely used during the 1970s. More recently however fluoroquinolones, which were initially considered low risk given that they have only a modest activity against the anaerobic component of the microbiota, have been associated with *C. difficile* outbreaks (11, 16, 30-32).

1.2.1.2 Age-related susceptibility

C. difficile is primarily considered a disease of the elderly (> 65 years old), with the incidence and severity of symptomatic CDAI increasing most notably in this population (21, 33, 34). However, elderly patients in long-term care facilities (LTCF) are often asymptomatic, indicating that advanced age, per se, may not be predispose disease (34). Rather, factors

typically associated with the elderly, such as immune senescence, reduced colonisation resistance and multiple underlying co-morbidities may underpin the apparent predilection for disease in this population.

1.2.1.3 Immunosuppression and other risk factors

Although antibiotics are the most notable source of microbiota perturbation, other medications or gastrointestinal procedures can precipitate *C. difficile* disease. These include, but are not limited to, surgical stress, antacids and stool softeners (35, 36). Proton pump inhibitors (PPIs), which act as gastric acid suppressors, have also been linked with increased CDAI. The rationale is that by reducing gastric acid secretion to treat conditions such as acid reflux, PPI therapy concomitantly reduces a major host defense mechanism, thus increasing the abilities of *C. difficile* to become established (37, 38). Although biologically plausible, indeed the ability of gastric acid to kill *C. difficile* has been demonstrated (39), the spores of the pathogen are resistant to gastric acid. This has created a paradox, and some argue that the link between PPI therapy and increased *C. difficile* infection is contentious (40).

More recently, cases of CDAI have increased in immunocompromised patients with severe underlying illnesses (41). For example, *C. difficile* is associated with increased mortality in nephrology patients with advanced chronic renal failure (42) and in patients undergoing bone marrow transplantation (43). Similarly, patients receiving antineoplastic chemotherapy are now considered at increased risk (44). There have also been increased reports of CDAI in

patients with idiopathic inflammatory bowel disease (IBD) (45, 46). Since CDAI can both mimic IBD and initiate an episode, it is critically important to rapidly identify whether *C. difficile* is an infectious complication of IBD.

It is important to note that these factors (whether age- or disease-related) share the common characteristic of perturbing the gut barrier or intrinsic microbiota composition. Intuitively, prolonged hospital stays also carry an increased disease risk, given that this may reflect increased exposure to a *C. difficile*-contaminated environment (47). However, infection with *C. difficile* often requires hospitalisation and increased duration of admission, perpetuating the situation (48).

Given that the factors discussed here are known to incite *C. difficile* disease, it is easy to see how the hospital environment represents a continuous source of infection to patients for whom *C. difficile* infection is already an increased risk.

1.2.1.4 Adaptive immunity and protective factors

Not everyone colonised by *C. difficile* will develop disease and it is likely that it is host, rather than bacterial, factors that predominantly drive this heterogeneity in clinical presentation. As discussed above, the host microbiota, gastric acid and health status are major defenses against *C. difficile* disease. However, the adaptive immune response of the host is thought to

influence both the severity and duration of the disease, representing a second line of defense (49).

The enterotoxins, designated toxins A and B, are potent virulence determinants of *C. difficile* (discussed later in Chapter 1.2.3), and there is increasing evidence that anti-toxin antibodies may protect against both disease and recurrence (50, 51). It is thought that such antibodies are present in much of the adult population, even in the absence of persistent colonisation (49, 52). In a prospective study, Kyne *et al*. (2000) demonstrated that *C. difficile*-colonised patients with high serum anti-toxin A immunoglobulin G (IgG) titers were significantly more likely to remain asymptomatic carriers, compared to patients with a low serum antibody response to toxin A, who were more likely to develop diarrhoea and symptomatic disease (53). Thus, failure to mount a sufficient anamnestic IgG response to toxin A may predispose *C. difficile* disease (53, 54).

It has also been suggested that patients of advanced age are both less likely to mount an immune response against toxin A (55, 56), and have serum that is less efficient at neutralising toxin activity (52), possibly accounting somewhat for the higher number of CDAI cases and greater mortality rates in this population. This observation led to the development of a potential therapy based on intravenous administration of toxin neutralising monoclonal antibodies, which is currently in clinical trials (trial number NCT00350298) (57).

1.2.2 Acquisition and colonisation

Exposure to *C. difficile* typically occurs within a healthcare setting, where it is exogenously acquired from person-to-person contact via the faecal-oral route, or on exposure to contaminated fomites or healthcare workers (see Chapter 1.3.2.1 for a more detailed review). Disease may also have an endogenous origin, whereby previously acquired *C. difficile* persists within an infected person, acting as an infection reservoir when conditions become favourable for the pathogen (58-60).

Unlike many other healthcare pathogens, *C. difficile* produces an infective and highly resistant spore form that is excreted by infected patients, producing an environmental transmission reservoir that confounds standard disinfection regimens (61). It is the ability of *C. difficile* to produce spores that enables the pathogen to persist in the environment for extended periods, and it is in part for this reason that infection with *C. difficile* is now endemic in many hospitals. It has also been postulated that the spore is important for the endogenous persistence of *C. difficile* with an infected patient (60). The biology and role of spores in *C. difficile* resistance and transmission is reviewed in greater detail in Chapters 1.3.2.1 and 1.4.1.

Infection is initiated when *C. difficile* is ingested by the host from the environment into the stomach. In this harsh and acidic environment, the majority of vegetative *C. difficile* cells die, as shown in the hamster experiments of Wilson *et al*. (1985) (62). Unlike the vegetative form, however, spores are inherently resistant to stomach acid and readily pass though the stomach and become established in the small bowel (23). Here, if conditions are favourable and

competition in the enteric microbiota is sufficiently reduced, the spores of *C. difficile* rapidly germinate, in response to bile derivatives (specifically taurocholic acids) (63). Once vegetative growth has resumed, *C. difficile* migrates to the colon- the most frequent site of significant infection (Figure 1.3).

The expression of a number of virulence-associated factors have tentatively been implicated in *C. difficile* pathogenesis, either directly or indirectly (64). These include (i) flagella and fimbrae which can mediate movement and putative adherence to the gut mucous layer, (ii) proteolytic and hydrolytic enzymes, which may help breach the gut mucous barrier, (iii) adhesins and surface proteins such as FliC (flagellin), FliD (flagellar cap protein) and cell wall proteins such as Cwp66 (65), which facilitate *C. difficile* binding to enterocytes, and (iv) an anti-phagocytic capsule to hinder opsonisation and engulfment by polymorphonuclear leukocytes (66). Once *C. difficile* has become established in the colon, the bacterium can produce two key virulence-associated factors, toxin A and toxin B. A schematic representation of *C. difficile* pathogenesis is shown in Figure 1.3.

1.2.3 Toxins A and B and the expression of disease

C. difficile is to a great extent a toxin-mediated disease, and non-toxigenic strains are considered to be significantly attenuated or even non-pathogenic (25, 67). Disease is to a significant degree attributable to the production of the two enterotoxins, named toxin A and

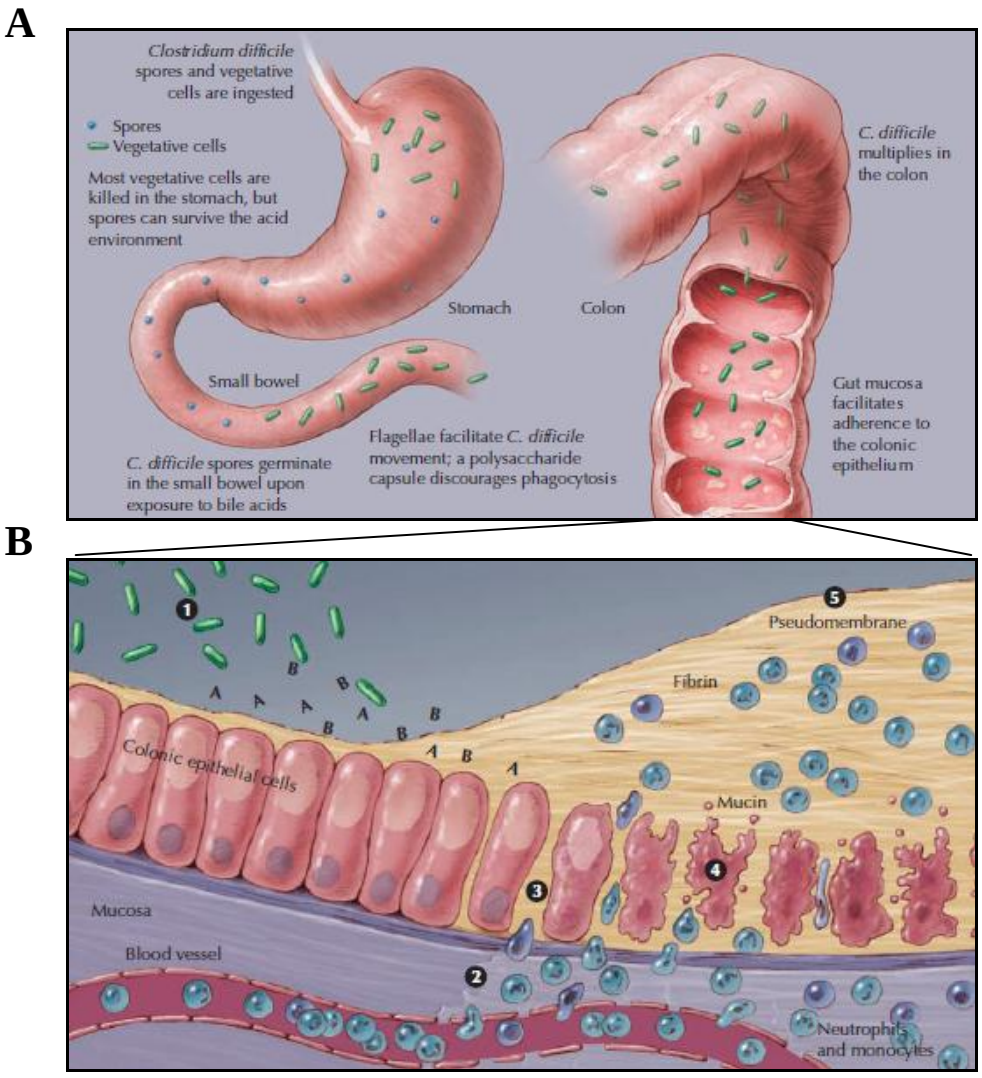


Figure 1.3. *C. difficile* colonisation and pathogenesis. A) Schematic representation of the gastrointestinal tract, describing the transit of *C. difficile* to the colon as the principle site of infection. *C. difficile* is ingested from the environment, and passes through the stomach where vegetative cells die. Spores, which are resistant to stomach acid, then pass through to the small bowel where they germinate in response to bile derivatives. This is followed by migration to the colon where (given a conducive environment) infection is established. This process is mediated by flagella and fimbrae, enzymic activity, adhesins and the bacterial capsule. **B)** Role of toxins A and B in the expression of disease. Once established in the colon, vegetative cells produce toxins A and B intraluminally which are internalised by colonocytes **(1)**. This induces TNF-α and interleukin release, resulting in neutrophil and monocyte extravasation, increased mucosal permeability **(2)**, loss of tight junction integrity **(3)** and cytoskeletal disregulation **(4)**. In severe cases, pseudomembranes comprising a purulent exudate of fibrin and immune cells can form on the colonic surface **(5)**. Adapted from: (23).

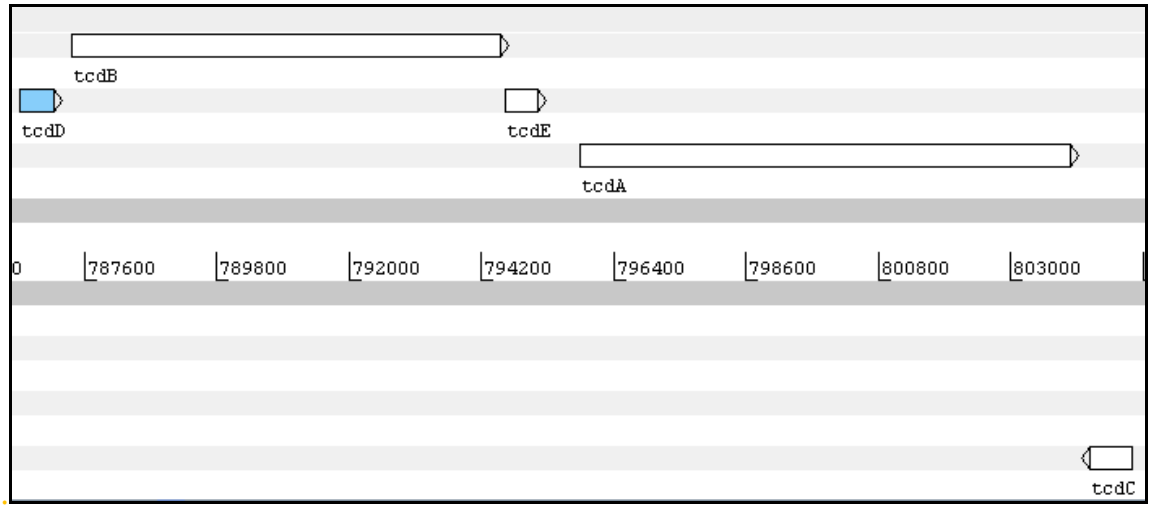
toxin B, which are regarded as the major virulence factors of the pathogen (68). There has been much debate surrounding the relative contribution of toxins A and B in recent years. It was initially thought that toxin A was the primary virulence determinant, and early experiments indicated that toxin A alone was sufficient to cause disease in hamsters, whereas toxin B only caused disease in the presence of previous mucosal damage induced by toxin A (69). It was thus thought that the toxins worked synergistically to bring about fulminant disease (69, 70). However, this view was challenged in 2009 when a report demonstrated that toxin B was essential for disease, and that this activity was not dependent on toxin A (71). This hypothesis was supported by the isolation of naturally occurring pathogenic A^-B^+ variant *C. difficile* strains (ribotype 017; See Figure 1.1).

The availability of directed mutagenesis techniques has challenged these paradoxical findings. Recent reports indicate that isogenic *C. difficile* mutants producing either toxin A or toxin B have pathogenic potential, though it is interesting to note that there have been no reports of clinically important A^+B^- isolates.

1.2.3.1 Genomic organisation and structure

Toxins A and B are high molecular weight proteins (308 and 269 kDa, respectively) belonging to the family of large clostridial toxins (LCTs) (72). As shown in Figure 1.4A, the toxin genes *tcdA* (toxin A) and *tcdB* (toxin B) are located within a pathogenicity locus (PaLoc), alongside three ancillary genes encoding positive (*tcdD*) and negative (*tcdC*) regulators, and a putative

A



B

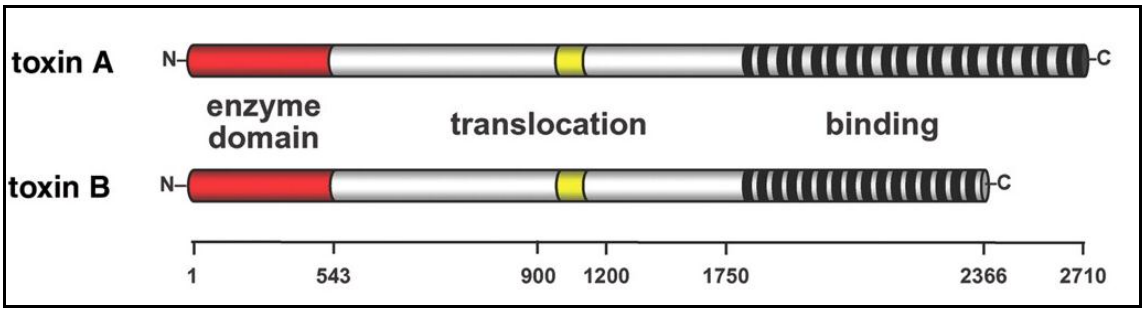


Figure 1.4. Organisation of the *C. difficile* PaLoc region and toxin structure. A) Artemis screenshot displaying the *C. difficile* PaLoc region from *C. difficile* 630, which harbours five genes (*tcdDBEAC*) responsible for the synthesis and regulation of toxins A and B. **B)** Structural organisation of *C. difficile* toxins A and B. Both toxins share a common tripartite structure, each comprising enzymic, translocation and binding domains. Adapted from: (72).

holin protein (*tcdE*) (72). At 19.6 kb, retaining the PaLoc region in the genome clearly represents a significant genetic investment by *C. difficile*.

Interestingly, ribotype 027 *C. difficile* harbour a single nucleotide deletion in the *tcdC* negative regulator at position 117, which introduces a frameshift and nonsense mutation and results in the production of a TcdC protein that is truncated from 232 to 65 amino acids (73). Some reports have postulated that this characteristic is one of the driving forces behind the emergence of 027 "hypervirulence", and results in increased production of toxins A and B, leading to increased disease severity, the need for surgical intervention and ultimately deaths (73). This idea, however, is not universally accepted (74, 75). For example, recent work has indicated that this deletion was present in the 027 lineage in 1985, prior to the emergence of the epidemic 027s (Miao He, in review). Additionally, *tcdC* genotyping has also identified a common 18 bp deletion in 027 variants (73, 76). However, these are largely regarded as silent and do not contribute to the hypertoxicity of such strains (77). Again, this genetic lesion was identified in both pre- and post-epidemic 027 isolates (Miao He, in review).

Toxins A and B are structurally somewhat similar and comprise a tripartite organisation with three main domains (72), as shown in Figure 1.4B. The carboxy-terminal comprises a series of repeating oligopeptides, which are thought to facilitate toxin - host cell receptor binding (78). Although specific human receptors for the toxins have not been fully characterised, animal studies have indicated that carbohydrates or glycoproteins are likely candidates (78-80). The large central domain harbours a small hydrophobic region (amino acids 956-1128), which is believed to function in membrane insertion and translocation of the toxins into the

host cell cytoplasm (72, 78). The amino-terminus (amino acids 1-543) possesses the full enzymatic (glucosyltransferase) activity, and it is only this catalytic domain that reaches the host cell cytoplasm (72, 81, 82).

1.2.3.2 Mode of action

Following colonisation, *C. difficile* releases toxins A and B intraluminally, which are then free to bind to receptors on the luminal wall of the colonocyte plasma membrane. Once internalised, toxin A has an enterotoxic and cytotoxic effect on the gastrointestinal tract, perturbing tight junction integrity and barrier function and increasing mucosal permeability (23, 67). This likely enables toxin B, a cytotoxin, to breach the epithelial barrier and enter the cell alongside toxin A via receptor-mediated endocytosis (83).

Once inside the cell, both toxins act as monoglucosyltransferases, catalysing the inactivation of small GTPases (such as Rho, Rac and Ras), which are regulators of host cell actin and myosin dynamics (84). Specifically, toxins A and B use cellular UDP-glucose as a co-substrate, and transfer the cleaved glucose group to threonine as the acceptor amino acid, forming an O-glycosydic bond and resulting in an inactive protein (81, 85, 86). Given the role of GTPases in actomyosin regulation, their inactivation results in cytoskeletal disorganisation, reduced trans-epithelial resistance, hemorrhage and the release of fluid in to the intestinal tract, ultimately producing the watery diarrhoea that is characteristic of *C. difficile* infection

(67). A schematic representation of the roles of toxin A and B in CDAI is shown in Figure 1.3B (23).

1.2.3.3 Immune activation

The pathophysiology caused by toxins A and B is compounded by the activity of the host immune system (49). Both toxins induce tumour necrosis factor-alpha (TNF-α) and interleukin (IL) production, particularly IL-1 and IL-6 from resident monocytes and macrophages in the lamina propria. As proinflammatory mediators this results in considerable neutrophil recruitment (83). In addition, the chemokine IL-8 is produced, resulting in a chemotactic gradient signaling prominent neutrophil extravasation to the site of inflammation (83, 87). Neutrophils play a central role in *C. difficile* pathophysiology, and neutrophil migration is a hallmark feature of *C. difficile* enterocolitis, as shown in Figure 1.3B (23).

1.2.4 Spectrum of disease

Historically considered to be a minor clinical problem, CDAI is now a common hospital-associated complication that has a wide-ranging disease spectrum resulting in a variety of clinical outcomes Interestingly, *C. difficile* can reside asymptomatically within the intestinal tract of up to 3% of the adult population and up to 70% of neonates (21). Although this sub-

population of infected people do not exhibit symptoms of disease, they do theoretically represent a clinically relevant reservoir of potential infection.

1.2.4.1 Asymptomatic carriage and newborn resistance to disease

C. difficile was essentially regarded as a non-pathogenic component of the intestinal microbiota for many years prior to the emergence of CDAI and in many hosts the organism still is. The intestinal carriage rates of *C. difficile* are highest in neonates, with up to 70% of newborns harbouring toxigenic strains of the pathogen (88). However, these cases are almost exclusively asymptomatic and clinical manifestations are rare.

It is not known why *C. difficile* is rarely pathogenic in many hosts even though it can produce potent enterotoxins. However, it has been suggested that although the immature gastrointestinal microbiota of newborns has an under-developed barrier effect and thus offers a microenvironment potentially conducive to *C. difficile* overgrowth, the immaturity of their colonocytes (which may lack toxin receptors) could account at least in part for the absence of disease in this demographic (89). During the process of ecological succession, the gastrointestinal microbiota matures and can exclude *C. difficile*. As such, the rates of asymptomatic *C. difficile* colonisation fall after the first year, at least in the non-healthcare community (88).

1.2.4.2 Pseudomembranous and fulminant colitis

C. difficile infections are predominantly gastrointestinal-associated, and extra-intestinal syndromes are infrequent. A common manifestation of *C. difficile* infection is mild "nuisance" diarrhoea. This is typically self-limiting and requires no treatment other than cessation of the inciting antibiotic (90). However, *C. difficile* infections can be severe and life-threatening. The description of the syndrome that led to the "re-discovery" of the pathogen in the late 1970's is now known as pseudomembranous colitis (PMC) (22). *C. difficile* is now well established as the aetiological agent of this syndrome, and is isolated from 95-100% of current PMC cases (21). PMC is defined clearly via histopathology and is characterised by the formation of pseudomembranes, evidenced as yellow-white raised plaques, which in extreme cases can coalesce over the whole colonic mucosa (91) (Figure 1.5). These oedematous lesions are filled with a purulent exudate containing leukocytes, neutrophils, fibrin, mucus and inflammatory debris (23).

Perhaps the most serious clinical expression of *C. difficile* infection is fulminant colitis, which occurs in around 1-3% of patients (23). This manifestation of CDAI is most common in the elderly and is associated with transmural inflammation leading to a range of symptoms including, but not limited to, constant abdominal pain, reduced bowel movements as a result of colonic dilation and paralytic ileus, toxic megacolon and colonic perforation (92, 93). Somewhat paradoxically, these more severe manifestations of disease often present without diarrhoea (93). Without prompt intervention, septicaemia, peritonitis and ultimately death can follow (93).

Figure 1.5. Psuedomembranous colitis due to *C. difficile* infection. Image of a resected colon, in which pseudomembranes are evidenced as raised yellowish exudative plaques over the colonic mucosa. A thickened bowel wall is also notable. Taken from: (91).

1.2.5 Diagnosis

The changing epidemiology and lack of traditional predisposing factors (Chapter 1.3.2.2) suggest that *C. difficile* should be considered as a causative agent in any patient with persistent diarrhoea. Indeed, healthcare workers should have a low threshold of suspicion in testing for *C. difficile* where the disease is endemic.

There are many laboratory tools that can be used to diagnose *C. difficile* infection, including testing for glutamate dehydrogenase via latex agglutination and stool culture (94, 95). These,

however are not widely used since they do not distinguish between toxigenic and non-toxigenic strains. The tissue culture cytotoxin assay for the detection of toxin B is largely considered in clinical circles as the 'gold standard', owing to its high sensitivity (94-100%) and superior specificity (99%) over other marketed tests (96, 97).

In reality, however ELISA assays for toxin A or toxin B are often used despite providing a lower sensitivity (70-90%), since they are less costly and technically demanding, and have a rapid turnaround time of only a few hours (97). Sigmoidoscopy and colonoscopy are highly sensitive and specific, and provide an immediate diagnosis in emergency cases. Since such procedures are expensive and carry a risk of bowel perforation, they are contraindicated in patients with fulminant colitis or toxic megacolon (96). Ultimately, refined forms of PCR- or sequence-based assays may find utility as general diagnostic assays.

1.2.6 Treatments

As already discussed, perturbation of the host protective microbiota consequent to antecedent antimicrobial exposure is the major predisposing factor for symptomatic *C. difficile* infection. As such, *C. difficile* can be treated rather conservatively in the first instance simply by removing the inciting antibiotic(s), and thus allowing the colonic microbiota to recover (98). In mild cases when used alongside supportive therapy (electrolyte and fluid replacement), this is often sufficient to resolve disease in otherwise healthy patients (59, 99). If it is medically necessary to continue antibiotic treatment for a primary infection, it is considered prudent to

use antibiotics with a narrower spectrum of activity, or those which do not have a strong correlation with CDAI (59).

1.2.6.1 Non-antimicrobial therapies for CDAI

Since Metchnikoff observed the potential health benefits of replacing "putrefactive" bacteria with beneficial lactic acid bacteria in 1907 (100), the concept of treating intestinal dysbiosis by modifying the hosts' microbiota has been widely studied. Faecal replacement therapy (FRT) and probiotic therapy are non-antimicrobial therapies that have specifically been explored as treatment options for CDAI.

1.2.6.1.1 Faecal replacement therapy

Faecal replacement therapy, also known as faecal bacteriotherapy, faecal microbiota therapy and intestinal microbiota transplantation, is not a new concept. The procedure was described as a treatment for human fulminant pseudomembraneous colitis in 1958 by Eiseman *et al*. (101) and its application in veterinary medicine is also well established (102). FRT refers to the delivery of healthy donor faeces (as a processed liquid suspension) into a patient via a faecal enema or nasogastral tube (103). Donor faeces are often from a close relative and are screened for pathogens prior to infusion (104).

By restoring the colonic microbiota, FRT aims to supplant *C. difficile* colonisation, restore colonic homeostasis and ultimately abrogate disease in the patient (103, 104). Although clinical trials are still incomplete, results suggest that FRT can be remarkably effective, demonstrating a success rate of approximately 90% in certain situations (103, 105). However, due to the nature of the therapy, many patients are inherently anxious about the procedure. It is thus likely that in the future such therapy will focus on the delivery of selected and refined bacterial strains, which resemble the composition of healthy human faeces but lack the aesthetic issues related to the use of crude, homogenised stools.

1.2.6.1.2 Probiotics

There have been multiple revised definitions of probiotics, however a current consensus is that probiotics are viable microorganisms which may exert a positive health benefit on the host when administered in adequate numbers (106). Such agents, including the bacteria *Lactobacillus*, *Bifidobacterium* and *Enterococcus* and the yeast *Saccharomyces*, have all been proposed for both the prophylaxis and treatment of CDAI (107). In the same way as FRT, the theoretical basis is that probiotics aim to restore the gastrointestinal equilibrium and thus limit or prevent *C. difficile* colonisation or overgrowth. However, in contrast to the perceived unpleasantness of FRT, probiotics are considered more aesthetically acceptable.

Although there is no definitive explanation of their mode of action, many mechanisms have been suggested, including direct competition for epithelial receptors, possibly via steric

hindrance (108), endogenous production of antimicrobials (such as bacteriocins and defensins), the stimulation of enhanced mucin production to obstruct mucosal adherence, and immunomodulation via immunoglobulin A (IgA) stimulation (107, 109, 110). It has also been suggested that *Saccharomyces boulardii* protects against CDAI by enzymatically degrading both the toxins and their enterocyte receptors (111-113).

Whilst alternatives to antibiotics are appealing, much of the evidence relating to the use of probiotics for the prevention and treatment of CDAI is suggestive but somewhat inconclusive (114, 115). Similarly, FRT appears to hold great promise for the treatment of CDAI but there is currently a lack of convincing clinical trial data. Thus, such therapies require both a robust understanding of the mechanism of action and large-scale clinical trials before definitive conclusions about their effectiveness and use in clinical practice is recommended. Nonetheless, preliminary observations suggest this as a direction of future research.

1.2.6.2 Standard antibiotic therapy

The current standard therapy for severe systemic *C. difficile* disease in many hospitals is a 7-10 day course of metronidazole or vancomycin. Oral delivery is the preferred method since bactericidal concentrations of the drug in the colon are more achievable via this route, although intravenous metronidazole therapy is an option (59, 116). Metronidazole is typically favoured due to its lower cost and the need to reduce the incidence of vancomycin-resistant enteropathogens, though both drugs are considered to be equally efficacious in treating mild

disease (117). Vancomycin is usually reserved for severe disease and for patients that have not responded to prior metronidazole therapy (117).

Whilst both vancomycin and metronidazole are highly effective at resolving symptoms during therapy, both are associated with a high frequency of recurrence once the treatment is removed (117, 118). This creates an unusual paradox in which the treatment for the disease can also precipitate it. The fundamental message for healthcare providers, however is to balance the risks and benefits when prescribing antibiotics, and indeed any therapy, for the treatment of CDAI.

1.2.7 Recurrent and relapsing disease: the antibiotic paradox

Recurrent infection after cessation of antibiotic therapy is a hallmark feature of *C. difficile* persistence, and recurring episodes of diarrheoa are rapidly becoming the norm rather than the exception in elderly patients. Recurrence can be defined by the complete subsidence of symptomatic disease followed by subsequent reappearance. The rates of recurrence are reported at 15-35% for the first relapse, and up to 65% for a second relapse (59, 60). Some patients experience chronic recurring enterocolitis, involving multiple diarrhoeic episodes over several years (119). It has been suggested that a reduced serum IgA response may underpin a patients' propensity to exhibit recurrent disease (120, 121).

Recurrent disease can be the result of either (i) relapse due to the endogenous persistence of the same strain that caused the initial infection within the host, or (ii) re-infection with a the same or a different strain of the pathogen from an exogenous source (122, 123). These are thought to occur at approximately similar rates, though the situation may vary from study-to-study and hospital-to-hospital (23, 60). Differentiating between relapse and re-infection can be challenging, expensive and time-consuming in the hospital environment, however the distinction is clearly pivotal if we are to understand the transmission dynamics of *C. difficile* and implement effective infection control strategies.

Tapered (decreasing antibiotic doses over an extended period) or pulsed (intermittent antibiotic delivery) dosing regimens are purported to reduce the incidence of *C. difficile* recurrence by gradually killing *C. difficile* as spores germinate and such approaches may also facilitate the recovery of the host microbiota (59). Ultimately, however, treating CDAI requires healthcare workers to be certain of the need to prescribe antimicrobials and to practice stringent and informed antimicrobial stewardship.

1.3 Epidemiology of *C. difficile*

1.3.1 Surveillance in the UK

In 1990, voluntary reporting of *C. difficile* to the Public Health Laboratory Service (and later to the Health Protection Agency; HPA) was introduced. At this time, there was a notable

increase in the number of reported cases of CDAI in England, from 1,172 cases in 1990, to 40,414 in 2004 (124). Due to concerns regarding increasing *C. difficile* cases and the apparent need to monitor episodes of CDAIs, surveillance of *C. difficile* in patients aged 65 years and over became mandatory for all acute National Health Service (NHS) Trusts in England from 2004. At this time, the Healthcare Commission was created to independently assess NHS performance. In response to the increased incidence and severity of *C. difficile* disease and major outbreaks (section 1.3.2), the Government's mandatory reporting scheme was extended to include patients 2 years and over from 2007. The introduction of the scheme was followed by a peak in the number of death certificates citing *C. difficile* (either as the underlying cause or mentioned; Figure 1.6) (125) and a total of 55,498 cases of *C. difficile* were reported by the HPA in 2007-2008 (126).

Intuitively, increased awareness and mandatory reporting brought with it an increase in the number of cases of *C. difficile*. However, it also led to the introduction of novel guidelines related to the management of patients and stricter infection control policies (section 1.3.3). As such, the incidence data reported by the HPA has steadily decreased in recent years, with an estimated 21,695 cases between 2010 and 2011, representing a 39% reduction from 2007-2008 (126). Similarly, there has been a notable reduction in the number of death certificates citing *C. difficile* (Figure 1.6).

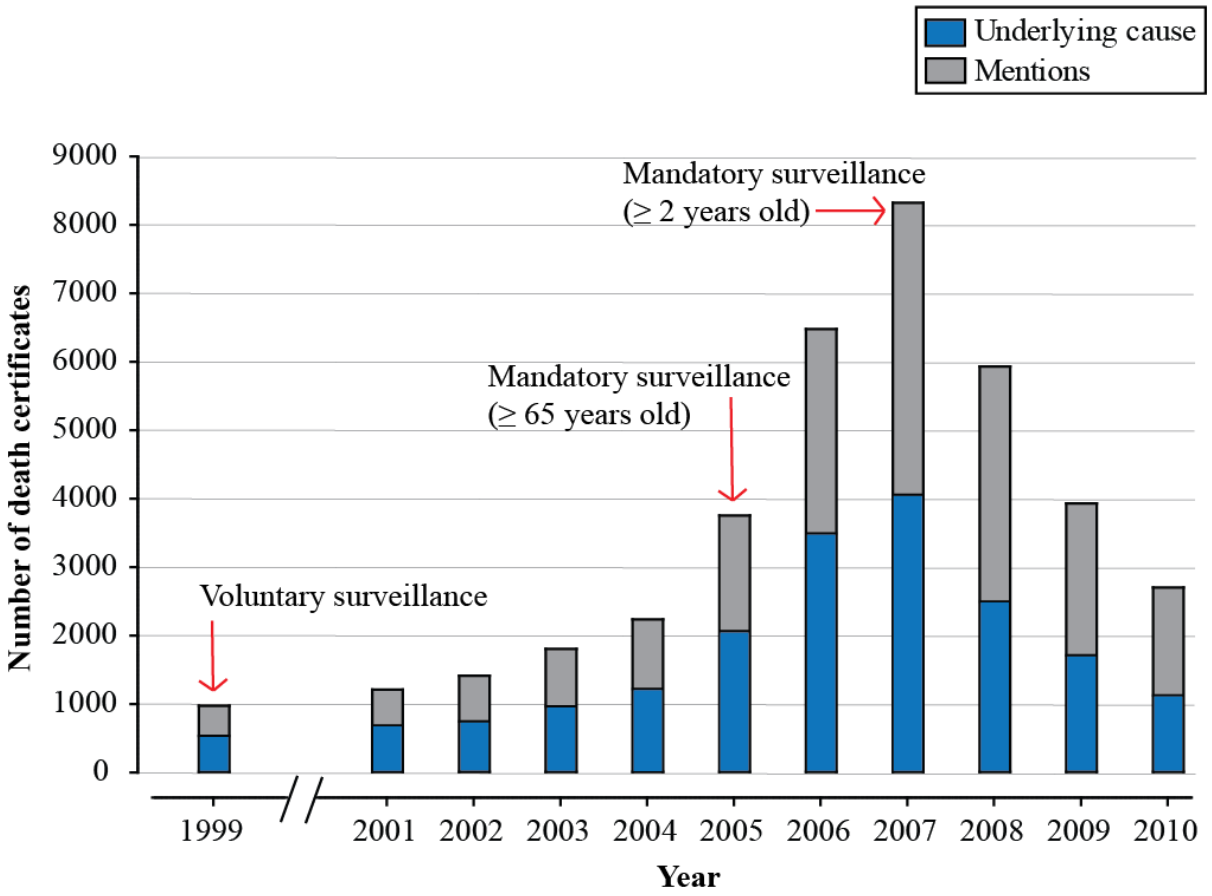


Figure 1.6. Number of death certificates citing *C. difficile*. Graph detailing the number of death certificates which either mention *C. difficile*, or cite it as the underlying cause of death. Data were collected from England and Wales, between 1999 and 2010 (data was not available for 2000). Also highlighted are the years at which key reporting scheme were introduced. The graph was generated from data published by the Office for National Statistics (ONS), (125).

1.3.2 Outbreaks and epidemics: the rise of the 027s in the UK

The global incidence of CDAI has risen considerably in the past decade, and public concern and intense political pressure regarding CDAI has been fuelled in light of recent and high-profile hospital outbreaks. The significant and highly publicised operational failings at Stoke Mandeville Hospital and Maidstone and Tunbridge Wells NHS Trust received considerable notoriety (127, 128). Following two outbreaks of *C. difficile* at Stoke Mandeville Hospital between 2003 and 2005, 334 new cases of *C. difficile* were identified which resulted in over 30 deaths (127).

In a strikingly similar case at the Maidstone and Tunbridge Wells NHS Trust, two outbreaks between 2004 and 2006 brought about over 500 new cases of CDAI and 90 deaths which were "probably or definitely" the result of *C. difficile* (128). According to the Healthcare Commission, ineffective surveillance, unnecessary antibiotic prescribing and failure to respond quickly to the index outbreak were the primary contributing factors (127, 128). Since this initial incursion, 027 strains have been reported (and remain present) in almost all NHS Trusts in the UK.

The increase in the rate and severity of CDAI has been attributed to the emergence of distinct genetic clades of *C. difficile*. Amongst the most notable of these clades is the so-called "hypervirulent" variant of *C. difficile*, commonly genotyped as PCR ribotype 027 (discussed in Chapter 1.1.2.2). *C. difficile* 027 is associated with higher rates of mortality, relapse, fluoroquinolone resistance as well as severe hospital outbreaks (8, 9, 11). In the UK, over

55% of *C. difficile* isolated from hospitals in 2007/2008 were 027 strains, though this number has progressively decreased, perhaps in part due to improved surveillance and reporting policies. In 2010/2011, 027 strains accounted for 12% UK of isolates (129). Although this represents a -43% reduction since 2007, *C. difficile* 027 remains the most commonly detected ribotype (129).

1.3.2.1 Hospital-acquired infection and transmission

The primary reservoir of *C. difficile* is believed to be the hospital. Within this setting several sources, including medical personnel, virtually all environmental surfaces, fomites and other infected patients may be contributing to the spread of the disease (Figure 1.7). Floors and bedrails are often associated with the heaviest *C. difficile* contamination (130). Intuitively, the risk of *C. difficile* acquisition increases proportionally with the duration of stay (131), which is a compelling reason to limit the length of hospital admissions to the absolute minimum required. However, given that CDAI often requires hospitalisation, it is easy to see how the cycle of infection self-perpetuates.

C. difficile is thought to be transmitted primarily via the faecal-oral route (by both symptomatic and asymptomatic patients), following contamination of the patients' local environment or from the hands of healthcare workers (132). Owing to the nature of *C. difficile* disease, diarrhoea can be explosive and unexpected, increasing *C. difficile* shedding and necessitating the use of commodes resulting in greater environmental contamination. It is

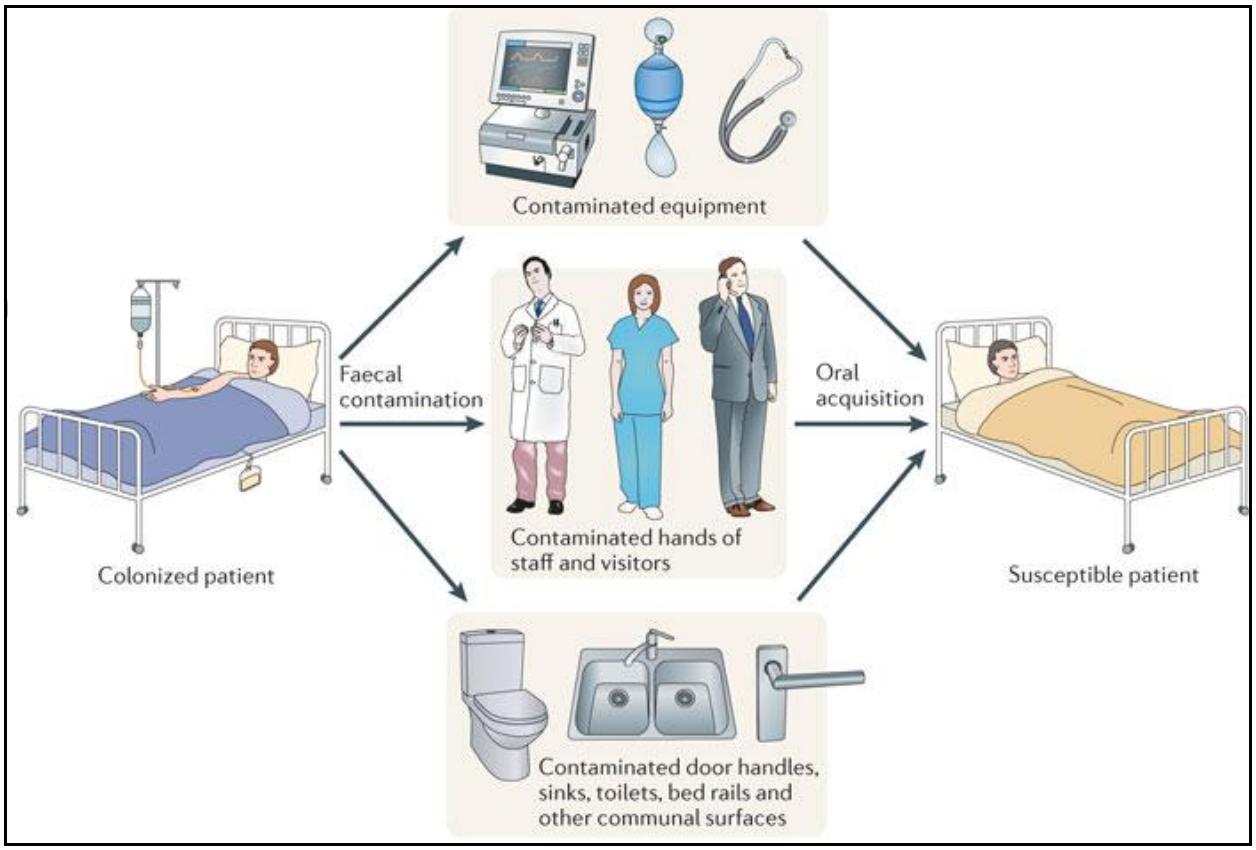


Figure 1.7. Primary reservoirs of healthcare-associated *C. difficile* infection. *C. difficile* is widely disseminated within the healthcare environment, and is able to survive for extended periods on medical personnel, patients, visitors, fomites, and almost all environmental surfaces such as bedrails, floors and door handles. The pathogen is primarily transmitted faecal-orally, and owing to the nature of the disease diarrhoea is often explosive, increasing environmental spore contamination. Aggressive surveillance activity and rigorous infection control strategies are the most prudent ways of breaking this chain of transmission. Adapted from: (133).

often difficult to determine whether environmental contamination of *C. difficile* is the consequence or cause of diarrhoeal shedding. In this sense, the physical proximity of infected and susceptible patients may be a key risk factor for the horizontal transmission of *C. difficile*. For example, outbreaks often occur in spatial clusters (134, 135). Similarly, an observational study by McFarland (2002) noted that within one week, a commode contaminated by an index patient resulted in eight further cases of *C. difficile* disease (136). Against this backdrop of endemic and epidemic CDAI, there is clearly a public health imperative to understand the transmission dynamics of *C. difficile*.

1.3.2.2 Changing epidemiology: community-acquired *C. difficile*

The new era of CDAIs has seen a shift in epidemiology. Undoubtedly, the majority of *C. difficile* cases are of a nosocomial and iatrogenic origin, and are most prevalent in the hospitalised elderly with a recent history of antibiotic therapy, as discussed. However, this association is not exclusive and the incidence of CDAI is considered by some to be increasing in populations previously considered to be low risk (137). These include community-dwelling individuals with no recent history of hospital contact or antimicrobial exposure. Additionally, there have been reports of severe community-acquired *C. difficile* disease in otherwise healthy children and peripartum women (138).

Many of these community-associated cases have been attributed to *C. difficile* ribotype 078, which like 027 has been linked with hypervirulence (139). Ribotype 078 strains are now also

considered to be an emerging threat in UK hospitals however the increasing incidence of CDAI due to ribotype 078 is most notable in The Netherlands (9, 140). Interestingly, a study by Goorhuis *et al*. (2008) conducted in The Netherlands demonstrated that 078 is also the dominating ribotype isolated from food-producing animals such as pigs and calves (139, 141), indicating that isolates from pigs and Dutch patients are highly related to the point of genetic clonality (139). This suggests that for ribotype 078 there may be a limited interspecies barrier and thus zoonotic potential.

1.3.2.3 Economic burden of *C. difficile* disease

C. difficile is a disease of high economic significance as well medical importance, representing a considerable drain on finite healthcare resources. However, data regarding the true economic cost of CDAI is limited. In 1996, a prospective case-controlled study estimated that CDAIs in the UK resulted in an increased management cost of £4107 per case per patient (142). More than 94% of this increased cost was the result of increased duration of stay, estimated at 21 days in a side room (142). Based on these estimates, *C. difficile* cost the UK £229 million in 2007 (at the peak of CDAI reports, see Figure 1.6) and £113 million in 2009 (143). These figures however, are likely to underestimate the full effect of CDAI, since the data only includes components for which precise costings were possible (such as the cost of medication, laboratory tests, length of stay). Ascertaining the costs associated with physiotherapy, increased laundry/cleaning, use of specialist isolation rooms, or indeed costs

incurred from disrupted service, ward closures and the unavailability of such amenities to other patients is more challenging to quantify (142, 144).

1.3.2.4 Infection control

There is a wealth of data regarding both the aetiology of *C. difficile* and its significance as a nosocomial pathogen. However, there remain many challenges and questions relating to the infection control practices required to reduce the incidence of CDAIs. The pathogen is now endemic in many hospitals and its management has become an infection control emergency. Clearly, infection control guidelines based on scientific acumen are critical if we are to reduce CDAIs, as well as the associated burden on healthcare providers.

The literature indicates that no single approach has been uniformly successful in perturbing the spread of *C. difficile* in hospitals. Rather, a multidisciplinary approach to infection control has been adopted, which takes into account all of the principle reservoirs of infection: the patient, their environment, and medical personnel and equipment. The Department of Health (DOH) recommends that patients with potentially infectious diarrhoea are immediately transferred to an isolation room (with self-contained toilet), or to a dedicated *C. difficile* unit where cohort nursing can be applied (145). In addition, the DOH advises (i) a restrictive antibiotic policy, (ii) environmental cleaning with a chlorine-based disinfectant, and (iii) rigorous hand washing before and after each patient contact (145). It is noteworthy that alcohol-based hand gels are not effective at killing the spores of *C. difficile* (61).

Further research has indicated that barrier nursing (which includes the use of disposable gloves and aprons) and having medical equipment dedicated to each patient can reduce *C. difficile* infections and cross-infections (146). A study by Zafar *et al*. (1998) indicated that enforcement of this multi-faceted policy together with intensive education and aggressive surveillance activity successfully reduced the frequency of CDAIs (147).

Ultimately, the successful management of *C. difficile* requires both the treatment of symptomatic patients and established infection control practices to manage the spread of the pathogen. Clearly, in the event of an outbreak, identifying the source is critical for effective disease control, since this is the most prudent way of breaking the chain of transmission. This is a key concept in any infection control strategy. However, it is critically important to note that asymptomatic patients and healthy carriers (such as relatives, patients and medical staff) remain an important reservoir of *C. difficile* infection (148).

1.4 *C. difficile* life cycle

1.4.1 Biology of spores

When conditions are permissive, *C. difficile* undergoes a cycle of vegetative replication via binary fission, which represents the metabolically active stage of growth and is a characteristic of all bacteria. However, in contrast to non-spore-forming bacteria, *C. difficile* can also enter a cycle of sporulation, resulting in a metabolically dormant cellular structure known as a spore

(Figure 1.8). Spores are exquisitely resistant to environmental insults (such as extremes of temperature, desiccation, and radiation) and are arguably one of the most robust life forms on Earth (149). It is the capacity of *C. difficile* to produce spores that is the principle feature in conferring its resistance to many environmental disinfectants (61), thus enabling the pathogen to persist in the environment for extended periods, facilitating its transmission.

1.4.1.1 Role of *C. difficile* spores

Unlike many hospital pathogens, *C. difficile* produces highly infective spores that are excreted by infected patients, allowing the oxygen sensitive pathogen to retain viability outside of the host (150). Spores of *C. difficile* can also resist commonly used hospital disinfectant regimens and as a result are able to persist in the environment generating a potential transmission reservoir that confounds many infection control measures (61, 151, 152). Moreover, spores are often able to resist antimicrobial therapy and persist in the host, contributing to recurrent disease.

C. difficile is unusual amongst some spore forming pathogens, including other clostridia, in being highly transmissible between humans. In contrast, there is little evidence of person-to-person transmission of *Clostridium botulinum* (153), *Clostridium tetani* (154) or *Bacillus anthracis* (155). Thus, transmissibility may be key to the continued survival and persistence of *C. difficile* in the human population, and suggests that the pathogen is highly adapted to its niche.

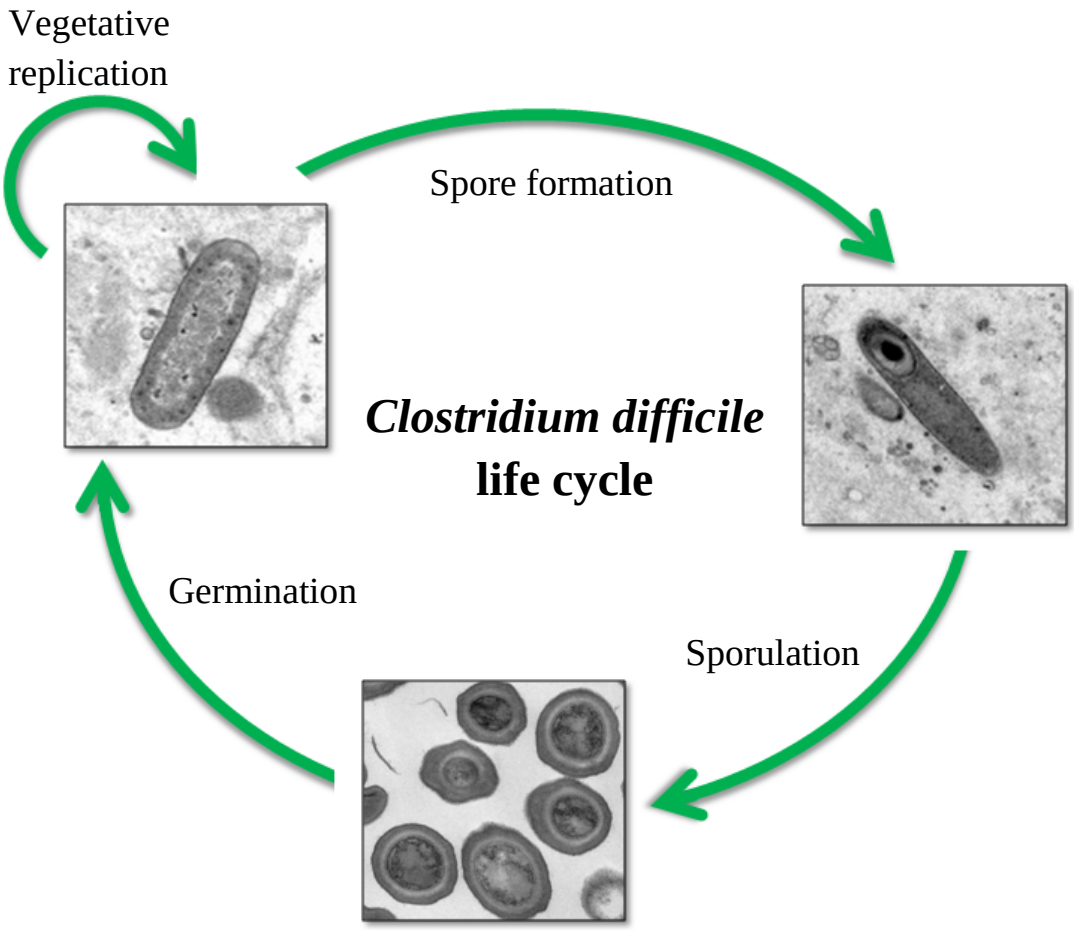


Figure 1.8. Schematic representation of the *Clostridium difficile* life cycle. When conditions are permissive, *C. difficile* is metabolically active and undergoes vegetative replication. In periods of stress, *C. difficile* is able to enter a cycle of spore formation, during which a metabolically dormant spore is formed within the mother cell. Upon maturation, the mother cell lyses and the spore is released into the environment. When conditions become favourable once more, the spore germinates, outgrows, and vegetative growth is resumed. TEM images taken by David Goulding (WTSI).

1.4.2 Structure of the spore

The structure of *C. difficile* spores follows a common concentric architectural plan typical of spores in general, and their hardy nature can largely be explained by their unique cellular anatomy, as shown in Figure 1.9. At the centre of the spore is the core, which is the site of cellular DNA and ribosomes. The core is maintained in a semi-dehydrated state in high concentrations of dipicolinic acid (DPA), which together are thought to maintain dormancy and contribute to the heat resistance of the spore (156-159). Small acid-soluble proteins (SASPs) are also believed to saturate the spore DNA and help protect against heat, desiccation and genotoxic assaults (157, 160, 161).

In a mature spore, the core is surrounded by a thick mantle of concentric layers (Figure 1.9). Together, these protect the inner core and make the spore resistant to factors that would generally be regarded as bactericidal and which would readily kill the vegetative form of *C. difficile* (162). The core is delineated by an inner membrane (a putative permeability barrier) (149, 163), the germ cell wall (which becomes the cell wall of the vegetative cell at germination) (164, 165), and the cortex, which in contrast to the core is a highly hydrated matrix rich in peptidoglycan. It is the hydrated nature of the cortex that maintains the dehydrated nature of the core via physical and/or osmotic pressure (166).

The spore coat lies over the outer membrane and is a hardened proteinaceous and carbohydrate-rich layer comprising several distinct striations which protects the underlying spore layers from enzymatic and chemical agents (167, 168). De-coated spores do not survive

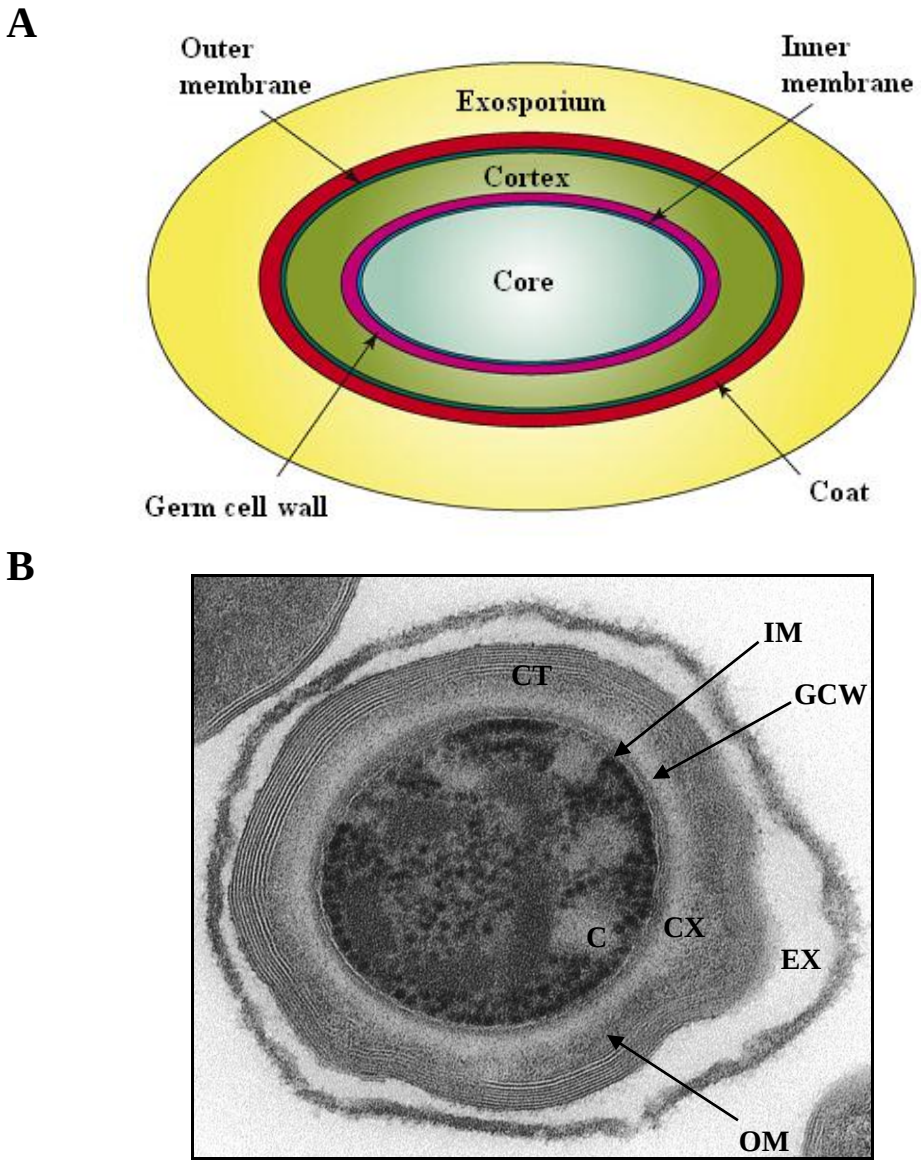


Figure 1.9. Ultrastructure of the bacterial endospore. A) Schematic representation of the salient features of a bacterial spore. Layers are not drawn to scale. Adapted from: (161). **B)** TEM image of a mature *C. difficile* spore, comprising a thick mantle of concentric layers. Shown innermost is the core (C) which contains nuclear DNA and ribosomes, and is delineated by the continuous inner membrane (IM). Over this lies the germ cell wall (GCW) and a thick, peptidoglycan-rich cortex (CX), both of which are important for spore integrity, and are encircled by the outer membrane (OM). The spore coat (CT) comprises several distinct striated layers and is situated inside the exosporium (EX; outermost layer). TEM image taken by David Goulding (WTSI).

exposure to gastric conditions, suggesting that the spore coat may also confer protection during transit through the gastrointestinal tract (168). The outermost layer of the spore is the exosporium, which is a clearly defined shell rich in glycoproteins. A further function of the exosporial layer is likely to be in adhesion and colonisation (169, 170) and, although present in *C. difficile*, may be absent in other spore forming bacteria.

1.4.3 General process of spore formation, sporulation and germination

Spore formation is a property of several bacterial genera and is historically described using *Bacillus* as a canonical system. The cycle of spore formation, sporulation and germination is complex and is marked by a series of biochemical and physiological changes that mediate the sequential development of new structures (171). An overview of spore formation, sporulation and germination is given in Figure 1.10, which uses major morphological changes as indicators to define the principle stages of development. The entire cycle can be loosely defined by seven cytological "stages", as described in Figure 1.10 and below.

1.4.3.1 Stages of sporulation

Stage 0 represents a preparatory phase and defines a vegetatively growing cell, with a no discrete spore structures visible. At present, stage I is not used to describe a physiological state, and as such developing vegetative cells proceed from stage 0 to stage II, which is the

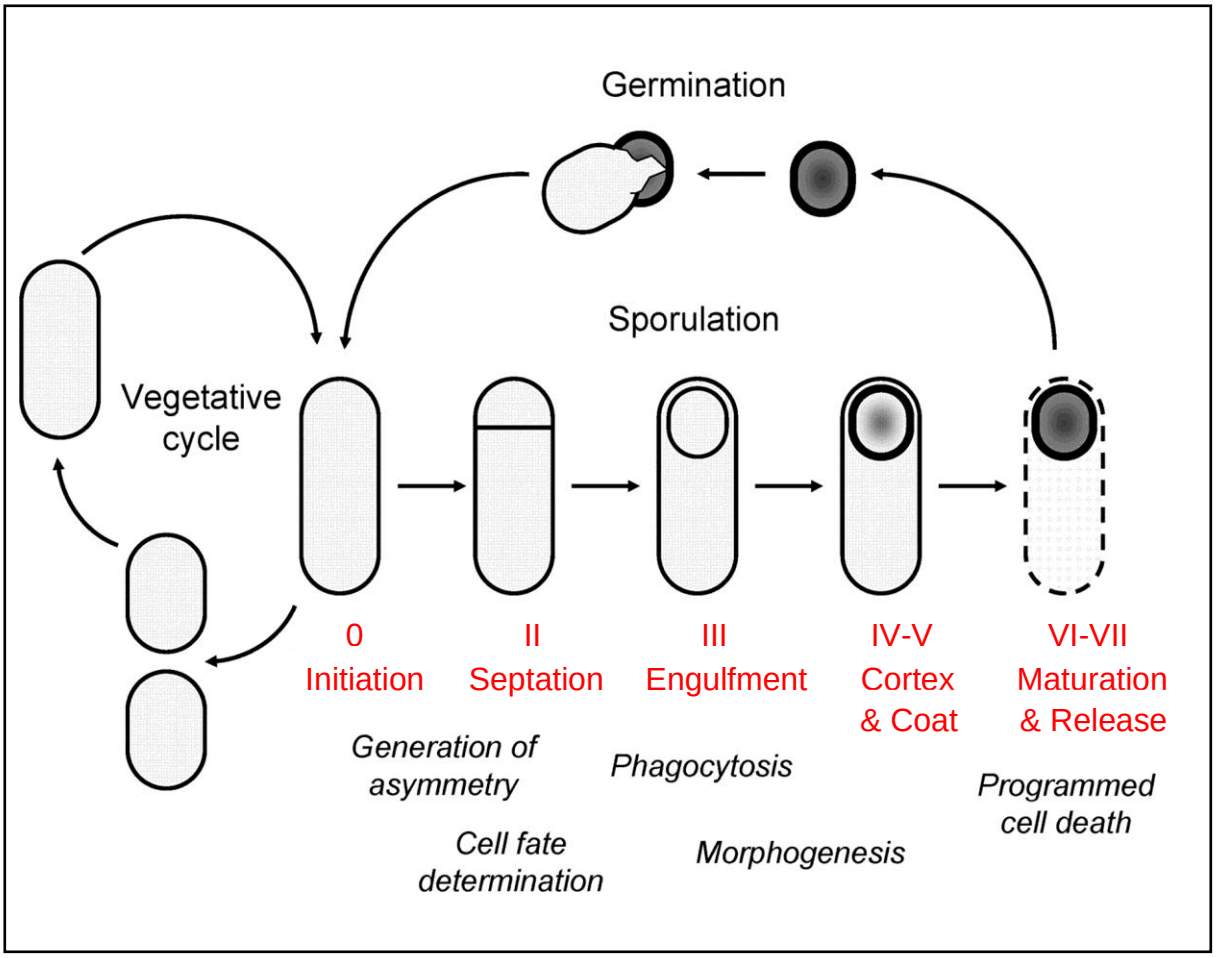


Figure 1.10. Schematic overview of spore formation, sporulation and germination. The cycle of spore formation is defined by 7 cytological stages (I-VII, excluding stage I), in which key morphological indicators (shown in red) are used to define the physiological state of the cell. Vegetative growth is resumed following germination, which occurs in response to currently unknown molecular cues. During germination the spore sheds its protective layers and outgrows, becoming metabolically active once more. Adapted from: (172).

first morphologically identifiable point in the cycle (172). Since spore formation arises from asymmetric cell division, stage II represents the point at which the cell favours a polar division site to initiate septum formation, rather than the central division site typical of binary fission. The resulting asymmetric septation generates two unequal compartments with distinct developmental fates: the smaller pre-spore and the larger vegetative mother cell from which it was derived (172).

During stage III, the mother cell engulfs the pre-spore via phagocytosis, generating a cell-within-a-cell and an immature pre-spore with a double membrane. After engulfment, a peptidoglycan-rich cortex is deposited between the inner and outer pre-spore membranes (stage IV) and the spore begins to develop its refractive properties (172). This is followed by stage V during which the multi-layer spore coat forms on the outer pre-spore membrane. Stage VI is the final stage of spore maturation and involves DPA synthesis. Although this is a spore-specific compound, it appears to be synthesised by and transported from the mother cell (173). The process culminates at stage VII with the liberation of a mature spore from the vegetative mother cell (171).

1.4.3.2 Stages of germination

Once growth conditions are favourable, the dormant spore is able to shed its protective layers and resume vegetative growth in a process known as germination. Recent research on the specific cues that signal germination initiation in *C. difficile* have identified taurocholate (a

bile salt derivative) and glycine as putative germinants (63). In a comparable way to spore formation, germination is characterised by work in *Bacillus* and can be arbitrarily divided into three principle stages (165). Stage I is defined by the release of the DPA pool in the spore core and its replacement with water, thus reversing spore dormancy. This is a hallmark feature of spore germination. In stage II, cortical peptidoglycan is hydrolysed, the core continues to hydrate and swell and the germ cell wall begins to expand (165). This is followed by the final stage, outgrowth, in which the new, viable vegetative cell is released from the remnants of the spore coat. At this stage, SASPs are also degraded rendering the vegetative cell susceptible to environmental insults (161, 165). Thus, spores of *C. difficile* are highly adapted for long-term survival in harsh environmental conditions, but are able to germinate in response to specific molecular cues (165, 174).

1.4.4 *Bacillus*: the paradigm spore former

B. subtilis is the most extensively studied model of bacterial sporulation, and the molecular basis of spore formation is comparably well characterised within this genus (175). It was initially believed that the grand scheme of spore formation in *Clostridium* and *Bacillus* species comprised similar processes inherited from a common ancestor that were mechanistically homologous. However, subsequent genome comparisons have highlighted fundamental differences, as shown in Figure 1.11. Additionally, many genes encoding classical *Bacillus* spore coat proteins were not identified in the *C. difficile* genome (168). Similarly, many canonical germination receptors found in *Bacillus* are absent in *C. difficile* (10).

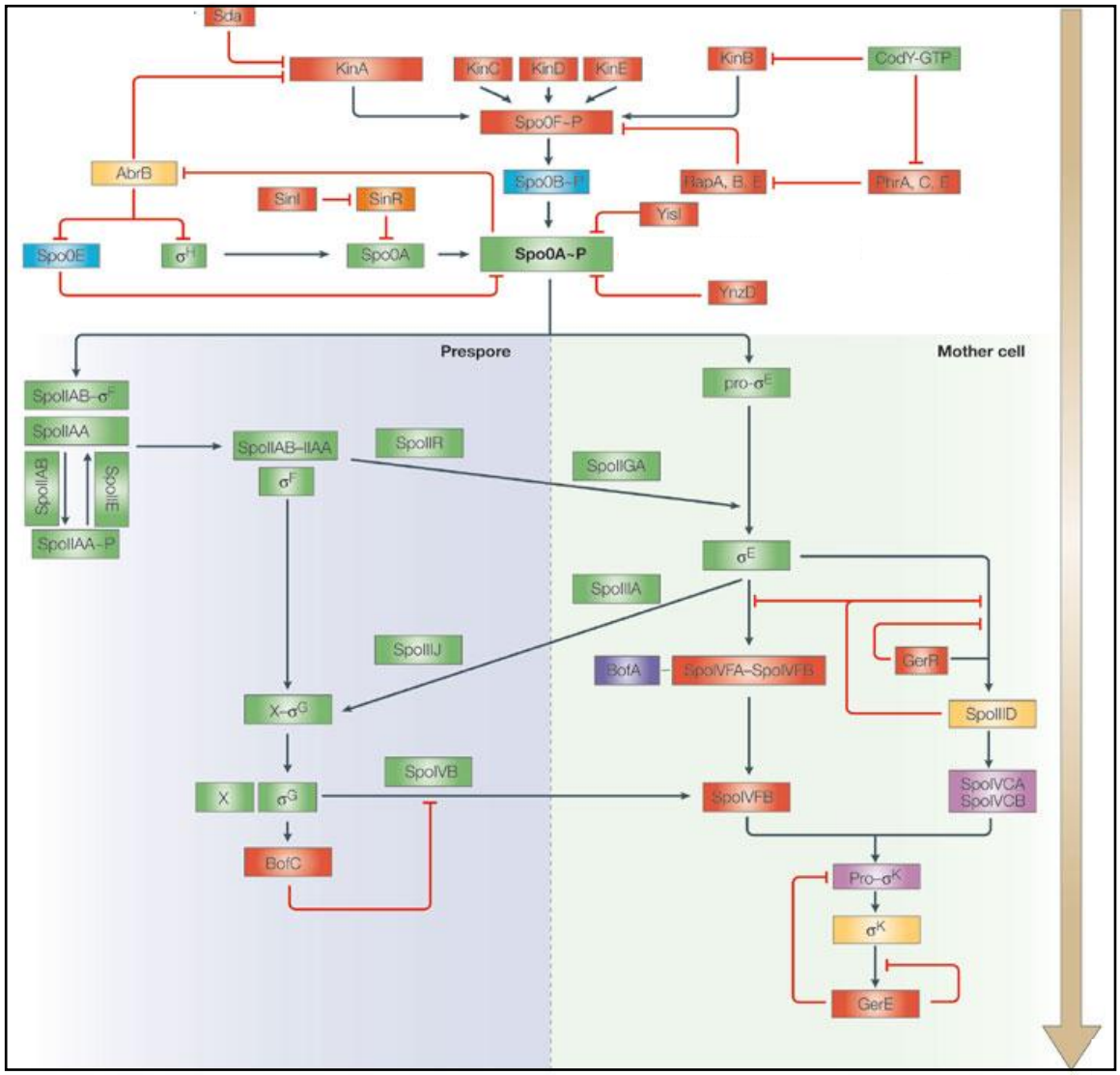


Figure 1.11. Canonical overview of clostridial spore formation. Schematic representation of the *B. subtilis* sporulation cascade, with associated clostridial orthologs. Orthologs are coloured according to presence or absence as follows: red, absent in all clostridia; green, present in all clostridia; blue, possible presence in *Clostridium tetani* only; yellow, present in all clostridia except *C. tetani*; orange, present in *C. tetani* and *Clostridium acetobutylicum* only; purple, present in *Clostridium botulinum* and *C. acetobutylicum* only; pink, requires post-translational processing to form a functional protein in *B. subtilis* and *C. difficile*. Adapted from: (176).

These differences are likely related to the time these classes are estimated to have diverged; clostridia emerged approximately 2.7 billion years ago (BYA), prior to the great oxidation event, whereas bacilli are comparatively younger, appearing as a separate class approximately 2.3 BYA, around the same time as the great oxidation event (176). This is suggestive of a divergent sporulation programme and is in line with the observation that characteristics of sporulation in *Bacillus* species do not completely overlap with the related process in *C. difficile*.

1.5 Stage 0 sporulation protein A

The initiation of sporulation is orchestrated to a significant extent by the transcription factor and master response regulator, Spo0A (176). In *C. difficile*, the DNA binding protein, Spo0A is encoded by the 825 bp gene, *spo0A*, which demonstrates a high degree of amino acid sequence conservation across different *C. difficile* ribotypes (Figure 1.12) and is present in all clostridia (Figure 1.11). Inactivation of *spo0A* results in a sporulation null phenotype in all species studied, including *C. difficile* (177). However, much of the research into Spo0A has derived from *Bacillus* species, and as such the role of Spo0A in *C. difficile* is largely unexplored.

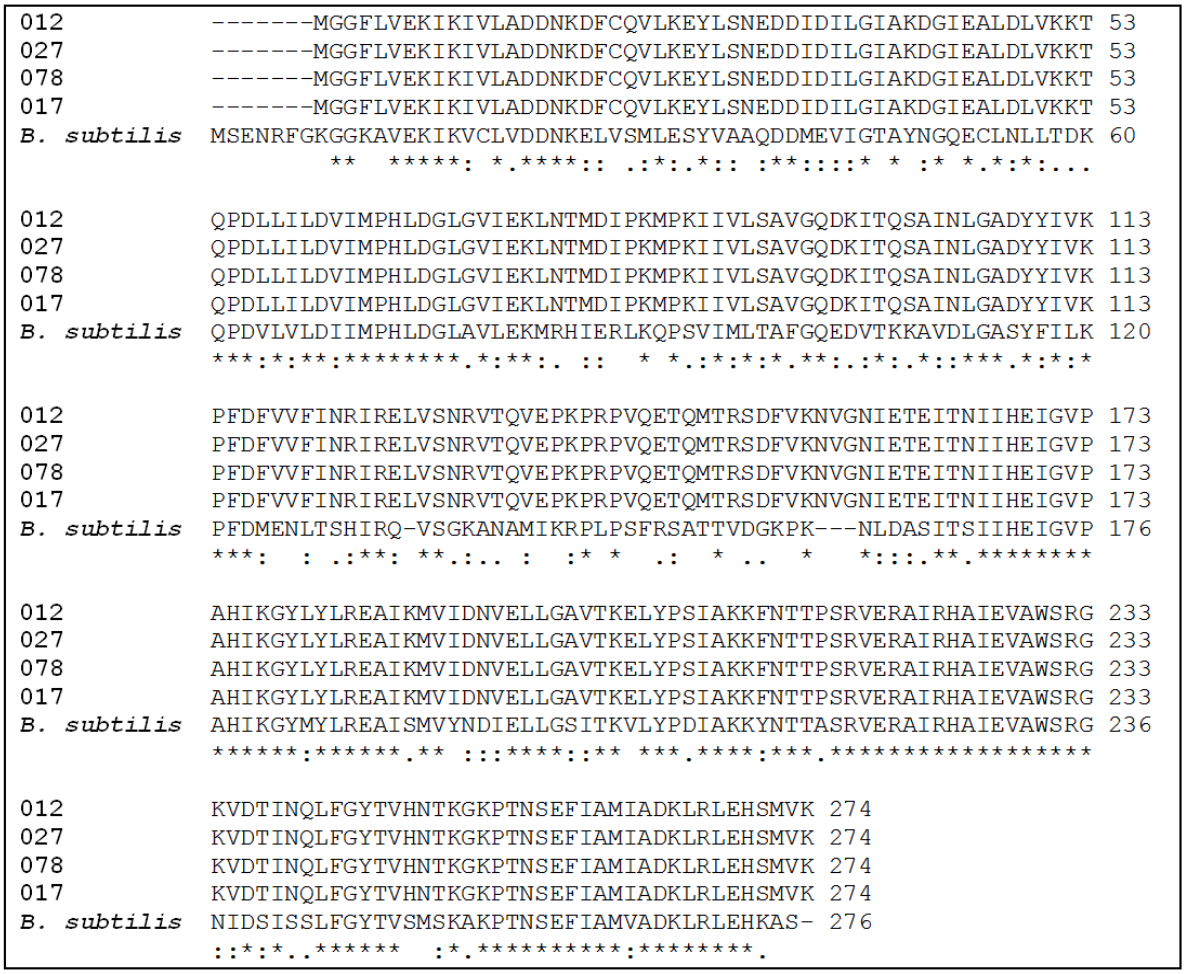


Figure 1.12. Amino acid sequence alignment of Spo0A from common *C. difficile* ribotypes. ClustalW2 (2.1) multiple sequence alignment of Spo0A from commonly isolated human virulent *C. difficile* ribotypes. *B. subtilis* Spo0A is included for reference. Symbols: "*", identical; ".", conserved substitution; ":", semi-conserved substitution.

1.5.1 Pathway to Spo0A activation

As already discussed, *Bacillus* and *Clostridium* are divergent genera, which differ most notably in the earlier part of their sporulation programmes (Figure 1.11) (176). For example, many of the genes involved in the signal transduction pathway of sporulation initiation in *Bacillus* are absent in *C. difficile* (176). In *Bacillus*, Spo0A activity is governed by a classical multicomponent phosphorelay system, comprising at least five histidine autokinases (KinA-KinE) which in turn relay phosphoryl groups via the Spo0F response regulator and the Spo0B phosphotransferase, and ultimately to Spo0A to modulate its activity (Figure 1.11).

However, orthologs of such kinases and phosphorelay proteins are not apparent in the genomes of sequenced clostridia, suggesting that the pathway to Spo0A activation is via a two-component system in which Spo0A is phosphorylated directly, possibly by a membrane bound orphan kinase (Dr W. K. Smits, personal communication). Although this hypothesis is widely accepted in the field, the signals that induce spore formation and the transcriptomic pathway leading to Spo0A activation are not well understood.

1.5.2 Structure and role of Spo0A as a transcription factor

In *Bacillus*, Spo0A has a distinct two-domain structure, comprising a helix-turn-helix DNA-binding function in the carboxy-terminus, and an amino-terminal phosphoacceptor domain. As a response regulator, Spo0A only becomes functional as a transcription factor following

the phosphorylation of a conserved aspartate residue of the latter domain, resulting in dimerisation which enables the protein to bind to cognate DNA sequences (178). Spo0A binds specifically to a target recognition sequence known as a '0A' box comprising the 7 bp sequence 5'-TGNCGAA-3' which is found in or near the promoters of genes in the regulon (179). Depending on the exact location of binding, Spo0A serves as either an activator or repressor of gene transcription.

1.5.2.1 Mutual regulation of gene expression by Spo0A and σ^H

Previous work in *Bacillus* has indicated that following the phosphorylation (activation) of Spo0A, a regulatory circuit is rapidly initiated (180, 181). In *B. subtilis*, this feedback loop has been shown to involve the alternative sigma factor, σ^H, which is encoded by the *sigH* gene, and the transition state transcriptional regulator AbrB. The *sigH* gene is present in all sequenced clostridia, and its translation is also essential for sporulation (182). There is no *abrB* orthologue in *C. difficile*. It has been reported that, in *B. subtilis*, Spo0A and SigH mutually regulate their transcription, however this is mediated by AbrB (183). For example, phosphorylated Spo0A activates its own transcription (both directly and indirectly) via the de-repression of the AbrB-regulated gene, *sigH*.

The reciprocal control of SigH and Spo0A synthesis has also been reported in *C. difficile.* Indeed, *C. difficile spo0A* is thought to be transcribed from a SigH-dependent promoter (182). Furthermore, a consensus '0A' box has been identified directly upstream of the both the *C.*

difficile spo0A and *sigH* genes (182). Since there is no evidence of a *C. difficile abrB* orthologue, it is thus possible that this mutual regulation is via a direct interaction, rather than indirectly via AbrB.

1.5.2.2 The *Bacillus* Spo0A regulon

In *B. subtilis,* Spo0A is known to temporally and spatially regulate the transcription of over 120 genes directly (184) and over 586 genes in total, representing > 10% of the *B. subtilis* genome (185). Furthermore, the genes under the control of Spo0A belong to multiple classes with varied cellular functions (184), indicating that Spo0A truly is a global regulator of gene transcription. Whilst many of the targets of Spo0A in *B. subtilis* are not present in the *C. difficile* genome (of which AbrB is the most extensively described), some of the key transcriptional units under Spo0A regulation which are required for sporulation (such as SpoIIA, SpoIIE), are present in both genera (Figure 1.11), suggesting that certain aspects of the Spo0A regulon may be conserved.

Data regarding the *C. difficile* Spo0A regulon is limited. Previous studies have tentatively indicated that Spo0A may modulate TcdA production, however the correlation is contentious (177, 182). Cataloging the *C. difficile* Spo0A regulon on a genome-wide basis is a focus of my thesis.

1.6 Genomics and genetic tools

1.6.1 Mutagenesis

Historically, it has been a challenge to manipulate *C. difficile* genetically, and there has been a longstanding inability to readily construct site-directed chromosomal mutations. Indeed, for a long time *C. difficile* was almost considered refractory to mutagenesis. Classical strategies of genetic modification, such as "knock-in" and "knock-out" approaches have largely proved unsatisfactory in *C. difficile* (due to extremely low transformation and recombination frequencies and mostly unstable single cross-over plasmid insertions). However, due to the clinical and economic prominence of CDAIs there is clearly a need for efficient means of mutagenesis. Recently, multiple methods which enable the generation of highly stable insertional mutants that are maintained in the chromosome have been developed (186-188).

1.6.1.1 ClosTron technology

Originally developed by the Lambowitz group (189) and marketed by Sigma Aldrich as "TargeTron" technology, the ClosTron technique uses a rapid and relatively simple recombination independent approach to gene inactivation, by exploiting the activity of a *Lactobacillus lactis* Ll.ltrB group II intron. Target recognition is governed by the base pairing specificity between intron RNA and target DNA, following which the intron inserts into the gene of interest and leads to the production of a truncated and non-functional protein.

Importantly, incorporated within the intron is a retrotransposition-activated marker (RAM) comprising the erythromycin resistance gene, *ermB*, thus enabling the selective isolation of integrants (186, 190). For a more comprehensive description of directed mutagenesis using ClosTron technology, please refer to Materials and methods chapter 2.2.5.9.

1.6.2 Transcriptomics and proteomics

The profound morphological and physical changes that occur during the course of spore formation in *C. difficile* will be reflected by a shift in gene expression and proteome profiles. As such, the ability to investigate the global behaviour of the *C. difficile* transcriptome and proteome in a quantitative manner during the course of spore formation is essential if we are to identify gene products, regulons and regulatory molecules associated with this phenomenon. Transcriptomics and proteomics mutually complement each other and parallel profiling of RNA transcripts and proteins is clearly necessary in order to gain a comprehensive and integrated overview of spore formation, and to derive a holistic picture of *C. difficile* biology and physiology.

1.6.2.1 ssRNA-Seq

DNA microarrays have been the mainstay of transcriptome analyses, and until recently their primacy had remained largely unchallenged. Recently, however the Illumina sequencing

platform has been adapted to enable direct, high-throughput screening of bacterial transcriptomes. The Wellcome Trust Sanger Institute has developed a method, termed high-density, strand-specific cDNA sequencing (or ssRNA-Seq) (191), which is a post-genomic strategy that is rapidly emerging as a powerful and reproducible tool to quantitatively survey global transcript abundance. The utility of RNA-Seq has previously been demonstrated in *Salmonella* Typhi (191), *B. anthracis* (192), *Mycobacterium tuberculosis* (193) and *Vibrio cholera* (194).

RNA-Seq is arguably a superior alternative to microarrays since it (i) avoids biases due to hybridisation, (ii) offers a higher sensitivity for less abundant transcripts, and (iii) enables the identification of transcripts mapping to previously unannotated regions of the genome. It does, however present certain algorithmic and bioinformatic challenges.

1.7 Aims of thesis

The aims of this study are broadly:

(i) To assess the role of the *C. difficile spo0A* gene in disease and transmission

(ii) To provide a molecular description of the *C. difficile* Spo0A regulon at the whole genome level.

2 Materials and methods

2.1 Materials

2.1.1 Bacteria and plasmids

Table 2.1 describes the bacteria and plasmids used in this study.

Table 2.1. Strains and plasmids used in this study.

Strain or plasmid	Characteristics	Source
Strains		
E. coli CA434	Electrocompetent conjugation donor	(195)
Top10	Electrocompetent *E. coli* for plasmid and amplicon cloning	Invitrogen
Rosetta(DE3) pLys	Competent *E. coli* for protein expression	Novagen
C. difficile 630	Virulent and multidrug resistant PCR-ribotype 012, isolated from a patient with pseudomembranous colitis in Zurich, Switzerland (1985)	(10)
C. difficile 630Δ*erm*	Erythromycin sensitive derivative of *C. difficile* 630	(196)
C. difficile R20291	Hypervirulent and epidemic PCR-ribotype 027, isolated from a hospital outbreak in Stoke Mandeville, UK (2004-2005)	(11)

630Δ*erm*Δ*spo0A*	*C. difficile* 630Δ*erm spo0A*::*erm*	This study
R20291Δ*spo0A*	*C. difficile* R20291 *spo0A*::*erm*	This study
630Δ*erm*Δ*spo0A*+p*spo0A*	Complemented 630Δ*erm*Δ*spo0A* mutant	This study
R20291Δ*spo0A*+p*spo0A*	Complemented R20291Δ*spo0A* mutant	This study

Plasmids

pMTL007	First generation ClosTron plasmid with *catP* marker and intron containing *erm* RAM. Intron expression is induced using IPTG	(186)
pMTL007C-E2	Second generation ClosTron plasmid with *catP* marker and intron containing *erm* RAM. Contains a constitutive *fdx* promoter to direct intron expression	(190)
pRPF101	*E. coli – C. difficile* shuttle vector	This study
p*spo0A*	pRPF101 containing the 825 bp *spo0A* coding region (and upstream promoter)	This study
pWKS1245	Plasmid for the production of full-length Spo0A protein with a C-terminus His_6-tag	This study

2.1.2 Oligonucleotides

Oligonucleotides used in this study were designed by Primer*3* (0.4.0) software, synthesised by Sigma Aldrich, and are described in Table 2.2. All were dissolved in 10 mM TE buffer to a concentration of 40 pmoles/µl. Oligonucleotide pairs were designed with approximately the same G/C content, length and annealing temperature (T_m).

Table 2.2. Oligonucleotides used in this study.

Oligonucleotide	**Sequence (5'-3')**
Intron retargeting and screening	
spo0A-178|179a-IBS	AAAAAAGCTTATAATTATCCTTATTATTCCATCTAG TGCGCCCAGATAGGGTG
spo0A-178|179a-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCC ATCTAGTTAACTTACCTTTCTTTGT
spo0A-178|179a-EBS2	TGAACGCAAGTTTCTAATTTCGGTTAATAATCGATA GAGGAAAGTGTCT
RAM-F	ACGCGTTATATTGATAAAAATAATAATAGTGGG
RAM-R	ACGCGTGCGACTCATAGAATTATTTCCTCCCG
EBS universal primer	CGAAATTAGAAACTTGCGTTCAGTAAAC
spo0A F	GCTAAGGATGGAATTGAAGCA
spo0A R	GCTCCTAGATTTATTGCGCTTT
spo0A complementation	
spo0A_A1F	(ATAT)GTCGACGGTGCAATAACTCATGTTTTTAGAG

spo0A_A2Rb	(ATAT)GTCGACGACTCTCATATTTAAACCTCCAC
DNA screening primers	
CD1498 F	GATTGCAGATGCATGTGGTT
CD1498 R	TTGGAGAGCAAGAACAGCAA
CD1455 F	GATGCAGAGGCAATTTCACA
CD1455 R	GCTAGAAGGATGCACGAAGG
CD0011 F	CCAGCTTTGCAACACCAACT
CD0011 R	GGCTATGGAGGCTTCTTATGG
CDadk F	TTACTTGGACCTCCAGGTGC
CDadk R	GCAGCCTTAGGAAGTGGAAA
C. difficile spo0A for protein purification	
oWKS-1122	TTTCATATGGGGGGATTTTTAGTGG
oWKS-1123a	TGCTCGAGTTTAACCATACTATGTTCTAGT

2.1.3 Mice

All experiments were carried out using 5-to-9 week old specific-pathogen-free mice from colonies maintained at the Wellcome Trust Sanger Institute. All mice were of C57BL/6 wild type genetic background. Mice were housed in sterile cages with *ad libitum* access to food

and water. All animal infections were performed in accordance with the UK Home Office Animals (Scientific Procedures) Act of 1986.

2.2 Methods

2.2.1 Bacterial culture

All *C. difficile* were grown for 24 to 48 h at 37°C under anaerobic conditions in a MACS MG-500 anaerobic workstation (Don Whitley Scientific). *C. difficile* was routinely cultured in Wilson's broth (197) with agitation (80 rpm) or on CCEY agar (Bioconnections) supplemented with cycloserine (250 µg/ml; Bioconnections), cefoxitin (8 µg/ml; Bioconnections) and 0.1% taurocholate (Sigma Aldrich). *E. coli* were grown at 37°C using Luria-Bertani (LB) broth with agitation (200 rpm) or LB agar with appropriate selection. For the enumeration of spores, *C. difficile* cultures were mixed with 100% ethanol (1:1 ratio) for 1 h at room temperature to kill vegetative cells, pelleted, washed in PBS and cultured as above.

2.2.2 *In vivo* methods

Infection monitoring, faeces and organ collection and plating were performed by L. J. Pettit. Mouse infections, bleeding and sacrifices were performed by Dr S. Clare (WTSI, Cambridge), with assistance from L. J. Pettit.

All procedures and mouse handling were performed aseptically in a biosafety cabinet to contain spore-mediated transmission. A clinical scoring system was employed to track the condition of the mice, and symptoms including abnormal/hunched gait, piloerection, lethargy and emaciation were monitored. Moribund mice or mice displaying overt signs of disease were sacrificed humanely in accordance with approved protocols. Mice were anaesthetised with isoflourane (IsoFlo) prior to surgical procedures or interventions, and were sacrificed via cervical dislocation.

2.2.2.1 Mouse infections

Mice were pre-treated with drinking water containing clindamycin (250 mg/L; Apollo Scientific) for 7 d. Clindamycin was then withdrawn for 24 h and donors were infected via oral gavage or transmission, as indicated.

2.2.2.1.1 Oral gavage with C. difficile

C. difficile cultures were grown overnight in broth as described above, after which the culture was diluted 1/4 in PBS. Mice were then orally inoculated (with anaesthetic) with 200 µl of the bacterial suspension using a sterilised blunt-tipped gavage needle. When infected via oral gavage, mice received 10^7 vegetative cells and 10^5 spores (R20291 and 630Δ*erm* parental

strains) or 10^7 vegetative cells (R20291Δ*spo0A* and 630Δ*erm*Δ*spo0A* mutant derivatives), as determined by plate counts of the inocula.

2.2.2.1.2 Enumeration of viable C. difficile from mouse faeces

C. difficile was enumerated from fresh faeces as previously described (98). Briefly, mice were placed under sterilised beakers and faeces were collected. Samples were homogenised in phosphate-buffered saline (PBS; 100 mg faeces/ml PBS), serially diluted in PBS and plated on CCEY agar with appropriate supplementation. This was always performed within 30 min of excretion. For the enumeration of spores, samples were mixed with 100% ethanol (1:1 ratio) for 1 h at room temperature to kill vegetative cells, pelleted, washed in PBS and cultured as above.

2.2.2.1.3 Competitive index infections

Mice (n = 5) were infected via oral gavage with 10^7 CFU in 0.2 ml PBS of an overnight culture in 0.2 ml PBS containing equal proportions of parental *C. difficile* and the respective isogenic *spo0A* mutant. In order to determine the inoculum dose and ratio of parent to *spo0A* mutant derivative, the inocula were grown on both CCEY plates (as a non-selective media) and CCEY supplemented with 20 mg/ml lincomycin (which selects for the *ermB* gene inserted into the *spo0A* gene of both mutant derivatives). Faecal samples were then diluted as above

and plated on CCEY agar and CCEY agar supplemented with lincomycin. The competitive index (CI) was determined by dividing the ratio of mutant to parent bacteria using the following formula:

CI = (*spo0A* mutant/parental strain)$_{output}$/(*spo0A* mutant/parental strain)$_{input}$.

All data were logarithmically converted prior to averaging and statistical analysis. The Mann-Whitney test was applied to the $\log_{10}$ values of the CI ratios to determine the statistical significance of the results.

2.2.2.1.4 Transmission experiments

Donor mice (n = 5) were infected via oral gavage as described above. To measure transmission efficiency, naïve recipient mice (n = 5) were pre-treated with clindamycin as above and used in one of four types of transmission assays, as described in Table 2.3 and Figure 2.1. Following transmission, recipient mice were individually housed and the transmission efficiency was determined after 4 d by the isolation of the same strain of *C. difficile* in the faeces that was used to infect the respective donor mice.

Table 2.3. Summary of transmission routes.

Type of transmission (1 h)	Characteristics
Mingling	Donor and recipient mice were able to freely touch and interact, with no restriction in movement. Potential for coprophagy
Contact	Donor and recipient mice were separated by a single porous barrier, and were able to touch each other. No potential for coprophagy
Airborne	Donor and recipient mice were separated by a double porous barrier, and were unable to touch each other. No potential for coprophagy
Environmental	Recipient mice were housed in a contaminated donor cage, which had been left for 24 h under ambient oxygen conditions. Donor faeces were removed to prevent transmission via coprophagy

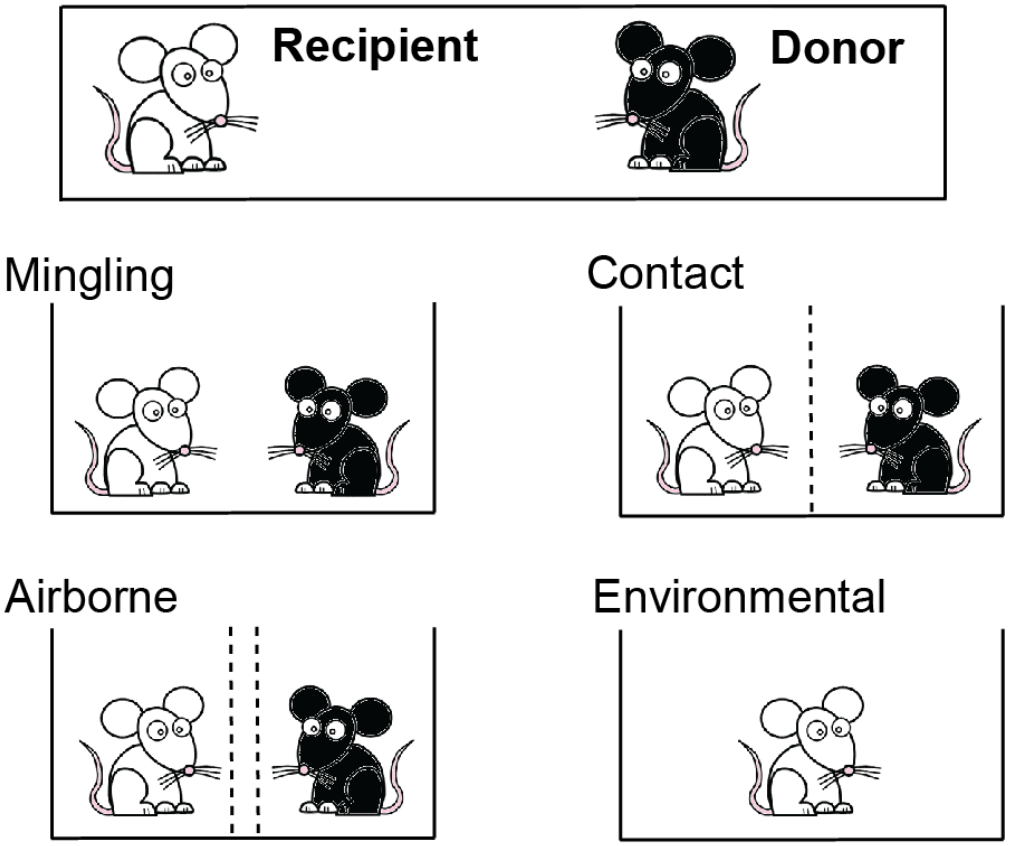


Figure 2.1. Schematic diagram demonstrating the experimental models used to investigate the role of the *C. difficile spo0A* gene in host transmission. Donor mice infected with comparable levels of *C. difficile* R20291, 630Δ*erm* or their equivalent *spo0A* mutant derivatives were exposed to susceptible naïve recipient mice via distinct transmission routes. Dashed vertical lines represent porous barriers between donor and recipient mice. A definition of the transmission routes is given in Table 2.3.

2.2.2.2 Antibody generation and serum extraction

Antibodies were generated via subcutaneous delivery of *C. difficile* 630Δ*erm*Δ*spo0A* into mice (for anti-vegetative cell sera; $n = 5$) or *C. difficile* 630Δ*erm* pure spores into rabbits (for anti-spore sera; $n = 2$). The latter was performed by Cambridge Research Biochemicals, UK. Animals received 100 µl of a ~10^7 CFU/ml culture in each flank, followed by a series of boost immunisations and test bleeds. To generate anti-Spo0A antibodies, mice ($n = 4$) received 30 µl (10 µg) purified Spo0A protein via intranasal delivery (15 µl/nare). Additionally, mice received *E. coli* heat-labile toxin (LT) to potentiate the immune response to Spo0A. Identical boosts were administered on days 7 and 21, without LT. Whole blood was collected, centrifuged at full speed (13,000 rcf) for 15 min, and sera were collected and stored at 4ºC.

2.2.3 Tissue methods

2.2.3.1 Paraffin embedding and sectioning of caecum tissue

For pathological analysis, mice were anaesthetised with isoflourane, sacrificed via cervical dislocation and surface sterilised with 70% ethanol. Caecum tissue (0.5 cm tubular sections) was then carefully excised, opened and incubated in 4% paraformaldehyde for 24 h at room temperature. Tissues were processed in a Shandon Excelsior Tissue Processor (Fisher Scientific) and paraffin wax embedded. Sections (5 µm) were cut using a RM2125 microtome (Leica) and slide mounted.

2.2.3.1.1 Histology

Paraffin wax embedded sections were processed for haematoxylin and eosin staining, as previously described (98). Briefly, sections were de-paraffinised and rehydrated stepwise as follows: Histoclear (Fisher Scientific; 2 min), two washes of 100% ethanol (2 min each), 90% ethanol, 70% ethanol (2 min) and water (5 min). Sections were stained with Mayer's haematoxylin (2 min) followed by washing in water and 1% ethanol to remove excess dye, followed by staining with eosin (5 min), and repeated washing as above. Sections were then dehydrated as follows: 70% ethanol (2 min), 90% ethanol (2 min), 100% ethanol (2 min) and Histoclear (2 min). Finally, slides were mounted with DPX resin, dried and visualised using a LSM510 Meta confocal microscope (Carl Zeiss Ltd.).

2.2.4 Microscopy

2.2.4.1 Indirect immunofluorescence

Cultures of *C. difficile* vegetative cells and spores were washed in PBS and seeded onto microscope slides. Briefly, slides were then fixed in acetone and blocked with 1% BSA in PBS. Wells were incubated with polyclonal rabbit sera (raised against pure wild-type *C. difficile* 630Δ*erm* pure spores) and polyclonal mouse sera (raised against 630Δ*erm*Δ*spo0A* vegetative cells), washed in PBS and reacted with Cy3-conjugated goat anti-rabbit IgG and FITC-conjugated goat anti-mouse IgG (each 1:1000) in a humified chamber, in the dark.

Slides were then mounted in ProLong Gold with DAPI (Invitrogen), sealed and immunofluorescence was assessed using a LSM510 Meta confocal microscope.

2.2.4.2 Preservation for transmission electron microscopy

Tissue preservation and visualisation was performed by David Goulding (WTSI, Cambridge), assisted by L. J Pettit.

Mice were sacrificed and surface sterilised as above. Following excision, ceacum tissue (0.5 cm tubular sections) was immediately incubated in a primary fixative (2% paraformaldehyde, 2% glutaraldehyde 0.1% magnesium chloride, 0.05% calcium chloride in 0.1 M sodium cacodylate buffer [pH 7.4]) for 2 h on ice. Tissue was then washed 3x in sodium cacodylate buffer supplemented with magnesium and calcium chlorides, after which they were incubated in 1% osmium tetroxide in sodium cacodylate buffer for 1 h. Following washing, the tissue was incubated with 1% tannic acid (a mordant) for 1 h, washed with 1% sodium sulfate for 10 min, and dehydrated in an ethanol series for 30 min each (20%, 30%, 50%, 70%, 90%, 95%) additionally staining with 2% uranyl acetate at the 30% ethanol wash stage, followed by a final dehydration stage in 100% ethanol (3x for 20 min each). The tissue was incubated 2x in propylene oxide for 15 min each, then for 1 h in a 1:1 mixture of propylene oxide and Epoxy resin (Epon-812; Sigma Aldrich) for 1 h, and finally in pure Epon-812 resin overnight. Ceacum tissue was embedded and cured in Epon-812 in a flat-moulded tray at 65^{o}C. 50 nm ultrathin sections were cut, loaded onto Formvar-carbon-coated grids (Agar Scientific),

contrast stained with uranyl acetate and lead citrate and ultimately visualised on a FEI 120 kV Spirit Biotwin transmission electron microscope (TEM) fitted with a F415 digital Teitz camera.

2.2.4.3 Flagellar negative staining

Negative staining and visualisation was performed by David Goulding (WTSI, Cambridge), assisted by L. J. Pettit.

Grids were prepared by briefly submerging slides into Formvar (0.1%) in dry chloroform. Formvar-carbon-coated support films were then floated onto distilled water, after which grids were placed onto the film before lifting onto parafilm and air-drying. Fresh bacterial colonies were picked, suspended in ammonium acetate and loaded onto the film side of the grid. An equal volume of ammonium molybdate (1%) was added to the film and immediately drained with filter paper. Samples were allowed to air-dry and were visualised via TEM as described above.

2.2.4.4 Tissue processing for ImmunoGold Electron Microscopy

Tissue processing and visualisation was performed by David Goulding (WTSI, Cambridge), assisted by L. J. Pettit.

Caecum tissue was excised and fixed, opened and incubated in 4% paraformaldehyde, 0.2% gluteraldehyde in 0.1 M phosphate buffer (pH 7.4) for 10 min at 37°C, followed by room temperature for 2 h. Samples were then washed in PBS and low-temperature embedded in Lowicryl HM20 resin (Agar Scientific). Briefly, tissue was dehydrated in an ethanol series for 30 min each (30% ethanol at 4°C, 50% ethanol at 1°C, 70% ethanol at -20°C, 90% ethanol at -20°C, 100% ethanol at -30°C, and finally100% ethanol at -50°C) and impregnated with a Lowicryl-ethanol series for 1 h each (1/3, 1/1, 3/1) followed by an overnight incubation with pure Lowicryl. Samples were then embedded and UV polymerised at -50°C. 50 nm ultrathin sections were cut, loaded onto Formvar-carbon-coated grids and then blocked with 0.02 M glycine in PBS for 10 min followed by fetal calf serum (10%) for 1 h. Tissues were then labelled with anti-spore and anti-vegetative cell antisera (1/200 dilution), washed in PBS x3 and probed with protein A-gold conjugates (Sigma Aldrich) for 20 min as appropriate. Following thorough washing in PBS, tissues were contrast stained with uranyl acetate and lead citrate, and visualised via TEM as described above. IGEM-localised *C. difficile* (and background bacterial cells) were then counted and recorded.

2.2.5 Molecular methods

2.2.5.1 Polymerase Chain Reaction (PCR)

For general amplification of template DNA, a standard *Taq* polymerase PCR reaction with Platnium PCR Supermix (Invitrogen) was used. When products were to be cloned, high

fidelity PCR was performed with *Pfu* DNA polymerase (Promega) or Go*Taq*® PCR mix (Promega). All amplifications were performed in accordance with the manufacturer's instructions, in a total volume of 50 µl. Reactions were performed using 0.2 ml thin wall PCR tubes (ABgene) on a DNA Engine DYAD Thermal Cycler (MJ Research). Template quantity was dependent on the source, but was typically 100 ng for genomic DNA, and up to 50 ng for plasmid DNA. For negative controls, template DNA was omitted. DNA amplifications were performed using the following thermal cycling conditions as a guide: 95°C for 2 min × 1; 95°C for 30 sec, 50°C* for 30 sec, 68°C for 2 min** × 35 cycles; and 68°C for 10 min*** × 1.

* Annealing temperature is dependent primers T_m and was calculated as

** Elongation parameters dependent on product size (approximately 1 min/kb) and DNA polymerase used

*** Extension parameters are dependent on type of DNA polymerase used

2.2.5.2 Cloning PCR amplicons

Amplicons were cloned using TOPO TA Cloning Kits (Invitrogen), according to the manufacturer's instructions. Briefly, 2 µl of PCR product was incubated at room temperature for 25 min with 1 µl salt and 1 µl vector in a final volume of 6 µl. DNA was then transformed via electroporation as described in section 2.2.5.7.

2.2.5.3 Plasmid DNA extraction

Plasmid DNA was isolated using QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer's instructions. Briefly, 3 ml of *E. coli* culture was grown overnight in LB supplemented with chloramphenicol (20 µg/ml), and was harvested by centrifugation. Plasmid DNA was then adsorbed onto a QIAprep membrane, washed and eluted in nuclease-free water.

2.2.5.4 Restriction digestion

All restriction endonucleases were purchased from New England Biolabs. Digestion was performed according to manufacturer's instructions for 2 h at 37°C. All plasmids were designed to enable cleavage with *SalI*, *HindIII* or *BsrG1*. Briefly, 20 units of enzyme were used per µg plasmid DNA, along with 3 µl of the appropriate 10x buffer made up to a final volume of 30 µl with water. Following digestion, samples were examined by gel electrophoresis. In addition, digests to be sub-cloned were extracted with phenol:chloroform:isoamyl alcohol (25:24:1; Sigma Aldrich) and dephosphorylated with Shrimp Phosphatase (Promega) to prevent self-ligation. Samples were heated to 70°C for 1 h to inactivate the phosphatase enzyme prior to ligation.

2.2.5.5 DNA ligation

Ligations of *spo0A* and pRPF101 vectors were performed using T4 DNA ligase (Roche), according to the manufacturer's instructions. Briefly, reactions were carried out in a total volume of 10 µl in the presence of ligation buffer, and were incubated at 14°C overnight. Following ligation, plasmids were cloned into electrocompetent *E. coli* Top10. Plasmids in which ligation and transformation had been successful were then transformed into electrocompetent *E. coli* CA434, which were used as conjugation donors.

2.2.5.6 Preparation of electrocompetent cells

An overnight *E. coli* CA434 culture was diluted 1/100 into 30 ml LB broth and grown as described to OD_{600} = 0.5. Bacteria were then harvested by centrifugation (4000 rpm, 10 min), washed x3 in ice cold 10% glycerol and pelleted as above. The pellet was resuspended in 10% glycerol to an approximate density of 10^{10} CFU/ml, and stored at -80°C.

2.2.5.7 Transformation via electroporation

Electroporation was used for the transformation of *E. coli* Top10 or *E. coli* CA434 with plasmid DNA. Electrocompetent cells were aliquoted (50 µl) into pre-chilled 0.2 cm electroporation cuvettes (Equibio), to which approximately 50 ng plasmid DNA was added.

Cells were then electroporated using a Bio-Rad Gene Pulsar^{TM} (Bio-Rad) as follows: 2.5 kV, 25 µF, 600 Ωhms. Following electroporation, cells, were immediately transferred to 400 µl pre-warmed recovery media (SOC; Invitrogen), and incubated with agitation at 37^{0}C for 2 h. Cells were then incubated overnight on LB supplemented with chloramphenicol (20 µg/ml) or kanamycin (50 µg/ml) as required, after which plasmid DNA was extracted, digested as described above and examined by gel electrophoresis.

2.2.5.8 Conjugation

Donor (*E. coli* CA434) and recipient (*C. difficile* R20291 or 630) strains were grown overnight with appropriate antibiotic supplementation, as described. Donor cells (1 ml) were then *gently* pelleted (4,000 rpm, 1 min), to which 200 µl of recipient cells were added and *gently* resuspended. The latter was performed in an anaerobic environment. The resulting slurry was plated as 20 µl spots onto non-selective BHI agar, and incubated for 24 – 48 h. Colonies were screened by plating onto cycloserine (250 µg/ml) and cefoxitin (8 µg/ml) to counter-select for *E. coli* CA434, and thiamphenicol (15 µg/ml) to select for the *catP* marker.

2.2.5.9 ClosTron mutagenesis

The *C. difficile* 630Δ*erm*Δ*spo0A* and *C. difficile* R20291Δ*spo0A* mutants were made by Dr L. F. Dawson (London School of Hygiene and Tropical Medicine, London) and Dr R. P. Fagan

(Imperial College, London), respectively, prior to the commencement of this study.

ClosTron technology was used to make targeted mutants in the *spo0A* gene of *C. difficile* strains R20291 (PCR ribotype 027) and 630Δ*erm* (PCR ribotype 012), as previously described (186, 190). Figures 2.2 and 2.3 describe the fundamentals of ClosTron mutagenesis. Briefly, the group II Ll.LtrB intron was retargeted to *spo0A* by splicing by overlapping extension (SOE) PCR, as previously described (186). The retargeted intron, an antisense insertion between nucleotides 178/179, was then cloned into the *HindIII* and *BsrGI* sites of pMTL007 (*C. difficile* 630Δ*erm*) or pMTL007C-E2 (*C. difficile* R20291) and transformed into the *E. coli* conjugation donor strain CA434.

C. difficile strains 630Δ*erm* and R20291 were mated with the *E. coli* donor harboring the retargeted pMTL007 or pMTL007C-E2, respectively, as described above. Transconjugants were selected for in the presence of thiamphenicol (15 μg/ml; Sigma Aldrich). For pMTL007, mobilisation of the intron to the *spo0A* gene of *C. difficile* 630Δ*erm* was induced using IPTG (1 mM) as previously described (186). pMTL007C-E2 is a version of pMTL007, in which the promoter of the group II intron is modified for a constitutive *fdx* promoter (190), and as such does not require IPTG induction to direct intron expression to the *spo0A* gene of *C. difficile* R20291.

Potential transconjugants were then serially diluted onto plates containing lincomycin (15 μg/ml; *C. difficile* R20291; Sigma Aldrich) or erythromycin (5 μg/ml; *C. difficile* 630Δ*erm*; Sigma Aldrich) to select for chromosomal intron insertion. Replica plating with lincomycin

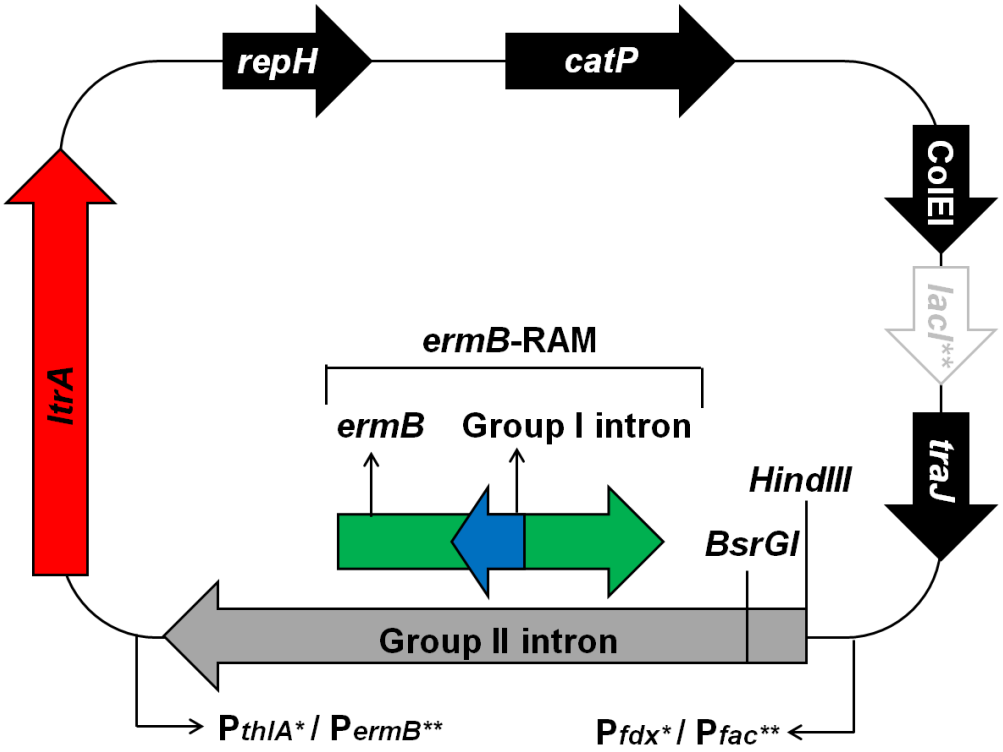


Figure 2.2. Schematic representation of ClosTron mutagenesis plamids. ClosTron plasmids share a common backbone of a group II intron (gray), within which lies a RAM based on the erythromycin resistance gene, *ermB* (green). The *ermB*-RAM is itself inactivated by a group I intron from the phage T4 *td* gene (blue), and is only activated upon its retroposition. *ltrA*, an intron encoded protein (IEP) is critical for intron mobility, RNA splicing, recognition of target site DNA and reverse transcription. *HindIII* and *BsrGI* restriction sites represent the insertion site of the retargetted group II intron. Also included are the replication gene, *repH*, the clostridial *catP* gene conferring chloramphenical/thiamphenicol resistance, the Gram-negative ColEI replicon and conjugal transfer region, *traJ*. Plasmid pMTL007 has a *fac* promoter derived from the *Clostridium pasteurianum* ferridoxin gene, which contains a *lacI* operator sequence to repress transcription from *fac*. Transcription from *fac* is induced by IPTG. Plasmid pMTL007C-E2 comprises a constitutive *fdx* promoter derived from *Clostridium sporogenes*, such that *lacI* is not required in this plamid. Intron *ermB* expression is driven by a native *ermB* promoter or a stronger *thlA* promoter derived from the *Clostridium acetobutylicum* thiolase gene, in plasmids pMTL007 and pMTL007C-E2, respectively. * Specific to plasmid pMTL007C-E2, ** Specific to plasmid pMTL007.

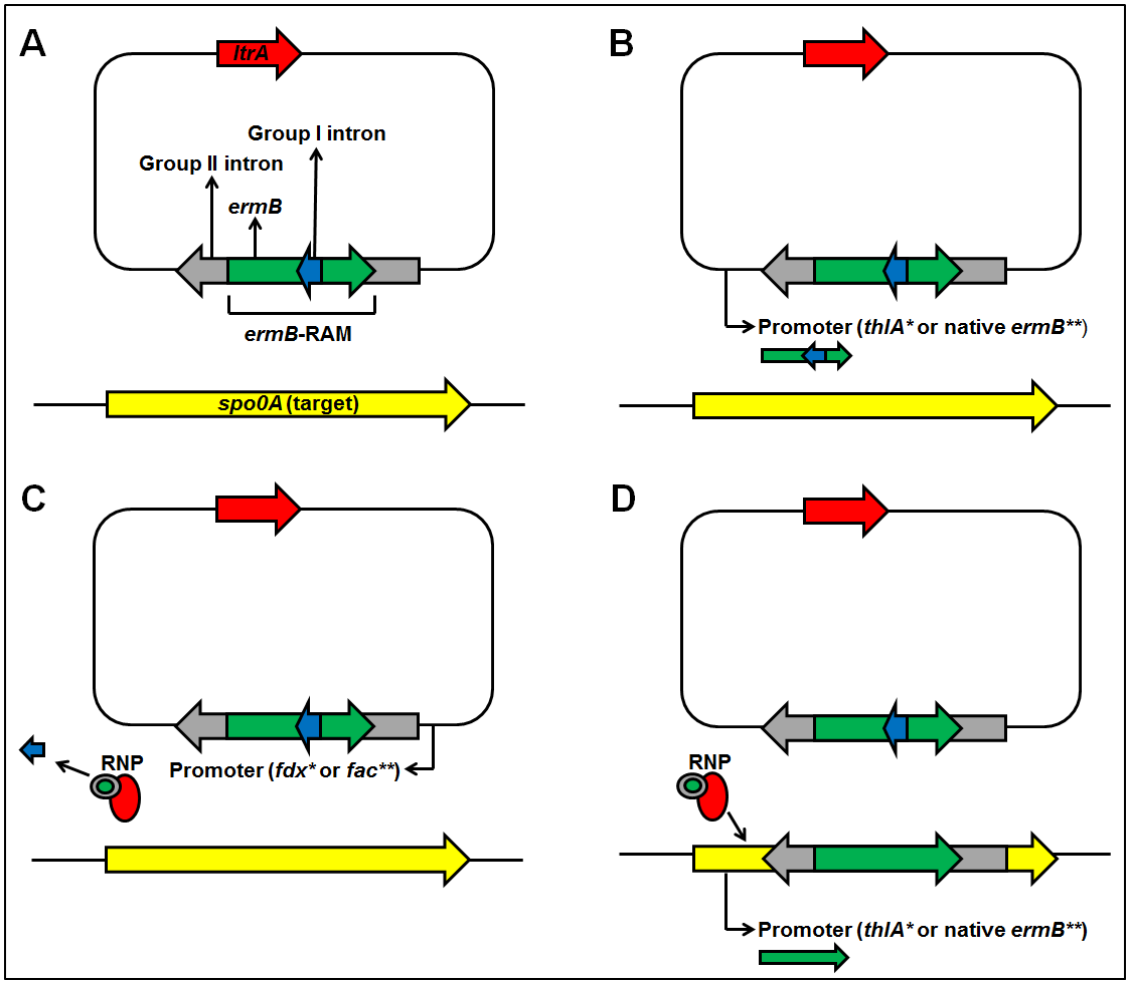


Figure 2.3. Schematic representation of the ClosTron mutagenesis system. A) ClosTron plasmids as described in Figure 2.2. Present are a group II intron (gray), an *ermB*-RAM (green), inactivated by a group I intron (blue), and *ltrA*, an intron encoded protein (IEP). **B)** Transcription of *ermB* produces a transcript containing the group I intron. However, it is in the incorrect orientation, such that self-catalytic splicing cannot occur and *ermB* remains innactive. **C)** On transcription of the group II intron (from the opposite strand), it binds LtrA producing a ribonuclear protein complex (RNP). Since the group I intron is now in the correct orientation, it is spliced out on transcription. **D)** LtrA recognises target site DNA and inserts the group II intron RNA into the host chromosome, and initiates reverse transcription for complementary DNA synthesis. As the group I intron has already spliced out, a functional *ermB* gene resides in the target gene of integrants. * Specific to plasmid pMTL007C-E2, ** Specific to plasmid pMTL007.

(15 μg/ml) or erythromycin (5 μg/ml) and thiamphenicol (15 μg/ml) was performed to select for the restored *ermB* retrotransposition-activated marker (RAM) that signals integration into the genome, and loss of the *catP* gene encoding thiamphenicol resistance carried on the plasmid.

2.2.5.9.1 Screening and verification of mutants

Mutants were screened by PCR and sequencing to confirm the chromosomal integration of the intron within the desired genes and loss of plasmids pMTL007 and pMTL007C-E2.

Briefly, three PCRs were performed to screen putative mutants using the following oligonucleotides (Table 2.2): i) RAM-F and RAM-R, to screen for loss of the group I intron, which insertionally inactivated the *ermB* RAM prior to chromosomal integration of the group II intron; ii) a gene specific primer (spo0A F) and the group II intron specific EBS universal primer, to screen for insertion of the intron into the desired location in the genome; and iii) gene specific forward and reverse primers (spo0A F and R) that flank the insertion site. Additionally, sequencing was performed across the junction of the gene to intron using gene specific primers and the EBS universal primer to verify insertion site.

2.2.5.10 Genetic complementation of the *spo0A* mutation

To complement the *spo0A* mutations, *C. difficile* - *E. coli* shuttle vector pRPF101 was constructed using the ColE1 replicon from pBlueScript II SK (+) (Agilent Technologies), *catP* from pJIR418 (198), the *C. difficile* pCD6 replicon (195) and oriT from pJB665 (199). The wild-type *C. difficile* 630 *spo0A* gene and upstream promoter region was amplified using oligonucleotide pair spo0A_A1F and spo0A_A2Rb (Table 2.2), which were designed to enable cleavage with *SalI*, and cloned into pRPF101 yielding p*spo0A*. Since the amino acid sequence of Spo0A is identical in both *C. difficile* R20291 and 630Δ*erm*, only one plasmid was constructed for the complementation of both strains.

p*spo0A* was initially transformed into *E. coli* CA434 followed by conjugation into the *spo0A* mutant derivatives as described above, generating strains R20291Δ*spo0A*+p*spo0A* and 630Δ*erm*Δ*spo0A*+p*spo0A*, respectively. Potential transconjugants were screened as above. Plasmid transfer was verified by p*spo0A* isolation and restriction analysis.

2.2.5.11 TcdA and TcdB quantification

C. difficile cultures were growth in Wilson's broth as described for 30 h, pelleted by centrifugation and the supernatant was removed for TcdA and TcdB quantification. Total and spore counts were determined to ensure equal numbers of vegetative cells in the cultures.

For TcdA quantification, microtitre plates were coated with capture antibody by adding 50 µl/well of a 2 µg/ml solution of anti-TcdA (tgcBiomics, GmbH) in PBS, and incubating overnight at 4ºC. Plates were then washed x3 in 0.05% Tween20 in PBS (PBS-T) and blocked with 200 µl 1% BSA in PBS for 2 h at room temperature. Purified TcdA (tgcBiomics) was diluted in 1% BSA-PBS (50 µl/well) and used to construct a standard curve. Culture filtrates (50 µl/well) were diluted as above in order to generate readings within the linear range of the standard curve. Plates were then incubated at room temperature for 2 h, followed by washing in PBS-T as above. The detection antibody (rabbit anti-*Clostridium difficile* TcdA; antibodies-online, GmbH) was diluted 1:5000 in 1% BSA-PBS, added to wells (50 µl/well) and incubated for 2 h at room temperature. After washing, polyclonal swine anti-rabbit IgG conjugated to horseradish peroxidase (HRP; Dako) was diluted 1:1000 in 1% BSA-PBS, added to the wells (50 µl/well) and incubated for 2 h at room temperature. Finally, plates were washed and 100 µl 3, 3′, 5, 5′-tetramethylbenzidine (TMB; Sigma-Aldrich) substrate was added for 30 min at room temperature in the dark. 50 µl 0.5 M H_2SO_4 was added to stop the reaction.

TcdB quantification was determined using a TcdB-specific ELISA kit performed according to manufacturer's instructions (tgcBiomics, GmbH). Briefly, microtitre plates pre-coated with anti-TcdB capture antibody were co-incubated with either *C. difficile* culture filtrates or purified TcdB (as described above) and anti-*Clostridium difficile* TcdB-HRP, and incubated for 1 h at 37ºC. Plates were then washed and 100 µl TMB substrate was added for 30 min at room temperature in the dark. The reaction was stopped with 50 µl 0.5 M H_2SO_4. Absorbance

of all plates was measured at 450 nm on a FLUOStar Omega (BMG Labtech). Data for TcdA and TcdB are from three independent experiments performed in triplicate.

2.2.5.12 Butyrate quantification

Butyrate quantification was performed in triplicate by Dr S. Duncan (Rowett Institute of Nutrition and Health, Aberdeen). Briefly, spent culture supernatants of exponentially growing *C. difficile* 630Δ*erm* strains were acidified, converted to *t*-butyldimethylsilyl derivatives as previously described (200) and quantified by capillary gas chromatography.

2.2.6 RNA methods

2.2.6.1 RNA stabilisation

Three biological replicates of *C. difficile* RNA was stabilised using *RNA*Protect Bacteria Reagent (Qiagen) according to manufacturer's instructions. Briefly,the *C. difficile* culture (~10^{10} cells) from exponentially growing cells was diluted 1/3 in pre-reduced *RNA*Protect and incubated under anaerobic conditions for 30 min. Bacteria were then harvested via centrifugation (4,500 rpm, 30 min) after which the supernatant was removed and the pellet was dried fully before storage at -80°C.

2.2.6.2 RNA extraction

RNA was isolated in a 2-step process. Initially, stabilised pellets underwent chemical and mechanical lysis using a FastRNA Pro Blue Kit (MP biomedicals) and FastPrep ribolyser, according to the manufacturer's recommendations. RNA quantification and intergrity was determined using both a ND-1000 (NanoDrop Technologies) and 2100 Bioanalyser (Agilent). The product of the first step was then transferred to a SV RNA Isolation Purification Kit (Promega), and RNA was isolated from the spin column assembly (washing) step, according to the manufacturer's instructions. RNA was quantified as described above.

2.2.6.3 DNA removal

DNA was removed from RNA samples using TURBO DNase (Ambion). Briefly, 2 units of DNase were used per 10 µg RNA in a final volume of 50 µl. This was incubated at 37°C for 30 min, after which a further 1 unit of DNase per 10 µg RNA was added to the reaction, followed by a second 37°C incubation for 30 min. DNAse was inactivated according to the manufacturer's recommendations. Samples were precipitated overnight at -80°C in 2.5 volumes of 100% ethanol, harvested by centrifugation (13,000 rpm, 1 h at 4°C), washed in ice cold 70% ethanol, harvested as before, resuspended in nuclease-free water (Ambion) and stored at -80°C. Samples were screened for the presence of DNA using primer pairs CD1498, CD1455, CD0011 and CDadk.

2.2.6.4 Reverse transcription of isolated RNA

RNA was reverse transcribed to complementary DNA (cDNA) as follows. All reagents were purchased from Invitrogen unless otherwise stated. 20 µg RNA was incubated with 3 µg random hexamers and RNaseOUT ribonuclease inhibitor in a total volume of 16.4 µl, at 70°C for 10 min and then cooled on ice. For cDNA synthesis, 6 µl First Strand buffer, 0.6 µl dNTP mix (25 mM each dATP, dCTP, dGTP, dATP), 0.4 µl actinomycin D (1.2 mg/ml), 3 µl DTT (0.1 M) and 2 µl Superscript III were added to a total volume of 33 µl. Samples were then incubated for 2 h at 42°C, following which RNA was hydrolysed with 1.5 µl NaOH (1 M) for 20 min at 70°C. Finally, samples were neutralised with 1.5 µl HCL (1M) and cDNA was purified using a G50-Sephadex column (Sigma-Aldrich), according to manufacturer's instructions. Samples were screened for the presence of cDNA using primer pairs MM_1498, MM_1455, MM_0011 and MM_adk.

2.2.7 DNA sequencing

2.2.7.1 Library construction and Illumina HiSeq cDNA sequencing

Samples for Illumina sequencing were submitted to the project co-ordinator, Dr David Harris (WTSI, Cambridge). Libraries were prepared, sequenced and passed through a quality-control pipeline under the direction of Dr Michael Quail (WTSI, Cambridge). All enzymes were

purchased from NEB and were used according to manufacturer's instructions except were stated.

Briefly, libraries were constructed by shearing the purified cDNA using a Covaris LE220 focused ultrasonicator to give fragments in the range of 150-250 bp. This was followed by an end-repair incubation with T4 DNA polymerase, Klenow polymerase and T4 polynucleotide kinase (to phosphorylate blunt-ended fragments) for 30 min at 20°C. cDNA samples were then 3' adenosine-tailed via the addition of Klenow exo- and dATP for 30 min at 37°C to reduce concatamerisation. Illumina adaptors (containing complementary sites to oligonucleotide anchors on the flow cell surface and primer sites for sequencing) were then ligated onto the cDNA repaired ends, and ligated fragments were electrophoretically separated from any unligated adapters based on size-selection. Fragments were then isolated via gel extraction. Libraries were amplified via PCR (18 cycles), quantified and denatured with 2 M NaOH to generate single stranded cDNA for sequencing. Samples were then loaded onto an Illumina flow cell to which the samples hybridise to the lawn of complementary oligonucleotide primers. Flow cell primers were then extended for 75 sequencing cycles, ultimately yielding clusters of clonally amplified cDNA templates. All steps were performed according to the manufacturer's recommendations.

2.2.7.2 Transcript mapping

Transcript mapping was performed by the pathogen informatics team as part of the ssRNA-

Seq computational analysis pipeline (WTSI, Cambridge). All sequence reads were aligned to the *C. difficile* 630 genome as a reference, using BWA (201) with a quality parameter (-q) of 15. Reads that did not align to the genome were discarded. Directional coverage plots were generated with mpileup, using:

```
run-mpileup +verbose +loop 600 +maxjobs 200 +config my.conf
+mail your@email -o outdir/ 2>&1 | tee -a my.log.
```

These data could then be read into Artemis (202) using the command "Graph, Add User Plot".

2.2.7.3 ssRNA-Seq transcript normalisation

Due to variations in sequencing depths, the total number of Illumina read counts were not identical between samples. To adjust for this and to reduce the systematic technical variation, all data were normalised using the DEseq "R" package for Bioconductor, and was accessed using:

```
source("http://www.bioconductor.org/biocLite.R")
biocLite("DESeq").
```

This uses a scaling factor based on the median of the ratios between the read count per gene and the geometric mean per gene. Normalised data were additionally transformed on a variance stabilised scale to yield homoscedastic data. This was then $\log_2$ transformed with a

false discovery rate (FDR) of 10% to adjust for multiple hypothesis testing. Data were then filtered for a *P* value of ≤ 0.01.

2.2.8 Protein methods

2.2.8.1 Spo0A purification and generation of anti-Spo0A antibodies

Spo0A purification was performed by Dr W. K. Smits (Leiden University Medical Centre, Leiden). Briefly, plasmid pWKS1245 was used to generate full-length Spo0A carrying a C-terminal His_6-tag, using primer pairs oWKS-1122 and oWKS-1123a. This was then cloned into *E. coli* Rosetta(DE3)pLys (Novagen), and transformants were selected for using chloramphenicol (34 µg/ml). Protein production was induced using 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG), after which cells were harvested, washed in PBS and stored at -80°C prior to Spo0A purification.

Protein purification was performed at 4°C. Briefly, cells were disrupted in lysis buffer (2 mM phenylmethylsulfonyl fluoride [PMSF], 10mM imidazole, 5mM β-mercaptoethanol, 300 mM NaCl, 50 mM NaH_2PO_4, pH 7.9), and lysates were incubated with 50% TALON His-Tag resin slurry (Clonech) for 1 h. The resin was allowed to settle on a Poly-Prep column (BioRad), and was washed with buffer (20 mM imidazole, 300 mM NaCl, 50 mM NaH_2PO_4, pH 7.9). Spo0A was eluted, assayed for purity and yield, and was dialysed against the final buffer (50mM Tris-HCl pH8, 1mM EDTA, 0.5mM DTT). Purified Spo0A concentrations were

determined using Bradford Reagent (BioRad), according to the manufacturer's instructions. Anti-Spo0A antibodies were generated and collected as described in section 2.2.2.2.

2.2.8.2 Protein preparation and cell lysis antibodies

Three biological replicates and three technical replicates of both *C. difficile* 630Δ*erm* and *C. difficile* 630Δ*erm*Δ*spo0A* cultures were prepared for mass spectrometry. Briefly, cells (~10^{10}) from exponentially growing *C. difficile* were harvested by centrifugation, resuspended in 300 µl EBT reducing lysis buffer (8 M urea, 2 M thiourea, 4% sodium dodecyl sulphate [SDS], 20 mM tris(2-carboxyethyl)phosphine [TCEP]) and incubated at 70°C for 10 min. Cells were then mechanically disrupted using acid-washed glass beads (size 425-600 µm; Sigma Aldrich) and a FastPrep ribolyser. The lysate supernatant was collected from the glass beads, centrifuged (14,000 rpm, 30 min) and separated from the resulting DNA pellet. Finally, samples were alkylated with a final concentration of 5 mM idoacetamide (IAA; Sigma-Aldrich), and diluted in lithium dodecyl sulfate (LDS) loading buffer (Pierce).

2.2.8.3 One-dimensional gel electrophoresis of proteins

Samples with an equivalency to ~2x10^7 cells per lane were loaded onto a pre-cast 1 mm, 12% polyacrylamide gel (Invitrogen), and electrophoresed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions in MOPS running

buffer (Invitrogen). All gels were run for 1 h at 200 V. These parameters were experimentally verified as providing the optimal resolution of *C. difficile* proteins. Following separation, proteins were fixed for 30 min (40% methanol, 2% acetic acid).

2.2.8.4 Western blotting

Proteins resolved by SDS-PAGE were electrophoretically transferred to a nitrocellulose membrane run for 1 h at 30 V. Protein transfer was visualised by staining in Ponseau-S Red (Sigma Aldrich) for 2 min, and membranes were blocked in blocking buffer (5% milk powder in 0.1% PBS-T) for 1 h at room temperature. Membranes were then probed with a specific primary antibody (1/10,000) overnight at 4°C, washed, and detected with an appropriate HRP-conjugated secondary antibody (1/10,000) for 1 h at room temperature. Proteins were revealed by chemiluminescence detection according to the Amersham ECL system (GE Healthcare), as per the manufacturer's instructions.

2.2.8.5 Coomassie staining

To visualise the proteins resolved by SDS-PAGE, gels were stained overnight with Coomassie Brilliant Blue (BioRad), prepared according to manufacturer's instructions. The gel was then de-stained in 30% methanol until the background stain was reduced, but protein bands were visible. The gel was then scanned and points of excision were marked.

2.2.8.6 Protein digestion and peptide extraction

Protein bands were carefully excised, cut into 1 mm^3 pieces, placed into wells of a 96 well plate (18 wells per lane of gel), and washed repeatedly in 40% acetonitrile/50 mM triethylammonium bicarbonate (TEAB) until completely de-stained. Gel pieces were then dehydrated in 100% acetonitrile, and digested with 0.15 µg trypsin/well (sequencing grade; Roche) (203) for 2 h at 37°C followed by 25°C overnight. For peptide extraction, gel pieces were repeatedly dehydrated in 50% acetonitrile/% formic acid to extract all peptides, and the resultant trypsin digests were collected in a clean 96 well plate and concentrated using a SpeedVac until the plate was dry.

2.2.8.7 Isotopomeric dimethyl labeling

Prior to mass-spectrometry peptides were chemically tagged with isotopomeric dimethyl labels (204). Specifically, primary amines exposed by tryptic digestion were converted to dimethylamines via a reaction with different combinations of formaldehyde and cyanoborohyde isotopes, resulting in 4 Da mass shifts (Table 2.4). This allows samples to be multiplexed and facilitates quantification and direct comparison of protein expression profiles between mutant and parental *C. difficile*.

Briefly, samples were dissolved in 100 µl 100 mM TEAB, after which 2 µl of the appropriate formaldehyde isotope (4%) was added, followed by 2 µl of the corresponding

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

	Label		
	"Light"	**"Intermediate"**	**"Heavy"**
Formaldehyde isotope	CH_20 Formaldehyde	CD_20 Deuterated	$^{13}CD_20$ Deuterated and ^{13}C-labelled
Cyanoborohyride isotope	$NaBH_3CN$ Cyanoborohyride	$NaBH_3CN$ Cyanoborohyride	$NaBD_3CN$ Cyanoborodeuteride
Mass increase/label (Da)	+ 28.0313	+ 32.0564	+ 36.0757

cyanoborohyriode isotope (600 mM), see Table 2.4. The reaction was incubated for 1 h with agitation in a fume hood to label all available primary amines, quenched with 2 µl ammonium (8 % v/v) and acidified with 5 µl formic acid (5% v/v). Wells of corresponding differentially labelled samples were pooled prior to mass spectrometry, and stored at -20ºC prior to quantification. To prevent biases associated with labeling, different biological and technical replicates of mutant and parental *C. difficile* peptides were labelled with different isotopes. The differentially labelled samples were then combined and simultaneously analysed by liquid

chromatography with tandem mass spectrometry detection (LC-MS/MS), whereby the 4 Da mass shifts were used to compare relative peptide abundance.

2.2.8.8 LC-MS/MS analysis

Protein sequencing and database searching was performed by Dr Lu Yu (WTSI, Cambridge). Briefly, dimethyl labelled peptides were analysed with an on-line nanoLC-MS/MS on an Ultimate 3000 RSLCnano System (Dionex) coupled to a LTQ Orbitrap Velos (Thermo Fisher) hybrid mass spectrometer equipped with a nanospray source. The system was controlled by Xcalibur 2.1 (Thermo Fisher) and DCMSLink 2.08 (Dionex). Only 1/3 of total volume of each sample was submitted for analysis. On the RSLCnano, samples were first loaded and desalted on a PepMap C18 trap (0.3 mm id x 5 mm, 5µm, Dionex) at 10 µL/min for 15 min, then peptides were separated on a PepMap RSLC column with 2 µm particle size (Dionex) at a 75 µm id x 50 cm column over a 120 min linear gradient of 4–32% CH_3CN/0.1% FA at a flow rate at 300 nL/min.

The LTQ Orbitrap Velos was operated in the "Top 10" data-dependent acquisition mode. The 10 most abundant and multiply-charged precursor ions in the MS survey scan in the Orbitrap (m/z 400 – 1500, with the lock mass at 445.120025) were dynamically selected for CID fragmentation (MS/MS) in the LTQ Velos ion trap. The ions must have a minimal signal above 2000 counts. The preview mode of FT master scan was disabled. The Orbitrap resolution was set at 60,000 at m/z 400 with one microscans. The isolation width for the

precursor ion was set at 2 Th. The normalised collision energy was set at 35% with activation Q at 0.250 and activation time for 10 msec. The dynamic exclusion mass width was set at ±20ppm and exclusion duration for 1 min. To achieve high mass accuracy, the Automatic Gain Control (AGC) were set at $1x10^6$ for the full MS survey in the Orbitrap with a maximum injection time at 100 msec, and 5000 for the MS/MS in the LTQ Velos with a maximum injection time at 300 msec.

The raw files were analyzed with MaxQuant Software (v. 1.2.2.5; http://maxquant.org) for protein quantification (205). The Andromeda search engine was used to search the MS/MS spectra using the following parameters: trypsin/P with maximum 2 missed cleavages sites; peptide mass tolerance at first search was set at 20 ppm; MS/MS fragment mass tolerance at 0.49 Da, and top 6 MS/MS peaks per 100 Da and a minimum peptide length of 6 amino acids were required. The mass accuracy of the precursor ions was improved by the time-dependent recalibration algorithm of MaxQuant. Fixed modification for carbamidomethyl and variable modifications for deamidated (NQ) and oxidation (M) were used, and a maximum of three labelled amino acids per peptide were allowed. The protein databases were extracted from annotated genome databases of *C. difficile* 630 *(November 2008)* and the contaminate database were supplemented by MaxQuant.

False discovery rates (FDR) were estimated based on matches to reversed sequences in the concatenated target-decoy database. The maximum FDR at 1% was allowed for proteins and peptides. Peptides were assigned to protein groups, a cluster of a leading protein(s) plus additional proteins matching to a subset of the same peptides. Protein groups with posterior

error probability (PEP) value over 0.01 or containing matches to reversed database or contaminants were discarded. Protein ratios of different labelled status were calculated after summing the corresponding peptide ions' intensity for a specific protein. A minimum of two peptides per protein was required.

2.3 Statistical analysis

Statistical significance of results was determined as follows; CI ratios were analysed using a non-parametric two-tailed Mann-Whitney test. In all other cases, a two-tailed Student's *T*-test was used. In all cases, a *P* value of < 0.05 was considered to be significant. Tests were performed using GraphPad Prism 5 graphing and statistical software (GraphPad Software, Inc.)

2.3.1 R

R was used for the differential analysis of transcriptome data. This is a freely available statistical programming language (http://www.r-project.org/).

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

3.1 Publications arising from this chapter

The key findings from this chapter have resulted in the following publication (see Appendix 1):

Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, Wren BW, Fairweather NF, Dougan G, Lawley TD. (2012). The *Clostridium difficile spo0A* gene is a persistence and transmission factor. *Infect Immun* **80**(8): 2704-2711.

3.2 Introduction

3.2.1 Human-virulent *C. difficile*

During the past decade, distinct genetic clades of *C. difficile* have emerged that are responsible for epidemics within North America and mainland Europe and continue to disseminate globally (see Chapter 1.1.2.2). Amongst the most notable of these clades is the so called "hypervirulent" variant of *C. difficile*, commonly genotyped as PCR ribotype 027, which is associated with high rates of mortality and disease recurrence as well as severe hospital outbreaks (7-9, 206). In addition, other genetic variants such as *C. difficile* PCR ribotype 012 continue to be endemic in many healthcare systems (207).

3.2.2 *C. difficile* transmission

C. difficile produces highly infective endospores that are excreted by infected patients, allowing the oxygen sensitive pathogen to retain viability outside of the host (150). Spores of *C. difficile* can also resist commonly used disinfectant regimens and as a result are able to persist in the environment generating a potential transmission reservoir that confounds standard infection control measures (61, 151, 152).

The ability of a bacterium to transmit within an ecological niche is essential for its continued survival within the population. Spores are often cited as the primary transmission factor for *C. difficile*, and it has indeed been demonstrated that spores of the pathogen are highly infective and resistant to commonly used disinfectant regimens (61). The incidence of *C. difficile* increases within the hospital environment, whereby shedding of *C. difficile* by infected patients facilitates the spread of infection. Spores have been isolated from environmental surfaces as well as fomites including medical equipment and staff clothing (59, 208).

3.2.3 Antibiotic treatment of *C. difficile* infection: disease recurrence

First line antibiotic treatments for *C. difficile* disease are vancomycin or metronidazole (117), however recurrent disease (relapse with the same strain or reinfection with a different strain) occurs in ~15 to 35% of patients (60). In this sense, recurrence and transmission are relatively

poorly understood aspects of *C. difficile* biology that likely underpin its persistence within healthcare systems.

In the backdrop of endemic and epidemic CDAI, there is clearly a public health imperative to understand the transmission dynamics of the pathogen and a need to reduce the incidence of recurrent episodes that have now become a hallmark feature of *C. difficile* infection.

3.2.4 Animal models of *C. difficile* infection

A number of animal species have been used to model *C. difficile* infection, including piglets (209), guinea pigs (210), rats (111), rabbits (211), hamsters (143, 212, 213) and mice (98, 214). The golden Syrian hamster model of infection is one of the most widely used and has been employed for several decades to study CDAI. However, these animals are exquisitely susceptible to infection and can rapidly develop fulminant disease with severe enterocolitis, which is often lethal (143). Although great efforts have been taken to refine this model (213), it can be argued that it does not capture the spectrum of *C. difficile* disease observed in humans (215). In addition, due to the severe nature of the disease in hamsters, it is inherently more challenging to study some aspects of *C. difficile* disease, such as long-term persistence and recurrence in this model.

3.4.2.1 Murine model of *C. difficile* disease

In contrast to hamsters, mice are a potentially more accessible model through which to study infectious diseases. Molecular tools and reagents are more readily available for mice and there are thousands of genetically modified lines available for investigations (http://www.sanger.ac.uk/mouseportal/). Against this backdrop, the Wellcome Trust Sanger Institute has developed a murine model of *C. difficile* disease that has proved to be a valuable tool in studying aspects of *C. difficile* disease such as pathogenesis, transmission and virulence.

3.3 Aims of the work described in this chapter

In this chapter, I use a murine model of *C. difficile* infection to investigate the role of the *spo0A* gene in distinct PCR ribotypes 027 and 012 during infection and transmission.

3.4 Results

3.4.1 *C. difficile spo0A* is essential for spore formation

The WTSI and collaborators have previously sequenced and fully annotated the complete genomes of individual isolates of *C. difficile* ribotypes 027 (strain R20291) (11) and 012 (strain 630) (10). *C. difficile* R20291 and 630 both produce ~10^5 spores/ml after 48 h of growth in Wilson's broth, representing < 1% of the total bacterial cells (Figure 3.1A). To

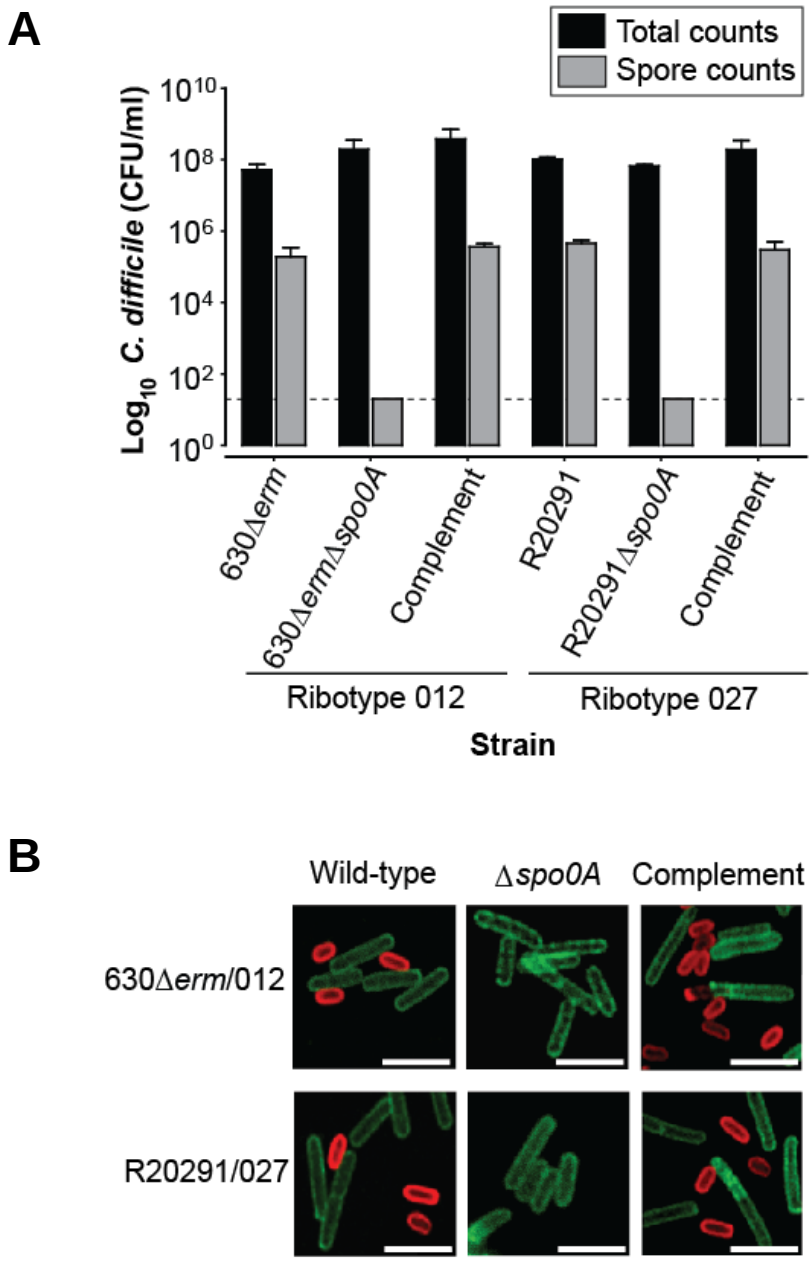


Figure 3.1. Genetic and phenotypic characterisation of *C. difficile* 630Δ*erm* and R20291 *spo0A* mutants. A) *C. difficile* cultures were grown in Wilson's broth for 48 h under anaerobic conditions and then total cell and spore counts were determined. Genetic complementation of the *spo0A* mutation restored spore formation to levels statistically comparable to the parental strains (P = > 0.05), according to the Student's t test. The dashed horizontal line indicates the detection limit. **B)** Representative indirect immunofluorescence images of *C. difficile* cultures stained with vegetative cell specific (green) and spore specific (red) polyclonal antibodies, and visualized with FITC-conjugated and Cy3-conjugated secondary antibodies, respectively. Scale bar = 5 μm.

evaluate the impact of *spo0A* on growth and sporulation, defined insertional mutations were generated in the *spo0A* genes of either *C. difficile* R20291 or *C. difficile* 630Δ*erm*, yielding mutant derivatives R20291Δ*spo0A* and 630Δ*erm*Δ*spo0A*, respectively, as described in Materials and methods chapter 2.2.5.9.

The vegetative cells of both R20291Δ*spo0A* and 630Δ*erm*Δ*spo0A* were morphologically indistinguishable from the respective parent strains when examined by light microscopy, and displayed similar general growth kinetics. However, both *C. difficile spo0A* mutant derivatives failed to produce any detectable spores or spore-like elements (177, 216), as determined by anaerobic culturing after ethanol shock (spores are resistant to ethanol) and indirect immunofluorescence using spore-specific anti-sera (Figure 3.1A and B).

To demonstrate that the asporogenous phenotype was directly attributable to the *spo0A* mutation, we complemented the mutated genes using the plasmid p*spo0A* harbouring the wild-type *C. difficile spo0A* gene under the control of the native promoter. We found that the complemented *spo0A* mutants produced spores at levels comparable to the parent ($P = > 0.05$) and that these spores were also morphologically indistinguishable (Figure 3.1A and B). Collectively, these observations indicate that the *spo0A* mutation is non-polar for spore formation and that expression of the *spo0A* gene is critical for vegetative cells to differentiate into spores in both *C. difficile* R20291 and 630Δ*erm*.

3.4.2 Antibiotic-induced murine model of *C. difficile* disease

CDAIs are precipitated primarily by the administration of antibiotics for therapy. As such, we pre-treated mice with clindamycin (a lincosomide antibiotic known to incite *C. difficile* disease in humans) to disrupt the intestinal microbiota. Groups of mice were subsequently independently orally infected with 10^7 CFU of either *C. difficile* strains R20291 or 630Δ*erm*, and colonisation was monitored by assessing *C. difficile* viable counts recovered from faeces (Figure 3.2).

Mice orally infected with *C. difficile* R20291 shed approximately 10^8 CFU/g fresh faeces and this level remained consistently high for the entire monitoring period (Figure 3.2). Thus, *C. difficile* R20291 infection can result in a long-term, persistent colonisation in this model. This is termed the "supershedder" state, and has been reported in *C. difficile* ribotype 017 (strain M68) (98) .

In contrast, *C. difficile* 630Δ*erm* did not establish a long-term persistent infection in this model. Following the initial infection, faecal shedding of *C. difficile* 630Δ*erm* remained high and comparable with R20291-infected mice for approximately 14 days. However, by day 15 post-infection the levels of *C. difficile* in the faeces had begun to decline, and by day 30 post-infection all mice had resolved the infection (Figure 3.2). Upon clearance, *C. difficile* 630Δ*erm* was undetectable in the faeces for the duration of the monitoring period.

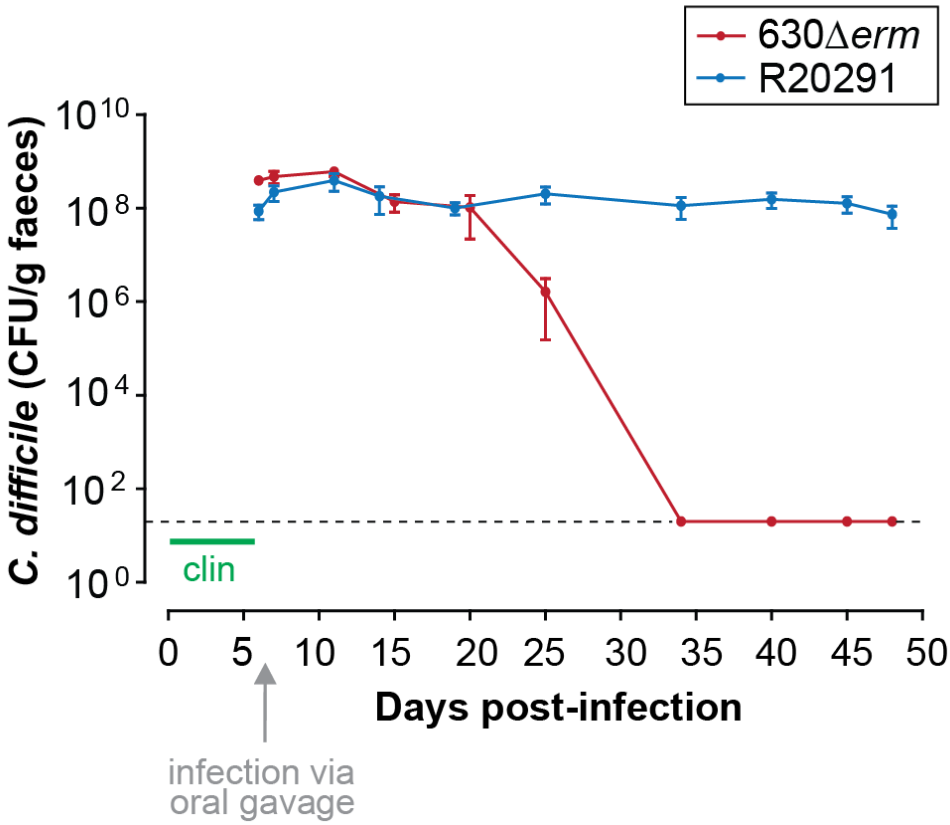


Figure 3.2. *C. difficile* R20291 can cause a long-term, persistent infection in mice. Representative faecal shedding profiles from mice (*n* = 5 per group) infected with *C. difficile* 630Δ*erm* or R20291. Mice were pre-treated with clindamycin (represented as a green line) for 7 d prior to infection via oral gavage. Following infection, colonisation was monitored by assessing viable counts recovered from faeces. The dashed horizontal line indicates the detection limit.

3.4.3 Role of *spo0A* in *C. difficile* pathogenesis

Spo0A is known to regulate the expression of many genes, including known virulence factors in other spore-forming pathogens such as *B. thuringiensis* (217) and *B. anthracis* (218).

However, the role of Spo0A in *C. difficile* disease, persistence and transmission had not yet been explored. A greater understanding of these aspects of *C. difficile* infection may help guide infection control measures and such information could have practical implications related to the management of hospital patients.

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

We previously described a *C. difficile* infection model in mice that mimics several aspects of *C. difficile* disease, persistence and transmission in humans (61, 98, 150). In order to investigate the role of the *spo0A* gene during the acute phase of *C. difficile* infection (day 2-4 post infection), we pre-treated mice with clindamycin and subsequently infected them with *C. difficile* R20291, *C. difficile* 630Δ*erm* or their equivalent *spo0A* mutant derivatives. Groups of mice infected orally with *C. difficile* 630Δ*erm* or 630Δ*erm*Δ*spo0A* did not exhibit any overt signs of disease and had a 100% survival rate (Figure 3.3).

In contrast, mice infected with *C. difficile* R20291 demonstrated notable signs of disease, such as hunched posture and ruffled fur; however these clinical symptoms resolved by day 5 post-infection and all mice survived. Mice infected with *C. difficile* R20291Δ*spo0A* also displayed signs of overt disease, but this was more exaggerated compared to R20291-infected mice with mice displaying hunched posturing, piloerection, lethargy, dehydration and emaciation. Within 5 days post-infection, 80% of R20291Δ*spo0A*-infected mice were considered to be

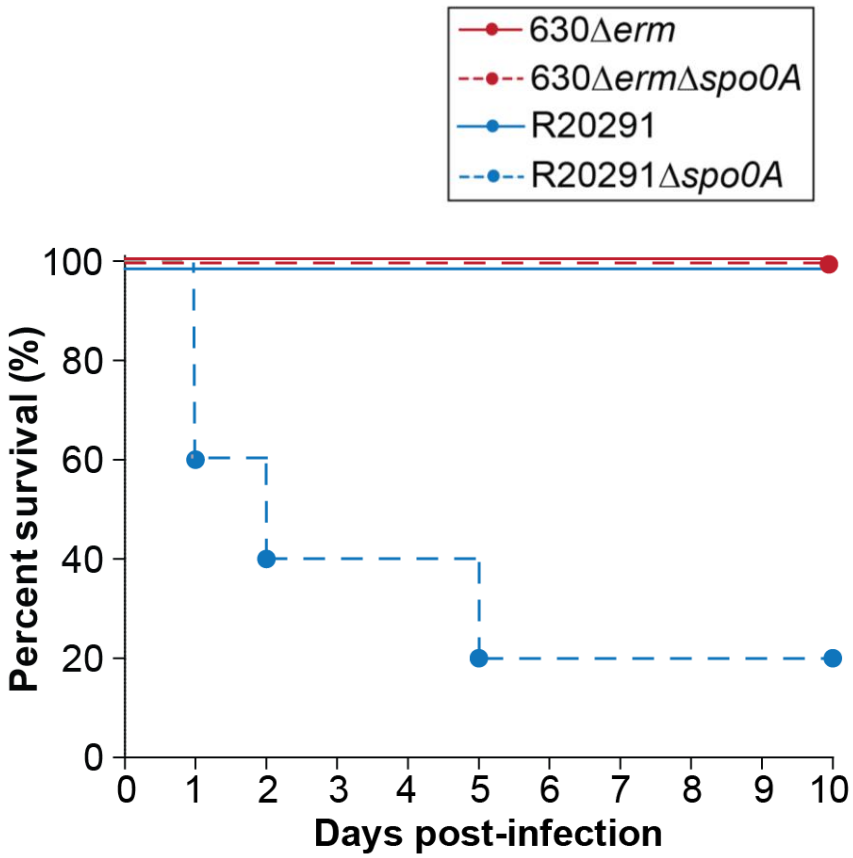


Figure 3.3. *C. difficile* R20291 *spo0A* mutants demonstrate increased virulence in mice. Kaplan-Meier survival curve of mice infected with *C. difficile* 630Δ*erm*, R20291, 630Δ*erm*Δ*spo0A* or R20291Δ*spo0A*. By day 5 post-infection, 80% of mice orally infected with *C. difficile* R20291Δ*spo0A* demonstrated severe signs of disease, including a hunched posture and lethargy and succumbed to infection.

moribund according to a clinical scoring system and were consequently sacrificed (Figure 3.3). The remaining 20% resolved disease and survived.

3.4.3.2 *C. difficile* Spo0A mediates pathological responses in mice

In order to identify any increased pathology associated with *C. difficile spo0A* mutant infections, we performed histologic analyses on the caeca of mice infected with *C. difficile* R20291, 630Δ*erm* or their respective *spo0A* mutant derivatives at day 2-4 post-infection. At these time-points, mice shed equivalent levels (~10^8 CFU/gram fresh faeces) of vegetative *C. difficile* R20291, R20291Δ*spo0A*, 630Δ*erm* or 630Δ*erm*Δ*spo0A*. Mice infected with *C. difficile* R20291 or 630Δ*erm* also shed spores (~10^6 CFU/gram fresh faeces) in contrast to R20291Δ*spo0A*- and 630Δ*erm*Δ*spo0A*-infected mice which did not, at least at a detectable level.

3.4.3.2.1 Pathological response to C. difficile R20291 infection

Uninfected, clindamycin control mice did not exhibit any pathology. Conversely, at day 2 post-infection mucosal damage was evident in R20291 and R20291Δ*spo0A*-infected mice, which included oedema and immune cell infiltrate within the caecal mucosa (Figure 3.4). These observations however were much more notable in mice infected with R20291Δ*spo0A*, which exhibited more pronounced oedema, epithelial surface damage and acute infiltration (Figure 3.4). Putative pseudomembranes were also visible on the epithelial surface of R20291Δ*spo0A*-infected mice.

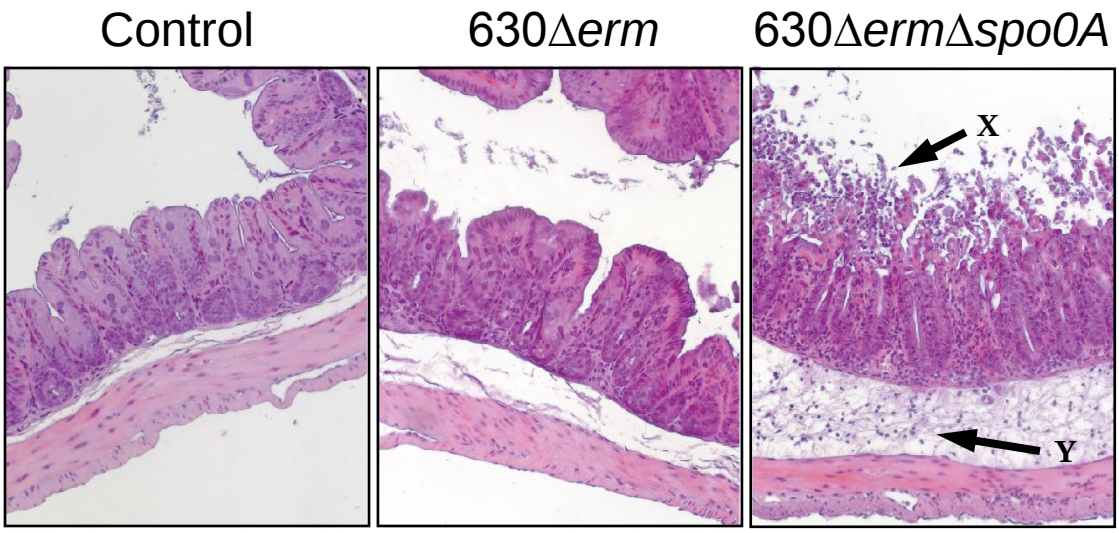


Figure 3.4. ***C. difficile*** **R20291** ***spo0A*** **mutants cause increased mucosal damage in mice.** Representative images demonstrating epithelial cell damage (arrow X) and acute infiltration arrow Y) in hematoxylin and eosin-stained caecum sections of mice infected with *C. difficile* R20291 or R20291Δ*spo0A* at day 2 post-infection. Magnification, x20.

Given the increased pathology associated with R20291Δ*spo0A*-infection, we decided to exploit electron and immunoelectron microscopy to both verify the observed histopathology and to provide a more detailed analysis of the infection. Using Transmission Electron Microscopy (TEM), we found that, again, control mice did not demonstrate any obvious pathology or signs of infection (Figure 3.5; top panel). R20291-infected mice, however exhibited some damage of the mucosal lining accompanied with immune cell infiltrate (Figure 3.5; top panel). Numerous *C. difficile* cells were evident, however these were largely contained to the lumen and no obvious invasion was noted.

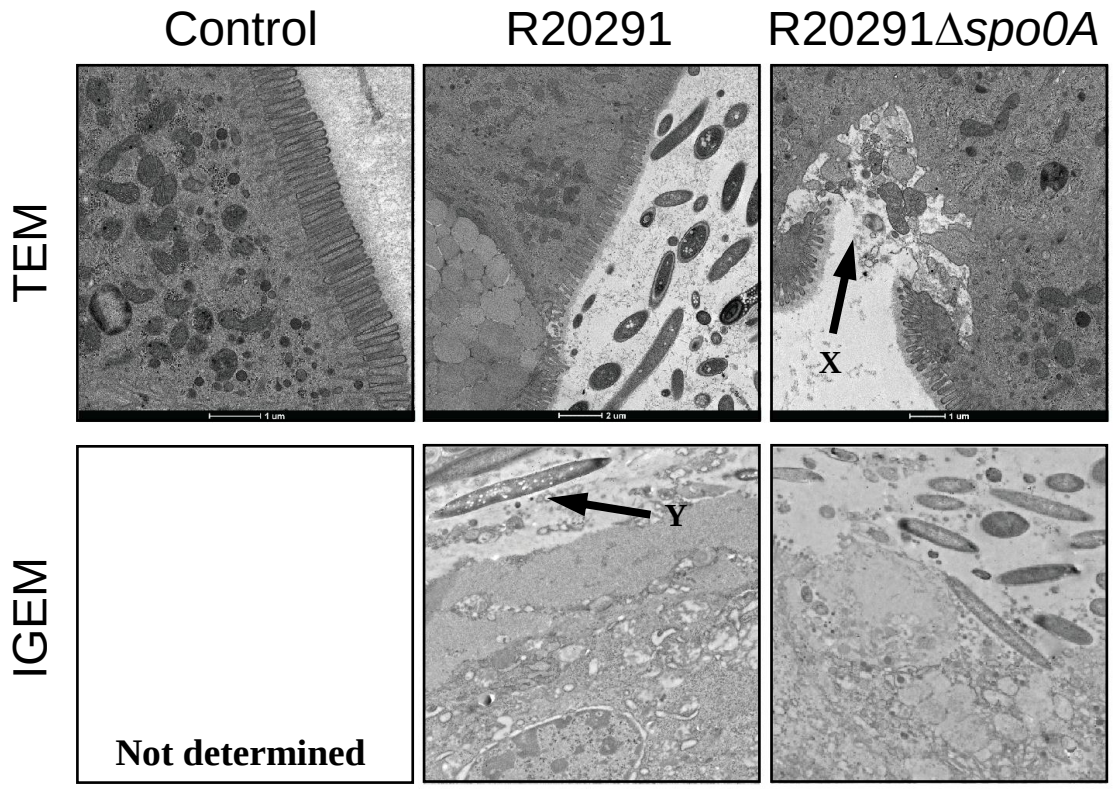


Figure 3.5. *C. difficile* R20291 Spo0A mediates pathological responses in mice. Top panel (TEM): transmission electron microscopy images, and bottom panel (IGEM): immune-gold electron microscopy images of mice infected with *C. difficile* R20291 or R20291Δ*spo0A* at day 2 post-infection. Post-infection damage was most notable in R20291Δ*spo0A*-infected mice, which was associated with microvillus effacement, endothelial cell necrosis and superficial invasion (arrow X). A greater ratio of *C. difficile* to non-*C. difficile* was also evident in the lumen of mice infected with *C. difficile* R20291Δ*spo0A*. Arrow Y indicates an immune-gold-labelled *C. difficile* R20291 vegetative cell.

The post-infection damage was greatest in R20291Δ*spo0A*-infected mice, which demonstrated acute mucosal damage with microvillus effacement and denudation (Figure 3.5; top panel). Endothelial cell necrosis and acute infiltration was also prominent. As with R20291-infected mice, numerous *C. difficile* cells were also visible, however this also appeared to be associated

with superficial invasion in mice infected with R20291Δ*spo0A* (Figure 3.5; top panel).

Using Immuno-Gold Electron Microscopy (IGEM), we were also able to estimate the number of *C. difficile* present, and thus also the ratios of *C. difficile* to non-*C. difficile* cells in local sections of caecum tissue (Figure 3.5; bottom panel). We found that the total numbers of *C. difficile* per unit area were approximately the same for R20291- and R20291Δ*spo0A*-infected mice. However, we noted that the ratios of *C. difficile* to non-*C. difficile* in the lumen were different, with a greater ratio of *C. difficile* present in mice infected with R20291Δ*spo0A*.

3.4.3.2.2 Pathological response to C. difficile 630Δerm infection

In contrast to *C. difficile* R20291, *C. difficile* 630Δ*erm* infection incited a much more modest pathological response. As before, uninfected, clindamycin control mice did not exhibit any pathology. At day 4 post-infection, some mucosal damage was apparent in mice infected with *C. difficile* 630Δ*erm*, such as oedema and some infiltration (Figure 3.6). These observations were more apparent in 630Δ*erm*Δ*spo0A*-infected mice, which demonstrated greater oedema and more pronounced immune cell infiltrate within the caecal mucosa compared to 630Δ*erm*-infected mice (Figure 3.6). Nonetheless, this was to a much more modest degree than was noted for ribotype 027 (R20291)-infected mice. Thus, in addition to not causing a long-term persistent infection, *C. difficile* 630Δ*erm* is also less pathogenic in our murine model of infection.

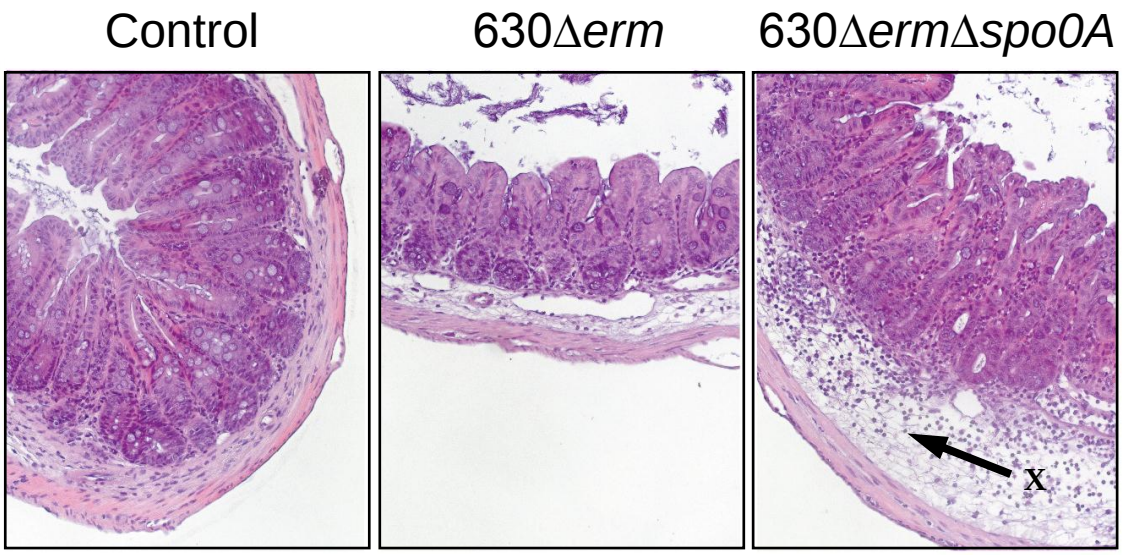


Figure 3.6. ***C. difficile* 630Δ*erm*Δ*spo0A* mutants cause increased mucosal damage in mice.** Representative images demonstrating epithelial cell damage in hematoxylin and eosin-stained caecum sections of mice infected with *C. difficile* 630Δ*erm* or 630Δ*erm*Δ*spo0A* at day 4 post-infection. Magnification, x20. Arrow X represents submucosal oedema and immune cell infiltrate, including the recruitment of polymorphonuclear neutrophils (PMNs).

3.4.3.3 Association between Spo0A and toxin synthesis

The association between spore formation and toxin production has been described previously, although the correlation is contentious. Reports describing positive regulatory associations between the two phenomena are opposed by those describing negative or minimal regulatory links. For example, in a recent publication, Underwood *et al*. (2009) outlined that *spo0A* deficient *C. difficile* produce less TcdA (177). On the other hand, a more recent paper by Saujet *et al*. (2011) demonstrated that SigH (an alternative sigma factor implicated in

sporulation initiation) markedly reduced *spo0A* transcription and concomitantly increased TcdA synthesis (182).

3.4.3.3.1 C. difficile R20291produces more TcdA and TcdB than 630Δerm strains

Using highly sensitive anti-TcdA and anti-TcdB specific ELISA assays, we demonstrate that *C. difficile* 630Δ*erm* produces less toxin than *C. difficile* R20291 under the growth conditions investigated, as shown in Figure 3.7A and B. This characteristic may theoretically be a contributory factor to R20291 "hypervirulence" and could explain, at least in part, why *C. difficile* 630Δ*erm* is less virulent in our model of infection. The difference in toxin production levels between 027 (R20291) and 012 (630Δ*erm*) ribotypes is consistent with previous observations (219). Additionally, purified TcdB from R20291 has previously been shown to have greater cytotoxic potency than 630 (11).

3.4.3.3.2 spo0A is a negative regulator of toxin production in C. difficile R20291

The increased virulence exhibited by *C. difficile* R20291Δ*spo0A* in our mouse model was not anticipated. We hypothesised that these observations may be related to increased toxin production. We therefore compared the levels of the major toxins TcdA and TcdB from broth grown cultures of wild-type *C. difficile* R20291 and R20291Δ*spo0A*. Low-levels of TcdA and TcdB were detected during mid-exponential growth in cultures of both R20291 and

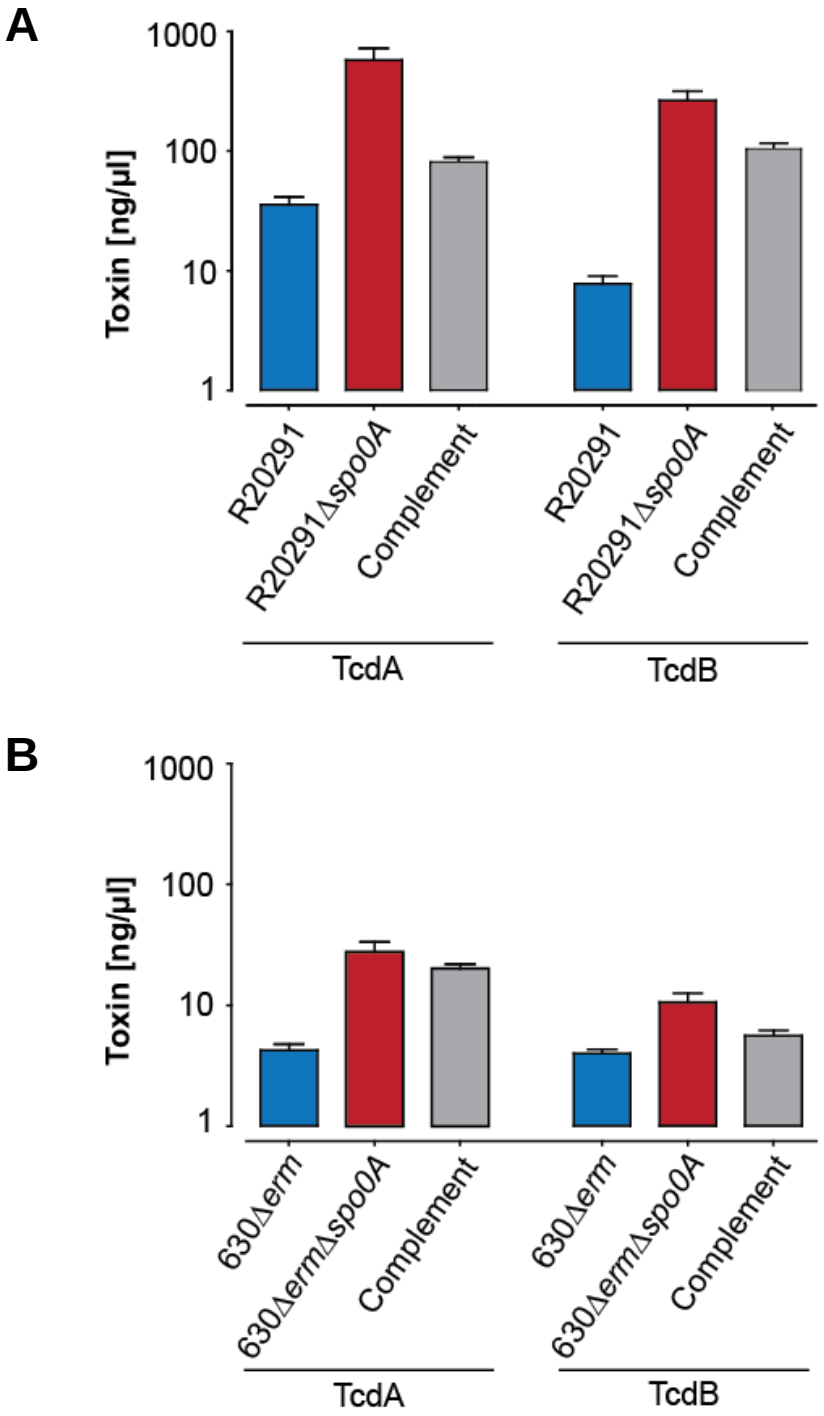


Figure 3.7. *C. difficile spo0A* mutants produce elevated levels of TcdA and TcdB. Sandwich ELISA indicating the relative levels of TcdA and TcdB produced by **A)** *C. difficile* R20291 derivatives and **B)** *C. difficile* 630Δ*erm* derivatives after 30 h of growth in Wilson's broth under anaerobic conditions. Data are from 3 independent experiments performed in triplicate. The error bars indicate standard deviations.

R20291Δ*spo0A* (data not shown). However, during stationary growth (30 h), significantly higher levels of both TcdA (P = 0.0006) and TcdB (P = 0.0005) were detected in *C. difficile* R20291Δ*spo0A* compared to R20291 cultures (Figure 3.7A). Complementation of the *spo0A* gene in R20291Δ*spo0A* reduced the levels of toxins produced, although both TcdA and TcdB were still slightly elevated compared to that produced by wild-type *C. difficile* R20291 (P = < 0.05).

Therefore, we demonstrate that R20291 and R20291Δ*spo0A* are virulent in mice and that exaggerated virulence by R20291Δ*spo0A* during the acute phase of infection is at least associated with increased TcdA and TcdB production. The increased toxin production by *spo0A* mutant derivatives may also be linked to the reduced bacterial diversity observed via IGEM in R20291Δ*spo0A*-infected mice, though this hypothesis clearly needs to be explored further.

3.4.3.3.3 Role of C. difficile 630Δerm Spo0A in toxin production

Low-levels of TcdA and TcdB were detected during mid-exponential growth by both 630Δ*erm* and 630Δ*erm* Δ*spo0A* (data not shown). During stationary growth (30 h), TcdA and TcdB were detected in all *C. difficile* 630Δ*erm* derivatives (Figure 3.7B). The levels of TcdA and TcdB were significantly higher in cultures of the *spo0A* mutant compared to those of the *C. difficile* parental strain (both P = < 0.05). The levels of TcdA and TcdB were both reduced upon genetic complementation with the *spo0A* gene, although this was not to the levels of

C. difficile 630Δ*erm* (P = < 0.05; Figure 3.7B).

We thus demonstrate that although *C. difficile* 630Δ*erm* produces less TcdA and TcdB than R20291 strains, the *spo0A* gene may nonetheless negatively regulate toxin production in this ribotype.

3.4.3.3.4 Attempt to quantify toxin from faecal samples

We have quantified both TcdA and TcdB levels *in vitro* as this is the most likely explanation for the exaggerated virulence phenotype in R20291Δ*spo0A*-infected mice. Additionally, we made multiple attempts at measuring the levels of TcdA and TcdB directly from the faeces of *C. difficile* infected mice, but have had no success.

It is important to note that precise and specific quantification of *C. difficile* toxin levels from faeces has never been reported, likely due to technical issues with faecal matter interfering with antibody-toxin interactions. Although these data would clearly be interesting, we do not believe that they would alter our conclusions about the role of *C. difficile spo0A* in virulence.

3.4.4 Role of *spo0A* in persistence and recurrence

Recurrent infection after cessation of antibiotic therapy is now a salient feature of *C. difficile*

disease and recurring episodes of diarrheoa are rapidly becoming the norm rather than the exception in elderly patients. As discussed in Chapter 1, recurrent disease can be the result of either (i) relapse with same strain that caused the initial infection, or (ii) re-infection with a different strain of the pathogen (122, 123). In this sense, recurrent disease is often considered to be a form of persistent infection, representing a constant background burden for the host, rather than a fortuitous sequelae of infection.

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

We previously demonstrated that *C. difficile* isolated from infected hospital patients can asymptomatically persist in mice for several months (98). Therefore, we next assessed the contribution of the *C. difficile spo0A* gene to colonisation and persistent infection. To assess this phenotype, competitive index (CI) experiments were performed, whereby groups of mice were infected with a mixture of equal numbers of parental strain *C. difficile* and the respective Δ*spo0A* mutant, enabling direct fitness comparisons *in vivo*.

In the R20291 background, the Δ*spo0A* mutant was recovered at a significantly reduced level compared to wild-type R20291 (Figure 3.8A). We observed that genetic disruption of *spo0A* resulted in reduced fitness by day 2 post-infection and continued to day 15 post-infection (Figure 3.9A). Additionally, the difference in relative fitness for the 630Δ*erm*Δ*spo0A* mutant was statistically significant by day 5 post-infection (Figure 3.8B). Although *C. difficile* 630Δ*erm* does not cause a long-term persistent infection in our murine model, we found that

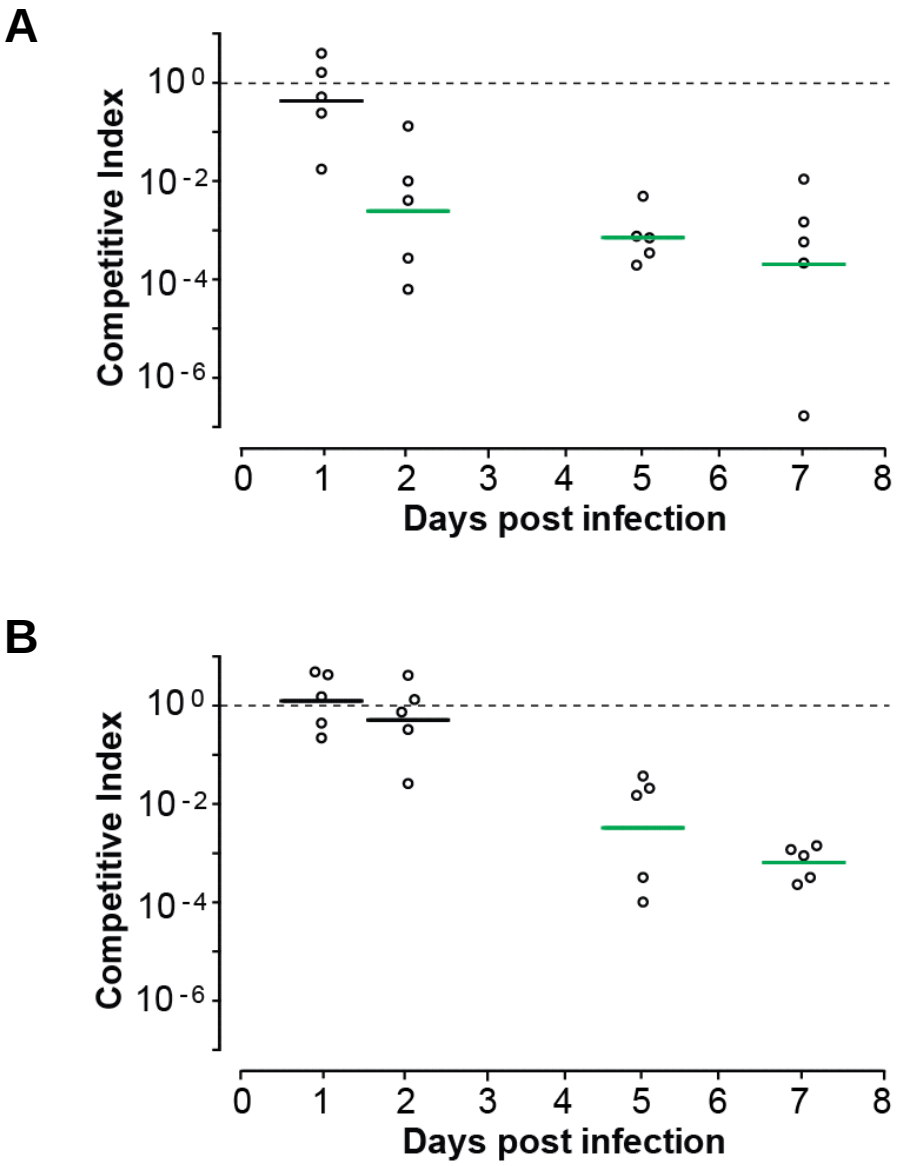


Figure 3.8. *C. difficile* Spo0A mediates intestinal persistence in mice. Competitive index (CI) time-course of **A)** *C. difficile* R20291 and R20291Δ*spo0A*, and **B)** 630Δ*erm* and 630Δ*erm*Δ*spo0A*. Individual CI values are illustrated as open circles, and horizontal bars represent the geometric means. Geometric means in green indicate that the *spo0A* mutant is attenuated, as determined by the Mann-Whitney test.

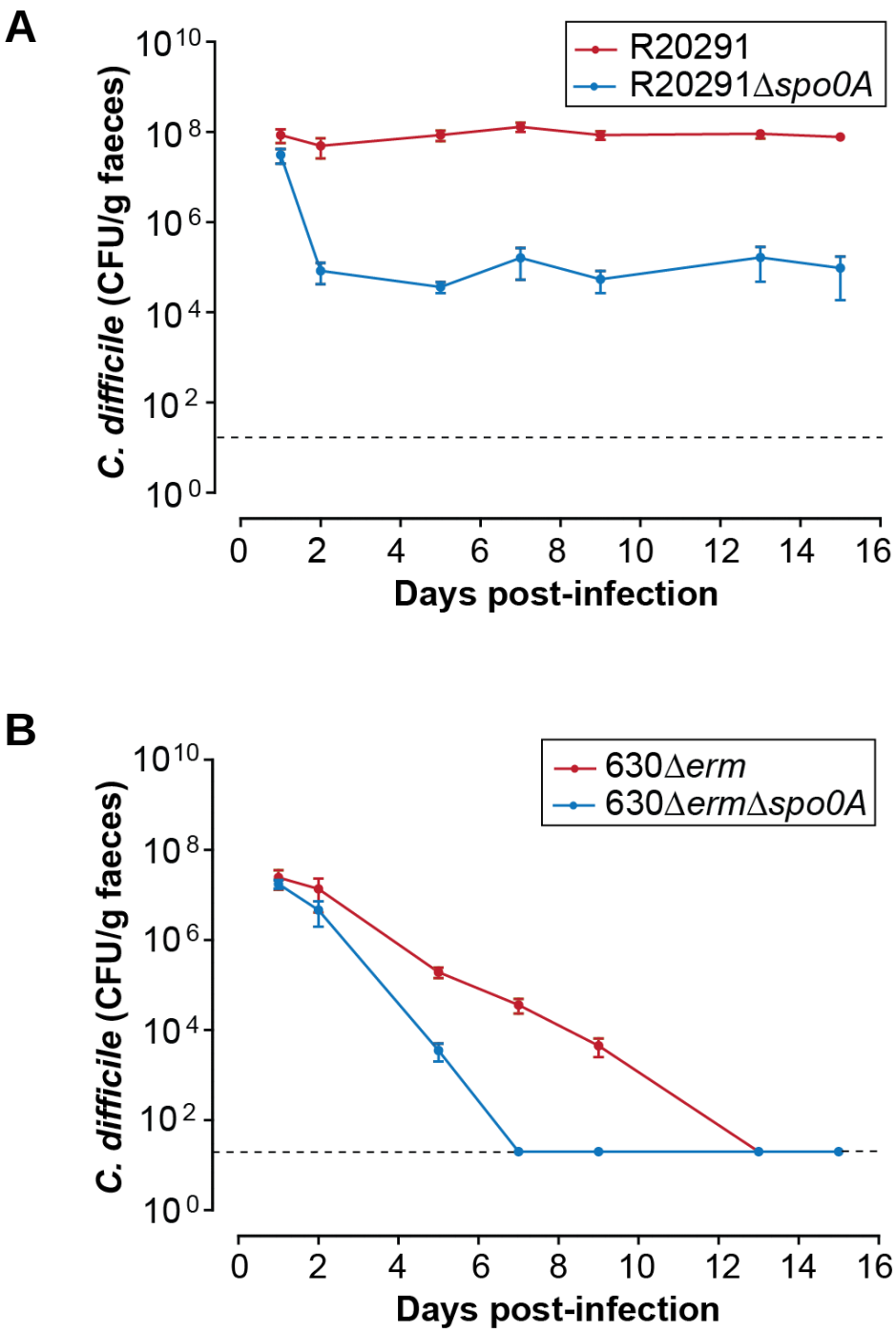


Figure 3.9. *C. difficile spo0A* gene is required for persistent colonisation in mice. Representative faecal shedding profiles from mice (*n* = 5 per group) infected with **A)** *C. difficile* R20291 or R20291Δ*spo0A*, or **B)** *C. difficile* 630Δ*erm* or 630Δ*erm*Δ*spo0A*. In both *C. difficile* ribotypes, genetic disruption of *spo0A* resulted in a reduced fitness *in vivo*.

the *C. difficile* 630Δ*erm*Δ*spo0A* mutant was reproducibly cleared ~6 days earlier than the parental strain (Figure 3.9B). These results indicate that the *spo0A* gene contributes to the intestinal colonisation and persistent infection of *C. difficile* in mice.

3.4.4.2 *C. difficile spo0A* mediates disease recurrence

Recurrent infection is defined by the complete subsidence of symptomatic disease followed by its subsequent reappearance. Given that *C. difficile spo0A* may affect persistence, we were interested in whether Spo0A also influences recurring infection. To test this, mice infected with R20291, R20291Δ*spo0A*, *C. difficile* 630Δ*erm* or 630Δ*erm*Δ*spo0A* were treated with a 7 day course of oral vancomycin. This is a standard therapy for patients experiencing severe systemic *C. difficile* disease. During this period the cages were regularly changed to reduce the potential for environmental transmission.

We found that vancomycin therapy rapidly decreased the levels of faecal shedding of *C. difficile* to below the detection limit in all infected groups of mice (Figure 3.10A and B). This was expected since these strains are sensitive to vancomycin. However, 3 days after vancomycin withdrawal the mice infected with parental *C. difficile* R20291 or *C. difficile* 630Δ*erm* exhibited recurrent infection with the same isolate.

With *C. difficile* 630Δ*erm*, this recurrence phenotype was observed after the first two courses of vancomycin, however after a third treatment the recurrent symptoms did not occur (Figure

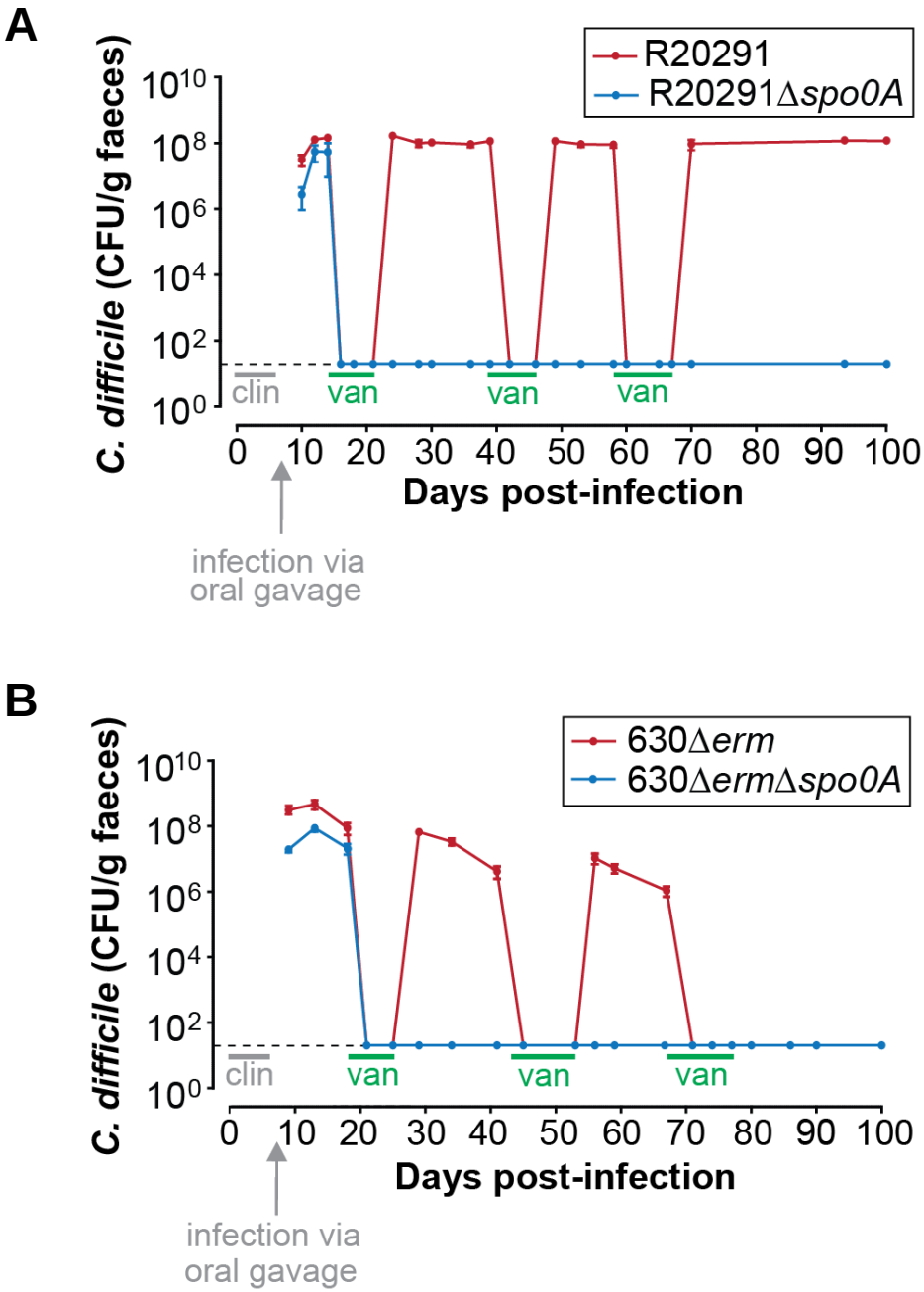


Figure 3.10. *C. difficile spo0A* gene is required for relapsing disease in mice. A and **B)** Representative faecal shedding profiles from mice (n = 5 per group) infected with *C. difficile* R20291, R20291Δ*spo0A*, 630Δ*erm* or 630Δ*erm*Δ*spo0A*. Mice were pre-treated with clindamycin (represented as a grey line) for 7 d prior to infection via oral gavage. Following infection, mice received a series of 7-10 d courses of vancomycin (represented as green lines) during which faecal shedding of *C. difficile* decreased to below the detection limit of the assay in all groups of mice. The dashed horizontal line indicates the detection limit.

3.10B). In contrast, *C. difficile* R20291-infected mice consistently exhibited disease recurrence on cessation of treatment for up to 100 days (Figure 3.10A). Significantly, mice infected with either *C. difficile* R20291Δ*spo0A* or 630Δ*erm*Δ*spo0A* cleared the infection during the first course of vancomycin and did not demonstrate recurrence at any point during the 100 day monitoring period (Figure 3.10A and B).

3.4.4.3 Disease recurrence is likely the result of relapse

Previous publications have indicated that approximately 50% of *C. difficile* recurrences are the result of re-exposure to the pathogen from an exogenous source such as other patients, healthcare workers, or from the environment, indicating that this is a key risk factor for disease recurrence (23). The remaining 50% of recurrences are estimated to be the result of a reactivation of the same strain that caused the initial infection, which had endogenously persisted within the host.

Thus, we hypothesised that recurrence in R20291 and 630Δ*erm*-infected mice could be the result of either persistence of the bacterium within the host during vancomycin therapy, or re-colonisation from environmental spores. To address this, mice infected with *C. difficile* R20291 or *C. difficile* 630Δ*erm* were treated with a 5 day course of vancomycin after which the entire intestinal tract (small and large intestine and cecum) was removed and cultured for *C. difficile*.

We consistently failed to detect any *C. difficile* in the intestinal tracts of mice receiving vancomycin therapy. However, we could readily culture *C. difficile* from mouse chow, cage shavings and bedding, as well as the fur, mouth and feet of mice receiving vancomycin treatment, even though cages where changed during the vancomycin treatment. Spores were detected in all locations, indicating that Spo0A acts as a persistence factor, enabling *C. difficile* to remain in the environment generating an infection reservoir on cessation of antibiotic therapy.

In a study of hospital patients, Walters *et al*. (1983) observed that disease recurrence was associated with incomplete clearance of *C. difficile* from the bowel of infected patients receiving vancomycin treatment, though the authors acknowledged that a contaminated environment was a potential reason for relapse (58). Our findings, however are in direct contrast with Onderdonk *et al*. (1980) who reported that the levels of spores actually increased in the ceacal contents of *C. difficile*-infected mice during vancomycin treatment (220). Nonetheless, our data suggest that it is relapse (i.e., recurrence with the same strain) from an exogenous reservoir that was the source of infection in our model of infection.

3.4.5 Role of *spo0A* in *C. difficile* transmission

The central importance of *C. difficile* transmission is recognised (61, 98, 150). However, the molecular factors that influence transmission are unknown. A greater understanding of the

transmission dynamics of *C. difficile* is potentially of great importance to infection control and clinical practices.

3.4.5.1 *C. difficile spo0A* is required for efficient host-to-host transmission

We previously used the murine infection model to demonstrate host-to-host transmission and the highly transmissible nature of various *C. difficile* ribotypes (61, 98, 150). Here, we use this model to investigate the role of the *C. difficile spo0A* gene in host transmission via distinct transmission routes, mimicking the situation commonly observed in hospital patients (see Materials and methods Table 2.3 and Figure 2.1). During transmission experiments, the donor mice shed equivalent levels (~10^8 CFU/gram fresh faeces) of vegetative *C. difficile* R20291, 630Δ*erm* or their respective *spo0A* mutant derivatives (Figure 3.11A). Mice infected with *C. difficile* R20291 or *C. difficile* 630Δ*erm* also shed spores (~10^6 CFU/gram fresh faeces) in contrast to mice infected with the *spo0A* mutants where spores were undetectable (Figure 3.11A).

Mingling donor mice that were shedding high-levels of *C. difficile* R20291 or 630Δ*erm* with naïve recipient mice for 1 h (in cages without bedding) resulted in 100% transmission (Figure 3.11B; n = 10 for each strain). In contrast, mingling donors excreting *C. difficile* R20291Δ*spo0A* or 630Δ*erm*Δ*spo0A* with naïve recipient mice resulted in only 40% or 20% transmission efficiency, respectively. In these experiments coprophagy was regularly observed and this behaviour likely promotes *C. difficile* transmission. To further define the

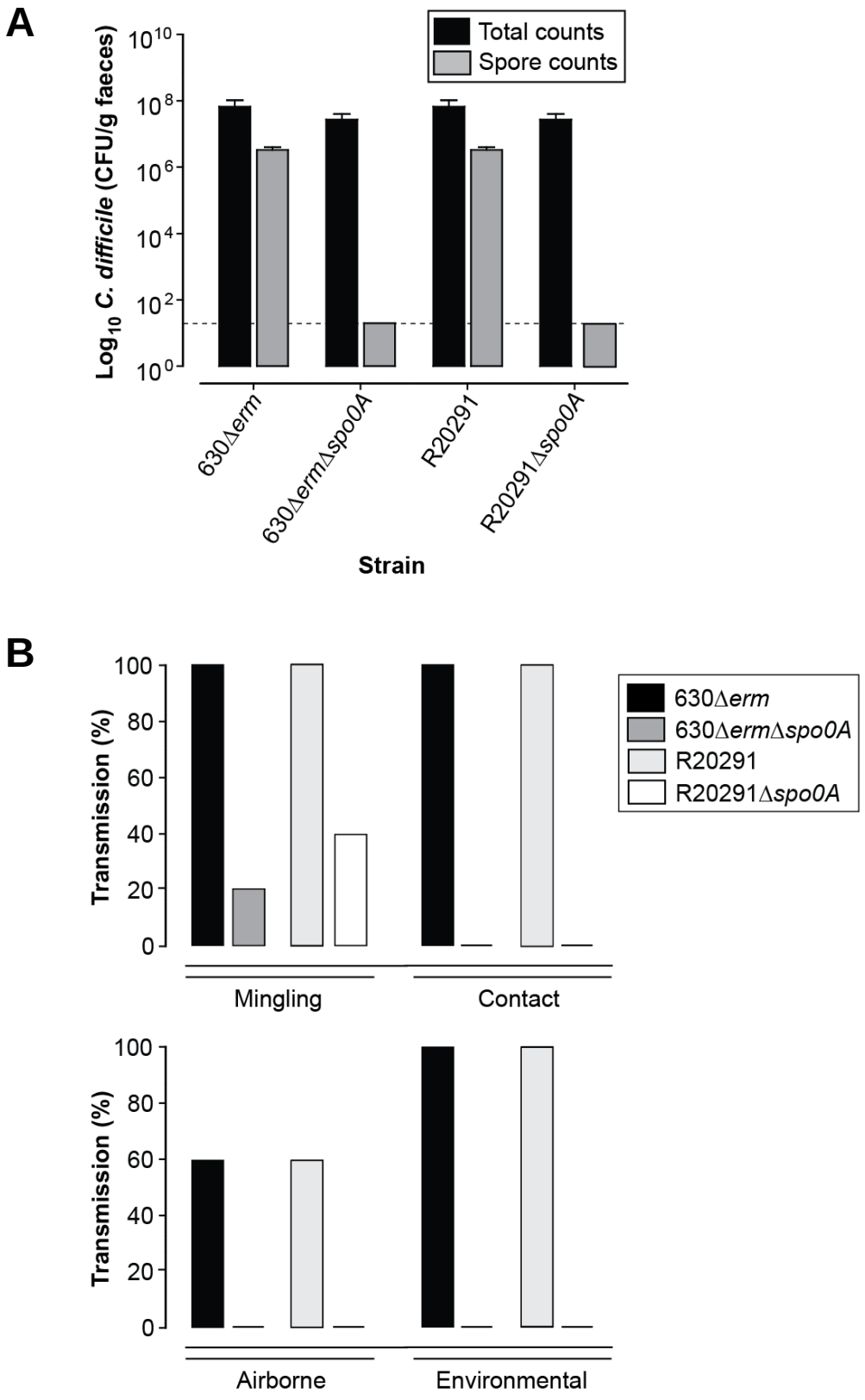


Figure 3.11. Host-to-host transmission of *C. difficile* is mediated by Spo0A. A) Average

faecal shedding of *C. difficile* by mice (n = 5 per group) at 5 d post-infection. The dashed horizontal line indicates the detection limit. **B)** Transmission efficiency of *C. difficile* 630Δ*erm*, R20291, 630Δ*erm*Δ*spo0A* or R20291Δ*spo0A*, demonstrating the percentage of naïve recipient mice that acquired infection following exposure to infected donor mice via mingling, contact, airborne or environmental transmission. Efficiency of transmission was determined as described in the Materials and Methods.

transmission route, we placed a porous wall between donor and recipient mice so that mice could come into contact but coprophagy was blocked. During these experiments *C. difficile* R20291 and 630Δ*erm* were still transmitted at 100% efficiency, whereas the *spo0A* mutant derivatives failed to transmit at a detectable level (Figure 3.11B; contact transmission).

Next, we tested transmission between donor and recipient mice that were separated by a double porous wall that prevents coprophagy and direct contact, but would still allow airborne transmission. In these experiments, the *C. difficile* R20291 and 630Δ*erm* were transmitted at 60% efficiency whereas the respective Δ*spo0A* derivatives failed to transmit (Figure 3.11B; airborne transmission). Thus, Spo0A plays a key role in the transmission of *C. difficile* between hosts sharing a contained environment but with no direct contact.

Environmental *C. difficile* spores are proposed to be a significant transmission reservoir (61, 221). To mimic this situation, we contaminated cages (without bedding) by placing donor mice shedding *C. difficile* in cages for 1 h, and then removed faeces and allowed the cages to stand overnight (~16 h) in a sterile environment such that only the aerobic resistant spores survived. Exposure of naïve recipient mice to the environment contaminated with *C. difficile*

spores resulted in 100% transmission to naïve recipient mice. In contrast, none of the naïve recipient mice exposed to *C. difficile* R20291Δ*spo0A* or 630Δ*erm*Δ*spo0A*-contaminated cages became colonised (Figure 3.11B; environmental transmission).

3.5 Discussion

The availability of a murine infection model opens the way to study the genetic basis of *C. difficile* disease, relapse, persistence and transmission. Here, we demonstrate for the first time that the *spo0A* gene of clinically relevant genetic variants of *C. difficile* is essential for persistent infection and efficient host-to-host transmission. Importantly, we also show that a functional *spo0A* gene is required for relapsing infection after vancomycin therapy, and we further illustrate that the local environment (i.e. cage, food, mouse fur, etc.) may serve as a reservoir of *C. difficile*.

Despite multiple attempts we could not culture *C. difficile* from the intestinal tract of challenged mice during vancomycin treatment, indicating that maintenance of *C. difficile* within the intestinal tract was not a likely cause of disease recurrence, although we cannot rule out that *C. difficile* persisted at very low levels or in a non-culturable form. We could, however, readily culture *C. difficile* from inside the cage during treatment.

Given the highly transmissible nature of environmental spores, it is possible that relapsing infection was in fact re-infection *with the same strain* that contaminated the local environment

before antibiotic treatment and persisted until after antibiotic therapy ended. The inability of *C. difficile spo0A* mutants to cause relapsing infection and form spores is consistent with this model. Indeed, antibiotic treatment of mice perturbs the intestinal microbiota and makes them extremely susceptible to *C. difficile* infection via environmental spores (98).

As expected, genetic inactivation of *spo0A* resulted in an asporogenous phenotype in *C. difficile* R20291 (ribotype 027) and 630Δ*erm* (ribotype 012) (177). In addition to not being able to form spores, we demonstrate *C. difficile spo0A* mutants produced elevated levels of the toxins TcdA and TcdB *in vitro* compared to the respective parental strains. Therefore, Spo0A may negatively regulate toxin production, which was associated with increased virulence for mice infected with the *C. difficile* 027 variant. This finding was in contrast to the observations of Underwood *et al*. (2009), which indicated that a *C. difficile* 630Δ*erm spo0A* mutant produced < 10% of the parental levels of TcdA (177). This discrepancy may reflect the different methods used to quantify toxin levels. However, it is also noteworthy that there are several further flaws that could explain these differences, the most obvious being the lack of genetic complementation to restore the sporulation and toxin expression phenotypes. Thus, it is not explicit what genetic lesions are contributing to these phenotypes.

The data presented in this chapter strongly suggests that Spo0A is required for persistence, disease recurrence and transmission of *C. difficile* in our clindamycin-induced mouse model of infection. However, it is important to note that *C. difficile* R20291Δ*spo0A* and 630Δ*erm*Δ*spo0A* derivatives are inherently resistant to clindamycin (owing to the ClosTron-mediated insertion of the *ermB* RAM). *C. difficile* R20291 and 630Δ*erm* parental strains do

not carry a functional copy of this gene and are consequently susceptible to clindamycin (see chapters 1.6.1.1 and 2.2.5.9). Although clindamycin has a short elimination half-life of approximately 2 - 4 hours (222), we were nonetheless unable to truly compare the colonisation, persistence and transmission kinetics of such strains in a like-for-like manner. Additionally, given the sensitivity *C. difficile* vegetative cells to oxygen, enumeration of the bacterium from the faeces inherently favours the recovery of the spore. In this sense, it may have been prudent to quantify the number of vegetative cells and spores directly from the site of infection. Clearly, however this would become the end-point of the experiment.

Furthermore, in the well-characterised model organism *B. subtilis*, Spo0A regulates processes other than spore formation, such as efflux pumps and metabolism (184). Consequently, it is possible that in addition to spore formation, there are other functions controlled by Spo0A that may play a role in persistence and transmission. Future experiments that define the *C. difficile* Spo0A regulon at the transcriptional and proteomic level should identify persistence and transmission factors controlled by Spo0A in *C. difficile*. Additionally, this approach will enable us to identify any pleiotropic effects of the *spo0A* mutation besides sporulation, and link this back to its role in disease, relapse, persistence and transmission.

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

4.1 Introduction

The *C. difficile spo0A* gene codes for a highly conserved transcriptional regulator of sporulation that is required, as shown in Chapter 3, for relapsing disease and transmission in mice. Defining the *spo0A* regulon using transcriptional and proteomic approaches should improve our understanding of the *C. difficile* sporulation programme at the whole genome level and could identify persistence and transmission factors controlled by Spo0A in *C. difficile*.

Furthermore, the profound morphological and physiological changes that occur during the course of spore formation in *C. difficile* will be accompanied by a shift in gene expression and proteome profiles. As such, the ability to investigate the global behaviour of the *C. difficile* transcriptome and proteome in a quantitative manner during sporulation is essential if we are to identify gene products, regulators and regulons associated with this phenomenon.

4.1.1 Transcriptomics

The genome sequence of *C. difficile* 630 was completed at the WTSI (http://www.sanger.ac.uk/Projects/C_difficile/) and was available to support this analysis. Additionally, the Illumina sequencing platform has been adapted to enable high-throughput screening of bacterial transcriptomes. One method, termed ssRNA-Seq, is a particularly

powerful and reproducible tool to quantitatively survey global transcript abundance at a single strand level (223).

4.1.2 Proteomics

Understanding the Spo0A regulon will clearly require a multi-layered understanding of the interactions between functional gene products, including proteins, and the mechanisms by which their expression is regulated. Proteomic analysis is particularly important since any phenotype observed should be a direct consequence of the activity of a cells protein composition. Recently, the liquid chromatography-tandem mass spectrometry (LC-MS/MS) platform at the WTSI has been adapted to allow accurate and comparative quantification of protein expression between biological samples.

4.1.3 An integrated approach to studying spore formation

Transcriptomic and proteomic approaches are mutually complementary (224) and an integrative approach enabling the parallel profiling of both RNA transcripts and proteins is clearly an advantage to gaining a comprehensive and integrated overview of the Spo0A regulon. Such a combined approach can be used to derive a holistic picture of *C. difficile* biology and physiology that could not be readily achieved using each approach independently

(224, 225). Taking advantage of these methods, we sought to apply mRNA and protein profiling to define the *C. difficile* Spo0A regulon.

4.2 Aims of the work described in this chapter

In this chapter, we aimed to provide a molecular description of the *C. difficile* Spo0A regulon at the whole genome level. Cataloguing transcripts and proteins should facilitate our understanding of the role of *C. difficile* Spo0A in disease and may lead to the identification of persistence and transmission factors controlled by Spo0A. Additionally, this approach will enable us to identify any pleiotropic effects of the *spo0A* mutation beyond sporulation and link these back to any role in disease and transmission.

4.3 Results

4.3.1 *C. difficile spo0A* is expressed during exponential growth

We previously demonstrated that Spo0A is essential for spore formation in *C. difficile* 630Δ*erm* (Chapter 3). To determine the effect of *spo0A* on *C. difficile* growth kinetics, we generated defined insertional mutations in the *spo0A* gene of *C. difficile* 630Δ*erm*, generating 630Δ*erm*Δ*spo0A*, and then performed growth curves in Wilson's broth. We found that *C.*

difficile 630Δ*erm* and 630Δ*erm*Δ*spo0A* displayed comparable growth kinetics, however the *spo0A* mutant failed to produce spores or spore-like elements as determined by anaerobic culturing after ethanol shock (spores but not vegetative cells of *C. difficile* are resistant to ethanol), which is in agreement with previously published data (Figure 4.1) (177).

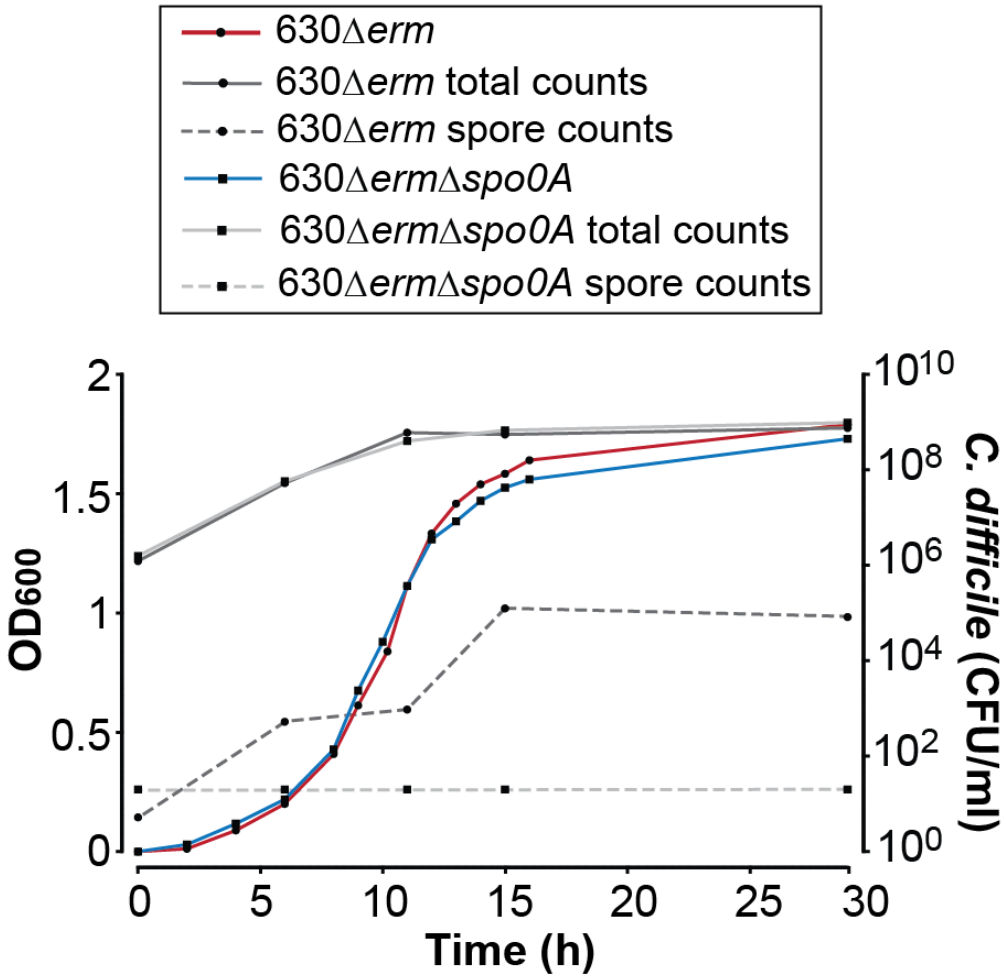


Figure 4.1. Growth and sporulation properties of *C. difficile* 630Δ*erm* derivatives. *C. difficile* cultures were grown in Wilson's broth under anaerobic conditions. Samples were taken at regular intervals throughout growth, during which the optical density, total cell and spore counts were determined. Optical density (OD_{600}) is plotted on the left *Y*-axis; bacterial counts (CFU/ml) are plotted on the right *Y*-axis.

Using *C. difficile* 630Δ*erm* cultures, we found that spores were produced during exponential phase, most notably at around 12 h of growth. This was surprising since spore formation is generally considered to be a stationary-phase phenomenon (associated with starvation or stress) and is somewhat conflicting with other models of bacterial sporulation (176, 177, 216). To demonstrate that the production of spores during exponential phase was directly attributable to the production of Spo0A, we performed a Western blot using *C. difficile* Spo0A-specific antibodies. We found that Spo0A was produced during both exponential and stationary growth in *C. difficile* 630Δ*erm* but was not produced by *C. difficile* 630Δ*erm*Δ*spo0A* at any time point (Figure 4.2). Given our interest in defining the *C. difficile spo0A* regulon and identifying genes involved in spore formation and sporulation, we chose to focus on exponentially growing (sporulating) *C. difficile*.

4.3.2 *C. difficile spo0A* is a global regulator of gene expression

Several studies in *Bacillus* have demonstrated that Spo0A is a global transcriptional regulator (184, 185, 226). In contrast, the role of Spo0A in *C. difficile* gene expression has not been well defined at the molecular level. For the first time, using high-density, strand-specific cDNA sequencing (ssRNA-seq), we surveyed gene expression profiles of exponentially growing *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A* at the whole-genome level in order to identify global changes in gene expression profiles and to identify genes, regulators and regulons associated with this phenomenon.

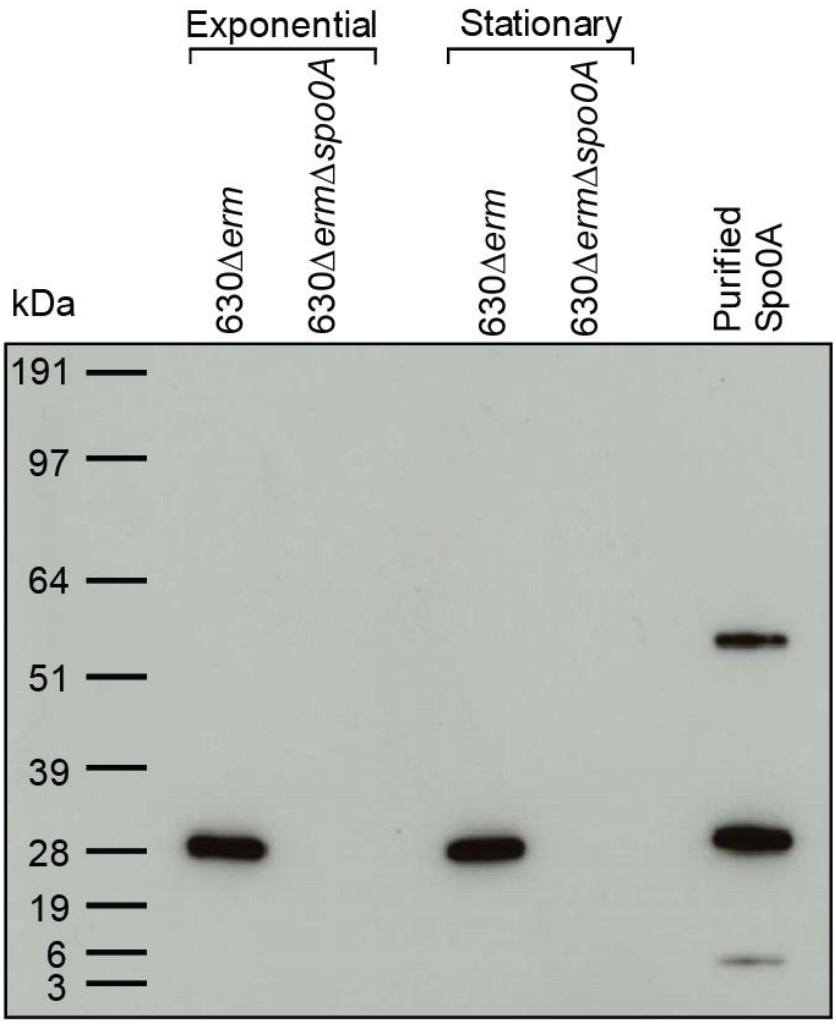


Figure 4.2. Western blot analysis of Spo0A expression in *C. difficile* 630Δ*erm* derivatives. *C. difficile* cultures were grown in Wilson's broth under anaerobic conditions, and samples were taken during mid-exponential and stationary growth. Extracted proteins with an equivalency to ~10^7 cells/lane were probed with anti-*C. difficile* Spo0A-specific antibodies. Purified Spo0A (~30 kDa) was included for reference. Lower and higher molecular weight bands in this lane likely represent degradation and Spo0A dimerisation, respectively.

4.3.2.1 Genome-wide identification of genes significantly regulated by Spo0A

We performed ssRNA-Seq on three biological replicates of both *C. difficile* 630Δ*erm* and *C. difficile* 630Δ*erm*Δ*spo0A*. However, since the total number of Ilumina read counts were not identical between the samples due to varying sequencing depths, the data were normalised using the DESeq package, as described in Materials and methods chapter 2.2.7.3. This reduces the systematic technical variation and thus enables a direct comparison between Spo0A mutant and parental *C. difficile* gene expression levels.

In order to identify and quantify genes that are controlled by Spo0A, normalised data were transformed on a variance stabilised scale resulting in homoscedastic data. This was additionally $\log_2$ transformed and filtered for a P value of ≤ 0.01. This stringent cutoff was chosen to reduce the likelihood of false positives and to generate a computationally manageable set of differentially expressed genes. Thus, it is likely that identified genes actually represent an underestimate of the true scale of Spo0A regulation in *C. difficile*.

Based on these parameters, we found that 261 genes were significantly differentially expressed in 630Δ*erm*Δ*spo0A* compared to the parental strain, of which 126 and 135 were up- and down-regulated, respectively (see Appendices 2 and 3) (Figure 4.3). We then mapped the relative gene expression values of all CDSs to the published *C. difficile* 630 genome, and overlaid the differentially expressed genes, as shown in Figure 4.4. In addition to mapping the differentially expressed genes, we also mapped an expanded list of 990 spore-associated polypeptides onto the annotated *C. difficile* 630 genome (Figure 4.4) in order to identify any

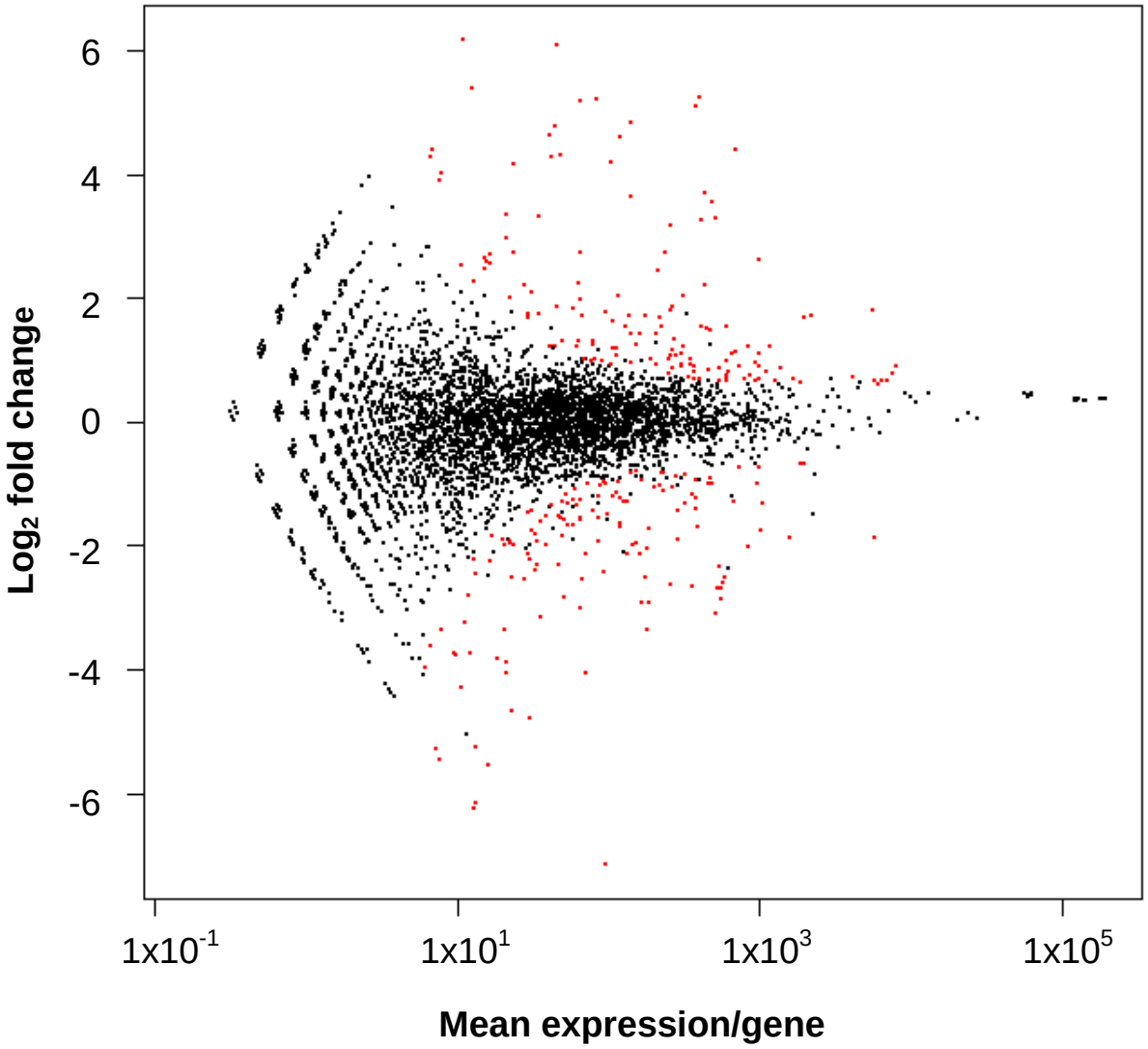


Figure 4.3. Identification of differentially expressed genes in *C. difficile* 630Δ*erm*Δ*spo0A* by transcriptional profiling. Scatter plot of the $\log_2$ fold changes against the normalised mean read abundance per gene (calculated at the base level). Red dots represent genes considered to be significantly differentially expressed using a *P* value of $P = \leq 0.01$. Black dots signify genes not deemed to be significantly differentially expressed according to these criteria.

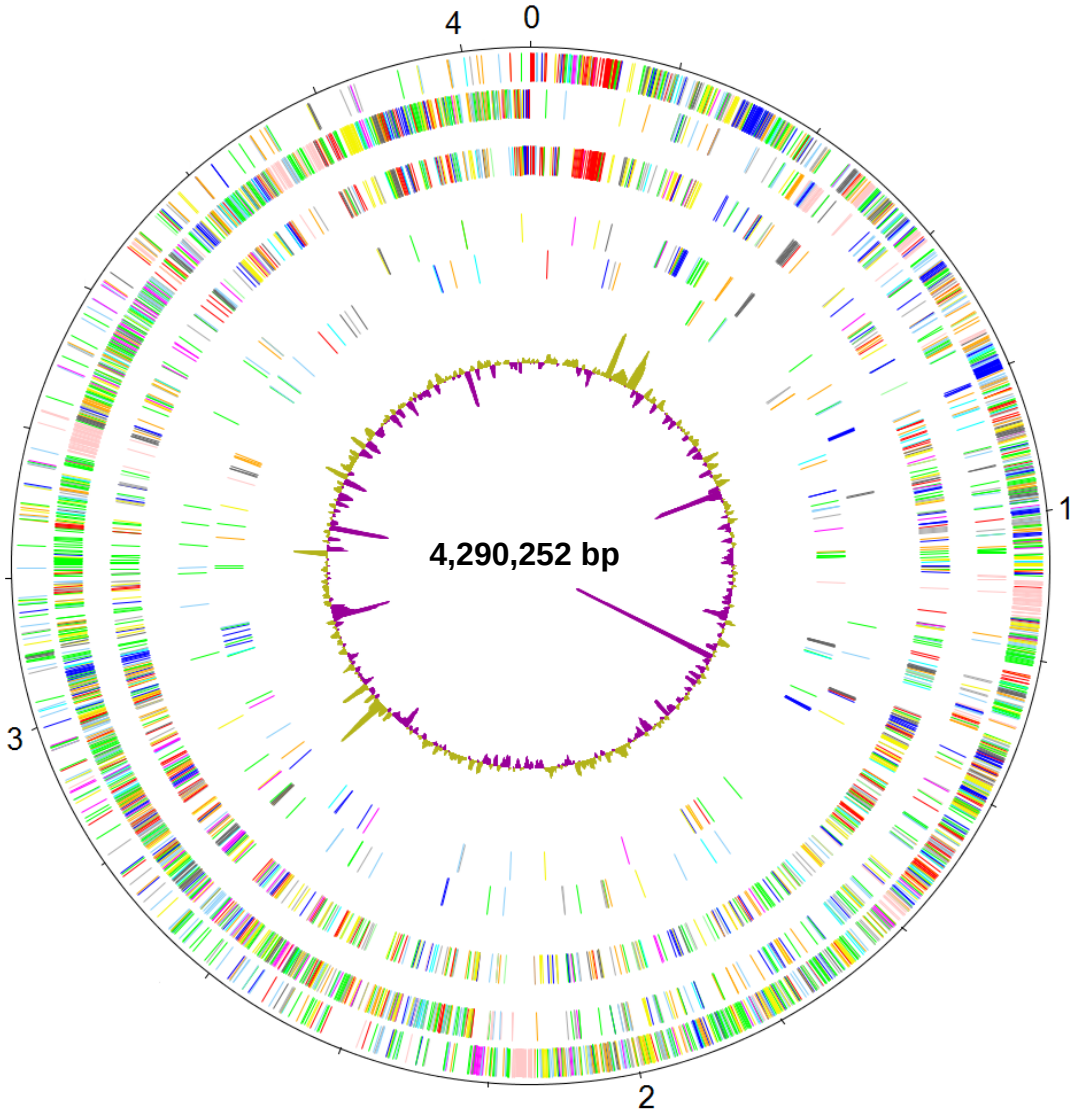


Figure 4.4. Global overview of the *C. difficile* 630Δ*erm* Spo0A-regulated transcriptome. The scale is marked in megabases (0-4) on the outermost circle. The concentric circles represent the following (from the outside in): 1+2, all CDS transcribed on the forward (clockwise) and reverse (counterclockwise) strands respectively; 3+4, CDS that are up and down regulated (P = 0.01) in 630Δ*erm*Δ*spo0A* compared to parental 630Δ*erm*, respectively; 5, plot of the $\log_2$ fold changes averaged over a 10-kbp window, in which green and purple shading corresponds to up and down regulation in 630Δ*erm*Δ*spo0A* compared to parental 630Δ*erm*, respectively. CDS are coloured according to predicted function as follows: dark blue, pathogenicity/adaptation; dark grey, energy metabolism; red, information transfer; dark green, surface-associated; cyan, degradation of large molecules; magenta, degradation of small molecules; yellow, central/intermediary metabolism; pale green, unknown; pale blue, regulators; orange, conserved hypothetical; brown, pseudogenes; pink, phage and insertion sequence (IS) elements; grey, miscellaneous.

genes that are critical for sporulation and/or germination. Genes regulated by Spo0A (either positively or negatively) were evenly distributed around the genome and were present on both the forward and reverse strands. A heat map of the top 40 most differentially expressed genes according to $\log_2$ fold-changes is shown in Figure 4.5.

In order to identify any global effects of the *spo0A* mutation, differentially expressed genes were assigned to a functional class, as designated in the annotated genome. These were then expressed as a percentage of the entire genome, which revealed a number of interesting trends, as shown in Figure 4.6, and discussed below. For example, of the 63 *in silico* predicted CDSs in the sporulation / germination class in the entire *C. difficile* 630 genome, almost a quarter (~22%) were found to be positively regulated by Spo0A, that is, their expression was attenuated in 630Δ*erm*Δ*spo0A* Figure 4.6A. None were found to be negatively regulated by Spo0A (Figure 4.6A). However, it is important to note that these data are based on the current annotation of the *C. difficile* 630 genome, which was published six years ago.

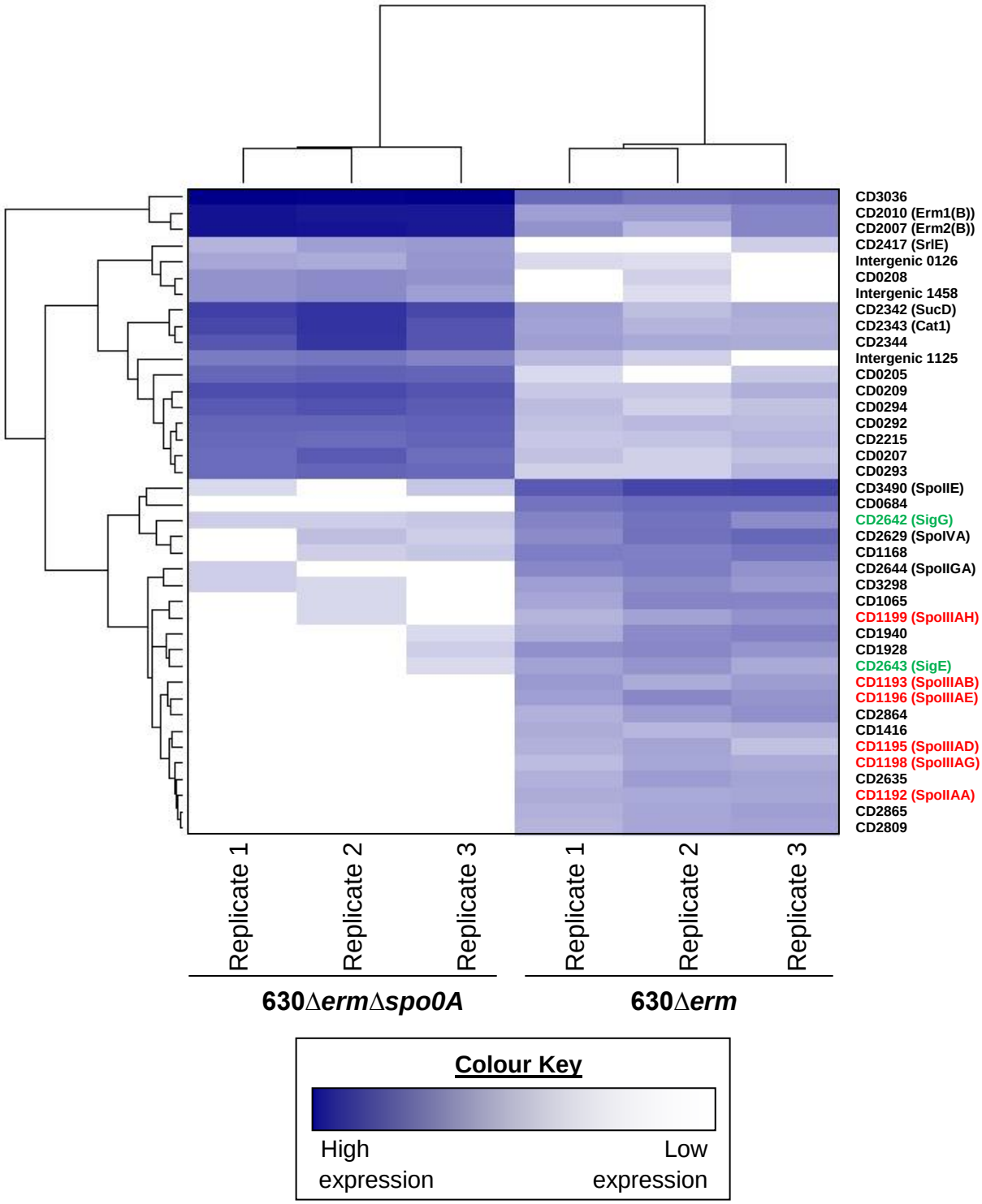


Figure 4.5. Heat map illustrating the expression data of the top 40 differentially expressed genes. Raw count data were normalised to enable a direct comparison and quantification of differentially expressed genes. Red, *spoIIIA* operon; green, *sigG* (CD2642) and *sigE* (CD2643) sigma factors.

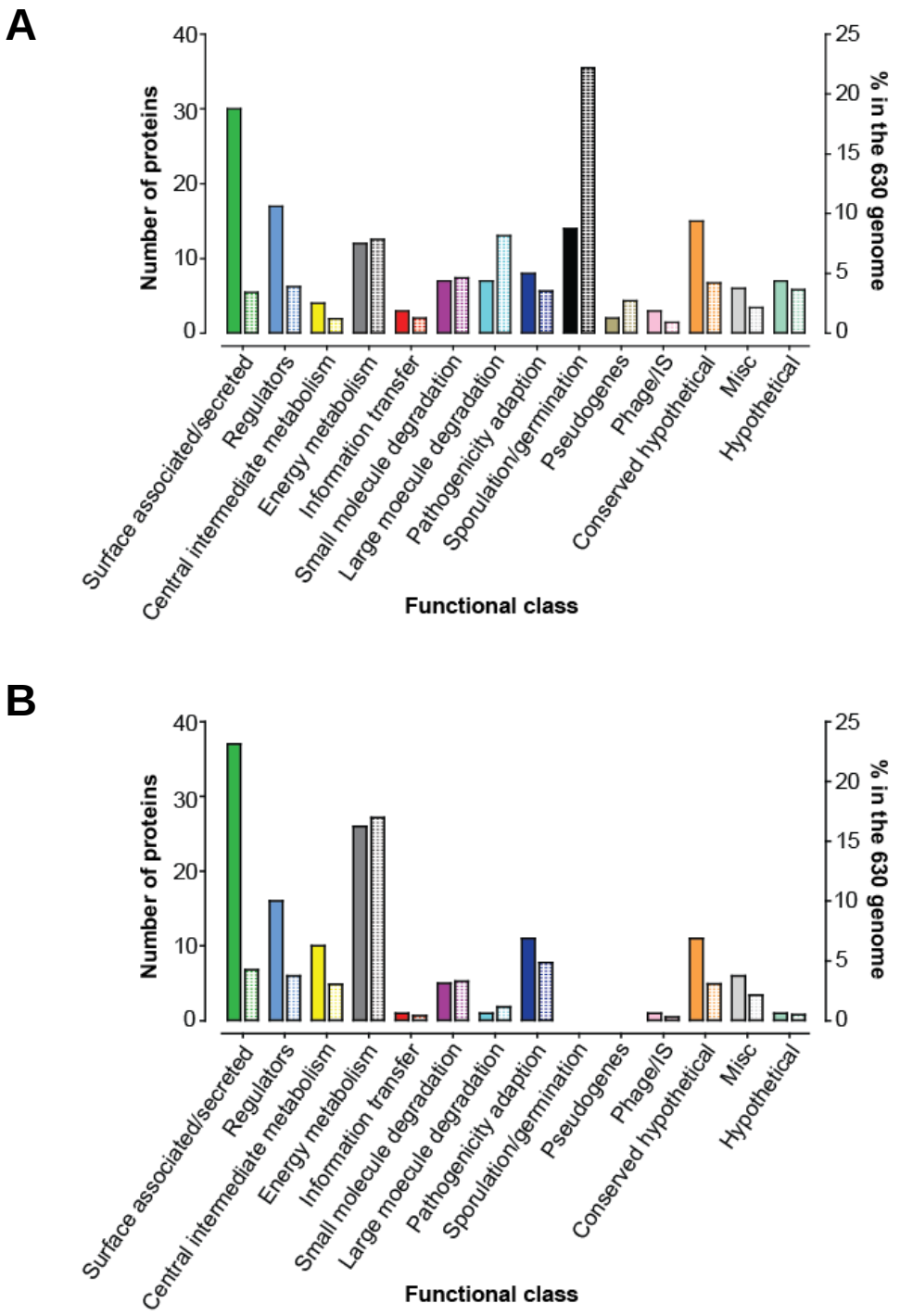


Figure 4.6. Distribution and relative abundance of *C. difficile* 630Δ*erm* Spo0A-regulated transcripts, according to functional classification. **A)** Transcripts down-regulated in 630Δ*erm*Δ*spo0A*, and **B)** transcripts up-regulated in 630Δ*erm*Δ*spo0A*. Transcripts were assigned to functional classes based on current *C. difficile* 630 annotation (10). Indicated are the absolute numbers of transcripts (filled bars, left *Y*-axis) and the percentage of the functional class relative to the total 630 genome (hashed bars, right *Y*-axis).

4.3.2.2 Confirmation of toxin production and sporulation phenotypes by ssRNA-Seq

We noted that Spo0A regulates the transcription of genes with significant phenotypes, including genes for TcdA production and spore formation, which are discussed in detail in Chapter 3.

4.3.2.2.1 Spo0A is an indirect negative regulator of TcdA production

tcdA expression was significantly up-regulated in 630Δ*erm*Δ*spo0A* at the mRNA level (Figure 4.7). This was anticipated given the increased levels of TcdA produced *in vitro* by both *C. difficile* 630Δ*erm*Δ*spo0A* and *C. difficile* R20291Δ*spo0A* (discussed in Chapter 3). However, we failed to identify an '0A' box upstream of the *tcdA* start site, implying that regulation could be via indirect control. In addition, *in vitro* binding assays did not indicate that *tcdA* was a direct target of Spo0A (Wiep Klass, personal communication) and thus some *tcdA* regulators may be encoded by genes which are under direct Spo0A control. For example, Saujet *et al*. (2011) demonstrated that the deletion of *sigH* markedly reduced *spo0A* expression and concomitantly increased TcdA expression (182).

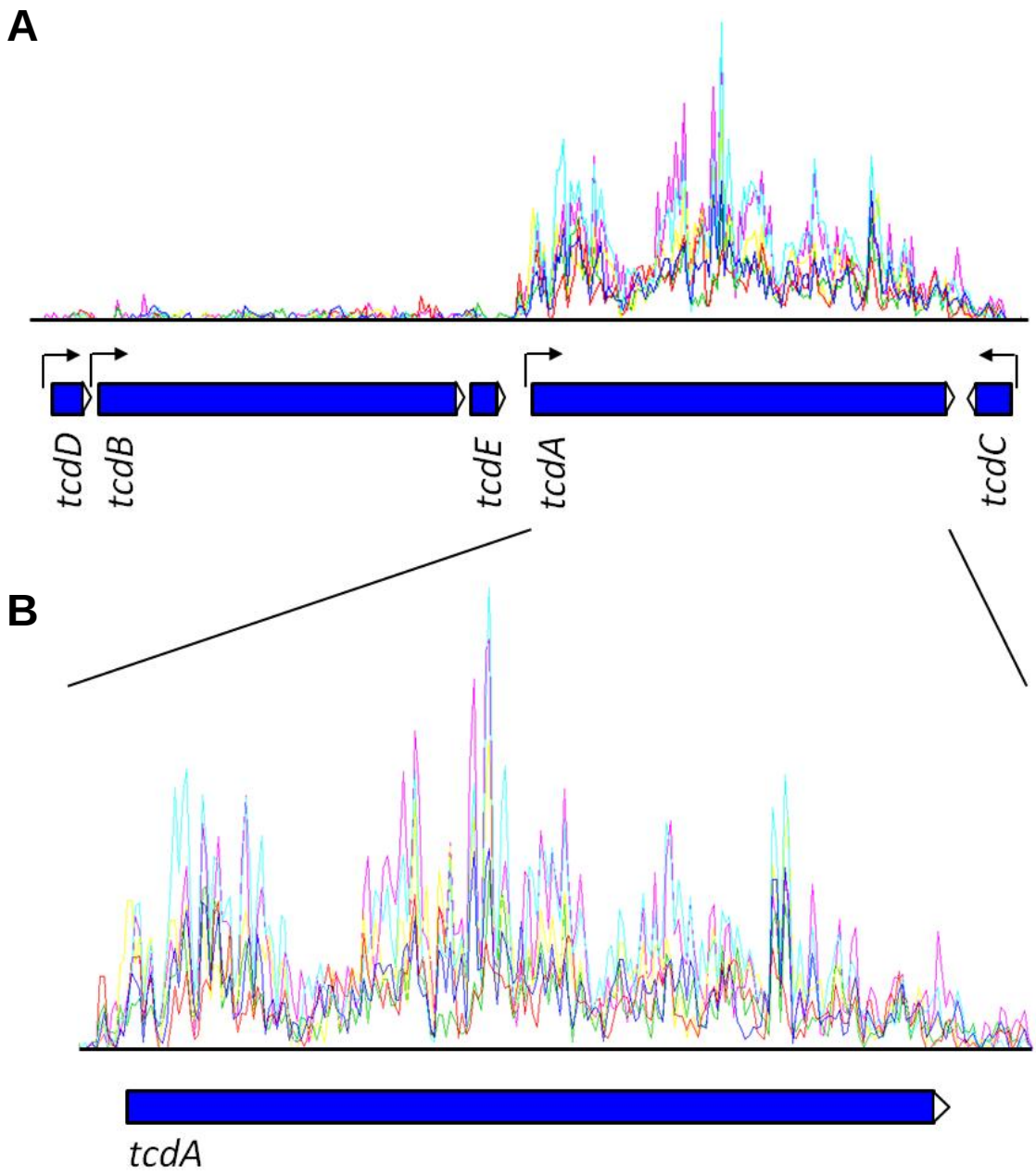


Figure 4.7. Linear representation of the *C. difficile* 630 genome with transcriptome reads mapping to the PaLoc region. A) Shown is the 19.6 kbp PaLoc region, which harbours five genes (*tcdDBEAC*) responsible for the synthesis and regulation of toxins A and B. **B)** Inset of the *tcdA* gene. Coloured lines represent transcriptome reads aligned to the 630 genomic sequence, and are coloured to represent biological replicates as follows: red, blue, green = *C. difficile* 630Δ*erm*; purple, turquoise, yellow = *C. difficile* 630Δ*erm*Δ*spo0A*. Arrows represent promoters (227). The expression of *tcdA* was significantly up-regulated at the mRNA level ($P = \leq 0.01$) in 630Δ*erm*Δ*spo0A*.

4.3.2.2.2 Spo0A is required for the expression of later stage sporulation genes

The mother-cell-specific *spoIIIA* polycistronic eight gene operon, which is involved in stage III spore formation, was significantly down-regulated in 630Δ*erm*Δ*spo0A* at the mRNA level (Figure 4.8). Many of the genes belonging to this operon (*CD1192-CD1199*) were also present in the heat map of the top 40 differentially regulated genes (Figure 4.5, highlighted in red). This was perhaps expected given the asporogenous phenotype observed in *C. difficile* 630Δ*erm*Δ*spo0A* (Chapter 3). Other sporulation genes including *spoIIIG*, *spoIIID*, *spoIVA*, *spoIIGA*, *spoVD*, *spoIIE* and the tricistronic *spoIIA* operon were also down-regulated in 630Δ*erm*Δ*spo0A*. Logically, *cspC* and *cspBA* encoding putative germination-specific proteases that are present in the spore proteome, and *CD3569* encoding a spore cortex-lytic enzyme were also down-regulated in the *spo0A* mutant. Interestingly, however, *ccpA*, a carbon catabolite repressor required for spore formation in *C. perfringens* was up-regulated in 630Δ*erm*Δ*spo0A*.

Sigma factors involved in the regulation of spore formation, such as *sigE* ($P = < 0.01$), *sigG* ($P = < 0.01$), *sigF* ($P = < 0.01$) and *sigH* ($P = 0.03$) were all significantly down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A* compared to parental *C. difficile* 630Δ*erm*. In fact, *sigG* (*CD2642*) and *sigE* (*CD2643*) were present in the heat map of the top 40 differentially regulated genes (Figure 4.5, highlighted in green). *sigH* also has a consensus '0A' box located ~84 bp upstream of the start site, suggesting that regulation of *sigH* expression by Spo0A may be via a direct interaction (Figure 4.9A). This has been supported by evidence of direct binding using *in vitro* binding assays (Wiep Klass, personal communication).

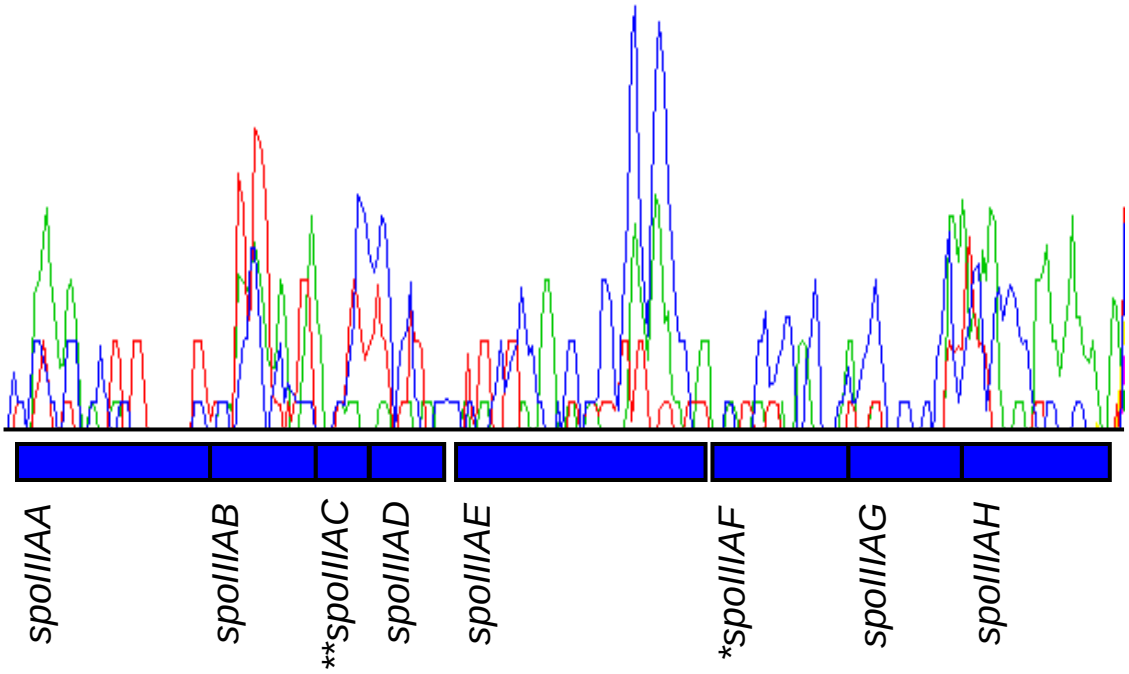


Figure 4.8. Linear representation of the *C. difficile* 630 genome with transcriptome reads mapping to the *spoIIIA* operon. Shown is the polycistronic eight gene s*poIIIA* operon, which is involved in stage III spore formation. Coloured lines represent transcriptome reads aligned to the 630 genomic sequence, and are coloured to represent biological replicates as follows: red, blue, green = *C. difficile* 630Δ*erm*; purple, turquoise, yellow = *C. difficile* 630Δ*erm*Δ*spo0A*. Expression of this operon was significantly attenuated in *C. difficile* 630Δ*erm*Δ*spo0A* ($P = \leq 0.01$) at the mRNA level. *$P = 0.04$; **$P = 0.07$.

4.3.2.2.3 Spo0A may be auto-regulated in C. difficile via a feedback loop

In *Bacillus*, *spo0A* transcription is auto-regulated via the direct binding of Spo0A to an '0A' box upstream of *spo0A*. This binding in turn represses *abrB* (a transition state regulator),

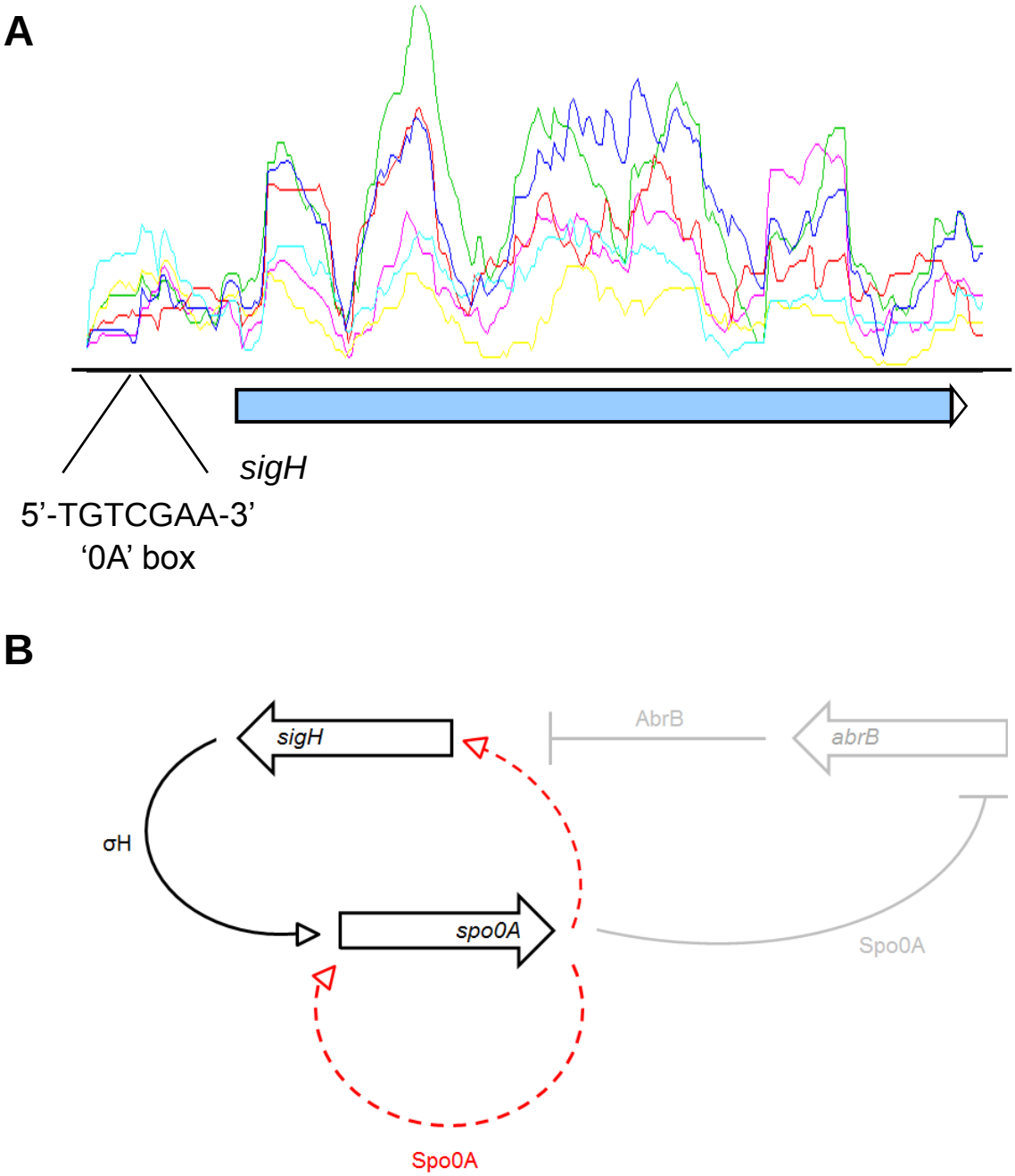


Figure 4.9. Spo0A may regulate *sigH* expression via a direct interaction. A) Expression of *sigH* was significantly attenuated in *C. difficile* 630Δ*erm*Δ*spo0A* (*P* = 0.03) at the mRNA level. Coloured lines represent transcriptome reads aligned to the 630 genomic sequence, and are coloured to represent biological replicates as follows: red, blue, green = *C. difficile* 630Δ*erm*; purple, turquoise, yellow = *C. difficile* 630Δ*erm*Δ*spo0A*. *sigH* also has a consensus '0A' box located directly upstream of its start site suggestive of direct Spo0A regulation. **B)** Schematic representation of a predicted network of *spo0A* and *sigH* regulation. Grey shading indicates regulation networks in *B. subtilis* that are absent in *C. difficile*. Red dashed lines indicate potential regulatory links in *C. difficile*.

resulting in *sigH* transcription and a concomitant activation of *spo0A*, as described in Figure 4.9B. However, based on sequence identity there is no *abrB* orthologue in *C. difficile*, indicating that the regulation loop may occur via a different mechanism.

It has previously been demonstrated that *C. difficile spo0A* is transcribed from a SigH-dependent promoter (182). Similarly, in *C. difficile* there is a consensus '0A' box directly upstream of *spo0A*, indicating direct auto-regulation by Spo0A. This is also supported by *spo0A* gel shift binding assays (Wiep Klass, personal communication). It is thus possible that *spo0A* transcription may be auto-regulated in a way comparable with *Bacillus*, but that *sigH* expression is directly driven by Spo0A rather than indirectly via AbrB (Figure 4.9B). Over-expression or ChIP-Seq studies could be used to test this hypothesis.

4.3.2.3 Spo0A is a regulator of *C. difficile* metabolism

We noted that Spo0A regulates the transcription of many genes of multiple functions. However, a large set of genes that were differentially expressed appeared to function in metabolism. For example, several genes involved in glycolysis were down-regulated in 630Δ*erm*Δ*spo0A*. These included the glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase (*gapB*) and glucose-6-phosphate isomerase (*pgi*), a central glycolytic genes regulator (*cggR*), beta-glucoside utilisation genes (*bglP*, *bglG*, *bglF*), and putative 6-phospho-alpha- and beta-glucosidases (*bglA3* and *bglA2*, respectively).

Additionally, components of the phosphotransferase (PTS) system (*CD0816*, *crr*, *ptsG* and *CD3137*) were down-regulated in 630Δ*erm*Δ*spo0A*. Similarly, two genes encoding a putative bi-functional glycine dehydrogenase (*CD1657*) and glycine cleavage system P-protein (*gcvPB*) were down-regulated in 630Δ*erm*Δ*spo0A*. These are thought to be components of the glycine cleavage system, whereby certain strict anaerobes are able to oxidise glycine to serve as an electron donor in the Stickland reaction (228-230).

In contrast, several mediators of amino group metabolism were up-regulated in 630Δ*erm*Δ*spo0A*. These included putative amino acid aminotransferases (*aspC*, *CD3664* and *CD2382*), a histidol-phosphate aminotransferase (*CD2502*) a glutamate dehydrogenase (*gluD*), a phosphoserine phosphatase (*CD0241*), an acetylornithine deacetylase (*argE*), a pyrroline-5-carboxylate reductase (*proC1*) and a D-alanine phosphoribitol ligase (*dltA*). The *hadAIBC* operon putatively involved in L-leucine reduction (231) was also found to be negatively regulated by Spo0A.

Genes involved in acidogenesis such as butyrate kinase (*buk*), phosphate butyryltransferases (*ptb*, *CD0715*), and the butyrate metabolism genes NAD-dependent 4-hydroxybutyrate dehydrogenase (*abfH*), 4-hydroxybutyrate CoA transferase (*abfT*), gamma-aminobutyrate dehydratase (*abfD*), succinate-semialdehyde dehydrogenase (*sucD*), succinyl-CoA transferase (*cat1*), acyl-CoA dehydrogenase (*bcd1*) and electron transport flavoprotein (*etfB1*) were also up-regulated in 630Δ*erm*Δ*spo0A*. Interestingly, all of these metabolic genes (excluding *CD2502*, *CD0241*, *argE* and *cat1*) encode proteins that are also present in the spore proteome.

4.3.2.3.1 Spo0A regulation of energy metabolism

Of the 153 *in silico* predicted CDSs belonging to the energy metabolism class in *C. difficile* 630, ~17% were found to be up-regulated in 630Δ*erm*Δ*spo0A* based on ssRNA-seq analyses (Figure 4.6B). These included the *rnf* complex (*rnfCDGEAB*), which is purported to be a membrane-bound multi-component system involved in NADH-dependent electron transport. This operon has been identified in many clostridial genomes, and was found to be negatively regulated by Spo0A at the mRNA level (excluding *rnfE* which was not differentially expressed). The complex is thought to oxidise NADH to NAD+ and reduce ferredoxin, however it has also been suggested that electrons can be transferred in the reverse direction (232). Previous publications have suggested the the *rnf* complex may have a role in generating NADH for butyrate fermentation (232).

Additionally, genes encoding metabolic enzymes such as oxidoreductases (including two indolepyruvate oxidoreductases, *iorA* and *iorB*), hydrogenases (*hymB* and *hymC*), flavoproteins and a cardiolipin synthetase (*cls*) were also up-regulated in 630Δ*erm*Δ*spo0A*, suggesting a possible role of Spo0A as a negative regulator of energy metabolism. Indeed, *spo0A* mutations have been shown to increase the metabolic maintenance demands of *B. subtilis* (233). However, *ntpB*, *ntpA* and *ntpI* which are members of the 8 gene ATPase synthase operon (*ntpIKECGABD*) were significantly down-regulated in 630Δ*erm*Δ*spo0A*.

4.3.2.4 Spo0A regulates the expression of surface-associated transcripts

Secreted/surface associated transcripts were well represented in our dataset (Figure 4.6). In fact, these were the most common functional class detected, and constituted ~22% and ~29% of detected transcripts which were positively and negatively regulated by Spo0A, respectively. Such proteins could be implicated in intestinal adherence, colonisation, or subversion of the hosts' immune system.

4.3.2.4.1 Surface-associated transcripts positively regulated by Spo0A

In silico predicted transcripts belonging to the secreted/surface-associated class whose expression was down-regulated in 630Δ*erm*Δ*spo0A* included some encoding ABC transporters such as ATP-binding proteins and permeases. *ssuA*, a spore-associated polypeptide, which encodes a putative periplasmic ABC transporter of aliphatic sulfonates, was also positively regulated by Spo0A. Sulfonates may have a role as electron acceptors in anaerobic respiration (234). Additionally, transcripts involved in cobalt transport (*cbiM* and *cbiQ*) were down-regulated in 630Δ*erm*Δ*spo0A*.

Adhesive cell surface proteins were also found to be potentially positively regulated by Spo0A. For example, *CD0440* and *CD2518*, which encode SLP-like proteins, are both similar to the *C. difficile* Cwp66 cell surface protein, suggesting adhesive activity. The down

regulation of such transcripts may explain the inability of *C. difficile* to cause a long-term persistent infection in our mouse model of infection (discussed in Chapter 3).

4.3.2.4.2 Surface-associated transcripts negatively regulated by Spo0A

Secreted/surface-associated transcripts whose expression was up-regulated in 630Δ*erm*Δ*spo0A* also included ABC transporters and permeases. Additionally, many transcripts encoding components of the phosphotransferase (PTS) carbohydrate transport system were also up-regulated. This system is important for anaerobic glycolysis (235) and up-regulated transcripts included sugar transporters and sugar-specific enzymes. One transcript (*CD0244*) encoding a CDP-glycerophosphotransferase involved in the biosysthesis of teichoic acids in *B. subtilis* was also negatively regulated by Spo0A in *C. difficile* 630Δ*erm*.

4.3.3 Global relationship between the transcriptome and proteome of exponentially growing *C. difficile*

According to the central dogma of molecular biology, the transcriptome is viewed as a precursor for the proteome. Whilst this is true, indeed the flow of information is from DNA to RNA to protein, it is being increasingly recognised that the concordance between transcript and protein expression is not always intuitive (224). In order to gain an accurate and thorough representation of the role of Spo0A in *C. difficile* biology and physiology, we integrated both

transcriptome and proteome data from exponentially growing *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*.

4.3.3.1 Analysis of proteins from exponentially growing *C. difficile*

4.3.3.1.1 SDS-PAGE

To ensure that proteomic data were comparable with transcriptome analyses, proteins were harvested from exponentially growing *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*, which were then electrophoretically separated by SDS-PAGE, as shown in Figure 4.10. Many polypeptide bands were resolved, with molecular weights ranging from 6 to > 100 kDa. Visual observations indicated that there was a common backbone of proteins shared between *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*, with both having similar mobilities and distribution profiles.

4.3.4.1.2 Mass-spectrometry

We then performed LC-MS/MS on three biological replicates and three technical replicates of both *C. difficile* 630Δ*erm* and *C. difficile* 630Δ*erm*Δ*spo0A*. In order for proteins to be quantified and to thus enable direct comparison of protein expression between mutant and parental *C. difficile*, samples were also dimethyl labelled prior to mass-spectrometry, as

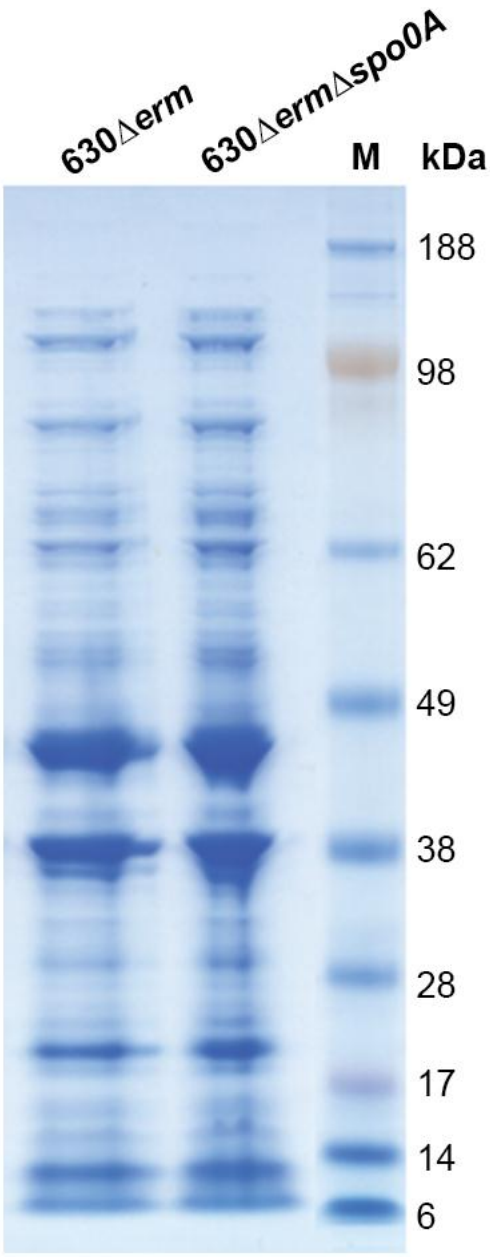


Figure 4.10. SDS-PAGE of cellular proteins of *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*. Secreted and cellular proteins with an equivalency to ~2x10^7 cells per lane were harvested from exponentially growing cultures, resolved by 12% SDS-PAGE and visualised by Coomassie-blue staining. Lane: 630, proteins extracted from *C. difficile* 630Δ*erm*; Δ*spo0A*, proteins extracted from *C. difficile* 630Δ*erm*Δ*spo0A*; M, protein molecular weight markers.

described in Materials and methods chapter 2.2.8.7. Based on a FDR of 1% and a minimum positive match of two peptides, we identified a total of 1,276 proteins by LC-MS/MS analysis. To identify a subset of these proteins that were considered to be significantly differentially regulated by Spo0A, all 1,276 proteins were $\log_2$ transformed and ranked based on relative expression between *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*, as shown in Figure 4.11. Based on these parameters, 150 proteins considered to be significantly differentially expressed were identified, of which 86 and 64 were up- and down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*, respectively (see Appendices 4 and 5, respectively).

4.3.3.2 Correlation between transcriptome and proteome expression

The level of concordance between the Spo0A-regulated transcriptome and proteome of *C. difficile* 630Δ*erm* is shown in Figure 4.12. Interestingly, only a small subset of CDSs were found to be differentially regulated by Spo0A at both the transcriptional and translational levels, indicating that transcript quantification alone may be insufficient to generate a comprehensive overview of this biological system. Furthermore, one CDS, *CD0865*, encoding a conserved hypothetical protein was down-regulated as a transcript, but up-regulated as a protein in 630Δ*erm*Δ*spo0A*. Thus, the key assumption that mRNA expression accurately informs protein expression may not hold true in *C. difficile*. Additionally, molecular events such as the efficiency of translation or protein degradation can notably alter protein levels independently of mRNA expression (236). However, it is possible that some of these observed differences are due to the technical limitations of our assays.

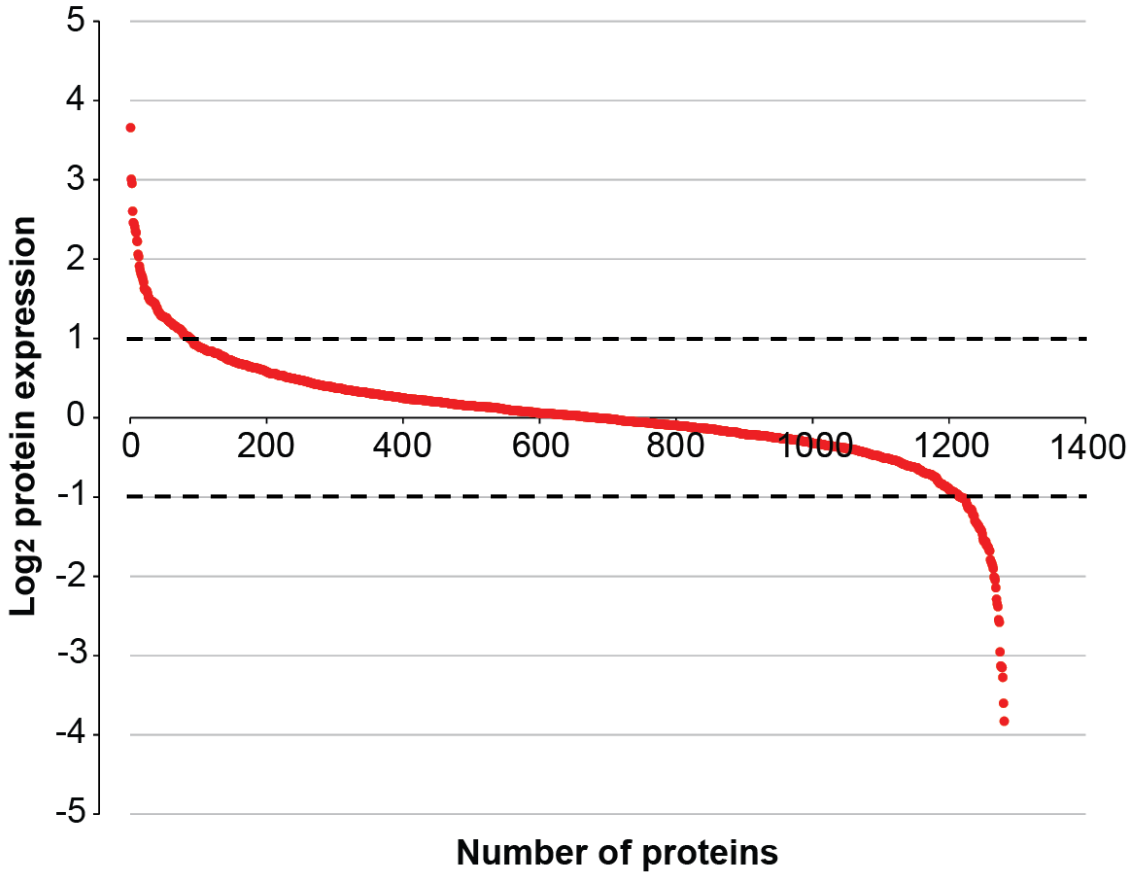


Figure 4.11. Total proteins identified by LC-MS/MS in *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*. A total of 1,276 proteins were identified by dimethyl labelled LC-MS/MS analysis, and are shown ranked based on relative expression between *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*. 86 proteins were considered to be significantly up-regulated in *C. difficile* 630Δ*erm*Δ*spo0A* and 64 proteins were considered to be significantly down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*, according to our parameters of significance (represented by a dashed line).

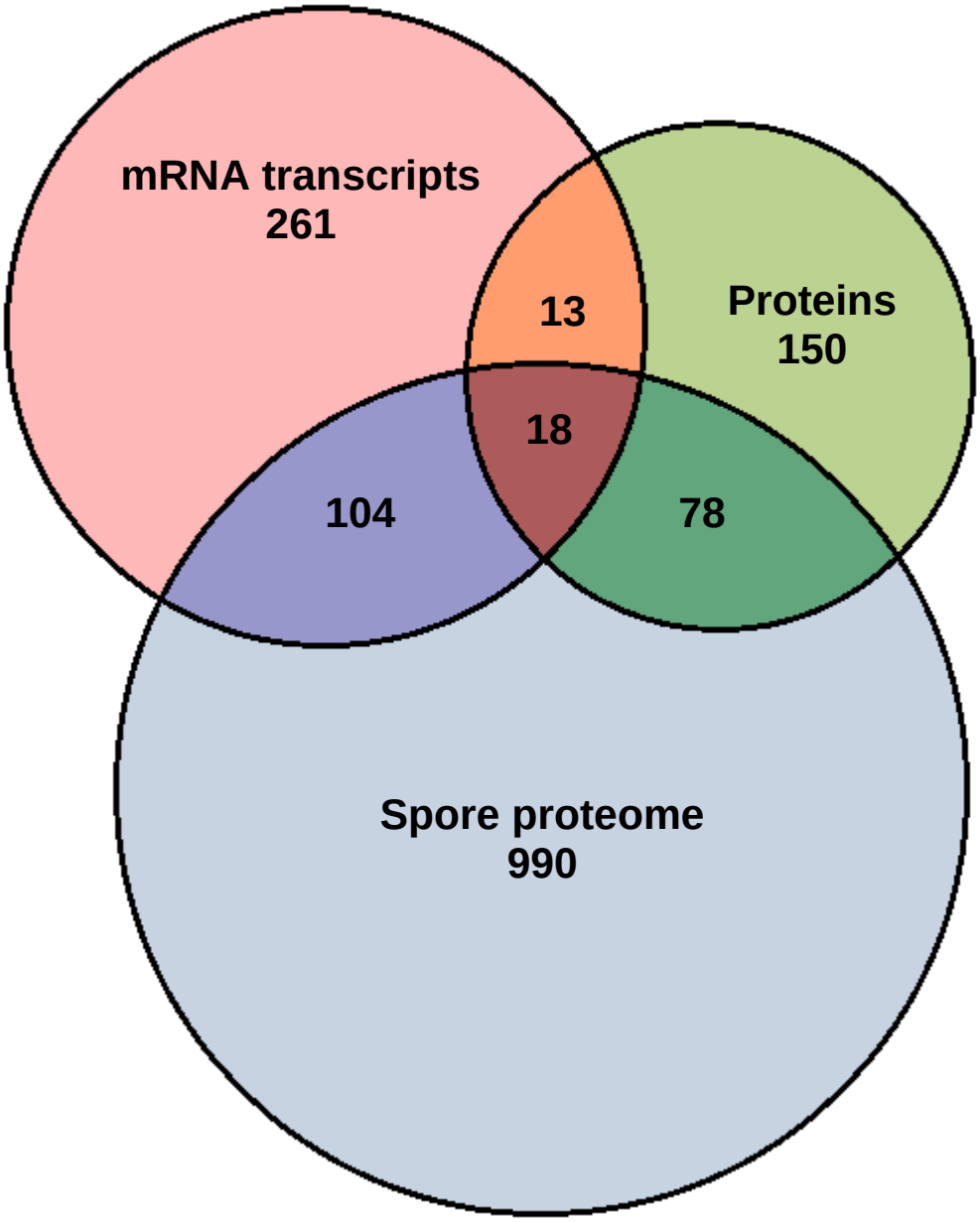


Figure 4.12. Venn diagram demonstrating the correlation between the Spo0A-regulated transcriptome and proteome of *C. difficile* 630Δ*erm*. 261 mRNA transcripts and 150 proteins under Spo0A regulation were identified, of which 31 "core" overlapping genes were detected. 18 of these "core" genes were also found to be present in the spore proteome, which is shown for reference.

*4.3.3.2.1 Proteins up-regulated in C. difficile 630Δ*erm*Δ*spo0A*

86 proteins were found to be negatively regulated by Spo0A, that is, their expression was increased in 630Δ*erm*Δ*spo0A*. Of these 86, 17 (~20%) were predicted to belong to the central/intermediate metabolism class of proteins. These include numerous metabolic enzymes such as AroK (shikimate kinase), CoaD (phosphopantetheine adenylyltransferase), Cmk (cytidylate kinase), Gmk (guanylate kinase) and LplA (lipoate-protein ligase), of which the latter was also significantly up-regulated as mRNA.

In addition, several amino group metabolism proteins were up-regulated in 630Δ*erm*Δ*spo0A*, such as GabT (4-aminobutyrate aminotransferase), NanA (N-acetylneuraminate lyase), ProC1 and ProC2 (pyrroline-5-carboxylate reductases). ProC1 was also negatively regulated by Spo0A at the transcript level, and is present in the spore proteome. The glycolytic enzyme GapA (glyceraldehyde-3-phosphate dehydrogenase) was also up-regulated in 630Δ*erm*Δ*spo0A*. This was unexpected since several genes involved in glycolysis were down-regulated in 630Δ*erm*Δ*spo0A* as transcripts (discussed earlier). In addition, FliM (flagellar motor switch protein) and FleN (flagellar number regulator) were up-regulated.

In a similar trend to the transcriptome data, enzymes which function in energy metabolism such as Hcp (hydroxylamine reductase), GrdA and GrdD (glycine/sarcosine/betaine reductases) HymA (iron-only hydrogenase) and oxidoreductases (CD1797 and CD2100) were also up-regulated in 630Δ*erm*Δ*spo0A*. Interestingly, the expression of *CD0065* encoding a NADP-dependent 7-alpha-hydroxysteroid dehydrogenase was significantly increased in

630Δ*erm*Δ*spo0A* as both mRNA and protein, and was also present in the spore proteome. This is thought to function in the catabolism of bile acids (237, 238). Thus although there is a low concordance between the transcriptome and proteome data, both appear to point towards an altered metabolic capacity.

*4.3.3.2.2 Proteins down-regulated in C. difficile 630Δ*erm*Δ*spo0A*

64 proteins were down-regulated in *C. difficile* 630Δ*erm*Δ*spo0A*. Proteins assigned to the surface-associated functional class were well represented and accounted for ~27% of polypeptides positively regulated by Spo0A. They included ABC transporters (CD0873, CD2311, CD2818, CD2989) and cell wall hydrolases (CD0183, CD3567). Several components of the PTS system were also represented (Crr, CD0492, CD3127) including BglF, which was also down-regulated as a transcript in 630Δ*erm*Δ*spo0A*.

Although *sigE*, *sigG* and *sigH* were down-regulated in 630Δ*erm*Δ*spo0A* as transcripts, they were not significantly differentially expressed as proteins. However, SigB (RNA polymerase sigma-B factor), which was not represented as a differentially regulated transcript was down-regulated in 630Δ*erm*Δ*spo0A* as a protein. SpoVG (stage V sporulation protein G) was the only known sporulation gene found to be potentially positively regulated by Spo0A at the protein level, despite several others being positively regulated as transcripts (see section *4.3.3.2.2*). SspA (small acid-soluble spore protein A) was also down-regulated in 630Δ*erm*Δ*spo0A*. This was logical since 630Δ*erm*Δ*spo0A* does not make spores and SASPs

are abundant in the spore core and function to protect DNA during periods of dormancy (157, 160, 161) (see Chapter 1.4.2).

4.3.3.2.3 Identification of "core" Spo0A-regulated genes

Despite the modest correlation between our transcriptome and proteome datasets, we were able to identify a set of 31 "core" genes that are differentially regulated by Spo0A at both the RNA and protein levels (Figure 4.12). These "core" genes are listed in Table 4.1. Based on our extensive transcriptome and proteome data, we decided to focus on two phenotypes associated with Spo0A:

1. Butyrate production
2. Flagellar assembly

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

Identifier (gene name)	Product	Regulation in 630Δ*erm*Δ*spo0A*
CD0065*	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase	Up
CD0142*	Putative RNA-binding protein	Down
CD0208*	PTS system, IIb component	Up
CD0209*	Putative sugar-phosphate kinase	Up
CD0242	Conserved hypothetical protein	Up
CD0440	Cell surface protein	Down
CD0865*	Conserved hypothetical protein	Down as mRNA but up as protein
CD1054 (Bcd2)*	Butyryl-CoA dehydrogenase	Down
CD1055 (EtfB2)*	Electron transfer flavoprotein beta-subunit	Down
CD1056 (EtfA2)*	Electron transfer flavoprotein alpha-subunit	Down
CD1057 (Crt2)*	3-hydroxybutyryl-CoA dehydratase	Down
CD1058 (Hbd)*	3-hydroxybutyryl-CoA dehydrogenase	Down
CD1059 (ThlA1)*	Acetyl-CoA acetyltransferase	Down
CD1484 (SsuA)*	Putative aliphatic sulfonates ABC transporter, substrate-binding lipoprotein	Down
CD1494	Putative transcriptional regulator	Up
CD1495 (ProC1)*	Pyrroline-5-carboxylate reductase	Up

CD1522	Putative polysaccharide deacetylase	Up
CD1654 (LplA)	Putative lipoate-protein ligase	Up
CD1717	Conserved hypothetical protein	Up
CD1797*	Putative pyridine nucleotide-disulfide oxidoreductase	Up
CD1930	Putative exported protein	Up
CD2010 (Erm1(B))	rRNA adenine N-6-methyltransferase (erythromycin resistance)	Up
CD2127	Putative membrane protein	Down
CD2181*	Putative aromatic compounds hydrolase	Down
CD2195*	Ferritin	Down
CD2765	Putative penicillin-binding protein repressor	Up
CD2796	Cell surface protein	Up
CD2797	Putative exported protein	Up
CD3116 (BglF)	PTS system, beta-glucoside-specific IIabc component	Down
CD3137*	PTS system, IIabc component	Down
CD3555*	Conserved hypothetical protein	Up

* CDS that is also present in the spore proteome.

Note that CD0865* is up-regulated as a protein, but down-regulated as a transcript. Grey shaded area represents an entire operon (CD1054-CD1059) involved in butyrate production that is down-regulated in 630Δ*erm*Δ*spo0A*.

4.3.4 Spo0A positively regulates butyrate production in *C. difficile* 630Δ*erm*

Butyrate, a short-chain fatty acid (SCFA) is the product of the fermentative metabolism by anaerobic bacteria in the colon (239). The role of butyrate in colonic health is established, and it is well recognised that butyrate serves as an important substrate for colonocyte metabolism (239). The production of butyrate is a common feature of anaerobic, Gram positive, colon-dwelling bacteria, but is perhaps best studied in the solventogenic clostridia, such as *C. acetobutylicum* and *C. butyricum* (240, 241).

In *C. difficile* 630, butyrate production is encoded by a 6 kbp polycistronic operon (Figure 4.13). This encodes butyryl-CoA dehydrogenase (*bcd2*), electron transfer flavoprotein β-subunit (*etfB2*), electron transfer flavoprotein α-subunit (*etfA2*), 3-hydroxybutyryl-CoA dehydratase (*crt2)*, 3-hydroxybutyryl-CoA dehydrogenase (*hbd)* and acetyl-CoA acetyltransferase (*thlA1)*. Our analyses have indicated that this entire operon (*bcd2, etfB2, etfA2, crt2, hbd, thlA1*) was significantly down-regulated in 630Δ*erm*Δ*spo0A*, at both the RNA and protein levels (Table 4.1). Figure 4.13 describes this operon and its positive regulation by Spo0A. We were unable to identify an '0A' box upstream of the butyrate operon, implying that its regulation is most likely indirect via an intermediate mediator. In addition, *in vitro* binding assays did not indicate that any genes in this operon were direct targets of Spo0A (Wiep Klass, personal communication). Figure 4.14 illustrates the principle components involved in butyrate metabolism, overlaid with the *C. difficile* Spo0A regulated

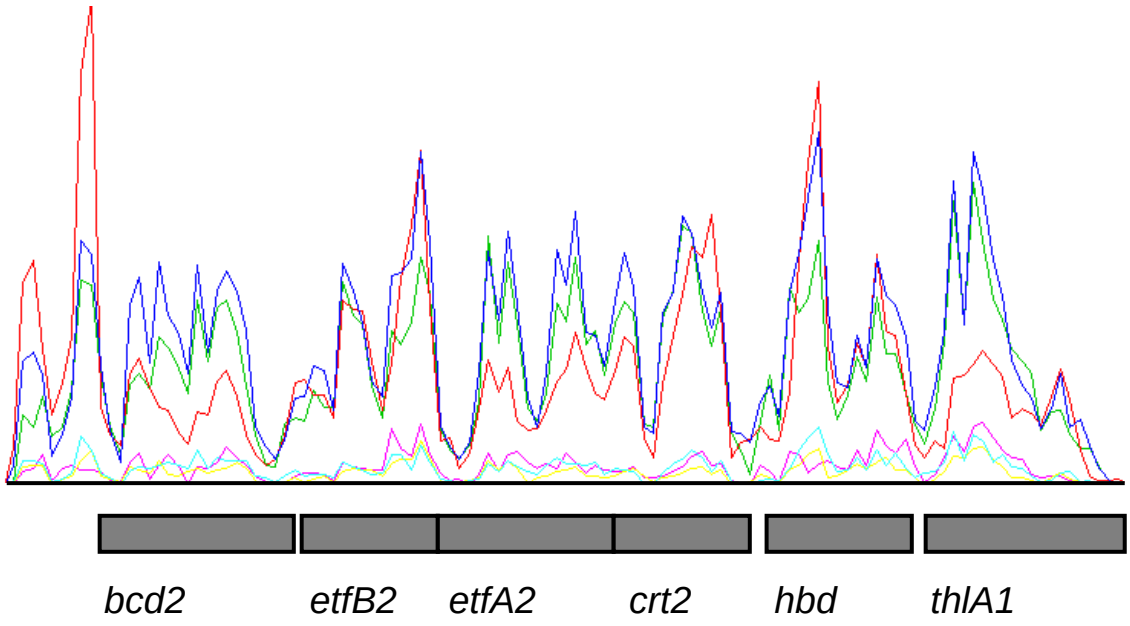


Figure 4.13. Linear representation of the *C. difficile* 630 genome with transcriptome reads mapping to a region involved in butyrate production. The 6 kbp region comprises a polycistronic, six gene operon (*CD1054-CD1059*) involved in butyrate metabolism in *C. difficile* 630. *bcd2*, butyryl-CoA dehydrogenase; *etfB2*, electron transfer flavoprotein beta-subunit; *etfA2*, electron transfer flavoprotein alpha-subunit; *crt2*, 3-hydroxybutyryl-CoA dehydratase; *hbd*, 3-hydroxybutyryl-CoA dehydrogenase; *thlA1*, acetyl-CoA acetyltransferase. Coloured lines represent transcriptome reads aligned to the 630 genomic sequence, and are coloured to represent biological replicates as follows: red, blue, green = *C. difficile* 630Δ*erm*; purple, turquoise, yellow = *C. difficile* 630Δ*erm*Δ*spo0A*. Expression of this operon was significantly attenuated in *C. difficile* 630Δ*erm*Δ*spo0A* at both the mRNA and protein levels. All proteins were also identified in the *C. difficile* 630Δ*erm* proteome.

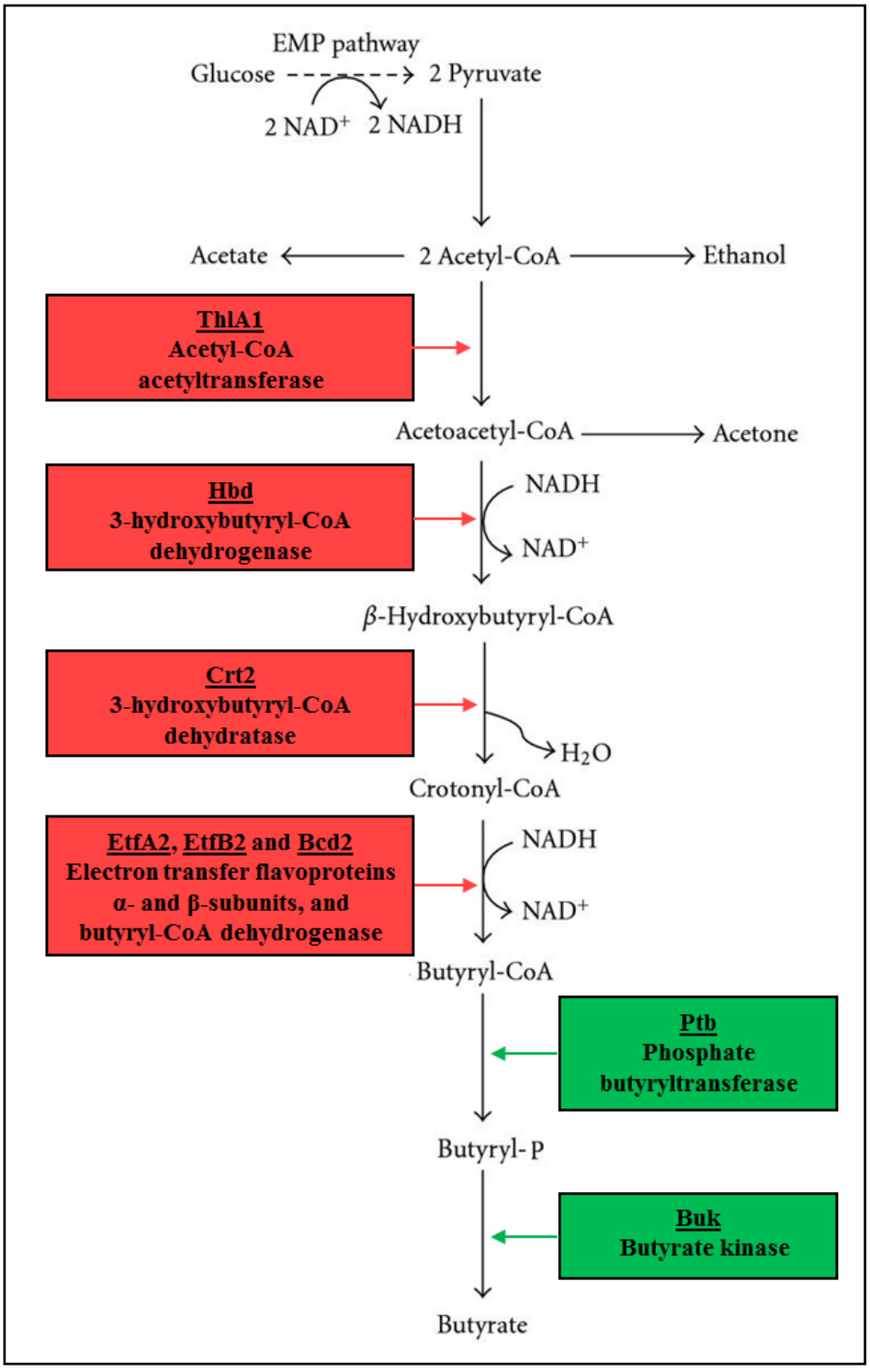


Figure 4.14. Metabolic pathway of butyrate biosynthesis. The pathway is based on data from *C. acetobutylicum* and *C. butyricum.* Components are coloured as follows: red, enzymes that are positively regulated by *C. difficile* Spo0A at both the transcriptional and translational levels; green, enzymes negatively regulated by *C. difficile* Spo0A at the transcriptional level. Adapted from: (240, 241).

enzymes. Interestingly, *ptb* and *buk*, encoding a phosphate butyryltransferase and a butyrate kinase, respectively, were up-regulated in 630Δ*erm*Δ*spo0A* at the mRNA level.

Proteins encoded by this operon are also present in the *C. difficile* 630Δ*erm* spore proteome, perhaps suggesting a possible association between butyrate and spore formation and/or germination in *C. difficile* 630Δ*erm*. Indeed, it has been previously reported that butyrate is a spore lipid precursor and that it's fermentation is linked to anaerobic sporulation (242). Similarly, the incorporation of butyrate into germinating spores of *B. thuringiensis* has been described (243).

4.3.4.1 Spo0A increases butyrate production in *C. difficile*

Given that our molecular data appeared to suggest that Spo0A positively regulates butyrate biosynthesis in *C. difficile* 630Δ*erm*, we decided to quantify butyrate taken from exponentially growing *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*.

We found that *C. difficile* 630Δ*erm*Δ*spo0A* produced significantly less butyrate than the parental 630Δ*erm* strain (P = 0.0005; Figure. 4.15). Moreover, the 630Δ*erm*Δ*spo0A* complemented mutant produced levels of butyrate that were statistically comparable with the 630Δ*erm* parental strain (P = 0.6; Figure 4.15). Our phenotypic and genotypic data are concordant with previous studies, which have demonstrated that asporogenic *C. botulinum* do not produce butyrate (242).

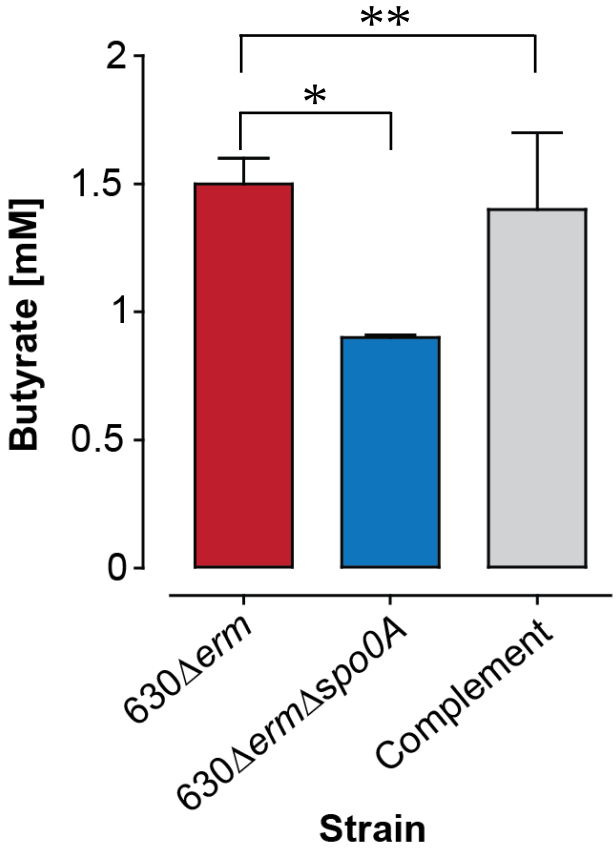


Figure 4.15. Butyrate quantification formed by *C. difficile* 630Δ*erm* derivatives. Cultures of *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A* were grown in Wilson's broth under anaerobic conditions, and samples were taken during mid-exponential growth. The level of butyrate (mM) in the culture supernatants was then quantified via capillary gas chromatography. **P* = 0.0005; ***P* = 0.6.

4.3.5 Spo0A negatively regulates components of the *C. difficile* flagellar assembly apparatus

Although components of the flagellar assembly apparatus did not appear in our list of "core" genes, we did identify several differentially expressed flagella-associated transcripts and proteins, suggesting that this phenotype may be negatively regulated by Spo0A.

In *C. difficile* 630, components of the flagellar assembly apparatus are encoded by two loci that are divided by an inter-flagellar locus that may be involved in flagellin glycosylation (11). In *Bacillus*, motility and sporulation show mutual regulation (176). Furthermore, in *C. acetobutylicum* motility genes were shown to be inhibited by Spo0A (244). Our analyses indicated that seven genes encoding flagellar proteins (*fliC, fliF, fliG, fliH, fliJ, fliK* and *CD0255A*) were up-regulated in 630Δ*erm*Δ*spo0A*. Interestingly, *CD0744* and *CD0745*, both encoding putative chemotaxis proteins, were down-regulated in 630Δ*erm*Δ*spo0A*.

Surprisingly, this was not in concordance with our proteome dataset and none of the differentially regulated transcripts were considered to be differentially regulated proteins, according to our parameters of significance. Nonetheless, all of these transcripts (excluding *fliJ*) were represented in the proteome dataset alongside other flagellar proteins including FliA, FliS1, FliD, FliK and FliN, but at a statistically comparable level to parental *C. difficile* 630Δ*erm*. In addition, FliM, a flagellar motor switch protein and FleN, a flagellar number regulator were both found to be up-regulated in 630Δ*erm*Δ*spo0A* at the protein level, but were not considered to be significantly differentially regulated as transcripts. Figure 4.16 indicates

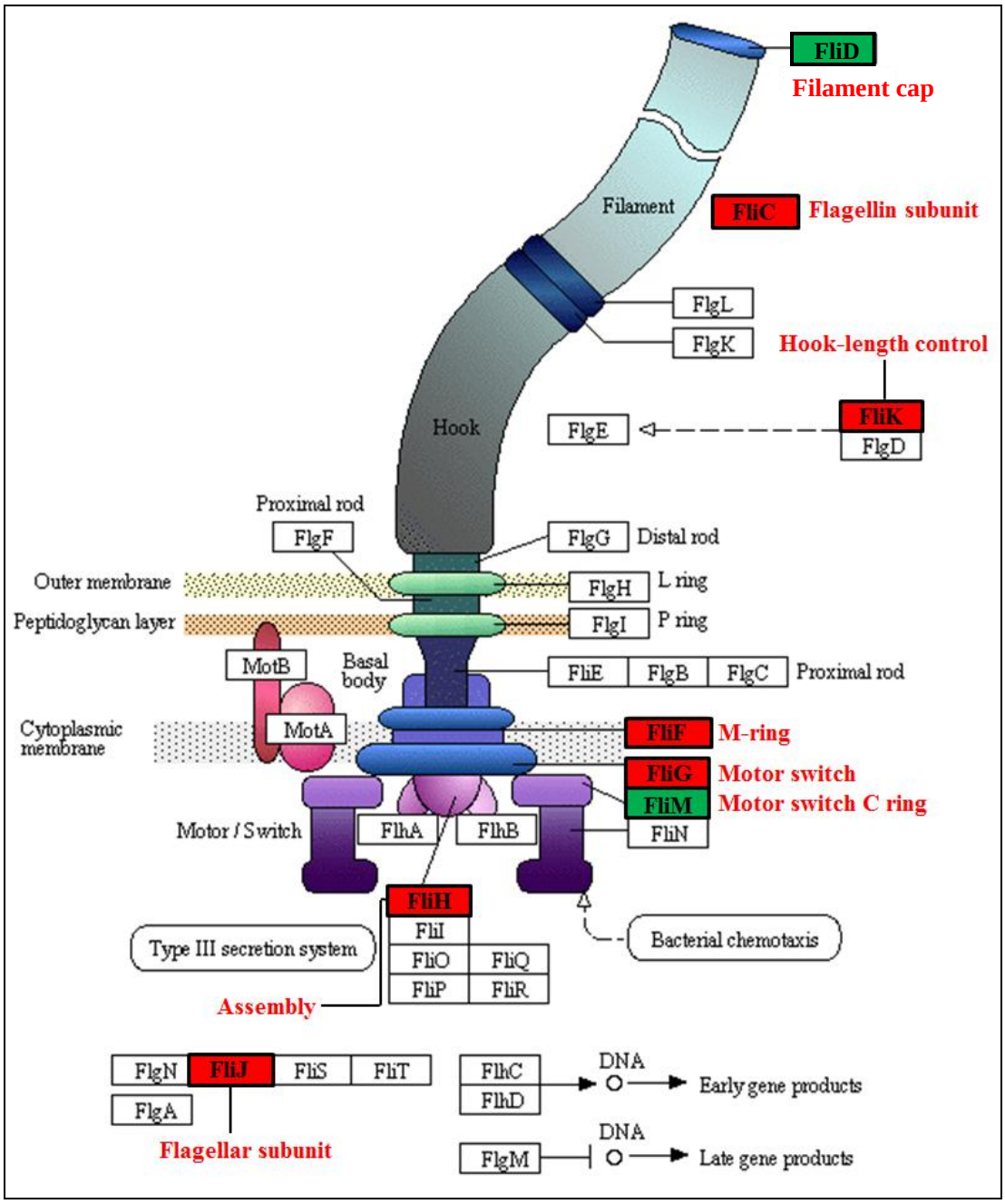


Figure 4.16. Schematic model of the bacterial flagellar assembly apparatus. Overview of the components involved in flagellar production, based on KEGG pathway analysis. CDS highlighted in red indicate those which are significantly up-regulated in *C. difficile* 630Δ*erm* Δ*spo0A* at the mRNA level, those in green indicate significant up-regulation in 630Δ*erm*Δ*spo0A* at the protein level. Note that CD0255A encoding a putative flagellar protein, and FleN encoding a flagellar number regulator are not represented in this model, although they are differentially expressed at the mRNA and protein levels, respectively, in our datasets. Courtesy: http://www.genome.jp/kegg/

the principle components of flagellar assembly, with differentially expressed genes given in red, and differentially expressed proteins given in green.

The expression of *CD0226*, encoding a putative transglycosylase, was up-regulated in 630Δ*erm* Δ*spo0A*, suggesting a possible link between Spo0A and flagellar glycosylation. This idea is further supported by the protein-protein interactions predicted by the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) 9.0 database, as shown in Figure 4.17.

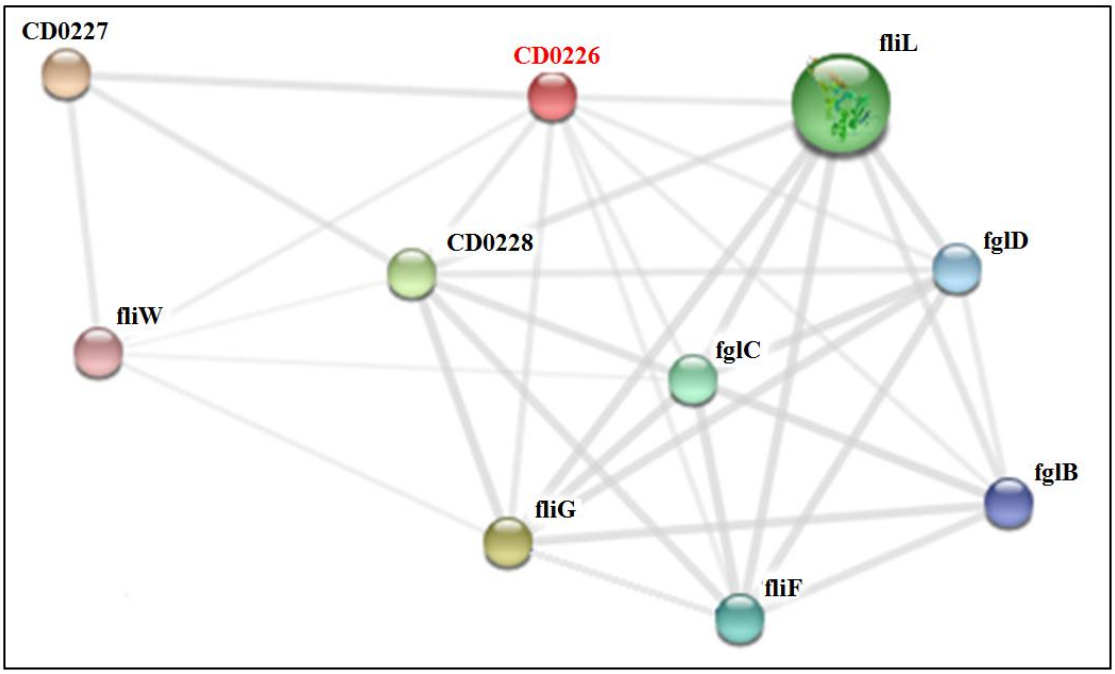


Figure 4.17. Schematic representation of the protein-protein interactions among components of the flagellar apparatus. Interactions are predicted by the STRING 9.0 database. CD0226 (shown in red) encodes a putative transglycosylase predicted to interact with components of the flagellar assembly apparatus that are significantly up-regulated in 630Δ*erm*Δ*spo0A*. Courtesy: http://string-db.org/

Furthermore, the four transcripts (*CD0241-CD0244*) that constitute the inter-flagellar glycosylation locus (11, 245) were also up-regulated in 630Δ*erm*Δ*spo0A* at the mRNA level. Three of these (CD0241, CD0243 and CD0244) were present in the proteome but were not considered to be significantly differentially regulated by Spo0A, whereas CD0242 was deemed be negatively regulated by Spo0A at both the mRNA *and* protein levels. This again indicates some negative association between Spo0A and flagellar glycosylation.

4.3.5.1 Spo0A represses *C. difficile* flagellum assembly

Given the debate surrounding the role of Spo0A in the regulation of motility, we analysed *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A* for the presence of flagella by negative staining and TEM. We also examined *C. difficile* R20291 and R20291Δ*spo0A* in the same manner, since previous publications have suggested that 630 derivatives are inherently less motile than ribotype 027 strains (11).

We found that *C. difficile* has peritrichous flagella, which is consistent with previous reports. However, there were notable differences between the parent and *spo0A* mutant derivatives, in both 012 (630Δ*erm*) and 027 (R20291) ribotypes. TEM negative stains are shown in Figure 4.18. Most striking was the increased number of flagella in the 630Δ*erm*Δ*spo0A* and R20291Δ*spo0A* mutant derivatives in comparison with their respective parental strains (Table 4.2). In fact, in *C. difficile* 630Δ*erm* images, no cell anchored flagella were visible at all (this observation was consistent over three biological replicates).

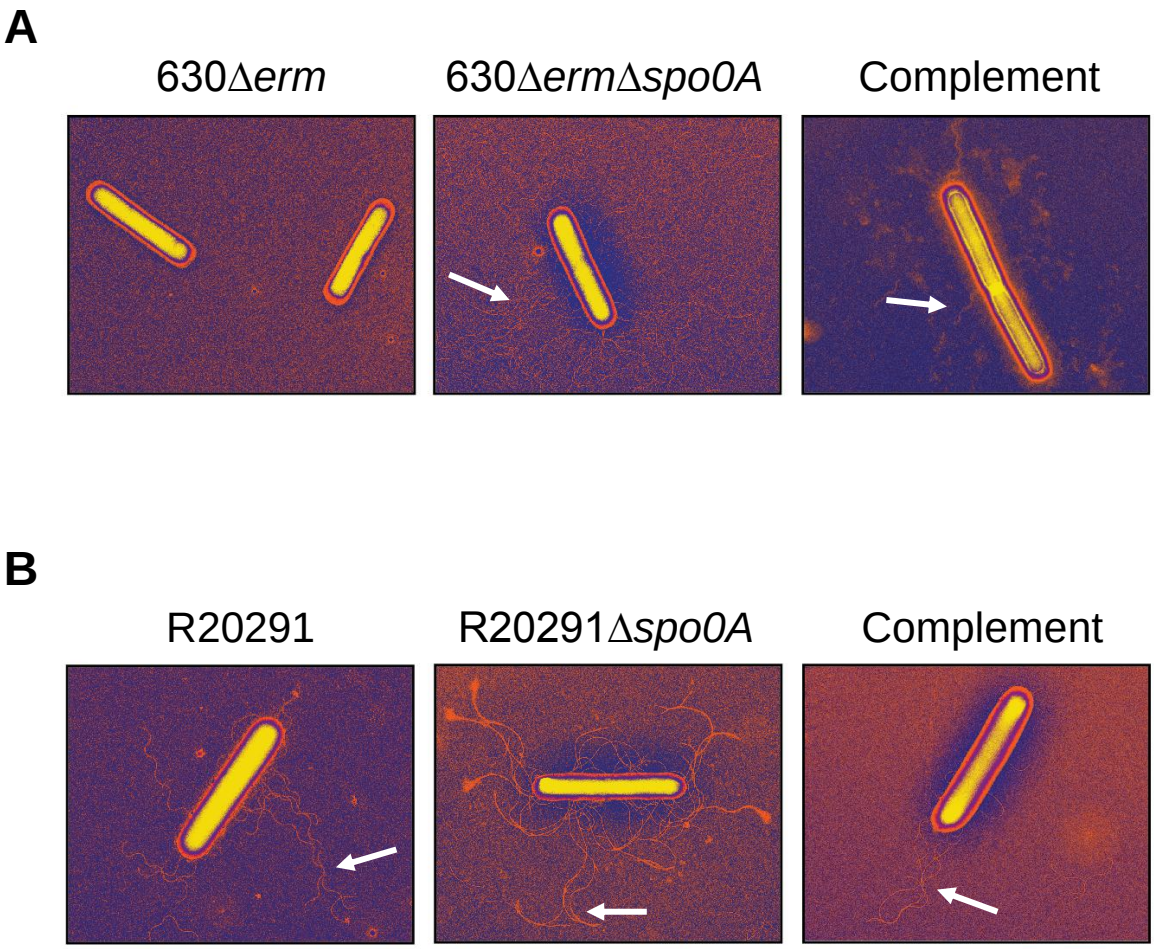


Figure 4.18. Spo0A is a negative regulator of flagellar synthesis. Representative Transmission Electron Micrographs of negatively stained *C. difficile* 630Δ*erm* **(A)** and R20291 **(B)** derivatives. Genetic complementation of the *spo0A* mutation restored flagella synthesis to levels comparable with the parental strains. Note that we were consistently unable to detect any cell anchored flagella in *C. difficile* 630Δ*erm*. Images taken by David Goulding, WTSI.

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

	Flagellar abundance* (average number/cell)	**Flagellar length (average µm)**	**Flagellar diameter (average nm)**
Ribotype 012			
630Δ*erm*	0	Not determined	Not determined
630Δ*erm*Δ*spo0A*	8.1	8.8	15.2
630Δ*erm* complement	4.8	3.5	14.6
Ribotype 027			
R20291	4.6	5.3	15.1
R20291Δ*spo0A*	12.6	5.6	13.2
R20291 complement	6.2	6.8	12.0

* Note that only flagella which appeared to be anchored to the cell surface were recorded.

Increased levels of the flagella number regulator, FleN, at the protein level may account for the overproduction of flagella in 630Δ*erm*Δ*spo0A*. Since we could not detect any flagella on

the parental 630Δ*erm* strain, we were unable to draw conclusions regarding the role of Spo0A in flagellar length and diameter regulation in this genetic background.

In the R20291 background, there were almost three times as many flagella in the Spo0A mutant (average of 12.6) compared to its parent R20291 strain (average of 4.6). However, there were no notable differences in flagellar length or diameter, as described in Table 4.2. In both *C difficile* 630Δ*erm* and *C. difficile* R20291, the complemented *spo0A* mutants appeared to have a flagellar phenotype that most closely resembled its respective parental strain (Figure 4.18 and Table 4.2).

In addition, we performed a Western blot with FliC (significantly regulated by Spo0A at the mRNA level) and FliD (significantly regulated by Spo0A at the protein level) antibodies. Samples were taken at mid-exponential phase, in order to give the highest level of concordance between the transcriptome and proteome datasets. We found that *C. difficile* 630Δ*erm* produced lower levels of both FliC and FliD compared to the 630Δ*erm*Δ*spo0A* mutant derivative (Figure 4.19). Somewhat paradoxically, the 630Δ*erm*Δ*spo0A* complemented strain produced levels of FliC and FliD that were more comparable with the 630Δ*erm*Δ*spo0A* mutant (Figure 4.19), despite producing a reduced number of flagella as observed via TEM (Figure 4.18). However, given that the Western blot was performed on whole cell lysates of the *C. difficile* cultures, it is possible that these proteins may have been present intracellularly in the complemented strain, rather than being expressed as functional components of the *C. difficile* flagellar apparatus on the cell surface. Furthermore, it may be that glycosylation plays an important role in the activity of these proteins.

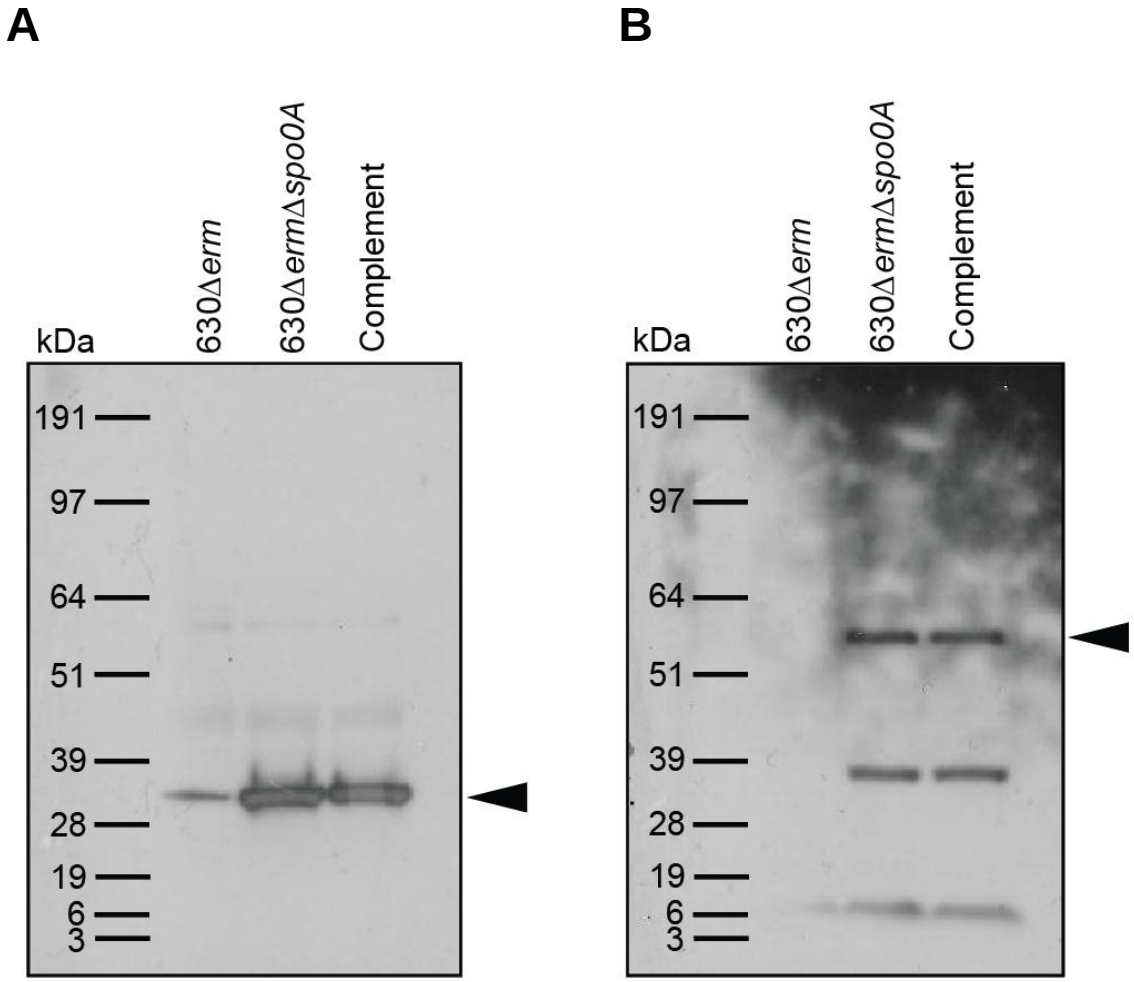


Figure 4.19. Western Blot analysis of flagellar gene expression in *C. difficile* 630Δ*erm* derivatives. *C. difficile* cultures were grown in Wilson's broth under anaerobic conditions, and samples were taken during mid-exponential growth. Extracted proteins with an equivalency to ~10^7 cells/lane were probed with **(A)** anti-*C. difficile* FliC and **(B)** anti-*C. difficile* FliD antibodies. Arrow heads indicate the approximate predicted molecular masses of FliC (30.9 kDa) and FliD (56 kDa).

4.4 Discussion

We previously demonstrated that the *C. difficile spo0A* gene is a persistence and transmission factor *in vivo* (Chapter 3). The availability of a defined *spo0A* mutant facilitated studies on the Spo0A regulon and its role in growth, sporulation and pathogenesis, at the whole genome level. Here, using a combination of transcriptomic and proteomic approaches, we demonstrate that the *spo0A* gene of *C. difficile* 630Δ*erm* plays a highly pleiotropic role and is a global transcriptional regulator that coordinates multiple virulence, sporulation and metabolic phenotypes.

Importantly, we also validated Spo0A as a transcriptional regulator of a number of sporulation genes, which are often demarked by canonical Spo0A binding sites, as shown in Figure 4.20. Furthermore, we confirmed that Spo0A negatively regulates toxin production in *C. difficile* 630Δ*erm* at the transcriptional level. In addition, we identified novel phenotypes associated with the Spo0A regulon. For example, we demonstrated that Spo0A negatively regulates key components of the *C. difficile* flagellar assembly apparatus and positively regulates several metabolic pathways, including the fermentation of carbohydrates leading to the production of butyrate. These phenotypes may have implications relating to the colonisation and persistence of *C. difficile* 630Δ*erm in vivo*.

The role of flagella in the colonisation and pathogenesis of several enteropathogens, including *V. cholera* (246), *Campylobacter jejuni* (247) and *Helicobacter pylori* (248) is well recognised. Additionally, recent work in *C. difficile* has described the requirement of

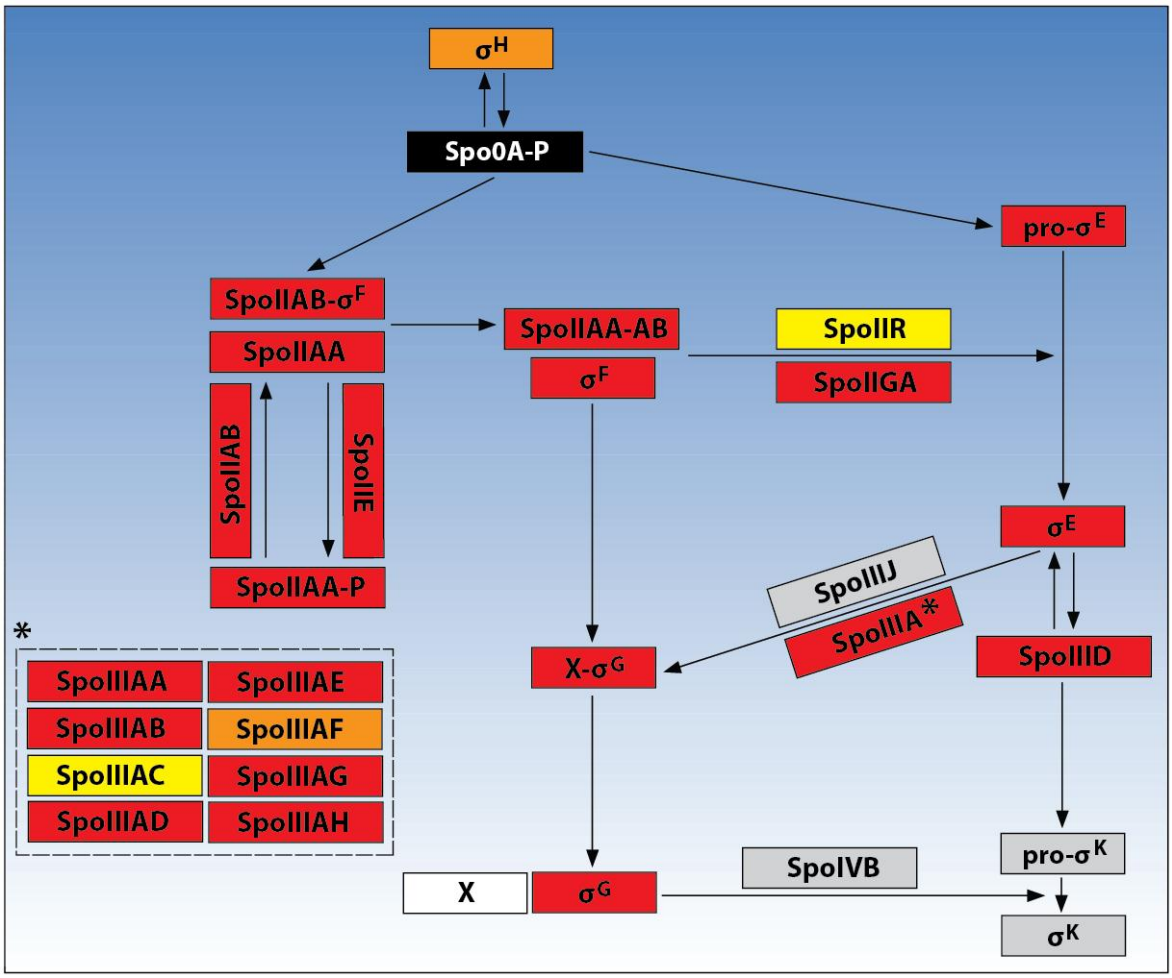


Figure 4.20. Schematic representation of the *C. difficile*-specific sporulation cascade, indicating genes belonging to the Spo0A regulon at the transcriptional level. Sporulation-associated genes belonging to the Spo0A regulon are colour coded as follows: red, positively regulated by Spo0A at $P = \leq 0.01$; orange, positively regulated by Spo0A at $P = \leq 0.05$; yellow, positively regulated by Spo0A at $P = \leq 0.08$); grey, not significantly regulated by Spo0A according to our parameters of significance; white, unknown factor. Generated from: (176, 249)

glycosylation in flagellar assembly and function (245). We found that Spo0A represses flagellum assembly and that *C. difficile* 630Δ*erm*Δ*spo0A* derivatives produce increased levels of FliC, FliD and genes encoding putative transglycosylases compared to parental *C. difficile* 630Δ*erm*, suggesting a possible link between Spo0A and flagellar assembly and glycosylation in *C. difficile* 630Δ*erm*.

It is possible that the enhanced flagellar phenotype of *C. difficile* 630Δ*erm*Δ*spo0A* may contribute to the inability of this strain to cause a long-term persistent infection in our murine model of infection (as discussed in Chapter 3). For example, it has been demonstrated that for *V. cholerae* to cause a persistent, long-term infection, the *loss* of motility is important (though flagella are important for the initial infection) (250, 251). Additionally, flagellin proteins are recognised by Toll-like receptor (TLR) 5, which acts as a receptor that can stimulate induction of host innate immunity (252, 253). Thus, the loss of flagella may be linked to a mechanism of immune subversion. Given that the role of flagella in *C. difficile* pathogenesis is currently unclear, we are attempting to make mutations in the *fliC* (flagellin subunit) and *fliD* (filament cap) genes of *C. difficile*. Using our mouse model of infection, we will then determine whether the repression of flagellar biosynthesis by *C. difficile* is required for long-term persistence, and whether this is associated with virulence *in vivo*.

The role of butyrate in human colonic health is well established (239). However, the role of butyrate production in the pathogenesis of human-virulent *C. difficile* is currently unknown. We noted that *C. difficile* 630Δ*erm*Δ*spo0A* produced significantly less butyrate than its parental strain, that is, Spo0A positively regulates butyrate production in 630Δ*erm.* Given the

role of butyrate in human colonic health, we hypothesise that the production of this SCFA by 630Δ*erm* may facilitate the establishment of a favourable environment to initiate an infection. For example, there is increasing evidence that butyrate has anti-inflammatory properties on the host (254, 255). Following colonisation, Spo0A may then function to increase the fermentation of glucose into SCFAs (such as butyrate) which concomitantly results in a de-repression of TcdA and TcdB production (glucose represses toxin synthesis). Clearly, however, this hypothesis would need to be explored further. In addition, we are currently profiling other SCFAs and alcohols from exponentially growing *C. difficile* to determine whether 630Δ*erm*Δ*spo0A* shifts its fermentative metabolism towards the production of other end-products (such as ethanol or acetate), whilst limiting its net carbon flow to butyrate. The generation of mutants in the *thlA1* and *hbd* genes of the butyrate biosynthesis operon is currently underway, and should facilitate such studies.

The degree of concordance between the Spo0A-regulated transcriptome and proteome of *C. difficile* 630Δ*erm* was low. Indeed, only a small subset of CDSs were found to be differentially regulated by Spo0A at both the transcriptional *and* translational levels. This disconnect is likely the result of an interplay between 2 major factors: (i) inherent biological variation in mRNA to protein correlation, and (ii) the technical limitations of our assay. Firstly, previous publications have reported that transcripts and proteins only demonstrate a weak positive correlation, and that the underlying principle that mRNA abundance accurately informs protein expression may not hold true in all biological systems (256, 257). Moreover, molecular events such as the efficiency of translation or protein degradation can notably alter protein levels independently of mRNA expression (236).

Secondly, the lack of correlation between the profiles may be explained somewhat by the technical limitations of our assay. For example, the excision of protein bands is a potential source of error prior to LC-MS/MS analysis. Thus, the use of a 2D gel electrophoresis approach may have increased the number of proteins recovered. In addition, the proteomic data presented in this thesis does not include complementary data on the proteins secreted by *C. difficile*. An attempt was made to catalogue the secreted proteins (secretome) of exponentially growing *C. difficile* 630Δ*erm* and 630Δ*erm*Δ*spo0A*, though this fraction was regrettably contaminated during LC- MS/MS analysis and was unavailable for comparative analyses. This may explain, for example, why TcdA (a secreted toxin) was significantly up-regulated as a transcript, but not as a protein according to LC-MS/MS analysis. Furthermore, given that secreted proteins can mediate important host - pathogen interactions *in vivo*, and may also represent novel candidate virulence factors, it would be advantageous to have a full proteomic (cellular *and* secreted) description of *C. difficile* during exponential growth.

In conclusion, we have demonstrated that Spo0A regulates multiple virulence, sporulation and metabolic phenotypes in *C. difficile* 630Δ*erm*, which informs on the intricate pleiotropy of a *spo0A* mutation. The phenotypes described in this chapter (butyrate and flagellar biosynthesis) may have important implications relating to *C. difficile* colonisation and persistence *in vivo*, and will inform the direction of future research. In addition, future experiments that involve blocking spore formation at a later stage in the sporulation cascade may reduce the number of pleiotropic effects in comparison to those observed for the *spo0A* mutation, which may further define the events that occur during the course of spore formation in *C. difficile*.

5 Final discussion

The work presented in this thesis represents an attempt to define the *C. difficile spo0A* regulon, and to underpin its role in disease and transmission. We took a two-step approach to this end. Firstly, we exploited a murine model of infection that mimics several aspects of *C. difficile* disease, persistence and transmission in humans, thus enabling us to assess the role of the *spo0A* gene in *C. difficile* pathogenicity. Secondly, we adopted an integrated transcriptomic and proteomic approach in order to provide a molecular description of the *C. difficile* Spo0A regulon at the whole genome level, and link this back to any role in disease and transmission.

Data from this study have demonstrated that the *C. difficile spo0A* gene is a highly pleiotropic global transcriptional regulator that coordinates multiple virulence, sporulation and metabolic phenotypes during *C. difficile* disease and transmission. A summary of the phenotypes under Spo0A regulation is given in Figure 5.1.

5.1 A murine model of *C. difficile* infection: insights into the role of Spo0A in persistence and transmission

The availability of a murine infection model opens a tractable route to study the genetic basis of *C. difficile* disease, relapse, persistence and transmission. In the work described in Chapter 3, we demonstrate for the first time that the *spo0A* gene of clinically relevant *C. difficile* ribotypes is essential for persistent infection and efficient host-to-host transmission in mice. Importantly, we also show that a functional *spo0A* gene is required for relapsing infection

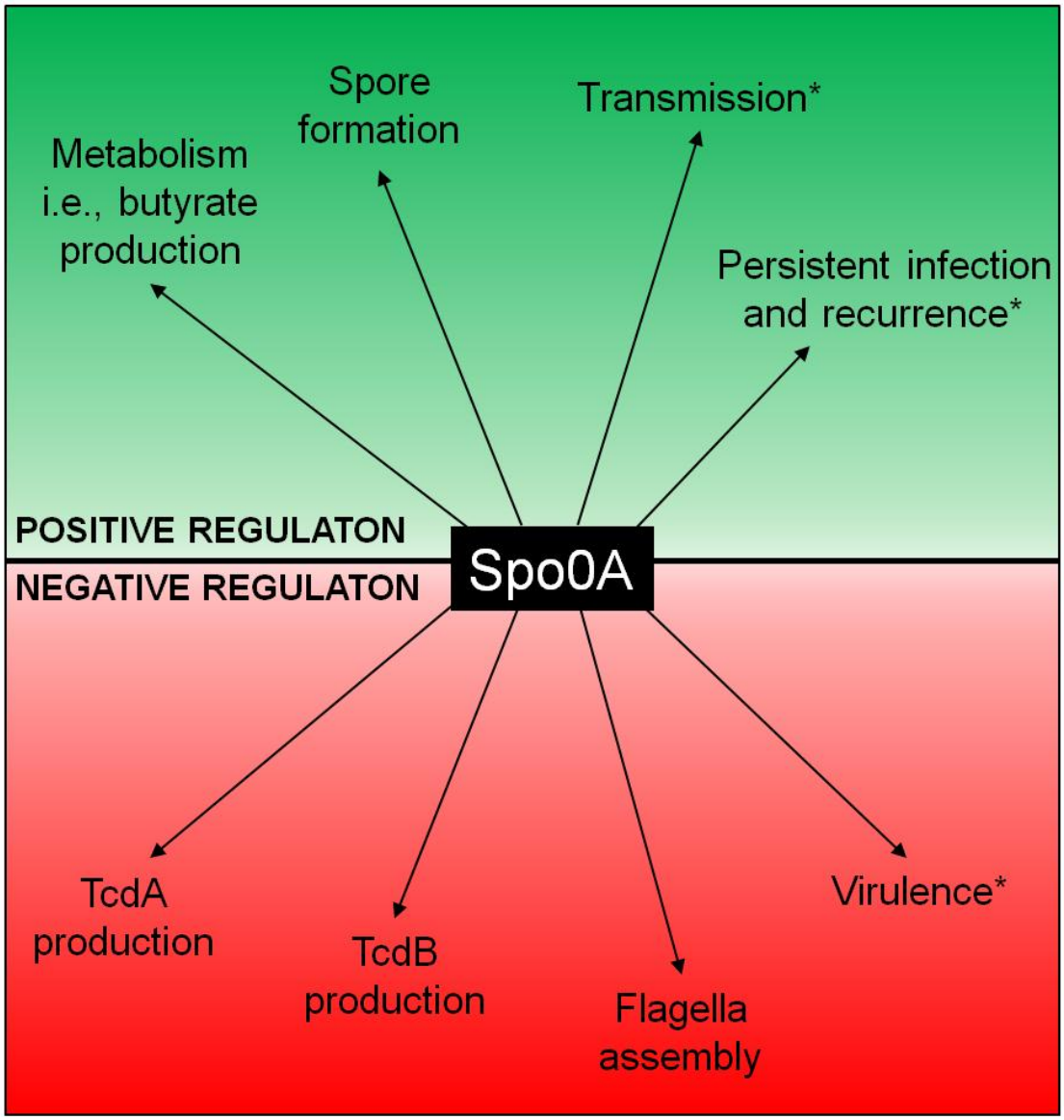


Figure 5.1. Summary of the phenotypes associated with the *C. difficile* Spo0A regulon. *C. difficile* Spo0A is a highly pleiotropic global transcriptional regulator that coordinates multiple phenotypes associated with metabolism, disease, persistence and transmission in mice. Phenotypes were identified using both the mouse model of *C. difficile* infection and transcriptomic and proteomic approaches. *According to our murine model of *C. difficile* infection.

after vancomycin therapy, and we further illustrate that the local environment may serve as a reservoir of *C. difficile* infection. These observations have potential practical implications relating to the management of hospital patients.

Additionally, we provide evidence that *C. difficile* Spo0A mutant derivatives produce higher levels of the toxins TcdA and TcdB *in vitro* compared to the parental equivalents, which was associated with enhanced virulence *in vivo*. Thus, Spo0A may negatively regulate *C. difficile* toxin production. Although this finding is potentially contradictory to the observations of Underwood *et al*. (2009), we believe that our approach was comparatively more robust and quantitative. Additionally, our observed phenotype was supported by the transcriptional data described in Chapter 4.

5.2 Insights from an "omics" approach to studying the *C. difficile* Spo0A regulon

The work described in Chapter 4 exploits a combined transcriptomic and proteomic approach to define the *C. difficile spo0A* regulon at the whole genome level. Such an integrated approach allowed us to derive a holistic picture of *C. difficile* biology that could not have been readily achieved using each approach independently and identified multiple novel phenotypes associated with the Spo0A regulon. We have determined that Spo0A is a negative regulator of *C. difficile* flagellar assembly and a positive regulator of several metabolic pathways, including the fermentation of carbohydrates leading to the production of the SCFA butyrate.

Importantly, we have also validated Spo0A as a transcriptional regulator of a number of sporulation genes, and confirm that Spo0A negatively regulates toxin production in *C. difficile* 630Δ*erm* at the transcriptional level.

5.3 Final conclusions and future directions

C. difficile is an enteropathogen that in the past two decades has emerged from relative obscurity to become a dominant healthcare-associated pathogen. Human-virulent ribotypes continue to disseminate globally and appear to be associated with increased incidence and severity, representing a major medical burden and significant economic drain on finite resources.

In the present study, we have demonstrated that Spo0A regulates multiple *C. difficile* phenotypes associated with disease, persistence and transmission in mice. However, many questions still remain and should inform the direction of future research. For example, due to the complex pleiotropy of the *spo0A* mutation, we were unable to provide a full molecular description of the events that occur during the course of spore formation in *C. difficile*. Work to generate and characterise a *C. difficile sigE* mutant is currently in progress. Such a mutation should also generate an asporogenous phenotype, but *sigE* is active at a later stage in the sporulation cascade and thus the number of pleiotropic effects may be reduced compared to *spo0A*. In addition, we do not at present have data on the repertoire of proteins secreted by

C. difficile. This information may prove useful in identifying novel secreted virulence-associated factors or mediators of host – pathogen interactions.

During this study, we identified novel phenotypes associated with the *C. difficile* Spo0A regulon. For example, we demonstrated that Spo0A negatively regulates several components of the *C. difficile* flagellar assembly apparatus. In addition, we also established Spo0A as a positive regulator of several metabolic pathways, including the fermentation of carbohydrates leading to the production of the SCFA, butyrate. These phenotypes have potential implications relating to the persistence and colonisation of *C. difficile in vivo* (discussed in chapter 4.4). As such, work is currently underway to generate defined insertional mutations in the *fliC* and *fliD* (flagella) genes, as well as the *thlA1* and *hbd* (butyrate biosysthesis) genes of both *C. difficile* R20291 and *C. difficile* 630Δ*erm*. The role of such genes in *C. difficile* disease, persistence and colonisation can then be determined using our mouse model of infection.

References

1. **Riley TV, Bowman RA, Carroll SM.** 1983. Diarrhoea associated with *Clostridium difficile* in a hospital population. *Med J Aust* **1**: 166-9

2. **Kelly CP, LaMont JT.** 2008. *Clostridium difficile*- more difficult than ever. *N Engl J Med* **359**: 1932-40

3. S**tabler RA, Dawson LF, Valiente E, Cairns MD, Martin MJ, Donahue EH, Riley TV, Songer JG, Kuijper EJ, Dingle KE, Wren BW.** 2012. Macro and micro diversity of *Clostridium difficile* isolates from diverse sources and geographical locations. *PLoS One* **7**: e31559

4. **Muto CA, Pokrywka M, Shutt K, Mendelsohn AB, Nouri K, Posey K, Roberts T, Croyle K, Krystofiak S, Patel-Brown S, Pasculle AW, Paterson DL, Saul M, Harrison LH.** 2005. A large outbreak of *Clostridium difficile*-associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. *Infect Control Hosp Epidemiol* **26**: 273-80

5. **Thomas C, Stevenson M, Riley TV.** 2003. Antibiotics and hospital-acquired *Clostridium difficile*-associated diarrhoea: a systematic review. *J Antimicrob Chemother* **51**: 1339-50

6. **O'Toole H.** 1935. Intestinal flora in newborn infants with a description of a new pathogenic anaerobe, *Bacillus difficilis. Am J Dis Child* **49**: 390-402

7. **Valiente E, Dawson LF, Cairns MD, Stabler RA, Wren BW.** 2012. Emergence of new PCR ribotypes from the hypervirulent *Clostridium difficile* 027 lineage. *J Med Microbiol* **61**: 49-56

8. **Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, Frost E, McDonald LC.** 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *Lancet* **366**: 1079-84

9. **Goorhuis A, Van der Kooi T, Vaessen N, Dekker FW, Van den Berg R, Harmanus C, van den Hof S, Notermans DW, Kuijper EJ.** 2007. Spread and epidemiology of *Clostridium difficile* polymerase chain reaction ribotype 027/toxinotype III in The Netherlands. *Clin Infect Dis* **45**: 695-703

10. **Sebaihia M, Wren BW, Mullany P, Fairweather N, Minton NP, Stabler RA, Thomson N, Roberts AP, Cerdeno-Tarraga AM, Wang H, Holden MT, Wright A, Churcher C, Quail MA, 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. *Nat Genet* **38**: 779-86

11. **Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, Lawley TD, Sebaihia M, Quail MA, Rose G, Gerding DN, Gibert M, Popoff MR, Parkhill J, Dougan G, Wren BW.** 2009. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol* **10**: R102

12. **Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeno-Tarraga AM, Wang H, Holden MT, Wright A, Churcher C, Quail MA, 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. *Nat Genet* **38**: 779-86

13. **He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, Holt KE, Seth-Smith HM, Quail MA, Rance R, Brooks K, Churcher C, Harris D, Bentley SD, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren BW, Dougan G, Parkhill J.** 2010. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proc Natl Acad Sci USA* **107**: 7527-32

14. **Zidaric V, Zemljic M, Janezic S, Kocuvan A, Rupnik M.** 2008. High diversity of *Clostridium difficile* genotypes isolated from a single poultry farm producing replacement laying hens. *Anaerobe* **14**: 325-7

15. **Dallal RM, Harbrecht BG, Boujoukas AJ, Sirio CA, Farkas LM, Lee KK, Simmons RL.** 2002. Fulminant *Clostridium difficile*: an underappreciated and increasing cause of death and complications. *Ann Surg* **235**: 363-72

16. **O'Connor JR, Johnson S, Gerding DN.** 2009. *Clostridium difficile* infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology* **136**: 1913-24

17. **Linder JA, Huang ES, Steinman MA, Gonzales R, Stafford RS.** 2005. Fluoroquinolone prescribing in the United States: 1995 to 2002. *Am J Med* **118**: 259-68

18. **Spigaglia P, Carattoli A, Barbanti F, Mastrantonio P.** 2010. Detection of *gyrA* and *gyrB* mutations in *Clostridium difficile* isolates by real-time PCR. *Mol Cell Probes* **24**: 61-7

19. **Kuijper EJ, Coignard B, Brazier JS, Suetens C, Drudy D, Wiuff C, Pituch H, Reichert P, Schneider F, Widmer AF, Olsen KE, Allerberger F, Notermans DW, Barbut F, Delmee M, Wilcox M, Pearson A, Patel BC, Brown DJ, Frei R, Akerlund T, Poxton IR, Tull P.** 2007. Update of *Clostridium difficile*-associated disease due to PCR ribotype 027 in Europe. *Euro Surveill* **12**: E1-2

20. **Kuijper EJ, Coignard B, Tull P.** 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. *Clin Microbiol Infect* **12** Suppl 6: 2-18

21. **Barbut F, Petit JC.** 2001. Epidemiology of *Clostridium difficile*-associated infections. *Clin Microbiol Infect* **7**: 405-10

22. **Gerding DN, Johnson S, Peterson LR, Mulligan ME, Silva J, Jr.** 1995. *Clostridium difficile*-associated diarrhea and colitis. *Infect Control Hosp Epidemiol* **16**: 459-77

23. **Poutanen SM, Simor AE.** 2004. *Clostridium difficile*-associated diarrhea in adults. *CMAJ* **171**: 51-8

24. **Sullivan A, Edlund C, Nord CE.** 2001. Effect of antimicrobial agents on the ecological balance of human microflora. *Lancet Infect Dis* **1**: 101-14

25. **Williams OM, Spencer RC.** 2009. The management of *Clostridium difficile* infection. *Br Med Bull* **91**: 87-110

26. **Borriello SP, Barclay FE.** 1986. An *in vitro* model of colonisation resistance to *Clostridium difficile* infection. *J Med Microbiol* **21**: 299-309

27. **Reeves AE, Theriot CM, Bergin IL, Huffnagle GB, Schloss PD, Young VB.** 2011. The interplay between microbiome dynamics and pathogen dynamics in a murine model of *Clostridium difficile* Infection. *Gut Microbes* **2**: 145-58

28. **Hurley BW, Nguyen CC.** 2002. The spectrum of pseudomembranous enterocolitis and antibiotic-associated diarrhea. *Arch Intern Med* **162**: 2177-84

29. **Wilcox M, Hoy C.** 2001. Broad-spectrum antibiotics in ORACLE. *Lancet* **358**: 503-4

30. **Owens RC, Jr., Donskey CJ, Gaynes RP, Loo VG, Muto CA.** 2008. Antimicrobial-associated risk factors for *Clostridium difficile* infection. *Clin Infect Dis* **46** Suppl 1: S19-31

31. **McCusker ME, Harris AD, Perencevich E, Roghmann MC.** 2003. Fluoroquinolone use and *Clostridium difficile*-associated diarrhea. *Emerg Infect Dis* **9**: 730-3

32. **Pepin J, Saheb N, Coulombe MA, Alary ME, Corriveau MP, Authier S, Leblanc M, Rivard G, Bettez M, Primeau V, Nguyen M, Jacob CE, Lanthier L.** 2005. Emergence of fluoroquinolones as the predominant risk factor for *Clostridium difficile*-associated diarrhea: a cohort study during an epidemic in Quebec. *Clin Infect Dis* **41**: 1254-60

33. **Brazier JS, Fitzgerald TC, Hosein I, Cefai C, Looker N, Walker M, Bushell AC, Rooney P.** 1999. Screening for carriage and nosocomial acquisition of *Clostridium difficile* by culture: a study of 284 admissions of elderly patients to six general hospitals in Wales. *J Hosp Infect* **43**: 317-9

34. **Simor AE, Bradley SF, Strausbaugh LJ, Crossley K, Nicolle LE.** 2002. *Clostridium difficile* in long-term-care facilities for the elderly. *Infect Control Hosp Epidemiol* **23**: 696-703

35. **McFarland LV, Surawicz CM, Stamm WE.** 1990. Risk factors for *Clostridium difficile* carriage and *C. difficile*-associated diarrhea in a cohort of hospitalized patients. *J Infect Dis* **162**: 678-84

36. **Southern WN, Rahmani R, Aroniadis O, Khorshidi I, Thanjan A, Ibrahim C, Brandt LJ.** 2010. Postoperative *Clostridium difficile*-associated diarrhea. *Surgery* **148**: 24-30

37. **Dalton BR, Lye-Maccannell T, Henderson EA, Maccannell DR, Louie TJ.** 2009. Proton pump inhibitors increase significantly the risk of *Clostridium difficile* infection in a low-endemicity, non-outbreak hospital setting. *Aliment Pharmacol Ther* **29**: 626-34

38. **Howell MD, Novack V, Grgurich P, Soulliard D, Novack L, Pencina M, Talmor D.** 2010. Iatrogenic gastric acid suppression and the risk of nosocomial *Clostridium difficile* infection. *Arch Intern Med* **170**: 784-90

39. **Gurian L, Ward TT, Katon RM.** 1982. Possible foodborne transmission in a case of pseudomembranous colitis due to *Clostridium difficile*: influence of gastrointestinal secretions on *Clostridium difficile* infection. *Gastroenterology* **83**: 465-9

40. **Lowe DO, Mamdani MM, Kopp A, Low DE, Juurlink DN.** 2006. Proton pump inhibitors and hospitalization for *Clostridium difficile*-associated disease: a population-based study. *Clin Infect Dis* **43**: 1272-6

41. **Marra AR, Edmond MB, Wenzel RP, Bearman GM.** 2007. Hospital-acquired *Clostridium difficile*-associated disease in the intensive care unit setting: epidemiology, clinical course and outcome. *BMC Infect Dis* **7**: 42

42. **Cunney RJ, Magee C, McNamara E, Smyth EG, Walshe J.** 1998. *Clostridium difficile* colitis associated with chronic renal failure. *Nephrol Dial Transplant* **13**: 2842-6

43. **Bilgrami S, Feingold JM, Dorsky D, Edwards RL, Bona RD, Khan AM, Rodriguez-Pinero F, Clive J, Tutschka PJ.** 1999. Incidence and outcome of *Clostridium difficile* infection following autologous peripheral blood stem cell transplantation. *Bone Marrow Transplant* **23**: 1039-42

44. **Anand A, Glatt AE.** 1993. *Clostridium difficile* infection associated with antineoplastic chemotherapy: a review. *Clin Infect Dis* **17**: 109-13

45. **Issa M, Ananthakrishnan AN, Binion DG.** 2008. *Clostridium difficile* and inflammatory bowel disease. *Inflamm Bowel Dis* **14**: 1432-42

46. **Issa M, Vijayapal A, Graham MB, Beaulieu DB, Otterson MF, Lundeen S, Skaros S, Weber LR, Komorowski RA, Knox JF, Emmons J, Bajaj JS, Binion DG.** 2007. Impact of *Clostridium difficile* on inflammatory bowel disease. *Clin Gastroenterol Hepatol* **5**: 345-51

47. **Bignardi GE.** 1998. Risk factors for *Clostridium difficile* infection. *J Hosp Infect* **40**: 1-15

48. **Forster AJ, Taljaard M, Oake N, Wilson K, Roth V, van Walraven C.** 2012. The effect of hospital-acquired infection with *Clostridium difficile* on length of stay in hospital. *CMAJ* **184**: 37-42

49. **Kelly CP, Kyne L.** 2011. The host immune response to *Clostridium difficile*. *J Med Microbiol* **60**: 1070-9

50. **Giannasca PJ, Zhang ZX, Lei WD, Boden JA, Giel MA, Monath TP, Thomas WD, Jr.** 1999. Serum antitoxin antibodies mediate systemic and mucosal protection from Clostridium difficile disease in hamsters. *Infect Immun* **67**: 527-38

51. **Leav BA, Blair B, Leney M, Knauber M, Reilly C, Lowy I, Gerding DN, Kelly CP, Katchar K, Baxter R, Ambrosino D, Molrine D.** 2010. Serum anti-toxin B antibody correlates with protection from recurrent *Clostridium difficile* infection (CDI). *Vaccine* **28**: 965-9

52. **Viscidi R, Laughon BE, Yolken R, Bo-Linn P, Moench T, Ryder RW, Bartlett JG.** 1983. Serum antibody response to toxins A and B of *Clostridium difficile*. *J Infect Dis* **148**: 93-100

53. **Kyne L, Warny M, Qamar A, Kelly CP.** 2000. Asymptomatic carriage of *Clostridium difficile* and serum levels of IgG antibody against toxin A. *N Engl J Med* **342**: 390-7

54. **Kyne L, Warny M, Qamar A, Kelly CP.** 2001. Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. *Lancet* **357**: 189-93

55. **Bacon AE, 3rd, Fekety R.** 1994. Immunoglobulin G directed against toxins A and B of Clostridium difficile in the general population and patients with antibiotic-associated diarrhea. *Diagn Microbiol Infect Dis* **18**: 205-9

56. **Rupnik M, Wilcox MH, Gerding DN.** 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol* **7**: 526-36

57. **Lowy I, Molrine DC, Leav BA, Blair BM, Baxter R, Gerding DN, Nichol G, Thomas WD, Jr., Leney M, Sloan S, Hay CA, Ambrosino DM.** 2010. Treatment with monoclonal antibodies against *Clostridium difficile* toxins. *N Engl J Med* **362**: 197-205

58. **Walters BA, Roberts R, Stafford R, Seneviratne E.** 1983. Relapse of antibiotic associated colitis: endogenous persistence of *Clostridium difficile* during vancomycin therapy. *Gut* **24**: 206-12

59. **McFarland LV.** 2005. Alternative treatments for *Clostridium difficile* disease: what really works? *J Med Microbiol* **54**: 101-11

60. **Barbut F, Richard A, Hamadi K, Chomette V, Burghoffer B, Petit JC.** 2000. Epidemiology of recurrences or reinfections of *Clostridium difficile*-associated diarrhea. *J Clin Microbiol* **38**: 2386-8

61. **Lawley TD, Clare S, Deakin LJ, Goulding D, Yen JL, Raisen C, Brandt C, Lovell J, Cooke F, Clark TG, Dougan G.** 2010. Use of purified *Clostridium difficile* spores to facilitate evaluation of health care disinfection regimens. *Appl Environ Microbiol* **76**: 6895-900

62. **Wilson KH, Sheagren JN, Freter R.** 1985. Population dynamics of ingested *Clostridium difficile* in the gastrointestinal tract of the Syrian hamster. *J Infect Dis* **151**: 355-61

63. **Sorg JA, Sonenshein AL.** 2008. Bile salts and glycine as cogerminants for *Clostridium difficile* spores. *J Bacteriol* **190**: 2505-12

64. **Borriello SP.** 1998. Pathogenesis of *Clostridium difficile* infection. *J Antimicrob Chemother* **41** Suppl C: 13-9

65. **Waligora AJ, Hennequin C, Mullany P, Bourlioux P, Collignon A, Karjalainen T.** 2001. Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infect Immun* **69**: 2144-53

66. **Dailey DC, Kaiser A, Schloemer RH.** 1987. Factors influencing the phagocytosis of *Clostridium difficile* by human polymorphonuclear leukocytes. *Infect Immun* **55**: 1541-6

67. **Voth DE, Ballard JD.** 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. *Clin Microbiol Rev* **18**: 247-63

68. **Borriello SP, Davies HA, Kamiya S, Reed PJ, Seddon S.** 1990. Virulence factors of *Clostridium difficile*. *Rev Infect Dis* **12** Suppl 2: S185-91

69. **Lyerly DM, Saum KE, MacDonald DK, Wilkins TD.** 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. *Infect Immun* **47**: 349-52

70. **Carter GP, Rood JI, Lyras D.** 2010. The role of toxin A and toxin B in *Clostridium difficile*-associated disease: Past and present perspectives. *Gut Microbes* **1**: 58-64

71. **Lyras D, O'Connor JR, Howarth PM, Sambol SP, Carter GP, Phumoonna T, Poon R, Adams V, Vedantam G, Johnson S, Gerding DN, Rood JI.** 2009. Toxin B is essential for virulence of *Clostridium difficile*. *Nature* **458**: 1176-9

72. **Jank T, Giesemann T, Aktories K.** 2007. Rho-glucosylating *Clostridium difficile* toxins A and B: new insights into structure and function. *Glycobiology* **17**: 15R-22

73. **Curry SR, Marsh JW, Muto CA, O'Leary MM, Pasculle AW, Harrison LH.** 2007. *tcdC* genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. *J Clin Microbiol* **45**: 215-21

74. **Murray R, Boyd D, Levett PN, Mulvey MR, Alfa MJ.** 2009. Truncation in the *tcdC* region of the *Clostridium difficile* PathLoc of clinical isolates does not predict increased biological activity of Toxin B or Toxin A. *BMC Infect Dis* **9**: 103

75. **Morgan OW, Rodrigues B, Elston T, Verlander NQ, Brown DF, Brazier J, Reacher M.** 2008. Clinical severity of *Clostridium difficile* PCR ribotype 027: a case-case study. *PLoS One* **3**: e1812

76. **McDonald LC, Killgore GE, Thompson A, Owens RC, Jr., Kazakova SV, Sambol SP, Johnson S, Gerding DN.** 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N Engl J Med* **353**: 2433-41

77. **Matamouros S, England P, Dupuy B.** 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Mol Microbiol* **64**: 1274-88

78. **von Eichel-Streiber C, Sauerborn M, Kuramitsu HK.** 1992. Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and *Streptococcus mutans* glucosyltransferases. *J Bacteriol* **174**: 6707-10

79. **Pothoulakis C, Gilbert RJ, Cladaras C, Castagliuolo I, Semenza G, Hitti Y, Montcrief JS, Linevsky J, Kelly CP, Nikulasson S, Desai HP, Wilkins TD, LaMont JT.** 1996. Rabbit sucrase-isomaltase contains a functional intestinal receptor for *Clostridium difficile* toxin A. *J Clin Invest* **98**: 641-9

80. **Tucker KD, Wilkins TD.** 1991. Toxin A of *Clostridium difficile* binds to the human carbohydrate antigens I, X, and Y. *Infect Immun* **59**: 73-8

81. **Hofmann F, Busch C, Prepens U, Just I, Aktories K.** 1997. Localization of the glucosyltransferase activity of *Clostridium difficile* toxin B to the N-terminal part of the holotoxin. *J Biol Chem* **272**: 11074-8

82. **Pfeifer G, Schirmer J, Leemhuis J, Busch C, Meyer DK, Aktories K, Barth H.** 2003. Cellular uptake of *Clostridium difficile* toxin B. Translocation of the N-terminal catalytic domain into the cytosol of eukaryotic cells. *J Biol Chem* **278**: 44535-41

83. **Poxton IR, McCoubrey J, Blair G.** 2001. The pathogenicity of *Clostridium difficile*. *Clin Microbiol Infect* **7**: 421-7

84. **Sehr P, Joseph G, Genth H, Just I, Pick E, Aktories K.** 1998. Glucosylation and ADP ribosylation of rho proteins: effects on nucleotide binding, GTPase activity, and effector coupling. *Biochemistry* **37**: 5296-304

85. **Jank T, Reinert DJ, Giesemann T, Schulz GE, Aktories K.** 2005. Change of the donor substrate specificity of *Clostridium difficile* toxin B by site-directed mutagenesis. *J Biol Chem* **280**: 37833-8

86. **Jank T, Aktories K.** 2008. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends Microbiol* **16**: 222-9

87. **Linevsky JK, Pothoulakis C, Keates S, Warny M, Keates AC, Lamont JT, Kelly CP.** 1997. IL-8 release and neutrophil activation by *Clostridium difficile* toxin-exposed human monocytes. *Am J Physiol* **273**: G1333-40

88. **Jangi S, Lamont JT.** 2010. Asymptomatic colonization by *Clostridium difficile* in infants: implications for disease in later life. *J Pediatr Gastroenterol Nutr* **51**: 2-7

89. **Eglow R, Pothoulakis C, Itzkowitz S, Israel EJ, O'Keane CJ, Gong D, Gao N, Xu YL, Walker WA, LaMont JT.** 1992. Diminished *Clostridium difficile* toxin A sensitivity in newborn rabbit ileum is associated with decreased toxin A receptor. *J Clin Invest* **90**: 822-9

90. **Verma P, Makharia GK.** 2011. *Clostridium difficile* associated diarrhea: new rules for an old game. *Trop Gastroenterol* **32**: 15-24

91. **Wolf PL, Kasyan A.** 2005. Images in clinical medicine. Pseudomembranous colitis associated with *Clostridium difficile*. *N Engl J Med* **353**: 2491

92. **Girotra M, Kumar V, Khan JM, Damisse P, Abraham RR, Aggarwal V, Dutta SK.** 2012. Clinical predictors of fulminant colitis in patients with *Clostridium difficile* infection. *Saudi J Gastroenterol* **18**: 133-9

93. **Jaber MR, Olafsson S, Fung WL, Reeves ME.** 2008. Clinical review of the management of fulminant *Clostridium difficile* infection. *Am J Gastroenterol* **103**: 3195-203

94. **Goldenberg SD, Cliff PR, French GL.** 2010. Glutamate dehydrogenase for laboratory diagnosis of *Clostridium difficile* infection. *J Clin Microbiol* **48**: 3050-1

95. **Wren MW, Sivapalan M, Kinson R, Shetty NR.** 2009. Laboratory diagnosis of *Clostridium difficile* infection. An evaluation of tests for faecal toxin, glutamate dehydrogenase, lactoferrin and toxigenic culture in the diagnostic laboratory. *Br J Biomed Sci* **66**: 1-5

96. **Kelly CP, Pothoulakis C, LaMont JT.** 1994. *Clostridium difficile* colitis. *N Engl J Med* **330**: 257-62

97. **Doern GV, Coughlin RT, Wu L.** 1992. Laboratory diagnosis of *Clostridium difficile*-associated gastrointestinal disease: comparison of a monoclonal antibody enzyme

immunoassay for toxins A and B with a monoclonal antibody enzyme immunoassay for toxin A only and two cytotoxicity assays. *J Clin Microbiol* **30**: 2042-6

98. **Lawley TD, Clare S, Walker AW, Goulding D, Stabler RA, Croucher N, Mastroeni P, Scott P, Raisen C, Mottram L, Fairweather NF, Wren BW, 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. *Infect Immun* **77**: 3661-9

99. **Hedge DD, Strain JD, Heins JR, Farver DK.** 2008. New advances in the treatment of *Clostridium difficile* infection (CDI). *Ther Clin Risk Manag* **4**: 949-64

100. **Metchnikoff E.** 1907. The prolongation of life. Optimistic studies. William Heineman, London, United Kingdom

101. **Eiseman B, Silen W, Bascom GS, Kauvar AJ.** 1958. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery* **44**: 854-9

102. **Rager KD, George LW, House JK, DePeters EJ.** 2004. Evaluation of rumen transfaunation after surgical correction of left-sided displacement of the abomasum in cows. *J Am Vet Med Assoc* **225**: 915-20

103. **Bakken JS.** 2009. Fecal bacteriotherapy for recurrent *Clostridium difficile* infection. *Anaerobe* **15**: 285-9

104. **Bakken JS, Borody T, Brandt LJ, Brill JV, Demarco DC, Franzos MA, Kelly C, Khoruts A, Louie T, Martinelli LP, Moore TA, Russell G, Surawicz C.** 2011. Treating *Clostridium difficile* infection with fecal microbiota transplantation. *Clin Gastroenterol Hepatol* **9**: 1044-9

105. **MacConnachie AA, Fox R, Kennedy DR, Seaton RA.** 2009. Faecal transplant for recurrent *Clostridium difficile*-associated diarrhoea: a UK case series. *QJM* **102**: 781-4

106. **Hempel S, Newberry SJ, Maher AR, Wang Z, Miles JN, Shanman R, Johnsen B, Shekelle PG.** 2012. Probiotics for the prevention and treatment of antibiotic-associated diarrhea: a systematic review and meta-analysis. *JAMA* **307**: 1959-69

107. **Surawicz CM.** 2003. Probiotics, antibiotic-associated diarrhoea and *Clostridium difficile* diarrhoea in humans. *Best Pract Res Clin Gastroenterol* **17**: 775-83

108. **Lee YK, Puong KY.** 2002. Competition for adhesion between probiotics and human gastrointestinal pathogens in the presence of carbohydrate. *Br J Nutr* **88** Suppl 1: S101-8

109. **Qamar A, Aboudola S, Warny M, Michetti P, Pothoulakis C, LaMont JT, Kelly CP.** 2001. *Saccharomyces boulardii* stimulates intestinal immunoglobulin A immune response to *Clostridium difficile* toxin A in mice. *Infect Immun* **69**: 2762-5

110. **Rolfe RD.** 2000. The role of probiotic cultures in the control of gastrointestinal health. *J Nutr* **130**: 396S-402S

111. **Pothoulakis C, Kelly CP, Joshi MA, Gao N, O'Keane CJ, Castagliuolo I, Lamont JT.** 1993. *Saccharomyces boulardii* inhibits *Clostridium difficile* toxin A binding and enterotoxicity in rat ileum. *Gastroenterology* **104**: 1108-15

112. **Castagliuolo I, LaMont JT, Nikulasson ST, Pothoulakis C.** 1996. *Saccharomyces boulardii* protease inhibits *Clostridium difficile* toxin A effects in the rat ileum. *Infect Immun* **64**: 5225-32

113. **Castagliuolo I, Riegler MF, Valenick L, LaMont JT, Pothoulakis C.** 1999. *Saccharomyces boulardii* protease inhibits the effects of *Clostridium difficile* toxins A and B in human colonic mucosa. *Infect Immun* **67**: 302-7

114. **Pochapin M.** 2000. The effect of probiotics on *Clostridium difficile* diarrhea. *Am J Gastroenterol* **95**: S11-3

115. **Lawrence SJ, Korzenik JR, Mundy LM.** 2005. Probiotics for recurrent *Clostridium difficile* disease. *J Med Microbiol* **54**: 905-6

116. **Bolton RP, Culshaw MA.** 1986. Faecal metronidazole concentrations during oral and intravenous therapy for antibiotic associated colitis due to *Clostridium difficile*. *Gut* **27**: 1169-72

117. **Zar FA, Bakkanagari SR, Moorthi KM, Davis MB.** 2007. A comparison of vancomycin and metronidazole for the treatment of *Clostridium difficile*-associated diarrhea, stratified by disease severity. *Clin Infect Dis* **45**: 302-7

118. **Vardakas KZ, Polyzos KA, Patouni K, Rafailidis PI, Samonis G, Falagas ME.** 2012. Treatment failure and recurrence of *Clostridium difficile* infection following

treatment with vancomycin or metronidazole: a systematic review of the evidence. *Int J Antimicrob Agents* **40**: 1-8

119. **McFarland LV, Surawicz CM, Rubin M, Fekety R, Elmer GW, Greenberg RN.** 1999. Recurrent *Clostridium difficile* disease: epidemiology and clinical characteristics. *Infect Control Hosp Epidemiol* **20**: 43-50

120. J**ohal SS, Lambert CP, Hammond J, James PD, Borriello SP, Mahida YR.** 2004. Colonic IgA producing cells and macrophages are reduced in recurrent and non-recurrent *Clostridium difficile* associated diarrhoea. *J Clin Pathol* **57**: 973-9

121. **Warny M, Vaerman JP, Avesani V, Delmee M.** 1994. Human antibody response to *Clostridium difficile* toxin A in relation to clinical course of infection. *Infect Immun* **62**: 384-9

122. **Sisto F, Scaltrito MM, Zago M, Bonomi A, Cocce V, Frugoni S.** 2011. Molecular analysis of relapses or reinfections of *Clostridium difficile*-associated diarrhea. *New Microbiol* **34**: 399-402

123. **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. *J Clin Microbiol* **50**: 915-21

124. **Health Protection Agency.** 2011 [online]. Voluntary surveillance of *Clostridium difficile* in England, Wales and Northern Ireland, 2011. Available: http://www.hpa.org.uk/web/HPA web&Page&HPAwebAutoListName/Page/1179745282413

125. **Office for National Statistics.** 2011 [online]. Deaths involving *Clostridium difficile,* England and Wales, 1999 to 2010. Available: http://www.google.co.uk/#hl=en&sclient=psy-ab&q=ONS+death+certificates+Clostridium+ difficile+England+Wales+1999+2010&oq=ONS+death+certificates+Clostridium+difficile+England+and+Wales+1999+2010&gs_l=serp.3...1825.6411.0.6645.13.13.0.0.0.0.218.1606.2 j10j1.13.0…1.0...1c.nV6cbLAOZ_o&pbx=1&bav=on.2,or.r_gc.r_pw.r_qf.&fp= f9ed1d9b74c502ec&biw=1058&bih=492

126. **Health Protection Agency.** 2012 [online]. Summary points on *Clostridium difficile* infection (CDI). Available: http://www.hpa.org.uk/ webc/HPAwebFile/HPAweb_C/1278944283388

127. **Healthcare Commission.** 2006 [online]. Investigation into outbreaks of *Clostridium difficile* at Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust. Available: http://www.buckinghamshirehospitals.nhs.uk/healthcarecommision/HCC-Investigation-into- the-outbreak-of-Clostridium-Difficile.pdf

128. **Healthcare Commission.** 2007 [online]. Investigation into outbreaks of *Clostridium difficile* at Maidstone and Tunbridge Wells NHS Trust. Available: http://news.bbc.co. uk/1/shared/ bsp/hi/pdfs/11_10_07maidstone_and_tunbridge_wells_investigation_ report_oct_2007.pdf

129. **Health Protection Agency.** 2010/2011 [online]. *Clostridium difficile* ribotyping network (CDRN) for England and Northern Ireland: 2010/2011 annual report. Available: http://www. hpa.org.uk/webc/ HPAwebFile/HPAweb_ C/1317133396963

130. **Verity P, Wilcox MH, Fawley W, Parnell P.** 2001. Prospective evaluation of environmental contamination by *Clostridium difficile* in isolation side rooms. *J Hosp Infect* **49**: 204-9

131. **Johnson S, Gerding DN.** 1998. *Clostridium difficile*-associated diarrhea. *Clin Infect Dis* **26**: 1027-34

132. **Sunenshine RH, McDonald LC.** 2006. *Clostridium difficile*-associated disease: new challenges from an established pathogen. *Cleve Clin J Med* **73**: 187-97

133. **Arias CA, Murray BE.** 2012. The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* **10**: 266-78

134. **Foulke GE, Silva J, Jr.** 1989. *Clostridium difficile* in the intensive care unit: management problems and prevention issues. *Crit Care Med* **17**: 822-6

135. **Chang VT, Nelson K.** 2000. The role of physical proximity in nosocomial diarrhea. *Clin Infect Dis* **31**: 717-22

136. **McFarland LV.** 2002. What's lurking under the bed? Persistence and predominance of particular *Clostridium difficile* strains in a hospital and the potential role of environmental contamination. *Infect Control Hosp Epidemiol* **23**: 639-40

137. **Khanna S, Pardi DS, Aronson SL, Kammer PP, Orenstein R, St Sauver JL, Harmsen WS, Zinsmeister AR.** 2012. The epidemiology of community-acquired

Clostridium difficile infection: a population-based study. *Am J Gastroenterol* **107**: 89-95

138. **Hookman P, Barkin JS.** 2009. *Clostridium difficile* associated infection, diarrhea and colitis. *World J Gastroenterol* **15**: 1554-80

139. **Goorhuis A, Bakker D, Corver J, Debast SB, Harmanus C, Notermans DW, Bergwerff AA, Dekker FW, Kuijper EJ.** 2008. Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction ribotype 078. *Clin Infect Dis* **47**: 1162-70

140. G**oorhuis A, Debast SB, van Leengoed LA, Harmanus C, Notermans DW, Bergwerff AA, Kuijper EJ.** 2008. *Clostridium difficile* PCR ribotype 078: an emerging strain in humans and in pigs? *J Clin Microbiol* **46**: 1157

141. **Keel K, Brazier JS, Post KW, Weese S, Songer JG.** 2007. Prevalence of PCR ribotypes among *Clostridium difficile* isolates from pigs, calves, and other species. *J Clin Microbiol* **45**: 1963-4

142. **Wilcox MH, Cunniffe JG, Trundle C, Redpath C.** 1996. Financial burden of hospital-acquired *Clostridium difficile* infection. *J Hosp Infect* **34**: 23-30

143. **Buckley AM, Spencer J, Candlish D, Irvine JJ, Douce GR.** 2011. Infection of hamsters with the UK *Clostridium difficile* ribotype 027 outbreak strain R20291. *J Med Microbiol* **60**: 1174-80

144. **Spencer RC.** 1998. Clinical impact and associated costs of *Clostridium difficile*-associated disease. *J Antimicrob Chemother* **41** Suppl C: 5-12

145. **Department of Health.** 2009 [online]. *Clostridium difficile* infection: how to deal with the problem. Available: http : // www.dh.gov.uk/prod_ consum _ dh / groups / dh_digitalassets/documents/digitalasset/dh_093218.pdf

146. **Gould CV, McDonald LC.** 2008. Bench-to-bedside review: *Clostridium difficile* colitis. *Crit Care* **12**: 203

147. **Zafar AB, Gaydos LA, Furlong WB, Nguyen MH, Mennonna PA.** 1998. Effectiveness of infection control program in controlling nosocomial *Clostridium difficile*. *Am J Infect Control* **26**: 588-93

148. **Walker AS, Eyre DW, Wyllie DH, Dingle KE, Harding RM, O'Connor L, Griffiths D, Vaughan A, Finney J, Wilcox MH, Crook DW, Peto TE.** 2012. Characterisation of *Clostridium difficile* hospital ward-based transmission using extensive epidemiological data and molecular typing. *PLoS Med* **9**: e1001172

149. **Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P.** 2000. Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**: 548-72

150. **Lawley TD, Croucher NJ, Yu L, Clare S, Sebaihia M, Goulding D, Pickard DJ, Parkhill J, Choudhary J, Dougan G.** 2009. Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *J Bacteriol* **191**: 5377-86

151. **Dawson LF, Valiente E, Donahue EH, Birchenough G, Wren BW.** 2011. Hypervirulent *Clostridium difficile* PCR-ribotypes exhibit resistance to widely used disinfectants. *PLoS One* **6**: e25754

152. **Gerding DN, Muto CA, Owens RC, Jr.** 2008. Measures to control and prevent *Clostridium difficile* infection. *Clin Infect Dis* **46** Suppl 1: S43-9

153. **Sobel J.** 2005. Botulism. *Clin Infect Dis* **41**: 1167-73

154. **Engelkirk PG, Duben-Engelkirk J.**2010.Major infectious diseases of humans, p 344. *In* Engelkirk PG, Duben-Engelkirk J (ed), Burton's microbiology for the health sciences, 9th ed. Lippincott Williams & Wilkins, Philadelphia, PA

155. **Shafazand S, Doyle R, Ruoss S, Weinacker A, Raffin TA.** 1999. Inhalational anthrax: epidemiology, diagnosis, and management. *Chest* **116**: 1369-76

156. **Paidhungat M, Setlow B, Driks A, Setlow P.** 2000. Characterization of spores of *Bacillus subtilis* which lack dipicolinic acid. *J Bacteriol* **182**: 5505-12

157. **Setlow B, Atluri S, Kitchel R, Koziol-Dube K, Setlow P.** 2006. Role of dipicolinic acid in resistance and stability of spores of *Bacillus subtilis* with or without DNA-protective alpha/beta-type small acid-soluble proteins. *J Bacteriol* **188**: 3740-7

158. **Young SB, Setlow P.** 2004. Mechanisms of *Bacillus subtilis* spore resistance to and killing by aqueous ozone. *J Appl Microbiol* **96**: 1133-42

159. **Beaman TC, Gerhardt P.** 1986. Heat resistance of bacterial spores correlated with protoplast dehydration, mineralization, and thermal adaptation. *Appl Environ Microbiol* **52**: 1242-6

160. **Setlow P.** 1995. Mechanisms for the prevention of damage to DNA in spores of *Bacillus* species. *Annu Rev Microbiol* **49**: 29-54

161. **Setlow P.** 2007. I will survive: DNA protection in bacterial spores. *Trends Microbiol* **15**: 172-80

162. **Riesenman PJ, Nicholson WL.** 2000. Role of the spore coat layers in *Bacillus subtilis* spore resistance to hydrogen peroxide, artificial UV-C, UV-B, and solar UV radiation. *Appl Environ Microbiol* **66**: 620-6

163. **Cowan AE, Olivastro EM, Koppel DE, Loshon CA, Setlow B, Setlow P.** 2004. Lipids in the inner membrane of dormant spores of *Bacillus* species are largely immobile. *Proc Natl Acad Sci USA* **101**: 7733-8

164. **Pearce SM, Kaethler AH.** 1977. Extent of cross-linking of germ cell wall of a variant of *Bacillus cereus*. *Biochem Biophys Res Commun* **77**: 1251-6

165. **Setlow P.** 2003. Spore germination. *Curr Opin Microbiol* **6**: 550-6

166. **Moir A, Smith DA.** 1990. The genetics of bacterial spore germination. *Annu Rev Microbiol* **44**: 531-53

167. **Driks A.** 1999. *Bacillus subtilis* spore coat. *Microbiol Mol Biol Rev* **63**: 1-20

168. **Henriques AO, Moran CP, Jr.** 2007. Structure, assembly, and function of the spore surface layers. *Annu Rev Microbiol* **61**: 555-88

169. **Panessa-Warren BJ, Tortora GT, Warren JB.** 1997. Exosporial membrane plasticity of *Clostridium sporogenes* and *Clostridium difficile*. *Tissue Cell* **29**: 449-61

170. **Lequette Y, Garenaux E, Tauveron G, Dumez S, Perchat S, Slomianny C, Lereclus D, Guerardel Y, Faille C.** 2011. Role played by exosporium glycoproteins in the surface properties of *Bacillus cereu*s spores and in their adhesion to stainless steel. *Appl Environ Microbiol* **77**: 4905-11

171. **Wolska KI, Grudniak AM, Kraczkiewicz-Dowjat A.** 2007. Genetic and physiological regulation of bacterial endospore development. *Pol J Microbiol* **56**: 11-7

172. **Errington J.** 2010. From spores to antibiotics via the cell cycle. *Microbiology* **156**: 1-13

173. **Tovar-Rojo F, Chander M, Setlow B, Setlow P.** 2002. The products of the *spoVA* operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis*. *J Bacteriol* **184**: 584-7

174. **Moir A.** 2006. How do spores germinate? *J Appl Microbiol* **101**: 526-30

175. **Errington J.** 2003. Regulation of endospore formation in *Bacillus subtilis*. *Nat Rev Microbiol* **1**: 117-26

176. **Paredes CJ, Alsaker KV, Papoutsakis ET.** 2005. A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* **3**: 969-78

177. **Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, Wilcox MH, Stephenson K.** 2009. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J Bacteriol* **191**: 7296-305

178. **Lewis RJ, Brannigan JA, Muchova K, Barak I, Wilkinson AJ.** 1999. Phosphorylated aspartate in the structure of a response regulator protein. *J Mol Biol* **294**: 9-15

179. **Greene EA, Spiegelman GB.** 1996. The Spo0A protein of *Bacillus subtilis* inhibits transcription of the *abrB* gene without preventing binding of the polymerase to the promoter. *J Biol Chem* **271**: 11455-61

180. **Strauch M, Webb V, Spiegelman G, Hoch JA.** 1990. The SpoOA protein of *Bacillus subtilis* is a repressor of the *abrB* gene. *Proc Natl Acad Sci USA* **87**: 1801-5

181. **Healy J, Weir J, Smith I, Losick R.** 1991. Post-transcriptional control of a sporulation regulatory gene encoding transcription factor sigma H in *Bacillus subtilis*. *Mol Microbiol* **5**: 477-87

182. **Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I.** 2011. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. *J Bacteriol* **193**: 3186-96

183. **Chastanet A, Vitkup D, Yuan GC, Norman TM, Liu JS, Losick RM.** 2010. Broadly heterogeneous activation of the master regulator for sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **107**: 8486-91

184. **Molle V, Fujita M, Jensen ST, Eichenberger P, Gonzalez-Pastor JE, Liu JS, Losick R.** 2003. The Spo0A regulon of *Bacillus subtilis*. *Mol Microbiol* **50**: 1683-701

185. **Fawcett P, Eichenberger P, Losick R, Youngman P.** 2000. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **97**: 8063-8

186. **Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP.** 2007. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods* **70**: 452-64

187. **Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K, Minton NP.** 2012. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res* **40**: e59

188. **Carter GP, Lyras D, Poon R, Howarth PM, Rood JI.** 2010. Methods for gene cloning and targeted mutagenesis. *Methods Mol Biol* **646**: 183-201

189. **Zhong J, Karberg M, Lambowitz AM.** 2003. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res* **31**: 1656-64

190. **Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP.** 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J Microbiol Methods* **80**: 49-55

191. **Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, Assefa SA, He M, Croucher NJ, Pickard DJ, Maskell DJ, Parkhill J, Choudhary J, Thomson NR, Dougan G.** 2009. A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. *PLoS Genet* **5**: e1000569

192. **Martin J, Zhu W, Passalacqua KD, Bergman N, Borodovsky M.** 2010. *Bacillus anthracis* genome organization in light of whole transcriptome sequencing. *BMC Bioinformatics* **11** Suppl 3: S10

193. **Pellin D, Miotto P, Ambrosi A, Cirillo DM, Di Serio C.** 2012. A genome-wide identification analysis of small regulatory RNAs in *Mycobacterium tuberculosis* by RNA-Seq and conservation analysis. *PLoS One* **7***:* e32723

194. **Dong TG, Mekalanos JJ.** 2012. Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52. *Nucleic Acids Res* **40**: 7766-75

195. **Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP.** 2002. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol Microbiol* **46**: 439-52

196. **Hussain HA, Roberts AP, Mullany P.** 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δ*erm*) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. *J Med Microbiol* **54**: 137-41

197. **Wilson KH.** 1983. Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination. *J Clin Microbiol* **18**: 1017-9

198. **Sloan J, Warner TA, Scott PT, Bannam TL, Berryman DI, Rood JI.** 1992. Construction of a sequenced *Clostridium perfringens-Escherichia coli* shuttle plasmid. *Plasmid* **27**: 207-19

199. **Blatny JM, Brautaset T, Winther-Larsen HC, Karunakaran P, Valla S.** 1997. Improved broad-host-range RK2 vectors useful for high and low regulated gene expression levels in gram-negative bacteria. *Plasmid* **38**: 35-51

200. **Schooley DL, Kubiak FM, Evans JV.** 1985. Capillary gas chromatographic analysis of volatile and non-volatile organic acids from biological samples as the t-butyldimethylsilyl derivatives. *J Chromatogr Sci* **23**: 385-90

201. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754-60

202. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. *Bioinformatics* **16**: 944-5

203. **Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M.** 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**: 2856-60

204. **Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ.** 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**: 484-94

205. **Cox J, Mann M.** 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**: 1367-72

206. **Dupuy B, Govind R, Antunes A, Matamouros S.** 2008. *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. *J Med Microbiol* **57**: 685-9

207. **Bauer MP, Notermans DW, van Benthem BH, Brazier JS, Wilcox MH, Rupnik M, Monnet DL, van Dissel JT, Kuijper EJ.** 2011. *Clostridium difficile* infection in Europe: a hospital-based survey. *Lancet* **377**: 63-73

208. **Hull MW, Beck PL.** 2004. *Clostridium difficile*-associated colitis. *Can Fam Physician* **50**: 1536-40, 43-5

209. **Steele J, Feng H, Parry N, Tzipori S.** 2010. Piglet models of acute or chronic *Clostridium difficile* illness. *J Infect Dis* **201**: 428-34

210. **Rothman SW.** 1981. Presence of *Clostridium difficile* toxin in guinea pigs with penicillin-associated colitis. *Med Microbiol Immunol* **169**: 187-96

211. **Kelly CP, Becker S, Linevsky JK, Joshi MA, O'Keane JC, Dickey BF, LaMont JT, Pothoulakis C.** 1994. Neutrophil recruitment in *Clostridium difficile* toxin A enteritis in the rabbit. *J Clin Invest* **93**: 1257-65

212. **Goulding D, Thompson H, Emerson J, Fairweather NF, Dougan G, Douce GR.** 2009. Distinctive profiles of infection and pathology in hamsters infected with *Clostridium difficile* strains 630 and B1. *Infect Immun* **77**: 5478-85

213. **Douce G, Goulding D.** 2010. Refinement of the hamster model of *Clostridium difficile* disease. *Methods Mol Biol* **646**: 215-27

214. **Chen X, Katchar K, Goldsmith JD, Nanthakumar N, Cheknis A, Gerding DN, Kelly CP.** 2008. A mouse model of *Clostridium difficile*-associated disease. *Gastroenterology* **135**: 1984-92

215. **Keel MK, Songer JG.** 2006. The comparative pathology of *Clostridium difficile*-associated disease. *Vet Pathol* **43**: 225-40

216. **Huang IH, Sarker MR.** 2006. Complementation of a *Clostridium perfringens spo0A* mutant with wild-type *spo0A* from other *Clostridium* species. *Appl Environ Microbiol* **72**: 6388-93

217. **Lereclus D, Agaisse H, Grandvalet C, Salamitou S, Gominet M.** 2000. Regulation of toxin and virulence gene transcription in *Bacillus thuringiensis*. *Int J Med Microbiol* **290**: 295-9

218. **Saile E, Koehler TM.** 2002. Control of anthrax toxin gene expression by the transition state regulator *abrB*. *J Bacteriol* **184**: 370-80

219. **Vohra P, Poxton IR.** 2011. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. *Microbiology* **157**: 1343-53

220. **Onderdonk AB, Cisneros RL, Bartlett JG.** 1980. *Clostridium difficile* in gnotobiotic mice. *Infect Immun* **28**: 277-82

221. **Riggs MM, Sethi AK, Zabarsky TF, Eckstein EC, Jump RL, Donskey CJ.** 2007. Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic *Clostridium difficile* strains among long-term care facility residents. *Clin Infect Dis* **45**: 992-8

222. **Lell B, Kremsner PG.** 2002. Clindamycin as an antimalarial drug: review of clinical trials. *Antimicrob Agents Chemother* **46**: 2315-20

223. **Croucher NJ, Fookes MC, Perkins TT, Turner DJ, Marguerat SB, Keane T, Quail MA, He M, Assefa S, Bahler J, Kingsley RA, Parkhill J, Bentley SD, Dougan G, Thomson NR.** 2009. A simple method for directional transcriptome sequencing using Illumina technology. *Nucleic Acids Res* **37**: e148

224. **Hegde PS, White IR, Debouck C.** 2003. Interplay of transcriptomics and proteomics. *Curr Opin Biotechnol* **14**: 647-51

225. **Wasinger V.** 2006. Holistic biology of microorganisms: genomics, transcriptomics, and proteomics. *Methods Biochem Anal* **49**: 3-14

226. **Fujita M, Gonzalez-Pastor JE, Losick R.** 2005. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. *J Bacteriol* **187**: 1357-68

227. **Hundsberger T, Braun V, Weidmann M, Leukel P, Sauerborn M, von Eichel-Streiber C.** 1997. Transcription analysis of the genes *tcdA-E* of the pathogenicity locus of *Clostridium difficile*. *Eur J Biochem* **244**: 735-42

228. **Andreesen JR.** 1994. Glycine metabolism in anaerobes. *Antonie Van Leeuwenhoek* **66**: 223-37

229. **Sebaihia M, Peck MW, Minton NP, Thomson NR, Holden MT, Mitchell WJ, Carter AT, Bentley SD, Mason DR, Crossman L, Paul CJ, Ivens A, Wells-Bennik MH, Davis IJ, Cerdeno-Tarraga AM, Churcher C, Quail MA, Chillingworth T, Feltwell T, Fraser A, Goodhead I, Hance Z, Jagels K, Larke N, Maddison M, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, White B, Whithead S, Parkhill J.** 2007. Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes. *Genome Res* **17**: 1082-92

230. **Fonknechten N, Chaussonnerie S, Tricot S, Lajus A, Andreesen JR, Perchat N, Pelletier E, Gouyvenoux M, Barbe V, Salanoubat M, Le Paslier D, Weissenbach J, Cohen GN, Kreimeyer A.** 2010. *Clostridium sticklandii*, a specialist in amino acid degradation: revisiting its metabolism through its genome sequence. *BMC Genomics* **11**: 555

231. **Dickert S, Pierik AJ, Buckel W.** 2002. Molecular characterization of phenyllactate dehydratase and its initiator from *Clostridium sporogenes*. *Mol Microbiol* **44**: 49-60

232. **Herrmann G, Jayamani E, Mai G, Buckel W.** 2008. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *J Bacteriol* **190**: 784-91

233. **Tannler S, Decasper S, Sauer U.** 2008. Maintenance metabolism and carbon fluxes in Bacillus species. *Microb Cell Fact* **7**: 19

234. **Lie TJ, Pitta T, Leadbetter ER, Godchaux W, 3rd, Leadbetter JR.** 1996. Sulfonates: novel electron acceptors in anaerobic respiration. *Arch Microbiol* **166**: 204-10

235. **Romano AH, Eberhard SJ, Dingle SL, McDowell TD.** 1970. Distribution of the phosphoenolpyruvate: glucose phosphotransferase system in bacteria. *J Bacteriol* **104**: 808-13

236. **Beyer A, Hollunder J, Nasheuer HP, Wilhelm T.** 2004. Post-transcriptional expression regulation in the yeast *Saccharomyces cerevisiae* on a genomic scale. *Mol Cell Proteomics* **3**: 1083-92

237. **Franklund CV, de Prada P, Hylemon PB.** 1990. Purification and characterization of a microbial, NADP-dependent bile acid 7-alpha-hydroxysteroid dehydrogenase. *J Biol Chem* **265**: 9842-9

238. **Coleman JP, Hudson LL, Adams MJ.** 1994. Characterization and regulation of the NADP-linked 7-alpha-hydroxysteroid dehydrogenase gene from *Clostridium sordellii*. *J Bacteriol* **176**: 4865-74

239. **Topping DL, Clifton PM.** 2001. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* **81**: 1031-64

240. **Sakuragi H, Kuroda K, Ueda M.** 2011. Molecular breeding of advanced microorganisms for biofuel production. *J Biomed Biotechnol* **2011**: 416931

241. **Cai G, Jin B, Saint C, Monis P.** 2011. Genetic manipulation of butyrate formation pathways in *Clostridium butyricum*. *J Biotechnol* **155**: 269-74

242. **Emeruwa AC, Hawirko RZ, Halvorson H, Suzuki I.** 1974. Comparison of butyric type of fermentation in sporogenic and asporogenic mutants of *Clostridium botulinum*. *J Bacteriol* **120**: 74-80

243. **Nickerson KW, Bulla LA.** 1980. Incorporation of Specific Fatty Acid Precursors During Spore Germination and Outgrowth in *Bacillus thuringiensis*. *Appl Environ Microbiol* **40**: 166-8

244. **Alsaker KV, Spitzer TR, Papoutsakis ET.** 2004. Transcriptional analysis of spo0A overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *J Bacteriol* **186**: 1959-71

245. **Twine SM, Reid CW, Aubry A, McMullin DR, Fulton KM, Austin J, Logan SM.** 2009. Motility and flagellar glycosylation in *Clostridium difficile*. *J Bacteriol* **191**: 7050-62

246. **Richardson K.** 1991. Roles of motility and flagellar structure in pathogenicity of *Vibrio cholerae*: analysis of motility mutants in three animal models. *Infect Immun* **59**: 2727-36

247. **Yao R, Burr DH, Doig P, Trust TJ, Niu H, Guerry P.** 1994. Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni*: the role of motility in adherence and invasion of eukaryotic cells. *Mol Microbiol* **14**: 883-93

248. **Eaton KA, Morgan DR, Krakowka S.** 1992. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *J Med Microbiol* **37**: 123-7

249. **Piggot PJ, Hilbert DW.** 2004. Sporulation of *Bacillus subtilis*. *Curr Opin Microbiol* **7**: 579-86

250. **Tsou AM, Frey EM, Hsiao A, Liu Z, Zhu J.** 2008. Coordinated regulation of virulence by quorum sensing and motility pathways during the initial stages of *Vibrio cholerae* infection. *Commun Integr Biol* **1**: 42-4

251. **Syed KA, Beyhan S, Correa N, Queen J, Liu J, Peng F, Satchell KJ, Yildiz F, Klose KE.** 2009. The *Vibrio cholerae* flagellar regulatory hierarchy controls expression of virulence factors. *J Bacteriol* **191**: 6555-70

252. **Ramos HC, Rumbo M, Sirard JC.** 2004. Bacterial flagellins: mediators of pathogenicity and host immune responses in mucosa. *Trends Microbiol* **12**: 509-17

253. **Yoon SI, Kurnasov O, Natarajan V, Hong M, Gudkov AV, Osterman AL, Wilson IA.** 2012. Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**: 859-64

254. **Segain JP, Raingeard de la Bletiere D, Bourreille A, Leray V, Gervois N, Rosales C, Ferrier L, Bonnet C, Blottiere HM, Galmiche JP.** 2000. Butyrate inhibits inflammatory responses through NFʁB inhibition: implications for Crohn's disease. *Gut* **47**: 397-403

255. **Meijer K, de Vos P, Priebe MG.** 2010. Butyrate and other short-chain fatty acids as modulators of immunity: what relevance for health? *Curr Opin Clin Nutr Metab Care* **13**: 715-21

256. **Sun N, Pan C, Nickell S, Mann M, Baumeister W, Nagy I.** 2010. Quantitative proteome and transcriptome analysis of the archaeon *Thermoplasma acidophilum* cultured under aerobic and anaerobic conditions. *J Proteome Res* 9: 4839-50

257. **Dressaire C, Gitton C, Loubiere P, Monnet V, Queinnec I, Cocaign-Bousquet M.** 2009. Transcriptome and proteome exploration to model translation efficiency and protein stability in *Lactococcus lactis*. *PLoS Comput Biol* **12**: e1000606

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD0065	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase	3.629862769	2.36E-36
CD0107 (AspC)	aspartate aminotransferase	0.716100658	0.006971664
CD0112 (Ptb)	phosphate butyryltransferase	0.667104568	0.007921769
CD0113 (Buk)	butyrate kinase	0.663624022	0.008116098
CD0118	putative subunit of oxidoreductase	0.654815836	0.006973221
CD0179 (GluD)	NAD-specific glutamate dehydrogenase	0.654479793	0.003816227
CD0205	putative transcription antiterminator, PTS operon regulator	6.09623072	1.45E-28
CD0206	PTS system, IIa component	3.886913238	0.000858762
CD0207	PTS system, IIc component	4.770150537	1.22E-10
CD0208	PTS system, IIb component	5.3760065	3.90E-08
CD0209	putative sugar-phosphate kinase	5.205948535	1.31E-40
CD0226	putative transglycosylase	0.971434538	0.001330514
CD0239 (FliC)	flagellin subunit	0.64993773	0.005144783
CD0241	phosphoserine phosphatase	1.106961294	4.19E-06
CD0242	conserved hypothetical protein	1.004382201	7.69E-05
CD0243	conserved hypothetical protein	1.14223852	1.05E-07
CD0244	putative CDP-Glycerol:Poly(glycerophosphate) glycerophosphotransferase	1.090609331	5.26E-07
CD0248 (FliF)	flagellar M-ring protein	0.691937512	0.008339449
CD0249 (FliG)	flagellar motor switch protein	0.704163503	0.007552048
CD0250 (FliH)	flagellar assembly protein	1.068085331	0.0006144
CD0252 (FliJ)	flagellar protein	1.005917227	0.000186141
CD0253 (FliK)	putative flagellar hook-length control protein	0.894553529	0.000337475
CD0255A	putative flagellar protein	1.020396977	0.009696913
CD0283	putative transcription antiterminator	1.217971016	0.001556487
CD0285	PTS system, IIb component	4.027424265	0.00051567
CD0288	PTS system, IIc component	2.471850245	0.000560242
CD0289	PTS system, IId component	2.653617296	0.002597411
CD0292	DNA-binding protein / transcriptional regulator	4.302473109	2.44E-22
CD0293	ABC transporter, ATP-binding protein / bacitracin multidrug famity ATP-binding protein	4.630995313	2.99E-21
CD0294	putative ABC transporter, permease protein / bacitracin/multidrug family permease	5.189801595	1.26E-33
CD0301 (RbsA)	ribose ABC transporter, ATP-binding protein	2.724347261	8.87E-07
CD0302 (RbsC)	ribose ABC transporter, permease protein	2.55763691	0.000166061
CD0303 (ArgE)	putative acetylornithine deacetylase	2.58181175	0.000223451
CD0340	conserved hypothetical protein	1.229288604	3.71E-09
CD0341	conserved hypothetical protein	1.639166338	2.62E-08
CD0395 (HadA/FldA)	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	0.66373869	0.002918154
CD0396 (HadI/FldI)	activator of 2-hydroxyisocaproyl-CoA dehydratase	0.711457875	0.001199908
CD0397 (HadB/FldB)	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	0.781900155	0.000219189
CD0398 (HadC/FldC)	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	0.604290822	0.00918678
CD0399 (AcdB/Bcd1)	acyl-CoA dehydrogenase, short-chain specific	0.67187765	0.002613865
CD0400 (EtfB1)	electron transfer flavoprotein beta-subunit	0.887548384	1.60E-05
CD0529	putative membrane protein	0.952512987	0.00163008
CD0587	conserved hypothetical protein	1.713471592	1.23E-05
CD0588	hypothetical protein	1.300071341	0.000560242
CD0616	MerR-family transcriptional regulator	1.96436041	1.35E-08
CD0617	putative membrane protein	1.786565014	0.00288371
CD0618	putative transcriptional regulator	2.091158099	1.58E-05

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD0663 (TcdA)	toxin A	0.967520109	6.42E-06
CD0715	putative phosphate butyryltransferase	0.933229347	0.00491507
CD0758 (PlfA)	pyruvate formate-lyase activating enzyme	1.252297367	0.000201317
CD0759 (PlfB)	formate acetyltransferase	1.670841387	8.45E-18
CD0893	iron-dependent hydrogenase	1.514860175	3.05E-12
CD0939	hypothetical phage protein	1.550823074	5.17E-06
CD1022	putative membrane protein	1.200660139	0.009155858
CD1064 (CcpA)	LacI-family transcriptional regulator (catabolite control protein)	1.006380139	0.004894689
CD1080	putative lipoprotein	2.045597495	4.79E-13
CD1137 (RnfC)	electron transport complex protein	0.679398699	0.00342762
CD1138 (RnfD)	electron transport complex protein	0.901879171	5.21E-05
CD1139 (RnfG)	electron transport complex protein	0.897363148	0.000249919
CD1141 (RnfA)	electron transport complex protein	0.86340992	5.84E-05
CD1142 (RnfB)	electron transport complex protein	0.832619392	0.000572698
CD1144 (RadC)	DNA repair protein	0.794502153	0.003404511
CD1148	putative penicillin-binding protein	0.751729324	0.001715852
CD1191 (Fbp)	putative fructose-1,6-bisphosphatase	0.939868722	0.00012966
CD1260	branched chain amino acid transport system carrier protein	0.814368658	0.000254892
CD1348	putative lipoprotein	0.974119915	0.005090336
CD1494	putative transcriptional regulator	1.49340953	7.86E-09
CD1495 (ProC1)	pyrroline-5-carboxylate reductase	1.336135433	2.14E-08
CD1522	putative polysaccharide deacetylase	1.858413598	9.97E-16
CD1579	putative two-component histidine kinase	2.729188221	7.67E-15
CD1654 (LplA)	putative lipoate-protein ligase	1.706598706	3.27E-10
CD1708	conserved hypothetical protein	1.434839086	1.43E-07
CD1716	probable permease	1.697555732	2.24E-12
CD1717	conserved hypothetical protein	1.531136235	0.0002403
CD1718	putative hydantoinase	1.063994467	1.73E-05
CD1768	putative membrane protein	1.702253209	1.33E-18
CD1796	putative nitrite and sulfite reductase subunit	1.408566164	2.81E-06
CD1797	putative pyridine nucleotide-disulfide oxidoreductase	1.533215901	7.86E-09
CD1893	putative regulatory protein	1.21619771	0.009604073
CD1930	putative exported protein	1.308492935	0.001715852
CD2007 (Erm2(B))	erythromycin resistance protein- rRNA adenine N-6-methyltransferase	5.242933225	2.02E-90
CD2010 (Erm1(B))	erythromycin resistance protein- rRNA adenine N-6-methyltransferase	5.10006667	2.25E-61
CD2052	putative lipoprotein	1.184949348	0.005176864
CD2134	putative signaling protein	2.242591734	1.08E-10
CD2156	radical SAM-superfamily protein	1.01433679	2.11E-05
CD2175	probable amino-acid ABC transporter, permease protein	0.69702847	0.003237738
CD2177	probable amino-acid ABC transporter, substrate-binding protein	0.734721854	0.00219319
CD2214	putative regulatory protein	3.535392369	6.44E-56
CD2215	putative regulatory protein	4.287173128	5.64E-20
CD2305	putative pilin protein	2.220923226	0.001789576
CD2338 (AbfH)	NAD-dependent 4-hydroxybutyrate dehydrogenase	3.259168739	0.000438995
CD2339 (AbfT)	4-hydroxybutyrate CoA transferase	3.157366059	5.71E-05
CD2340	conserved hypothetical protein	2.96442692	4.81E-06
CD2341 (AbfD)	gamma-aminobutyrate metabolism dehydratase/isomerase [includes: 4- hydroxybutyryl-coa dehydratase; vinylacetyl-coa-delta- isomerase]	3.689464172	2.63E-05

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD2342 (SucD)	succinate-semialdehyde dehydrogenase [NAD(P)+]	4.826850593	3.07E-15
CD2343 (Cat1)	succinyl-CoA:coenzyme A transferase	4.595958513	9.49E-05
CD2344	putative membrane protein	4.186502319	0.001866395
CD2379	butyrate kinase	1.24419252	6.32E-06
CD2380 (IorB)	putative indolepyruvate oxidoreductase subunit	1.830018405	6.19E-07
CD2381 (IorA)	indolepyruvate oxidoreductase subunit	1.814405546	1.03E-14
CD2382	putative aspartate aminotransferase	1.54588669	2.13E-10
CD2417 (SrlE)	PTS system, glucitol/sorbitol-specific IIbc component	4.280426592	0.000944418
CD2502	putative histidinol-phosphate aminotransferase	0.927784357	0.000680281
CD2625	putative membrane protein	1.298949868	0.000111432
CD2711	LysR-family transcriptional regulator	1.711595038	6.76E-07
CD2755 (PtsI)	phosphoenolpyruvate-protein phosphotransferase	0.689931477	0.002490116
CD2757	putative phospholipase	0.797211818	0.001397758
CD2765	putative penicillin-binding protein repressor	2.199301337	2.18E-24
CD2796	cell surface protein	2.439933041	1.01E-13
CD2797	putative exported protein	2.608806934	5.22E-25
CD2830	putative exported protein	2.041806735	1.33E-19
CD2853 (DltA)	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)	0.693180152	0.008974023
CD3036	proton-dependent oligopeptide transporter/major facilitator superfamily	4.388137962	1.44E-83
CD3127	PTS system, IIabc component	1.221580142	0.008082314
CD3133	transcription antiterminator	1.00605468	0.008970866
CD3158	conserved hypothetical protein	3.304133839	2.18E-12
CD3165	conserved hypothetical protein	1.678616513	0.00088117
CD3166	putative transcriptional regulator	2.265313701	0.005200464
CD3246	putative surface protein	1.095665394	1.48E-07
CD3404 (Cls)	putative cardiolipin synthetase	0.741429305	0.001244253
CD3406 (HymB)	putative iron-only hydrogenase, electron-transferring subunit	0.903922783	3.17E-05
CD3407 (HymC)	putative iron-only hydrogenase, catalytic subunit	0.689917122	0.00337708
CD3458	putative membrane protein	0.820554311	0.000409987
CD3555	conserved hypothetical protein	1.423465908	1.10E-08
CD3556	putative membrane protein	1.227409985	1.43E-07
CD3664	putative amino acid aminotransferase	1.794066616	3.90E-08

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD0022	putative translation elongation factor	-1.435177186	7.27E-10
CD0126 (SpoIIID)	stage III sporulation protein D	-3.975328311	0.003988481
CD0142	putative RNA-binding protein	-1.057363331	7.96E-05
CD0148	conserved hypothetical protein	-2.800686067	0.001061682
CD0157	putative membrane protein (pseudogene	-2.921178976	3.82E-28
CD0311	hypothetical protein	-4.063410917	2.84E-08
CD0324 (CbiM)	putative cobalt transport protein	-0.662863884	0.003886953
CD0326 (CbiQ)	cobalt transport protein	-0.687187275	0.00247577
CD0330	conserved hypothetical protein	-0.9778842	0.00479394
CD0388 (BglP)	PTS system, beta-glucoside-specific IIabc component	-1.198511551	0.000409987
CD0390 (BglG)	beta-glucoside bgl operon transcription antiterminator	-1.460948584	0.007552048
CD0440 (similar to Cwp66)	cell surface protein	-2.317288463	6.46E-07
CD0556	putative endonuclease	-1.158315896	0.005729237
CD0557	putative phosphoribulokinase/uridine kinase	-0.960163231	0.002774647
CD0571	conserved hypothetical protein	-1.822649957	0.000112801
CD0572	conserved hypothetical protein	-1.980269174	0.0014809
CD0684	putative ATP-dependent peptidase	-29.329564	4.96E-25
CD0692	conserved hypothetical protein	-2.117895493	2.78E-05
CD0744 (MotD)	putative chemotaxis protein	-1.297760076	0.001494966
CD0745 (MotB)	putative chemotaxis protein	-1.336759872	0.000419336
CD0760	Ca2+/Na+ antiporter	-1.935407687	0.001134617
CD0770 (SpoIIAA)	anti-sigma F factor antagonist	-3.012104306	2.61E-06
CD0771 (SpoIIAB)	anti-sigma F factor	-2.219236598	7.24E-05
CD0772 (SpoIIAC/SigF)	RNA polymerase sigma-F factor	-2.398756207	3.59E-07
CD0788	putative ATP/GTP-binding protein	-1.21715207	3.77E-05
CD0816	PTS system, IIabc component	-0.862058451	0.000684304
CD0818	putative 6-phospho-beta-glucosidase	-0.926227909	0.001384179
CD0865	conserved hypothetical protein	-1.127403392	0.000217448
CD0866	putative exported protein	-1.209772024	8.10E-05
CD0874	ABC transporter, ATP-binding protein	-1.312243261	2.98E-08
CD0875	ABC transporter, permease protein	-1.287872019	0.002535493
CD0877	ABC transporter, ATP-binding protein	-0.990828424	0.00913806
CD0878	ABC transporter, permease protein	-1.176685562	3.59E-07
CD0940A	hypothetical phage protein	-1.989048324	6.32E-06
CD0941	phage protein	-1.569897167	1.20E-05
CD1054 (Bcd2)	butyryl-CoA dehydrogenas	-2.5961265	1.83E-22
CD1055 (EtfB2)	electron transfer flavoprotein beta-subunit	-2.683497328	2.36E-36
CD1056 (EtfA2)	electron transfer flavoprotein alpha-subunit	-2.686059258	1.05E-12
CD1057 (Crt2)	3-hydroxybutyryl-CoA dehydratase	-3.090695365	3.57E-45
CD1058 (Hdb)	3-hydroxybutyryl-CoA dehydrogenase	-2.518020502	8.99E-33
CD1059 (ThlA1)	acetyl-CoA acetyltransferase	-2.34691265	2.43E-08
CD1065	hypothetical protein	-6.164172544	6.62E-05
CD1079	LysR-family transcriptional regulator	-2.12449871	2.87E-10
CD1085	putative membrane protein	-1.691085931	2.69E-09
CD1086	putative peptidase	-1.667458643	1.58E-05
CD1168	putative membrane protein	-4.675401845	1.83E-12
CD1192 (SpoIIIAA)	stage III sporulation protein AA	-26.578922	0.000979086
CD1193 (SpoIIIAB)	stage III sporulation protein AB	-27.081907	2.35E-05
CD1195 (SpoIIIAD)	stage III sporulation protein AD	-26.249808	0.005629832
CD1196 (SpoIIIAE)	stage III sporulation-related protein	-27.731001	1.97E-07
CD1198 (SpoIIIAG)	stage III sporulation protein AG	-26.32129	0.004082152
CD1199 (SpoIIIAH)	putative stage III sporulation protein AH	-5.280906841	0.0044163

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD1388	putative transcriptional regulator (pseudogene)	-1.428025955	0.006009262
CD1416	putative membrane protein	-26.226648	0.006999155
CD1484 (SsuA)	putative aliphatic sulfonates ABC transporter, substrate-binding lipoprotein	-2.432897557	2.54E-15
CD1492	putative two-component sensor histidine kinase	-2.923130388	1.72E-29
CD1493	putative 3-methyladenine DNA glycosylase	-3.241054631	0.00022301
CD1498 (RpoD2/SigA2)	RNA polymerase sigma factor rpoD	-0.818609088	0.00288371
CD1511	conserved hypothetical protein	-3.145446944	3.48E-06
CD1513 (PanB)	3-methyl-2-oxobutanoate hydroxymethyltransferase	-1.607863456	0.000611659
CD1514	conserved hypothetical protein	-1.260956275	0.001575776
CD1524	putative rubrerythrin	-1.321718368	1.22E-08
CD1526 (PyrC)	dihydroorotase	-1.285512307	7.13E-06
CD1657	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein	-0.88258665	0.000657741
CD1658 (GcvPB)	glycine cleavage system P protein	-1.227982087	5.38E-08
CD1826 (MetA)	homoserine O-succinyltransferase	-1.298920427	0.001775008
CD1883	AraC-family transcriptional regulator	-1.515866812	0.00163008
CD1928	putative membrane protein	-5.249334334	4.55E-08
CD1940	putative membrane protein	-6.248545995	0.000735822
CD1941	hypothetical protein	-2.503401151	1.44E-05
CD1967	hypothetical protein	-2.539605421	3.10E-13
CD1993	putative decarboxylase	-2.307234533	6.86E-09
CD1996	AraC-family transcriptional regulator	-1.267759294	0.001235725
CD2127	putative membrane protein	-0.999574513	1.07E-05
CD2181	putative aromatic compounds hydrolase	-2.124774595	9.05E-15
CD2195	ferritin	-0.746373738	0.001544703
CD2216	hypothetical protein	-1.83442299	3.06E-06
CD2246 (CspC)	putative germination-specific protease	-2.540054266	2.78E-05
CD2247 (CspBA)	putative germination-specific protease	-1.935437381	3.60E-05
CD2261	peptidase	-1.046223325	6.90E-05
CD2295	putative membrane protein	-1.741024125	0.000419336
CD2373	putative carbon starvation protein	-1.924325732	8.51E-10
CD2396	conserved hypothetical protein	-1.587879958	5.45E-05
CD2445	probale sensory protein	-2.448449672	0.001924126
CD2509	probable 6-phospho-alpha-glucosidase	-1.325519571	0.001064996
CD2518 (similar to Cwp66)	cell surface protein	-1.548613884	6.70E-05
CD2612	putative amino acid permease	-1.90022516	1.36E-16
CD2613	probable peptidase	-1.738939315	3.05E-12
CD2629 (SpoIVA)	stage IV sporulation protein A	-4.777663355	4.13E-05
CD2635	putative membrane protein	-26.813826	0.000186406
CD2642 (SigG)	RNA polymerase sigma-G factor	-4.068242153	3.05E-05
CD2643 (SigE)	RNA polymerase sigma-E factor	-5.468253694	7.40E-05
CD2644 (SpoIIGA)	sporulation sigma-E factor processing peptidase	-5.55102365	8.32E-09
CD2656 (SpoVD)	stage V sporulation protein D (sporulation specific penicillin-binding protein)	-3.348406137	2.25E-05
CD2664 (MurE)	putative UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	-1.683182898	2.52E-14
CD2665	AraC-family transcriptional regulator	-2.02285346	3.23E-19
CD2666 (Crr)	PTS system, glucose-specific IIa component	-1.75518228	3.93E-18
CD2667 (PtsG)	PTS system, glucose-specific IIbc component	-1.873657473	8.56E-23
CD2668 (LicT)	putative transcription antiterminator	-0.949017301	8.76E-05
CD2762 (UppS)	undecaprenyl pyrophosphate synthetase	-1.962619932	0.001002204
CD2767	cell surface protein	-0.922492114	6.61E-05
CD2809	conserved hypothetical protein	-26.622358	0.000740048

Identifier (gene name)	**Product**	**Log_2 fold change: 630Δ*spo0A* vs 630**	***P* value**
CD2833	putative transport-related ATPase	-1.841321914	0.008058772
CD2864	putative esterase/halogenase	-27.230417	0.000219189
CD2865	putative bacterioferritin	-26.713903	0.000418474
CD2925	putative phage protein	-1.55946053	0.000106244
CD2955 (NtpB)	V-type sodium ATP synthase subunit B	-1.111520268	1.44E-05
CD2956 (NtpA)	V-type sodium ATP synthase subunit A	-1.010969553	0.000695137
CD2960 (NtpI)	V-type sodium ATP synthase subunit I	-0.825045552	0.00315282
CD2976	conserved hypothetical protein	-1.006181833	0.003230324
CD2978	conserved hypothetical protein	-0.806426976	0.009956438
CD2979	conserved hypothetical protein	-1.097398909	0.008285489
CD2981	hypothetical protein	-1.011829422	0.004867338
CD2982	conserved hypothetical protein	-1.336389137	0.003404511
CD3024	hypothetical protein	-3.727826742	0.000245753
CD3115 (BglA2)	putative 6-phospho-beta-glucosidase	-1.967807773	1.07E-12
CD3116 (BglF)	PTS system, beta-glucoside-specific IIabc component	-1.641631203	1.04E-08
CD3117 (BglG)	putative beta-glucoside bgl operon transcription antiterminator	-1.541172813	6.42E-06
CD3136 (BglA3)	6-phospho-beta-glucosidase	-2.654221294	4.92E-08
CD3137	PTS system, IIabc component	-2.865759926	1.59E-11
CD3138	transcription antiterminator	-1.6703225	7.13E-06
CD3145	putative serine-aspartate-rich surface anchored fibrinogen-binding protein	-2.129318471	2.66E-16
CD3174 (GapB)	glyceraldehyde-3-phosphate dehydrogenase 2	-0.982134737	5.22E-06
CD3175 (CggR)	central glycolytic genes regulator	-0.741566216	0.001307703
CD3235	single-stranded DNA binding protein	-3.810725337	2.20E-08
CD3257	putative polysaccharide deacetylase	-3.722115835	6.84E-05
CD3285 (Pgi)	glucose-6-phosphate isomerase	-1.399851788	3.27E-10
CD3298	putative ATP/GTP-binding protein	-4.276778621	1.44E-05
CD3314 (HydA)	hydrogenase	-0.831654059	0.008974023
CD3457	putative exported protein	-3.773602179	0.00397641
CD3489	putative oligopeptidase	-2.837469603	4.84E-06
CD3490 (SpoIIE)	stage II sporulation protein E	-7.151965308	4.41E-17
CD3522	conserved hypothetical protein	-3.870995792	4.04E-06
CD3563	putative spore cortex-lytic enzyme	-3.364344455	0.00280839
CD3569	putative peptidase	-1.517845811	0.004165547

Identifier (gene name)	**Product**
CD2010 (Erm1(B))	rRNA adenine N-6-methyltransferase (erythromycin resistance protein)
CD2352 (GrdA)	glycine/sarcosine/betaine reductase complex component A
CD1717	conserved hypothetical protein
CD2227	putative radical SAM superfamily lipoprotein
CD1767 (GapA)	glyceraldehyde-3-phosphate dehydrogenase
CD0065	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase
CD1177 (FapR)	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)
CD1799	tellurium resistance protein
CD1037	hypothetical protein
CD2348 (GrdD)	glycine/sarcosine/betaine reductase complex component C alpha subunit
CD1712 (MoaB)	putative molybdenum cofactor biosynthesis protein
CD1226	putative exported aminodeoxychorismate lyase
CD1165	conserved hypothetical protein
CD3438 (CobU)	bifunctional adenosylcobalamin biosynthesis protein [includes: adenosylcobinamide kinase and adenosylcobinamide-phosphate guanylyltransferase]
CD3015	PTS system, IIa component
CD2128 (IspG)	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
CD0153 (HpdB)	4-hydroxyphenylacetate decarboxylase, catalytic subunit
CD1316 (RpsO)	30S ribosomal protein S15
CD2558 (CoaD)	phosphopantetheine adenylyltransferase
CD3090 (TreR)	GntR-family transcriptional regulator
CD0865	conserved hypothetical protein
CD1838 (AroK)	shikimate kinase
CD3157 (SmpB)	SsrA-binding protein
CD2114	two-component system response regulator
CD2622	conserved hypothetical protein
CD0835	MarR-family transcriptional regulator
CD2764	putative hydrolase
CD2796	cell surface protein
CD2008	putative plasmid replication protein
CD0265 (FleN)	flagellar number regulator
CD0829	metallo-beta-lactamase superfamily protein
CD1676 (Pcp)	pyrrolidone-carboxylate peptidase
CD3290	putative protein translocase
CD3039	conserved hypothetical protein
CD0208	PTS system, IIb component
CD1209 (RecN)	DNA repair protein
CD3405 (HymA)	putative iron-only hydrogenase, electron-transferring subunit
CD1816 (Cmk)	cytidylate kinase
CD0279	conserved hypothetical protein
CD1979	ABC transporter, substrate-binding protein
CD3555	conserved hypothetical protein
CD1618	ABC transporter, ATP-binding protein
CD1400	conserved hypothetical protein
CD0081 (RpsQ)	30S ribosomal protein S17
CD2748	putative transferase
CD2588 (Gmk)	guanylate kinase
CD2158 (GabT)	4-aminobutyrate aminotransferase
CD1930	putative exported protein
CD2797	putative exported protein
CD1728	conserved hypothetical protein
CD0709	putative DNA mismatch repair protein
CD2211	ABC transporter, ATP-binding/permease protein

Identifier (gene name)	**Product**
CD2745 (Apt)	adenine phosphoribosyltransferase
CD3281 (ProC2)	pyrroline-5-carboxylate reductase
CD2240 (NanA)	N-acetylneuraminate lyase
CD1335	MarR-family transcriptional regulator
CD1931 (LexA)	SOS regulatory protein
CD1146 (MreC)	putative rod shape-determining protein
CD2805 (RuvB)	holliday junction DNA helicase
CD0104 (RplM)	50S ribosomal protein L13
CD1654 (LplA)	putative lipoate-protein ligase
CD0242	conserved hypothetical protein
CD1715	putative molybdenum cofactor sulfurase
CD1797	putative pyridine nucleotide-disulfide oxidoreductase
CD3456	putative 5-formyltetrahydrofolate cyclo-ligase
CD2217 (AroD)	3-dehydroquinate dehydratase
CD0209	putative sugar-phosphate kinase
CD1207 (Dxs)	1-deoxy-D-xylulose 5-phosphate synthase
CD3517 (PurR)	putative transcriptional repressor
CD0589	hypothetical protein
CD1522	putative polysaccharide deacetylase
CD2100	putative [2Fe-2S]-binding subunit of oxidoreductase
CD1507	putative thioredoxin
CD0768 (SrlD)	sorbitol-6-phosphate 2-dehydrogenase
CD0114	putative ATP/GTP-binding protein
CD2646 (FtsZ)	cell division protein
CD2765	putative penicillin-binding protein repressor
CD0111	putative exported protein
CD2168 (Hcp)	hydroxylamine reductase
CD0270 (FliM)	putative flagellar motor switch protein
CD3035	conserved hypothetical protein
CD3616	putative Na+/H+ exchanger
CD1495 (ProC1)	pyrroline-5-carboxylate reductase
CD2332 (MtlF)	PTS system, mannitol-specific IIa component
CD3245 (PrdR)	sigma-54-dependent transcriptional activator
CD1494	putative transcriptional regulator

Identifier (gene name)	**Product**
CD0907	CD0907 putative phage regulatory protein
CD3116 (BglF)	PTS system, beta-glucoside-specific IIabc component
CD0687 (RplT)	50S ribosomal protein L20
CD2195	ferritin
CD0142	putative RNA-binding protein
CD2961	conserved hypothetical protein
CD2259	conserved hypothetical protein
CD1264	hypothetical protein
CD0756	putative reductase
CD0183	putative cell wall hydrolase
CD1660	conserved hypothetical protein
CD0869 (ModA)	putative molybdenum ABC transporter, substrate-binding protein
CD0735	hypothetical protein
CD3027 (Crr)	PTS system, glucose-specific IIa component
CD0873	ABC transporter, substrate-binding lipoprotein
CD1822 (Bcp)	putative thiol peroxidase (bacterioferritin comigratory protein)
CD1771	conserved hypothetical protein
CD2818	ABC transporter, ATP-binding/permease protein
CD3188	hypothetical protein
CD2778	putative polysaccharide biosynthesis protein
CD3315 (HydN2)	electron transport protein
CD1225 (DeoA)	pyrimidine-nucleoside phosphorylase
CD2964	conserved hypothetical protein
CD2989	ABC transporter, substrate-binding protein
CD2181	putative aromatic compounds hydrolase
CD3220	putative methyltransferase
CD1769	conserved hypothetical protein
CD3567	putative cell wall hydrolase
CD2213	putative carbonic anhydrase
CD0011 (SigB)	RNA polymerase sigma-B factor
CD2258	hypothetical protein
CD1471	putative membrane protein (putative phage infection protein)
CD0704	putative phosphohydrolase
CD2165	conserved hypothetical protein
CD0007	putative exported protein
CD2355 (TrxA2)	thioredoxin
CD3087	phosphosugar-binding transcriptional regulator
CD0850	conserved hypothetical protein
CD0882 (GlgC)	glucose-1-phosphate adenylyltransferase
CD1056 (EtfA2)	electron transfer flavoprotein alpha-subunit
CD1459	putative 5-nitroimidazole reductase
CD1054 (Bcd2)	butyryl-CoA dehydrogenase
CD3137	PTS system, IIabc component
CD1055 (EtfB2)	electron transfer flavoprotein beta-subunit
CD0492	PTS system, IIb component
CD3516 (SpoVG)	stage V sporulation protein G
CD3094	putative sigma-54-dependent transcriptional regulator
CD1126A	putative transcriptional regulator
CD1058 (Hbd)	3-hydroxybutyryl-CoA dehydrogenase
CD2873	putative signaling protein
CD2127	putative membrane protein
CD1484 (SsuA)	putative aliphatic sulfonates ABC transporter, substrate-binding lipoprotein

Identifier (gene name)	**Product**
CD1837 (AroE)	shikimate dehydrogenase
CD1059 (ThlA1)	acetyl-CoA acetyltransferase
CD0440	cell surface protein
CD0942	hypothetical phage protein
CD2311	ABC transporter, substrate-binding protein
CD1057 (Crt2)	3-hydroxybutyryl-CoA dehydratase
CD2688 (SspA)	small acid-soluble spore protein A
CD1214 (Spo0A)	stage 0 sporulation protein A
CD0810 (FloX)	Putative flavodoxin/nitric oxide synthase
CD3093	putative glutamine amidotransferase
CD3092	putative amino acid permease
CD0451	Putative dioxygenase