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).

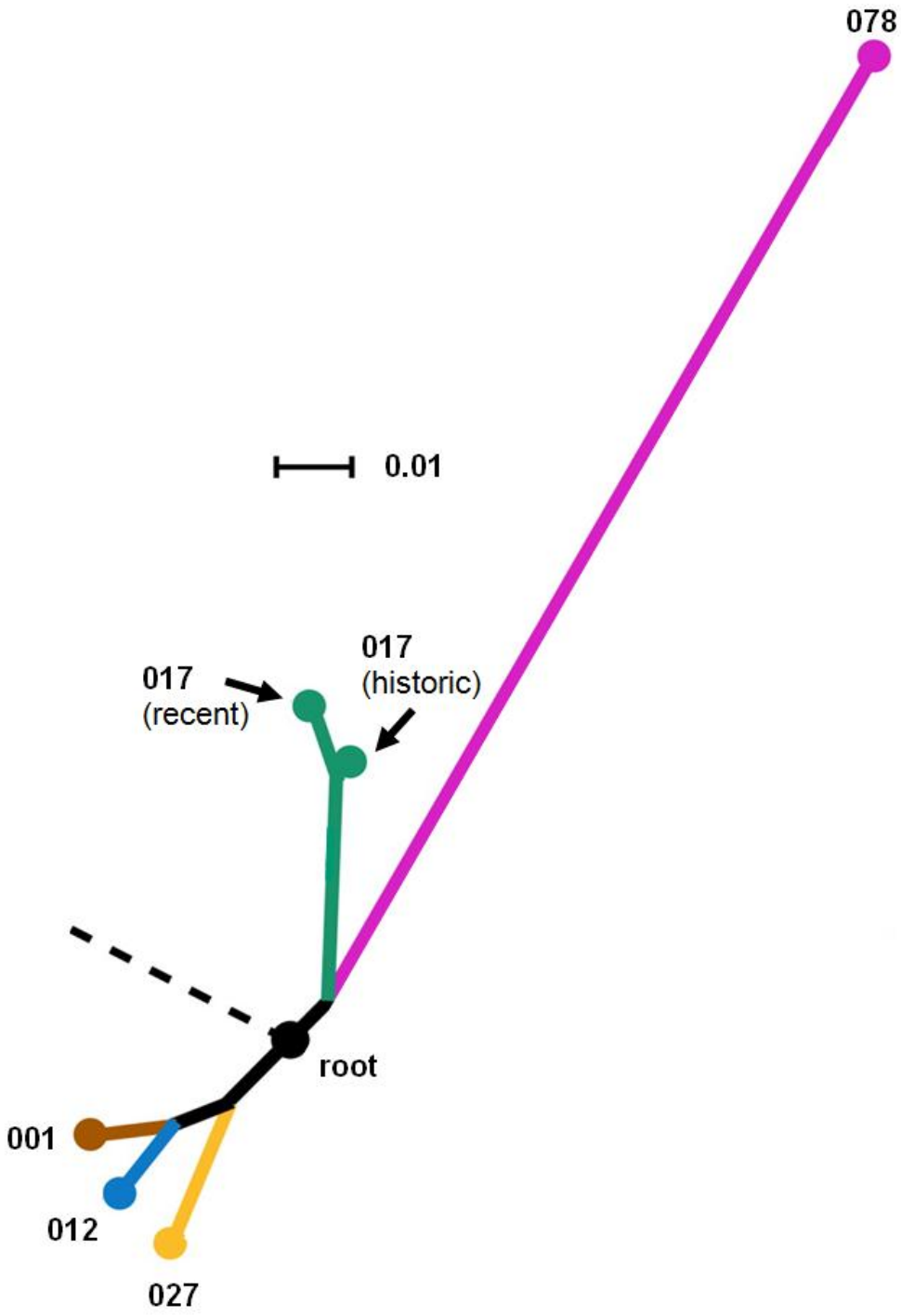


Figure 1.1. Deep branching phylogenetic tree of human virulent *C. difficile*. Phylogeny of *C. difficile* based on whole-genome sequences, describing the association between common human virulent *C. difficile* ribotypes. The root connects to *Clostridium bartlettii* and *Clostridium hiranonis*. Pathogenic 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

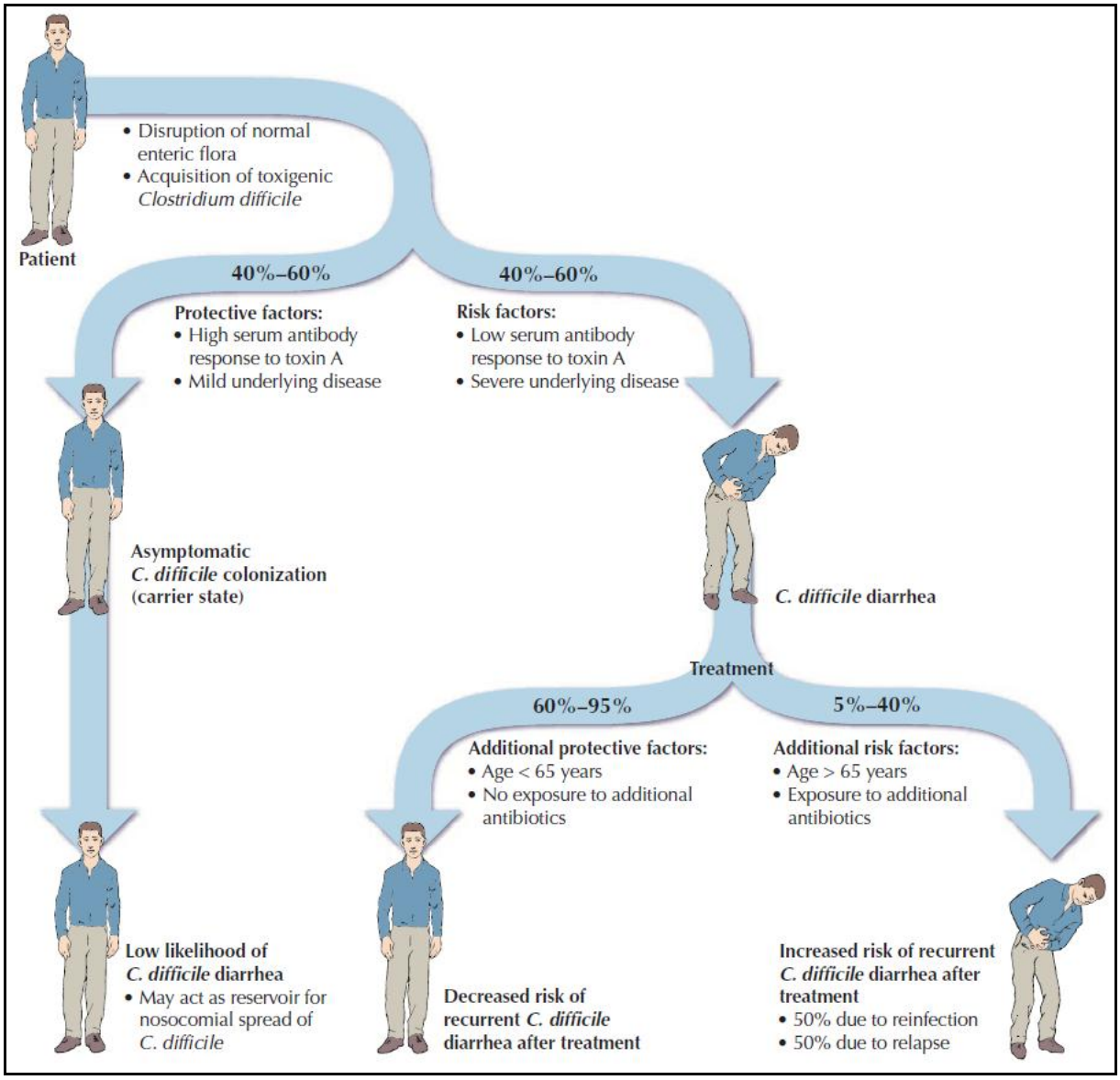


Figure 1.2. Key determinants in the acquisition and expression of CDAI. Flow diagram describing the events leading to *C. difficile* colonisation, and the development of systematic disease. A triad of events are typically associated with the development of *C. difficile* disease, and include (i) perturbation 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

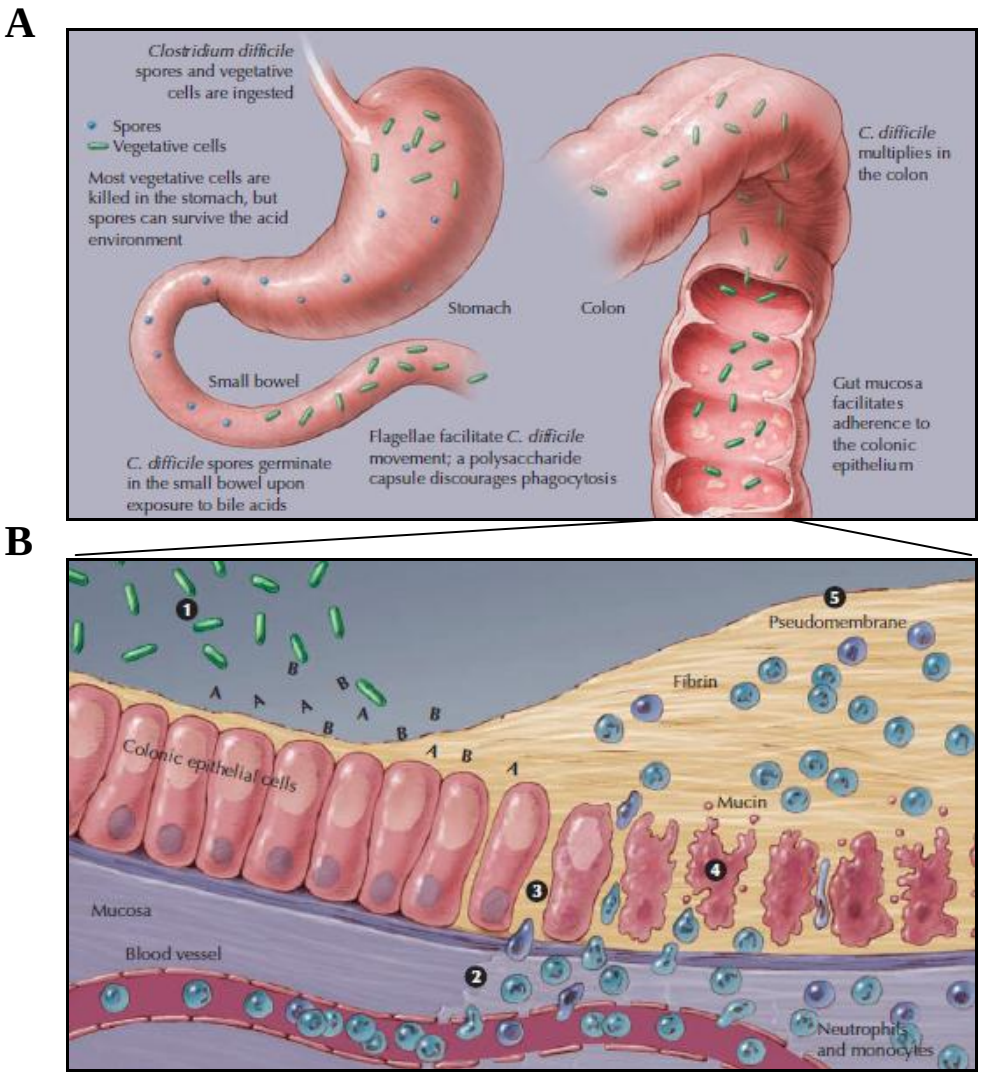


Figure 1.3. *C. difficile* colonisation and pathogenesis. A) Schematic representation of the gastrointestinal tract, describing the transit of *C. difficile* to the colon as the principle site of infection. *C. difficile* is ingested from the environment, and passes through the stomach where vegetative 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).

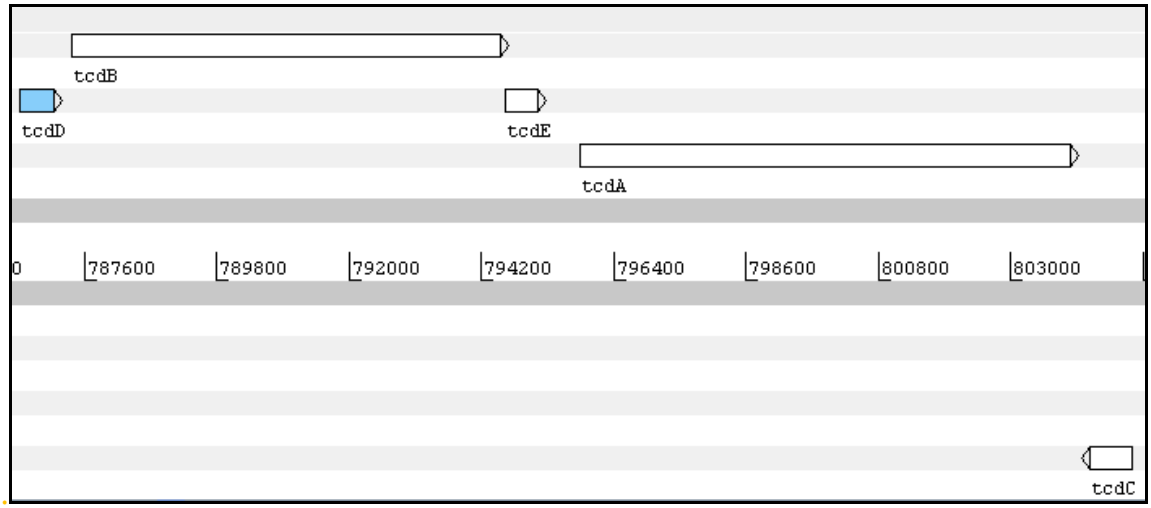
toxin B, which are regarded as the major virulence factors of the pathogen (68). There has been much debate surrounding the relative contribution of toxins A and B in recent years. It was initially thought that toxin A was the primary virulence determinant, and early experiments indicated that toxin 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

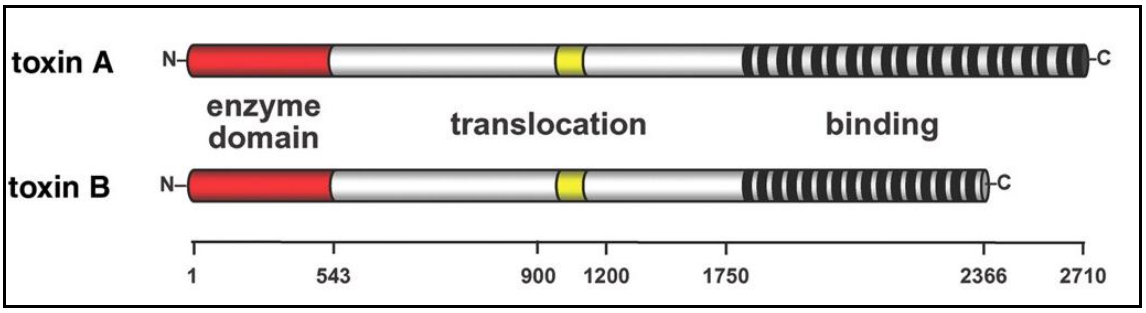


Figure 1.4. Organisation of the *C. difficile* PaLoc region and toxin structure. A) Artemis screenshot displaying the *C. difficile* PaLoc region from *C. difficile* 630, which harbours five genes (*tcdDBEAC*) responsible for the synthesis and regulation of toxins A and B. **B)** Structural organisation 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).

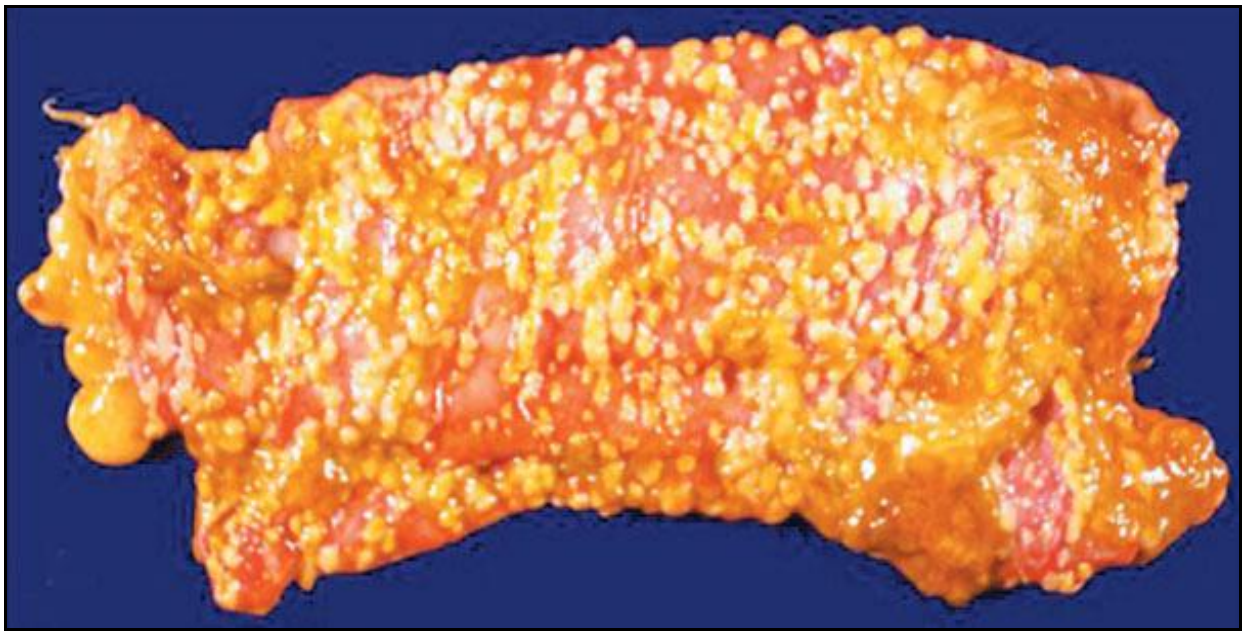


Figure 1.5. Psuedomembranous colitis due to *C. difficile* infection. Image of a resected colon, in which pseudomembranes are evidenced as raised yellowish exudative plaques over the colonic mucosa. A thickened bowel wall is also notable. Taken from: (91).

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).

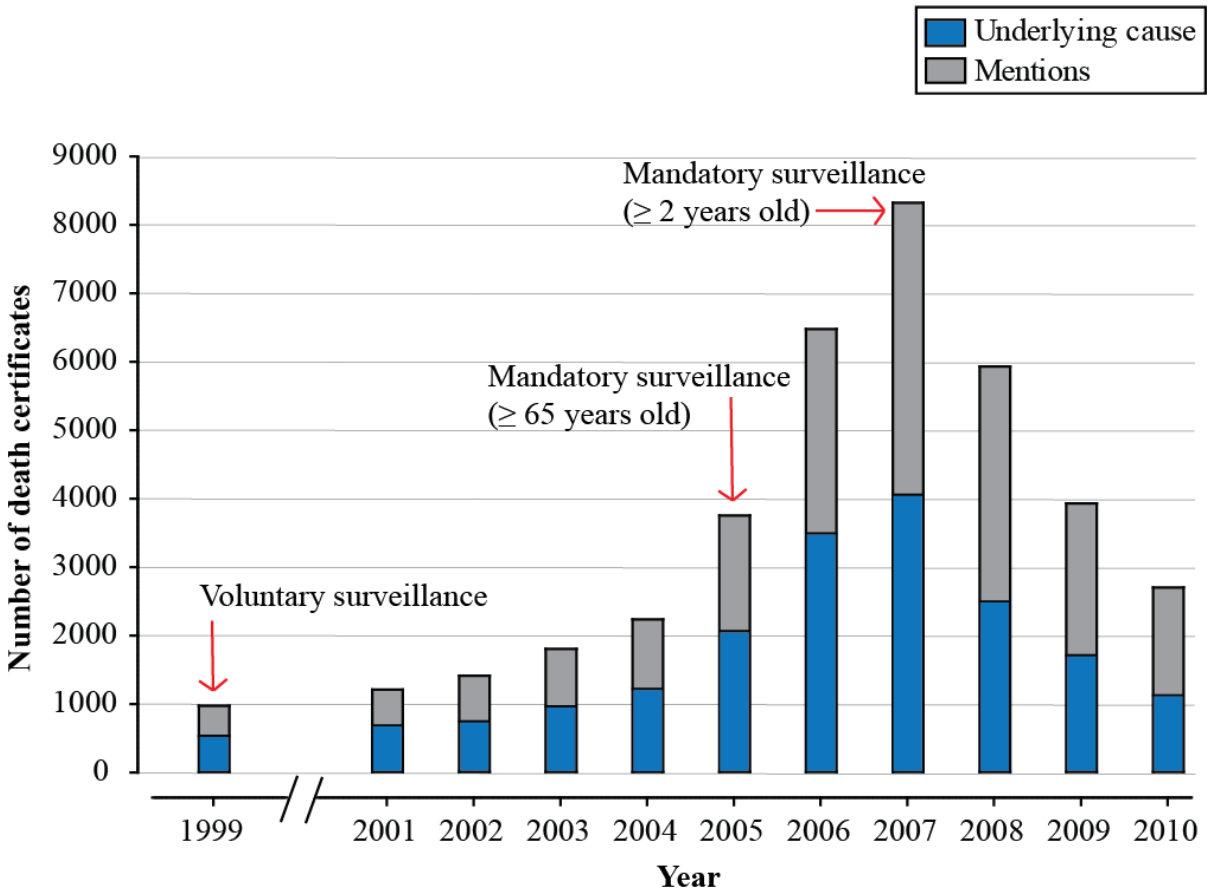


Figure 1.6. Number of death certificates citing *C. difficile*. Graph detailing the number of death certificates which either mention *C. difficile*, or cite it as the underlying cause of death. Data were collected from England and Wales, between 1999 and 2010 (data was not available for 2000). Also 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

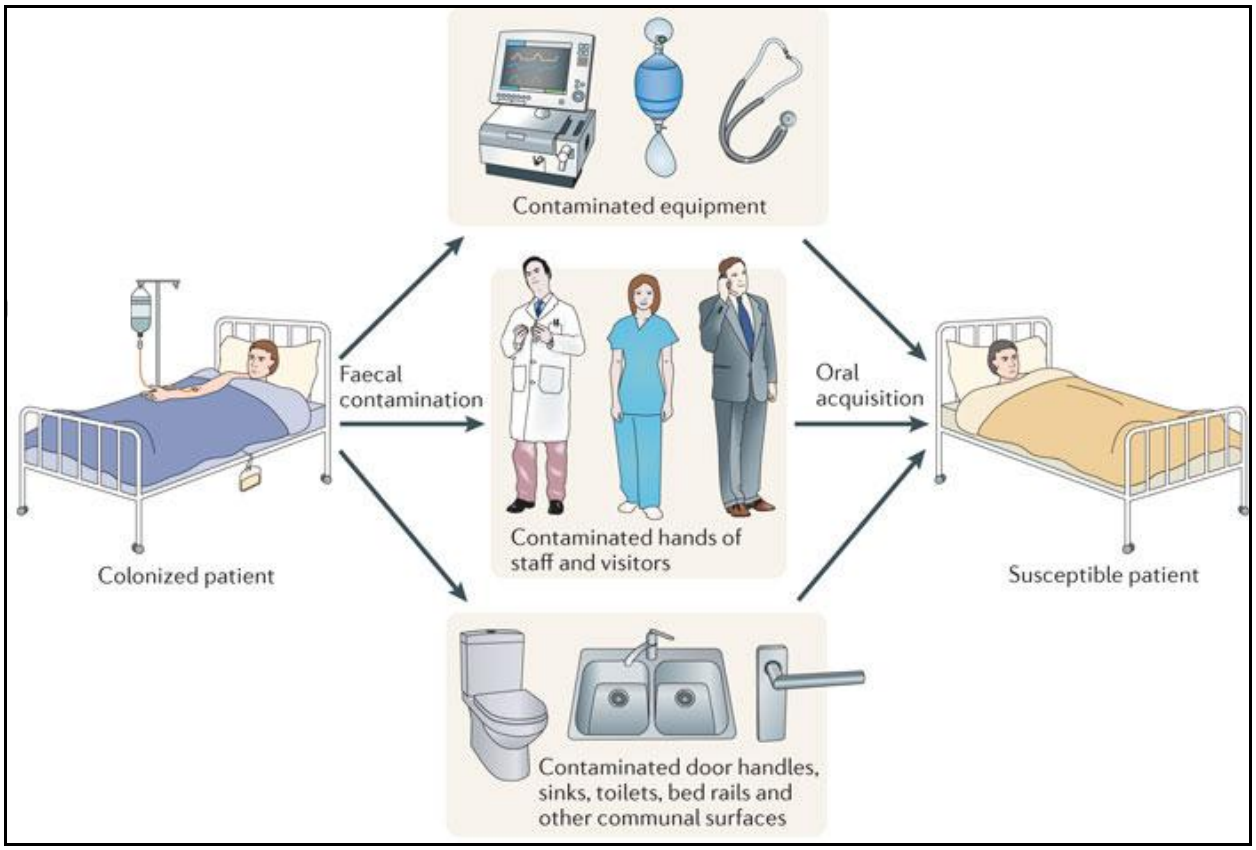


Figure 1.7. Primary reservoirs of healthcare-associated *C. difficile* infection. *C. difficile* is widely disseminated within the healthcare environment, and is able to survive for extended periods on medical personnel, patients, visitors, fomites, and almost all environmental surfaces such as bedrails, 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.

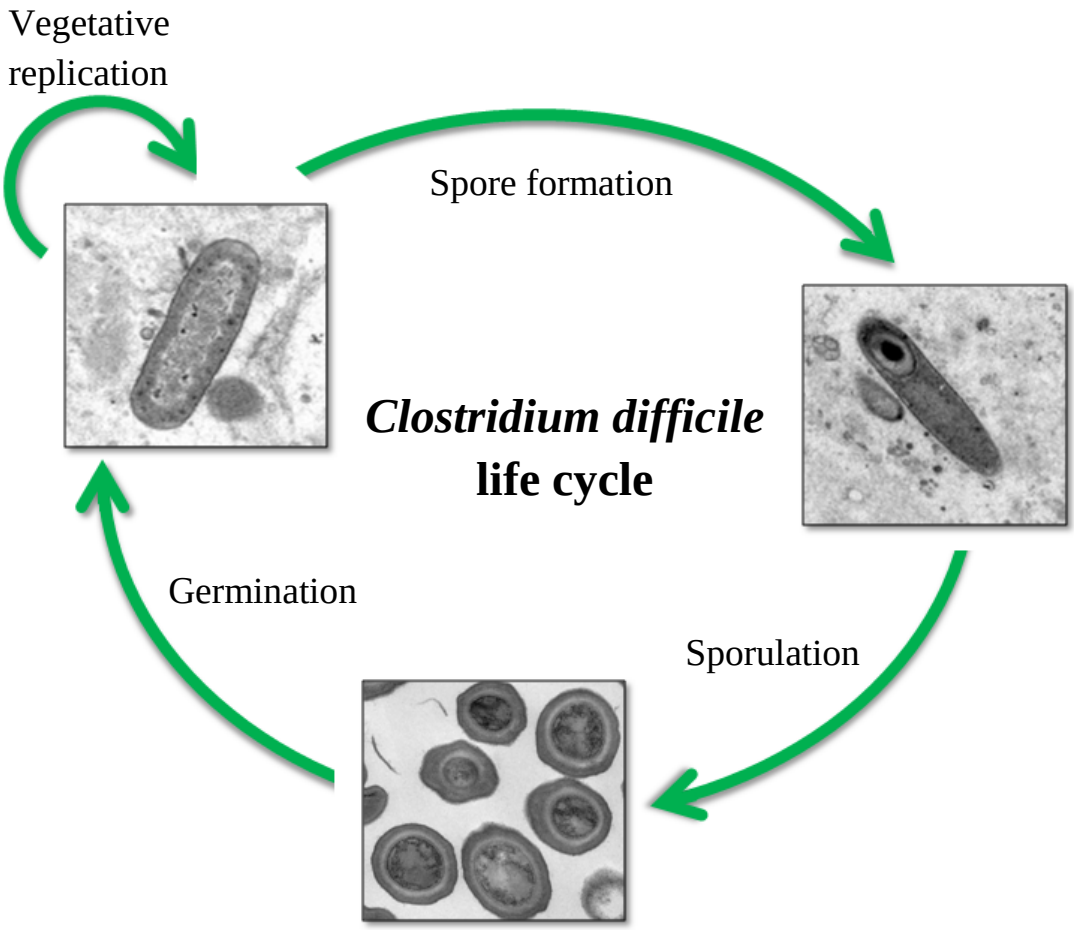


Figure 1.8. Schematic representation of the *Clostridium difficile* life cycle. When conditions are permissive, *C. difficile* is metabolically active and undergoes vegetative replication. In periods of stress, *C. difficile* is able to enter a cycle of spore formation, during which a metabolically dormant 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

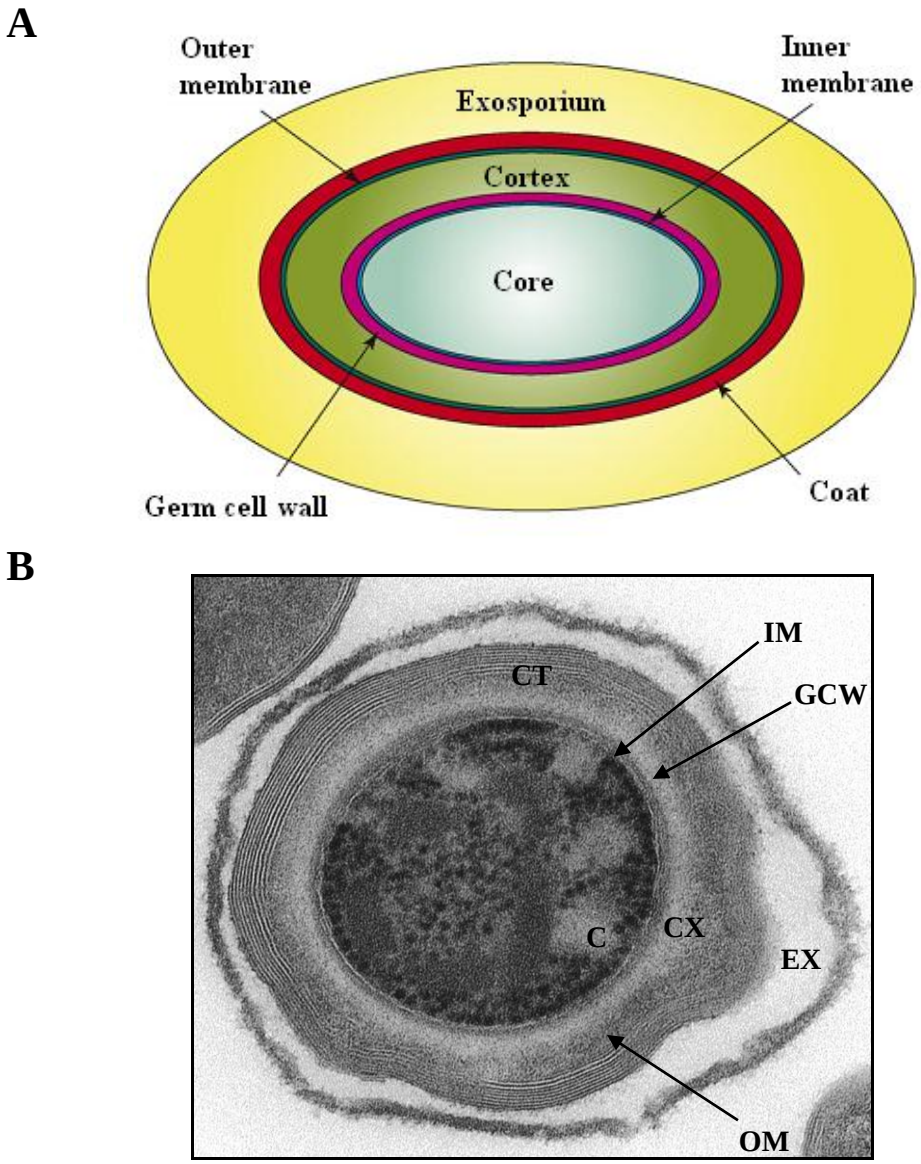


Figure 1.9. Ultrastructure of the bacterial endospore. A) Schematic representation of the salient features of a bacterial spore. Layers are not drawn to scale. Adapted from: (161). **B)** TEM image of a mature *C. difficile* spore, comprising a thick mantle of concentric layers. Shown innermost is the 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

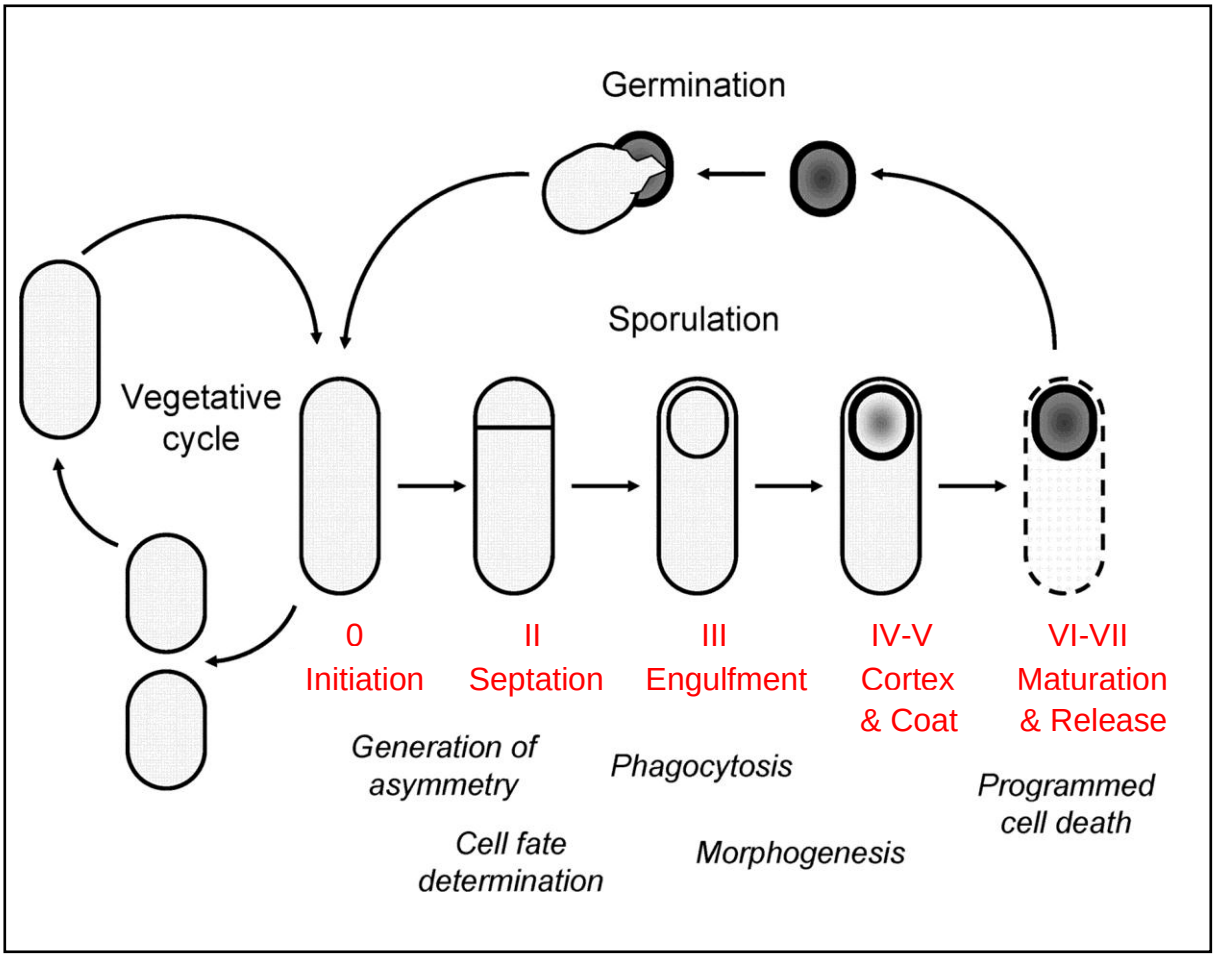


Figure 1.10. Schematic overview of spore formation, sporulation and germination. The cycle of spore formation is defined by 7 cytological stages (I-VII, excluding stage I), in which key morphological indicators (shown in red) are used to define the physiological state of the cell. Vegetative growth is 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).

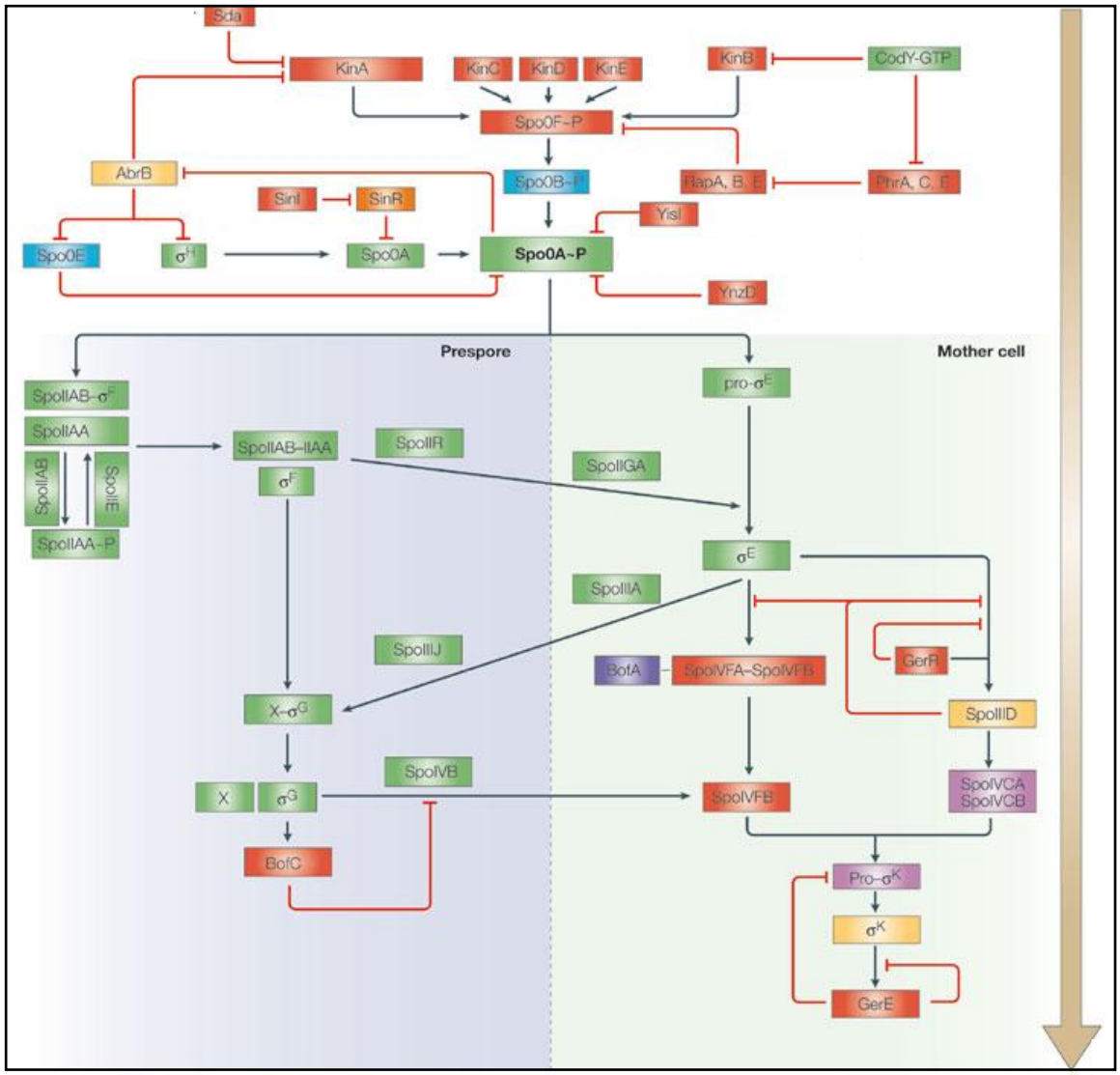


Figure 1.11. Canonical overview of clostridial spore formation. Schematic representation of the *B. subtilis* sporulation cascade, with associated clostridial orthologs. Orthologs are coloured according to presence or absence as follows: red, absent in all clostridia; green, present in all clostridia; 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.

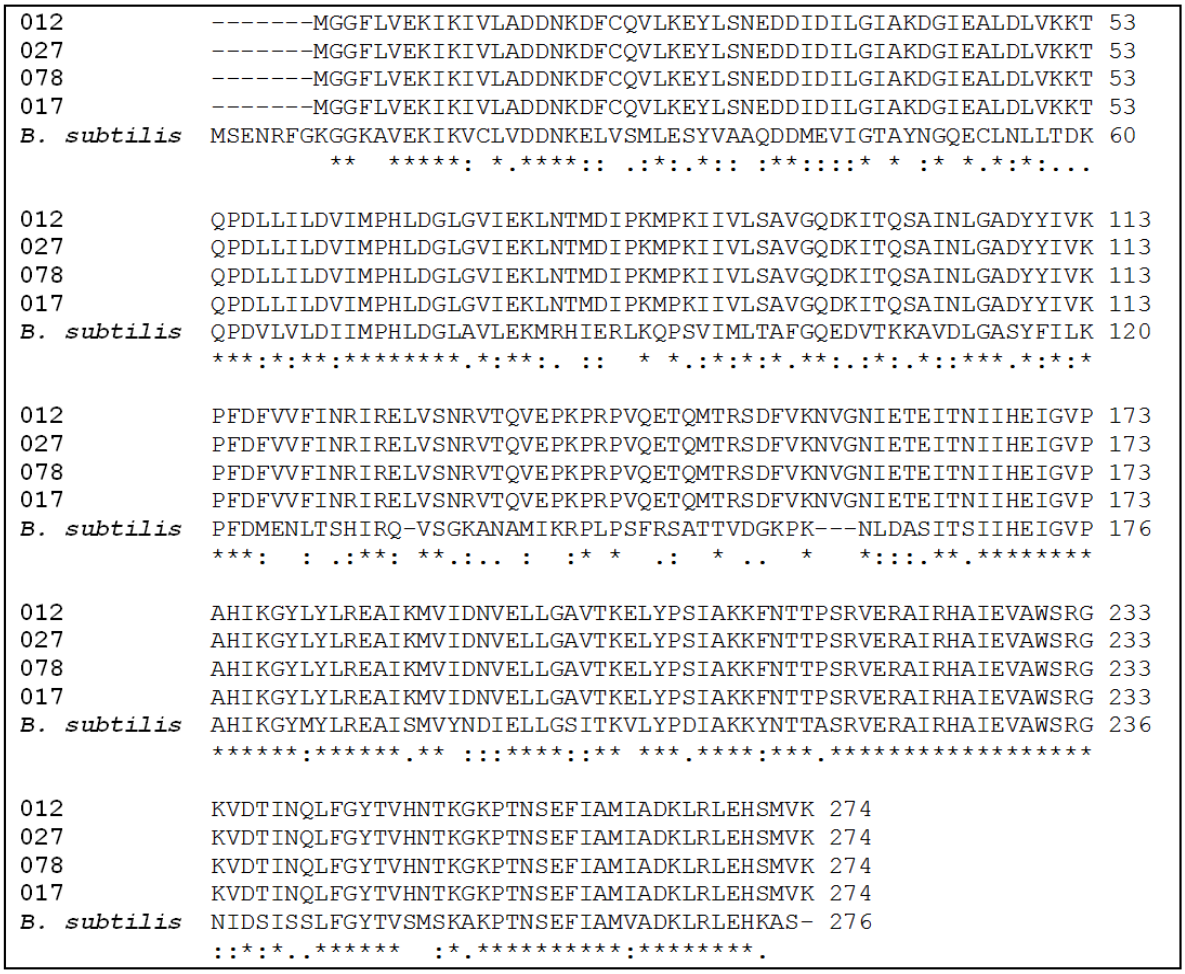


Figure 1.12. Amino acid sequence alignment of Spo0A from common *C. difficile* ribotypes. ClustalW2 (2.1) multiple sequence alignment of Spo0A from commonly isolated human virulent *C. difficile* ribotypes. *B. subtilis* Spo0A is included for reference. Symbols: "*", identical; ".", conserved substitution;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