Studies on the gene regulatory networks that control sporulation initiation in *Clostridioides difficile*

by

Babita Adhikari Dhungel

B. Tech in Biotechnology, Kathmandu University, Nepal, 2013

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

Clostridioides difficile is a Gram-positive, spore-forming anaerobic bacillus and is the causative agent of *C. difficile* Infection (CDI). CDI is the consequence of antibiotic therapy, which disrupts the commensal gut microbiota, thus providing a favorable environment for an opportunistic pathogen like *C. difficile* to outgrow and cause infection. *Clostridioides difficile* produces toxins A & B, that damage host epithelial cells leading to diarrheal symptoms. *Clostridioides difficile* also produces metabolically dormant spores that are resistant to antibiotics and detergents and are the major cause of transmission and persistence of the disease. Hyper-sporulation observed in the strains that emerged in the early 2000s is attributed to their higher transmission rates and, hence, increased CDI incidence in Europe and North America. Sporulation pathways are poorly characterized in *C. difficile* and the mechanism by which sporulation initiation occurs in this pathogen is not well understood.

Here, I present data from two different studies on *C. difficile* sporulation. In our first study, we determined the regulatory relationship between Spo0A, the sporulation master regulator with SinR, a pleiotropic regulator in *C. difficile*. By using genetic and biochemical approaches, we demonstrated that Spo0A directly binds to the upstream of *sin* locus and represses the expression of *sinR*.

In the second part of the study, we identified the novel link between intracellular c-di-GMP concentration and sporulation. We found that the upregulation of *pdcB* coding for a phosphodiesterase reduces the intracellular concentration of c-di-GMP. This change in the level of c-di-GMP positively influenced sporulation in *C. difficile* UK1 strain. By showing that a DNA inversion regulates *pdcB* gene expression, we have discovered the phase variable regulation of intracellular c-di-GMP concentration. Additionally, by finding that CodY, a global nutritional

regulator, binds to and represses the expression of *pdcB*, we redefined CodY mediated gene regulatory networks of sporulation initiation in *C. difficile*.

Together, findings from our studies have added new players and links in the gene regulatory network of *C. difficile* sporulation.

Studies on the gene regulatory networks that control sporulation initiation in *Clostridioides difficile*

by

Babita Adhikari Dhungel

B. Tech in Biotechnology, Kathmandu University, Nepal, 2013

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Revathi Govind, PhD

Copyright

Abstract

Clostridioides difficile is a Gram-positive, spore-forming anaerobic bacillus and is the causative agent of *C. difficile* Infection (CDI). CDI is the consequence of antibiotic therapy, which disrupts the commensal gut microbiota, thus providing a favorable environment for an opportunistic pathogen like *C. difficile* to outgrow and cause infection. *Clostridioides difficile* produces toxins, A & B that damage host epithelial cells leading to diarrheal symptoms. *Clostridioides difficile* also produces metabolically dormant spores that are resistant to antibiotics and detergents and are the major cause of transmission and persistence of the disease. Hyper-sporulation observed in the strains that emerged in the early 2000s is attributed to their higher transmission rates and, hence, increased CDI incidence in Europe and North America. Sporulation pathways are poorly characterized in *C. difficile* and the mechanism by which sporulation initiation occurs in this pathogen is not well understood.

Here, I present data from two different studies on *C. difficile* sporulation. In our first study, we determined the regulatory relationship between Spo0A, the sporulation master regulator with SinR, a pleiotropic regulator in *C. difficile*. By using genetic and biochemical approaches, we demonstrated that Spo0A directly binds to the upstream of *sin* locus and represses the expression of *sinR*.

In the second part of the study, we identified the novel link between intracellular c-di-GMP concentration and sporulation. We found that the upregulation of *pdcB* coding for a phosphodiesterase reduces the intracellular concentration of c-di-GMP. This change in the level of c-di-GMP positively influenced sporulation in *C. difficile* UK1 strain. By showing that a DNA inversion regulates *pdcB* gene expression, we have discovered the phase variable regulation of intracellular c-di-GMP concentration. Additionally, by finding that CodY, a global nutritional

regulator, binds to and represses the expression of *pdcB*, we redefined CodY mediated gene regulatory networks of sporulation initiation in *C. difficile*.

Together, findings from our studies have added new players and links in the gene regulatory network of *C. difficile* sporulation.

Table of Contents

List of figures

List of supplementary figures

List of tables

List of supplementary tables

Acknowledgements

I would like to express my deepest appreciation to my major Professor, Dr. Revathi Govind, for her continuous support and supervision in my research work in the past three years. Without her guidance and persistent help, I would not have been able to progress in my projects and put together this dissertation. I am also very grateful to Dr. Revathi for giving me the freedom to balance my personal and professional life.

I would like to thank my committee members Dr. Philip Hardwidge, Dr. Stephanie Shames, and Dr. Tom Platt for their continuous support and valuable inputs in my projects.

I am very grateful to my lab members Brintha, Kamrul, Yusuf, Carolina, and Thomas for being friends and compatible colleagues to work with. I am especially grateful to Brintha who trained me during my initial days in the lab.

In addition, I would like to thank Dr. Stella Lee and the then committee members Dr. Anna Zolkiwieska, Dr. Jeroen Roelofs, Dr. Ruth Welti, and Dr. Zhilong Yang for the support and guidance they provided me in the first 2 years of my graduate school career when I was a graduate student in Dr. Lee's lab.

I would like to thank Division of Biology for the funding and opportunity to conduct research in its prestigious program.

I would like to thank my family members especially my Dad, Mr. Chiranjibi Adhikari, and my Mom, Late Mrs. Bhagawati Adhikari, and Kamala Adhikari for raising me with good values and work ethics. Being the youngest one in the family, I was privileged to go to the best academic institutions in Nepal and get the best education which has made me who I am today. I am forever grateful to my Dad for all the sacrifices he has made for me and my family. I am grateful to have

siblings like my brother Arun, sister-in-law Asmita, sisters Anita and Sunita, brothers-in-laws Dip and Sudhir and nephews Aayan, Adrit and Alvin. They have been my rock and I am proud to call them my family. I want to remember and express my deepest gratitude to my grandparents, Mr. Umakanta Adhikari and Late. Mrs. Kamala Adhikari, who instilled in me the spirituality and moral values which I always abide by. Above all, I am thankful to Pragyesh, my best friend, my life partner for being there for me physically, mentally and emotionally during my happy and sad days. You mean so much to me, I love you!

Dedication

I want to dedicate this work to my family who have always believed in me and have supported me in every possible ways!

Chapter 1 - Introduction

1.1 Introduction to *Clostridioides difficile*

Clostridioides (Clostridium) difficile is a Gram-positive, rod-shaped, motile, spore-forming obligate anaerobe (Fig 1.1). All strains of *C. difficile* grow in neutral to alkaline pH (pH 5.5 -9.0). They are oxidase and catalase-negative and are mostly mesophilic. *Clostridioides difficile* has rapidly emerged as a leading cause of nosocomial infection worldwide and is the number one cause of hospital- acquired infections in North America. It has been categorized as an urgent threat to public health by Center for Disease Control (CDC) because of its ability to cause significant morbidity and mortality in patients and recurrent infection (1–3).

In this chapter, I will review the emergence of *C. difficile* as a human pathogen, its mode of infection, epidemiology, pathogenesis, treatment, and prevention approaches. I will also review the literature on the gene regulatory networks of sporulation which is the main focus of my study.

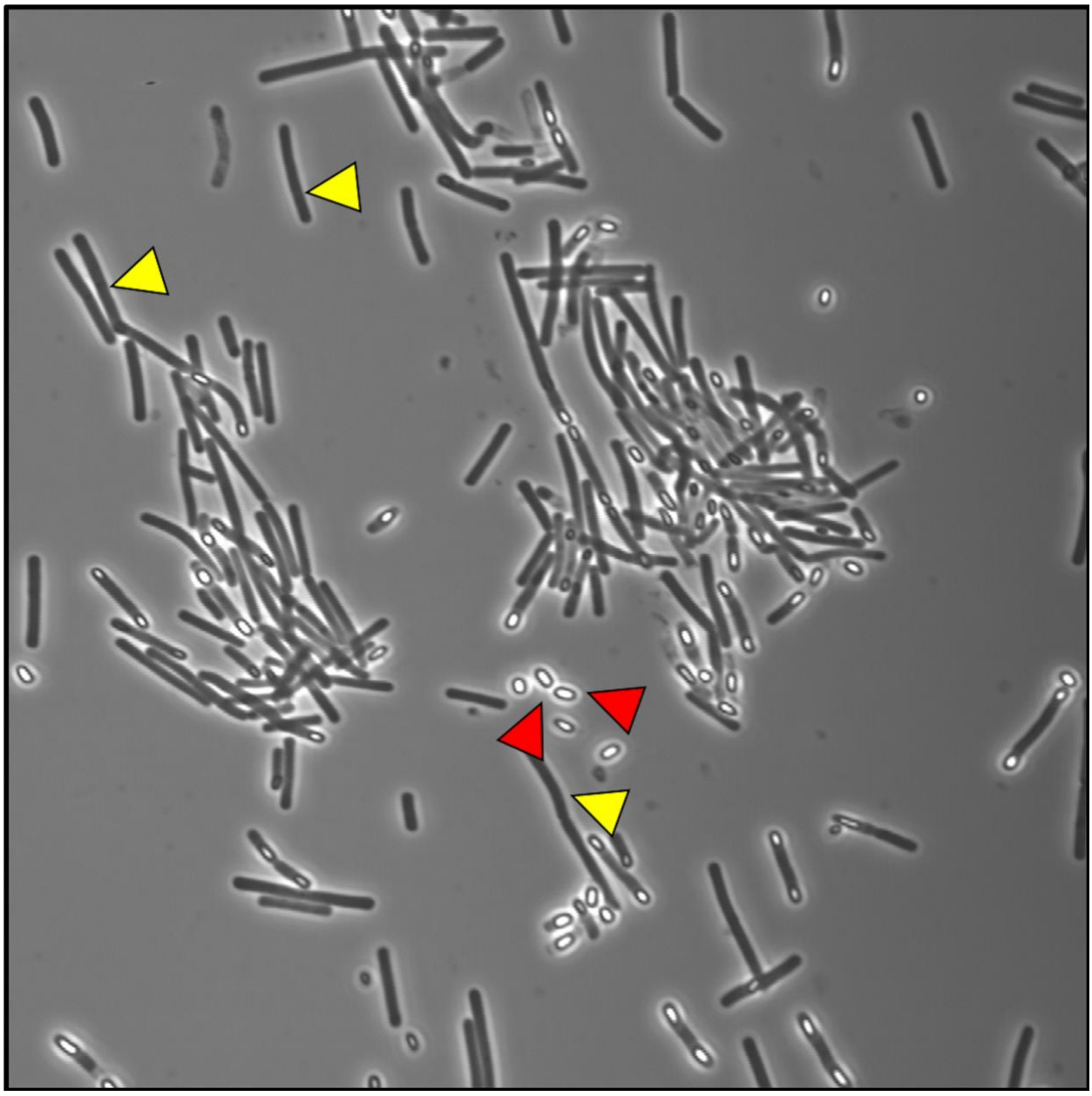


Figure 1.1. Phase contrast *microscopy* image of *C. difficile*

The image depicts the rod shaped vegetative cells (Yellow triangle) and exospores (Red triangles) of *C. difficile* UK1 strain.

1.2 History of *C. difficile*

Clostridioides difficile was first identified in the feces of breast-fed infants by Hall and O'Toole in 1935 (4). It was initially named as *Bacillus difficilis*, "bacillus" for its elongated rod shape cell morphology and "difficilis" for its difficulty in isolation and study (4–6). *Bacillus difficilis* was reclassified as *Clostridium difficile* in 1938 based upon its ability to produce toxins that exhibited symptoms in animals that were very similar to those of tetanus that is produced by *Clostridium tetani* (6). *Clostridioides difficile* was later suggested to be renamed as *Peptoclostridium difficile* based on its close relatedness to *Peptostreptococcaceae,* as demonstrated by the next-generation sequencing technologies (7). But to avoid the confusion that could be created by this new name, *C. difficile* has recently been named as *Clostridioides difficile,* meaning an organism that is similar to *Clostridium* (7).

Until the 1970s, few reports were published, which demonstrated isolation of *C. difficile* from the human gut from different cases of infections such as gas gangrene, abscess of bone fracture and vaginal vault, and peritoneal and pleural fluid (8–10). However, there was no direct evidence of *C. difficile* being pathogenic to humans. Several animal studies demonstrated that *C. difficile* toxin was not absorbed by the intestinal tract of animals and the toxin was pathogenic only upon injection into tissues (11). These observations suggested that *C. difficile* was a part of the normal human gut flora. Pathogenicity of *C. difficile* in humans was recognized only in 1978 when three simultaneous studies identified *C. difficile* as a causative agent of pseudomembranous colitis (PMC) (Fig 1.2). One of the studies found increased titers of *C. difficile* toxins in the feces of all patients with PMC (12). Another study attributed clindamycin-associated PMC, also known as "clindamycin-colitis" to *C. difficile* toxin (13). The third study reported that previous exposure to antibiotics increased susceptibility to *C. difficile* infection and identified *C. difficile* as a causative

agent of PMC (14). Pathogenicity of *C. difficile* further came into perspective, retrospectively, from the penicillin toxicity study, which reported cytopathic toxin produced by an uncultivable virus in response to penicillin treatment (15).

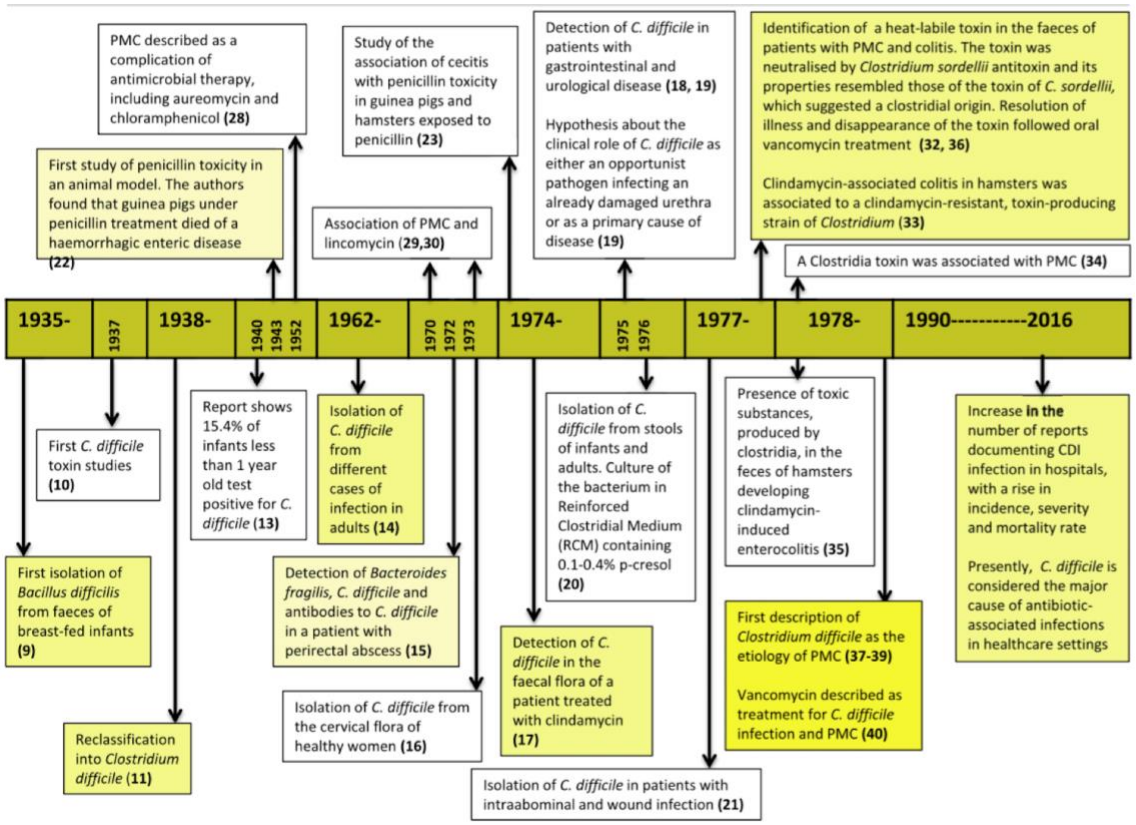


Figure 1.2. History of *C. difficile*

Image depicts the timeline of first isolation of *C. difficile* in humans and its recognition as a human pathogen associated with PMC. Image adapted from (6).

1.3 Epidemiology

Clostridioides difficile infection has become a global health problem. *C. difficile* infection has been reported in Australia, Asia, North America and Europe. Nationwide study has shown that CDI has increased significantly in Korea from 2004 to 2008 (16). Similarly, there was an increased incidence of CDI in Australia during 2011-2012 (17). A territory-wide population based study showed increasing CDI in Hong Kong and China from 2006-2014 (18). There is an increase in incidence, severity, and mortality due to CDI in North America (1,19) and Europe (20) in the past decade. The increased incidence of CDI has been attributed to the emergence of hypervirulent strains of *C. difficile* like PCR ribotype 027 and 078 (21–23). There were 15,512 CDI cases in 10 states of the United States in 2017, with an estimated national burden of 495,600 total infections (1). Increased incidence of CDI has also been observed in children (24,25) and peripartum women (26,27) who were initially considered to be at low risk.

1.4 Risk factors associated with CDI

1.4.1 Antibiotic treatment

CDI occurs primarily due to broad-spectrum antibiotic treatments that disrupt not only the pathogenic bacteria but also the commensal gut flora and provides a favorable environment for *C. difficile* to cause infection. Antibiotics such as clindamycin, cephalosporins, and fluoroquinolones are suggested to impose a high risk for CDI (Table 1.1) (28,29). These antibiotics have been associated with increased expression of colonization factors in *C. difficile* (30). Treatment with sub-inhibitory concentration (half of the minimum inhibitory concentration) of clindamycin upregulated the expression of *cwp84* that encodes colonization factor (30). This suggests the direct role of an antibiotic in inducing *C. difficile* infection. Antibiotic treatment shifts the gut microbial

community structure with decreased *Bacteroidetes* and increased *Proteobacter* populations that are part of the beneficial gut flora (30–32). Commensal gut microbes like *Bacteroidetes* are suggested to keep check on *C. difficile* growth and absence of these bacteria provides a conducive environment for *C. difficile* to cause infection (31).

Table 1.1. Antibiotics and their associated risk with *Clostridioides difficile* Infection (CDI) (19)

High Risk	Medium Risk	Low risk
Clindamycin	Sulfonamides	Aminoglycosides
Ampicillin	Trimethoprim	Bacitracin
Amoxicillin	Trimethroprim-sulphamethoxazole	Metronidazole
Cephalosporins	Macrolides	Teicoplanin
Fluroquinolones	Other penicillins	Rifampin
		Chloramphenicol
		Carbapenems
		Daptomycin
		Tigecycline
		Vancomycin

1.4.2 Proton Pump Inhibitors (PPI)

PPIs are drugs used to treat gastroesophageal reflux disease, peptic ulcer, and *Helicobacter pyroli* infection (31). PPIs have been shown to modulate the composition of gut microbiota with decreased *Bacteroidetes,* and increased *Firmicutes*, suggesting the role of PPIs in predisposing the development of CDI in people who take these drugs (31,33).

1.4.3 Age

Elderly population who are of age >65 years are at high risk of getting CDI. Although the major reason for CDI susceptibility in elderly people is unclear, weak immune system and increased comorbidity may play a role (34). Ageing causes a lot of physiological changes in the body including innate and adaptive immune dysfunction (35). Weak immune system causes upregulation of pro-inflammatory cytokines and reduction of phagocytic activity of immune cells like neutrophils, macrophages, and natural killer cells, both of which contribute to susceptibility to infection (36). Weak immune system fails to produce anti-toxin serum immunoglobulin G (IgG) in response to *C. difficile* toxins (28). Similarly, age related organ specific changes contribute to increase risk of infection (35). Reduced gut barrier function because of decreased intestinal motility and gastric acids, and alteration in the gut microbiome may also contribute to higher chance of getting CDI. Peniche et al. showed that aging suppress neutrophil and IL-22 mediated host immunity during CDI (37). Therefore, decline in immune system and organ functionality may make elderly population more susceptible to CDI.

1.4.4 Hospitalization

Long term stay at hospital or health care setting exposes an individual to spore containing environment like toilets, clinic furnishings, medical devices, etc., exposure to antibiotics and inadequate health hygiene by the health care workers, all of which impose a high risk of getting CDI (28,38).

1.4.5 Community-associated infections

Even though antibiotic treatments remain as significant causes of CDI, community-acquired infections are on the rise. Community-acquired infection occurs if a person has CDI without being or visiting the health care settings or after 4 weeks of getting discharged from the hospital or after being hospitalized for less than 48 hours (38). Although community CDI rates are very low in the United States, 7.7 cases per 100,000 people annually (28), CDI infection has also been observed in young individuals and pregnant women who were initially thought to be at low risk of getting the infection. The exact source of community acquired CDI is not known. Animals used as food like pigs which harbor *C. difficile* spores in their gut are speculated to be the source of *C. difficile* spores (39). However, no any foodborne illness outbreaks of *C. difficile* has been reported. Other environmental such as soil, water, pets, food, meats, and vegetables are linked to CDI (23).

1.5 Clinical manifestation

The clinical manifestation of CDI is very heterogenous as it ranges from asymptomatic carrier state to mild diarrhea to life-threatening fulminant pseudomembranous colitis (PMC) (40–42). PMC is a complication that arises due to severe inflammation of the colon that leads to the

formation of lesions and yellow-white plaques over the colonic mucosa, which is filled with host immune cells, such as leukocytes, neutrophils, in addition to fibrin, mucous and inflammatory debris (12,43). In extreme conditions, CDI leads to toxic megacolon during which the colon is dilated and is unable to remove its content leading to perforations and rupture of the colon. This ultimately leads to sepsis, shock, and death of the infected individual. Other symptoms such as malaise, fever, nausea, anorexia presence of mucus or blood in the stool, cramping, abdominal discomfort, peripheral leukocytosis have also been reported in several cases of CDI (28). Since CDI is caused primarily by *C. difficile* toxins, extra-intestinal manifestations are rare. However, there are reports that have described arthritis or bacteremia like symptoms associated with CDI (44,45).

1.6 Economic burden

CDI causes a substantial economic burden, and most of it comes from prolonged hospitalization and re-hospitalization. The hospitalization cost of CDI in one state of the United States with 10,254 CDI cases was calculated to be $16,217,295 (2). Similarly, one of the studies estimated $10,528 per person per CDI infection (3). The annual healthcare cost of CDI is $4 billion (3). Because of increasing number of cases of CDI and the high cost associated with it, CDC has categorized CDI as an urgent threat to public health (46). Therefore, studying *C. difficile* pathogenesis is important from public health as well as economic prospect.

1.7 Prevention and treatment

1.7.1 Prevention

Following regular hand health hygiene practices such as washing hands with soap and water and the wearing of personal protective equipment by medical personnel can prevent transmission of *C. difficile* spores from the environment to the susceptible individual (47). Since some antibiotics are shown to impose high risk on CDI, prescribing low-risk antibiotics can help lower the chances of getting CDI (47). Cell-free supernatant of *Lactobacillus casei* sub sp. *rhamnosus* has been shown to have bacteriostatic but not bactericidal effects on the growth of several human pathogenic bacteria, including *C. difficile in vitro* (48). In a similar study, *Lactobacillus gasseri*, when co-cultured with *C. difficile*, has been shown to inhibit the adhesion of *C. difficile* in a mucus model (49). Although *in vitro* studies suggest the potential of using probiotics to keep a check on CDI, this has not been endorsed by any randomized studies (48,49). Chlorine-based cleaning reagents with 1,000-5,000 ppm of chlorine concentration can be used to clean abiotic surfaces (28).

1.7.2 Treatment

Vancomycin was shown to improve pseudomembranous colitis caused by CDI and has been used as the cornerstone of severe CDI infection treatment (50). Subsequent studies demonstrated metronidazole and fidaxomicin to be as effective as vancomycin for CDI (51–53). Because Vancomycin and metronidazole are less likely to disrupt the gut microflora and are thus categorized as low risk antibiotics, recurrence rates of infection are low (51).

Several studies have highlighted the possibility of immunotherapy against CDI. Immunotherapy is a form of treatment which uses substances that stimulates the immune system

of the person who receives the treatment. Immunotherapy is based upon the knowledge that CDI can be asymptomatic and asymptomatic *C. difficile* carriers have antibodies against *C. difficile* toxins in their serum (54). Bezlotoxumab was the first anti-toxin antibody that was approved by FDA in 2016 (55). Although the registration clinical trials had some promising results, there were some side effects associated with Bezlotoxumab treatment such as abdominal pain, diarrhea, nausea, vomiting, fatigue, pyrexia, urinary tract infection and headache (55). Since immunotherapy can protect from recurrent CDI and does not disrupt the commensal gut flora as does the antibiotic treatments, it can be a promising approach to treat CDI.

1.7.2.1 Fecal microbiota transplantation (FMT)

FMT aims to restore the gut flora using the fecal microbiota from a healthy individual. FMT has been shown to be effective in treating recurrent and refractory CDI (56). Although positive outcomes are seen with vancomycin or metronidazole, 15-35% of CDI patients develop recurrent disease either with the same strain or reinfection with a different strain (28). CDI is the consequence of the disruption of the commensal gut microbiome. One way in which *C. difficile* is suggested to persist in the gut environment is by manipulating the gut environment to inhibit the reconstitution of the commensal gut microbiota (57). *C. difficile* induces indole production when co-cultured with indole producing bacteria. It has been shown that when *E. coli*, an indole producer, and *C. difficile* are co-cultured, expression of *tnaA* was upregulated in *E. coli*. *tnaA* encodes an enzyme that converts tryptophan to indole. This suggests that *C. difficile* induces indole producing bacteria to produce more indole so that the growth of the indole sensitive commensal gut microbes are reduced and do not exert colonization resistance to *C. difficile* growth (57). Another study showed that clearance of *C. difficile* from murine gastrointestinal tract was

dependent on the gut microbiota and was independent of the adaptive immune system (32). This further highlights the imperative role of gut microbiota in clearing *C. difficile* infection. All these evidences suggest a beneficial role for the microbiota in *C. difficile* infection.

Although the exact mechanism by which FMT contributes to reducing CDI is still not clear, there are few suggested roles of FMT. As secondary bile acid and short-chain fatty acids (SCFAs) is shown to increase with the administration of FMT, this suggests a metabolic role of FMT in generating compounds that can inhibit *C. difficile* spore germination (58). Similarly, gut microbiome also scavenges the metabolite like primary bile salts that are needed for *C. difficile* spore germination (59). *Saccharomyces boulardii* and *Bacillus claussii* are shown to secrete proteases that inhibit TcdA activity (60,61) . The concerted action of gut microbiota is also thought to exert colonization resistance on *C. difficile*. So, FMT is speculated to exert healthy and spatial competition against *C. difficile* and promote colonization resistance (31).

1.8 Pathogenesis of *C. difficile*

1.8.1 Toxins

1.8.1.1 *Clostridioides difficile* toxins

1.8.1.1.1 Toxin A and B

C. difficile produces two exotoxins: TcdA and TcdB. The genes *tcdA* and *tcdB* that encode for these toxins are present in a pathogenicity locus (PaLoc), which is a 19.5 kb chromosomal region found only in the toxigenic strains of *C. difficile* (Fig 1.3) (62). Besides the toxin genes, PaLoc consists of *tcdR*, *tcdC* and *tcdE*. TcdR is the alternate sigma factor required for the transcription of *tcdA* and *tcdB* (63). TcdC is the anti-sigma factor and therefore, is the negative

regulator of toxin genes (64). TcdE is a holin like protein which has been shown to be required for the secretion of toxins A and B (65).

1.8.1.1.2 Binary toxin CDT

The binary toxin CDT is produced in addition to Toxin A and B by non-pathogenic *C. difficile* isolates (66). CDT encoding genes, *cdtA,* and *cdtB,* are present in the 6.2 kb CDT locus (CdtLoc) and encode two polypeptides, CdtA helps in attachment of the toxin to the host gut epithelial cells (binding component) and CdtB possess actin-specific ADP-ribosyltransferase activity (active component) that promotes actin depolymerization (Fig 1.3) (67). CDT is shown to depolymerize the actin cytoskeleton (66), and it has been shown to enhance the adherence of *C. difficile* cells to host gut epithelial cells *in vitro* (68). However, the pathological role of CDT is very underexplored, primarily because it is not produced by the virulent *C. difficile* strains.

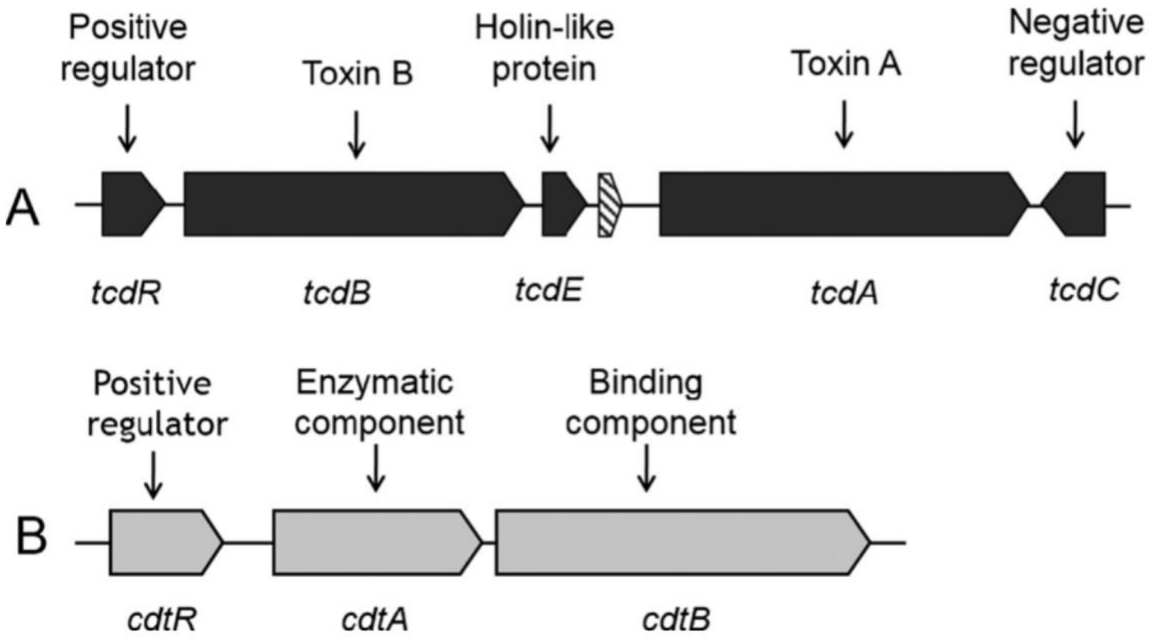


Figure 1.3. Toxin loci of *C. difficile*

Clostridioides difficile consists of two toxin loci. **A.** PaLoc, that encodes for Toxin A and B. **B.** CdtLoc, that encodes for the binary toxin CDT. Both of the loci also encode accessory products required for the expression of toxins. Image taken from (69).

1.8.1.2 Cytopathic effect of *C. difficile* toxin

Clostridioides difficile Toxin A and B are the major virulence factors that contribute to the colonic inflammation that is a characteristic feature of CDI. Both the toxin A and B are glucosyltransferases and glycosylate Rho family GTPases in the host cell (70), leading to disruption of the actin cytoskeleton, cell rounding, and eventually cell death and inflammatory responses. *C. difficile* toxins A and B consist of four major domains: N-terminal active Glucosyltransferase domain (A), middle translocation section consisting of Cysteine protease domain (C) and hydrophobic domain (D), and the C-terminal receptor-binding domain (B) (Fig1.4) (71). *Clostridioides difficile* toxins bind to the receptor in the gut epithelium using their C-terminal receptor binding domain and are endocytosed (Fig 1.5) (72). Several receptor for TcdB has been suggested. Tao et al. found Frizzled proteins FZD1, 2 and 7 as the receptors of Toxin B at the colonic epithelium (73). Similarly, LaRance et al. identified Poliovirus receptor-like 3 (PVRL3) which is found in the surface of the human colon as the receptor of TcdB (74). Frizzled proteins are the key players of Wnt signaling pathway which is important for repair and renewal of colonic epithelial cells, suggesting the specificity of TcdB toxin to the colonic cells (73). The fate of the toxin depends on the acidification of the compartment. At low pH, C. difficile toxins change conformation such that its hydrophobic regions are exposed to insert itself the membrane (75). This insertion leads to the pore formation in the endosomal membrane (75). Whether this pore formation allows the delivery of the toxin to the cell cytoplasm is not known. However, recent studies have shown that C. difficile toxin A undergoes autoproteolytic cleavage and toxin B undergoes host cell inositol hexakisphosphate (InsP6) mediated proteolytic cleavage, releasing the N-terminal active glucosyltransferase domain in the cytosol (71). The active domain in the cytoplasm glycosylates and inactivates the Rho GTPases such as Rho, Rac, and Cdc42, and results

in cell rounding, actin depolymerization, and loss of epithelial barrier function leading to diarrhea (76–78).

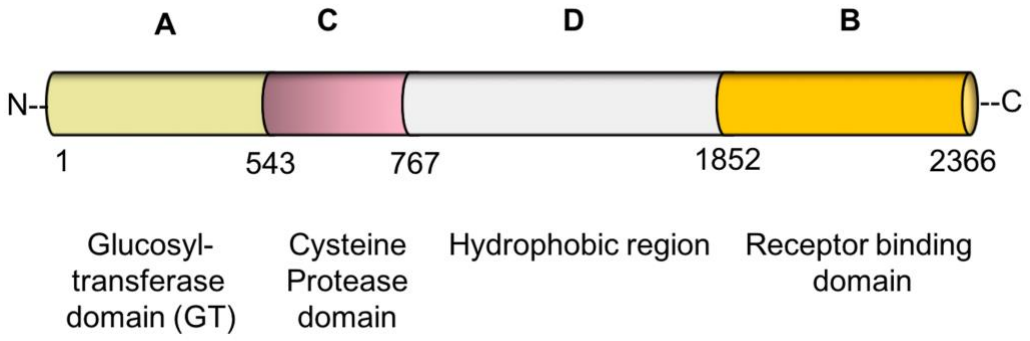


Figure 1.4. Domain structure of *C. difficile* toxins

Clostridioides difficile toxins A and B consist of four major domains: N-terminal active Glucosyltransferase domain (A), middle translocation section consisting of Cysteine protease domain (C) and hydrophobic domain (D), and the C-terminal receptor-binding domain (B).

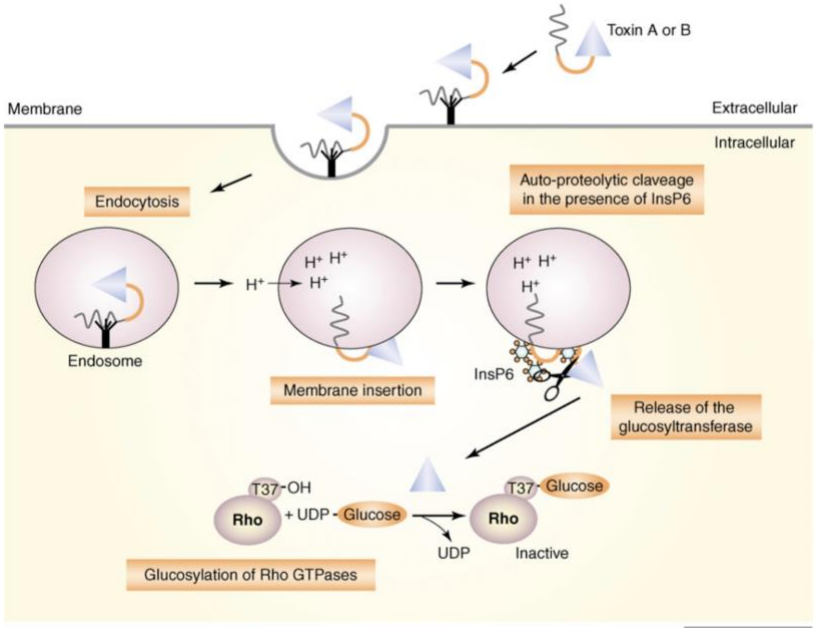


Figure 1.5. The mechanism of action of *C. difficile* toxin A or B on a host epithelial cell

Upon binding with the receptor on the surface of host gut epithelial cells, *C. difficile* toxin is endocytosed, proteolytically processed and released to the cytoplasm where it glycosylates and hence inactivates the Rho proteins. Image taken from (71).

The toxins also induce intestinal epithelial cells and the immune cells such as neutrophils to secrete pro-inflammatory cytokines such as Interleukine-1beta, TNF-alpha and Interleukine-8 (79). These cytokines encourage neutrophil influx to the site of infection, which further deteriorates the intestinal lining. In severe conditions, the neutrophils enter in the blood stream (leukocytosis), which causes poor prognosis (80) .

1.8.1.3 Regulation of toxins

TcdR : The role of TcdR in regulating the expression of *C. difficle* toxin A and B came into perspective when *C. diffiile tcdR* was expressed *in trans* in *E.coli* and it activated the reporter fusions of *C. difficile* toxin encoding genes *tcdA* and *tcdB* (81). TcdR demonstrated significant amino acid sequence similarity to TetR, which regulates the toxin genes in *Clostridium botulinum* and *Clostridium perfringens (82)*. TcdR contained a helix-turn-helix motif, which suggests its DNA binding ability. Mani and Dupuy demonstrated that TcdR is the alternate sigma factor that directly binds to the RNA polymerase and directs the holoenzyme to the promoter region of *tcdA* and *tcdB*, thus regulating the expression of toxin genes. (63). TcdR is the only known regulator that directly binds to the promoter region and regulates the expression of *C. difficile* toxin genes.

TcdC: Transcriptional analysis of PaLoc genes demonstrated that toxin genes are not expressed when *tcdC* is expressed and vice versa (83), suggesting that TcdC could be the negative regulator of toxin genes. The observation that the hypervirulent epidemic strains that produced a high level of toxin had deletions or frameshift mutations in the *tcdC* gene further supports this speculation (20,84). TcdC does not have a DNA binding motif. It has been shown that its regulation of toxin genes is based on its anti-sigma factor activity to modulate TcdR activity and its ability to interact directly with the RNA polymerase core enzyme (64). However, several studies have shown

no association between *tcdC* and increased toxin production, highlighting the need for additional studies to attribute the role of TcdC in *C. difficile* conclusively (85–87).

Carbon catabolite repression: Glucose negatively regulates *tcdR,* thereby repressing toxin production in *C. difficile* (82). It has been shown that glucose-dependent regulation of toxin production is carried out by the global regulator CcpA by binding to the *cre* (cis-acting catabolite responsive element) sites present in the upstream region of *tcdR* and other PaLoc genes (88,89).

Proline and glycine: Both of these amino acids have inhibitory effect on *C. difficile* toxin production through the regulator PrdR (90). However, the exact mechanism by which PrdR regulates toxin production is not understood. Alternatively, proline positively regulates toxin production. Proline dependent PrdR mediated toxin regulation is further mediated by another regulator, Rex, through an alternative pathway that leads to the accumulation of butyrate, a compound known to increase toxin production (91). Cysteine also downregulates toxin synthesis in *C. difficile* (92). However, further investigation revealed that toxin production in *C. difficile* might be regulated by by-products of cysteine degradation like pyruvate and sulfide and not by cysteine (93).

CodY: has been shown to bind directly to the upstream of *tcdR* and repress the expression of PaLoc genes (94). Consistent with this study, *codY* mutant produces higher toxins than wild-type (95). CodY is also known to repress genes in the butyrate synthesis pathway, which positively regulates toxin synthesis (95).

Quorum sensing: Autoinducer-2 (AI-2), the quorum-sensing molecule identified in different bacterial systems, when added to the growth medium, upregulates the expression of PaLoc genes in *C. difficile* (96). Consistent with this study, a decrease of AI-2 production is shown to downregulate the expression of toxin genes (97). However, the exact molecular mechanism by

which AI-2 and the associated quorum-sensing system regulates toxin production in *C. difficile* is poorly understood. This suggests that quorum sensing positively regulates toxin production and thus contributes to increased virulence traits.

c-di-GMP: Increase in intracellular c-di-GMP concentration by ectopic expression of the c-di-GMP synthesis gene, *dccA*, has been shown to decrease the expression of toxin genes (98,99). c-di-GMP has also been shown to negatively regulate the flagellar operon by binding to the upstream riboswitch (99). Flagellar operon consists of *sigD* which is known to positively regulate *tcdR* expression and toxin production positively (100). Therefore, c-di-GMP mediated regulation of toxin production is likely mediated by SigD. Additionally, flagellin FliC and cap protein FliD are shown to regulate toxin production (101,102). It has been suggested that the physical presence of flagellin proteins rather than the active presence regulates the expression of toxin genes (103).

1.8.2 Sporulation

Bacterial cells enter into sporulation (Fig 1.6) when they are not able to sustain vegetative growth (104). This failure to maintain vegetative growth arises due to various stress conditions like exhaustion of nutrients, exposure to harsh environmental conditions like oxygen, temperature, pH, etc. or unfavorable conditions like antibiotic treatment or attack from host immune responses (104–106).

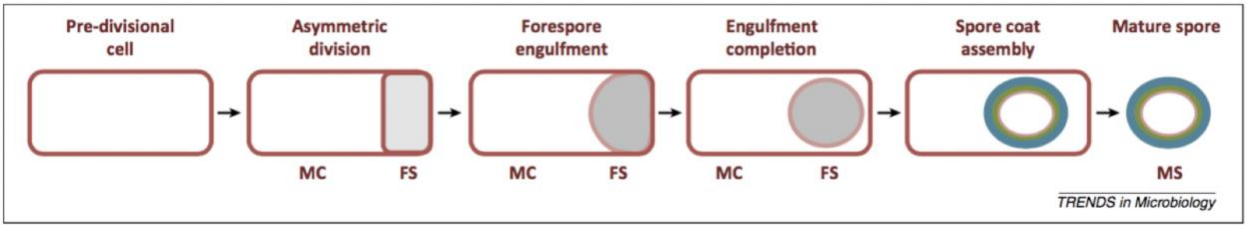


Figure 1.6. Schematic overview of general process of bacterial sporulation

In response to a sporulation trigger, the larger mother cell (MC) undergoes asymmetric division to generate a smaller forespore (FS) compartment. During the asymmetric division, DNA segregation occurs after the completion of which FS in engulfed by the MC. After the production and assembly of the cortex and inner and outer coats a mature spore is formed. Mature spore is then released into the environment by the lysis of MC. Upon encountering a favorable environment, the spore germinates to the vegetative cell. Image taken from (105).

1.8.2.1 *Clostridioides difficile* spore, a factor for persistence and transmission

Clostridioides difficile spores are the major factor for transmission and persistence (107). *C. difficile* spores are excreted in the feces of the infected individual. Because the spores are resistant to disinfectants and oxygen (106), they are viable in the environment outside the host. Spores act as a vehicle for transmitting the disease through fecal oral route in the health care setting. Spores are ingested, and because they are resistant to mammalian gastric juice, they make their way to the anaerobic mammalian intestine where they germinate with the help of bile salts (105) into active vegetative cells and secrete toxin A and B, that cause tissue damage and inflammation. These spores enable the bacteria to be persistent and thus are difficult to eradicate.

1.8.2.2 Sporulation initiation in *C. difficile*

The sporulation initiation pathway and the associated signal transduction system has been studied in *Bacillus subtilis*, a model Gram-positive organism. In *B. subtilis* sporulation initiation pathway consists of a multicomponent phosphorelay system composed of sensor histidine kinases

and response regulators (108–110). Sensor histidine kinases, in response to the environmental or cellular signals, undergo autophosphorylation in their specific histidine residue in their catalytic domain and transfer this phosphate group, in a phosphorelay manner, to the response regulators Spo0F-Spo0B-Spo0A (111–114). Spo0A is the master regulator of sporulation, and phosphorylated Spo0A is capable of binding to the specific target sequences called "0A boxes" (discussed below) in the promoter region of the target gene, thus regulating the expression of the gene (109). *B. subtilis* has multiple sensor kinases, KinA- KinE, allowing the bacteria to sense various environmental and cellular signals to activate the sporulation pathway (112). Phosphorylated Spo0A activates the expression of the downstream cascade of sporulation specific RNA polymerase sigma factors like sigE, sigF, sigG, sigK (108).

Several facts suggest that sporulation initiation pathway in C. difficile is very similar to that of B. subtilis. Spo0A of B. subtilis is highly conserved in C. difficile (115). "0A boxes" are identified in the promoter region of genes that are likely regulated by Spo0A (116). The residues of Spo0A that interact with the "0A boxes", and the sporulation specific sigma factors like SigE, sigF, sigG and sigK are conserved in C. difficile (117). However, the expression and activation of these sigma factors are very different compared to B. subtilis and C. difficile genome do not encode the orthologs of response regulators Spo0F and Spo0B which suggests that sporulation initiation pathway of *C. difficile* is divergent from B. subtilis sporulation initiation pathway (118). The common speculation in the field is sporulation initiation in C. difficile does not consist of phosphorelay system, instead, it is composed of two-component systems where the extracellular and cellular signals are sensed by sensor kinases and phosphorylate Spo0A directly (119).

1.8.2.3 Spo0A, the master regulator of sporulation

Spo0A, the master regulator of sporulation, is required for spore formation, persistent, recurrent infection, and for host-to-host transmission (118,119). Spo0A regulates the expression of early sporulation sigma factor genes and thus lies in the heart of the sporulation regulon of *C. difficile*. Spo0A is a transcriptional regulator that consists of the N-terminal phosphorylation and dimerization domain and the C-terminal DNA binding domain (109). Spo0A is shown to bind with high affinity to specific sites called "0A boxes". Spo0A binding "0A box" comprises of TGTCGAA where G at the second position and C and G at the fourth and fifth position, respectively, are the important nucleotides that make specific contacts with the target DNA (109,119). However, Spo0A regulates the expression of many target genes that contain one or more degenerate motifs (119,120).

1.8.2.4 Regulation of sporulation initiation in *C. difficile*

Despite the critical importance of *C. difficile* spores in causing CDI, very little is known about the regulatory network of sporulation initiation. The specific environmental and nutritional signals that trigger sporulation initiation in *C. difficile* are largely uncharacterized. The exact mechanism by which the regulators are responding to these signals to control sporulation is poorly understood.

1.8.2.4.1 Transcriptional regulation of Spo0A

Several regulators such as opp and app (121) , sigH (122), ccpA (88), rstA (123,124), sinRI (125) , tcdR (126) and codY (127) influence sporulation in C. difficile (Fig 1.7).

CodY governs the global response to GTP and branched-chain amino acids (BCAAs) availability (127). CodY is a transcriptional regulator and has been shown to negatively regulate sporulation (127,128). It has been shown that a codY mutant has increased expression of sporulation associated genes such as *sigF*, *sigE,* and *sigG* (127). However, the exact mechanism by which CodY regulates sporulation is poorly understood mainly because it shows strain specific differential regulation of other sporulation associated genes like *sinR* and *opp* and *app* (127).

CcpA governs the global response to carbon availability in *C. difficile* by regulating the expression of genes involved in sugar uptake, fermentation and amino acid metabolism. CcpA directly regulates the expression of *tcdR* and, therefore, is known to influence sporulation indirectly (89). CcpA represses the expression of spo0A, and spo0IIAA by directly binding to their promoter region. *ccpA* mutant is shown to have a hyper-sporulation phenotype (88). It has been shown that CcpA negatively regulates sporulation initiation by inhibiting the expression of *spo0A*, *opp,* histidine kinase CD1579, and *sinR (88).*

Opp and App are oligopeptide permeases transport peptides which can be utilized as a source of carbon and nitrogen to generate ATP (121). *opp* and *app* null mutants are shown to have earlier and upregulated expression of sporulation genes and hypersporulation phenotype (121). It is suggested that *opp* and *app* enhance intracellular nutritional availability and inhibit sporulation. Since CcpA and CodY are the global nutritional regulators, it is more likely that *opp* and *app* regulation of sporulation is indirect and is mediated through CcpA and CodY. However, it is shown that CcpA and CodY regulated genes are not differentially expressed in *opp* and *app* null mutants except for *sinRI* (121). More evidence is required to attribute CcpA and CodY mediated regulation of sporulation to *opp* and *app*.

TcdR is an alternate sigma factor that directs the transcription of toxins genes in *C. difficile*. Reduced sporulation and transcript levels of sporulation associated sigma factors in *tcdR* mutant suggest that tcdR positively regulates sporulation in *C. difficile* strain R20291 (126). However, tcdR mutant in the 630Δ*erm* strain demonstrated hypersporulation phenotype, suggesting strain-specific regulation of sporulation (126).

SinR has been shown to positively regulate the expression of *spo0A.* A *sinR* mutant is asporogenic and has demonstrated decreased transcription of *spo0A* (125). However, the exact regulatory relationship between SinR and Spo0A and how *sinR* regulates sporulation is not understood.

SigH is the transition phase sigma factor and is shown to positively regulate sporulation initiation. A sigH mutant was asporogenic and had reduced expression of sporulation associated genes such as *spo0A*, histidine kinase CD2492, *sinRI*, *spo0J* and *soj* and increased expression of the negative regulator of sporulation, *app* (122).

RstA is a transcriptional regulator and is shown to positively regulate sporulation in *C. difficile* R20291 and 630Δ*erm* strains. In both of these strains, a *rstA* mutant demonstrated significantly reduced sporulation (123,124). The exact mechanism by which RstA regulates sporulation is not known. However, based on the study that demonstrated unchanged *spo0A* transcript level and reduced transcription of sporulation associated genes like *sigE, sigF, sigG* and *spoIIE* in *rstA* mutant suggests a role for RstA in the Spo0A activation pathway (123). A *rstA* mutant had increased transcription of histidine kinases CD1492, CD1579, and CD2492, suggesting the possible role of RstA in regulating the activity of these histidine kinases (128).

1.8.2.4.2 Post-translational regulation of Spo0A

Post-translational regulation of Spo0A controls the activity of Spo0A. It controls the expression and activity of histidine kinases and phosphatases, that phosphorylate and dephosphorylate Spo0A, respectively, thus determining the availability of phosphorylated Spo0A (128).

Histidine kinases: The *C. difficile* genome encodes five orphan histidine kinases CD1492, CD1579, CD2492, CD1352, CD1949 (118). CD1949 has not been characterized yet. All the other histidine kinases, except for CD1579, are integral membrane proteins, which suggests their ability to sense and transduce extracellular signals (118). A CD2492 mutant has been shown to reduce sporulation by 3.5-fold compared to the *C. difficile* parent strain, and purified CD1597 is shown to directly phosphorylate Spo0A *in vitro* (118). However, CD1492 is shown to negatively regulate Spo0A dependent sporulation associated genes such as *sigF* and *sigE,* which suggest that CD1492 is a negative regulator of sporulation (129).

Phosphatases: *Clostridioides difficile* encodes of the orthologues of Rap phosphatases, which dephosphorylate Spo0F in *B. subtilis*, of KipI, an inhibitor of histidine kinase kinA, and of KipA, the inhibitor of KipI. However, there is no evidence of these orthologues influencing sporulation of *C. difficile* (128).

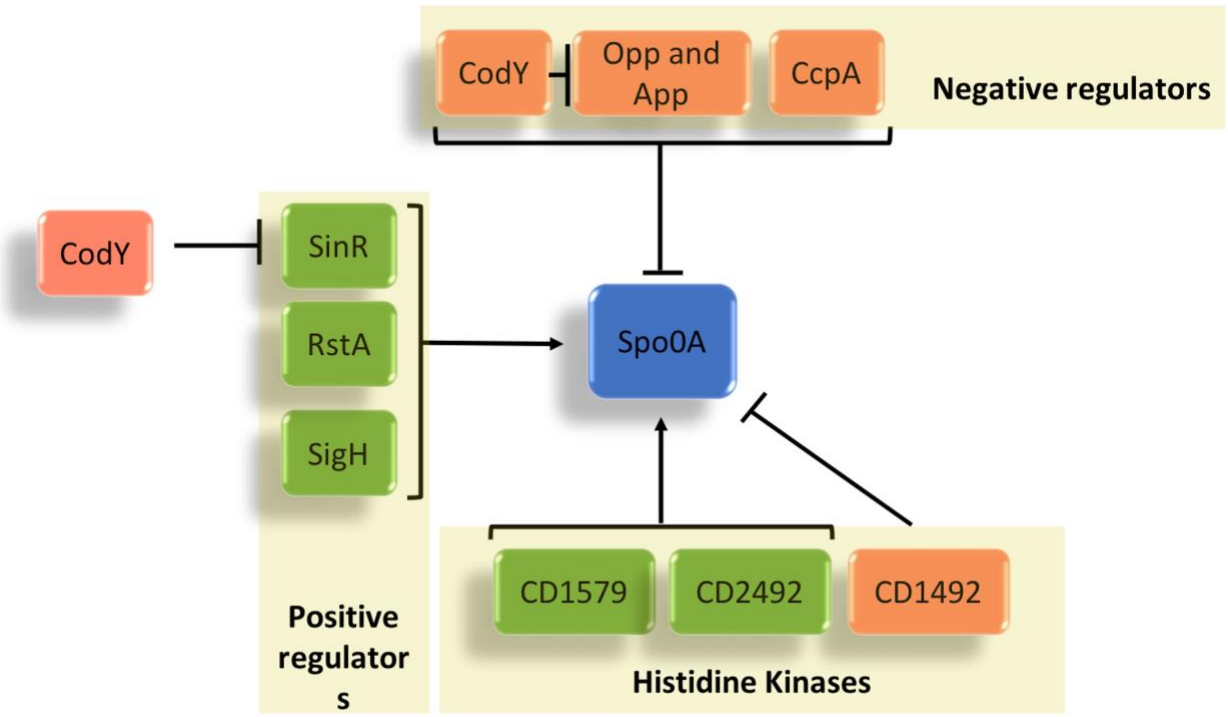


Figure 1.7. Gene regulatory network of Spo0A regulation in *C. difficile*

Spo0A is transcriptionally regulated by the global regulators SinR, RstA, SigH, CodY, Opp and App and CcpA and post translationally regulated by the histidine kinases CD1579, CD2492 and CD1492.

1.8.3 Adherence and colonization

Adhesion is the process by which a bacteria attaches itself to a surface. A gut pathogen adheres itself with the mucosal environment of the host gut. The ability to adhere to host tissues facilitates colonization and influence bacterial pathogenesis. Therefore, adhesion has been suggested to be a pre-requisite for colonization of the gut and is considered as the first step in the initiation of infection by a pathogen (130). Adhesion factors may include fimbrial or afimbrial microbial adhesins (131), surface protein, and secretory enzymes or toxins.

1.8.3.1 Microbial surface components

Several microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) have been characterized in Gram-positive bacteria, which are usually localized in the peptidoglycan cell wall and bind to extracellular matrices (ECM) like fibronectin, laminin, collagen and elastin (132). These adhesins not only help the bacteria in binding to host Extra Cellular Matrix (ECM) and establishing themselves in the host, but also have been shown to induce host immune responses (133). *Clostridioides difficile* has been shown to encode adhesins as well as proteolytic enzymes and their role in adherence and colonization are reported. *C. difficile* CbpA (Collagen Binding Protein A) is shown to have MSCRAMM property as it is found to be expressed in the bacterial surface and binds to collagen and human fibroblast and gut tissues *in vitro* (132). Similarly, Fbp68 (Fibrinogen Binding Protein 68) is also localized on the *C. difficile* surface and is shown to bind to fibrinogen and vitronectin (133,134). Surface layer proteins (SLPs) encoded by *slpA* is also shown to bind to collagen I, thrombospondin and vitronectin *in vitro* (135–137). Cell Wall binding Proteins (Cwps) like Cwp84 and Cwp66 (138), which show homology to SLPs are also shown to have a role in *C. difficile* adherence. Cwp84 has been characterized as a cysteine

protease and is likely to degrade some surface proteins required for adhesions or host-associated proteins (139–141). *Clostridioides difficile* lipoprotein CD0873 has been recently characterized as a surface lipoprotein. CD0873 mutant has been shown to have reduced adherence, and purified CD0873 has been shown to bind to the intestinal cell line *in vitro*, suggesting its role as an adhesin (142). Recently, *C. difficile cwp2* knock out mutant has been shown to have reduced adherence to Caco-2 cells *in vitro*, suggesting the potential role of Cwp2 as an adhesin (143). Competitive inhibition assay demonstrated that *C. difficile* GroEL serves as an adhesin (144). Therefore, *C. difficile* adherence is one of the factors that contributes to its pathogenesis.

1.8.3.2 Bacterial secreted proteins (Non-Toxins)

CD2830 is an extracellular metalloprotease that is secreted by *C. difficile* and has high specificity towards proline residues (145). CD2830 is shown to cleave putative adhesins CD2831 and CD3246, suggesting its role in the negatively regulating *C. difficile* adhesion (145). Interestingly, CD2831 and CD3246 are shown to have riboswitch type II, which turns the gene on at high concentration of c-di-GMP and CD2830 itself has riboswitch type I, which turns the gene off at high concentration of c-di-GMP (146). Recently, CD2831 has been characterized as collagen-binding protein, and overexpression of CD2831 is shown to have enhanced the biofilm forming ability of *C. difficile*, suggesting the role of CD2831 in adhesion (147).

1.8.3.3 Flagella

Tasteyre et al. demonstrated reduced tissue association *in vivo* of non-flagellated strains and *in vitro* adherence of FliC and FliD to mouse cecum mucus in *C. difficile* strains 79685, ATCC 43593, EX560, ATCC 43598, and 6058, suggesting the role of flagella in adhesion (148).

However, mutagenic analysis study on *C. difficile* 630 strain has shown that *fliC* and *fliD* mutants, which are non-flagellated, are better in adhering to the intestinal cell line Caco-2 *in vitro* and more virulent *in vivo* compared to the parent WT strain, suggesting that flagella do not contribute to *C. difficile* adherence (101). This study is consistent with an earlier study in *C. difficile* that has shown adherence is not mediated by flagellum-associated adhesins (149). Similar studies in *C. diffiicle* R20291 strain demonstrated that flagellar mutants adhere less to intestinal cell lines *in vitro* and flagellar structure protein is required for adhesion and colonization, suggesting strain-specific role of flagella in adherence and colonization in *C. difficile* (150).

1.8.3.4 Type IV Pili (TFP)

Clostridioides difficile TFP has been shown to promote adherence to intestinal cell lines *in vitro*, and TFP-null mutant strain has reduced association with cecal mucosa, suggesting the role of TFP in adherence and colonization of *C. difficile* (151). The same study also demonstrated that the production of TFP dependent on intracellular c-di-GMP concentration.

1.8.3.5 *Clostridioides difficile* toxin

Earlier studies have suggested the potential contribution of *C. difficile* toxin to adherence and colonization (149). *C. difficile* toxins A and B is shown to have 30% similarity to *efa1* of Enterohaemorrhagic *E. coli* (EHEC), which is shown to enhance *in vitro* adhesion to epithelial cells (152). Similarly, *Clostridium difficile* transferase (CDT), the binary toxin, is shown to induce microtubule-based cell protrusion and is suggested to have a role in bacterial adhesion and colonization (68). These evidences suggest the potential role of *C. difficile* toxin in adherence and colonization.

The three major pathogenic factors of *C. difficile* infection and the CDI cycle is illustrated below in Fig 1.8.

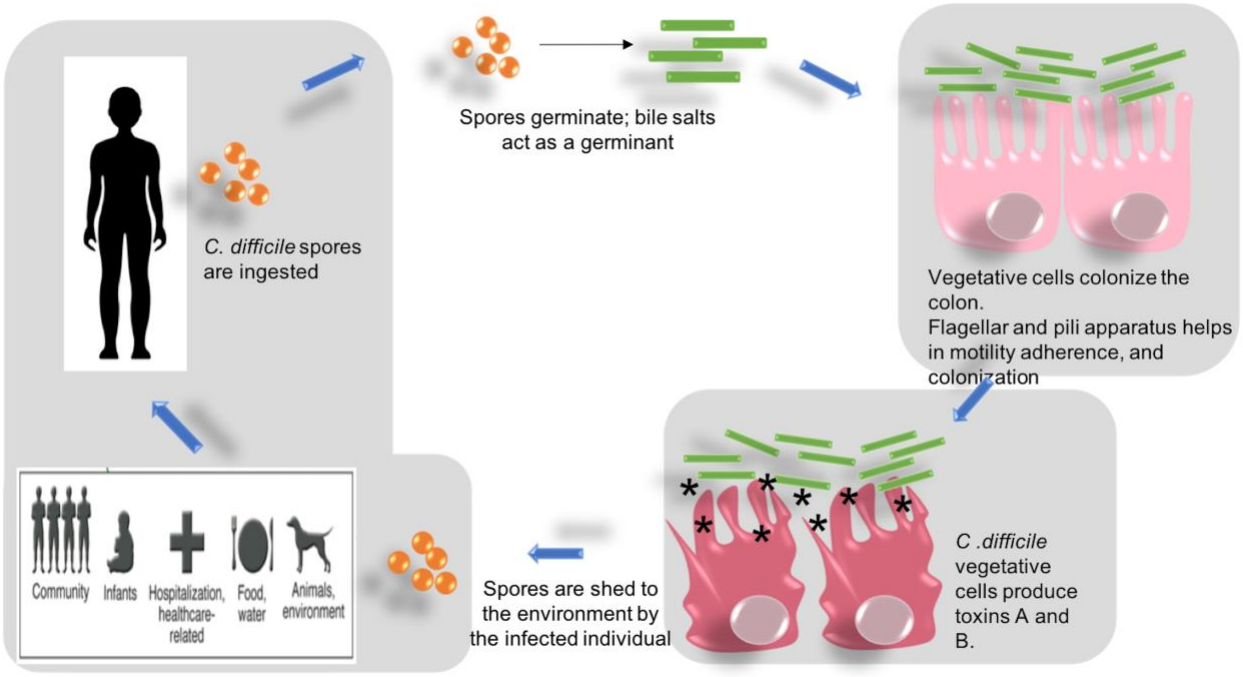


Figure 1.8. A schematic overview of *C. difficile* infection cycle

Sporulation, Adherence and colonization, and toxin production are the three major pathogenesis factors of *C. difficile* infection and are highlighted in the image. *Clostridioides difficile* infection cycle begins with the ingestion of spores by the susceptible individual. Spores germinate to active vegetative cells which then establish themselves in the host with the help of adherence and colonization factors and produce toxins. Image taken and modified from (153).

1.9 *Clostridioides difficile* genome and genotypes used in this study

Clostridioides difficile 630 genome was the first *C. difficile* genome to be sequenced and published in 2006 (154). *Clostridioides difficile* 630 belongs to PCR ribotype 012 and is a clinical isolate from a patient in an outbreak at a hospital in Zurich, Switzerland (155). According to the sequence results, *C. difficile* 630 strain has a chromosomal genome size of 4,290,252 bp with a G+C content of 29.06%. *Clostridioides difficile* 630 genome also consists of a plasmid DNA, pCD630, which is 7,881 bp in size. *Clostridioides difficile* 630 chromosomal genome has 3,776 coding sequences (CDSs), among which 80% are encoded on a leading strand. *Clostridioides difficile* 630 shows a 15% similarity in its CDSs with other sequenced *Clostridium* species, while 50% of the CDSs are unique to *C. difficile.* 11% of the *C. difficile* genome consists of mobile genetic elements, and the remaining genome consists of genes for virulence factors and surface proteins, antimicrobial resistance, sporulation and germination, metabolism and adaptation in host gut and regulatory genes. *Clostridioides difficile* plasmid DNA, pCD630, has 11 CDSs whose functions are not characterized yet. After the complete sequencing, *C. difficile* 630 was adapted as a reference strain by laboratories studying *C. difficile.*

1.9.1 JIR8094 and 630Δ*erm*

Clostridioides difficile 630 strain consists of *erm1*(B) and *erm2*(B) in its mobilizable transposon Tn5398, which provides erythromycin resistance to the strain. Mullany laboratory at University College, London generated *C. difficile* 630Δ*erm* strain, which is sensitive to erythromycin, by spontaneous mutation caused by serial passaging of 630 strain for 30 days in a non-selective media (156). Rood laboratory at Monash University, Australia independently isolated another erythromycin sensitive strain 630E, which is commonly known as JIR8094 (157).

Even though, both 630Δ*erm* and JIR8094 are susceptible to erythromycin; numerous other genetic differences exist between them. These strains are widely used in labs around the world to study molecular mechanisms of *C. difficile* pathogenesis, as they are more amenable to genetic manipulations.

1.9.2 R20291 and UK1

R20291 and UK1 fall under the ribotype (RT) 027 and are the epidemic strains that caused the CDI outbreak at Stoke Mandeville in the United Kingdom in 2003-2006 (158). These strains are considered hypervirulent strains because of their high cytotoxicity and hyper sporulation (158). The global rate of CDI has been shown to increase after the emergence and dissemination of RT027 strains (23).

1.10 Aims of the study

Clostridioides difficile spore is the major cause of the transmission and persistence of the disease. Therefore, it is essential to understand the exact mechanism by which sporulation initiation occurs and the gene regulatory networks that regulate spore formation. There are only a handful of genes that are shown to influence sporulation, but the exact mechanism by which these genes regulate sporulation is poorly understood. So, the aim of this study is to:

1. Identify new genes that have a role in sporulation of *C. difficile* and investigate the mechanism and the associated gene regulatory network.

2. Determine the regulatory relationship between Spo0A and *sinR*, the previously characterized regulator that is already shown to influence sporulation.

1.11 References

1. Guh AY, Mu Y, Winston LG, Johnston H, Olson D, Farley MM, et al. Trends in U.S. Burden of *Clostridioides difficile* Infection and Outcomes. N Engl J Med. 2020 Apr 2;382(14):1320–30.

2. Younas M, Royer J, Weissman SB, Waites KS, Dash S, Rac H, et al. Burden of community-associated *Clostridioides difficile* infection in southeastern United States: a population-based study. Infection. 2020 Feb;48(1):129–32.

3. Mollard S, Lurienne L, Heimann SM, Bandinelli P-A. Burden of *Clostridium* (*Clostridioides*) *difficile* infection during inpatient stays in the USA between 2012 and 2016. Journal of Hospital Infection. 2019 Jun 1;102(2):135–40.

4. Hall IC, O'toole E. Intestinal flora in new-born infants: with a description of a new pathogenic anaerobe, *Bacillus difficilis*. Am J Dis Child. 1935 Feb 1;49(2):390–402.

5. Bartlett JG. *Clostridium difficile* infection: Historic review. Anaerobe. 2009 Dec 1;15(6):227–9.

6. Rodriguez C, Van Broeck J, Taminiau B, Delmée M, Daube G. *Clostridium difficile* infection: Early history, diagnosis and molecular strain typing methods. Microbial Pathogenesis. 2016 Aug 1;97:59–78.

7. Lawson PA, Citron DM, Tyrrell KL, Finegold SM. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. Anaerobe. 2016 Aug 1;40:95–9.

8. Hafiz S, Morton RS, Mcentegart MG, Waitkins SA. *Clostridium difficile* in the urogenital tract of males and females. The Lancet. 1975 Feb 22;305(7904):420–1.

9. HAFIZ S, OAKLEY CL. *Clostridium difficile*: isolation and characteristics (Plate VIII). Journal of Medical Microbiology,. 1976;9(2):129–36.

10. Gorbach SL, Thadepalli H. Isolation of Clostridium in Human Infections: Evaluation of 114 Cases. J Infect Dis. 1975 May 1;131(Supplement):S81–5.

11. Snyder ML. Further Studies on *Bacillus Difficilis* (Hall and O'Toole). Journal of Infectious Diseases. 1937 Mar 1;60(2):223–31.

12. George RH, Symonds JM, Dimock F, Brown JD, Arabi Y, Shinagawa N, et al. Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. Br Med J. 1978 Mar 18;1(6114):695–695.

13. Lance George W, Goldstein EllieJC, Sutter VeraL, Ludwig ShellyL, Finegold SydneyM. Ætiology of antimicrobial-agent-associated colitis. The Lancet. 1978 Apr 15;311(8068):802–3.

14. Larson HE, Price AB, Honour P, Borriello SP. *Clostridium difficile* and the ætiology of pseudomembranous colitis. The Lancet. 1978 May 20;311(8073):1063–6.

15. Green RH. The Association of Viral Activation with Penicillin Toxicity in Guinea Pigs and Hamsters. Yale J Biol Med. 1974 Sep;47(3):166–81.

16. Kim YS, Han DS, Kim YH, Kim WH, Kim JS, Kim HS, et al. Incidence and clinical features of *Clostridium difficile* infection in Korea: a nationwide study. Epidemiol Infect. 2013 Jan;141(1):189–94.

17. Slimings C, Armstrong P, Beckingham WD, Bull AL, Hall L, Kennedy KJ, et al. Increasing incidence of *Clostridium difficile* infection, Australia, 2011–2012. Medical Journal of Australia. 2014 Mar;200(5):272–6.

18. Ho J, Dai RZW, Kwong TNY, Wang X, Zhang L, Ip M, et al. Disease Burden of *Clostridium difficile* Infections in Adults, Hong Kong, China, 2006–2014. Emerg Infect Dis. 2017 Oct;23(10):1671–9.

19. Leffler DA, Lamont JT. *Clostridium difficile* Infection. Longo DL, editor. N Engl J Med. 2015 Apr 16;372(16):1539–48.

20. Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, et al. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. The Lancet. 2005 Sep 24;366(9491):1079–84.

21. Roo ACD, Regenbogen SE. *Clostridium difficile* Infection: An Epidemiology Update. Clinics in Colon and Rectal Surgery. 2020 Mar;33(2):49–57.

22. Rupnik M, Wilcox MH, Gerding DN. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol. 2009 Jul;7(7):526–36.

23. Freeman J, Bauer MP, Baines SD, Corver J, Fawley WN, Goorhuis B, et al. The Changing Epidemiology of *Clostridium difficile* Infections. Clinical Microbiology Reviews. 2010 Jul 1;23(3):529–49.

24. Lees EA, Carrol ED, Ellaby NAF, Roberts P, Corless CE, Lenzi L, et al. Characterization of Circulating *Clostridium difficile* Strains, Host Response and Intestinal Microbiome in Hospitalized Children With Diarrhea. Pediatr Infect Dis J. 2020;39(3):221–8.

25. Spruit JL, Knight T, Sweeney C, Salimnia H, Savaşan S. *Clostridium difficile* infection in a children's hospital with specific patterns among pediatric oncology and hematopoietic stem cell transplantation populations. Pediatr Hematol Oncol. 2020 Apr;37(3):211–22.

26. Meda M, Virgincar N, Gentry V, Walker A, Macdonald N, Hooper M, et al. *Clostridium difficile* infection in pregnant and postpartum women in 2 hospitals and a review of literature. Am J Infect Control. 2019;47(1):e7–14.

27. de Curraize C, Rousseau C, Corvec S, El-Helali N, Fihman V, Barbut F, et al. Variable spectrum of disease and risk factors of peripartum *Clostridium difficile* infection: report of 14 cases from French hospitals and literature review. Eur J Clin Microbiol Infect Dis. 2018 Dec;37(12):2293–9.

28. Czepiel J, Dróżdż M, Pituch H, Kuijper EJ, Perucki W, Mielimonka A, et al. *Clostridium difficile* infection: review. Eur J Clin Microbiol Infect Dis. 2019;38(7):1211–21.

29. Seekatz AM, Young VB. *Clostridium difficile* and the microbiota. J Clin Invest. 2014 Oct 1;124(10):4182–9.

30. Denève C, Deloménie C, Barc M-C, Collignon A, Janoir C. Antibiotics involved in *Clostridium difficile*-associated disease increase colonization factor gene expression. Journal of Medical Microbiology. 2008 Jun 1;57(6):732–8.

31. Samarkos M, Mastrogianni E, Kampouropoulou O. The role of gut microbiota in *Clostridium difficile* infection. European Journal of Internal Medicine. 2018 Apr 1;50:28–32.

32. Leslie JL, Vendrov KC, Jenior ML, Young VB. The Gut Microbiota Is Associated with Clearance of *Clostridium difficile* Infection Independent of Adaptive Immunity. 2019;4(1):12.

33. Clooney AG, Bernstein CN, Leslie WD, Vagianos K, Sargent M, Laserna-Mendieta EJ, et al. A comparison of the gut microbiome between long-term users and non-users of proton pump inhibitors. Alimentary Pharmacology & Therapeutics. 2016;43(9):974–84.

34. Donskey CJ. *Clostridium difficile* in Older Adults. Infectious Disease Clinics of North America. 2017 Dec;31(4):743–56.

35. El Chakhtoura NG, Bonomo RA, Jump RLP. Influence of Aging and Environment on Presentation of Infection in Older Adults. Infectious Disease Clinics of North America. 2017 Dec;31(4):593–608.

36. Gavazzi G, Krause K-H. Ageing and infection. The Lancet Infectious Diseases. 2002 Nov;2(11):659–66.

37. Peniche AG, Spinler JK, Boonma P, Savidge TC, Dann SM. Aging impairs protective host defenses against *Clostridioides (Clostridium) difficile* infection in mice by suppressing neutrophil and IL-22 mediated immunity. Anaerobe. 2018 Dec;54:83–91.

38. Turner NA, Smith BA, Lewis SS. Novel and emerging sources of *Clostridioides difficile* infection. PLOS Pathogens. 2019 Dec 19;15(12):e1008125.

39. Warriner K, Xu C, Habash M, Sultan S, Weese SJ. Dissemination of *Clostridium difficile* in food and the environment: Significant sources of *C. difficile* community-acquired infection? J Appl Microbiol. 2017 Mar;122(3):542–53.

40. Taege AJ, Adal KA. *Clostridium difficile* diarrhea and colitis: a clinical overview. Cleve Clin J Med. 1999 Sep;66(8):503–7.

41. Cleary RK. *Clostridium difficile*-associated diarrhea and colitis: clinical manifestations, diagnosis, and treatment. Dis Colon Rectum. 1998 Nov;41(11):1435–49.

42. Farrell RJ, LaMont JT. Pathogenesis and clinical manifestations of *Clostridium difficile* diarrhea and colitis. Curr Top Microbiol Immunol. 2000;250:109–25.

43. Hurley BW, Nguyen CC. The spectrum of pseudomembranous enterocolitis and antibiotic-associated diarrhea. Arch Intern Med. 2002 Oct 28;162(19):2177–84.

44. Legendre P, Lalande V, Eckert C, Barbut F, Fardet L, Meynard J-L, et al. *Clostridium difficile* associated reactive arthritis: Case report and literature review. Anaerobe. 2016 Apr;38:76–80.

45. Libby DB, Bearman G. Bacteremia due to *Clostridium difficile*--review of the literature. Int J Infect Dis. 2009 Sep;13(5):e305-309.

46. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of *Clostridium difficile* Infection in the United States. New England Journal of Medicine. 2015 Feb 26;372(9):825–34.

47. Song JH, Kim YS. Recurrent *Clostridium difficile* Infection: Risk Factors, Treatment, and Prevention. Gut Liver. 2019 15;13(1):16–24.

48. Forestier C, De Champs C, Vatoux C, Joly B. Probiotic activities of *Lactobacillus casei rhamnosus*: in vitro adherence to intestinal cells and antimicrobial properties. Research in Microbiology. 2001 Feb 1;152(2):167–73.

49. Ferreira CL, Grześkowiak Ł, Collado MC, Salminen S. In Vitro Evaluation of *Lactobacillus gasseri* Strains of Infant Origin on Adhesion and Aggregation of Specific Pathogens. Journal of Food Protection. 2011 Sep 1;74(9):1482–7.

50. Debast SB, Bauer MP, Kuijper EJ, European Society of Clinical Microbiology and Infectious Diseases. European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for *Clostridium difficile* infection. Clin Microbiol Infect. 2014 Mar;20 Suppl 2:1–26.

51. Johnson S, Louie TJ, Gerding DN, Cornely OA, Chasan-Taber S, Fitts D, et al. Vancomycin, metronidazole, or tolevamer for *Clostridium difficile* infection: results from two multinational, randomized, controlled trials. Clin Infect Dis. 2014 Aug 1;59(3):345–54.

52. Louie TJ, Miller MA, Mullane KM, Weiss K, Lentnek A, Golan Y, et al. Fidaxomicin versus Vancomycin for *Clostridium difficile* Infection. N Engl J Med. 2011 Feb 3;364(5):422–31.

53. Cornely OA, Crook DW, Esposito R, Poirier A, Somero MS, Weiss K, et al. Fidaxomicin versus vancomycin for infection with *Clostridium difficile* in Europe, Canada, and the USA:

a double-blind, non-inferiority, randomised controlled trial. Lancet Infect Dis. 2012 Apr;12(4):281–9.

54. Mulligan ME, Miller SD, McFarland LV, Fung HC, Kwok RY. Elevated levels of serum immunoglobulins in asymptomatic carriers of *Clostridium difficile*. Clin Infect Dis. 1993 Jun;16 Suppl 4:S239-244.

55. Wilcox MH, Gerding DN, Poxton IR, Kelly C, Nathan R, Birch T, et al. Bezlotoxumab for Prevention of Recurrent *Clostridium difficile* Infection. N Engl J Med. 2017 Jan 26;376(4):305–17.

56. Ianiro G, Maida M, Burisch J, Simonelli C, Hold G, Ventimiglia M, et al. Efficacy of different faecal microbiota transplantation protocols for *Clostridium difficile* infection: A systematic review and meta-analysis. United European Gastroenterology Journal. 2018 Oct 1;6(8):1232–44.

57. Darkoh C, Plants-Paris K, Bishoff D, DuPont HL. *Clostridium difficile* Modulates the Gut Microbiota by Inducing the Production of Indole, an Interkingdom Signaling and Antimicrobial Molecule. Turnbaugh PJ, editor. mSystems. 2019 Mar 19;4(2):e00346-18, /msystems/4/2/msys.00346-18.atom.

58. Weingarden AR, Vaughn BP. Intestinal microbiota, fecal microbiota transplantation, and inflammatory bowel disease. Gut Microbes. 2017 May 4;8(3):238–52.

59. Theriot CM, Young VB. Interactions Between the Gastrointestinal Microbiome and *Clostridium difficile*. Annual Review of Microbiology. 2015;69(1):445–61.

60. Koon HW, Su B, Xu C, Mussatto CC, Tran DH-N, Lee EC, et al. Probiotic *Saccharomyces boulardii* CNCM I-745 prevents outbreak-associated *Clostridium difficile* -associated cecal inflammation in hamsters. American Journal of Physiology-Gastrointestinal and Liver Physiology. 2016 Oct 1;311(4):G610–23.

61. Ripert G, Racedo SM, Elie A-M, Jacquot C, Bressollier P, Urdaci MC. Secreted Compounds of the Probiotic *Bacillus clausii* Strain O/C Inhibit the Cytotoxic Effects Induced by *Clostridium difficile* and *Bacillus cereus* Toxins. Antimicrob Agents Chemother. 2016;60(6):3445–54.

62. Braun V, Hundsberger T, Leukel P, Sauerborn M, von Eichel-Streiber C. Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. Gene. 1996 Nov 28;181(1):29–38.

63. Mani N, Dupuy B. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci U S A. 2001 May 8;98(10):5844–9.

64. Matamouros S, England P, Dupuy B. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Molecular Microbiology. 2007;64(5):1274–88.

65. Govind R, Dupuy B. Secretion of *Clostridium difficile* Toxins A and B Requires the Holin-like Protein TcdE. Cheung A, editor. PLoS Pathog. 2012 Jun 7;8(6):e1002727.

66. Gerding DN, Johnson S, Rupnik M, Aktories K. *Clostridium difficile* binary toxin CDT. Gut Microbes. 2014 Jan 1;5(1):15–27.

67. Martin-Verstraete I, Peltier J, Dupuy B. The Regulatory Networks That Control *Clostridium difficile* Toxin Synthesis. Toxins (Basel) [Internet]. 2016 May 14 [cited 2020 Apr 30];8(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4885068/

68. Schwan C, Stecher B, Tzivelekidis T, van Ham M, Rohde M, Hardt W-D, et al. *Clostridium difficile* Toxin CDT Induces Formation of Microtubule-Based Protrusions and Increases Adherence of Bacteria. Blanke SR, editor. PLoS Pathog. 2009 Oct 16;5(10):e1000626.

69. Elliott B, Androga GO, Knight DR, Riley TV. *Clostridium difficile* infection: Evolution, phylogeny and molecular epidemiology. Infection, Genetics and Evolution. 2017 Apr 1;49:1–11.

70. Just I, Selzer J, Wilm M, Eichel-Streiber C von, Mann M, Aktories K. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature. 1995 Jun;375(6531):500–3.

71. Jank T, Aktories K. Structure and mode of action of *clostridial* glucosylating toxins: the ABCD model. Trends in Microbiology. 2008 May;16(5):222–9.

72. Florin I, Thelestam M. Internalization of *Clostridium difficile* cytotoxin into cultured human lung fibroblasts. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 1983 Dec 19;763(4):383–92.

73. Tao L, Zhang J, Meraner P, Tovaglieri A, Wu X, Gerhard R, et al. Frizzled proteins are colonic epithelial receptors for C. difficile toxin B. Nature. 2016 Oct;538(7625):350–5.

74. LaFrance ME, Farrow MA, Chandrasekaran R, Sheng J, Rubin DH, Lacy DB. Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity. Proc Natl Acad Sci USA. 2015 Jun 2;112(22):7073–8.

75. Qa'Dan M, Spyres LM, Ballard JD. pH-Induced Conformational Changes in *Clostridium difficile* Toxin B. Infect Immun. 2000 May;68(5):2470–4.

76. Egerer M, Giesemann T, Jank T, Satchell KJF, Aktories K. Auto-catalytic Cleavage of *Clostridium difficile* Toxins A and B Depends on Cysteine Protease Activity. J Biol Chem. 2007 Aug 31;282(35):25314–21.

77. Reineke J, Tenzer S, Rupnik M, Koschinski A, Hasselmayer O, Schrattenholz A, et al. Autocatalytic cleavage of *Clostridium difficile* toxin B. Nature. 2007 Mar;446(7134):415–9.

78. Rupnik M, Pabst S, Rupnik M, von Eichel-Streiber C, Urlaub H, Söling H-D. Characterization of the cleavage site and function of resulting cleavage fragments after

limited proteolysis of *Clostridium difficile* toxin B (TcdB) by host cells. Microbiology. 2005 Jan 1;151(1):199–208.

79. Shen A. *Clostridium difficile* Toxins: Mediators of Inflammation. J Innate Immun. 2012 Feb;4(2):149–58.

80. Savidge TC, Pan W, Newman P, O'Brien M, Anton PM, Pothoulakis C. *Clostridium difficile* toxin B is an inflammatory enterotoxin in human intestine. Gastroenterology. 2003 Aug 1;125(2):413–20.

81. Carter GP, Cheung JK, Larcombe S, Lyras D. Regulation of toxin production in the pathogenic *clostridia*. Molecular Microbiology. 2014;91(2):221–31.

82. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, et al. Environmental Response and Autoregulation of *Clostridium difficile* TxeR, a Sigma Factor for Toxin Gene Expression. J Bacteriol. 2002 Nov;184(21):5971–8.

83. Hundsberger T, Braun V, Weidmann M, Leukel P, Sauerborn M, Eichel-Streiber CV. Transcription Analysis of the Genes tcdA-E of the Pathogenicity Locus of *Clostridium Difficile*. European Journal of Biochemistry. 1997;244(3):735–42.

84. MacCannell DR, Louie TJ, Gregson DB, Laverdiere M, Labbe A-C, Laing F, et al. Molecular Analysis of *Clostridium difficile* PCR Ribotype 027 Isolates from Eastern and Western Canada. J Clin Microbiol. 2006 Jun;44(6):2147–52.

85. Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP. Precise Manipulation of the *Clostridium difficile* Chromosome Reveals a Lack of Association between the *tcdC* Genotype and Toxin Production. Appl Environ Microbiol. 2012 Jul 1;78(13):4683–90.

86. Bakker D, Smits WK, Kuijper EJ, Corver J. TcdC Does Not Significantly Repress Toxin Expression in *Clostridium difficile* 630ΔErm. Popoff MR, editor. PLoS ONE. 2012 Aug 17;7(8):e43247.

87. Carter GP, Douce GR, Govind R, Howarth PM, Mackin KE, Spencer J, et al. The Anti-Sigma Factor TcdC Modulates Hypervirulence in an Epidemic BI/NAP1/027 Clinical Isolate of *Clostridium difficile*. Koehler TM, editor. PLoS Pathog. 2011 Oct 13;7(10):e1002317.

88. Antunes A, Camiade E, Monot M, Courtois E, Barbut F, Sernova NV, et al. Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. Nucleic Acids Research. 2012 Nov;40(21):10701–18.

89. Antunes A, Martin-Verstraete I, Dupuy B. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Molecular Microbiology. 2011;79(4):882–99.

90. Karlsson S, Burman LG, Åkerlund T. Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. Microbiology. 1999 Jul 1;145(7):1683–93.

91. Bouillaut L, Dubois T, Francis MB, Daou N, Monot M, Sorg JA, et al. Role of the global regulator Rex in control of NAD+-regeneration in *Clostridioides (Clostridium) difficile*. Molecular Microbiology. 2019;111(6):1671–88.

92. Karlsson S, Lindberg A, Norin E, Burman LG, Åkerlund T. Toxins, Butyric Acid, and Other Short-Chain Fatty Acids Are Coordinately Expressed and Down-Regulated by Cysteine in *Clostridium difficile*. Barbieri JT, editor. Infect Immun. 2000 Oct 1;68(10):5881–8.

93. Dubois T, Dancer-Thibonnier M, Monot M, Hamiot A, Bouillaut L, Soutourina O, et al. Control of *Clostridium difficile* Physiopathology in Response to Cysteine Availability. Young VB, editor. Infect Immun. 2016 Aug;84(8):2389–405.

94. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL. Repression of *Clostridium difficile* toxin gene expression by CodY. Molecular Microbiology. 2007;66(1):206–19.

95. Dineen SS, McBride SM, Sonenshein AL. Integration of Metabolism and Virulence by *Clostridium difficile* CodY. JB. 2010 Oct 15;192(20):5350–62.

96. Lee ASY, Song KP. LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. Biochemical and Biophysical Research Communications. 2005 Sep 30;335(3):659–66.

97. Yun B, Oh S, Song M, Hong Y-S, Park S, Park D-J, et al. Inhibitory Effect of Epigallocatechin Gallate on the Virulence of *Clostridium difficile* PCR Ribotype 027. Journal of Food Science. 2015;80(12):M2925–31.

98. Purcell EB, McKee RW, Courson DS, Garrett EM, McBride SM, Cheney RE, et al. A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls *Clostridium difficile* Biofilm and Toxin Production during Stationary Phase. Infect Immun. 2017;85(9).

99. McKee RW, Harvest CK, Tamayo R. Cyclic Diguanylate Regulates Virulence Factor Genes via Multiple Riboswitches in *Clostridium difficile*. mSphere. 2018 24;3(5).

100. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, et al. Characterization of the SigD Regulon of *C. difficile* and Its Positive Control of Toxin Production through the Regulation of tcdR. Nübel U, editor. PLoS ONE. 2013 Dec 16;8(12):e83748.

101. Dingle TC, Mulvey GL, Armstrong GD. Mutagenic Analysis of the *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence in Hamsters. McCormick BA, editor. Infect Immun. 2011 Oct;79(10):4061–7.

102. Aubry A, Hussack G, Chen W, KuoLee R, Twine SM, Fulton KM, et al. Modulation of Toxin Production by the Flagellar Regulon in *Clostridium difficile*. Infect Immun. 2012 Oct;80(10):3521–32.

103. Twine SM, Reid CW, Aubry A, McMullin DR, Fulton KM, Austin J, et al. Motility and Flagellar Glycosylation in *Clostridium difficile*. JB. 2009 Nov 15;191(22):7050–62.

104. Tocheva EI, Ortega DR, Jensen GJ. Sporulation, bacterial cell envelopes and the origin of life. Nat Rev Microbiol. 2016 Aug;14(8):535–42.

105. Paredes-Sabja D, Shen A, Sorg JA. *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends in Microbiology. 2014 Jul 1;22(7):406–16.

106. Dürre P. Physiology and Sporulation in *Clostridium.* Microbiol Spectr. 2014 Aug;2(4):TBS-0010-2012.

107. Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, et al. The *Clostridium difficile spo0A* Gene Is a Persistence and Transmission Factor. Bäumler AJ, editor. Infect Immun. 2012 Aug;80(8):2704–11.

108. Trach K, Burbulys D, Strauch M, Wu J-J, Dhillon N, Jonas R, et al. Control of the initiation of sporulation in *Bacillus subtilis* by a phosphorelay. Research in Microbiology. 1991 Jan 1;142(7):815–23.

109. Fujita M, González-Pastor JE, Losick R. High- and Low-Threshold Genes in the Spo0A Regulon of *Bacillus subtilis*. JB. 2005 Feb 15;187(4):1357–68.

110. Hoch JA. Regulation of the Phosphorelay and the Initiation of Sporulation in *Bacillus Subtilis*. Annual Review of Microbiology. 1993;47(1):441–65.

111. LeDeaux JR, Yu N, Grossman AD. Different roles for KinA, KinB, and KinC in the initiation of sporulation in *Bacillus subtilis*. J Bacteriol. 1995 Feb;177(3):861–3.

112. Jiang M, Shao W, Perego M, Hoch JA. Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. Molecular Microbiology. 2000;38(3):535–42.

113. Liu W, He Z, Gao F, Yan J, Huang X. Sensor kinase KinB and its pathway-associated key factors sense the signal of nutrition starvation in sporulation of *Bacillus subtilis*. Microbiologyopen. 2018;7(3):e00566.

114. Eswaramoorthy P, Duan D, Dinh J, Dravis A, Devi SN, Fujita M. The Threshold Level of the Sensor Histidine Kinase KinA Governs Entry into Sporulation in *Bacillus subtilis*. J Bacteriol. 2010 Aug;192(15):3870–82.

115. Molle V, Fujita M, Jensen ST, Eichenberger P, González-Pastor JE, Liu JS, et al. The Spo0A regulon of *Bacillus subtilis*. Mol Microbiol. 2003 Dec;50(5):1683–701.

116. Strauch MA, Trach KA, Day J, Hoch JA. Spo0A activates and represses its own synthesis by binding at its dual promoters. Biochimie. 1992 Aug;74(7–8):619–26.

117. Dürre P. Physiology and Sporulation in *Clostridium*. Microbiology Spectrum [Internet]. 2014 Aug 22 [cited 2020 Apr 19];2(4). Available from:

http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.TBS-0010-2012

118. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, et al. Characterization of the Sporulation Initiation Pathway of *Clostridium difficile* and Its Role in Toxin Production. JB. 2009 Dec 1;191(23):7296–305.

119. Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. C. difficile 630Δerm Spo0A Regulates Sporulation, but Does Not Contribute to Toxin Production, by Direct High-Affinity Binding to Target DNA. Chang Y-F, editor. PLoS ONE. 2012 Oct 31;7(10):e48608.

120. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol. 2011 Jul;193(13):3186–96.

121. Edwards AN, Nawrocki KL, McBride SM. Conserved Oligopeptide Permeases Modulate Sporulation Initiation in *Clostridium difficile*. Blanke SR, editor. Infect Immun. 2014 Oct;82(10):4276–91.

122. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The Key Sigma Factor of Transition Phase, SigH, Controls Sporulation, Metabolism, and Virulence Factor Expression in *Clostridium difficile*. Journal of Bacteriology. 2011 Jul 1;193(13):3186–96.

123. Edwards AN, Tamayo R, McBride SM. A Novel Regulator Controls *Clostridium difficile* Sporulation, Motility and Toxin Production. Mol Microbiol. 2016 Jun;100(6):954–71.

124. Edwards AN, Krall EG, McBride SM. Strain-Dependent RstA Regulation of *Clostridioides difficile* Toxin Production and Sporulation. Federle MJ, editor. J Bacteriol. 2019 Oct 28;202(2):e00586-19, /jb/202/2/JB.00586-19.atom.

125. Girinathan BP, Ou J, Dupuy B, Govind R. Pleiotropic roles of *Clostridium difficile* sin locus. Koehler TM, editor. PLoS Pathog. 2018 Mar 12;14(3):e1006940.

126. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, et al. Effect of tcdR Mutation on Sporulation in the Epidemic *Clostridium difficile* Strain R20291. Ellermeier CD, editor. mSphere. 2017 Feb 22;2(1):mSphere.00383-16, e00383-16.

127. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. CodY-Dependent Regulation of Sporulation in *Clostridium difficile*. Henkin TM, editor. J Bacteriol. 2016 Aug 1;198(15):2113–30.

128. Edwards AN, McBride SM. Initiation of sporulation in *Clostridium difficile* : a twist on the classic model. FEMS Microbiol Lett. 2014 Sep;358(2):110–8.

129. Childress KO, Edwards AN, Nawrocki KL, Anderson SE, Woods EC, McBride SM. The Phosphotransfer Protein CD1492 Represses Sporulation Initiation in *Clostridium difficile*. Young VB, editor. Infect Immun. 2016 Dec;84(12):3434–44.

130. Péchiné S, Denève-Larrazet C, Collignon A. *Clostridium difficile* Adhesins. Methods Mol Biol. 2016;1476:91–101.

131. Guerrant RL, Steiner TS, Lima AAM, Bobak DA. How Intestinal Bacteria Cause Disease. J INFECT DIS. 1999 Mar;179(s2):S331–7.

132. Tulli L, Marchi S, Petracca R, Shaw HA, Fairweather NF, Scarselli M, et al. CbpA: a novel surface exposed adhesin of *Clostridium difficile* targeting human collagen. Cellular Microbiology. 2013;15(10):1674–87.

133. Lin Y-P, Kuo C-J, Koleci X, McDonough SP, Chang Y-F. Manganese Binds to *Clostridium difficile* Fbp68 and Is Essential for Fibronectin Binding. J Biol Chem. 2011 Feb 4;286(5):3957–69.

134. Hennequin C, Janoir C, Barc M-C, Collignon A, Karjalainen T. Identification and characterization of a fibronectin-binding protein from *Clostridium difficile*. Microbiology. 2003 Oct 1;149(10):2779–87.

135. Calabi E, Calabi F, Phillips AD, Fairweather NF. Binding of *Clostridium difficile* Surface Layer Proteins to Gastrointestinal Tissues. IAI. 2002 Oct;70(10):5770–8.

136. Spigaglia P, Barketi-Klai A, Collignon A, Mastrantonio P, Barbanti F, Rupnik M, et al. Surface-layer (S-layer) of human and animal *Clostridium difficile* strains and their behaviour in adherence to epithelial cells and intestinal colonization. Journal of Medical Microbiology. 2013 Sep 1;62(9):1386–93.

137. Merrigan MM, Venugopal A, Roxas JL, Anwar F, Mallozzi MJ, Roxas BAP, et al. Surface-Layer Protein A (SlpA) Is a Major Contributor to Host-Cell Adherence of *Clostridium difficile*. Popoff MR, editor. PLoS ONE. 2013 Nov 12;8(11):e78404.

138. Waligora A-J, Hennequin C, Mullany P, Bourlioux P, Collignon A, Karjalainen T. Characterization of a Cell Surface Protein of*Clostridium difficile* with Adhesive Properties. Infection and Immunity. 2001 Apr 1;69(4):2144–53.

139. Janoir C, Grénery J, Savariau-Lacomme M-P, Collignon A. Caractérisation d'une protéase extracellulaire de *Clostridium difficile*. Pathologie Biologie. 2004 Oct 1;52(8):444–9.

140. Kirby JM, Ahern H, Roberts AK, Kumar V, Freeman Z, Acharya KR, et al. Cwp84, a Surface-associated Cysteine Protease, Plays a Role in the Maturation of the Surface Layer of *Clostridium difficile*. J Biol Chem. 2009 Dec 11;284(50):34666–73.

141. Janoir C, Péchiné S, Grosdidier C, Collignon A. Cwp84, a Surface-Associated Protein of *Clostridium difficile*, Is a Cysteine Protease with Degrading Activity on Extracellular Matrix Proteins. JB. 2007 Oct 15;189(20):7174–80.

142. Kovacs-Simon A, Leuzzi R, Kasendra M, Minton N, Titball RW, Michell SL. Lipoprotein CD0873 Is a Novel Adhesin of *Clostridium difficile*. Journal of Infectious Diseases. 2014 Jul 15;210(2):274–84.

143. Bradshaw WJ, Kirby JM, Roberts AK, Shone CC, Acharya KR. Cwp2 from *Clostridium difficile* exhibits an extended three domain fold and cell adhesion in vitro. FEBS J. 2017 Sep;284(17):2886–98.

144. Collignon A, Bourlioux P, Waligora-Dupriet A-J, Karjalainen T, Barc M-C, Hennequin C, et al. GroEL (Hsp60) of *Clostridium difficile* is involved in cell adherence. Microbiology. 2001 Jan 1;147(1):87–96.

145. Hensbergen PJ, Klychnikov OI, Bakker D, Ras N, Kemp AC, Cordfunke RA, et al. A Novel Secreted Metalloprotease (CD2830) from *Clostridium difficile* Cleaves Specific Proline Sequences in LPXTG Cell Surface Proteins□S. :14.

146. Peltier J, Shaw HA, Couchman EC, Dawson LF, Yu L, Choudhary JS, et al. Cyclic diGMP Regulates Production of Sortase Substrates of *Clostridium difficile* and Their Surface Exposure through ZmpI Protease-mediated Cleavage. J Biol Chem. 2015 Oct 2;290(40):24453–69.

147. Arato V, Gasperini G, Giusti F, Ferlenghi I, Scarselli M, Leuzzi R. Dual role of the colonization factor CD2831 in *Clostridium difficile* pathogenesis. Sci Rep. 2019 Dec;9(1):5554.

148. Tasteyre A, Barc M-C, Collignon A, Boureau H, Karjalainen T. Role of FliC and FliD Flagellar Proteins of*Clostridium difficile* in Adherence and Gut Colonization. DiRita VJ, editor. Infect Immun. 2001 Dec 1;69(12):7937–40.

149. Borriello SP, Welch AR, Barclay FE, Davies HA. Mucosal association by Clostridiurndificile in the hamster gastrointestinal tract. :6.

150. Baban ST, Kuehne SA, Barketi-Klai A, Cartman ST, Kelly ML, Hardie KR, et al. The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. Popoff MR, editor. PLoS ONE. 2013 Sep 23;8(9):e73026.

151. McKee RW, Aleksanyan N, Garrett EM, Tamayo R. Type IV Pili Promote *Clostridium difficile* Adherence and Persistence in a Mouse Model of Infection. Young VB, editor. Infect Immun. 2018 Feb 26;86(5):e00943-17, /iai/86/5/e00943-17.atom.

152. Nicholls L, Grant TH, Robins-Browne RM. Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic *Escherichia coli* to epithelial cells. Molecular Microbiology. 2000;35(2):275–88.

153. Seekatz AM, Young VB. *Clostridium difficile* and the microbiota. J Clin Invest. 2014 Oct 1;124(10):4182–9.

154. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, et al. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet. 2006 Jul;38(7):779–86.

155. Wüst J, Sullivan NM, Hardegger U, Wilkins TD. Investigation of an outbreak of antibiotic-associated colitis by various typing methods. J Clin Microbiol. 1982 Dec;16(6):1096–101.

156. Hussain HA, Roberts AP, Mullany P. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. J Med Microbiol. 2005 Feb;54(Pt 2):137–41.

157. Farrow KA, Lyras D, Rood JI. The macrolide-lincosamide-streptogramin B resistance determinant from *Clostridium difficile* 630 contains two erm(B) genes. Antimicrob Agents Chemother. 2000 Feb;44(2):411–3.

158. Knight DR, Elliott B, Chang BJ, Perkins TT, Riley TV. Diversity and Evolution in the Genome of *Clostridium difficile*. Clin Microbiol Rev. 2015 Jul;28(3):721–41.

Chapter 2 - Spo0A suppresses *sin* locus expression in *Clostridioides difficile*

2.1 Abstract

Clostridioides difficile is the leading cause of nosocomial infection and is the causative agent of antibiotic-associated diarrhea. The severity of the disease is directly associated with the production of toxins, and spores are responsible for the transmission and persistence of the organism. Previously we characterized *sin* locus regulators SinR and SinR', where SinR is the regulator of toxin production and sporulation, while the SinR' acting as its antagonist. In *Bacillus subtilis*, Spo0A, the master regulator of sporulation, regulates SinR, by regulating the expression of its antagonist *sinI*. However, the role of Spo0A in the expression of *sinR* and *sinR'* in *C. difficile* is not yet reported. In this study, we created *spo0A* mutants in three different *C. difficile* strains R20291, UK1, and JIR8094, to understand the role of Spo0A in *sin* locus expression. Western blot analysis revealed that *spo0A* mutants had increased SinR levels. The qRT-PCR analysis for its expression further supported this data. By carrying out genetic and biochemical assays, we have shown that Spo0A can bind to the upstream region of this locus to regulates its expression. This study provides vital information that Spo0A regulates *sin* locus, which controls critical pathogenic traits such as sporulation, toxin production, and motility in *C. difficile*.

2.2 Introduction

Clostridioides difficile is a Gram-positive, anaerobic bacillus and is the principal causative agent of antibiotic-associated diarrhea and pseudomembranous colitis (1–3). Antibiotic use is the primary risk factor for the development of *C. difficile* associated disease because it disrupts normal protective gut flora and provides a favorable environment for *C. difficile* to colonize the colon. Two major pathogenic traits of *C. difficile* are toxin (toxin A and B) and spores (3–5). Deaths related to *C. difficile* increased by 400% between 2000 and 2007, in part because of the emergence of more aggressive *C. difficile* strains (6,7). Robust sporulation and toxin production were suspected of contributing to the widespread of the *C. difficile* infections associated with these highly virulent strains (8–14). How *C. difficile* triggers toxin production and sporulation in the intestinal environment is only beginning to be understood.

Recently, we reported the identification and characterization of master regulator SinR in *C. difficile* that was found to regulate sporulation, toxin production, and motility (15). SinR in the Gram-positive model organism *B. subtilis* is well characterized and is known to regulate multiple pathways, including sporulation, competence, motility, and biofilm formation (16–18). In *B. subtilis*, SinR' is encoded by the first gene of the downstream operon called *sin* (sporulation inhibition) locus and its transcription is driven from two promoters. The second gene in the operon, *sinR*, is transcribed by an internal promoter and is constitutively expressed. SinR represses the first committed (stage II) genes in the sporulation pathway (17).The promoter upstream of the operon is activated by phosphorylated Spo0A, leading to the expression of *sinI*, along with *sinR*. The SinI protein binds and inhibits the DNA binding activity of SinR (19–21). The combined effect of positive regulation by Spo0A∼P and the inactivation of the negative regulator SinR activates the sporulation pathway. In *C. difficile* also the *sin* locus is a two-gene operon and encodes for SinR

and SinR'. In our initial characterization of the *C. difficile sin* locus, we have shown that disruption of *sin* locus (absence of both SinR and SinR') resulted in asporogenic, less toxic, and less motile phenotype (15). Another study which reports that *C. difficile sin* locus suppresses biofilm formation corroborates our finding (22). Further investigation demonstrated that among the two regulators, SinR positively influences sporulation, toxin production, and motility, while SinR' acts as an antagonist to SinR and control its activity. Since *sin* locus has a role in regulating various pathogenic traits in *C. difficile*, understanding the regulation of its expression is important. Earlier, we have shown that disrupting the first gene *sinR* in the operon to affects both the *sinR* and the downstream *sinR'* transcription. This led to the assumption that unlike *B. subtilis*, the *C. difficile sin* locus is transcribed from a single upstream promoter. Real-time RT PCR analysis of the cells grown *in vitro*, demonstrated *sin* locus expression at 10 hour time point, indicating its tight regulation (15). From various gene expression data, we can observe that mutations in *sigH* and *spo0A* positively influence the expression of *sinRR'* (23–26) and mutations in *tcdR* downregulate their expression (27). We have also demonstrated that CodY can directly bind to *sin* locus upstream DNA to transcriptionally repress its expression (15). In the same line of investigation, in this study, we discovered that Spo0A, the sporulation master regulator, represses *sinR* expression. The effect was directly caused by the specific binding of Spo0A to the promoter region upstream of the locus.

2.3 Material and methods

2.3.1 Bacterial strains and growth conditions

Clostridioides difficile strains (Table S 2.1) were grown in TY (Tryptose and Yeast extract) agar or broth culture in an anaerobic chamber which is maintained at 10% H_2 , 10% CO_2 and 80% N_2 as described previously (27–30). Lincomycin (Linc 20µg/ml) and thiamphenicol (Thio; 15 µg/ml) were added to the culture medium when required. S17-1, an *E.coli* strain used for conjugation (31), was cultured aerobically in LB (Luria-Bertani) broth or agar and was supplemented with ampicillin (100 µg/ml) or chloramphenicol (25 µg/ml) when necessary.

2.3.2 General DNA techniques

Chromosomal DNA was extracted from *C. difficile* cultures using DNeasy Blood and Tissue Kit (Qiagen). PCR reactions were carried out using gene specific primers (Table S 2.2). PCR products were extracted from the gel using Geneclean Kit (mpbio). Plasmid DNA was extracted using QIAprep Spin Miniprep Kit (Qiagen). Standard procedures were used to perform routine cloning.

Table S 2.1. Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Relevant features or genotype	Reference
Clostridium difficile JIR8094	Erm_S derivative of strain 630	(32)
Clostridium difficile R20291	Clinical isolate - NAP1/027 ribotype, isolated in 2006 following an outbreak in Stoke Mandeville Hospital, UK	(33)
Clostridium difficile UK1	Clinical isolate	(34)
Escherichia coli DH5α	*endA1 recA1 deoR hsdR17 (r_{K-} m_{K+})*	NEB
Escherichia coli S17-1	Strain with integrated RP4 conjugation transfer function; favors conjugation between *E. coli* and *C. difficile*	(35)
Clostridium difficile JIR8094::*spo0A*	JIR8094 with intron insertion within *spo0A*	This study
Clostridium difficile R20291::*spo0A*	R20291with intron insertion within *spo0A*	(36)
Clostridium difficile UK1::spo0A	UK1 with intron insertion within *spo0A*	This study
pMTL007-CE5	ClosTron plasmid	(37)
pMTL007-CE5:Cdi-*spo0A*-178-179a	pMTL007-CE5 with group II intron targeted to *spo0A*	This study
pRPF185	*E. coli*/ *C. difficile* shuttle plasmid	(38)
pBA009	pRPF185 containing ~340 bp *sinR* WT promoter	This study
pBA029	pRPF185 containing *spoIIAB* promoter	This study
pBA030	pRPF185 containing *sinR* RGI mutated promoter	This study
pBA031	pRPF185 containing *sinR* RGII mutated promoter	This study
pBA032	pRPF185 containing *sinR* RGI and RGII mutated promoter	This study
pBA038	pRPF185 containing ~600 bp *sinR* WT promoter	This study
pBA037	pRPF1 containing ~475 bp *sinR* WT promoter	This study
pBA040	pRPF185 containing *gusA* gene without the *tet* promoter	This study

2.3.3 Construction and complementation of *C. difficile spo0A* mutant strains

The *spo0A* mutants in JIR8094 and UK1 strains were created using ClosTron gene knockout system as described previously (24,25,39,40). Briefly, for *spo0A* disruption, the group II intron insertion site between nucleotides 178 and 179 in *spo0A* gene in the antisense orientation was selected using a web-based design tool called the Perutka algorithm. The designed retargeted intron was cloned into pMTL007-CE5 as described previously (41). The resulting plasmid pMTL007-CE5::*spo0A*-178-179a was transferred into *C. difficile* UK1 and JIR8094 cells by conjugation. The potential Ll.ltrB insertions within the target genes in the *C. difficile* chromosome was conferred by the selection of lincomycin resistant transconjugants in 20 μg/ ml lincomycin plates. PCR using gene-specific primers (Table S 2.2) in combination with the EBS-U universal was performed to identify putative *C. difficile* mutants. *Clostridioides difficile spo0A* mutants were complemented by introducing pRG312 which contains the *spo0A* gene with the 300 bp upstream region, through conjugation. Complementation was confirmed by PCR and western blot analysis.

2.3.4 Western blot analysis

C. difficile cultures for western blot analysis were harvested and washed in 1xPBS solution. The pellets were resuspended in sample buffer (Tris 80 mM; SDS 2%; Glycerol 10%) and lysed by sonication. The whole cell extracts were then centrifuged at 17,000 g at 4°C for 1 min. The lysate was heated at 100°C for 7 min and the proteins were separated by SDS-PAGE and electro-blotted onto PVDF membrane. The blots were then probed with specific primary and the secondary antibodies at a dilution of 1:10,000. Immuno-detection of proteins was carried out using ECL kit (Millipore) following the manufacturer's recommendations and were developed using G-Box

iChemi XR scanner. Blot images were overlapped with the original images of the membrane to visualize pre-stained marker.

2.3.5 Construction of reporter plasmids and beta-glucuronidase assay

The *sin* locus upstream DNA regions of various lengths were amplified by PCR using specific primers with *KpnI* and *SacI* (Table S 2.2) recognition sequences. R20291 strain chromosomal DNA was used as a template for this amplification. Plasmid pRPF185 carries a *gusA* gene for beta-glucuronidase under tetracycline-inducible (*tet*) promoter. The *tet* promoter was removed using *KpnI* and *SacI* digestion and was replaced with the *sin* locus upstream regions of various lengths to create plasmids pBA009, pBA037, and pBA038 (Table S 2.1). The control plasmid pBA040 with promoter less *gusA* was created by digesting with *KpnI* and *SacI* to remove *tet* promoter and then self-ligated after creating blunt ends. Plasmids were introduced into R20291 and R20291::*spo0A* strains through conjugation as described previously (15,27). The transconjugants were grown in TY medium in the presence of thiamphenicol (15 μg/ml) overnight. Overnight cultures were used as an inoculum at a 1:100 dilution to start a new culture. Bacterial cultures were harvested at 10 hr of growth and the amount of for beta-glucuronidase activity was assessed as described elsewhere (42,43). Briefly, the cells were washed and resuspended in 1 ml of Z buffer (60 mM $Na_2HPO_4.7H_20$ pH 7.0, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, 1mM $MgSO_4.7H_2O$ and 50mM 2ME) and lysed by homogenization. The lysate was mixed with 160 μl of 6mM p-nitrophenyl β-D-glucuronide (Sigma) and incubated at 37°C. The reaction was stopped by the addition of 0.4 ml of 1.0 M $NaCO_3$. β-Glucuronidase activity was calculated as descried earlier (42,43).

2.3.6 Mutagenesis of *sin* locus promoter region

Quick Change Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) was used to carry out site directed mutagenesis whereby G and C residues of the potential Spo0A binding '0A' boxes were substituted with A residues. The mutagenic oligonucleotide primers used are listed in Table S 2.2.

2.3.7 DNA binding

DNA binding was carried out as described elsewhere (44). Briefly, the promoter region of interest was biotin labelled and was coupled to immobilized Monomeric Avidin Resin (G Biosciences) in B/W Buffer (44). The DNA and the beads were incubated at room temperature for 30 min in a rotor. The bead-DNA complex was washed with TE Buffer to remove any unbound DNA. To prepare cell lysates, *C. difficile* R20291 strain was grown to late exponential phase (16 hour) in 500 ml TY medium, pH 7.4. After washing with 1XPBS, the cells were resuspended in BS/THES buffer (44) and lysed using French press. The whole lysate was centrifuged at 20,000 g for 30 min at 4°C and the supernatant was incubated with the bead-DNA complex and allowed to rotate at 4°C overnight. The bead-DNA-protein complex was washed with BS/THES Buffer (5 times). Elution was carried out with 50mM, 100 mM, and 200 mM NaCl in Tris-HCl pH 7.4. The eluates were analyzed by SDS-PAGE and Western Blotting using Spo0A specific antibody.

2.4 Results

2.4.1 Elevated level of SinR is present in *spo0A* mutant

In *C. difficile*, we have previously shown that *sin locus* mutant is asporogenic and this phenotype is associated with downregulation of *spo0A* expression. Interestingly, disruption of *sinR',* the second gene in the locus resulted in elevated levels of sporulation. This result suggested that SinR as a positive regulator of the sporulation. Gene expression data of *spo0A* mutants from different studies have shown elevated levels of *sin locus* expression when compared to their respective parents (25,39,45). These observations taken together suggest that these two master regulators, Spo0A and SinR regulate each other's transcription. To understand the possible regulatory relationship between SinR and Spo0A in *C. difficile*, we created *spo0A* mutant in two different *C. difficile* strains, JIR8094 and UK1 strain using the ClosTron mutagenesis technique. Mutation in *spo0A* was confirmed by PCR (Fig. S 2.1ABC) and western blot analysis using Spo0A specific antibodies (Fig. S 2.1D). The *spo0A* mutant in R20291 obtained from Dena Lyras Lab (46) was also included in the analysis. As previously reported, mutation in *spo0A* resulted in the asporogenic phenotype (24,25,40,46). For complementation, plasmid pRG312 carrying *spo0A* under its own promoter was introduced into the mutants through conjugation. Introduction of the *spo0A* expressing plasmid was successful in the JIR8094::*spo0A* and in R20291::*spo0A* mutants, but unsuccessful in the UK1::*spo0A* mutant. Heat resistant spores were observed in the complemented strains, however the levels were significantly lower than wild-type (Fig. S 2.2). To test whether Spo0A influences expression of the *sin* locus genes, we performed quantitative reverse transcription-PCR (qRT-PCR) analysis of the *sinR* and *sinR'* transcripts in *spo0A* mutants and their respective parent strains. As previously reported, the level of *sinR* and *sinR'* transcripts were increased several folds, (Fig. 2.1A) in all three *spo0A* mutants compared to their parent

strains. To further confirm this result, we performed western blot analysis using SinR specific antibodies. We grew the mutants and the respective parent strains in TY medium for 10 hours and observed the levels of SinR in their cytosol. We found that *spo0A* mutants in all three strains produced higher amounts of SinR compared to their respective parents (Fig. 2.1B, Fig S 2.1D). However, in our complementation of R20291::*spo0A* and JIR8094::*spo0A*, we did not see lower levels of SinR (Fig. 2.1B). Together, these data suggest that Spo0A negatively regulates the expression of *sinR*.

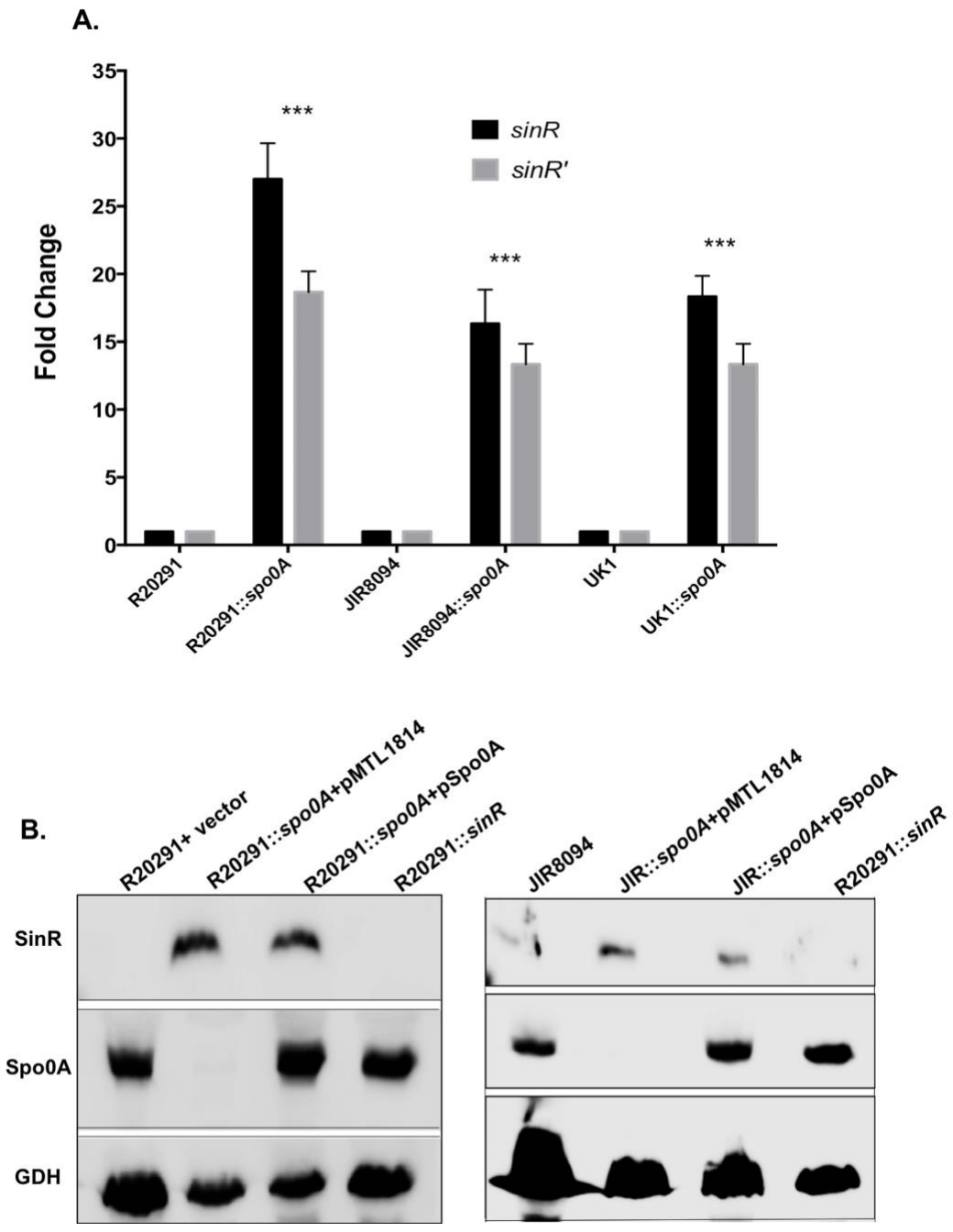


Figure 2.1. In the absence of Spo0A, *C. difficile* produces elevated levels of SinR

A. qRT-PCR results of *sin* locus transcripts in R20291; R20291::*spo0A* ; JIR8094 and JIR8094::*spo0A* strains collected at 10 hr time point. The representative results from three independent experiments are shown. The asterisks (***) indicate statistical difference at a *P* value of <0.005. **B.** Western blot analysis parents (R20291 and JIR8094), and their respective *spo0A* mutants using SinR and Spo0A specific antibodies demonstrating upregulated SinR in the absence of Spo0A. The SinR negative R20291::*sinRR'* mutant served as a negative control. GDH detection using anti-GDH antibodies was used as loading control.

2.4.2 Spo0A represses the expression of *sinR*

Spo0A is a transcriptional regulator and is a DNA binding protein. Spo0A binds to specific DNA sequences in the promoter region of its target gene to regulate their expression. To determine if the elevated levels of SinR observed in *spo0A* mutants is due to the repressor activity of Spo0A, we performed reporter fusion assays. We fused 600 bps of *sin* locus upstream DNA with the *gusA* reporter gene coding for beta-glucuronidase and the construct was introduced into the R2091::*spo0A* mutant and its parent strain. The plasmid carrying a promoter-less *gusA* was used as a negative control. We also cloned the promoter region of *spoIIAB* known to be regulated by Spo0A, with the *gusA* and used this construct as a positive control. The *spoIIAB* promoter is positively regulated by Spo0A and was found to be active only in the parent strain and not in the *spo0A* mutant (Fig. 2.2A). We observed significantly higher beta glucuronidase activity when it was expressed from the *sin* locus promoter in the R20291::*spo0A* mutant strain compared to the parent strain, where very minimal reporter activity was recorded. This observation is consistent with our western blot results, where we detected elevated levels of SinR in *spo0A* mutant strains. Taken together these results suggest that Spo0A represses the transcription of *sinR* either directly or indirectly. To narrow down the Spo0A controlled region in the *sin* locus promoter, we cloned the 475 and 340 bps of the upstream DNA with the *gusA* gene and performed the reporter fusion assays. The reporter gene activity was similar in the cultures carrying the 600 bp as well as 340 bp upstream fusions (Fig. 2.2B). This indicated that both the *sin* locus promoter and the Spo0A regulated regions are present within this 340 bp region.

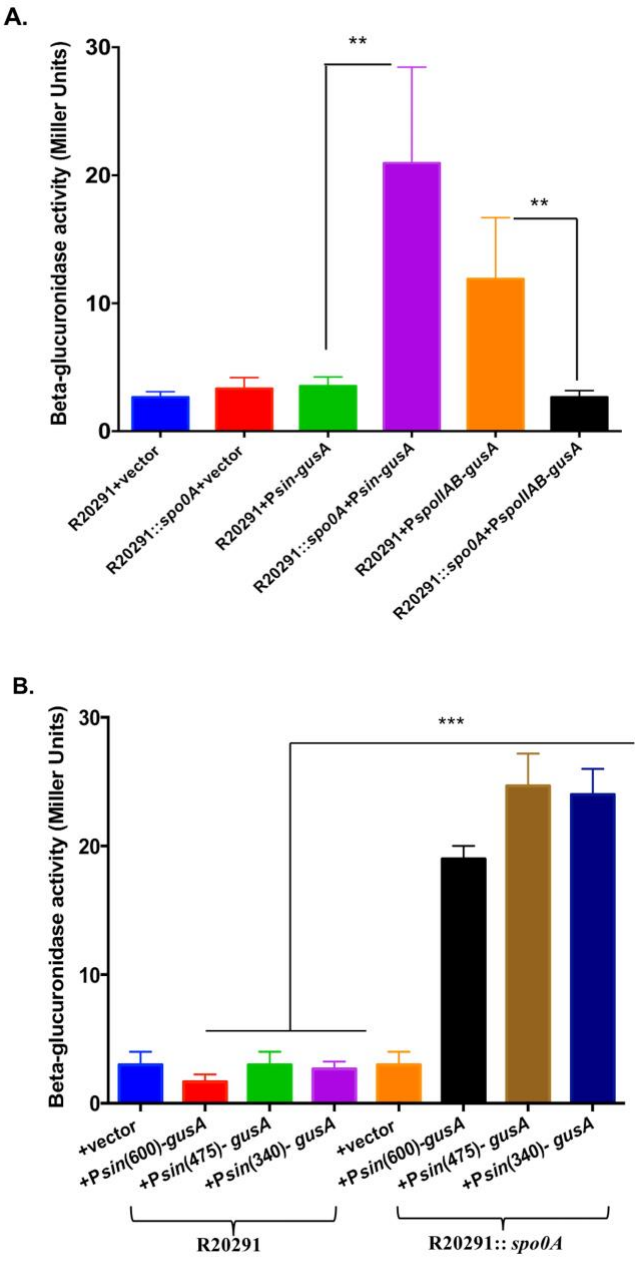


Figure 2.2. Spo0A represses *sin* locus expression

A. Beta-glucuronidase activity of the P*sin-gusA* fusions in the parent R20291 and R20291::*spo0A* mutant. Plasmid (pBA038) with *gusA* as the reporter gene fused to 600 bp of *sin* locus upstream. Plasmids carrying P*spoIIAB-gusA* (pBA029) and a promoter less *gusA* (pBA040) were used as positive and negative controls, respectively. **B.** Expression of beta-glucuronidase in parent R20291 and *spo0A* mutant carrying plasmids pBA037 (475 bps P*sin-gusA*), pBA038 (600 bp P*sin-gusA*) and pBA009 (340 bp P*sin-gusA*). The error bars in panels **A** and **B** correspond to standard errors of the means of results from 3 biological replicates, where ** and *** indicates $P<0.05$ and $P<0.005$, respectively (by two-tailed Student *t* test). At least three independent experiments were performed.

2.4.3 Spo0A binds to the promoter region of *sinR*

The results in Fig 2.2AB show that expression of P*sin-gusA* was less in the R20291 background while the expression of the reporter gene was at higher levels in R20291::*spo0A* background. To determine whether the repression of *sinR* by Spo0A is due to Spo0A binding specifically to the promoter region of *sinR*, we carried out a DNA binding experiment. Considering that Spo0A needs to be phosphorylated to bind to the target DNA and inability to purify Spo0A-P, we did not attempt the *in vitro* electrophoretic gel shift assay. Instead, we used a biotin labelled DNA pulldown assay to determine the DNA binding ability of Spo0A under native conditions. The DNA segment representing the promoter region of *sinR* was biotinylated and was coupled to Immobilized Monomeric Avidin Resin. This bead-DNA complex was incubated with the cell lysate from the parent R20291 strain. The bound proteins were eluted and were run in SDS page and immunoblotted with Spo0A antibody. We first standardized the binding experiment by using *spo0IIAB* promoter region as a positive control. Spo0A protein could be detected in the elutes when the *spo0IIAB* upstream DNA was used as the bait. The biotinylated *gluD* upstream DNA also processed similarly and served as a negative control. We applied the same protocol using the biotinylated 340 bp *sin* upstream DNA as bait. Results demonstrated that it could pull down Spo0A, suggesting that Spo0A binds specifically to the promoter region of the *sin* locus (Fig. 2.3A). Next, to narrow down the Spo0A binding site within that 340 bps we created three biotin-labelled fragments covering the first 118 bps (340-222 bps upstream), the last 140 bps and overlapping 135 bps mid-region (237 to 102 bps upstream) (Fig. 2.3B) and used them as bait in the pull-down experiment. More Spo0A was detected when the 140 bps mid-region was used as a bait compared to the first 118 bps (Fig. 2.3C). The Spo0A protein appears to bind with greater affinity to the middle fragment, when compared to the other two fragments. Since the biotin DNA

pulldown assay is semi quantitative in nature, this assumption needs further validation. Spo0A could not be recovered from the elute from the binding of *gluD* upstream region, suggesting the specificity of the Spo0A binding with the *sin* locus promoter.

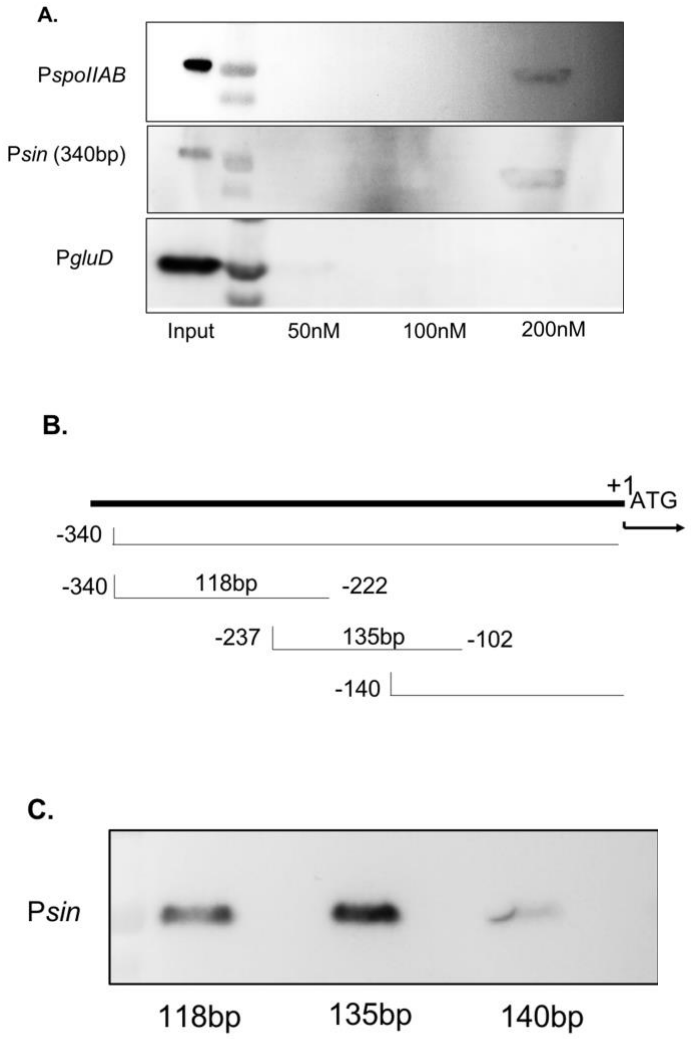


Figure 2.3. Spo0A binds to *sin* locus upstream DNA

A. Western blot analysis using Spo0A specific antibody to detect endogenous Spo0A in input and elute fractions. For the biotin-DNA pulldown assay, the promoter regions of *spoIIAB* and *gluD* were used as positive and negative controls, respectively. **B**. Schematics of the 340 bp upstream *sin* locus denoting the translation start (+1) of *sinR*. The lower lines indicate the location and the sizes of the DNA fragments used for the biotinylated DNA-pulldown assay. **C.** Three DNA fragments (118 bp; 135 bp; 140 bp) spanning different regions of the 340 bp *sin* locus upstream were used independently to carry out the binding and Spo0A was detected as in panel A.

2.4.4 Mutational analysis of the *sinR* upstream region

In *B. subtilis* Spo0A-P binds to 7 bp DNA element 5'-TGNCGAA-3', commonly known as Spo0A box (47). However, there are certain exceptions where Spo0A binds to degenerated Spo0A boxes with mismatches in the upstream of some targets (48,49). The DNA binding domain of *C. difficile* Spo0A is highly homologous to *B. subtilis* Spo0A and the key residues of Spo0A known to mediate the interaction with the bases of the 0A box are highly conserved in *Bacillus* and *Clostridium* species (24,50). In *C. difficile*, Spo0A is known to bind to *spo0A* upstream and *sigH* upstream. Both of these genes have the TGTCGAA consensus Spo0A box sequence (23,51). *Clostridioides difficile* Spo0A also bind to the upstream of *spoIIAA-spoIIE-spoIIGA* operon with low affinity, where the binding sequence is a degenerated Spo0A box with TACGACA sequence (23). We scanned the upstream region of *sinR* for potential Spo0A binding consensus sequence. From the DNA binding pull-down experiment we could predict that Spo0A binds to sequences within the 237-102 bps upstream of *sin* locus. Classical Spo0A binding boxes couldn't be identified in this region. However, two repeats with TATTGTAG sequences could be seen in this region. We mutated the TATTGTAG into TATTATAA, created reporter fusions with mutation in these sequences and analyzed their effect on the reported expression. None of the mutations affected the expression of the reporter gene (Fig 2.4). Mutations were introduced in the repeat sequence TAGTCTAT that occurred within the first 88 bps of *sin* locus upstream (data not shown). These changes didn't affect the *sin* locus expression. Although we could not identify the specific Spo0A binding boxes in the *sin* locus upstream, we have mapped the region to which it binds.

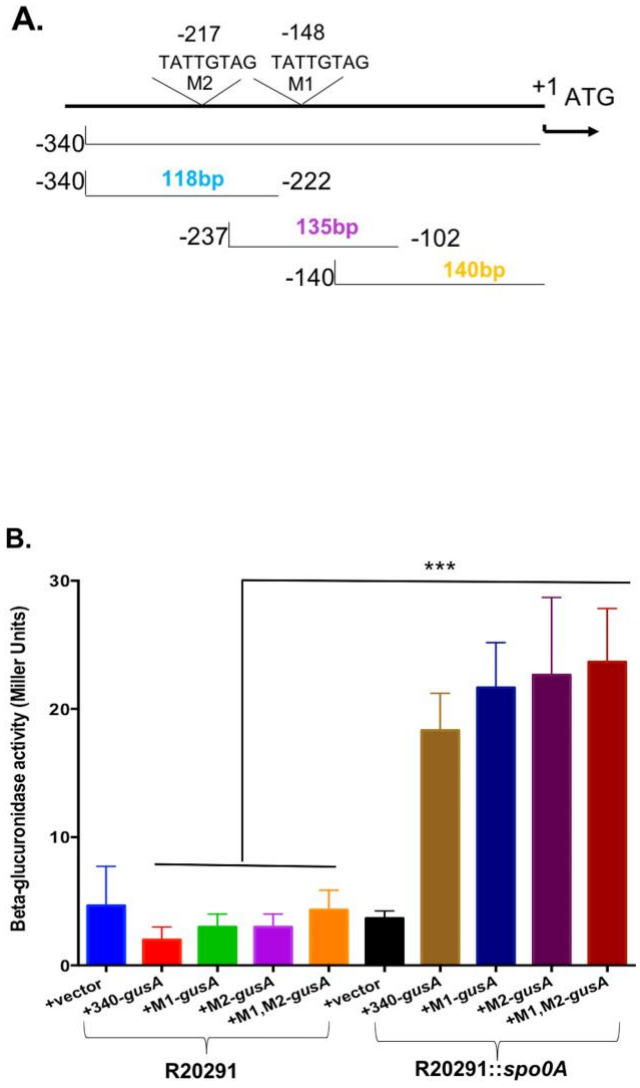


Figure 2.4. Spo0A represses the *sin* expression in M1 and M2 mutated promoter

A. Schematics of the 340 bp upstream *sin* locus denoting M1 and M2 mutations in the predicted Spo0A Box. **B.** Expression of beta glucuronidase in parent R20291 and R20291::*spo0A* mutant strains carrying plasmids with *gusA* as the reporter gene fused to the promoter of *sinR*. The predicted Spo0A box was mutagenized either on M1 or M2 site or both M1, M2. Strain carrying a promoter less *gusA* plasmid (pBA040) was used as control. Data represent the means ± standard errors of the means (SEM) (n = 3). The asterisks (***) in panel A indicate statistical difference at a P value of <0.005.

2.5 Discussion

Sporulation in a cell is an intense response to stress and is particularly expensive, in terms of both time and materials (52). The exact conditions and timing for sporulation are likely to be under strong selective pressure as both premature and belated spore production can have disastrous effects on cell growth and survival. In *B. subtilis*, the *sin* (sporulation *in*hibition) operon is central to the timing and early dynamics of this network (53–55), and its regulation is controlled by the sporulation master regulator Spo0A itself . In this study, we have demonstrated that similar to *B. subtilis*, the *sin* locus in *C. difficile* is also regulated by Spo0A. We created *spo0A* mutant in two different *C. difficile* strains JIR8094 and UK1, and included the previously created R20291::*spo0A* mutant in the analysis. As previously noted, we found both the JIR8094::*spo0A* and the UK1::*spo0A* strain to be asporogenic in nature. Complementation of sporulation phenotype was not complete in JIR8094::*spo0A* and we could not complement the UK1::*spo0A* mutant strain even after multiple attempts for reasons unknown. Western blot analysis detected higher amount of SinR in all three different *spo0A* mutant strains. However, similar to the sporulation phenotype complementation of R20291::*spo0A* and JIR8094::*spo0A* did not lower the level of SinR. Unlike *B. subtilis* SinR, *C. difficile* SinR acts as a positive regulator of sporulation (15). How SinR controls sporulation is yet to be characterized. As we noted previously, *sin* locus is expressed at a low level only at a short window of time at 10-12 h of growth, suggesting that the presence of these regulators in the right amount at a specific time point could be important to regulate sporulation initiation and progression. In *spo0A* mutant however, *sinR* is expressed at a very high level which can affect the temporal nature of the sporulation pathway. It is worth noting that asporogenic phenotype of *sinRR'* mutant could not be complemented either (15). Failures to complement a *spo0A* mutation have been previously observed in *C. difficile*. Two independent studies

demonstrated incomplete restoration of sporulation phenotype in R20291::*spo0A* (39,46). However, when Deakin *et al.* (2012) tested the *spo0A* mutants in 630Δ*erm* and R20291 strains and they found *in vitro* levels of sporulation to be restored to wild type levels in their complemented derivative (40). When they tested the R20291 strains for toxin production, however, the complemented strain still produced increased levels of toxin compared to wild type (40). These observations suggest that this method of introducing *spo0A* using a multicopy plasmid may not be the suitable method for complementation considering the complex nature of Spo0A regulatory networks.

Like in many Gram-positive bacteria, Spo0A is the master regulator of sporulation in *C. difficile* (24,40,46). When the post-exponential phase begins, Spo0A activates the expression of the genes involved in the sporulation initiation process and positively regulates the sigma factor cascade required for sporulation (45). In many other pathogenic spore-forming bacteria the gene regulatory networks that influence sporulation and virulence are closely linked with each other (56–60). In *C. difficile*, the mutation in *spo0A* affected many pathogenic traits, including toxin production, flagella expression and biofilm formation (5,39,40,46). Mackin *et. al* observed a clear increase in the production of toxin A and B upon disruption of *spo0A* in the ribotype 027 isolates R20291 and M7404 (46). In a similar study, Deakin *et al.* found that a R20291 *spo0A* mutant caused more severe disease in a murine model than the wild type strain, and associated this increase in severity with an increase in the amount of toxin A and toxin B produced by the mutant *in vitro (40).* Dawson et. al. demonstrated that Spo0A in 630Δerm strain promotes sporulation cascade and biofilm formation in addition to negatively regulating virulence factor expression (toxins and flagella) (39). We found UK1::*spo0A* strain to produce higher toxins compared to its parent strain, while no significant difference was observed between the JIR8094 parent and JIR8094::*spo0A*

mutant (Fig S 2.3A). This observation was consistent with the previous report, where mutation in *spo0A* influenced the toxin production only in the 027 ribotype, which includes UK1 and R20291 strains, but not in the 630Δerm strain that belongs to 012 ribotype as the JIR8094. We measured the cytosolic toxins in all the three *spo0A* mutants and observed increased toxin production only in the R20291 and UK1 background, but not in the JIR8094 strain (Fig. S 2.3A). Reduced biofilm formation was also found only in R20291::*spo0A*, UK1::*spo0A*, but not in JIR8094::*spo0A* strains (Fig. S 2.3BC). The mechanism of Spo0A regulation over these pathways remains to be answered. In the *C. difficile* genome >100 open reading frames have potential 0A boxes within 500 bp of their start codons, indicating direct regulation by Spo0A (24). However, *tcdA* and *tcdB*, encoding toxin A and toxin B, respectively, are not among them, indicating the indirect influence of Spo0A on toxin production (24). Motility and biofilm formation could also be indirectly controlled by Spo0A, since many candidate regulators are encoded by the genes putatively under the direct control of Spo0A in *C. difficile* (24,40). Our current finding of Spo0A mediated *sin* locus regulation can partly explain many of the phenotypes displayed by *spo0A* mutants, especially in the ribotype 027 strains (39,46). In our initial characterization of the *sin* locus, we have shown decreased toxin production and motility in the absence of SinR and SinR' (15). Expression of *sinR* alone was sufficient to complement these phenotypes and suggested SinR as a positive regulator of these pathways (15). We have further shown that SinR controls toxin production by regulating *sigD*, a sigma factor that positively regulates *tcdR*, which is needed for the transcription of toxin genes (15,61,62). SigD is also needed for the transcription of flagellar operon in *C. difficile* (61,62). In this study, we have shown an increased SinR production in the absence of Spo0A (Fig. 2.2B). Quantitative RT-PCR results demonstrated increased expression of *sigD*, *tcdR*, and *tcdB* in the R20291::*spo0A* and in UK1::*spo0A*, compared to their respective parent strains (Fig. S 2.4).

Increased *sigD* expression can lead to increased flagellar, toxin production and a reduced biofilm formation in the *spo0A* mutant (22, 32) (Fig. S 2.3BC).

In this study, we have shown that Spo0A binds to the promoter of the *C. difficile sin* locus and suppresses the expression of both *sinR* and *sinR'*. Previously, we have shown that disrupting *sinR* by insertion mutagenesis affects both *sinR* and *sinR'* transcription (15), suggesting that *sinRR'* is transcribe as a bicistronic message. Our QRT-PCR analysis detected lower levels of *sinR'* transcripts than the *sinR* transcripts. Since the reduction is observed both in the parent and the *spo0A* mutants, we can conclude that this effect is independent of Spo0A. The reason for the reduced level of SinR' over SinR is not clear. In *B. subtilis* the polycistronic *sinRI* transcripts are produced from two upstream promoters. The monocistronic *sinR* transcripts are driven from a promoter located within the coding region of *sinI*. In *B. subtilis*, Spo0A activates the expression of *sinI* by binding to the upstream promoter of the operon and indirectly regulating SinR activity (19,20) . Similar to *B. subtilis*, in *C. difficile* also there is a possibility that *sinR'* may have an independent promoter within the *sinR* coding sequence and could be controlled by an unknown regulator in *C. difficile*.

In summary, we have demonstrated that Spo0A, the master regulator of sporulation, regulates the expression of *sin* locus. We have further shown that Spo0A can bind to the upstream region of *sin* locus have successfully mapped the region to which it binds.

2.6 References

1. Guh AY, Kutty PK. *Clostridioides difficile* Infection. Ann Intern Med. 2018 02;169(7):ITC49–64.

2. Prete RD, Ronga L, Addati G, Magrone R, Abbasciano A, Decimo M, et al. *Clostridium difficile*. A review on an emerging infection. La Clinica Terapeutica. 2019 Feb 14;1(170):e41–7.

3. Awad MM, Johanesen PA, Carter GP, Rose E, Lyras D. *Clostridium difficile* virulence factors: Insights into an anaerobic spore-forming pathogen. Gut Microbes. 2014 Sep 3;5(5):579–93.

4. Schäffler H, Breitrück A. *Clostridium difficile* - From Colonization to Infection. Front Microbiol. 2018;9:646.

5. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, et al. Characterization of the Sporulation Initiation Pathway of *Clostridium difficile* and Its Role in Toxin Production. Journal of Bacteriology. 2009 Dec 1;191(23):7296–305.

6. CDC Press Releases [Internet]. CDC. 2016 [cited 2018 Oct 23]. Available from: https://www.cdc.gov/media/releases/2015/p0225-clostridium-difficile.html

7. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of *Clostridium difficile* Infection in the United States. New England Journal of Medicine. 2015 Feb 26;372(9):825–34.

8. Depestel DD, Aronoff DM. Epidemiology of *Clostridium difficile* infection. J Pharm Pract. 2013 Oct;26(5):464–75.

9. Akerlund T, Persson I, Unemo M, Noren T, Svenungsson B, Wullt M, et al. Increased Sporulation Rate of Epidemic *Clostridium difficile* Type 027/NAP1. Journal of Clinical Microbiology. 2008 Apr 1;46(4):1530–3.

10. O'Connor JR, Johnson S, Gerding DN. *Clostridium difficile* Infection Caused by the Epidemic BI/NAP1/027 Strain. Gastroenterology. 2009 May;136(6):1913–24.

11. Camacho-Ortiz A, López-Barrera D, Hernández-García R, Galván-De Los Santos AM, Flores-Treviño SM, Llaca-Díaz JM, et al. First report of *Clostridium difficile* NAP1/027 in a Mexican hospital. PLoS ONE. 2015;10(4):e0122627.

12. Giancola SE, Williams RJ, Gentry CA. Prevalence of the *Clostridium difficile* BI/NAP1/027 strain across the United States Veterans Health Administration. Clin Microbiol Infect. 2018 Aug;24(8):877–81.

13. Fatima R, Aziz M. The Hypervirulent Strain of *Clostridium Difficile*: NAP1/B1/027 - A Brief Overview. Cureus. 2019 Jan 29;11(1):e3977.

14. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biology. 2009 Sep 25;10(9):R102.

15. Girinathan BP, Ou J, Dupuy B, Govind R. Pleiotropic roles of *Clostridium difficile* sin locus. PLOS Pathogens. 2018 Mar 12;14(3):e1006940.

16. Kearns DB, Chu F, Branda SS, Kolter R, Losick R. A master regulator for biofilm formation by *Bacillus subtilis*. Molecular Microbiology. 2005;55(3):739–49.

17. Cervin MA, Lewis RJ, Brannigan JA, Spiegelman GB. The *Bacillus subtilis* regulator SinR inhibits spoIIG promoter transcription *in vitro* without displacing RNA polymerase. Nucleic Acids Res. 1998 Aug 1;26(16):3806–12.

18. Mandic-Mulec I, Doukhan L, Smith I. The *Bacillus subtilis* SinR protein is a repressor of the key sporulation gene spo0A. J Bacteriol. 1995 Aug;177(16):4619–27.

19. Bai U, Mandic-Mulec I, Smith I. SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction. Genes Dev. 1993 Jan;7(1):139–48.

20. Milton ME, Draughn GL, Bobay BG, Stowe SD, Olson AL, Feldmann EA, et al. The Solution Structures and Interaction of SinR and SinI: Elucidating the Mechanism of Action of the Master Regulator Switch for Biofilm Formation in *Bacillus subtilis.* J Mol Biol. 2020 Jan 17;432(2):343–57.

21. Shafikhani SH, Mandic-Mulec I, Strauch MA, Smith I, Leighton T. Postexponential Regulation of sin Operon Expression in *Bacillus subtilis*. Journal of Bacteriology. 2002 Jan 15;184(2):564–71.

22. Poquet I, Saujet L, Canette A, Monot M, Mihajlovic J, Ghigo J-M, et al. *Clostridium difficile* Biofilm: Remodeling Metabolism and Cell Surface to Build a Sparse and Heterogeneously Aggregated Architecture. Front Microbiol. 2018;9:2084.

23. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The Key Sigma Factor of Transition Phase, SigH, Controls Sporulation, Metabolism, and Virulence Factor Expression in *Clostridium difficile*. Journal of Bacteriology. 2011 Jul 1;193(13):3186–96.

24. Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. *C. difficile* 630Derm Spo0A Regulates Sporulation, but Does Not Contribute to Toxin Production, by Direct High- Affinity Binding to Target DNA. PLOS ONE. 2012;7(10):12.

25. Pettit LJ, Browne HP, Yu L, Smits W, Fagan RP, Barquist L, et al. Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics. 2014;15(1):160.

26. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. CodY-Dependent Regulation of Sporulation in *Clostridium difficile*. Journal of Bacteriology. 2016 Aug 1;198(15):2113–30.

27. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, et al. Effect of tcdR Mutation on Sporulation in the Epidemic *Clostridium difficile* Strain R20291. mSphere. 2017 Feb;2(1).

28. Girinathan BP, Braun SE, Govind R. *Clostridium difficile* glutamate dehydrogenase is a secreted enzyme that confers resistance to H2O2. Microbiology. 2014 Jan 1;160(Pt_1):47–55.

29. Girinathan BP, Braun S, Sirigireddy AR, Lopez JE, Govind R. Importance of Glutamate Dehydrogenase (GDH) in *Clostridium difficile* Colonization In Vivo. Paredes-Sabja D, editor. PLOS ONE. 2016 Jul 28;11(7):e0160107.

30. Govind R, Dupuy B. Secretion of *Clostridium difficile* Toxins A and B Requires the Holin-like Protein TcdE. PLOS Pathogens. 2012 Jun 7;8(6):e1002727.

31. Teng F, Murray BE, Weinstock GM. Conjugal transfer of plasmid DNA from *Escherichia coli* to *enterococci:* a method to make insertion mutations. Plasmid. 1998;39(3):182–6.

32. O'Connor JR, Lyras D, Farrow KA, Adams V, Powell DR, Hinds J, et al. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol. 2006 Sep;61(5):1335–51.

33. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102.

34. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol. 2010 Oct;192(19):4983–90.

35. Teng F, Murray BE, Weinstock GM. Conjugal transfer of plasmid DNA from *Escherichia coli* to *enterococci*: a method to make insertion mutations. Plasmid. 1998;39(3):182–6.

36. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. Spo0A differentially regulates toxin production in evolutionarily diverse strains of *Clostridium difficile*. PLoS ONE. 2013;8(11):e79666.

37. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010 Jan;80(1):49–55.

38. Fagan RP, Fairweather NF. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem. 2011 Aug 5;286(31):27483–93.

39. Dawson LF, Valiente E, Faulds-Pain A, Donahue EH, Wren BW. Characterisation of *Clostridium difficile* biofilm formation, a role for Spo0A. PLoS ONE. 2012;7(12):e50527.

40. Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, et al. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. Infect Immun. 2012 Aug;80(8):2704–11.

41. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: A universal gene knock-out system for the genus *Clostridium*. Journal of Microbiological Methods. 2007 Sep;70(3):452–64.

42. Dupuy B, Sonenshein AL. Regulated transcription of *Clostridium difficile* toxin genes. Molecular Microbiology. 1998;27(1):107–20.

43. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, et al. Environmental Response and Autoregulation of *Clostridium difficile* TxeR, a Sigma Factor for Toxin Gene Expression. Journal of Bacteriology. 2002 Nov 1;184(21):5971–8.

44. Jutras BL, Verma A, Stevenson B. Identification of novel DNA-binding proteins using DNA-affinity chromatography/pull down. Curr Protoc Microbiol. 2012 Feb;Chapter 1:Unit1F.1.

45. Fimlaid KA, Bond JP, Schutz KC, Putnam EE, Leung JM, Lawley TD, et al. Global analysis of the sporulation pathway of *Clostridium difficile*. PLoS Genet. 2013;9(8):e1003660.

46. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. Spo0A Differentially Regulates Toxin Production in Evolutionarily Diverse Strains of *Clostridium difficile*. Popoff MR, editor. PLoS ONE. 2013 Nov 13;8(11):e79666.

47. Asayama M, Yamamoto A, Kobayashi Y. Dimer form of phosphorylated Spo0A, a transcriptional regulator, stimulates the *spo0F* transcription at the initiation of sporulation in *Bacillus subtilis*. J Mol Biol. 1995 Jun 30;250(1):11–23.

48. Baldus JM, Green BD, Youngman P, Moran CP. Phosphorylation of *Bacillus subtilis* transcription factor Spo0A stimulates transcription from the spoIIG promoter by enhancing binding to weak 0A boxes. J Bacteriol. 1994 Jan;176(2):296–306.

49. Chen G, Kumar A, Wyman TH, Moran CP. Spo0A-Dependent Activation of an Extended −10 Region Promoter in *Bacillus subtilis*. Journal of Bacteriology. 2006 Feb 15;188(4):1411–8.

50. Molle V, Fujita M, Jensen ST, Eichenberger P, González-Pastor JE, Liu JS, et al. The Spo0A regulon of *Bacillus subtilis*. Mol Microbiol. 2003 Dec;50(5):1683–701.

51. Saujet L, Pereira FC, Henriques AO, Martin-Verstraete I. The regulatory network controlling spore formation in *Clostridium difficile*. FEMS Microbiology Letters. 2014 Sep;358(1):1–10.

52. Hutchison EA, Miller DA, Angert ER. Sporulation in Bacteria: Beyond the Standard Model. Microbiol Spectr. 2014 Oct;2(5).

53. Grossman AD. Genetic Networks Controlling the Initiation of Sporulation and the Development of Genetic Competence in *Bacillus Subtilis*. Annual Review of Genetics. 1995;29(1):477–508.

54. Gaur NK, Dubnau E, Smith I. Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies. J Bacteriol. 1986 Nov;168(2):860–9.

55. Gaur NK, Cabane K, Smith I. Structure and expression of the *Bacillus subtilis* sin operon. J Bacteriol. 1988 Mar;170(3):1046–53.

56. Dale JL, Raynor MJ, Ty MC, Hadjifrangiskou M, Koehler TM. A Dual Role for the *Bacillus anthracis* Master Virulence Regulator AtxA: Control of Sporulation and Anthrax Toxin Production. Front Microbiol. 2018;9:482.

57. Deng C, Peng Q, Song F, Lereclus D. Regulation of cry gene expression in *Bacillus thuringiensis*. Toxins (Basel). 2014 Jul 23;6(7):2194–209.

58. Dürre P. Physiology and Sporulation in *Clostridium*. Microbiol Spectr. 2014 Aug;2(4):TBS-0010-2012.

59. Li J, Paredes-Sabja D, Sarker MR, McClane BA. *Clostridium perfringens* Sporulation and Sporulation-Associated Toxin Production. Microbiol Spectr. 2016;4(3).

60. Mi E, Li J, McClane BA. NanR Regulates Sporulation and Enterotoxin Production by *Clostridium perfringens* Type F Strain F4969. Infect Immun. 2018 Oct;86(10).

61. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, et al. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of tcdR. PLoS ONE. 2013;8(12):e83748.

62. McKee RW, Mangalea MR, Purcell EB, Borchardt EK, Tamayo R. The second messenger cyclic Di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD. J Bacteriol. 2013 Nov;195(22):5174–85.

2.7 Supplemental data

2.7.1 Supplementary methods

2.7.1.1 Sporulation assay

Sporulation assay were performed as described previously (15). Briefly, *C. difficile* strains were grown on 70:30 sporulation agar. After 30 h of growth, cells were scraped from the plates and suspended in 70:30 sporulation liquid medium to an OD_{600} of 1.0. To enumerate viable vegetative cells and spores, cells were serially diluted and plated onto TY agar with 0.1% taurocholate and incubated at 37◦C for 24 to 48 hours. To enumerate the number of viable spores only, 500µl of the samples from each culture were mixed 1:1 with 95% ethanol and incubated for 1 hour at room temperature to kill all the vegetative cells. The ethanol-treated samples were then serially diluted, plated on TY agar with 0.1% taurocholate and incubated at 37◦C for 24 to 48 hours. The percentage of ethanol-resistant spores were calculated by dividing the number of CFU from spores by the total number of CFU and multiplying the value by 100. The results were based on a minimum of three biological replicates.

2.7.1.2 Quantitative Reverse Rranscription PCR (qRT-PCR)

Clostridioides difficile cultures were grown in TY medium for 10h and total RNA was extracted following a previously described protocol (15). Briefly, total RNA was treated with DNAse (Turbo; Ambion) for 2 h at 37◦C. After DNase treatment, a 25 µL reaction was set up with 1 µg of template RNA and was heated at 70 °C for 10 min. The same reaction was made to 50 µl by adding 10 µL of 5X reverse transcription (RT) reaction buffer, 1mM of dNTP; 1 µg of hexamer oligonucleotide primer (5 µg/µL pdN_6; Roche), and avian myeloblastosis virus (AMV) reverse transcriptase (Promega) to synthesize cDNA at 42 °C for 2 hours. Real-time quantitative PCR was

performed in iQPCR real-time PCR instrument (BioRad) by setting up 20 µL reaction volume containing 10 ng of cDNA, 400 nM gene-specific primers, and 10 µL of SYBR PCR master mix (BioRad).Quantity of cDNA of a gene in each sample was normalized to the quantity of *C. difficile* 16S rRNA gene and the ratio of normalized target concentrations (threshold cycle [$2^{-\Delta\Delta Ct}$] method) gives the relative change in gene expression. A minimum of three biological replicates were used per sample.

2.7.1.3 Toxin ELISA

Cytosolic toxins from 12h old *C. difficile* cultures grown in TY medium were measured as described previously . In brief, one ml of *C. difficile* cultures were harvested and suspended in 200 µl of sterile PBS, sonicated and centrifuged to harvest the cytosolic protein. One hundred µg of cytosolic proteins was used to measure the relative toxin levels using *C. difficile* premier Toxin A &B ELISA kit from Meridian Diagnostics Inc. (Cincinnati, OH).

2.7.1.4 Biofilm assay

Biofilm assay was performed as described by Poquet et al. with minor modifications. Briefly, *C. difficile* strains grown in TY medium for 16h and were diluted two-fold in fresh TY. Two ml of the diluted over-night cultures were added to 12-well micro-titer plates and was incubated for 3 hr at 37°C. After carefully removing the liquid, fresh TY medium was added onto the adhesive cells and were incubated at 37°C. After 24 hr, the supernatant was removed from the wells, being careful not to touch the bottom of the plate. The biofilms were then washed once with phosphate-buffered saline (PBS), stained for 30 min in 0.1% filtered crystal violet, and washed

twice with PBS. To quantify the biofilm formation, 100% ethanol was used to solubilized the dye and quantified at 570 nm.

2.7.2 Supplementary tables

Table S 2.2. Oligonucleotides used for PCR amplification

Name	Sequence (5' → 3')	Description
EBS-U	CGAAATTAGAAACTTGCGTTAGT AAAC	Group II intron specific primer
ORG559	CTTCTTATTTTTATGGTACCGGTG CAATAACTCATGTTTTTAG	*spo0A* upstream + *spo0A*-Forward with *Kpn1*–pMTL84151 (pRG312) cloning
ORG560	GGGCATCGAAATAAAAAACTAGT GACTCTCATATTTAAACCTCCAC	*spo0A* reverse with *EcoR1*–pMTL84151 (pRG312 cloning)
ORG551	GTTAACAGATCTGAGCTCCTGTA ATAAGAAGATGTTTTTTAATGG	*spo0A*-Forward with SacI-pRPF185 (pRG301) cloning
ORG552	AAGTTTTATTAAAACTTATAGGAT CCTTATTTAACCATACTATGTTCT AGTC	*spo0A*-Reverse with BamH1-pRPF185 (pRG301) cloning
ORG749	GGTACCATTTACATAAATTGTTCA ATATATAAAATAGAAAA	*sin* LOCUS promoter (340 bps) to clone in pRPF185 (Forward with KpnI)
ORG750	GAGCTCAATTATTATCCCTCCACT TTAGATTATATTCAT	*sin* LOCUS promoter (340 bps) to cline in PRPF185 (Reverse with SacI)
ORG755	GTTGTTTTATATATATTATTTATA CTTTTTTAGTCTAAATATATATTA TATATGTTTTTATACTGAAATAAG TGCTATTTATTTTTGTA	Spo0A RGI and RGII mutagenesis primer (forward)
ORG756	TACAAAAATAAATAGCACTTATT TCAGTATAAAAACATATATAATA TATATTTAGACTAAAAAAGTATA AATAATATATATAAAACAAC	Spo0A RGI and RGII mutagenesis primer (reverse)
ORG809	5Biosg/TTTGGAAGCATTGATAGAA TAAATAAAATCAATATAGAG	Forward 350 pb upstream GluD (GDH) R20291_5' biotin
ORG810	5Biosg/ TTTGAAAGCCCCCTTATAAATAC GTTATAATTATGTATACTCC	Reverse 350 pb upstream GluD (GDH) R20291_5' biotin
ORG811	5Biosg/ATTTACATAAATTGTTCAA TATATAAAATAGAAAA	Forward 340 bp upstream *sinR* 5' biotin labelled
ORG812	5Biosg/AATTATTATCCCTCCACTT TAGATTATATTCAT	Reverse 340 bp upstream *sinR* 5' biotin labelled
ORG825	GGTACCGTAGGTAATTATATAGT AAAAATGTATGTTGGTTGGTGGA AGAAACATGG	SpoIIAB upstream Forward with KpnI
ORG826	GAGCTCCAAAAATCCCTCCTTCA ATAGTTTTGTAAAAATAGTAAC	SpoIIAB upstream Reverse with SacI
ORG 827	AGAAATATCTTCATAATTTCATA GCCATATTATAATAATAAAACGA ATTAATAATATGTTAATTATAGCA TTTTTTATTGTAATTACA	340 bp upstream sinR RG_I mutagenesis Forward
ORG 828	TGTAATTACAATAAAAAATGCTA TAATTAACATATTATTAATTCGTT TTATTATTATAATATGGCTATGAA ATTATGAAGATATTTCT	340 bp upstream sinR RG_I mutagenesis Reverse

ORG 829	GTGCTATTTATTTTTGTATAGTCT GTATTTATAATAAACAAATATGG ATAAATCATATTTTTAATTAGAGA AATATCTTCATAAT	340 bp upstream sinR RG_II mutagenesis Forward
ORG 830	ATTATGAAGATATTTCTCTAATTA AAAATATGATTTATCCATATTTGT TTATTATAAATACAGACTATACA AAAATAAATAGCAC	340 bp upstream sinR RG_II mutagenesis Reverse
ORG 850	5/Biosg/GTAGGTAATTATATAGTA AAAATGTATGTTGGTTGGTGGAA GAAACATGG	SpoIIAB promoter_5' biotinylated _Forward
ORG 851	5/Biosg/CAAAAATCCCTCCTTCAA TAGTTTTGTAAAAATAGTAAC	SpoIIAB promoter_5' biotinylated _Reverse
ORG 873	5/Biosg/GCACTTATTTCAGTATAA AAACATATATAGTCTAT	SinR Forward @ last 110 bps upstream
ORG 872	5/Biosg/CTGAAATAAGTGCTATTT ATTTTTGTATAGTCTGTATC	SinR promoter Reverse @ middle 140 bps
ORG 871	5/Biosg/CATATTATTAATTCGTTTT ATTATTGTAGTATGGC	SinR promoter Forward @ middle 140 bps
ORG 870	5/Biosg/CGAATTAATAATATGTTA ATTATAGCAT	SinR promoter Reverse @ first 100 bps
ORG 889	TTTAGGTACCTTAAATTATTTTAT AAGATTATTACTCTACTATAAATC TTGTATATAACT	SinR upstream 610 bps Forward with KpnI
ORG 890	TTTAGGTACCTTCTAAATGCCTTA CTTATAATTAATTTTTTATTTCAC CTATATATAATT	SinR upstream 475 bps Forward with KpnI

Table S 2.3. Oligonucleotides used for qRT-PCR amplification

Primer	Sequence (5' → 3')	Gene target
RG-RT23 (F)	GAGGAGAGTGGAATTCCTAGTGTAG	*16srRNA*
RG-RT24 (R)	GGACTACCAGGGTATCTAATCCTGT	*16srRNA*
RG-RT25 (F)	AGGCAGGTTTACATCCAACATA	*sinR*
RG-RT26 (R)	AGTGGTATGTCTAAAGCAGTAGC	*sinR*
RG-RT27 (F)	AAAGACTTAAAGAAGAACGGAAAA	*sinR'*
RG-RT28 (R)	TTGGATTCTTTTTACCACTTTCG	*sinR'*
RG-RT1 (F)	CAAGAAATAACTCAGTAGATGATTTGCAA	*tcdR*
RG-RT2 (R)	TCTCCCTCTTCATAATGTAAAACTCTACTA	*tcdR*
sigD-RT(F)	TGATAGAGAAGAGGAAGCTCCA	*sigD*
sigD-RT(R)	TCTGAAACACCTAGCACTTTTCC	*sigD*

2.7.3 Supplemental figures

Figure S 2.1. Construction and confirmation of the spo0A mutant in *C. difficile* JIR8094 and UK1 strain

(A) Schematic representation of ClostTron (group II intron) mediated insertional inactivation of *spo0A* gene in *C. difficile*. **(B)** PCR verification of the intron insertion and complementation of *spo0A* in JIR8094 with intron-specific primer EBS universal [EBS(U)] with *spo0A* specific primers ORG 551 and ORG 552. **(C)** PCR verification of the intron insertion in UK1 strain with EBS(U), ORG 551 and ORG 552. **(D)** Western blot analysis of Spo0A and SinR production in UK1 and UK1::*spo0A*. GDH was used as loading control.

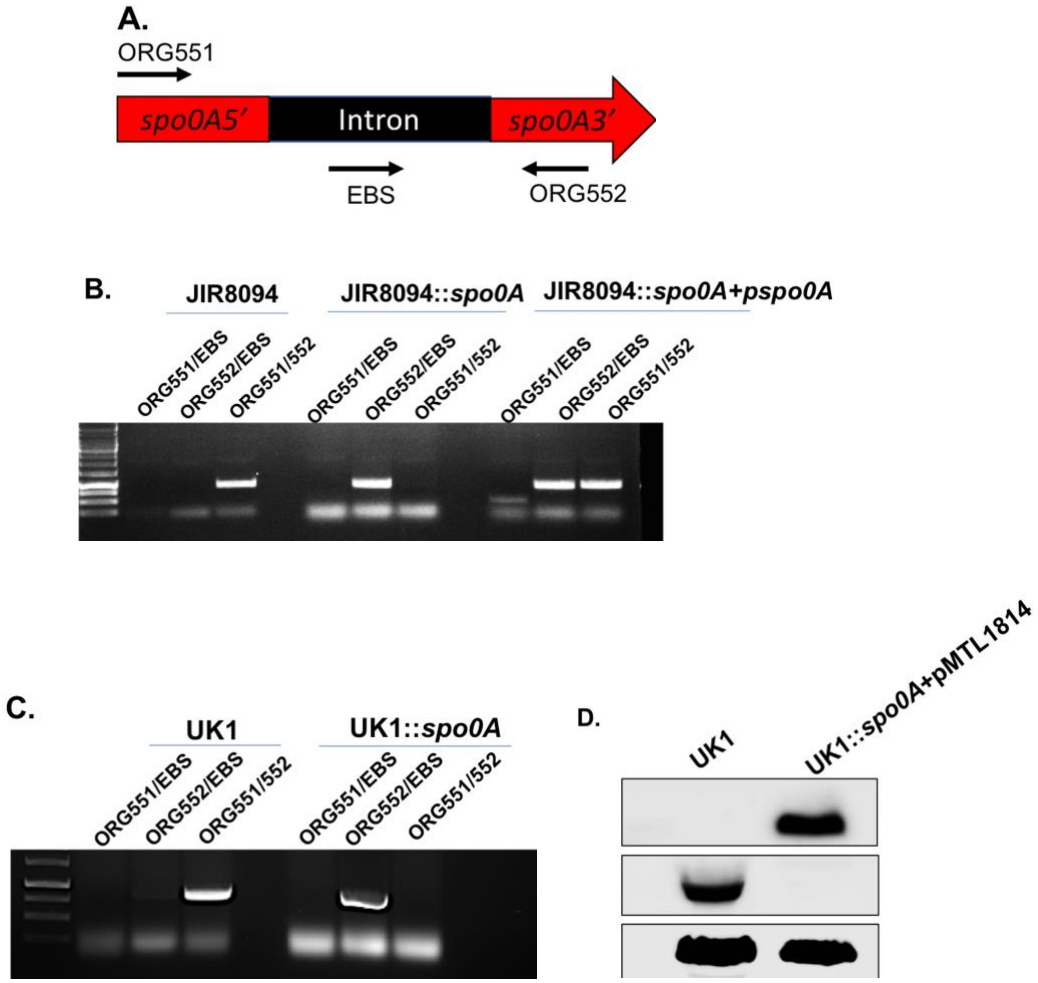


Figure S 2.2. Sporulation in spo0A mutants

Percentage sporulation (the CFU/ml from ethanol resistant spores) of the parent and *spo0A* mutants. The representative results from three independent experiments are shown.

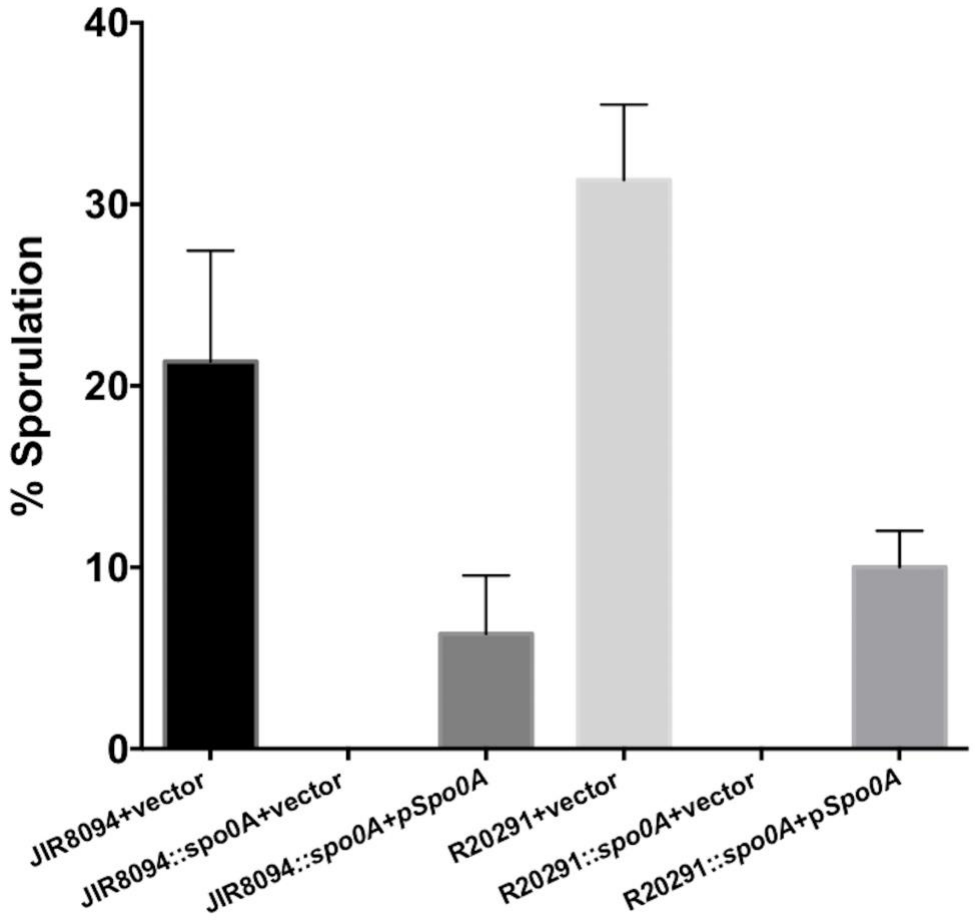


Figure S 2.3. Toxin gene transcription and biofilm formation in R20291::spo0A mutant strain

A) Toxin production measured by ELISA. Statistical analysis was performed using student t-test. (**; <0.05 *P* value). **(B)** Crystal violet stained biofilm in the 12 wells tissue culture plate, showing poor biofilm formation in R20291::*spo0A* and in UK1::*spo0A* mutants.(**C**) Quantification of crystal violet dye attached to the cells forming biofilms.

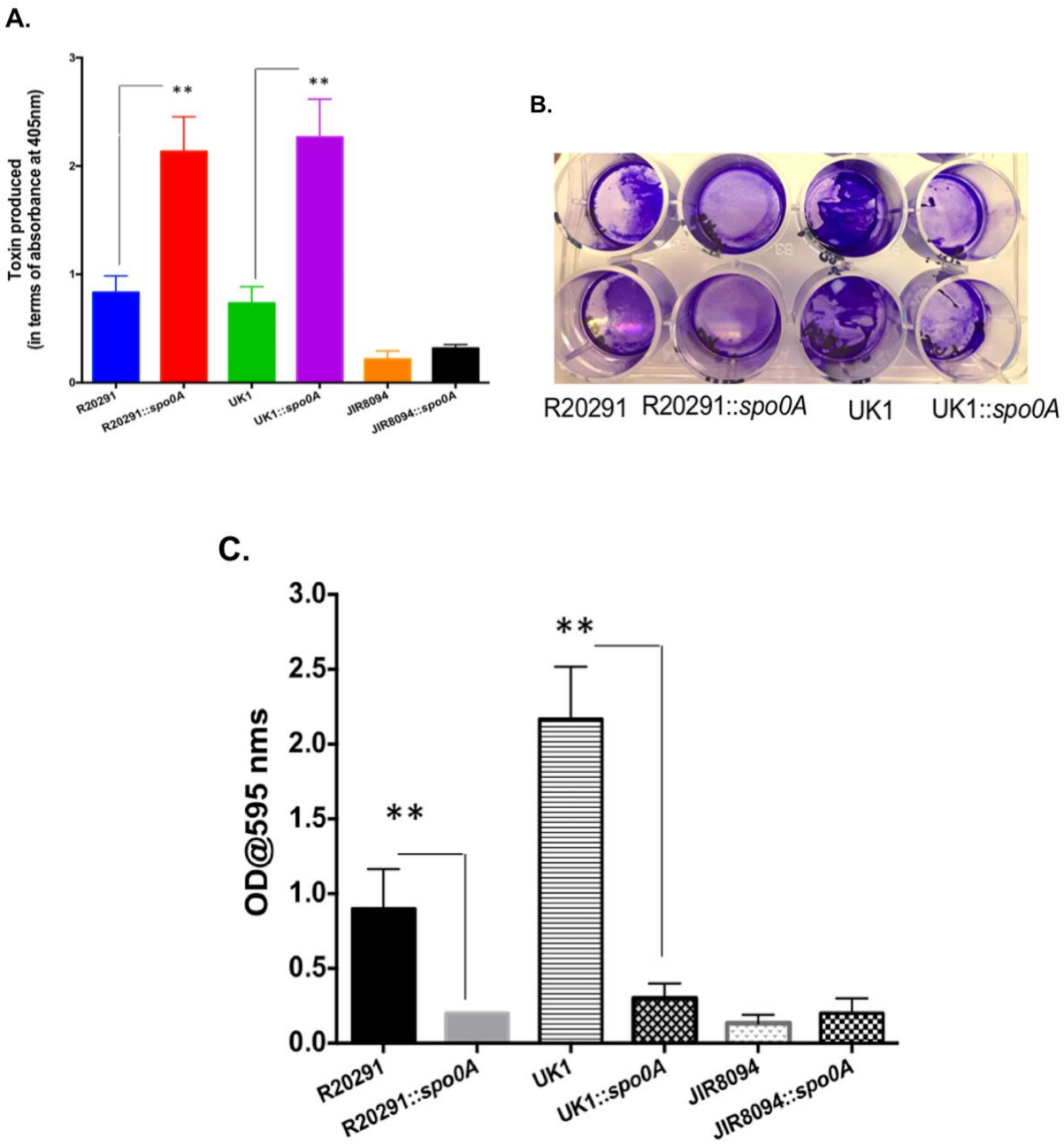
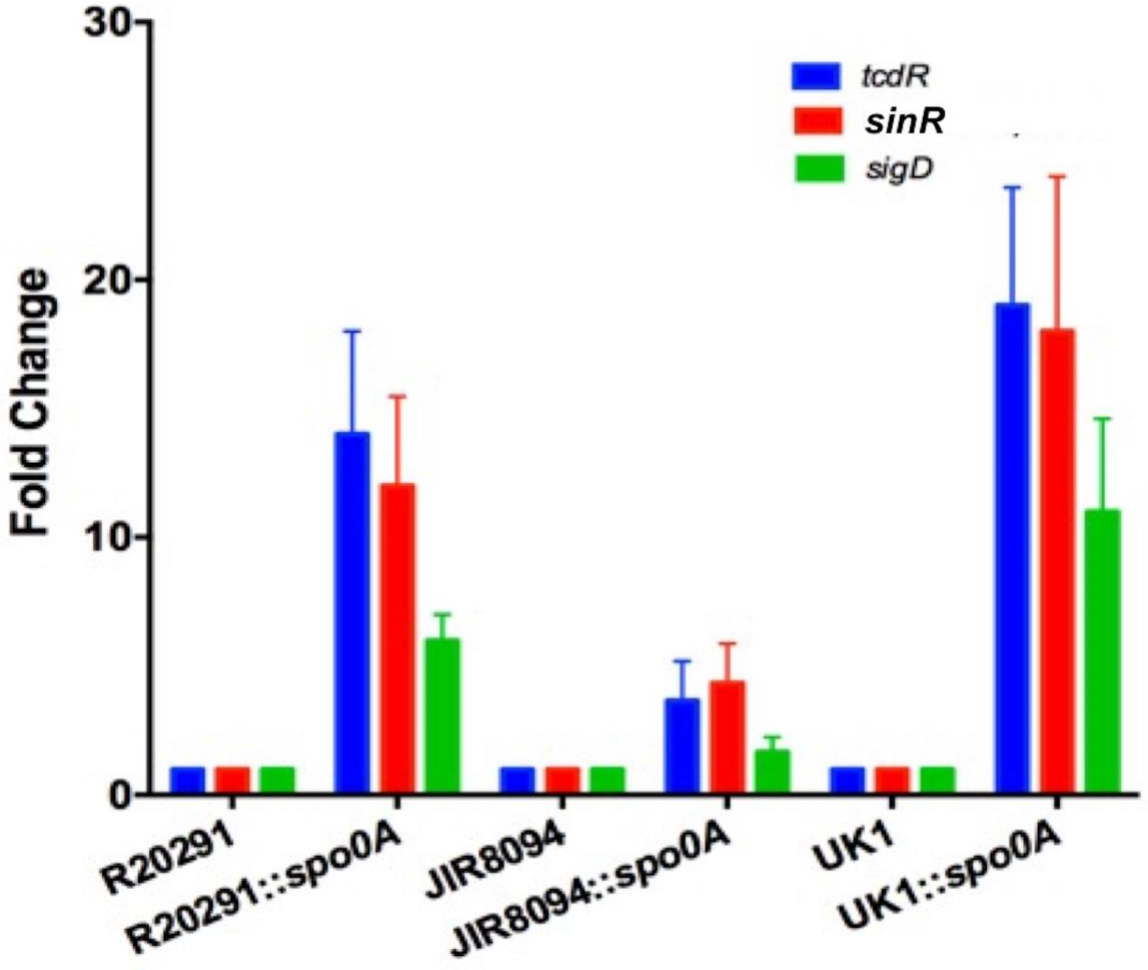


Figure S 2.4. qRT-PCR analysis

Relative expression of the transcripts of *tcdR, sigD, and sinR* genes from *C. difficile* parent and *spo0A* mutant strains in JIR8094, R20291 *and* UK1 background. RNA was collected at 10 hr time point.



Chapter 3 - Phase variable expression of *pdcB* and the role of intracellular c-di-GMP in *Clostridioides difficile* sporulation.

3.1 Abstract

Clostridioides difficile is the causative agent of antibiotic-associated diarrhea and is the leading cause of nosocomial infection in developed countries. Increasing number of *C. difficile* infection is attributed to the emergence of hypervirulence strains which produce more toxins and spores. *C. difficile spores* are the major factor for transmission and persistence of the organism. Previous studies have identified global regulators that influence sporulation in *C. difficile*. In this study, we identified PdcB as a novel regulator that positively regulates sporulation in *C. difficile* UK1 strain. Through genetic and biochemical assays, we have shown that phase variable expression of *pdcB* results in hypo- and hyper-sporulation phenotype. By showing that the "ON" orientation of the phase variable switch of *pdcB* in hyper-sporulating strain reduces the intracellular c-di-GMP concentration, we have shown a novel link between c-di-GMP and sporulation in *C. difficile*. Additionally, we demonstrated that CodY binds to the upstream region of *pdcB* and represses its expression and CodY mediated repression is partially relieved by DNA inversion. This study provides a novel pathway of c-di-GMP mediated regulation of sporulation in *C. difficile.*

3.2 Introduction

Phase variation is a phenomenon where two distinct phenotypes are established within a clonal bacterial population in a reversible manner. This phenomenon in bacteria has been long recognized and has been well studied in various bacterial pathogens. Heterogeneity within a bacterial population is readily visible as colony variation and is often associated with pathogenicity (1–3). Genes that are regulated by phase variation are often involved in pathogenesis such as those that encode fimbriae, flagella, cell surface protein, capsular polysaccharides, components that are responsible for establishing an interaction with the host which gives the bacteria an advantage against the challenging environment of host like colonization, establish an infection or evading from host immune response (2,4). Phase variation has been reported in several Gram-negative bacteria pathogens including *Salmonella* (1,5–10), *E.coli* (11–14), *Klebsiella* (15), *Neisseria* (16,17) and *Campylobacter jejuni* (18,19) where it has been employed to evade host immune system, increase adhesion, colonization and virulence. While there are ample reports on phase variable regions and their role in pathogenesis of Gram-negative bacteria, little is known about the role of phase variation in Gram-positive pathogens. *Streptococcus pneumoniae* utilize phase variation to switch colony morphology and to express pilus (4). Phase variation has been reported more recently in *Clostridioides difficile*, where two distinct colony morphologies: smooth and circular and a rough filamentous were observed and were found to be associated with motility and distinct cell shapes (20). In this study, we initiated investigation to determine the molecular mechanisms involved in the phase variable sporulating and non-sporulating colonies observed in *C. difficile* UK1 strain.

Clostridioides difficile is a Gram-positive and anaerobic pathogen that lives in the mammalian GI tract (21). *Clostridioides difficile's* growth in the GI tract is kept on check by the commensal gut microbiota. But during antibiotic treatment, a healthy bacterial population of the gut are disrupted thereby creating an environment that is favorable for *C. difficile* to proliferate and cause infection (21). *Clostridioides difficile* infection is caused primarily by the toxin produced by the bacteria which causes inflammation of gut epithelium leading to severe diarrhea, commonly known as antibiotic-associated diarrhea, and ultimately causing fetal pseudomembranous colitis (22,23). Each year >200,000 *C. difficile* infection (CDI) and >12,000 deaths due to CDI are reported in the United States (24). Due to the increased number of infections and a massive cost of health care of ~ 4.2 billion dollars, CDI has been categorized by the CDC as an urgent threat to public health (25).

The increasing number of *C. difficile* infections has been attributed to the emergence of hypervirulent *C. difficile* strains, which produce more toxins and spores (26). *Clostridioides difficile* spores are the major cause of transmission and persistence and are the major factor for making it the number one cause of hospital-acquired infection in North America. However, many aspects of *C. difficile* sporulation and its regulation remain poorly understood. *Clostridioides difficile* produces metabolically dormant spores that are resistant to oxygen, heat, disinfectants, antibiotics and gastric acid of the mammalian stomach, allowing the bacteria to survive the unfavorable environment both inside and outside the host (21,27,28). These spores enable the bacteria to be persistent and thus are difficult to eradicate. Spores act as a vehicle for transmitting the disease through fecal oral route in the health care setting. Spores are ingested and make their way to anaerobic mammalian intestine where they germinate with the help of bile salts into active vegetative cells and secrete cytotoxins that cause tissue damage and inflammation (27). So, it is

important to understand the regulation of sporulation in *C. difficile*. In this study, by characterizing the phase variation involving sporulating phenotypes, we have uncovered new link between intracellular c-di-GMP concentrations and sporulation.

C-di-GMP is a common bacterial second messenger found in many bacterial species. Two classes of enzymes are important for c-di-GMP homeostasis in a bacterial cell, diguanylate cyclases (DGCs) that synthesize c-di-GMP from 2 GTP molecules and phosphodiesterases (PDEs) that degrade c-di-GMP. C-di-GMP has been shown to regulate a variety of physiological processes in many Gram-negative bacteria like changes in cellular envelope and the transition from planktonic growth to biofilm formation (29). c-di-GMP also co-ordinates entry into stationary phase, expression of virulence genes, resistance to antibiotics, and host immune responses and cell development shifts. In *C. difficile*, c-di-GMP has been shown to negatively regulate toxin production (29–33) and flagella mediated swimming motility (29,31,32,34), whereas positively regulate *pilA* mediated swarming motility (29,31,32,35–37), biofilm formation (30,31), cell aggregation (34,37) and expression of cell wall anchor structures (38,39). While c-di-GMP regulates a plethora of physiological processes, its role in sporulation has been poorly understood. In this study, we discovered the novel link between the reduced intracellular level of c-di-GMP and the hyper sporulation phenotype in *C. difficile* UK1 strain. Reduction in intracellular c-di-GMP level was brought about by the overexpression of *pdcB*, which encodes for a predicted phosphodiesterase. We also show that *pdcB* mutant is hypo sporulating and overexpression of the PdcB-EAL domain complements the sporulation phenotype. Our data shows that *pdcB* is a mid-exponential phase gene that is negatively regulated by CodY. When the *pdcB* gene is suppressed by CodY, it results in a non-sporulating colony. DNA inversion in the upstream of *pdcB* gene

partially relieves the repression of CodY and expression of *pdcB* results in sporulation initiation in *C. difficile* UK1 strain.

3.3 Material and methods

3.3.1 Bacterial strains and growth conditions

Clostridioides difficile UK1 strains (Table S 3.1) were grown in rich TY (Tryptose and Yeast extract) medium, agar or broth culture, in an anaerobic chamber which is maintained at 10% H_2, 10% CO_2 and 80% N_2. Transconjugants were selected and grown in TY agar with lincomycin (Linc 20µg/ml) or thiamphenicol (Thio; 15 µg/ml) or both. An *E.coli* strain optimized for conjugation S17-1 was used for conjugation (40) and *E.coli* DH5α strain was used for cloning . *E.coli* strains were cultured aerobically in LB (Luria-Bertani) broth or agar and was supplemented with ampicillin (100 µg/ml) or chloramphenicol (25 µg/ml) when necessary.

3.3.2 Phase-contrast microscopy

Clostridioides difficile strains were grown in TY medium as described above. 1 ml of overnight culture was centrifuged at 17,000g for 1 min and washed with 30µl of sterile PBS. The resulting pellets ware fixed for 2 hours in room temperature using 2% paraformaldehyde in 1X PBS. A thin layer of 0.7% agarose was applied to the surface of a microscopic slide and 2µl of concentrated culture was placed on it. The cells were imaged using Zeiss LSM-5 PASCAL (objective lens 100x/1.4 oil). The LSM 5 service pack was used to acquire the images of at least three fields for each strain.

Table S 3.1. Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Relevant features or genotype	Reference
Clostridioides difficile UK1	Clinical isolate	(41)
Escherichia coli DH5α	*endA1 recA1 deoR hsdR17 (r_{K-} m_{K+})*	NEB
Escherichia coli S17-1	Strain with integrated RP4 conjugation transfer function; favors conjugation between *E. coli* and *C. difficile*	(42)
Clostridioides difficile UK1::pdcB	UK1 with intron insertion within *pdcB* (*CDR20291_0685*)	This study
pMTL007-CE5	ClosTron plasmid	(43)
pMTL007-CE5:Cdi-*pdcB*-840-841s	pMTL007-CE5 with group II intron targeted to *pdcB*	This study
pMTL007-CE5:Cdi-*recV*-144a	pMTL007-CE5 with group II intron targeted to *recV*	This study
pRPF185	*E. coli*/ *C. difficile* shuttle plasmid	(44)
pBA042	pRPF185 containing 1.5 kb upstream region of *pdcB* in Translucent orientation.	This study
pBA043	pRPF185 containing ~1.5 kb *pdcB* upstream region locked in Translucent orientation	This study
pBA045	pRPF185 containing 1.5upstream region of *pdcB* in Opaque orientation.	This study
pBA046	pRPF185 containing ~1.5 kb *pdcB* upstream region locked in Opaque orientation	This study
pBA048	pRPF185 containing PdcB-EAL domain	This study
pBA050	pRPF185 containing PdcA-EAL domain	This study
pBA051	pRPF185 containing the *gusA less tet* promoter	This study

3.3.3 Sporulation assay

Sporulation assays were performed in TY medium as described previously (45). To induce germination of any spores that were present initially, *C. difficile* cultures were grown overnight in TY broth supplemented with 0.1% taurocholate. Cells were then diluted in TY medium to an OD600 of 0.5, and 100 µl the OD600 adjusted culture was spread on TY agar. Plates were incubated at 37°C and after 24 hours cells were harvested from the plates and looked under the microscope. The total number of cells and the total number of sporulating cells of at least 10 fields per strain were counted and percentage sporulation was calculated by dividing the average number of sporulating cells by the average number of total cells, multiplied by 100. Experiments were performed at least three independent times.

3.3.4 Toxin ELISA

Clostridioides difficile cytosolic toxins were measured as described previously (46,47). In brief, one ml of *C. difficile* 16 hour cultures were harvested. The cells were centrifuged at 17,000 g for 1 min and washed with sterile 1X PBS. The resulting pellets were resuspended in 200 µl of sterile 1X PBS, sonicated and centrifuged to harvest the cytosolic protein. 100µg of cytosolic proteins was used to measure the relative toxin levels using *C. difficile* premier Toxin A &B ELISA kit from Meridian Diagnostics Inc. (Cincinnati, OH).

3.3.5 Motility assay

Clostridioides difficile cultures were grown until mid-exponential phase in TY broth at 37◦C. After adjusting their OD600 to 0.5, 10µl of each strain was spotted in TY media plates with 0.5% w/v agar for swimming motility and 1.8% w/v agar for swarming motility. The plates were

incubated at 37◦C for 72 hours and the motility was quantified by measuring the diameter of the spotted cultures. Since the spotted cultures were not a perfect circle, multiple diameters were taken for a single spotted culture and calculated the average. Motility assay was independently repeated at least three times.

3.3.6 Biofilm assay

Biofilm assay was carried out as described previously (30) with slight modification. *Clostridioides difficile* cultures were grown until mid-exponential phase in TY broth at 37◦C. After adjusting their OD600 to 0.5, cultures were seeded at 1:100 dilution to an untreated 6-tissue culture plates containing 2 ml of TY medium. The plates were incubated at 37◦C in an anaerobic chamber for 72 hours. The culture medium was carefully removed and the adherent biofilm was washed once with 1XPBS. The biofilm was stained for 30 min at room temperature with 0.1% (wt/vol) crystal violet in water. Excess crystal violet was carefully removed, and the biofilm were washed twice with 1XPBS. Quantification of biofilm was done by solubilizing the biofilm bound crystal violet with 95% ethanol and measuring the absorbance at 570 nm. Experiments were done at least three times.

3.3.7 Genomic DNA extraction

Genomic DNA was extracted as previously (48). Briefly, *C. difficile* cultures were allowed to grow overnight in a 37°C anaerobic chamber. 10 ml of the overnight culture was harvested by centrifuging at 3,000g for 30 min at 4°C. Cell pellets were washed with 1 ml of TE (Tris-EDTA) buffer. After centrifugation, cell pellets were resuspended in 600 µl of genomic DNA solution. 60 µl of 50 mg/ml of lysozyme solution was added and the mixture was allowed to incubate for 2

hours at 37°C. 100 µl of 20% Sarkosyl, 15 µl of RNAase A (10mg/ml) and 15 µl of proteinase K (10 mg/ml) were added and incubated for additional 30 min at 37°C. 600 µl of 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol was added and mixed by vortexing. The mixture was centrifuged at maximum speed for 10 min. The upper aqueous phase was transferred into a new clean tube. The process was repeated 3 times to collect as much of the upper aqueous phage. 200 µl of chloroform was added to the collected upper aqueous phage, mixed and centrifuged at maximum speed for 10 min. The resulting upper phase was transferred into a clean tube. DNA precipitation was carried out by adding 50 µl of 3M sodium acetate, pH 5.2 and 150 µl of 95% ethanol. The mixture was stored overnight at -20°C . DNA pellet was recovered by centrifuging at maximum speed for 5 min. The pellet was washed with 500 µl of 70% ethanol. The ethanol was allowed to dry at room temperature and DNA was dissolved in 50 µl of Nuclease Free Water (NFW).

3.3.8 General DNA techniques

Chromosomal DNA was extracted from *C. difficile* cultures using DNeasy Blood and Tissue Kit (Qiagen). PCR reactions were carried out using gene specific primers (Table S 3.2). PCR products were extracted from the gel using Geneclean Kit (mpbio). Plasmid DNA was extracted using QIAprep Spin Miniprep Kit (Qiagen). Standard procedures were used to perform routine cloning.

3.3.9 RNA isolation and quantitative Reverse Rranscription PCR (qRT-PCR)

Clostridioides difficile cultures were grown in TY medium and 6 ml of cells were harvested at different time points by centrifugation at 3000g for 30 min at 4 °C. Total RNA was extracted

from the harvested cells following the protocol described previously (49). Total RNA was treated with DNase (Turbo; Ambion) for 2 hours at 37 °C. 30 µL of final reaction volume comprising of 5 µg of template RNA, 4 µL of deoxynucleoside triphosphates (dNTP; 10mM each), 1 µg of hexamer oligonucleotide primer (5 µg/µL pdN_6; Roche), and 6 µL of reverse transcription (RT) buffer was heated at 80 °C for 5 min and cDNA was synthesized at 42 °C for 2 hours using avian myeloblastosis virus (AMV) reverse transcriptase (Promega). 20 µL reaction volume containing 10 ng or 10 pg (for 16S rRNA) of cDNA, 400 nM gene-specific primers, and 12.75 µL of SYBR PCR master mix (BioRad) was used to perform Real-time quantitative PCR using iQPCR real-time PCR instrument (BioRad). Amplification and detection was performed as described previously (49). Quantity of cDNA of a gene in each sample was normalized to the quantity of *C. difficile* 16S rRNA gene and the ratio of normalized target concentrations (threshold cycle [$2^{-\Delta\Delta Ct}$] method) (49,50) gives the relative change in gene expression.

3.3.10 Mutagenesis of *CDR20291_0685* upstream region

Quick Change Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) was used to carry out site directed mutagenesis whereby the Left Inverted Repeat (LIR) TAGTTGTAAAAGGGTT that flanked the DNA region that undergoes inversion was converted to TACACATGCGAGGGTT. The mutagenic oligonucleotide primers used are listed in Table S 3.2.

3.3.11 Construction of reporter plasmids and beta-glucuronidase assay

The 1.5 kb upstream DNA regions of *CDR20291_0685* was amplified by PCR using specific primers with *KpnI* and *SacI* (Table S 3.2) recognition sequences. UK1_T and UK1_O

chromosomal DNA was used as a template to amplify the upstream region in respective orientations. The upstream region was cloned in pRPF185 shuttle vector using standard cloning procedures. Plasmid pRPF185 carries a *gusA* gene for beta-glucuronidase under tetracycline-inducible (*tet*) promoter (44). The *tet* promoter was removed using *KpnI* and *SacI* digestion and was replaced with the *CDR20291_0685* upstream region to create plasmids pBA042 and pBA045 (Table S 3.1). The inverted repeats in the upstream region were mutagenized as described earlier to create plasmid pBA043 containing the upstream region locked in translucent orientation and plasmid pBA046 containing the upstream region locked in opaque orientation. The control plasmid pBA040 with promoter less *gusA* was created by digesting with *KpnI* and *SacI* to remove *tet* promoter and then self-ligated after creating blunt ends. Plasmids were introduced into *C. difficile* strains through conjugation. The transconjugants were grown overnight in TY medium in the presence of thiamphenicol (15 µg/ml). Overnight cultures were used as an inoculum at a 1:100 dilution to start a new culture. Bacterial cultures were harvested at different time points of growth and the amount of beta-glucuronidase activity was assessed as described elsewhere (51). Briefly, the cells were washed and resuspended in 1 ml of Z buffer (60 mM $Na_2HPO_4.7H_20$ pH 7.0, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, 1mM $MgSO_4.7H_2O$ and 50mM 2ME) and lysed by homogenization. The lysate was mixed with 160 µl of 6mM p-nitrophenyl β-D-glucuronide (Sigma) and incubated at 37°C. The reaction was stopped by the addition of 0.4 ml of 1.0 M $NaCO_3$. β-Glucuronidase activity was calculated as descried earlier (51,52).

3.3.12 Electrophoretic mobility shift assay (EMSAs)

For the CodY binding experiment, upstream region of the pdcB with the predicted CodY binding sequence 5’

CATAGATAATTTTTAGAAAAGTATCTAAATTTTCAATAAATAGTAAC 3' was synthesized and was labeled with [γ- 32 P]dATP-6000 Ci/mmol (PerkinElmer Life Sciences) using T4 polynucleotide kinase. It was then annealed with the complementary oligo to generate double-stranded DNA probe. Known CodY binding sequence upstream of the tcdR gene was similarly synthesized (Table S 3.2) and used as a positive control. A non-specific double-stranded DNA was used as negative control (Table S 3.2) . The DNA-protein binding reactions were carried out at room temperature for 30 min in 10µl volume containing 1x binding buffer [10mM Tris pH 7.5, 50mM KCl, 50µg BSA, 0.05% NP40, 10% Glycerol, 10 mM GTP and 2mM ILV (Isoleucine, Leucine and Valine), 100 µg/ml poly dI-dC and 800nM of DNA probe with varying concentration of purified CodY protein. DNA probe in reaction buffer was incubated for 10 min at RT before adding purified CodY-6His protein. The reaction was stopped by adding 5µl of gel loading buffer and electrophoresed at 100V for 1.5 h using 6% 1XTBE gel in 0.5X TBE buffer containing 10 mM ILV. Gels were then dried, and the autoradiography was performed with Molecular Dynamics Phosphor-Imager technology.

3.3.13 Reverse Transcriptase PCR (RT-PCR)

Total RNA was purified from *C. difficile* UK1_O strain cells collected at 4 hr time point as described above for qRT-PCR. DNAase I treated RNA was used as a template to generate *pdcB* specific cDNA using the *pdcB* specific oligo ORG846 and Promega-AMV reverse transcriptase following the manufacturer's instruction. This cDNA was used as a template for the subsequent PCR amplification using primers ORG925, ORG926, ORG921, ORG922 and ORG 853.

3.3.14 Construction and complementation of *C. difficile* UK1::*pdcB* mutant strains

The UK1::*pdcB* mutant in UK1 strain was created using ClosTron gene knockout system as described previously (53). Briefly, for *pdcB* disruption, the group II intron insertion site between nucleotides 840 and 841 in *pdcB* gene in the sense orientation was selected using a web-based design tool called the Perutka algorithm. The designed retargeted intron was cloned into pMTL007-CE5 as described previously (47,54). The resulting plasmid pMTL007-CE5::*pdcB*-840-841s was transferred into *C. difficile* UK1 cells by conjugation. The potential Ll.ltrB insertions within the target genes in the *C. difficile* chromosome was conferred by the selection of lincomycin resistant transconjugants in 20 µg/ ml lincomycin plates. PCR using gene-specific primers (Table S 3.2) in combination with the EBS-U universal was performed to identify putative *C. difficile* mutants. *Clostridioides difficile pdcB* mutants were complemented by introducing pBA048 which contains the PdcB-EAL domain and pBA050 which contains the PdcA-EAL domain under inducible promoter, through conjugation.

3.3.15 Transmission Electron Microscopy (TEM) and negative staining

Clostridioides difficile UK1_T and UK1_O strains were grown for 16 hr the anaerobic chamber. Pellets from 100 µl of the culture were fixed overnight with 2% glutaraldehyde and 2% paraformaldehyde made in 1X PBS buffer. The cells were washed with 1X PBS. Cells were negatively stained with 2% uranyl acetate and observed under the transmission electron microscopy. Images were analyzed for cell length using ImageJ.

3.3.16 c-di-GMP measurement

c-di-GMP was measured as described previously (34), with few modifications. Briefly, *C. difficile* strains were grown in 50 ml TY medium for 16 hours. OD_{600} was recorded and dilution plating was carried out to determine the number of CFU. Cells were centrifuged and pellets were washed once with TE buffer (10 mM Tris [pH 7.5], 1 mM EDTA, pH 8). Pellets were resuspended in extraction solvent consisting of methanol (MeOH):acetonitrile:distilled water (dH_2O) in 40:40:20 ratio plus 0.1 N formic acid. The mixture was incubated in -20°C for 30 min. The extract was harvested by centrifugation. Samples were sent to Kansas University Biochemistry Core facility to be analyzed by high-pressure liquid chromatography (HPLC)-tandem mass spectrometry (MS/MS). To determine the amount of intracellular c-di-GMP, a standard curve was first obtained by analyzing the serial dilution of pure c-di-GMP (Sigma-Aldrich). The peak intensity of each samples were fitted to the standard curve and the value was divided by the total intracellular volume of bacteria extracted. To determine the total intracellular volume of bacteria, volume of one cell was multiplied by the number of cells extracted, which was based on CFU counts. Volume of one cell was estimated as a cylinder and was determined by measuring the radius and length of the bacterial cell from the electron micrographs.

3.4 Results

3.4.1 *Clostridioides difficile* UK1 strain exhibits phenotypic heterogeneity with two distinguishable colony morphologies

Several studies have shown that *C. difficile* ribotype 027 which consists of clinically isolated strains associated with epidemic infections (R20291 and UK1), have phenotypic heterogeneity thereby exhibiting two distinct colony morphologies *in vitro* (20,55,56). Consistent

with these studies, we observed two different isolated colony types in UK1 strain during routine plating on rich TY (tryptophan and yeast extract) medium (Fig. 3.1A). One colony type is round, has circular and smooth edge with a pale yellowish pigmentation and is termed as an Opaque colony (UK1_O) while the other colony type is irregular with rough edge and is termed as Translucent colony (UK1_T). UK1_O and UK1_T colony types were also distinctly different in cell morphology when observed through phase contrast microscope (Fig. 3.1B). UK1_T cells were longer with an average length of 5.5 µm while the UK1_O cells are significantly shorter with an average length of 2.5 µm (Fig 3.1C). Isolated UK1_T and UK1_O colonies were streaked on TY agar plates and their morphologies observed over time. Opaque colonies could maintain the colony morphology even after 24 h while the translucent colony gave rise to Opaque colonies within 24 hours. Similar results were observed when they were grown in TY broth medium. Opaque colonies in the TY plate that were kept inside the anaerobic chamber for a prolonged time (>72 hrs) started exhibiting rough edges with pale yellowish center over time suggesting that opaque colonies could slowly revert back to translucent colony type. Previous studies have also shown that streaking a culture from the center of the undifferentiated R20291 strain resulted in smooth colony type while streaking from the edge resulted in rough colony types (20).

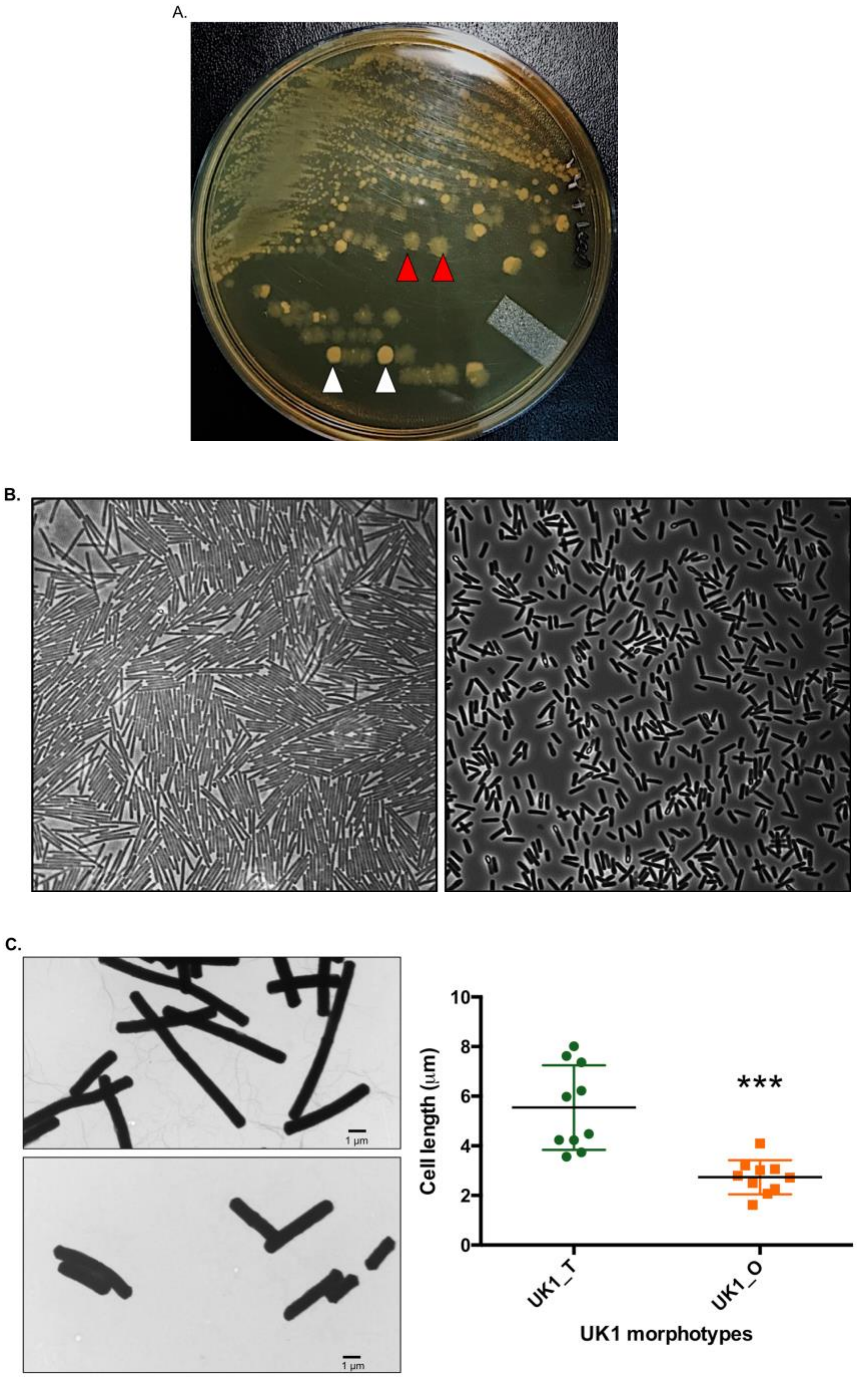


Figure 3.1. Colony and cell morphologies of UK1_T and UK1_O strain

A.UK1 strain when grown in TY medium gives rise to phenotypically distinguishable Translucent (UK1_T, shown in red) and Opaque (UK1_O, shown in white) colony morphotypes . **B.** Phage contrast microscopic images of UK1_T and UK1_O colony morphotypes. **C.** Transmission electron microscopy images showing UK1_T cells are significantly longer than UK1_O strains. Data analyzed using unpaired t-test with Welch's correction where *** indicates $P<0.0005$.

3.4.2 UK1_T and UK1_O colony morphologies exhibit distinct sporulation and swarming motility phenotype

In order to determine the physiological characteristics of these two colony types, we assayed for sporulation, toxin and motility of UK1_T and UK1_O colony types. To determine the sporulation phenotype, we streaked the isolated UK1_T and UK1_O colonies in solid TY medium and we found that UK1_O strain was hyper-sporulating and produced significantly more spores compared to UK1_T strain at 24 hours (Fig. 3.2A). We also found that UK1_T and UK1_O strains produce similar level of cytoplasmic toxins at 16 hr of growth (Fig. 3.2B). While UK1_T and UK1_O strains exhibited similar swimming motility even after 72 hours in 0.5% agar TY medium (Fig. 3.2C), we found that UK1_T strain exhibited enhanced swarming motility than the UK1_O strain when grown in 1.8% TY medium at 72 hr (Fig. 3.2D). Similarly, UK1_T strain demonstrated biofilm formation than UK1_O strain (Fig 3.2E). Together, these data demonstrate that UK1_T and UK1_O are phenotypically differ in sporulation, swarming motility and biofilm formation.

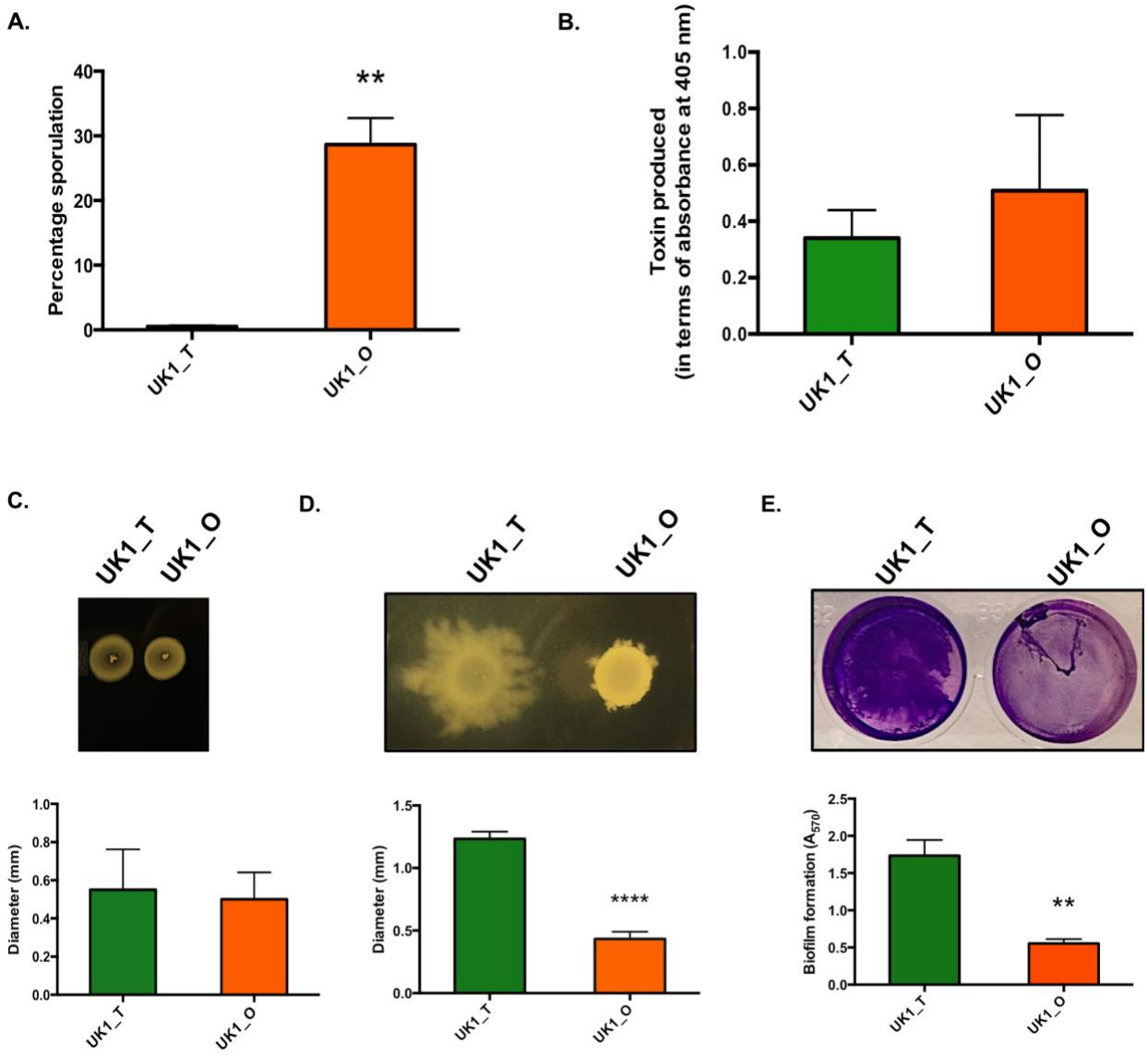


Figure 3.2. Phenotypic characterization of UK1_T and UK1_O strains

A. Percentage sporulation of UK1_T and UK1_O strains showing hyper sporulation phenotype in UK1_O morphotype. **B.** Cytoplasmic toxin levels produced by UK1_T and UK1_O morphotypes during stationary phase of growth. **C**. Swimming motility exhibited by UK1_T and UK1_O in TY+0.5% agar. **D.** Swarming motility exhibited by UK1_T and UK1_O in TY+1.8% agar. **E**. Biofilm formation of UK1_T and UK1_O morphotypes. OD_{560} nm was measured to quantify biofilm formation. Three biological replicates were used to carry out the experiments and data were analyzed using two tailed unpaired t test with Welch's correction where **** =P<0.0005. **=P<0.05

3.4.3 UK1_O strain undergoes phase variation by DNA inversion

To determine the genetic basis for the phenotypic heterogeneity that gave rise to translucent and opaque colonies in UK1 strain, we extracted genomic DNA from UK1_T and UK1_O colony morphotypes and performed whole genome sequencing at Tufts University Genomic core facility. Sequenced genome was assembled against the closely related and annotated R20291 genome (57). Unlike our expectation, the sequencing result did not show point mutations in any of the known sporulation associated genes. However, the intergenic region between *CDR20291_0684* and *CDR20291_0685* genes was listed among unassigned new junction evidence (Table 3.1).

In order to determine if there is any information in this intergenic region, we amplified and cloned ~1.5 kb intergenic region from UK1_T and UK1_O in a vector and sent 10 clones from each for sequencing. The sequences were first aligned to one another and analyzed with NCBI BLASTn. The results suggested a mismatch in ~200 bp segment of the intergenic region between UK1_T and UK1_O in all the 10 pair of clones (Fig 3.3A). The same BLASTn analysis also demonstrated that the mismatched segment would align perfectly if it were to be flipped and reverse complemented suggesting that the ~200 bp DNA segment of the intergenic region from UK1_O underwent DNA inversion (Fig. 3.3B). We further aligned the sequence of UK1_T and UK_O, independently, to the R20291 genome (57) and analyzed with NCBI BLASTn. The result suggested that all clones from UK1_T perfectly aligned with the published sequence of the R20291 while all the clones from UK1_O had a mismatch in ~200 bp segment of the intergenic region (Fig 3.3C).

The invertible DNA region is flanked by inverted repeats which are recognized by a group of enzyme called invertases that carry out the inversion. Invertases are specific recombinases and inversion is the result of site specific recombination of the inverted repeats (58). Consistent with

the previous study (2), the invertible region in the UK1_O intergenic region was flanked by the inverted repeat sequence GTTGTAAAAGGGTT-AACTTTTTTACAAC. Recent work on genome wide detection of site-specific recombination in *C. difficile* identified and characterized invertible regions Cdi1 to Cdi7 (2). Among these Cdi2 refers to the intergenic region *CDR20291_0684* and *CDR20291_0685*.

Site specific recombinases that carry out DNA inversion might be associated with the invertible site or encoded elsewhere in the genome (58). Prior work with *C. difficile* R20291 strain have shown that RecV, the tyrosine recombinase, catalyzes the inversion of Cdi1, Cdi4, Cdi6 switches, partial inversion of Cdi2, Cdi3, Cdi5 and has no effect on Cdi7 inversion (2). To determine if RecV has a role in inversion of Cdi2 in the UK1 strain we constructed a UK1::*recV* mutant using ClostTron mutagenesis (Fig S 3.1A and B). If RecV had a role in DNA inversion of Cdi2, we expected UK1::*recV* mutant to exist in both the colony UK1_T and UK1_O colony types that are locked in their respective morphotypes in the absence of RecV. However, UK1::*recV* mutant produced colonies that are unlike UK1_T or UK1_O (Fig S 3.1C and D). *UK1::recV* mutants have longer cells with percentage sporulation significantly higher than UK1_T strain but significantly lower than UK1_O strain. This result suggested that RecV might not have a role in DNA inversion and recombinases other than RecV are likely to mediate DNA inversion in Cdi2 in UK1 strain.

Table 3.1. Whole genome sequence data analysis of UK1_T and UK1_O strain

Frequency	Annotation	Gene	Product
100%	Intergenic (+727/-1097)	*nrdG/23S rRNA*	Anaerobic ribonucleoside-triphosphate reductase activating protein/ 23S ribosomal RNA
98.2%	Intergenic (+568-238)	*CDR20291_0247/flgB*	Putative uncharacterized protein/flagellar basal-body rod protein
76.8%	Intergenic (+225/-131)	*CDR20291_0492/CDR20291_0493*	Conserved hypothetical protein/putative outer membrane lipoprotein
98.9%	Intergenic (+375/-1038)	*CDR20291_0684/CDR202910685*	Putative reductase/putative signaling protein

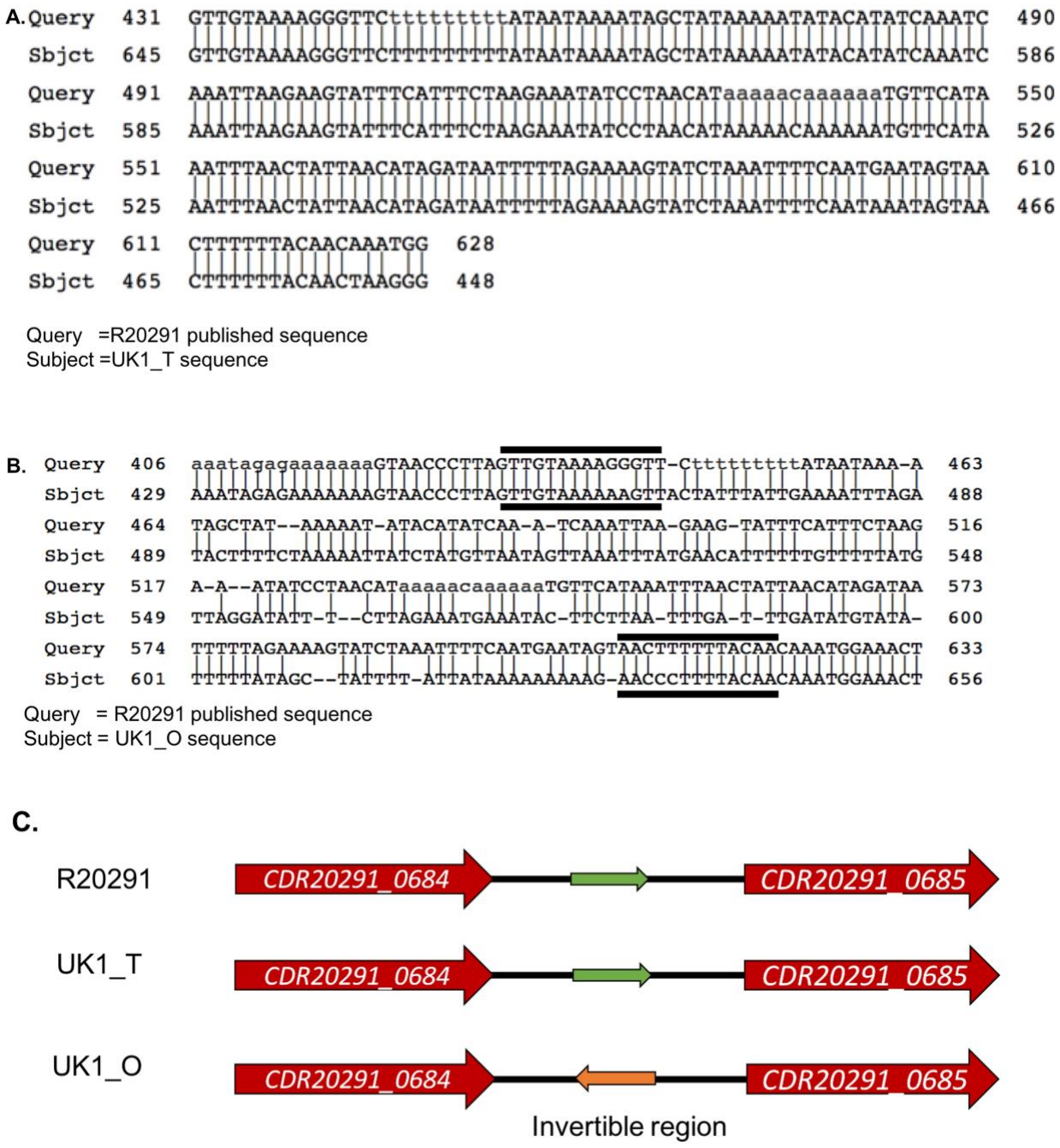


Figure 3.3. UK1_O undergoes DNA inversion in the upstream of *CDR20291_0685*

A. NCBI BLASTn result showing the perfect alignment of 168 bp segment of the upstream region of *CDR20291_0685* from UK1_T with the published R20291 sequence. **B.** NCBI BLASTn result showing the mismatch alignment of the 168 bp segment of the upstream region of *CDR20291_0685* from UK1_O strain with the published R20291 sequence. The invertible region is flanked by inverted repeats shown by the black lines**. C.** A schematic showing the sequence alignment of the intergenic region between *CDR20291_0684/CDR20291_0685* from R20291, UK1_T and UK1_O strain. The invertible region are shown by arrows in the intergenic region.

3.4.4 DNA inversion regulates the expression of the downstream gene *CDR20291_0685*

DNA inversion is known to regulate the expression of the downstream gene as it gives rise to two orientation of the loci where "ON" orientation enhances and "OFF" orientation represses the expression of the downstream gene or operon (1,3,58,59). In order to determine the effect of DNA inversion in the expression of the downstream *CDR20291_0685* gene, we performed a reporter fusion assay. We fused ~1.5 kb of the upstream DNA segment of *CDR20291_0685* containing the invertible region from both UK1_T and UK1_O strain with *gusA* reporter gene coding for beta-glucuronidase to create plasmid constructs pBA042 and pBA045 . Using site directed mutagenesis, we mutated the inverted repeat sequences that flanked the invertible region to lock the upstream region either in the translucent (pBA043) or in opaque (pBA046) orientation. The constructs were introduced into the UK1 undifferentiated parent strain by conjugation. The strains were grown anaerobically in TY broth medium at 37°C and cells were harvested at different time points to perform the reporter gene expression assay. We observed significantly higher beta glucuronidase activity at each time point with the reporter fusion that was locked in opaque orientation, with higher expression at early to mid-exponential phase (Fig. 3.4A). Very minimal reporter activity was recorded with the reporter fusion that was locked in the translucent orientation. This observation is consistent with our qRT-PCR results where we detected elevated levels of *CDR20291_0685* transcript in UK1_O when compared to that from UK1_T (Fig 3.4B). Taken together these results suggest that DNA inversion in the upstream region of the *CDR20291_0685* gene upregulated its expression in the opaque UK1_O colonies.

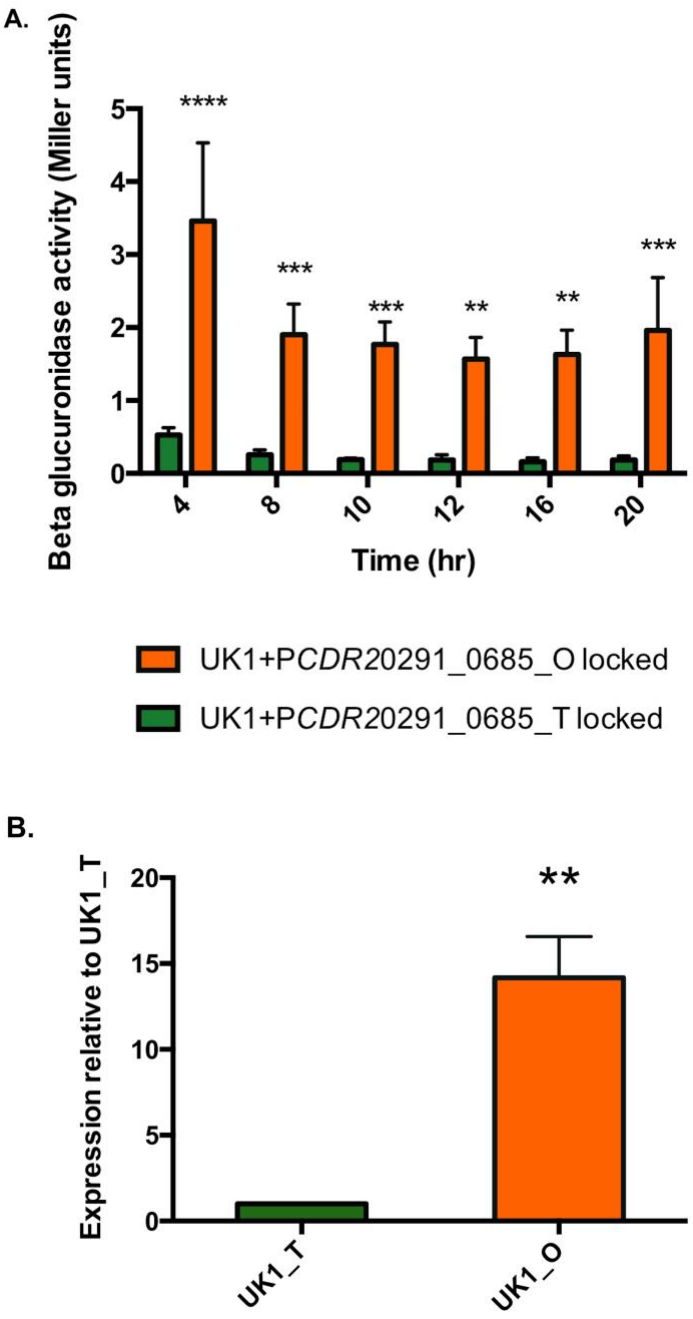


Figure 3.4. DNA inversion upregulates the expression of *CDR20291_0685*

A. Beta-glucuronidase activity of the P*CDR20291_0685* upstream-*gusA* fusions locked in UK1_T and UK1_O orientation in the UK1 parent strain. The data represents the results from 3 biological replicates. Data were analyzed using 2-way ANOVA (Sidak's multiple comparisons test) comparing mean of UK1_T and UK1_O at each time points. ****, ***, and ** indicates P<0.0005, P<0.005 and P<0.05 respectively. **B.** qRT-PCR results showing overexpression of *CDR20291_0685* in UK1_O strain. The representative results from three independent experiments are shown. The asterisks (**) indicate statistical difference at a P value of <0.05.

3.4.5 *CDR20291_0685* promoter lies within the invertible region.

The mechanism by which DNA inversion regulates the expression of the gene depends upon its location. If DNA inversion occurs in the region that contains the promoter, it will disrupt the promoter and the gene is not expressed (11,59). If the promoter is upstream of the switch then intrinsic terminator can be formed to terminate the transcription of gene (2).We know that the inversion sequence in *C. difficile* UK1 Cdi2 is upstream of the gene. In order to determine the mechanism by which this inversion is regulating the expression of the gene we sought to identify the promoter of *CDR20291_0685*. We hypothesized that the promoter of *CDR20291_0685* is within the invertible region. We extracted total RNA from 4 hr culture of UK1_O strain and synthesized *CDR20291_0685* specific cDNA. We carried out PCR using this cDNA as a template and several primers that were designed along the length of the invertible region (Fig 3.5A). Only the primer pairs ORG921/ORG853 and ORG922/ORG853 gave amplification product which suggests that the promoter region is located around the region spanned by the primer ORG921. 5' RACE assay to precisely map the promoter and determine the transcription initiation site is currently ongoing in our lab.

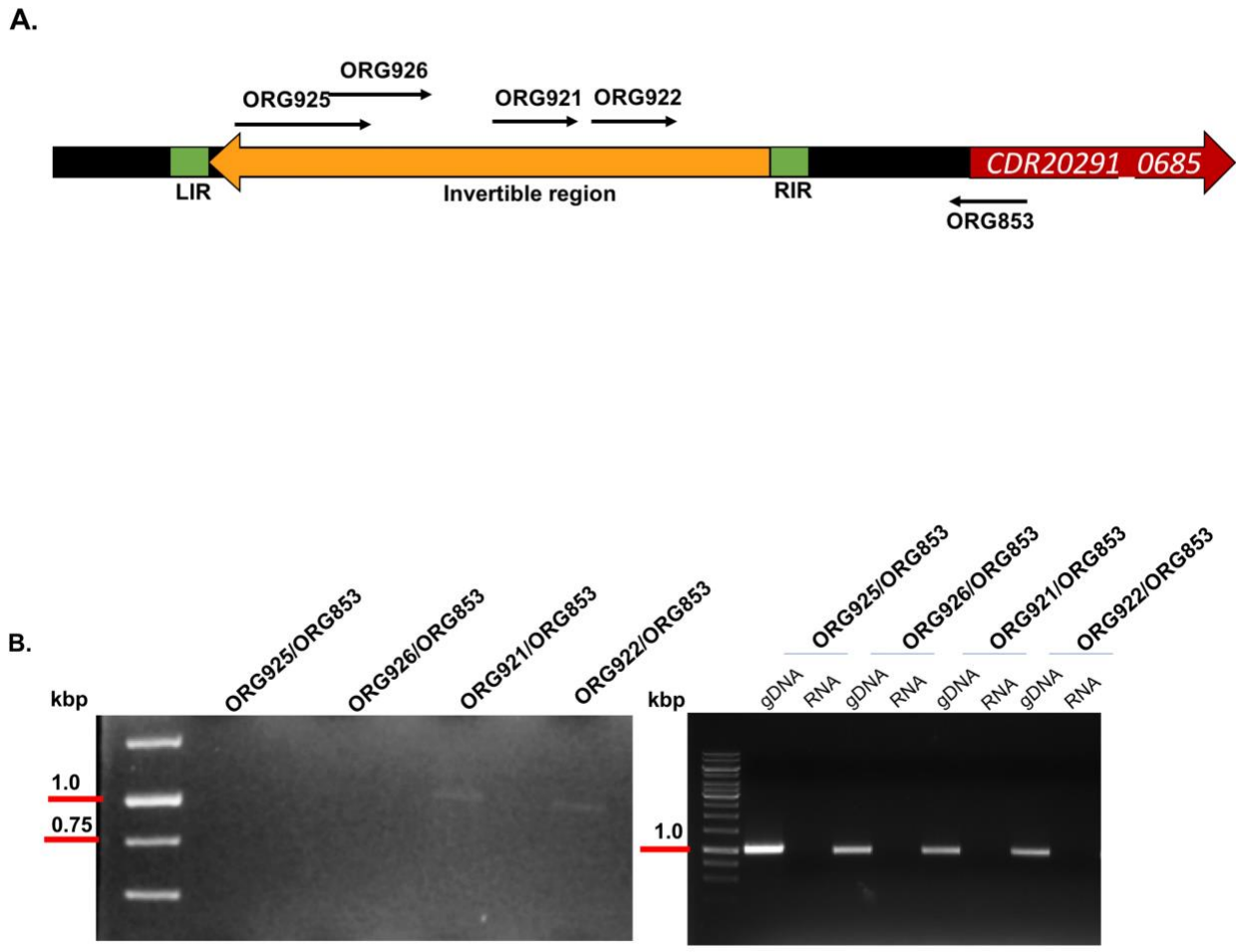


Figure 3.5. The promoter of *CDR20291_0685* is located within the invertible region

A. Schematic of the upstream region of *CDR20291_0685* from UK1_O strain (the orientation in which *CDR20291_0685* is expressed) depicting the invertible region flanked by the right and left inverted repeats (RIR and LIR). The forward primers designed along the length of the invertible region and the reverse primer are shown. **B.** Agarose gel electrophoresis of the reverse transcriptase PCR products. Genomic DNA (gDNA) and RNA were used as positive and negative controls, respectively. Only the primer pairs ORG921/ORG853 and ORG922/ORG853 gave amplification products of size 982bp and 942bp, respectively.

3.4.6 Overexpression of *CDR20291_0685* results in reduced c-di-GMP levels in UK1_O strain

CDR20291_0685 is the orthologous gene of *CD630_07570* in strain CD630. *CD630_07570* hydrolyzes c-di-GMP to pGpG *in vitro* and therefore is predicted to be a phosphodiesterase enzyme (60). Intracellular c-di-GMP homeostasis is maintained by two classes of enzymes, diguanylate cyclases that synthesize c-di-GMP and, phosphodiesterases that hydrolyze c-di-GMP. Phosphodiesterases consists of either EAL domain only or an inactive GGDEF and active EAL domain. *CDR20291_0685* shares 99.9% sequence similarity with *CD630_07570* and the GGDEF and EAL domains of *CD630_07570* are conserved in *CDR20291_0685* (Fig. 3.6A) and *in vitro* assays have shown *CDR20291_0685* cleaves c-di-GMP (60). Therefore, we predict *CDR20291_0685* to be a phosphodiesterase enzyme and named *CDR20291_0685* as *pdcB* (Phosphodiesterase of *C. difficile* B). We have demonstrated that DNA inversion in UK1_O strain leads to the overexpression of *pdcB* (Fig. 3.4). To determine if overexpression of *pdcB* alters the levels of intracellular c-di-GMP in *C. difficile* UK1_O strain, we first created UK1::*pdcB* mutant strain using ClosTron mutagenesis (Fig S 3.2) and extracted and measured the levels of intracellular c-di-GMP from UK1_T, UK1_O and UK1::*pdcB* strains. Our data suggests that UK1_T and UK1::*pdcB* strains have higher levels of c-di-GMP while UK1_O strain has significantly reduced c-di-GMP level (Fig. 3.6B). UK1::*sinR* mutant strain was used as a positive control as the previous study from our lab has shown that *sinR* mutant have elevated intracellular c-di-GMP level. This result suggests that DNA inversion overexpresses *pdcB* gene which cleaves c-di-GMP and reduces the global intracellular concentration of c-di-GMP in UK1_O strain.

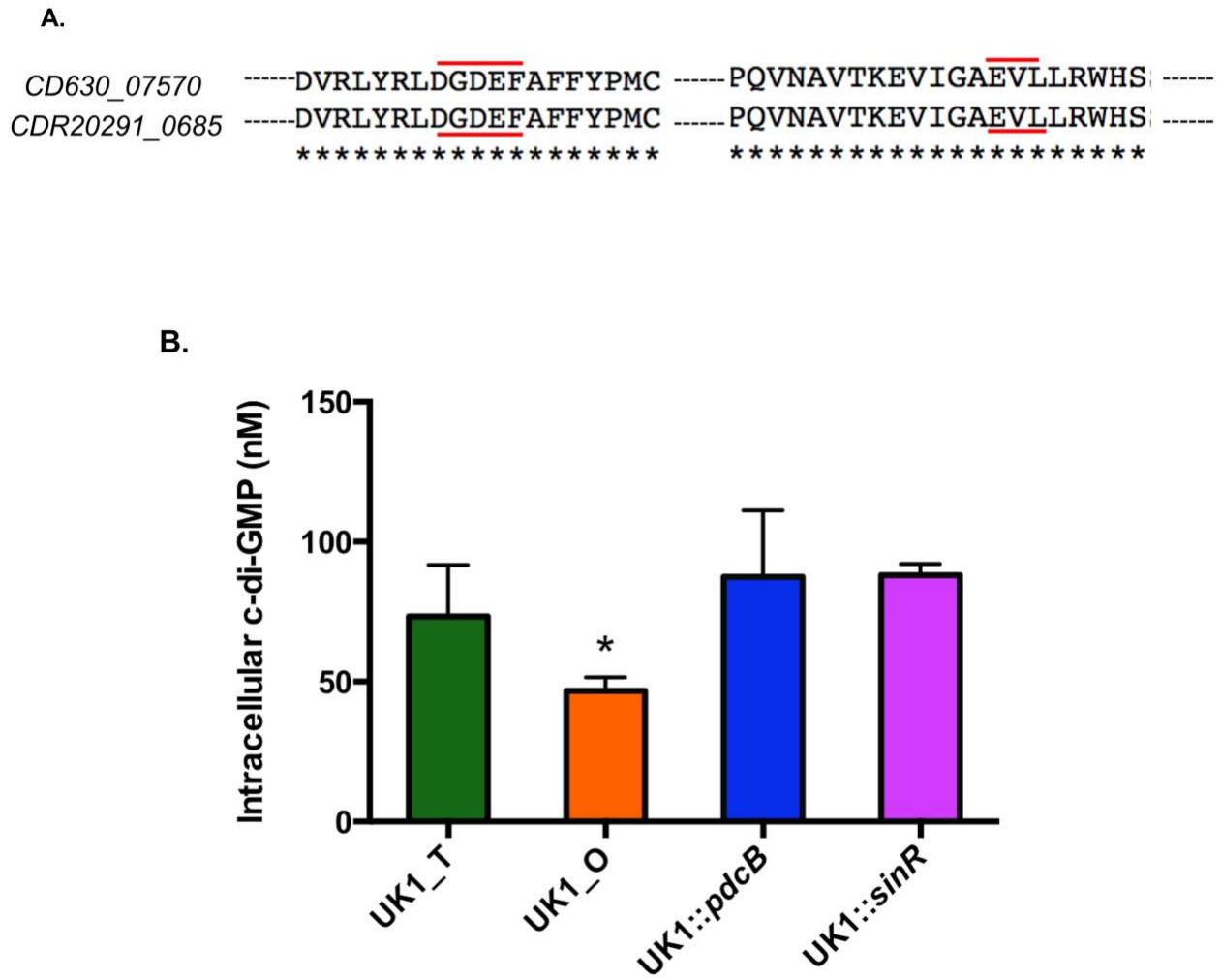


Figure 3.6. Overexpression of *pdcB* reduces the intracellular c-di-GMP levels in UK1_O strain

A.Sequence alignment between *CD630_07570* and *CDR20291_0685* showing the conserved GGDEF (DGDEF) and EAL (EVL) domains. **B.** The intracellular levels of c-di-GMP in UK1_T, UK1_O and UK1::*pdcB* mutant strains. UK1::*sinR* is used as a positive control. Three biological replicates were used to carry out the experiment and data was analyzed by one-way analysis of variance (ANOVA) where * indicates $P<0.05$.

3.4.7 CodY binds to the upstream region and represses the expression of *pdcB*

Several studies have shown CodY mediated regulation of c-di-GMP signaling proteins. Microarray analysis in CD630 strain have identified two cyclic di-GMP signaling proteins, *CD630_07570* (*pdcB*) and *CD1476* , among the 140 genes that were regulated by CodY (61). The study demonstrated enhanced expression of both of these proteins in the *codY* mutant and identified a consensus CodY binding site in the upstream region of *CD1476*. However, affinity purification in the same study did not report the presence of consensus CodY binding site in the upstream region of *CD630_07570*. Another study have shown that CodY binds directly to the promoter region and represses the expression of *pdcA,* a phosphodiesterase that cleaves c-di-GMP in JIR8094 strain (62). The effect of CodY on the regulation of c-di-GMP signaling protein is less studied in *C. difficile* 027 ribotype. (60). To determine if expression of *pdcB* is also dependent on CodY, we first determined whether a CodY binding consensus sequence is present in the upstream region of *pdcB*. The motif, AATTTTCWGAAAATT, has been identified in the upstream region of genes that are regulated by CodY in *B. subtilis* and *Lactococcus lactis* (63,64). We found two potential CodY binding sites, TTTTTAGAAAAGTA and AAATTTTCAATGAAT, in the upstream of *pdcB* within the region that undergoes inversion (Fig. 3.7A). These sites are within the DNA region that undergoes inversions. Interestingly, CodY binding site in the opaque orientation falls right upstream of the predicted promoter region. So, we hypothesize that CodY binds to the upstream region of *pdcB* and represses its expression in UK1_T strain. DNA inversion in UK1_O strain would shift the CodY binding site to the upstream region of the predicted promoter and thus relieves the repression leading to the overexpression of *pdcB*.

To determine if CodY regulates the expression of *pdcB*, we carried out the reporter fusion assay. We used the same reporter fusions (pBA040, pBA043, pBA046) that was used earlier in this study. Each of the constructs were introduced into the UK1 undifferentiated parent strain and UK1::*codY* mutant strain by conjugation. The strains were grown anaerobically in TY broth medium at 37°C and cells were harvested at different time points to perform the reporter assay. We observed increased beta glucuronidase activity at each time point when it was expressed from the *pdcB* upstream locked in opaque orientation, with higher expression at mid exponential phase from 4-12 hours for both UK1 parent and UK1::*codY* mutant strains (Fig 3.7B). Minimal reporter activity was observed with the upstream locked in translucent orientation. The expression of the reporter gene was significantly higher in UK1::*codY* at 8-12 hour time points compared to the UK1 parent strain suggesting that the presence of CodY in UK1 strain repressed the expression of *pdcB* during these time points and the repression was released in the absence of CodY in the UK1::*codY* mutant strain.

To further confirm the interaction of CodY with *pdcB* upstream DNA, we carried out Electrophoretic Mobility Shift Assay (EMSA) with purified CodY protein. We synthesized the DNA fragment of 47bp containing the potential CodY binding site and radio-labelled with γ-^{32}P and carried out the binding of the fragment with purified CodY protein with increasing concentration of the protein, both in the absence and presence of its cofactors ILV (Branched chain amino acids Isoleucine, Leucine and Valine) and GTP (Fig 3.7C). The shift in the bands at higher concentrations of CodY protein suggests that purified CodY binds to the DNA fragment and the binding was enhanced in the presence of its cofactors GTP and ILVs . These data suggests that CodY has a direct role in downregulating the expression of *pdcB*.

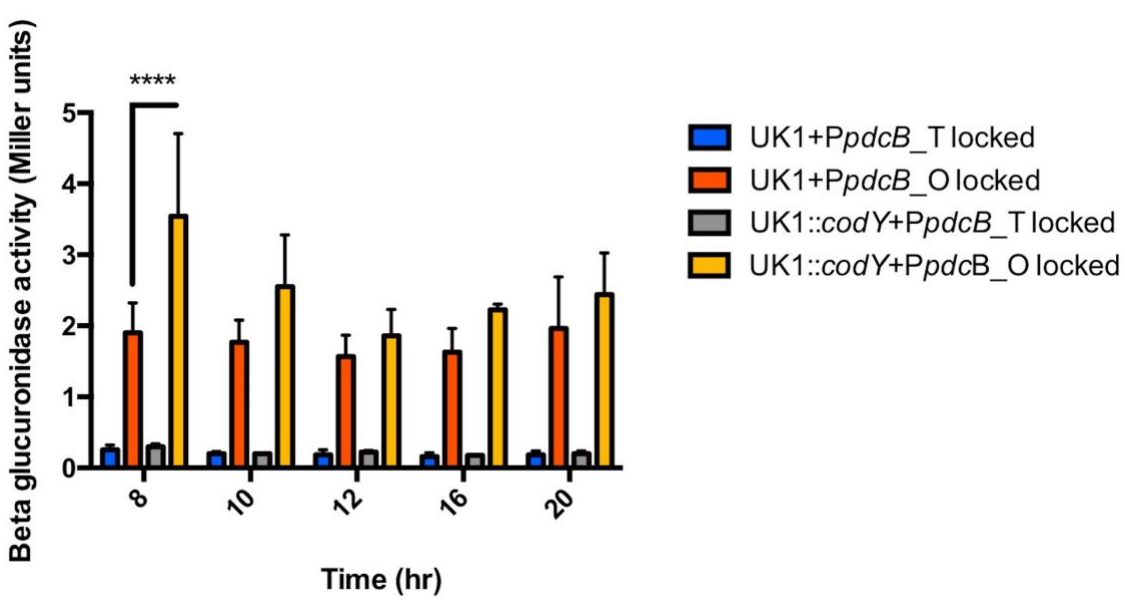
Taken together these results suggest that *pdcB* expression in UK1 is tightly regulated by both DNA inversion and the activity of CodY and DNA inversion partially releases the repression.

A.

```
Query  406  aaatagagaaaaaaaGTAACCCTTAGTTGTAAAAGGGTT-Cttttttttt ATAATAAA-A  463
Sbjct  429  AAATAGAGAAAAAAAGTAACCCTTAGTTGTAAAAAAGTTACTATTTATTGAAAATTTAGA  488
Query  464  TAGCTAT--AAAAAT-ATACATATCAA-A-TCAAATTAA-GAAG-TATTTCATTTCTAAG  516
Sbjct  489  TACTTTTCTAAAAATTATCTATGTTAATAGTTAAATTTATGAACATTTTTTGTTTTTATG  548
Query  517  A-A--ATATCCTAACATaaaaacaaaaaaTGTTCATAAATTTAACTATTAACATAGATAA  573
Sbjct  549  TTAGGATATT-T--CTTAGAAATGAAATAC-TTCTTAA-TTTGA-T-TTGATATGTATA-  600
Query  574  TTTTTAGAAAAGTATCTAAATTTTCAATGAATAGTAACTTTTTTACAACAAATGGAAACT  633
Sbjct  601  TTTTTATAGC--TATTTT-ATTATAAAAAAAAAG-AACCCTTTTACAACAAATGGAAACT  656
```

Query=published R20291 sequence
Subject=UK1_O sequence

B.



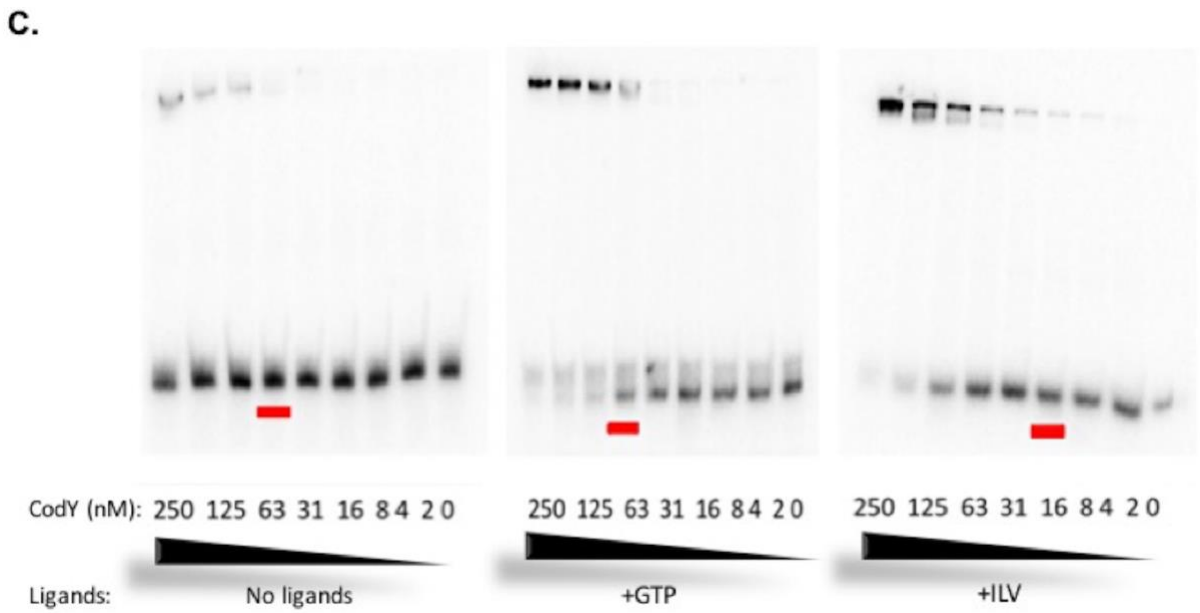


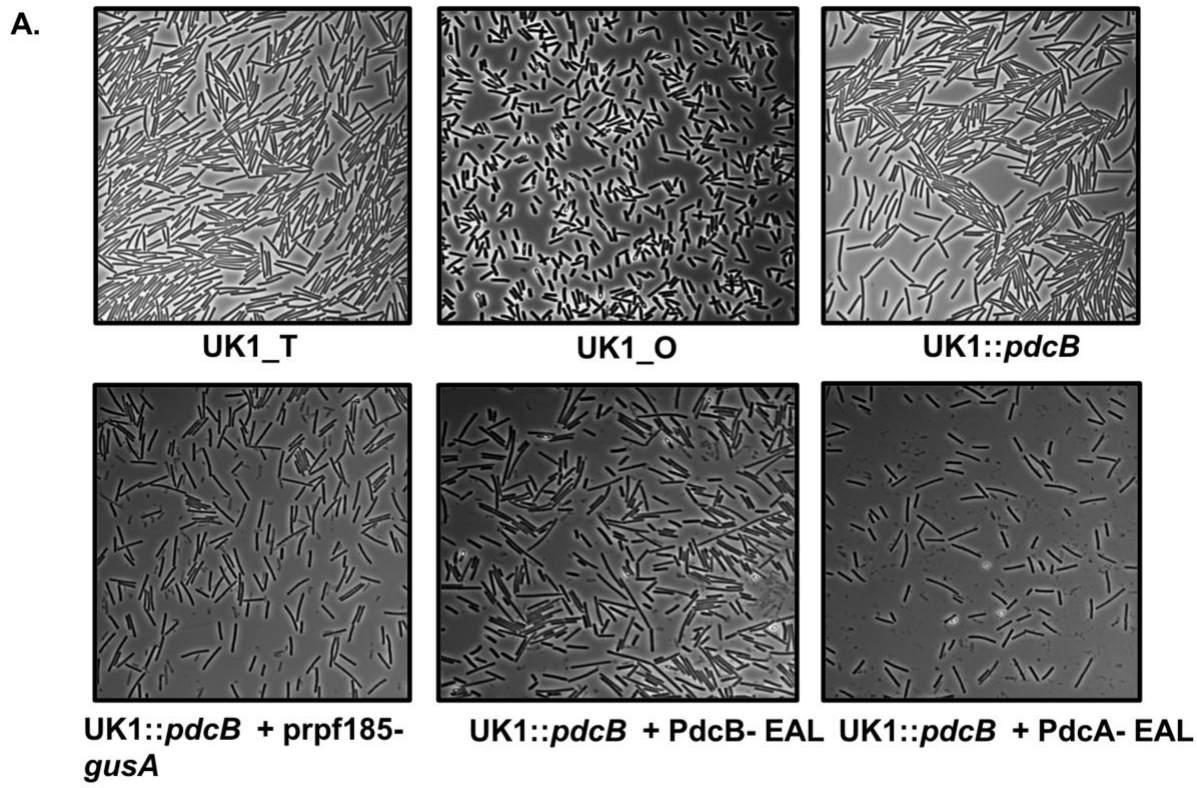
Figure 3.7. CodY represses the expression of *pdcB*, which is partially relieved by DNA inversion

A.The potential CodY binding sites that lie within the invertible region in the upstream of *pdcB.* The sites in the published R20291 sequence (Query strand) are near to the RIR and are highlighted in Red. The sites in the Opaque orientation (Subject strand) are near to the LIR and are highlighted in green. **B**. Beta-glucuronidase activity of P*pdcB-gusA* fusions locked in either translucent or opaque orientation in the parent UK1 and UK1::*codY* mutant. The data shown are the standard errors of the mean of three biological replicates. Statistical analysis was performed using two-way ANOVA with Tukey's multiple comparisons test comparing values to the average of each strain with another strain. ****=P<0.0005.**C.** Binding of purified CodY to the upstream region of *pdcB.* Binding was increased with increasing concentration of CodY. Presence of GTP and ILV further enhanced CodY binding.

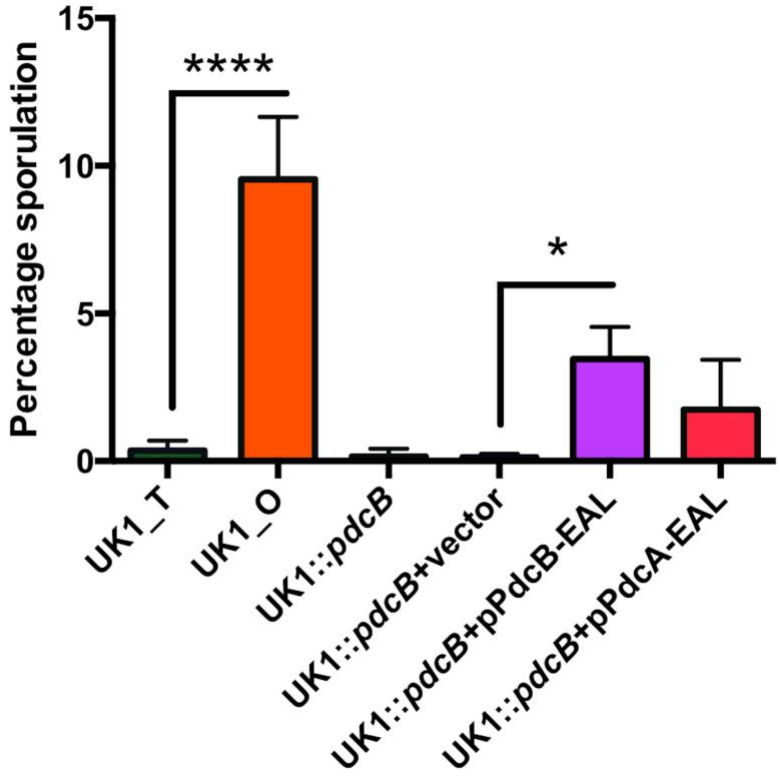
3.4.8 Sporulation in *C. difficile* UK1 is associated with the intracellular levels of c-di-GMP

DNA inversion lead to the overexpression of *pdcB* and hence reduced amount of intracellular levels of c-di-GMP in *C. difficile* UK1_O strain. We also observed hyper-sporulation in UK1_O strain. In order to determine if the intracellular levels of c-di-GMP is associated with a sporulation phenotype, we quantified percentage sporulation in UK1_T, UK1_O, UK1::*pdcB* mutant and UK1::*pdcB* mutant complemented with plasmid-encoded PdcB-EAL domain under a tetracycline inducible promoter. Despite multiple attempts, we were not able to complement the UK1::*pdcB* mutant strain with a full length plasmid-encoded *pdcB*. Other studies have shown that complementation by just the EAL domain with the phosphodiesterase activity of *pdcA* is sufficient to rescue the associated motility phenotypes (62). Since our objective was to determine the role of intracellular levels of c-di-GMP in regulating sporulation in *C. difficile* UK1 strain, we performed the sporulation assays with the mutant strain complemented with just the plasmid-encoded PdcB-EAL domain. Percentage sporulation data demonstrated hypo-sporulation in UK1::*pdcB* strain and hyper-sporulation in mutant was complemented with the PdcB-EAL domain expressed from a plasmid (Fig 3.8A and 3.8B). In order to further confirm that sporulation phenotype is associated with intracellular levels of c-di-GMP, we complemented UK1::*pdcB* mutant with the plasmid borne PdcA-EAL domain under an inducible promoter. PdcA is another well characterized phosphodiesterase of *C. difficile* and PdcB and PdcA share homology in the conserved GGDEF and EAL domain (Fig 3.8C). We observed that sporulation phenotype was complemented with PdcA-EAL domain as well.

Intracellular levels of c-di-GMP is known to positively regulate swarming motility in *C. difficile* (37). Consistent with previous studies, we observed increased swarming motility in UK1_T strain which has increased intracellular c-di-GMP levels compared with the UK_O strain. So, to further confirm that complementation of the UK1::*pdcB* mutant with PdcB-EAL and PdcA-EAL domains reduces the intracellular c-di-GMP and sporulation phenotype that we observed is associated with the intracellular c-di-GMP levels, we carried out swarming motility test assay for all the strains (Fig 3.8DE). As expected, we observed decreased swarming motility in the complemented strains similar to UK1_O strain which has reduced intracellular c-di-GMP levels. Taken together, these data suggest that sporulation in *C. difficile* is dependent on the intracellular levels of c-di-GMP.



B.



c.

```
CD630_07570 RLDGDEFAFFYPMC----VIGAEVLLRWHSSTYGEVS
pdcB        RLDGDEFAFFYPMC----VIGAEVLLRWHSSTYGEVS
pdcA        KLDGDEMGILVDNV----IIGVEVLLRWTNDKCKAIS
            :*****:.::        :**.****** ...   :*
```

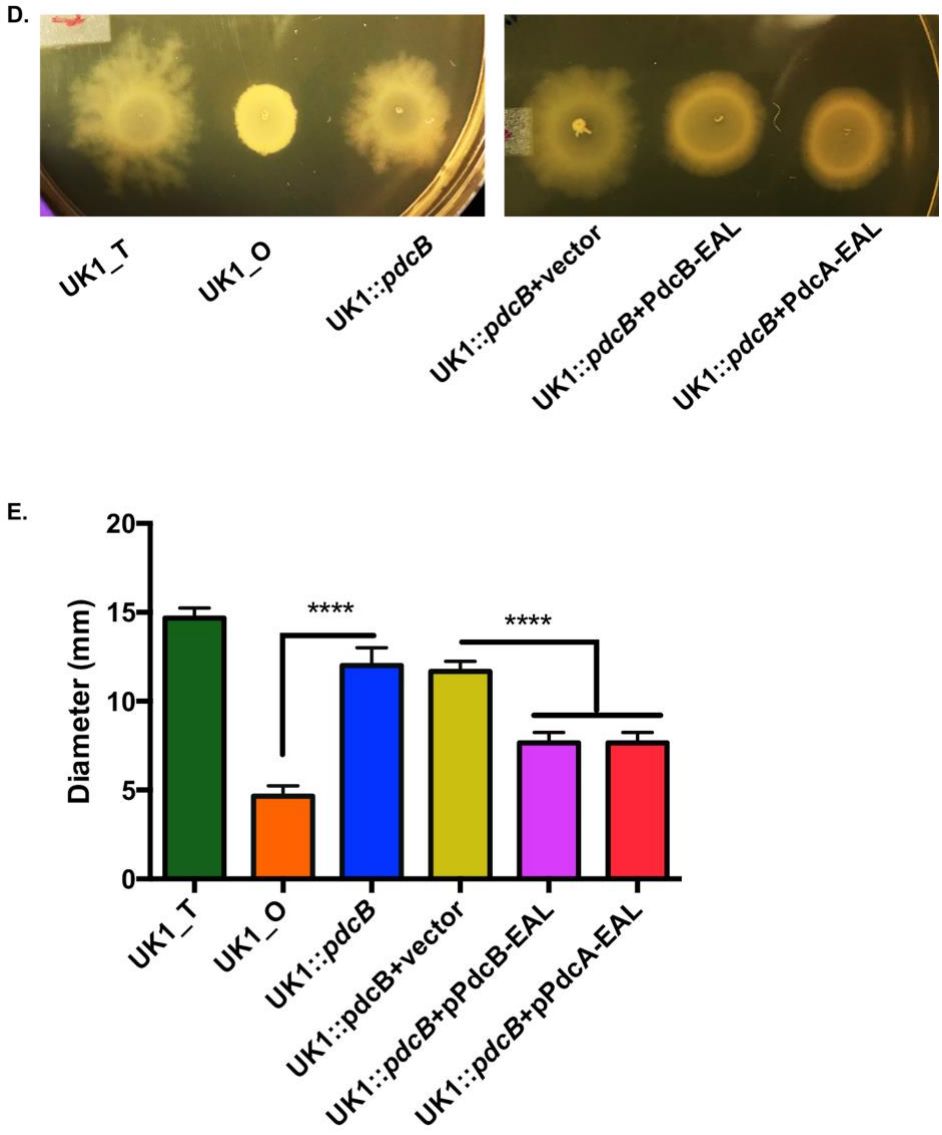


Figure 3.8. Reduced c-di-GMP is associated with hyper sporulation phenotype in in UK1_O strain

A. Phage contrast microscopy images of paraformaldehyde fixed UK1_T, UK1_O, UK1::*pdcB* mutant, UK1::*pdcB* + PdcB- EAL and UK1::*pdcB* + PdcA- EAL complemented strains showing their associated sporulation phenotypes. **B.** Percentage sporulation showing partial complementation of the hypo-sporulation phenotype in UK1::*pdcB* mutant strain byboth PdcB-EAL and PdcA-EAL complementation. **C.** Sequence alignment of the *CD630_07570*, *pdcB* and *pdcA* using ClustalW. The red lines show the conserved GGDEF and EAL domains. **D.** Representative swarming motility plates of each strain after 72 hours of incubation at 37°C. **E.** Quantitative measurements of swarming motilities of strains grown after 72 hours. The data shown are the standard errors of the mean of three biological replicates. Statistical analysis was performed using one-way ANOVA with Sidak's multiple comparisons test comparing values to the average of each strain with another strain. ****=P<0.0005 *and* *=P<0.05.

3.5 Discussion

The major objective of this study was to determine the role of phase variation in *C. difficile* physiology and pathogenesis. Genome wide analysis identified 7 different types of invertible sites in *C. difficile* genome, designated as Cdi1 to Cdi7 (2). Cdi1 is shown to control the expression of *cwpV*, a gene that encodes a surface expressed protein (65). Altering the expression of cell wall protein, which potentially helps *C. difficile* to adhere and interact with the host cells, can provide *C. difficile* an advantage in altered environmental condition during infection to detach from host substrates and disseminate, thus evading from the immune responses (65). Cdi4 is present in the upstream of the flagellar biosynthesis operon and is shown to regulate the expression of flagellar genes and also indirectly control the production of *C. difficile* toxins by regulating the expression of *sigD* which is a part of *flgB* operon and which directly regulates *tcdR*, the positive regulator of toxins TcdA and TcdB (55,66,67). Recent study demonstrated that Cdi6 regulates the expression of *cmrRST* which regulates colony morphology, motility and virulence in *C. difficile*. Cdi3, Cdi5, Cdi7 are poorly understood as they have not been studied yet. Our current work studies the role of Cdi2 in regulating pathogenesis and physiology in *C. difficile* and is the first study to show phase variation regulated c-di-GMP homeostasis in *C. difficile* and the novel implication of intracellular c-di-GMP in regulating sporulation in *C. difficile*.

Previous studies have shown that Cdi2 is present in the upstream region of *pcdB* (*CDR20291_0685*), which encodes one of the putative c-di-GMP phosphodiesterases (PDEs) in *C. difficile* UK1 strain (2). In this work, we have shown the role of Cdi2 in regulating the expression of *pdcB*. Our data shows that DNA inversion in the upstream of *pdcB* in UK1_O strain overexpresses *pdcB* and this genotype is associated with reduced intracellular c-di-GMP. Consistent with this result, UK1_T strain which had minimal *pdcB* expression and UK1::*pdcB*

mutant demonstrated relatively higher levels of c-di-GMP. *C. difficile* is predicted to have 18 predicted PDEs (60). Despite the redundancy, intracellular c-di-GMP levels seems to be greatly affected by the expression of *pdcB.* Increase in intracellular c-di-GMP by overexpression of the diguanylate cyclase gene *dccA* demonstrated decrease in *pdcB* orthologue in *C. difficile* 630Δerm strain (30). This result further strengthens our observation that *pdcB* is an important phosphodiesterase to maintain c-di-GMP homeostasis in *C. difficile*. Our study is also the first one to report the role of c-di-GMP concentration in regulating *C. difficile* sporulation.

While c-di-GMP regulates plethora of physiological processes in *C. difficile*, its role in sporulation had not been previously described. Only a few studies have investigated the role of c-di-GMP in sporulation. Recently, the relationship between c-di-GMP and sporulation has been explored in *Streptomyces venezuelae*, a Gram-positive soil bacterium, where low levels of c-di-GMP resulted in hyper sporulation phenotype (68–73). Gallagher KA, et al. reported that c-di-GMP bridges the binding of anti-sigma factor RsiG with sigma factor WhiG, thus strengthening their interaction and preventing sigma factor WhiG-dependent transcription of late sporulation genes (70). Similarly, another study demonstrated that c-di-GMP is required to control transition from vegetative to reproductive cells (74). Their study reported that c-di-GMP was required to form the active dimeric form of BldB, the master regulator of cell development, and represses the development of reproductive hyphae and keeps *Streptomyces* in their vegetative state (70–72,74). Taken together, these studies suggest that high c-di-GMP level traps the bacteria in vegetative growth and low c-di-GMP levels causes precocious hyper sporulation. In another study carried out in *Streptomyces coelicolor*, overexpression of CdgD, a diguanylate cyclase, resulted in completely blocked sporulation or aerial mycelium formation (69). Similarly, deletion of diguanylate cyclases, *cdgB* and *cdgC,* in *Streptomyces venezuelae* resulted in enhanced sporulation and deletion of

phosphodiesterases *rmdA* and *rmdB* delayed development (71). Single cell microscopy study carried out in *B. subtilis* have shown positive correlation of high c-di-GMP levels with sporulation (75). In *Myxococcus xanthus*, DmxB was identified as a diguanylate cyclase which resulted in increased c-di-GMP levels and was essential for completion of fruiting body formation and sporulation (76). Our current work is the first study to directly associate intracellular c-di-GMP concentration to sporulation in *C. difficile*. Hyper sporulation was observed when the intracellular c-di-GMP concentration was reduced by 1.5-fold in UK1_O strains compared to UK1_T strains grown *in vitro*. This suggests that sporulation phenotype is very sensitive to the intracellular levels of c-di-GMP concentration because even a small change can give robust sporulation. It is noteworthy that, in our previous study, R20291::*sinRR'* mutant strain, which is asporogenic had elevated levels of c-di-GMP compared with R20291 parent strain, which is hyper sporulating and had lower level of c-di-GMP (77). This study further corroborates our current observation. By analyzing percentage sporulation in UK1::*pdcB* mutant strain and PdcB-EAL and PdcA-EAL domains complemented strains we have shown that *pdcB* gene expression and hence the intracellular c-di-GMP level, is positively associated with sporulation phenotype. However, the overexpression of EAL domains of both PdcB and PdcA have partially complemented sporulation phenotype. This could be because of the overexpression of PdcB-EAL domain alone without the associated GGDEF domain. Although GGDEF domain of PdcB-EAL is catalytically inactive (2), it might be necessary to enhance the activity of the EAL domain and fully complement sporulation. In homologous proteins which contains both GGDEF and EAL domains, catalytically inactive GGDEF domains are capable of binding to GTP thus enhancing the PDE activity of the neighboring EAL domain (78).

Regulation of *C. difficile* sporulation can occurs either at the transcriptional level by altering the expression of *spo0A* or at the post translational level by interfering with the phosphorylation of Spo0A (21). There are two known mechanisms by which c-di-GMP is known to function. First, c-di-GMP mediated regulation can occur through RNA based effectors which includes direct binding of c-di-GMP to riboswitches present in the 5' UTR of the target gene transcript leading to premature termination of the transcript. Second, c-di-GMP mediated regulation can occur through protein effectors containing PilZ domain, GMP binding domain, diguanylate cyclases containing I-sites, GIL proteins, MshEN domains and the Cle subfamily of CheY proteins which sense and respond to the changes in the intracellular c-di-GMP concentration (31). Whether c-di-GMP binding riboswitches or c-di-GMP binding domains are present in the histidine kinases that phosphorylate Spo0A, and whether c-di-GMP regulates their expression or activity thus regulating sporulation needs to be further investigated. On the other hand, c-di-GMP could indirectly influence sporulation, by affecting the expression of regulators known to affect *spo0A.*

Previous studies have shown that *C. difficile* UK1::*codY* mutant exhibits a hyper-sporulation phenotype suggesting that CodY represses sporulation (79). The exact mechanism by which CodY regulates sporulation is not understood, however, studies have shown that CodY directly regulates the expression of *sinRI* (77) and *opp* and *app* (79,80), both of which are positive regulators of sporulation. By showing that CodY directly binds to the upstream of *pdcB* and represses the expression of *pdcB* during exponential phase of growth, thereby reducing intracellular level of c-di-GMP and causing hyper sporulation, our study has shown a novel pathway of CodY mediated regulation of sporulation. Intracellular GTP is known to regulate the activity of DGCs. The glycine at the second position and the aspartic acid at the fourth position

of catalytically active GGDEF domain of DGCs are suggested to bind to the guanine base of GTP and magnesium ion that interacts with phosphates, respectively (81,82). These residues are conserved in the GGDEF domain of PdcA and it has been shown that increasing concentration of GTP enhances the phosphodiesterase activity of PdcA, suggesting the role of intracellular GTP in the acitivity of PDE (62). The second and the fourth residues of GGDEF domain of PdcA are conserved in PdcB which suggests that intracellular level of GTP might be important for PdcB activity. Reporter fusion assays and qRT-PCR results from our study suggest that *pdcB*, like CodY, is expressed during the exponential phase of growth suggesting the repression of CodY on *pdcB* is likely driven by the competition for available intracellular GTP. Since the activity of CodY depends on the intracellular level of GTP, our study further strengthens the insight in the link between cytoplasmic levels of GTP and phosphodiesterase activity.

Previous studies have also shown the role of c-di-GMP in regulating cell aggregation thus leading to biofilm formation (34,82,83). In *C. difficile*, increased intracellular c-di-GMP concentration is shown to upregulate *pilA1* transcription leading to pilus formation, increased swarming motility and increased cell aggregation (37). Consistent with these studies, we observed elevated levels of *pilA* transcript (Fig S 3.3), increased swarming motility and biofilm formation in UK1_T strain compared to UK1_O strain. Phase contrast microscopy images demonstrated UK1_T and UK1::*pdcB* mutant cells, both of which have relatively higher level of c-di-GMP levels than UK1_O strain, were longer and UK1_O cells were shorter in length. Transmission electron microscopy images revealed that a longer looking cell existed as two cells divided by a septum but not separated into individual cells (Fig S 3.4). Overexpression of PdcB-EAL domain in UK1::*pdcB* complemented cell separation, swarming motility and biofilm formation. These observations corroborates that the intracellular level of c-di-GMP levels enhances *pilA* expression and has a

direct role in impairing cell separation after cell division and enhancing cell aggregation, swarming motility and biofilm formation. Phase variation mediated regulation of cell aggregation, biofilm formation and swarming motility has recently been described (34,37).

C. difficile wild type strains has been shown to produce smooth and rough colony morphologies (20,56). Phase variation is suggested to regulate variation in colony morphology in some *C. difficile* strains (20,65–67). Recent study on Cdi6 has suggested the role of c-di-GMP on colony morphology. By showing that the intracellular c-di-GMP concentration is regulated by the phase variable expression of *pdcB* , our study has identified *pdcB* as an upstream signal that most likely modulates Cdi6 mediated regulation of colony morphology. The same study also suggested that the colony morphology in *C. difficile* is dependent on the growth condition with swimming condition selective of smooth colonies and surface condition selective of rough colonies (20). Our result shows that colony morphology of UK1_T and UK1_O is very similar in swimming conditions but is different in swarming conditions with UK1_T maintaining the rough colony and UK1_O maintaining the smooth colony. However, we observed that UK1::*pdcB* mutant exhibited rough colony in the swarming plate and this phenotype was complemented by the overexpression of PdcB-EAL and PdcA-EAL domains. This suggests that Translucent and Opaque colony morphologies in UK1 strain are independent of growth condition and but are dependent on the intracellular c-di-GMP levels which is regulated by the phase variable expression of *pdcB.* Results from orientation specific PCR of the invertible region of *pdcB* and *cmrRST* further supports that translucent and opaque colony morphologies are specifically associated with *pdcB* expression (Fig S 3.5 A and B).

Our study has shown phase variation mediated regulation of intracellular c-di-GMP homeostasis and its effect on *C. difficile* physiology (Fig 3.9). Although both translucent and

opaque strains are shown to produce similar levels of toxins, they differ drastically in other phenotypes like sporulation and colony morphology. Based on the increased *pilA* expression and biofilm formation phenotype, it can be suggested that during initiation of infection, UK1 might prefer to exist as the translucent colony type so that they can better colonize the host and cause infection. Under the selective pressure exerted in the intestinal tract and host immune responses, UK1 strain might prefer to switch to Opaque colony type and activate its signaling cascade to produce more spores so that it can evade the host immune response and persist in the host to cause recurrent infection or in the environment once it is released outside in the feces of the infected individual. Therefore, understanding the importance of phase variable expression of *pdcB* and its associated phenotypes during in vivo infection model is a priority for future investigation and will shed light on the importance of *pdcB* in *C. difficile* pathogenesis.

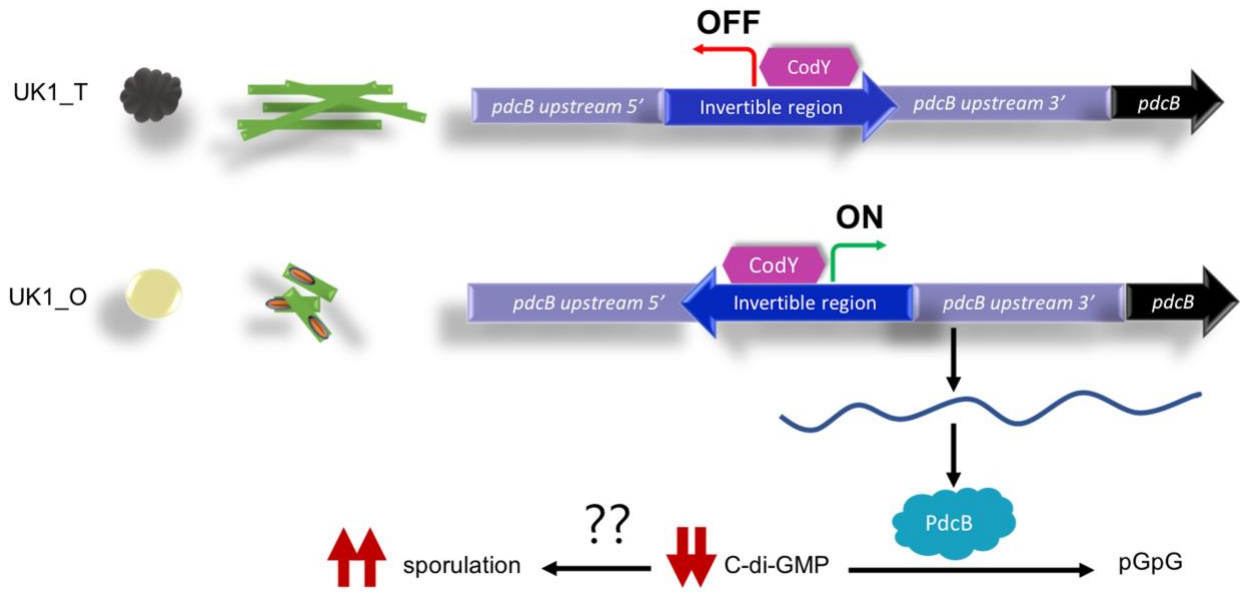


Figure 3.9. Proposed model of this study

CodY binds to the upstream region of *pdcB* and represses the expression of *pdcB*. The promoter region of *pdcB* is predicted to be upstream of the CodY binding site in translucent orientation. In opaque orientation repression by CodY is partially relieved by DNA inversion which results in switching the promoter the "ON" orientation and thus expresses the *pdcB* gene. Overexpression of *pdeB* reduces the intracellular c-di-GMP concentration which is associated with hyper-sporulation phenotype.

3.6 References

1. Silverman M, Simon M. Phase variation: genetic analysis of switching mutants. Cell. 1980 Apr;19(4):845–54.

2. Sekulovic O, Mathias Garrett E, Bourgeois J, Tamayo R, Shen A, Camilli A. Genome-wide detection of conservative site-specific recombination in bacteria. PLoS Genet. 2018;14(4):e1007332.

3. Jiang X, Hall AB, Arthur TD, Plichta DR, Covington CT, Poyet M, et al. Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. Science. 2019 Jan 11;363(6423):181–7.

4. Li J, Zhang J-R. Phase Variation of *Streptococcus pneumoniae*. Microbiol Spectr. 2019;7(1).

5. Smith NH, Selander RK. Sequence invariance of the antigen-coding central region of the phase 1 flagellar filament gene (fliC) among strains of *Salmonella typhimurium*. Journal of Bacteriology. 1990;172(2):603–9.

6. Salehi S, Howe K, Lawrence ML, Brooks JP, Bailey RH, Karsi A. *Salmonella enterica Serovar Kentucky* Flagella Are Required for Broiler Skin Adhesion and Caco-2 Cell Invasion. Appl Environ Microbiol. 2017 15;83(2).

7. Silverman M, Zieg J, Hilmen M, Simon M. Phase variation in *Salmonella*: genetic analysis of a recombinational switch. Proc Natl Acad Sci U S A. 1979 Jan;76(1):391–5.

8. Kutsukake K, Iino T. Inversions of specific DNA segments in flagellar phase variation of *Salmonella* and inversion systems of bacteriophages P1 and Mu. Proc Natl Acad Sci U S A. 1980 Dec;77(12):7338–41.

9. Scott TN, Simon MI. Genetic analysis of the mechanism of the *Salmonella* phase variation site specific recombination system. Mol Gen Genet. 1982;188(2):313–21.

10. Kutsukake K, Iino T. A trans- acting factor mediates inversion of a specific DNA segment in ftagellar phase variation of *Salmonella*. 1980;284:3.

11. Eisenstein BI. Type 1 fimbriae of *Escherichia coli*: genetic regulation, morphogenesis, and role in pathogenesis. Rev Infect Dis. 1988 Aug;10 Suppl 2:S341-344.

12. Klemm P. Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in *Escherichia coli*. The EMBO Journal. 1986 Jun 1;5(6):1389–93.

13. McClain MS, Blomfield IC, Eisenstein BI. Roles of fimB and fimE in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. Journal of Bacteriology. 1991;173(17):5308–14.

14. McClain MS, Blomfield IC, Eberhardt KJ, Eisenstein BI. Inversion-independent phase variation of type 1 fimbriae in *Escherichia coli*. Journal of Bacteriology. 1993;175(14):4335–44.

15. Struve C, Bojer M, Krogfelt KA. Characterization of *Klebsiella pneumoniae* Type 1 Fimbriae by Detection of Phase Variation during Colonization and Infection and Impact on Virulence. Infection and Immunity. 2008 Sep 1;76(9):4055–65.

16. Tauseef I, Ali YM, Bayliss CD. Phase Variation of PorA, a Major Outer Membrane Protein, Mediates Escape of Bactericidal Antibodies by *Neisseria meningitidis*. Camilli A, editor. Infect Immun. 2013 Apr;81(4):1374–80.

17. van der Ende A, Hopman CTP, Dankert J. Multiple Mechanisms of Phase Variation of PorA in *Neisseria meningitidis.* Infection and Immunity. 2000 Dec 1;68(12):6685–90.

18. Holst Sørensen MC, van Alphen LB, Fodor C, Crowley SM, Christensen BB, Szymanski CM, et al. Phase variable expression of capsular polysaccharide modifications allows *Campylobacter jejuni* to avoid bacteriophage infection in chickens. Front Cell Infect Microbiol. 2012;2:11.

19. Bacon DJ, Szymanski CM, Burr DH, Silver RP, Alm RA, Guerry P. A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. Molecular Microbiology. 2001;40(3):769–77.

20. Garrett EM, Sekulovic O, Wetzel D, Jones JB, Edwards AN, Vargas-Cuebas G, et al. Phase variation of a signal transduction system controls *Clostridioides difficile* colony morphology, motility, and virulence. Blokesch M, editor. PLoS Biol. 2019 Oct 28;17(10):e3000379.

21. Edwards AN, McBride SM. Initiation of sporulation in *Clostridium difficile* : a twist on the classic model. FEMS Microbiol Lett. 2014 Sep;358(2):110–8.

22. Mullish BH, Williams HR. *Clostridium difficile* infection and antibiotic-associated diarrhoea. Clinical Medicine. 2018 Jun 1;18(3):237–41.

23. Prete RD, Ronga L, Addati G, Magrone R, Abbasciano A, Decimo M, et al. *Clostridium difficile.* A review on an emerging infection. La Clinica Terapeutica. 2019 Feb 14;1(170):e41–7.

24. CDC. Antibiotic-resistant Germs: New Threats [Internet]. Centers for Disease Control and Prevention. 2020 [cited 2020 Apr 8]. Available from: https://www.cdc.gov/drugresistance/biggest-threats.html

25. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, et al. Burden of *Clostridium difficile* Infection in the United States. New England Journal of Medicine. 2015 Feb 26;372(9):825–34.

26. Denève C, Janoir C, Poilane I, Fantinato C, Collignon A. New trends in *Clostridium difficile* virulence and pathogenesis. International Journal of Antimicrobial Agents. 2009 Mar 1;33:S24–8.

27. Paredes-Sabja D, Shen A, Sorg JA. *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends Microbiol. 2014 Jul;22(7):406–16.

28. Dürre P. Physiology and Sporulation in *Clostridium*. Microbiol Spectr. 2014 Aug;2(4):TBS-0010-2012.

29. Hall CL, Lee VT. Cyclic-di-GMP regulation of virulence in bacterial pathogens. Wiley Interdiscip Rev RNA. 2018;9(1).

30. Purcell EB, McKee RW, Courson DS, Garrett EM, McBride SM, Cheney RE, et al. A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls *Clostridium difficile* Biofilm and Toxin Production during Stationary Phase. Infect Immun. 2017;85(9).

31. McKee RW, Harvest CK, Tamayo R. Cyclic Diguanylate Regulates Virulence Factor Genes via Multiple Riboswitches in *Clostridium difficile*. mSphere. 2018 24;3(5).

32. Bordeleau E, Burrus V. Cyclic-di-GMP signaling in the Gram-positive pathogen *Clostridium difficile*. Curr Genet. 2015 Nov;61(4):497–502.

33. McKee RW, Mangalea MR, Purcell EB, Borchardt EK, Tamayo R. The Second Messenger Cyclic Di-GMP Regulates *Clostridium difficile* Toxin Production by Controlling Expression of sigD. Journal of Bacteriology. 2013 Nov 15;195(22):5174–85.

34. Purcell EB, McKee RW, McBride SM, Waters CM, Tamayo R. Cyclic Diguanylate Inversely Regulates Motility and Aggregation in *Clostridium difficile*. Journal of Bacteriology. 2012 Jul 1;194(13):3307–16.

35. McKee RW, Aleksanyan N, Garrett EM, Tamayo R. Type IV Pili Promote *Clostridium difficile* Adherence and Persistence in a Mouse Model of Infection. Infect Immun. 2018;86(5).

36. Purcell EB, McKee RW, Bordeleau E, Burrus V, Tamayo R. Regulation of Type IV Pili Contributes to Surface Behaviors of Historical and Epidemic Strains of *Clostridium difficile*. Christie PJ, editor. J Bacteriol. 2016 Feb 1;198(3):565–77.

37. Bordeleau E, Purcell EB, Lafontaine DA, Fortier L-C, Tamayo R, Burrus V. Cyclic Di-GMP Riboswitch-Regulated Type IV Pili Contribute to Aggregation of *Clostridium difficile*. Parkinson JS, editor. J Bacteriol. 2015 Mar 1;197(5):819–32.

38. Arato V, Gasperini G, Giusti F, Ferlenghi I, Scarselli M, Leuzzi R. Dual role of the colonization factor CD2831 in *Clostridium difficile* pathogenesis. Sci Rep. 2019 Dec;9(1):5554.

39. Peltier J, Shaw HA, Couchman EC, Dawson LF, Yu L, Choudhary JS, et al. Cyclic diGMP Regulates Production of Sortase Substrates of *Clostridium difficile* and Their Surface Exposure through ZmpI Protease-mediated Cleavage. J Biol Chem. 2015 Oct 2;290(40):24453–69.

40. Teng F, Murray BE, Weinstock GM. Conjugal Transfer of Plasmid DNA from*Escherichia colito Enterococci*: A Method to Make Insertion Mutations. Plasmid. 1998 May 1;39(3):182–6.

41. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol. 2010 Oct;192(19):4983–90.

42. Teng F, Murray BE, Weinstock GM. Conjugal transfer of plasmid DNA from *Escherichia coli to enterococci*: a method to make insertion mutations. Plasmid. 1998;39(3):182–6.

43. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010 Jan;80(1):49–55.

44. Fagan RP, Fairweather NF. *Clostridium difficile* Has Two Parallel and Essential Sec Secretion Systems. J Biol Chem. 2011 Aug 5;286(31):27483–93.

45. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, et al. Effect of tcdR Mutation on Sporulation in the Epidemic *Clostridium difficile* Strain R20291. mSphere. 2017 Feb;2(1).

46. Govind R, Fitzwater L, Nichols R. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. Journal of Bacteriology. 2015 Aug 1;197(15):2600–9.

47. Govind R, Dupuy B. Secretion of *Clostridium difficile* Toxins A and B Requires the Holin-like Protein TcdE. Cheung A, editor. PLoS Pathog. 2012 Jun 7;8(6):e1002727.

48. Bouillaut L, McBride SM, Sorg JA. Genetic Manipulation of *Clostridium difficile*. Curr Protoc Microbiol. 2011 Feb;0 9:Unit-9A.2.

49. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, et al. Characterization of the SigD Regulon of *C. difficile* and Its Positive Control of Toxin Production through the Regulation of tcdR. Nübel U, editor. PLoS ONE. 2013 Dec 16;8(12):e83748.

50. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The Key Sigma Factor of Transition Phase, SigH, Controls Sporulation, Metabolism, and Virulence Factor Expression in *Clostridium difficile*. Journal of Bacteriology. 2011 Jul 1;193(13):3186–96.

51. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, et al. Environmental Response and Autoregulation of *Clostridium difficile* TxeR, a Sigma Factor for Toxin Gene Expression. Journal of Bacteriology. 2002 Nov 1;184(21):5971–8.

52. Dupuy B, Sonenshein AL. Regulated transcription of *Clostridium difficile* toxin genes. Mol Microbiol. 1998 Jan;27(1):107–20.

53. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010 Jan;80(1):49–55.

54. Girinathan BP, Braun S, Sirigireddy AR, Lopez JE, Govind R. Importance of Glutamate Dehydrogenase (GDH) in *Clostridium difficile* Colonization In Vivo. PLOS ONE. 2016 Jul 28;11(7):e0160107.

55. Anjuwon-Foster BR, Tamayo R. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. Kearns DB, editor. PLoS Genet. 2017 Mar 27;13(3):e1006701.

56. Lipovsek S, Leitinger G, Rupnik M. Ultrastructure of *Clostridium difficile* colonies. Anaerobe. 2013 Dec;24:66–70.

57. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102.

58. Van de Putte P, Goosen N. DNA inversions in phages and bacteria. Trends in Genetics. 1992 Dec 1;8(12):457–62.

59. Zieg J, Simon M. Analysis of the nucleotide sequence of an invertible controlling element. Proc Natl Acad Sci U S A. 1980 Jul;77(7):4196–200.

60. Bordeleau E, Fortier L-C, Malouin F, Burrus V. c-di-GMP Turn-Over in *Clostridium difficile* Is Controlled by a Plethora of Diguanylate Cyclases and Phosphodiesterases. PLOS Genetics. 2011 Mar 31;7(3):e1002039.

61. Dineen SS, McBride SM, Sonenshein AL. Integration of Metabolism and Virulence by *Clostridium difficile* CodY. Journal of Bacteriology. 2010 Oct 15;192(20):5350–62.

62. Purcell EB, McKee RW, Courson DS, Garrett EM, McBride SM, Cheney RE, et al. A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls *Clostridium difficile* Biofilm and Toxin Production during Stationary Phase. Young VB, editor. Infect Immun. 2017 Sep;85(9):e00347-17, /iai/85/9/e00347-17.atom.

63. Belitsky BR, Sonenshein AL. Genome-wide identification of *Bacillus subtilis* CodY-binding sites at single-nucleotide resolution. PNAS. 2013 Apr 23;110(17):7026–31.

64. Sonenshein AL. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Current Opinion in Microbiology. 2005 Apr 1;8(2):203–7.

65. Emerson JE, Reynolds CB, Fagan RP, Shaw HA, Goulding D, Fairweather NF. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol. 2009 Nov;74(3):541–56.

66. Anjuwon-Foster BR, Maldonado-Vazquez N, Tamayo R. Characterization of Flagellum and Toxin Phase Variation in *Clostridioides difficile* Ribotype 012 Isolates. Zhulin IB, editor. J Bacteriol. 2018 May 7;200(14):e00056-18, /jb/200/14/e00056-18.atom.

67. Anjuwon-Foster BR, Tamayo R. Phase variation of *Clostridium difficile* virulence factors. Gut Microbes. 2018 Jan 2;9(1):76–83.

68. Tschowri N, Schumacher MA, Schlimpert S, Chinnam NB, Findlay KC, Brennan RG, et al. Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development. Cell. 2014 Aug 28;158(5):1136–47.

69. Liu X, Zheng G, Wang G, Jiang W, Li L, Lu Y. Overexpression of the diguanylate cyclase CdgD blocks developmental transitions and antibiotic biosynthesis in *Streptomyces coelicolor*. Sci China Life Sci. 2019 Nov 1;62(11):1492–505.

70. Gallagher KA, Schumacher MA, Bush MJ, Bibb MJ, Chandra G, Holmes NA, et al. c-di-GMP Arms an Anti-σ to Control Progression of Multicellular Differentiation in *Streptomyces*. Molecular Cell. 2020 Feb 6;77(3):586-599.e6.

71. Al-Bassam MM, Haist J, Neumann SA, Lindenberg S, Tschowri N. Expression Patterns, Genomic Conservation and Input Into Developmental Regulation of the GGDEF/EAL/HD-GYP Domain Proteins in *Streptomyces*. Front Microbiol. 2018;9:2524.

72. Hengst CDD, Tran NT, Bibb MJ, Chandra G, Leskiw BK, Buttner MJ. Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. Molecular Microbiology. 2010;78(2):361–79.

73. Latoscha A, Wörmann ME, Tschowri N. Nucleotide second messengers in *Streptomyces*. Microbiology. 2019 Nov 1;165(11):1153–65.

74. Bush MJ, Tschowri N, Schlimpert S, Flärdh K, Buttner MJ. c-di-GMP signalling and the regulation of developmental transitions in *streptomycetes*. Nat Rev Microbiol. 2015 Dec;13(12):749–60.

75. Weiss CA, Hoberg JA, Liu K, Tu BP, Winkler WC. Single-Cell Microscopy Reveals That Levels of Cyclic di-GMP Vary among *Bacillus subtilis* Subpopulations. O'Toole G, editor. J Bacteriol. 2019 May 28;201(16):e00247-19, /jb/201/16/JB.00247-19.atom.

76. Skotnicka D, Smaldone GT, Petters T, Trampari E, Liang J, Kaever V, et al. A Minimal Threshold of c-di-GMP Is Essential for Fruiting Body Formation and Sporulation in *Myxococcus xanthus*. PLOS Genetics. 2016 May 23;12(5):e1006080.

77. Girinathan BP, Ou J, Dupuy B, Govind R. Pleiotropic roles of *Clostridium difficile* sin locus. Koehler TM, editor. PLoS Pathog. 2018 Mar 12;14(3):e1006940.

78. Christen M, Christen B, Folcher M, Schauerte A, Jenal U. Identification and Characterization of a Cyclic di-GMP-specific Phosphodiesterase and Its Allosteric Control by GTP. J Biol Chem. 2005 Sep 2;280(35):30829–37.

79. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. CodY-Dependent Regulation of Sporulation in *Clostridium difficile*. Henkin TM, editor. J Bacteriol. 2016 Aug 1;198(15):2113–30.

80. Edwards AN, Nawrocki KL, McBride SM. Conserved Oligopeptide Permeases Modulate Sporulation Initiation in *Clostridium difficile*. Blanke SR, editor. Infect Immun. 2014 Oct;82(10):4276–91.

81. Navarro MVAS, De N, Bae N, Wang Q, Sondermann H. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. Structure. 2009 Aug 12;17(8):1104–16.

82. Wang R, Wang F, He R, Zhang R, Yuan J. The Second Messenger c-di-GMP Adjusts Motility and Promotes Surface Aggregation of Bacteria. Biophys J. 2018 04;115(11):2242–9.

83. Kharadi RR, Sundin GW. Physiological and Microscopic Characterization of Cyclic-di-GMP-Mediated Autoaggregation in *Erwinia amylovora*. Front Microbiol. 2019;10:468.

3.7 Supplemental data

3.7.1 Supplementary methods

3.7.1.1 ClosTron mutagenesis of UK1 *recV*

The UK1::*recV* mutant in UK1 strain was created using ClosTron gene knockout system as described previously (53). Briefly, the group II intron insertion site at nucleotides 144a in *recV* gene in the antisense orientation was selected using a web-based design tool called the Perutka algorithm. The designed retargeted intron was cloned into pMTL007-CE5 as described previously (47,54). The resulting plasmid pMTL007-CE5::*recV*-144a was transferred into *C. difficile* UK1 cells by conjugation. The potential Ll.ltrB insertions within the target genes in the *C. difficile* chromosome was conferred by the selection of lincomycin resistant transconjugants in 20 μg/ ml lincomycin plates. PCR using gene-specific primers (Table S 3.2) in combination with the EBS-U universal was performed to identify putative *C. difficile* mutants.

3.7.1.2 Transmission Electron Microscopy (TEM)

Cell pellets from 1 ml of *C. difficile* strains that were grown for 16 hr were fixed overnight with 2% glutaraldehyde and 2% paraformaldehyde in 1X PBS. The cells were washed with 1X PBS. Cells were resuspended with 1% osmium tetroxide and incubated for 2 hour at room temperature with constant rotation. The samples were washed X3 times with 1X PBS, followed by an incubation for 1 hour with 2% Uranyl acetate in water. After washing the pellets several times with distilled water, cells were treated with acetone series (50% -100% vol/vol) for 5 minutes each and infiltrated in graded EMBED 812/Araldite resin (Electron Microscopy Sciences) at room temperature with rotation. The resin was cut into thin sections and were stained with 2% alcoholic

uranyl acetate and Reynolds' lead citrate on a copper grids. Sections were examined with a transmission electron microscope (Philips CM100) and images taken for analysis.

3.7.2 Supplementary tables

Table S 3.2. Oligonucleotides used for PCR reactions

Name	Sequence (5' → 3')	Description
EBS-U	CGAAATTAGAAACTTGCGTTCAGT AAAC	Group II intron specific primer
ORG771	GACTGAGCTCTGGGGAAATGTGTTTG ATGAGATAAAAAAATTAGTAAAATA	Forward primer with with SacI to clone *CDR20291_0685* (*pdcB*) in pRPF185
ORG772	AACTGGATCCTCACTTATCGTCGTCAT CCTTGTAATCCTTTGATAGTTGAAATT TATAAAATTCAGAGGCAC	Reverse primer with with BamHI to clone *CDR20291_0685* (*pdcB*) in pRPF185
ORG788	GAGCTCAAGAGGAGTGGTTGAAAATG GCAACAAGACCTATAGAAATAG	Forward primer with SacI to amplify *recV*
ORG789	GGATCCTTAATGATGATGATGATGATG ACCAATAAAGAAATTTTCACTAGCTTC ATTAATAGCTTGCTGAG	Reverse primer with BamHI and 6X His to amplify *recV*
ORG844	GTTGTAAAAAAGTTACTATTTATTGAA AATTTAGATACTTTTCTAAAATTATCT ATG	Forward primer for CodY binding site in *pdcB* upstream region in Opaque orientation
ORG845	CATAGATAATTTTAGAAAAGTATCTAA ATTTTCAATAAATAGTAACTTTTTTAC AAC	Reverse primer for CodY binding site in *pdcB* upstream region in Opaque orientation
ORG852	GGTACCAGTTTAGGATAAAGTATTGC AAGAACCAATCAG	Forward primer with KpnI site to clone ~1.5 kb upstream region of *pdcB* gene
ORG853	GAGCTCCTTTTCCCCCTACAATATTAC TATTAGTGTAGTTTAATCAAC	Reverse primer with SacI site to clone ~1.5 kb upstream region of *pdcB* gene
ORG866	ATGTATATTTTTATAGCTATTTTATTAT AAAAAAAAAGAACCCTCGCATGTGTA AGGGTTACTTTTTTTCTCTATTTTTTTT ATTATATCACTATTTTTTCCCTATTTCA ATATTTTGACATTTTTTCCATTATACTG A	Forward primer to mutagenize inverted repeat from translucent orientation
ORG867	TCAGTATAATGGAAAAAATGTCAAAA TATTGAAATAGGGAAAAAATAGTGAT ATAATAAAAAAAATAGAGAAAAAAA GTAACCCTTACACATGCGAGGGTTCTT TTTTTTTATAATAAAATAGCTATAAAA ATATACAT	Reverse primer to mutagenize inverted repeat from translucent orientation
ORG868	CATAAATTTAACTATTAACATAGATAA TTTTAGAAAAGTATCTAAATTTTCAAT AAATAGTAACTTTCGCATGTGTAAGG GTTACTTTTTTTCTCTATTTTTTTTATTA TATCACTATTTTTTCCCTATTTCAATAT TTTGAC	Forward primer to mutagenize inverted repeat from opaque orientation
ORG869	GTCAAAATATTGAAATAGGGAAAAAA TAGTGATATAATAAAAAAAATAGAGA	Reverse primer to mutagenize inverted repeat from opaque orientation

	AAAAAAGTAACCCTTACACATGCGAA AGTTACTATTTATTGAAAATTTAGATA CTTTTCTAAAATTATCTATGTTAATAG TTAAATTTATG	
ORG879	GTATTATTTTGGTAAATATATTGTTAC AAAAGGTTTATATTTTGC	Forward primer (*cmrRST* upstream): specific for orientation similar to published (410 bps)
ORG880	GGAAATATTGACAAAATAATATTACA ATGTTAGAATAA	Reverse primer common for both orientation or *cmrRST*.
ORG881	AGTATAATGCTATTATAATAAGAAAA TAACTTTTTTATAAACATTGAGAT	Forward primer (*cmrRST* upstream): specific for flipped orientation (540 bps)
ORG882	GAGCTCGGGGGAAAAGATGTATCCAG AAGATGGAGATAATTATTTAGATTTAT TTAAACA	Forward primer with SacI to clone PdcB-EAL in pRPF185
ORG883	GGATCCTCACTTTGATAGTTGAAATTT ATAAAATTCAGAGGCACTTACAGGTC TTCC	Reverse primer with BamHI to clone PdcB-EAL in pRPF185
ORG884	GAGCTCGGAGGAGATAAGATGCAAGA AATATTAAAAAATAAAATGTATTCAT	Forward primer with SacI to clone PdcA-EAL in pRPF185
ORG885	GGATCCTTAATTATCTAGCTTTAAAAG GTCAAAGATTTCTGTTGCTGTCTGCGG TTTTCC	Reverse primer with BamHI to clone PdcA-EAL in pRPF185
ORG886	TAATAAAATAGCTATAAAAATATACA TATCAAATCAAATTAAG	Forward primer to amplify the flip orientation of *pdcB* upstream in Translucent strain.
ORG887	CTTCTTAATTTGATTTGATATGTATATT TTTATAGC	Forward primer to amplify the flip orientation of *pdcB* upstream in Opaque strain.
ORG888	CTGGATTTTTAAATTTATGTACTTTAA AATGGATATCC	Reverse primer to amplify the flip orientation both translucent and opaque of *pdcB* upstream.
ORG846	GGTCTAAACTTAAAGAGTTCGATTCTT CTCTCATCTTTTCCCCCTACAATATTAC TATTAGTG	SP1 primer for 5’ RACE
ORG921	ATGTTAATAGTTAAATTTATGAACATT TTTTGTTTT	Forward primer to amplify *pdcB* promoter_1
ORG922	ATGTTAGGATATTTCTTAGAAATGAAA TACTTCTTAATTTG	Forward primer to amplify *pdcB* promoter_2
ORG925	ACTATTTATTGAAAATTTAGATACTTT TCTAAAATTATCTATG	Forward primer to amplify *pdcB* promoter_3
ORG926	ATACTTTTCTAAAATTATCTATGTTAA TAGTTAAAT	Forward primer to amplify *pdcB* promoter_3

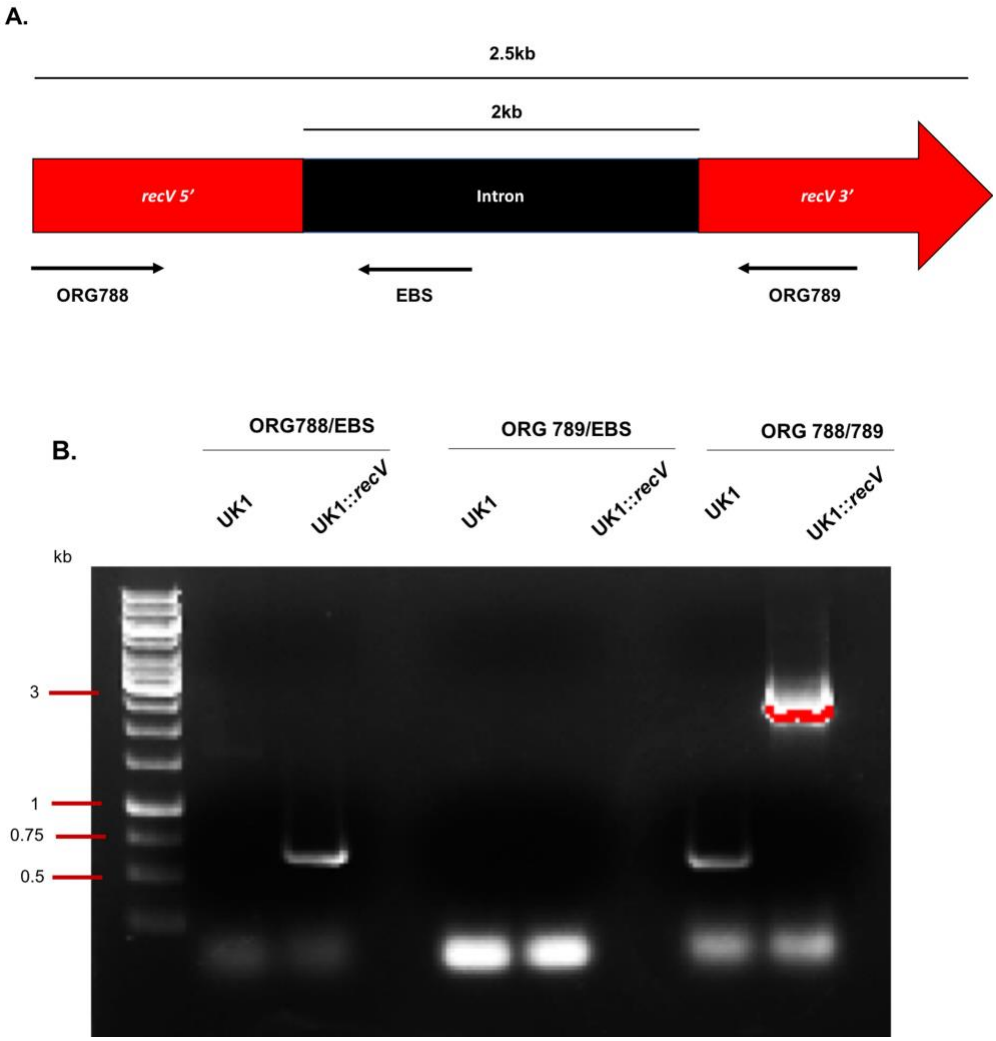
Table S 3.3. Oligonucleotides used for qRT-PCR reactions

Primer	Sequence (5' → 3')	Gene target
(Forward)	GAGGAGAGTGGAATTCCTAGTGTAG	*16srRNA*
(Reverse)	GGACTACCAGGGTATCTAATCCTGT	*16srRNA*
(Forward)	ATTGCACAGTGACTGCAGGAGTTGC	*pdcB*
(Reverse)	GTAACAGCATTTACCTGGGGTTGAAA	*pdcB*
(Forward)	CCAGTTTGACCATCTGGTGT	*pilA*
(Reverse)	CAGTAGTGGCAGTTCCAGCTT	*PilA*

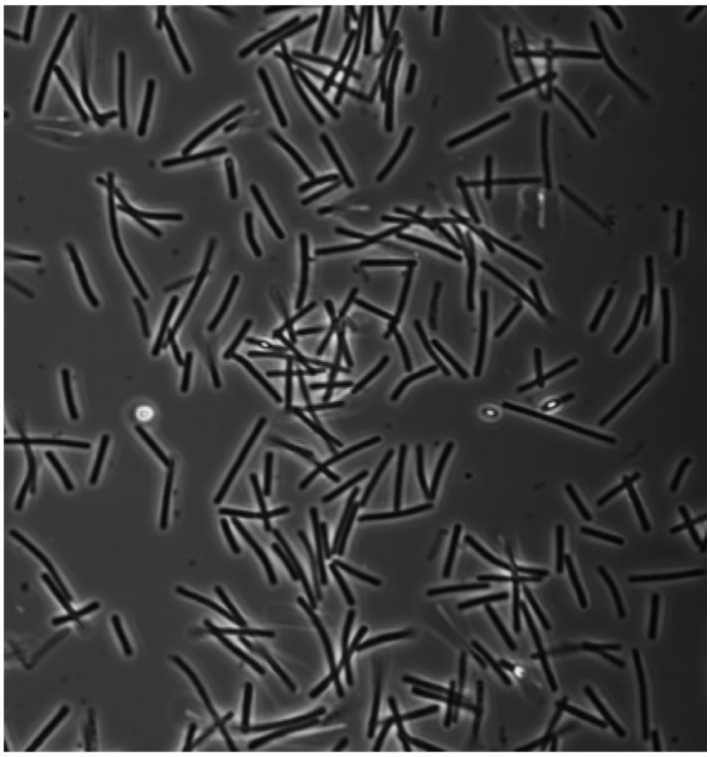
3.7.3 Supplementary figures

Figure S 3.1. Construction and confirmation of the recV mutation in *C. difficile* UK1 strain

A. Schematic representation of ClostTron (group II intron) mediated insertional inactivation of *recV* gene in *C. difficile*. **B**. PCR verification of the intron insertion in *recV in* UK1 strain, conducted with intron-specific primer EBS universal [EBS(U)] with *recV* specific primers ORG 788 and ORG 789. **C.** Phase contrast microscopy image of UK1::*recV* mutant strain. **D.** Percentage sporulation of UK1::*recV* mutant compared to UK1_T and UK1_O strains.



C.



UK1::*recV*

D.

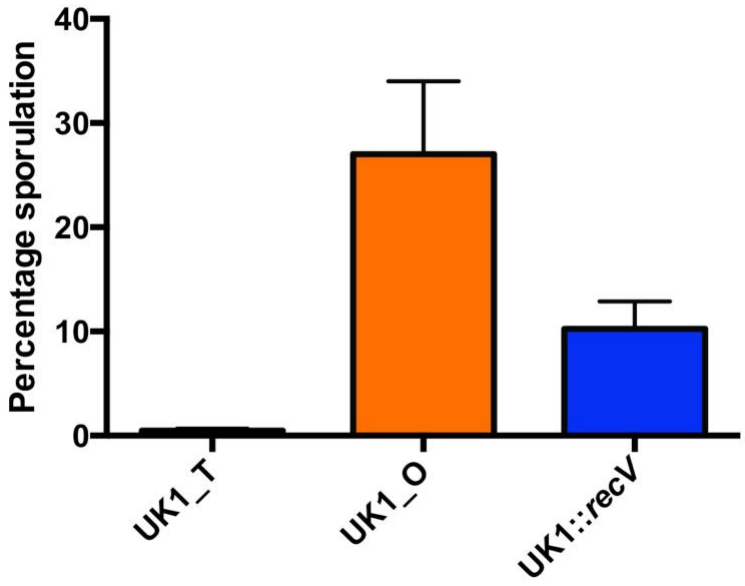
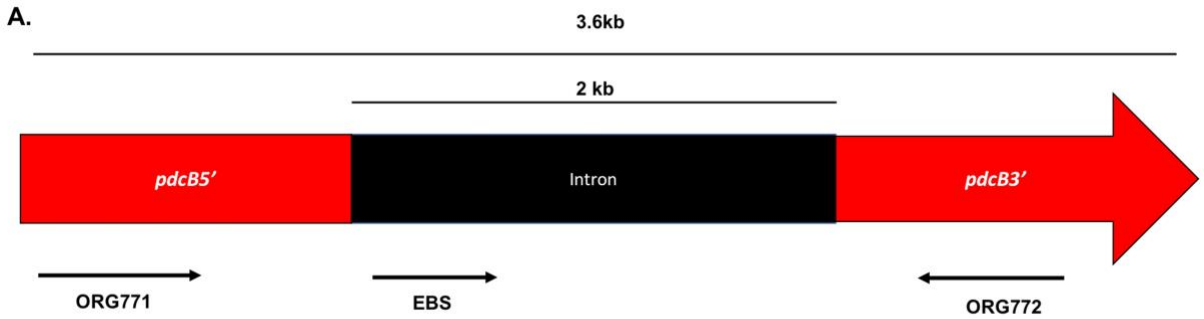


Figure S 3.2. Construction and confirmation of the pdcB mutation in *C. difficile* UK1 strain

A. Schematic representation of ClostTron (group II intron) mediated insertional inactivation of *pdcB* gene in *C. difficile*. **B.** PCR verification of the intron insertion in *pdcB in* UK1 strain, conducted with intron-specific primer EBS universal [EBS(U)] with *pdcB* specific primers ORG 771 and ORG 772.



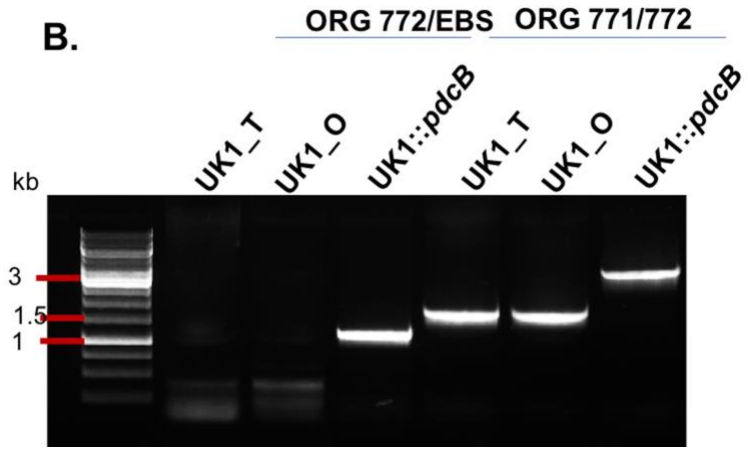


Figure S 3.3. Relative expression of pilA in UK1_T and UK1_O strains

qRT-PCR results showing overexpression of *pilA* in UK1_T strain. The representative results from three independent experiments are shown. The asterisks (**) indicate statistical difference at a P value of <0.005.

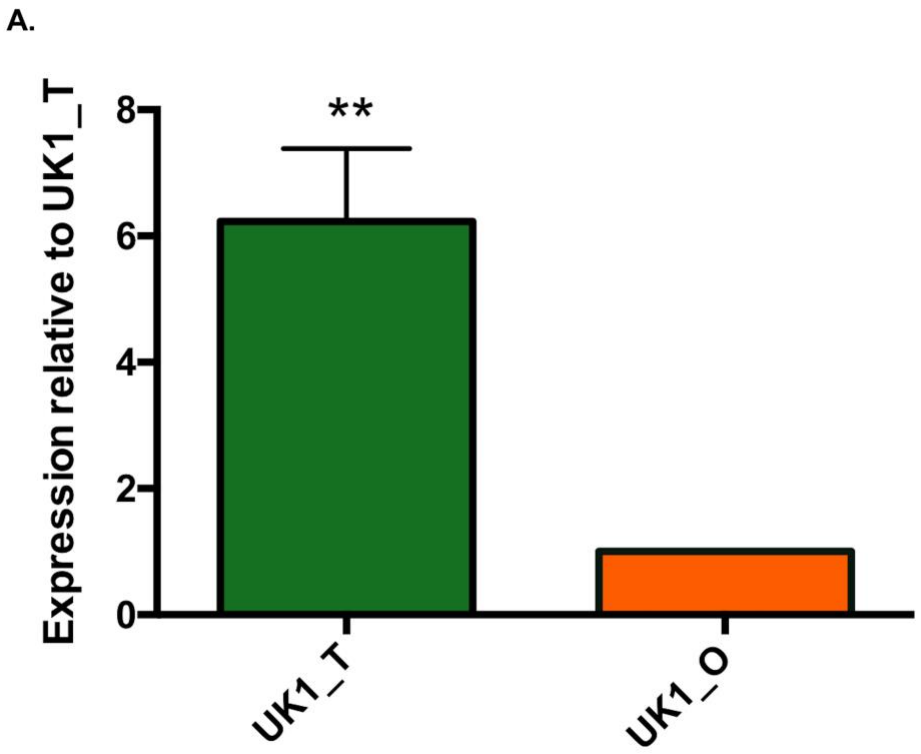


Figure S 3.4. Transmission electron microscope image of UK1_T cross section showing a septum between two dividing cells

A septum was observed in UK1_T cells, shown in red triangles, when observed through transmission electron microscope. TEM image was taken by Dr. Dan Boyle (K-State).

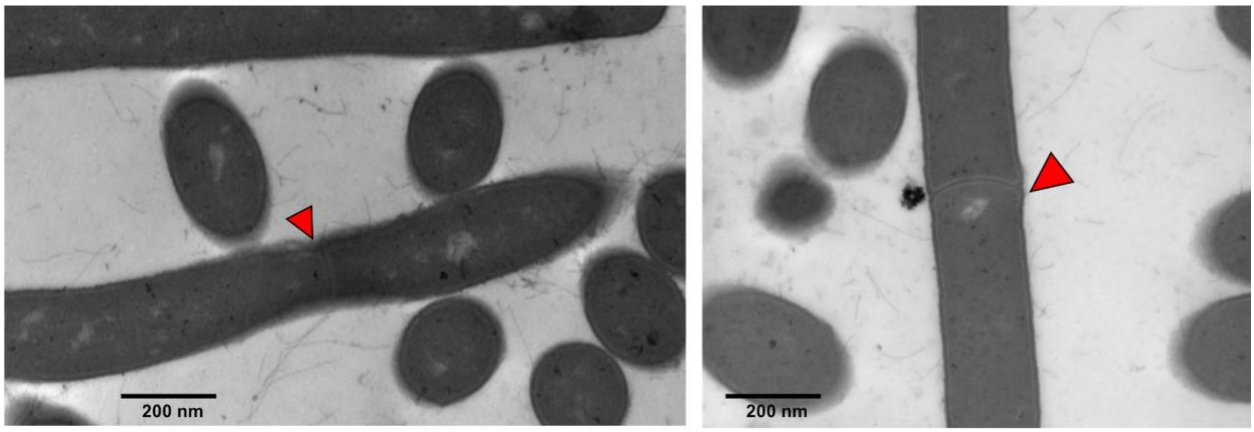
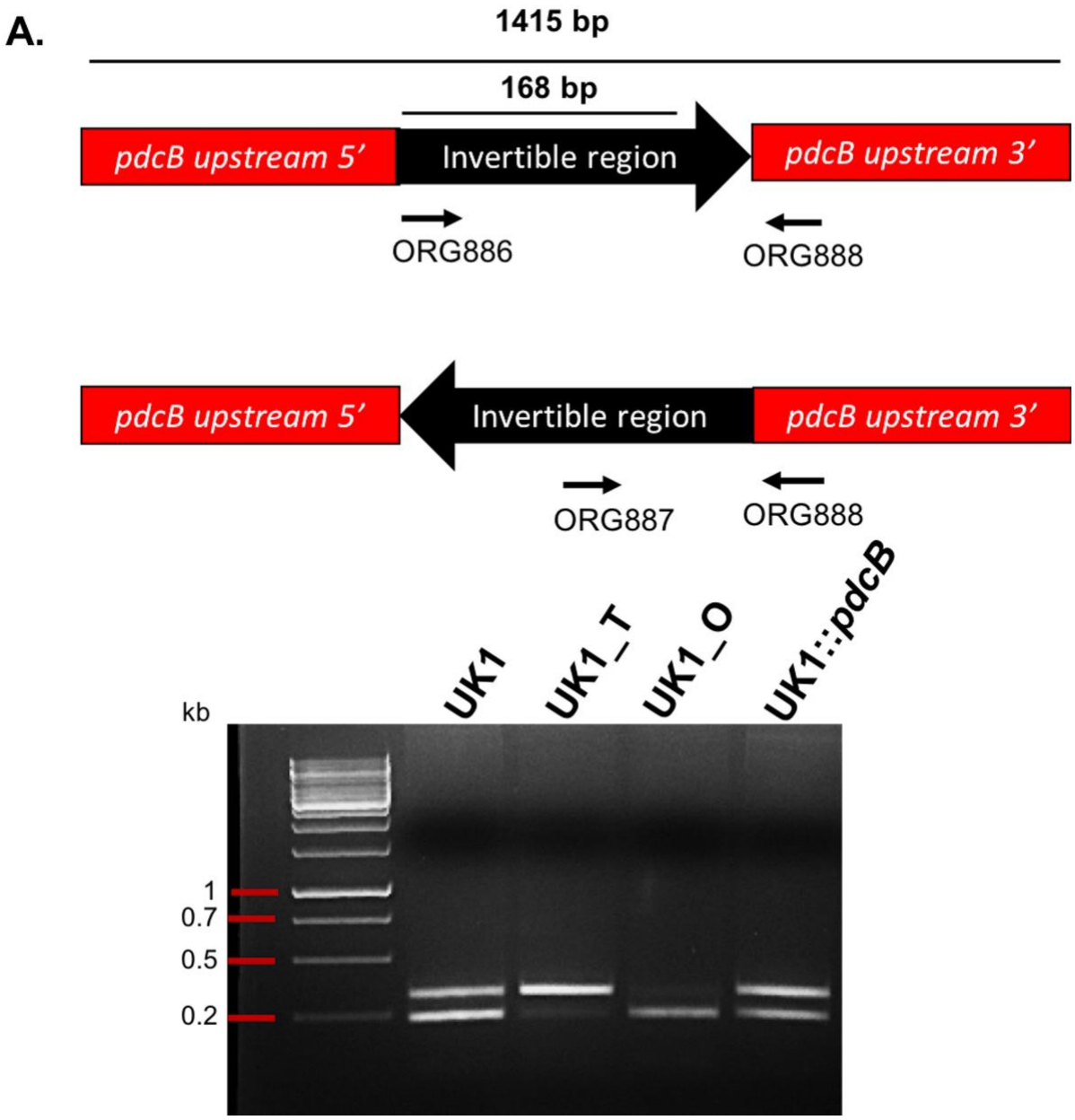
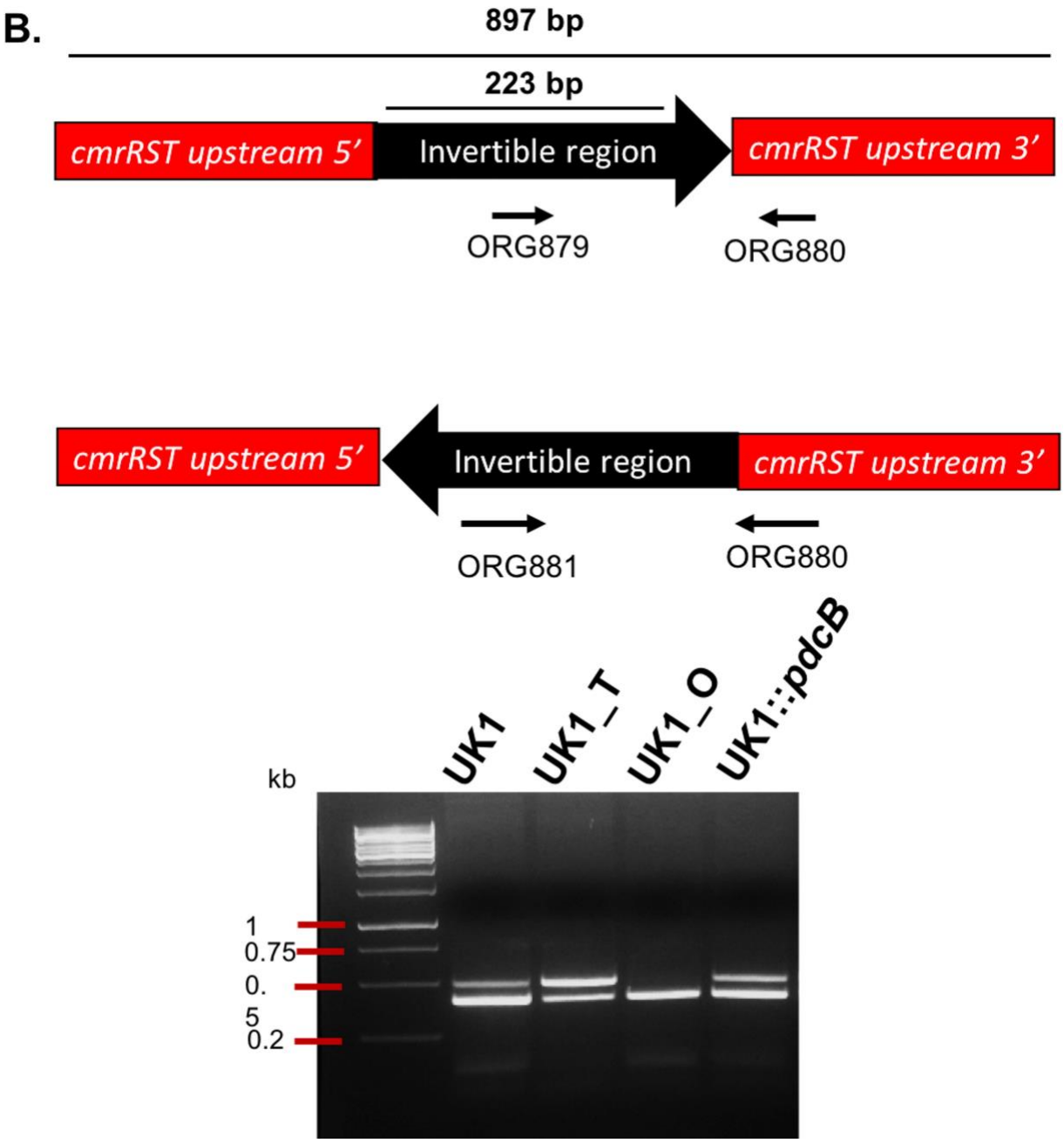


Figure S 3.5. Orientation specific PCR amplification of pdcB and cmrRST upstream

A.PCR amplification of the upstream region of *pdcB* using primer mixture containing primers ORG886/887/888. ORG886/888 resulted in a product size of 325 bp while ORG 887/888 resulted in a product size of 215 bp. **B.** PCR amplification of the upstream region of *cmrRST* using primer mixture containing primers ORG879/880/881. ORG879/880 resulted in a product size of 410 bp while ORG 880/881 resulted in a product size of 540 bp.




B.
897 bp
223 bp
cmrRST upstream 5'
Invertible region
cmrRST upstream 3'
ORG879
ORG880
cmrRST upstream 5'
Invertible region
cmrRST upstream 3'
ORG881
ORG880
UK1
UK1_T
UK1_O
UK1::pdcB
kb
1
0.75
0.
5
0.2

Chapter 4 - Conclusion

4.1 Spo0A suppresses sin locus expression in *C. difficile*

4.1.1 Major findings

- The work presented in Chapter 2 describes the role of Spo0A in regulating *sinR* expression. We created *spo0A* mutant in three different *C. difficile* strains and through western blot analysis we demonstrated that all *spo0A* mutants had higher SinR levels as compared to the respective parent strains.
- By carrying out genetic and biochemical assays we have shown that Spo0A directly binds to the upstream region of *sin* locus and represses the expression of both *sinR and sinR'.*

Previous independent studies have shown that SinR positively regulates motility and toxin production in *C. difficile* (1) and Spo0A negatively regulates motility and toxin production (2,3). Current study has bridged those two studies and demonstrated that Spo0A indirectly controls toxin production and motility by regulating SinR (Fig 4.1).

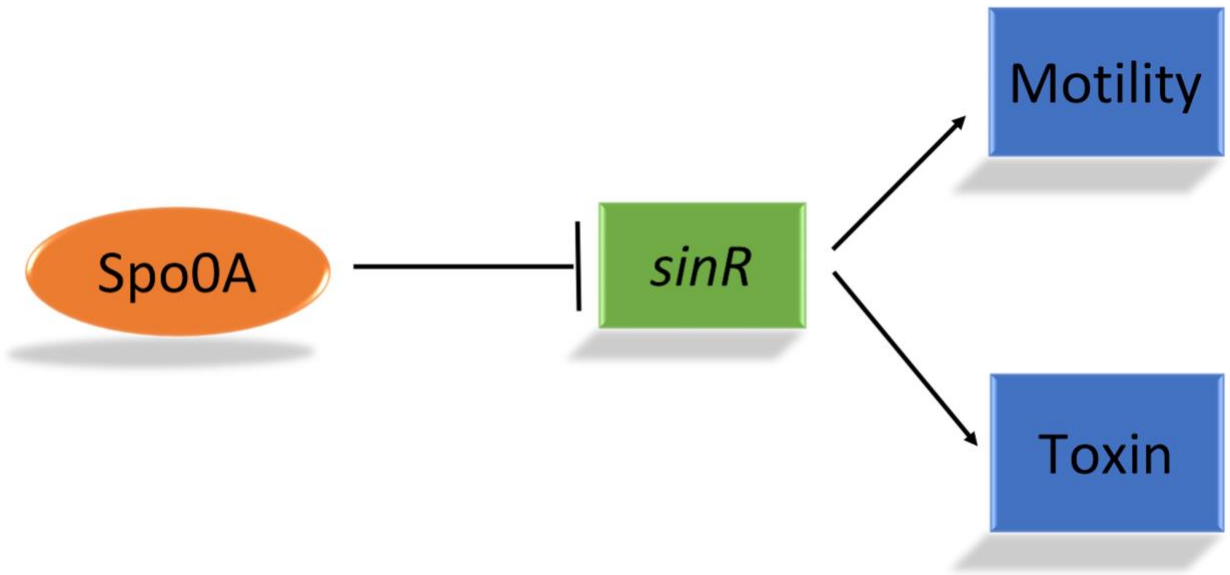


Figure 4.1. Spo0A regulates motility and toxin production in *C. difficile* via SinR

Schematic showing the overall contribution of Chapter 2. We have shown that Spo0A acts via SinR to regulate motility and toxin in *C. difficile*. Spo0A directly binds to the promoter region of *sinR* and represses its expression, thus negatively regulating motility and toxin production.

4.1.2 Future direction

Our data has shown the region where Spo0A can bind to in the upstream of *sin* locus. Spo0A is a transcriptional regulator and binds to the consensus sequences in the upstream of the target gene which are commonly known as the "0A box" . Work to determine the "0A box" in the upstream of *sinR* is currently going in our lab. Since SinR is the pleitropic regulator of sporulation, motility and toxin production, it is highly likely that there are multiple regulators involved in regulating the expression of *sinR.* Previous work from our lab has shown that purified CodY can directly bind to the upstream of SinR and repress its expression. Investigating the interplay between Spo0A and other regulators of *sinR* like CodY will provide more information on the Spo0A mediated repression of *sinR*. Similarly, previous work from our lab has shown that *sinR* mutant is asporogenic with reduced *spo0A* transcript but the exact mechanism by which SinR regulates sporulation is not understood. Future investigation to determine the pathway by which SinR regulates Spo0A will shed light on the regulatory relationship between these two important regulators. It has been shown that SinR binds to the upstream of *codY*. Investigating whether SinR is transcriptional regulator and has consensus binding site in the upstream of the target gene would shed light on the mechanism of SinR mediated regulation of toxin, motility and sporulation.

4.2 Phase variable expression of *pdcB* and the role of intracellular c-di-GMP in sporulation of *C. difficile*

4.2.1 Major findings

The work presented in chapter 3 describes the role of *pdcB* in sporulation of *C. difficile.*

- In this work, we first observed two colony morphotypes of UK1 strain, UK1_T and UK1_O, with hyper sporulation phenotype in UK1_O strain.
- Whole genome sequencing of these morphotype identified the upstream region of *pdcB* as a potential candidate contributing to the underlying genetic difference between UK1_T and UK_O.
- We found DNA inversion in 168 bp region in the upstream of *pdcB* in UK1_O strain. Further investigation demonstrated the DNA inversion upregulates the expression of *pdcB.*
- We also demonstrated that overexpression of *pdcB* reduced the concentration of intracellular c-di-GMP in UK1_O strain and by complementation of UK1::*pdcB* mutant we demonstrated that hyper sporulation in UK1_O is directly associated with intracellular c-di-GMP concentration.
- We demonstrated that CodY binds directly to the consensus site present within the invertible region and represses the expression of *pdcB.*

This study has added a new player, *pdcB,* in the gene regulatory network of sporulation and demonstrated a novel link between intracellular c-di-GMP and sporulation in *C. difficile*. Additionally, by showing that *pdcB* is regulated by CodY, we have redefined the role of CodY in sporulation of *C. difficile* (Fig. 4.2).

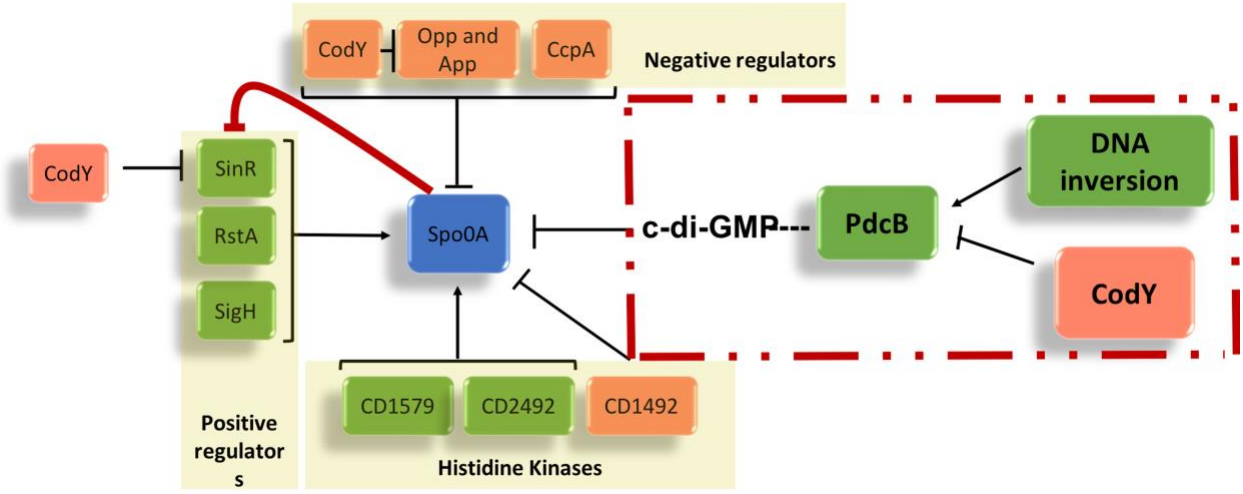


Figure 4.2. Updated gene regulatory network of *C. difficile* highlighting the contribution from our study in Chapter 2 and 3

The study from Chapter 3 has added new player, PdcB, in the gene regulatory network of sporulation in *C. difficile*, as highlighted by the red dotted rectangle. Our study has demonstrated that PdcB, by modulating the intracellular c-di-GMP concentration, positively regulates sporulation. We have also shown that *pdcB* expression is regulated by DNA inversion and by direct binding of CodY.

4.2.2 Future direction

Our work is the first to associate the role of c-di-GMP in sporulation initiation in any *Clostridium* species including *C. difficile*. This has opened a new avenue for future research to investigate many fundamental aspects of *C. difficile* sporulation. Our work has also shown phenotypic heterogeneity in *C. difficile* UK1 strain with hyper-sporulation in UK1_O strain resulting from reduced c-di-GMP concentration. Since sporulation has an implication in persistence and transmission of the disease, determining the pathogenic significance of these two colony morphotypes *in vivo* would shed more light to the importance of intracellular c-di-GMP mediated regulation of sporulation. c-di-GMP has been mostly associated with transition from sessile to motile state. *In vivo* studies will provide information which may prove useful in investigating the role of c-di-GMP in colonization of *C. difficile* which is the first step in pathogenesis, yet is under explored in *C. difficile.* Since the intracellular levels of c-di-GMP appear to regulate sporulation and other pathogenic factors in *C. difficile*, it is vital to understand the regulation of enzymes that maintain c-di-GMP homeostasis in bacterial cells. Our work has shown that *pdcB,* has dual regulation by DNA inversion and CodY. We have shown that CodY represses the expression of *pdcB* by directly binding to the consensus site within the invertible region. Since CodY and PdcB both require GTP for their activity, the regulatory relationship between these two proteins needs to be determined. It would be interesting to know whether PdcB indirectly regulates CodY activity by controlling the intracellular GTP content. Our reverse transcriptase PCR analysis have located the promoter region of *pcdB* to be downstream of CodY binding region (in the opaque orientation), but within the invertible region. 5' RACE experiment to determine the transcription initiation site will help to precisely map the promoter and this assay is currently ongoing in our

lab. Similarly, recombinase that carries out the inversion of *pcdB* upstream and its regulation has to be investigated to further understand the gene regulatory network that regulates sporulation.

4.3 References

1. Girinathan BP, Ou J, Dupuy B, Govind R. Pleiotropic roles of *Clostridium difficile* sin locus. Koehler TM, editor. PLoS Pathog. 2018 Mar 12;14(3):e1006940.

2. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. Spo0A Differentially Regulates Toxin Production in Evolutionarily Diverse Strains of *Clostridium difficile*. Popoff MR, editor. PLoS ONE. 2013 Nov 13;8(11):e79666.

3. Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. *C. difficile 630Δerm* Spo0A Regulates Sporulation, but Does Not Contribute to Toxin Production, by Direct High-Affinity Binding to Target DNA. Chang Y-F, editor. PLoS ONE. 2012 Oct 31;7(10):e48608.

Appendix A - Neuronal ceroid lipofuscinosis related ER membrane protein CLN8 regulates PP2A activity and ceramide levels

BBA - Molecular Basis of Disease 1865 (2019) 322–328

Contents lists available at ScienceDirect

BBA - Molecular Basis of Disease

journal homepage: www.elsevier.com/locate/bbadis

Neuronal ceroid lipofuscinosis related ER membrane protein CLN8 regulates PP2A activity and ceramide levels

Babita Adhikari[a,1], Bhagya De Silva[b,c,1,2], Joshua A. Molina[a], Ashton Allen[a], Sun H. Peck[a,3], Stella Y. Lee[a,b,*]

[a] Division of Biology, Kansas State University, Manhattan, KS 66506, USA
[b] Graduate Biochemistry Group, Kansas State University, Manhattan, Kansas 66506, USA
[c] Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, Kansas 66506, USA

ARTICLE INFO

Keywords:
CLN8
PP2A
I2PP2A
Ceramide
Neuronal ceroid lipofuscinosis

ABSTRACT

The neuronal ceroid lipofuscinoses (NCLs) are a group of inherited neurodegenerative lysosomal storage disorders. CLN8 deficiency causes a subtype of NCL, referred to as CLN8 disease. CLN8 is an ER resident protein with unknown function; however, a role in ceramide metabolism has been suggested. In this report, we identified PP2A and its biological inhibitor I2PP2A as interacting proteins of CLN8. PP2A is one of the major serine/threonine phosphatases in cells and governs a wide range of signaling pathways by dephosphorylating critical signaling molecules. We showed that the phosphorylation levels of several substrates of PP2A, namely Akt, S6 kinase, and GSK3β, were decreased in CLN8 disease patient fibroblasts. This reduction can be reversed by inhibiting PP2A phosphatase activity with cantharidin , suggesting a higher PP2A activity in CLN8-deficient cells. Since ceramides are known to bind and influence the activity of PP2A and I2PP2A, we further examined whether ceramide levels in the CLN8-deficient cells were changed. Interestingly, the ceramide levels were reduced by 60% in CLN8 disease patient cells compared to controls. Furthermore, we observed that the conversion of ER-localized NBD-C6-ceramide to glucosylceramide and sphingomyelin in the Golgi apparatus was not affected in CLN8-deficient cells, indicating transport of ceramides from ER to the Golgi apparatus was normal. A model of how CLN8 along with ceramides affects I2PP2A and PP2A binding and activities is proposed.

1. Introduction

The neuronal ceroid lipofuscinoses (NCLs) are a group of neurodegenerative lysosomal disorders that predominantly affect children [1,2]. They lead to severe pathological conditions such as progressive loss of motor neuron functions, loss of vision, mental retardation, epilepsy, ataxia, and eventually premature death. Thirteen genetically distinct subtypes of the NCLs have been identified to date [3]. Intriguingly, these genes encode a variety of unrelated proteins (ceroid-lipofuscinosis neuronal proteins, CLNs) that are localized to various cellular compartments. One of the CLN proteins, CLN8, is an ER resident protein with five predicted transmembrane regions and a C-terminal di-lysine ER retrieval signal [4].

The first hint that the function of CLN8 involves ceramide comes from a conserved domain of ~200 amino acids in CLN8 named TLC. The TLC domain was originally identified in the TRAM, Lag1, and CLN8 proteins [5]. It is postulated that the function of this domain is lipid binding/sensing. The human orthologues of yeast Lag1 is ceramide synthase family protein Lass1–6, later renamed CerS1–6 [6]. Even though the purported ceramide synthase active site is located within the TLC domain, and CLN8 contains two conserved histidine residues essential for the synthase activity, CLN8 does not possess such activity [7]. Decreases in ceramide levels have been observed in CLN8 disease patient cerebral samples and the CLN8 disease *mnd*-mouse model [8,9]. However, how CLN8 is linked to ceramide metabolism is not clear. Besides CLN8, there are also reports suggesting a role for sphingolipid metabolism in other CLN proteins. For instance, CLN3, an *endo*-lysosomal multi-spanning transmembrane protein, has been shown to bind galactosylceramide and this interaction is required for both galactosylceramide and CLN3 trafficking out from the Golgi [10,11]. Altered

* Corresponding author at: Division of Biology, Kansas State University, Ackert Hall 308, 1717 Claflin Road, Manhattan, KS 66506, USA.
E-mail address: sylee@ksu.edu (S.Y. Lee).
[1] These authors contributed equally.
[2] Present address: Department of Chemistry and Biochemistry, University of Nevada, Las Vegas, NV 89154, USA.
[3] Present address: University of Pennsylvania, Department of Orthopaedic Surgery, Philadelphia, PA 19104, USA.

https://doi.org/10.1016/j.bbadis.2018.11.011
Received 3 August 2018; Received in revised form 9 November 2018; Accepted 15 November 2018
Available online 16 November 2018
0925-4439/