PHASE VARIATION OF *CLOSTRIDIUM DIFFICILE* VIRULENCE FACTORS

Brandon R. Anjuwon-Foster

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor in Philosophy in the Department of Microbiology and Immunology in the School of Medicine.

Chapel Hill
2018

Approved by:

Rita Tamayo

Peggy Cotter

Virginia Miller

Edward Miao

Matthew Wolfgang

ABSTRACT

Brandon R. Anjuwon-Foster: Phase Variation of *Clostridium difficile* Virulence Factors
(Under the direction of Rita Tamayo)

Clostridium difficile is a Gram-positive spore-forming anaerobe and the leading cause of antibiotic-associated diarrheal disease in the United States. *C. difficile* produces two toxins, TcdA and TcdB, that are necessary for diarrheal disease symptoms. Colonization of the intestine is a necessary prerequisite to diarrheal disease symptoms. *C. difficile* produces flagella that aid not only in bacterial motility, but adherence to intestinal tissue. SigD, a flagellar alternative sigma factor in the early stage flagellar (*flgB*) operon, indirectly activates expression of the *tcdA* and *tcdB* genes. Both flagella and toxins are *C. difficile* virulence factors that synergistically promote diarrheal disease symptoms, pathology, and inflammation. Therefore, factors that regulate expression of the *flgB* operon affect not only motility, but toxin production and the virulence of *C. difficile*. The main objective of the research described in this dissertation was to identify and characterize a genetic mechanism controlling co-regulated flagellar and toxin gene expression in *C. difficile*. In Chapter 2, we identified a "flagellar switch" located upstream of the *flgB* operon, that mediates the phase variable production of flagella and toxins in *C. difficile*. Bacteria with the sequence in one orientation produced flagella, were motile and secreted the toxins ("*flg* ON"). Bacteria with the sequence in the inverse orientation were aflagellate and showed decreased toxin secretion ("*flg* OFF"). We determined that the tyrosine recombinase RecV is required for inversion of the flagellar switch in both directions. In Chapter 3, we found a single strain family, designated as "ribotype 012", of *C. difficile* exhibits low frequency inversion

of the flagellar switch in laboratory-adapted, environmental, and clinical isolates. In Chapter 4, we demonstrated that Rho factor is required for flagellar phase variation in *C. difficile*. We hypothesize that Rho factor directly terminates transcription in the leader RNA of the *flgB* operon in *flg* OFF bacteria. Future studies will assess the virulence contribution of flagellar and toxin phase variation to host infection dynamics, outcome, and transmission. Phase variable flagellar motility and toxin production suggests that these important virulence factors have both advantageous and detrimental effects during the course of infection.

In memory of my biggest supporters, Eulalia Pannell Foster and Margaret D. Foster Washington.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Rita Tamayo for her unconditional support and mentorship throughout this journey. Rita has devoted time and energy on numerous occasions to give constructive oral and written feedback on my ideas to ensure I develop into an inquisitive and thoughtful independent scientist. My dissertation project started with trying to find a "needle in a haystack" phenotype, but through the scientific challenges Rita remained patient. Rita has dramatically helped improve my writing as a scientist through constructive, iterative feedback on various manuscripts and essays. Joining Rita's lab also afforded me the opportunity to observe a young underrepresented minority navigate the path to professorship and establish herself in bacterial pathogenesis. Minorities in STEM often have a shared experience of having to work twice as hard to get half as far as their privileged colleagues. Dr. Tamayo has taught me you can make tremendous strides in research, academia, and life without having to sacrifice who you are as a person nor your culture to be successful.

I'd like to thank past and present members of the Tamayo Lab for feedback in lab meetings, thoughtful scientific discussions, and jovial banter in dealing with the ups and downs of research and life as a graduate student. I'd like to especially thank former colleague Dr. Ankunda Kariisa for not only sharp scientific feedback, but for always keeping me entertained with jokes and stories. Our shared laughter could often be heard echoing through labs and hallways in Mary Ellen Jones and Marisco Hall much to the amusement, or irritation, of others. I'd especially like to thank Dr. Robert McKee for insightful scientific discussions and keeping everyone entertained with his quick-wit and indelible puns.

I'd like to thank past and present students, postdoctoral fellows, faculty members, and staff of the Department of Microbiology and Immunology for my training both professionally and personally. I'd like to especially acknowledge future Dr. Brittany Miller, prodigy of the SecA2 OG', for entertaining and thoughtful discussions on protein chaperones and life. I'd like to thank my former late night and weekend lab partner in crime Dr. Nicholas Vitko for support through challenging experiments and lighthearted discussions about life. To my very patient dissertation committee for providing constant feedback and support, I thank Drs. Peggy Cotter, Virginia Miller, Edward Miao, Matt Wolfgang and former members Anthony Richardson and Thomas Kawula. A special thanks to Dr. Virginia Miller for being generous with her time to provide thoughtful research and career advice and attending ABRCMS in 2011 which lured me to UNC-CH. To Dr. Bill Goldman for leading a collegial department and helping me network within the microbiology community.

To the Initiative for Maximizing Student Diversity (IMSD) community for constant emotional support in dealing with the challenges of academia, STEM, and the current political environment. I'd like to thank the incomparable Dr. Ashalla Magee Freeman for directing IMSD and sacrificing her time and energy to ensure minority STEM PhD trainees at the UNC-CH School of Medicine have an inclusive environment to not only survive in but thrive and reach their highest goals. To the faculty and staff in the Office of Graduate Education, for devoting countless hours to ensuring trainees like myself have a productive and happy experience in doctoral sciences programs – Drs. Jean Cook, Patrick Brandt, Josh Hall, Jessica Harrell, Rebekah Layton, and former member Erin Hopper.

I would also like to acknowledge former teachers, professors, and research mentors who provided support and always challenged me to push beyond my comfort zone. A special thanks

to Tracy Fisher, Janice Graham, and Paul Rischard for advice and support in high school. To George Mason professors Dr. Kim Blackwell and Dr. Nadine Kabbani for providing me with my first opportunities to conduct bench research. To Dr. Marcia Coss for challenging me in an immunology class, setting high expectations, and teaching me the importance of speaking with confidence. To Dr. Tiffany Oliver for mentorship in the Left BRAIN summer program that sparked my long-term interest in research and her unwavering support both personally and professionally throughout my career. To the committee of the PREP program at Tufts University for admission and especially Dr. Henry Wortis for emotional and career support throughout PREP and my career. To Dr. Ralph Isberg for taking a chance to mentor a PREP student with very little experience in microbiology, giving me tremendous advice along the way, and introducing me to the microbiology community. To Dr. Linc Sonneshein for always answering my non-stop questions on graduate school and research and for always being a very humble and approachable scientist.

I'd like to thank family and friends who've been my backbone throughout this long PhD journey. To my mother Angelia Foster who made immeasurable sacrifices to ensure her children reach their highest potential in life. To my aunts Dorinda Thomas and Sherrie Withrow for always being loving and supportive. To last and certainly not least, my boyfriend Felix Olivares-Quintero who has made this last year of graduate school incredibly fun and memorable. I thank you for teaching me to appreciate life and learn to stop and smell the roses. Wakanda Forever!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

σ	Sigma
A	Alanine
Amp	Ampicillin
ATc	Anhydrotetracycline
BHIS	Brain Heart Infusion medium with yeast extract
bp	Base pair
C	Cysteine
D	Aspartic acid
E	Glutamic acid
Erm	Erythromycin
Flg ON	Flagellar phase ON
Flg OFF	Flagellar phase OFF
G	Glycine
Kan	Kanamycin
nt	Nucleotide
OD	Optical Density
R	Arginine
Tet	Tetracycline
Tm	Thiamphenicol

CHAPTER 1: INTRODUCTION

Overview

Bacillus difficilis was identified in 1935 in the feces of healthy newborns (1). As referenced in the species name, *B. difficilis* was "difficult" to culture because the bacterium required anaerobic conditions for growth (1). Hall and O'Toole demonstrated that injection of two day old cultures of *B. difficilis* into the peritoneum of guinea pigs and rabbits induced severe pathology, convulsions, and death (1). From the outcomes of the animal experiments, Hall and O'Toole hypothesized the bacterium produced a toxin (1). Lastly, Hall and O'Toole noted that *B. difficilis* was spore-forming and motile under the conditions tested (1). The key phenotypes observed by Ivan C. Hall and Elizabeth O'Toole in *B. difficilis* growth and virulence have stood the test of time and are under active investigation to date.

B. difficilis was renamed *Clostridium difficile* in the 1970s based on phenotypic similarity to other Clostridia (1-3). Like many Clostridia, the actively growing "vegetative" cell form of *C. difficile* is sensitive to atmospheric oxygen. Therefore, *C. difficile* vegetative cells reside in an anaerobic environment, such as the large intestine of a human host. The ability to form metabolically dormant spores confers the ability of *C. difficile* to transmit between hosts. *C. difficile* infection (CDI) manifests as a diarrheal disease that can range in severity from mild to a severe form known as pseudomembranous colitis. According to the Centers for Disease Control and Prevention, ~500,000 people are diagnosed with CDI resulting in ~30,000 deaths annually (4). CDI is largely considered a nosocomial infection whereby the most susceptible population

are patients in a hospital (5). In terms of colonization, *C. difficile* is not restricted to a human host and can colonize farm animals, such as cattle and pigs (6). The leading risk factor for CDI is prior antibiotic treatment for a preexisting bacterial infection. Broad spectrum antibiotics, such as clindamycin and ampicillin, increase susceptibility to infection through disruption of the microbiota in the intestine that normally afford resistance to *C. difficile* (7). CDI is treated with either metronidazole or vancomycin, for which *C. difficile* rarely exhibits resistance (8). Recurrent infections are common due to antibiotics failing to clear metabolically dormant spores. The most effective therapy is an experimental treatment called fecal microbial transplantation, whereby feces from a healthy donor is transferred into the intestine of a CDI patient for reconstitution of a protective microbiota (8).

C. *difficile* completes a life cycle of germination, vegetative cell growth, and sporulation during diarrheal disease (3, 9, 10). In a susceptible host, *C. difficile* spores transit the gastrointestinal tract and upon contact with primary bile salts and certain amino acids the spores germinate into vegetative cells (9, 11). Outgrowth of vegetative cells leads to the expression of different cell surface adhesins to promote *C. difficile* colonization of the colon (12, 13). Surface appendages include flagella (discussed later) and Type IV pili as contributing to *C. difficile* colonization (14, 15). The vegetative cells consume various host and microbial metabolites to grow to a high burden (16). Upon deprivation of nutrients within the intestine, the vegetative cells express and produce two toxins named TcdA and TcdB (10). Both toxins inactivate intracellular host factors responsible for assembly of the actin cytoskeleton resulting in disassembly of the cytoskeleton, stimulation of host immune defense genes, secretion of inflammatory cytokines and recruitment of neutrophils and monocytes that potentiate inflammation (10, 17). Nutrient deprivation and additional ill-defined signals trigger sporulation

of vegetative cells (9). The sporulation cascade is fairly well conserved from the model organism *Bacillus subtilis* in *C. difficile* (9, 18-20). Spores produce various surface proteins to restrict and permit passage of select molecules, adherence to host tissues, and detoxify host antimicrobials (9). Spores are shed into the environment where a susceptible host can become infected for the next cycle.

The remaining sections of the Introduction focus on topics directly relevant to understanding content in Chapters 2-4: toxin structure, regulation, and function; flagellar structure, regulation, and function; and an overview of phase variation.

Regulation & Function of the Glucosylating Toxins in *Clostridium difficile*

Genetic Organization of the Pathogenicity Locus

The genes encoding the glucosylating toxins are located within the Pathogenicity Locus (PaLoc) in toxigenic *C. difficile* strains (21). The PaLoc is a 19.6 kb chromosomal region that is conserved in toxigenic *C. difficile* strains and can be mobilized for transfer to non-toxigenic strains via a conjugation-like mechanism (22). The genetic locus contains genes that encode the glucosylating toxins, TcdA and TcdB, as well as the transcriptional regulators of the toxin genes, TcdR and TcdC, and an export apparatus, TcdE. Below, I briefly review the function of TcdR, TcdC, and TcdE. The structure, function, and regulation of TcdA and TcdB will be discussed later.

TcdR is an alternative sigma factor that is required for expression of the toxins in *C. difficile* (23, 24). Using *in vitro* transcription assays, Dupuy *et al.* demonstrated that recombinant TcdR alone cannot promote transcription from promoter sequences upstream of *tcdR*, *tcdA*, and *tcdB,* but can do so only when in association with RNAP polymerase (RNAP) (23). Recent work

from Girinathan *et al*. demonstrated that TcdR not only contributes to expression of the toxins, but also regulates genes outside the PaLoc contributing to sporulation and germination (25). However, the contribution of TcdR to *C. difficile* virulence during host infection has yet to be determined.

TcdC is a proposed negative regulator of toxin gene expression. TcdC is proposed to function as an anti-sigma factor by antagonizing TcdR interaction with RNAP prior to transcription initiation (26). Thus, TcdC indirectly inhibits transcription of *tcdA, tcdB*, and *tcdR* (26). Outside of the prescribed role in TcdR inhibition, TcdC can also interact with G-quadraplex DNA structures containing repetitive guanine nucleotides, but the functional contribution of this feature has yet to be assessed in *C. difficile* (27). Mutants of *tcdC* tell a different story depending on the strain background. Epidemic ribotype 027 strains contain nonsense mutations in *tcdC*, which were proposed to contribute to the hypervirulence of these strains. Overexpression of functional, intact TcdC in M7404, a ribotype 027 strain with nonsense mutations in *tcdC*, was sufficient to mediate repression of toxin gene transcription and production relative to vector control strains (28). In addition, hamsters infected with M7404 overexpressing TcdC succumb to infection significantly earlier than wildtype. In contrast to these data, Cartman *et al.* used allelic exchange to repair the mutations in *tcdC* in the R20291 background, a ribotype 027 strain, and found no difference in toxin activity by cytotoxicity assays between strains with a repaired and non-functional *tcdC* gene (29). Similar results were observed in a *tcdC* mutant and complemented strain of 630Δ*erm* (30).Taken together, the functional contribution of TcdC to *C. difficile* virulence has yet to be conclusively determined.

TcdE is a small protein suggested to contribute to toxin export. TcdE has three predicted transmembrane spanning domain and homology to holins (31). Holins are small membrane

proteins encoded in bacteriophage that insert in the membrane of bacterial cells to promote lysis and release of bacteriophage (32). Govind *et al.* generated a *tcdE* mutant in JIR8094 and found the mutant had a significant reduction in secreted TcdA and TcdB, but an increase in intracellular levels, relative to wildtype (31). Expression of *tcdE* in *E. coli* with an endolysin lead to bacterial lysis, providing *in vivo* evidence that TcdE can function as a holin (31). In contrast to these findings, Olling *et al.* demonstrated that a *tcdE* mutant of strain 630Δ*erm* secretes TcdA and TcdB to levels comparable to wildtype. TcdE function in TcdA and TcdB export is inconclusive based on the aforementioned reports that could be due to differences in *C. difficile* strains and assay conditions.

Roles of Toxins in Disease

Multiple animal models have been used to assess the functional contribution of TcdA and TcdB to diarrheal disease. In a rabbit ileal loop model, purified TcdA caused fluid accumulation, mild hemorrhaging, and immune cell trafficking, so TcdA became referred to as the enterotoxin (33). TcdB caused very little fluid accumulation and other symptoms in animal models, but in cell culture experiments promoted cell rounding and reduced viability, so it became referred to as a cytotoxin (33). In subsequent studies using the hamster model (34), multiple low doses of TcdA were sufficient to elicit fluid accumulation, diarrhea, and intestinal pathology, whereas even high doses of TcdB resulted in no observable effect. Combination studies using low doses of both TcdA and TcdB resulted in lethal intoxication with significant intestinal pathology and diarrhea (34). These results suggested that TcdA and TcdB act synergistically to cause diarrheal disease symptoms and that TcdA must cause damage to host tissue for TcdB to gain a foothold.

Infection studies using single or double mutants of *tcdA* and *tcdB* have provided new

insight into the specific contributions of each toxin during infection. Early studies used JIR8094,

a laboratory derivative of the ribotype 012 *C. difficile* strain 630 that was isolated from a patient

with fulminant diarrhea in 1982 (35). In JIR8094, the *tcdA* and *tcdB* genes were mutated

individually and in combination and then the resulting mutants were assayed for virulence in

hamsters pretreated with the antibiotic clindamycin (36). Lyras *et al.* found that hamsters

infected with the TcdB-only producing strain (A- B+) exhibited fulminant diarrheal disease and a

mean time to death similar to hamsters infected with the parent strain (36). By contrast, infection

with the TcdA-only producing strain (A+ B-) resulted in survival of at least 70% of the hamsters.

These results led to the early conclusion that TcdA is dispensable for disease, and TcdB is the

only essential toxin. However, in the subsequent year, Kuehne *et al.* produced conflicting data

using strain 630Δ*erm*, showing that both toxins individually contribute to diarrheal disease

development and significantly reduced time to death compared to uninfected or a double toxin

mutant (37). The primary difference between the two findings in these papers was the parental

strain used. Both studies used derivatives of strain 630 that were serially passaged on non-

selective media to obtain an erythromycin-susceptible strain (36-39). However, whole genome

sequencing and phenotypic analysis revealed multiple differences in the genome, including

single nucleotide polymorphisms, inversions, and other genomic alterations (40). The JIR8094

strain is also non-motile and produces less toxin than 630Δ*erm* and the mutual 630 parent (41,

42). These phenotypes, as well as the conflicting contributions of the toxins to disease in

hamsters, could be explained by the accumulated changes in the JIR8094 genome.

Toxin gene mutations in *C. difficile* ribotype 027 strains helped resolve the contribution

of each individual toxin to disease development and pathology. Using the hamster model of

infection, Kuehne *et al.* demonstrated the either TcdA or TcdB produced alone in R20291 (using

tcdB or *tcdA* mutants, respectively) is sufficient to cause fulminant diarrheal disease and lethality

compared to a double toxin mutant (43). The mean time to end point for hamsters infected with

TcdB only producing strain was 2.3 days, whereas an infection with TcdA only producing strain

was 5.9 days (43). These data are consistent with previous findings that TcdB is the more potent

toxin, but TcdA can also lead to disease, pathology, and lethality. To comprehensively determine

the contribution of each toxin to disease and to unify the field, Carter *et al.* assessed individual

and double mutants of *tcdA* and *tcdB* in ribotype 027 strain M7404 in three different animal

models in three different laboratories to control for differences in animal husbandry, the

intestinal microbiota, and antibiotic treatment protocols (44). In a mouse model at Monash

University where mice consistently develop severe diarrheal disease symptoms, all mice died

when infected with wildtype (TcdA+ TcdB+), and 95% of mice infected with the isogenic TcdA-

TcdB+ strain succumbed to infection by 48 hours (44). By contrast, only 20% of mice infected

with the isogenic TcdA+TcdB- strain died by 48 hours. In addition, histopathology scores and

weight loss for mice infected with the TcdA+ TcdB- strain more closely resembled infection

with the toxin double mutant than the wildtype or TcdA- TcdB+ strains (44). In a Sanger

Institute mouse model, where mice exhibit self-limiting inflammation and mild diarrhea,

infection with the TcdA-TcdB+ strain was more lethal than infection with wildtype or any of the

other mutants tested (44). Lastly, in the hamster model of infection at Hines VA hospital, the

infection with TcdA+ TcdB- was significantly attenuated based on mean time to moribund state

compared to infection with TcdA- TcdB+ or wildtype (44). Collectively, these data from

multiple ribotype 027 strains confirm the previous findings from Kuehne *et al.* in that both toxins

are sufficient for diarrheal disease development and lethality (37), but TcdA is less essential and

plays an adjunctive role. Recently, a TcdA+ TcdB- clinical isolate was found in a patient with fulminant diarrheal disease that resolved after treatment with antibiotics (45). The TcdA+ TcdB- only strain could not be detected by conventional diagnostic kits as they rely on TcdB activity. Contrary to the only TcdA+ TcdB- clinical isolate reported (45), TcdA-TcdB+ strains are frequently isolated from patients with diarrheal disease symptoms (46-48). Taken together, *C. difficile* produces toxins that are necessary for diarrheal disease development in humans and certain animal models.

Structure and Mechanism of Action of the Toxins

C. difficile toxins TcdA and TcdB are greater than 250 kDa in size and are roughly 50% identical in amino acid sequence (10, 49, 50). Both toxins have four functional domains designated as ACDB. The Active domain (A) contains the glucosyltransferase domain that modifies host substrates. The Cutting domain (C) becomes activated in the host cell cytosol to mediate autoprocessing and release of the A domain. The Delivery domain (D) mediates uptake and trafficking of the toxin through an endocytic pathway. The Binding domain (B) recognizes and binds a host cell surface receptor for subsequent intoxication. Briefly reviewed below are important domains, functional characterization, and additional details for each domain.

Step 1: Receptor Binding

The Combined Repetitive Oligopeptides (CROPS) domain is the C-terminal tail of the toxin and represents the Binding domain. The CROPS domain consists of short amino acid repeat sequences regularly interspersed between large amino acid repeat sequences. Protein crystallography and molecular modeling studies indicate that the CROPS domain forms a

cylindrical coil, or solenoid, built from the short and long repeat sequences at the junctions (51, 52).

Evidence for the CROPS domain functioning as the host cell receptor-binding domain came from studies evaluating the interaction of TcdA CROPS domain and glycan motifs present on rabbit erythrocytes (53), and various blood group glycans present in humans (54). In terms of protein receptors, TcdA can bind to gp96, a heat shock protein ER chaperone that is posttranslationally modified with carbohydrate sequences (55). Depletion and inhibition of gp96 expression and activity only partially protected cells from TcdA cytopathicity and intoxication, suggesting an additional set of receptors for TcdA (55). A mutant of TcdA lacking the CROPS domain remained capable of exerting cytopathic effect on various mammalian cells lines, suggesting alternative domains upstream of the CROPS that mediate receptor recognition (56).

Antibodies targeting the CROPS domain of TcdB reduces disease, suggesting that this domain contributes to intoxication (57). Three host cell receptors for TcdB have been identified. Yaan *et al.* identified Chondroitan Sulfate Proteoglycan 4 (CSPG4) as a receptor for TcdB using a knockdown system in HeLa cells and screening for resistant clones (58). CSPG4 interacts with TcdB via the CROPS domain based on pull down, blocking, and mutant experiments (58). However, loss of CSPG4 in mice does not completely protect against TcdB by intraperitoneal challenge, suggesting an additional receptor. Moreover, high concentrations of TcdB can kill CSPG4 deficient cells, again suggesting another receptor (58). In 2015, LaFrance *et al.* identified nectin-3 (or poliovirus receptor-like protein 3) as a TcdB receptor by screening cells with a spectrum of gene knockdown constructs with high concentrations of TcdB and evaluating cell viability (59). Knockdown and competitive inhibition of nectin-3 protected cells from TcdB destruction (59). CROPS-less TcdB could still infect host cells, suggesting that, like in TcdA, a

domain outside of the CROPS facilitates TcdB interaction with nectin-3 (59). Lastly in 2016, Tao *et al.* identified the Wnt receptor frizzled protein receptor family (1, 2 & 7) as cellular receptors for TcdB (60). TcdB bound frizzled proteins independent of the CROPS domain, and combined knockdown and knockout of essential and non-essential frizzled genes, respectively, mediates cellular resistance to TcdB (60). Mice deficient in non-essential frizzled proteins were less susceptible to TcdB pathology (60). If and how TcdB engages all three protein receptors during host infection remains to be determined.

Step 2: Toxin Uptake & Translocation

TcdA and TcdB bind their respective receptors, and subsequent uptake and translocation is dependent on the Delivery domain. The GTPase dynamin mediates toxin endocytosis (61). Dynamin is responsible for the scission of vesicles. Clathrin proteins mark the budding vesicles to specify the destination. TcdB depends on clathrin for uptake (61), whereas TcdA uses a clathrin-independent pathway mediated by PACSIN2 (62).

Upon entry, the toxin-containing vesicles are trafficked through endosomes to an acidified compartment. Both toxins undergo pH-dependent conformational changes, which promotes insertion into the membrane (52). Pharmacological agents that target the vesicular ATPase attenuate toxin activity, suggesting a requirement for low pH (63, 64).

Step 3: Activation

The Cutting domain of TcdA and TcdB mediates the activation and release of the Active domain into the cytosol. Cytosolic inositol hexakisphosphate (InsP6) induces an autocatalytic cleavage through binding to the autoprocessing domain (APD) (65, 66), which undergoes a

significant conformational change to release the A domain. Three essential residues – cysteine, histidine, and aspartate – coordinate a zinc ion for auto processing (67, 68).

Step 4: Target Modification

The Active domain of TcdA and TcdB contains the enzymatic domain for host substrate modification. The A domain is a glucosyltransferase (GTD) domain. The GTD uses uridine diphosphate (UDP)-glucose in the host cell cytosol for host substrate modification (69, 70). The GTD targets Rho family proteins, the GTP binding Ras superfamily that regulates the cytoskeleton, but also indirectly affects cell cycle progression and apoptosis. The GTD attaches the UDP-glucose to a threonine residue in the switch I region of Rho proteins (71). The switch I region is essential for GTP binding and interaction with downstream effectors. Modification with UDP-glucose prevents switch I from undergoing the conformational changes necessary for GTP binding, which is necessary for activity and keeps the protein inactive (71). Glucosylation of the Rho proteins is irreversible since humans do not possess glycosidases in the cytosol that can cleave the glucose-threonine bond. Both TcdA and TcdB target RhoA, RhoB, RhoC, Cdc42, and Rac1, and individually affect other minor substrates (10).

Functional Consequence of Toxin Activity

TcdA and TcdB indirectly disrupt the actin cytoskeleton, which ultimately activates a variety of host cell death pathways. In polarized and confluent cell lines, toxin activity compromises tight junctions and results in increased paracellular permeability (72-75). In terms of cell death, apoptosis in epithelial cells is dependent on GTD activity and characterized by DNA fragmentation, cell shrinkage, and caspase activation (76, 77). The toxins also activate the

inflammasome and induce pyroptosis. Inflammasome activation occurs via Pyrin sensing of the modification of RhoA, which eventually leads to cleavage of caspase-1 (78). Caspase 1 cleaves both pro-interleukin 1ß and interleukin-18, which are released from cells via gasdermin pores and the cell lyses from a disrupted concentration and ion gradient (79). Uniquely, TcdB can induce necrosis in epithelial cells, independent of the GTD, at high concentrations by activating the epithelial NADPH oxidase to make reactive oxygen species (80, 81).

Regulation of Toxin Production

The glucosylating toxins are subject to diverse regulatory mechanisms beyond TcdR and TcdC. How nutrition, sporulation, and quorum sensing control toxin production are reviewed below.

A. Nutritional Control

Nutrient limitation during stationary phase *in vitro* growth conditions of *C. difficile* are optimal for expression and production of the toxins. Early studies found that *tcdA*, *tcdB*, *tcdR*, and *tcdE* are transcribed at low levels in exponential phase, but highly induced upon entry into and during stationary phase, whereas *tcdC* has the inverse expression pattern (82). The addition of different nutrient sources to the media was found to activate or repress toxin gene expression and production (83). Glucose, fructose, and mannitol repress toxin gene expression during exponential and stationary phase (83). Glucose repression of toxin gene expression and production was observed for several *C. difficile* strains and therefore represents a generalizable phenomenon.

Carbon catabolite repression allows bacteria to inhibit expression of genes encoding machinery necessary for transport and utilization of non-preferred sugars when a preferred one is present (84). The process is mediated by the transcriptional regulator CcpA, the carbon catabolite control protein A. Bacteria can transport different carbon sources into the cell via the phosphoenolpyruvate-dependent carbohydrate: phosphotransferase system. In a *ccpA* mutant of *C. difficile*, toxin gene expression was constitutive and no longer responsive to glucose (85). In addition, mutations in genes encoding different components of the PTS made *C. difficile* unresponsive to glucose-mediated repression of toxin gene expression (85). CcpA was found to bind non-specifically to sites upstream of *tcdA* and *tcdB*, and two consensus binding sites for CcpA were found upstream of *tcdR* (85). Thus, carbon catabolite repression affects not only nutrient utilization, but also virulence factor production in *C. difficile* like many other pathogens (84).

Branched chain amino acids and GTP levels inversely correlate with toxin production in *C. difficile* (84). In low-GC Gram positive bacteria, the transcriptional regulator CodY directly senses GTP and branched amino acids (isoleucine, leucine, and valine) via the N-terminus, and the helix-turn-helix DNA binding domain in the C-terminus mediates repression of gene transcription during nutritional abundance (86). During nutrient deprivation, or conditions that mimic stationary phase where GTP and amino acid pools become limited, CodY becomes inactivated, resulting in derepression of genes that promote amino acids availability. A *codY* mutant in *C. difficile* showed increased expression of the toxin genes (87). In Chip-Seq and DNase I protection assays, CodY was found to bind to three regions upstream of *tcdR* – two regions in the *tcdR*-dependent promoter and one downstream in a σ^A- dependent promoter region (88, 89). In addition, CodY regulation of toxin production is conserved among ribotypes (90).

Thus, CodY activity and regulation are conserved in *C. difficile* as in other Gram-positive bacterial pathogens (84).

B. Sporulation

Nutrient limitation is a cue for certain bacteria to undergo sporulation, a process involving the formation of metabolically dormant cells that can persist until favorable growth conditions resume (91). Genes that are required for or contribute to initiation of sporulation in *C. difficile* also regulate toxin production (92). In *B. subtilis*, a phosphorelay involving multiple sensor kinases activates the master regulator of sporulation Spo0A to activate genes necessary for sporulation (91). In tandem, an alternative sigma factor, SigH, is required for expression of *spo0A* (93). In *C. difficile*, a *sigH* mutant showed increased toxin gene expression and toxin production, indicating that SigH represses toxin production (94). No SigH consensus sequence was found in the PaLoc, so the mechanism is undetermined and likely indirect. Mutation of *spo0A* in *C. difficile* resulted in an increase or decrease in toxin production depending on the ribotype and strain lineage (95-97). No Spo0A consensus sequence was found in the PaLoc, suggesting an indirect mechanism of regulation. Outside of sporulation initiation, toxin gene expression was unaffected in sporulation-specific sigma factor mutants (18, 20). These data suggest that the initiation process of sporulation serves as a cue, depending on the strain background, to promote toxin production, but upon commitment to the sporulation developmental program, toxin gene transcription remains unchanged. In agreement with this hypothesis, Edwards *et al.* found a new regulator of sporulation and toxin production, RstA, that represses toxin production and activates sporulation (98). Moreover, recent work from Girinathan *et al.* sought to determine the role of the *sin* locus in *C. difficile* (99). In *B. subtilis,*

SinR is a DNA-binding protein that negatively regulates sporulation and SinR activity is directly antagonized by SinI (91, 100). In *C. difficile*, mutation of *sinR* reduced expression of *spo0A*, which resulted in an asporogenous phenotype, as well as a significant reduction in production of TcdA and TcdB (99). In *C. difficile*, the SinI orthologue appears more like SinR paralogue and was named SinR' (99). A *sinR'* mutant produced significantly more toxin, yet the *sinR'* mutant still produces SinR (99). The mechanism identified by which SinR and SinR' control toxin production is through c-di-GMP and SigD (discussed later).

C. Quorum Sensing

Quorum sensing (QS) is a process by which bacteria coordinate group behaviors, and QS involves the detection, sensing, and transduction of a secreted molecule or peptide based on cell-density (101). *C. difficile* strains possess genes for synthesis of both a peptide based quorum sensing system called the accessory gene regulator (Agr) and a second class of small molecule called autoinducer-2 involved in interspecies interactions. Mutation of genes necessary for peptide quorum sensing attenuates toxin production and *C. difficile* virulence likely by modulating expression of the flagellar alternative sigma factor, SigD (discussed later) (102, 103). The contribution of autoinducer-2 to toxin production is inconclusive due to conflicting reports (104, 105). Nonetheless, the contribution of peptide-based QS to *C. difficile* virulence is consistent with studies done in other Gram-positive bacterial pathogens (101).

Flagellar Function & Regulation in *Clostridium difficile*

<u>Conservation and Genetic Organization of Flagellar Genes</u>

The flagellar genes are organized into three distinct loci in most *C. difficile* ribotype strains (106). The early stage flagellar biosynthesis locus, based on gene conservation and similarity to flagellar gene organization in *B. subtilis* (107), is the F3 or *flgB* operon. The *flgB* operon encodes proteins necessary for assembly of the flagellar basal body, rod, motor proteins, and the alternative sigma factor SigD (FliA or σ^{28} in Gram negative bacteria). Expression of the *flgB* operon is driven by a single promoter with consensus sequences that match the σ^{A} housekeeping sigma factor (108). The basal body serves as a structural anchor located in the cytoplasmic membrane and through the peptidoglycan. The basal body functions as a conduit for the secretion of flagellar proteins involved in late stage assembly of the hook and filaments, and also as a rotor for torque generation by the motor proteins. The rod serves as a drive shaft. The hook has a curved architecture, is located outside of the basal body, and functions as the joint between the basal body and filament. The SigD alternative sigma factor encoded within the *flgB* operon controls transcription of flagellar genes in the late stage F1 and F2 loci (41, 109). The F1 locus contains genes encoding the flagellar filament, chaperones, the flagellar cap, and a glycosyltransferase (106, 110). Regulation of the F1 locus is hypothesized to be restricted until hook-basal body completion by the anti-sigma factor FlgM as in *B. subtilis* (111). The F2 locus encodes flagellar glycosylation and other posttranslational modification proteins and is located between the F3 and F1 loci. The F2 locus consists of four genes in *C. difficile* 630, but six genes in ribotype 027 strains CD196 and R20291 (110).

Glycosylation and Posttranslational Modification of Flagella

Flagella can be subject to glycosylation and other modifications to alter function in numerous Gram-negative and -positive bacterial pathogens (112). Glycosylation is predicted to shield binding sites of flagellin from host immune receptor and antibody recognition, in addition to autoagglutination, flagellar assembly, and various other roles (112, 113). In 2009, Twine *et al.* showed using mass spectroscopy that flagellin from *C. difficile* 630 is glycosylated with a N-acetyl hexosamime (HexNAc) residue (114). Addition of the HexNAc residue was dependent on the glucosyltransferase encoded within the F1 locus. Mutation of the glucosyltransferase gene rendered the flagellin protein smaller and unmodified, and the corresponding mutant bacteria produced significantly fewer flagellar filaments and were non-motile (114). Analysis of flagellin glycan modification in other *C. difficile* isolates suggested greater complexity in flagellin modifications. Faulds-Pain *et al.* determined that the HexNAc modification was misidentified and, based on NMR analysis, was actually a ß-O-linked N-acetylglucosamine (GlcNAc) attached to either serine or threonine in the flagellin protein (115). The O-linked GlcNAc addition was further modified with a methylated phospho-threonine (115). Mutations of the genes involved in posttranslational modification of flagellin led to autoagglutination, increased binding to abiotic surfaces, and reduced colonization in a relapsing model of *C. difficile* infection (115). All of the aforementioned studies were done using *C. difficile* strains harboring 4 genes in the glycosylation F1 locus. A recent study from Valiente *et al.* evaluated the contribution of five out of seven genes predicted to contribute to flagellin glycosylation and PTM in ribotype 027 strains(116) . Mutation of three predicted glycosyltransferase genes were shown to contribute to the sequential addition of ß-O-linked GlcNac, two rhamnoses, and a novel sulfonated peptidyl-amido sugar moiety (116). Select F1 loci gene mutants in R20291 were attenuated for motility, had increased

biofilm formation and autoagglutination (116). Lastly, all of the F1 mutants tested could stimulate host Toll-like Receptor 5 (TLR5), an immune receptor for flagellin recognition, to secrete interleukin-8 at levels comparable to the wildtype strain demonstrating that glycosylation does not hinder recognition by TLR5 (116). Taken together, glycosylation and PTM of flagellin plays variable roles in *C. difficile* flagellar motility, adherence, and host infection.

Roles of Flagella

A. Motility

Flagella contribute to *C. difficile* swimming and swarming motility, but the degree of motility is ribotype-specific. Stabler *et al.* found that ribotype 027 strains were more motile than ribotype 012 strains (110). However, Valiente *et al.* revealed heterogeneity in the degree of motility even among strains of the 027 ribotype (117). Lastly, as predicted, mutations in individual structural and regulatory genes necessary for flagellar biosynthesis abrogated flagellum production and impair motility irrespective of ribotype (118-121).

B. Autoagglutination

Surface structures in bacteria can alter their ability to autoagglutinate, or the ability to clump together or with host cells. Among clinical isolates, Valiente *et al.* observed heterogeneity in autoagglutination among *C. difficile* ribotype 027, 176, 198, and 244 strains (117). Stabler *et al.* found that R20291, which has additional genes in the F2 locus, was attenuated for autoagglutination compared to the historical 027 ribotype strains CD196 and BI-1 and ribotype 012 strain 630 (110). The differences in flagellin posttranslational modification could contribute to these autoagglutination differences (115, 116). Many other cell surface factors may contribute

to autoagglutination independent of flagella because strain M120, which lacks the *flgB* operon locus, autoagglutinates to a significantly higher degree than all of the other strains tested (110). Nonetheless, autoagglutination by flagella could contribute to *C. difficile* virulence during host infection.

C. Adherence and Colonization

The flagellum structure and flagellar-driven motility contribute to bacterial adherence or colonization to host epithelial cells in many bacterial pathogens, such as the enteric pathogens *Campylobacter jejuni* and *Listeria monocytogenes* (113). Flagellar filaments are long structures that extend up to 20 micrometers beyond the bacterial cell and can thus probe the host cell surface for initial contact. Flagellar-dependent motility can drive filament interaction with great force onto the host cell surface (113). Beyond TLR5 (122), there are no reports of a specific flagellin or flagellar receptor for bacterial binding. Most of the features that promote flagellar-dependent adhesion are polymeric protein, glycans, and lipids (113).

Flagella contribute to *C. difficile* adherence to epithelial cells *in vitro* and host colonization in a ribotype-dependent manner. An early study from Tasteyre *et al.* showed recombinant flagellin FliC, the flagellar cap protein FliD, and crude flagella can bind to and interact with cecal mucus purified from axenic mice, but not with porcine mucus (123). To determine whether flagella contribute to colonization and adherence to mucus in the cecum, axenic mice were infected with flagellated or aflagellate *C. difficile* strains from the same serogroup, the mice were sacrificed, and cecum removed in an anaerobic chamber to plate for colony forming units. With two different serogroup strains, Tasteyre *et al.* found that flagella contribute to adherence to the cecum in axenic mice during infection (123). However, these

studies predated tools for targeted genetic modification of *C. difficile* and relied on comparisons of non-isogenic strains. Later work from Dingle *et al.* using 630Δ*erm*, an erythromycin sensitive derivative for gene inactivation using an *erm* resistance cassette marker, showed that mutation of *fliC* or *fliD* enhances bacterial adherence to epithelial cells compared to wildtype (118). In addition, both mutants colonized the lumen and tissue of the hamster cecum to wildtype levels, indicating that flagellar assembly is dispensable in the hamster model of infection. Surprisingly, hamsters infected with either the *fliC* or *fliD* mutant reached a moribund state sooner than wildtype (118). In agreement, Aubry *et al.* used 630Δ*erm* to generate mutants and observed a similar trend in which mutations that inactivate early stage flagellar genes result in attenuated virulence in the hamster model, whereas late stage mutants (*fliC, fliD*) were more virulent and animals reached a moribund state sooner (119). Results from both Dingle *et al.* (118) and Aubry *et al.* (119) suggested that mutation of late stage genes altered the immunogenicity of *C. difficile*. Both aforementioned studies revealed an increase in toxin production for late stage flagellar gene mutants (118, 119), but Aubry *et al.* noted toxin gene expression and production were attenuated for early stage mutants. How are flagellar gene expression and toxin production connected? The flagellar alternative sigma factor SigD was demonstrated to control expression and production of TcdA and TcdB by activating expression of *tcdR* (41, 109, 119). Thus, the mutations 5' of *sigD* in the *flgB* operon can exert a polar effect on *sigD* expression resulting in a decrease in toxin expression, production, and attenuated virulence of *C. difficile* during host infection (119).

Phenotypic characterization of flagellar mutants in R20291 yielded different results. Baban *et al.* found that flagellar gene mutants of R20291, in both early and late stage genes, were attenuated for adherence to Caco-2 cells. Allelic exchange was used to generate a paralyzed flagellar mutant through mutation of *motB,* a motor protein, in R20291 (29, 121). The paralyzed

flagellar mutant produced wildtype levels of flagellar filaments, but were unable to engage in

swimming motility due to an inactive motor (121). Survival curves between mice infected with a

paralyzed flagellar mutant and wildtype were comparable (121). However, wildtype R20291 was

attenuated for colonization based on bacterial burden relative to the paralyzed flagellar mutant

suggesting that motility-driven colonization is dispensable and instead the flagellar filaments

serves as a scaffold to promote mucus and host cell interactions. To support the filament-

colonization hypothesis, axenic mice were co-infected with the paralyzed flagellar mutant and a

fliC mutant and the paralyzed flagellar mutant outcompeted the *fliC* mutant for colonization in

the cecum (121). Notably, studies from this paper should be interpreted with caution, since their

R20291 strain produced a single flagellum (121), which has never been observed in our R20291

isolate, any other published ribotype 027 strains, or clinical isolates of *C. difficile*.

Regulation of Flagellar Gene Expression

Prior to our studies, a single *cis*-acting regulatory element was identified to control

flagellar gene expression in *C. difficile.* The nucleotide second messenger cyclic diguanylate (c-

di-GMP) is synthesized from two GTP molecules by diguanylate cyclases (DGC) and

hydrolyzed into pGpG by phosphodiesterases (PDE) (124). Cyclic diguanylate regulation

contributes to numerous phenotypes in bacteria, to include motility, biofilm formation, and

virulence factor production (124). Cyclic diguanylate binds directly to the Cd1 riboswitch in the

leader RNA of the *flgB* operon, resulting in an altered conformation of the mRNA leader that

controls downstream gene expression. (125). In the absence, or in low-levels, of c-di-GMP,

transcriptional elongation through Cd1 is uninterrupted and the downstream genes are

transcribed. In the presence of higher concentrations of c-di-GMP, the nucleotide binds to the

riboswitch, alters RNA base pairing in a manner that results in the formation of an intrinsic transcription terminator. Thus, c-di-GMP triggers premature transcription termination within the *flgB* leader sequence, precluding transcription of the downstream genes (108, 125). Consistent with this, artificially increasing intracellular c-di-GMP in *C. difficile* through the ectopic expression a DGC inhibited flagellar gene expression and impeded motility (108, 126).

Numerous transcriptional regulators control flagellar gene expression and motility in *C. difficile*. Activators of flagellar motility include AgrA (102), Hfq (127), and SinR (99) (Figure 1.1). As previously mentioned, AgrA is a quorum-sensing regulated transcription factor. An *agrA* mutant was shown to be attenuated in transcription of flagellar genes by RNA-seq and production of flagella by electron microscopy (102). The mechanism of flagellar gene regulation by Agr QS could be through the Cd1 riboswitch since several PDEs were differentially regulated in the *agr* mutant (102). An additional regulator, Hfq, was demonstrated by Boudry *et al.* to control flagellum production in *C. difficile* (128). Hfq is an RNA chaperone that promotes base-pairing of regulatory RNAs, specifically *trans*-acting that have limited sequence complementarity, to target messenger RNAs (129). Boudry *et al.* were unable to generate a mutation in *hfq* and thus relied on antisense RNA-mediated depletion of target RNA and protein synthesis (127). Data (not shown in the manuscript) suggested that depletion of Hfq results in reduced motility and the loss of flagella (127). RNA-seq experiments identified a variety of differentially expressed genes, however, no changes in flagellar gene expression were observed (127). Interestingly, depletion of Hfq resulted in a significant increase in expression of both *sinR* and a DGC gene, but a decrease in the expression of a PDE gene (127). Based on the Hfq depletion RNA-seq results, Hfq could possibly control flagellar motility through altering c-di-GMP metabolism genes that either transcriptionally, post-transcriptionally, or post-

translationally control flagellar motility. Lastly, SinR was recently described to serve as an activator of flagellar motility through repression of a DGC gene, *dccA* (99, 126). By HPLC analysis, the *sinR* mutant had an over three-fold increase in the intracellular concentration of c-di-GMP relative to the wildtype strain (99). Taken together, Agr QS, Hfq, and SinR appear to activate flagellar gene expression through modulation of genes that would reduce the intracellular concentration of c-di-GMP, which could promote transcriptional readthrough at the Cd1 riboswitch.

Repressors of flagellar motility and gene expression include RstA (98), SinR' (99), and SigH (94) (Figure 1.1). All three aforementioned regulators have been demonstrated to control toxin gene expression and production in *C. difficile* by regulation of *sigD* expression or activity. Briefly, Edwards *et al.* demonstrated that RstA affects transcription of *sigD* and *fliC*, suggesting RstA regulates transcription of the *flgB* operon either directly or indirectly (98). The Helix-Turn-Helix (HTH) domain of RstA was deemed essential for regulation of motility since complementation of the *rstA* mutant with RstAΔ-HTH failed to restores *fliC* expression (98). An additional repressor, SinR', was found by Girinathan *et al.* to negatively regulate flagellar motility by directly antagonizing the activity of SinR, a positive regulator of flagellar gene expression (99). Lastly, mutation of the alternative sigma factor, SigH, in *C. difficile* was found to have increased expression of *fliC* and *sigD* in stationary phase relative to wildtype (94). The regulatory mechanisms by which SigH and RstA in *C. difficile* repress flagellar gene expression have yet to be determined. Nonetheless, many of the positive (i.e., Agr QS and SinR) and negative (RstA, SinR', and SigH) regulators of flagellar motility also affect toxin production through *sigD* expression or activity, which highlights the intimate link between both virulence factors in *C. difficile*.

C. difficile expresses flagellar genes during host infection that result in an antibody response. Early studies from Pechin *et al.* found that sera from patients diagnosed with CDI had antibodies targeting different components of the flagellum, including the flagellin FliC and the flagellar cap protein FliD (130). An antibody response to FliC and FliD was observed during stages of CDI with fulminant diarrheal disease symptoms highlighting the link between flagella and toxin production (130).

A role for innate immune recognition of *C. difficile* flagellin has been largely inconclusive until recently. Host TLR5 can recognize bacterial FliC through specific interactions with a conserved domain of FliC as long as certain amino acids are conserved (122, 131). To date, all sequenced strains of *C. difficile* retain the conserved amino acids necessary for recognition by TLR5. However, Jarchum *et al.* found that TLR5 deficient mice were no more susceptible to infection than wildtype, suggesting that TLR5 does not play a role in recognition of *C. difficile* flagella during infection to mount a protective (or detrimental) immune response (132). Data from subsequent studies challenged the hypothesis that *C. difficile* flagellin fails to stimulate TLR5 during host infection. Yoshino *et al.* demonstrated in cell culture experiments with intestinal epithelial cells that recombinant FliC from *C. difficile* was sufficient to activate both NFkB and p38 MAPK signaling via TLR5 (133). In addition, recombinant FliC induced secretion of chemokines IL-8 and CCL20, which function in recruitment of neutrophils to the site of infection (133). More significantly, authors found that pretreatment of cell monolayers with TcdB to disrupt tight junction between epithelial cells further enhanced the host response to FliC (133). TLR5 is largely restricted to the basolateral face of intestinal epithelial cells and

damage to tight junctions enhances TLR5 engagement with FliC (134, 135). Work from Batah *et al.*(136) also corroborated findings from Yoshino *et al.* (133). The aforementioned studies were done in cell culture, but what about the host response to *C. difficile* FliC during infection? Later work from Batah *et al.* showed that a *fliC* mutant of R20291 failed to elicit diarrheal disease symptoms or clinical manifestation, with very little pathology observed compared to infections with the wildtype parent (137). In addition, a *fliC* mutant failed to activate NFkB signaling during infection (137), which corroborates previous cell culture phenotypes (136). Batah *et al.* further evaluated a paralyzed flagellar mutant during host infection and found the mutant elicited symptoms and pathology comparable to infection with the wildtype parent. In TLR5 deficient mice, very little pathology was observed in cecal tissue in terms of inflammation, and no animals developed symptoms of disease. When the cell culture and host infection studies are taken together (133, 136, 137), TLR5 activation in intestinal epithelial cells by *C. difficile* FliC stimulates a detrimental host response (Figure 1.2). TLR5 deficient mice were resistant to damage mediated by FliC stimulation in synergy with toxin activity (137). The aforementioned studies are contradictory to previous findings from Jarchum *et al* (132)*.* The differences among the studies may be attributable to the *C. difficile* strains used. The Jarchum *et al.* paper used *C. difficile* strain VPI10463, which overproduces the toxins relative to clinical or laboratory strains (138). The increased toxin production by VPI10463 may severely damage intestinal epithelial cells prior to engaging *C. difficile* FliC via TLR5, although this is conjecture. Nonetheless, *C. difficile* flagellin, in combination with the toxins, stimulates a detrimental host response during infection that leads to severe pathology and death.

Phase Variation of Virulence Factors in Bacteria

<u>Definition and Genetic Mechanisms of Phase Variation</u>

Phenotypic diversity in a clonal population of bacteria can be generated through a process known as phase variation (139-141). Phase variation involves the stochastic or regulated production of secreted or cell surface proteins in an ON/OFF manner. The overarching hypothesis is that phase variation contributes to host immune evasion during infection to promote bacterial persistence and transmission. Secreted and cell surface gene products in bacteria are often immunostimulatory, thus finely tuned expression and production of these antigens could prevent detection and clearance by the host immune system.

A variety of genetic mechanisms can generate phase variation of bacterial gene products in a population (139, 140, 142). An essential requirement is that the nucleotide identity in a DNA sequence remain unchanged relative to the parental clone, in contrast to a missense mutation. In addition, phase variation occurs at a frequency higher than that of a standard mutation. The ON/OFF expression of the bacterial gene product is heritable, meaning the daughter cells inherit the expression phase of the parent. Although seemingly contradictory, the expression is reversible, meaning that under select conditions expression of the gene product can switch to the opposite phase. In addition, the fitness benefit conferred to a population is a bet hedging strategy. If and when the population encounters conditions that favor one phase over another a small subpopulation can persist and continue to divide. Briefly, we review three genetic mechanisms of phase variation, to include slip strand mispairing, gene conversion, and conservative-site specific recombination, and provide a well-studied example in a bacterial pathogen.

A. Slip Strand Mispairing (SSM)

Single or short repeat sequences of 1 to 7 nucleotides can expand or contrac during DNA replication between the template and daughter strands in a process called slip strand mispairing (SSM) (143). The increase or decrease in the size of repeat elements in the operator, promoter, leader, or coding sequence of a gene can affect expression. In *Campylobacter jejuni,* flagellar motility is phase variable due to slip strand mispairing in the coding sequence for both the sensor histidine kinase and response regulator of the FlgSR two component system (144, 145). FlgSR activates transcription of genes with promoters that are σ^{54}-dependent, such as the flagellar basal body, rod, and hook genes. Flagella are essential for *C. jejuni* commensal colonization of chicks (146), and non-motile phase variants of *C. jejuni* are attenuated for chick colonization (144, 145). SSM is one mechanism present in bacteria that can generate phenotypic diversity (143).

B. Gene Conversion

Gene conversion involves the homologous recombination of a silent allele with the expression allele of a gene to generate diversity in gene products (139, 147). Gene conversion requires proteins involved in DNA repair, such as RecA, and recombination is unidirectional. Gene conversion occurs at a much higher frequency than general homologous recombination in that it requires less sequence identity to recombine and involves accessory factors. The best described example in the literature involves the type IV pilin *pilE* gene from *Neisseria gonorrheae,* which contribute to DNA transformation and host cell attachment (148). The pilin protein has a region of functionality reserved to the N-terminus (~2/3), whereas the C-terminus is variable and surface exposed. The gonococcal genome encodes up to 19 silent alleles of *pilE* known as *pilS,* which consist of full-length or truncated versions of *pilE* with limited homology (148). Type IV pilus biogenesis is phase variable in that select silent alleles encode frameshift or nonsense mutations that result in pilus OFF bacteria that evade host antibodies targeting pilin

proteins. Given the diversity of *pilS* in a gonococcal genome, gene conversion allows pilus ON

bacteria to evade host antibodies targeting antigenically distinct pilin proteins (149)

C. Conservative Site-specific Recombination

Conservative site-specific recombination involves recombination of a short region of

DNA with limited homology (139, 147). A serine or tyrosine recombinase can mediate inversion,

insertion, or excision of a DNA element based on the limited homology (150-152). DNA

inversion can occur in the regulatory or coding sequence of a gene to affect expression or

activity. The best studied system for DNA inversion has been the fimbrial switch (*fimS*) that

controls the Type I fimbrial operon (*fimA*) in *E. coli*. FimS consists of a 296 bp DNA element

flanked by 9 bp inverted repeats (153). The switch contains a promoter that in an orientation-

dependent manner facilitates transcription of the *fimA* operon and either fimbriae are produced or

not on the bacterial surface. Two recombinase genes located upstream of *fimS* recognize and

catalyze DNA inversion. FimB can perform inversion in both orientations whereas FimE

functions exclusively in recombination from the ON to OFF orientation (154). Functionally,

Type I fimbriae contribute to *E. coli* adherence and invasion of epithelial cells in the bladder.

Genetically phase locked *fimS* ON bacteria outcompete *fimS* OFF bacteria in colonization of the

bladder during mouse infection (155). Many other bacterial pathogens, including *Salmonella*

enterica serovar Typhimurium and *Streptococcus pneumoniae* (139, 156-158), use DNA

inversion elements to control production of surface proteins and molecules that contribute to

colonization, biofilm formation, and virulence during host infection.

Phase Variation in *Clostridium difficile*

Prior to our research, Cell wall protein Variable (CwpV) was the only gene product characterized to undergo phase variation in *C. difficile* (159). CwpV is a cell surface anchored protein with a unique C-terminal domain consisting of serine-glycine rich repeat elements (159). The amino acid composition and number of repeat elements in the C-terminus of CwpV was shown to be consistent within, but varies among, ribotypes (159, 160). Emerson *et al.* demonstrated that CwpV production was restricted to ~5% of bacterial cells in a population using flow cytometry and indirect immunoflurorescence (159). Sequencing of the region upstream of *cwpV* lead to the identification of a genetic switch (159). A 195 bp element, termed the *cwpV* switch, was flanked by imperfect 21 bp inverted repeats (159). Emerson *et al.* identified a σ^A –dependent promoter upstream of the *cwpV* switch and found that the switch regulates *cwpV* post-transcription initiation. In the *cwpV* ON orientation, transcription initiates from the σ^A –dependent promoter and elongation proceeds uninterrupted into the *cwpV* coding sequence for expression (159) (Figure 1.3). In the *cwpV* OFF orientation, transcriptional initiation is unchanged, but the mRNA at the junction of the distal inverted repeat and the switch adopts the stem-loop structure an intrinsic terminator to prematurely terminate *cwpV* expression (Figure 1.3). Furthermore, a single tyrosine recombinase was identified to be necessary and sufficient for inversion at the *cwpV* switch and renamed recombinase variable (RecV) (159, 160).

Phase variable production of CwpV in *C. difficile* has been shown to contribute to bacterial aggregation and bacteriophage resistance (159, 160). Reynolds *et al.* demonstrated that CwpV overexpression, specifically the repeat elements in the C-terminus, promotes bacterial aggregation *in vitro* (159, 160). Based on this phenotype, Reynolds *et al.* hypothesized that CwpV promotes bacterial colonization during host infection. In a tangentially related study,

Sekulovic *et al.* found that lysogenic conversion of *C. difficile* strain R20291 with CD38 phage resulted in all *cwpV* phase ON cells (161). Based on these findings, Sekulovic *et al.* hypothesized that CwpV is a remnant of a prophage that prevents phage superinfection by blocking additional phage from binding to the cell surface (161). In subsequent work, Sekulovic *et al.* tested this hypothesis and demonstrated that CwpV mediates evasion to bacteriophage (162). Specifically, CwpV was found to reduce phage adsorption (partially) and through an unknown mechanism prevent phage tail spike DNA injection (162) (Figure 1.3). Beyond characterization of CwpV and RecV, to date, no other genetic switches had been characterized.

Stabler *et al.* identified two additional DNA inversion sites by comparing *C. difficile* genomes and named them *C. difficile* inversion sites 2 and 3 (110). Cdi1 was the *cwpV* genetic switch (110, 159). Cdi2 and Cdi3 were found upstream of c-di-GMP phosphodiesterases (163, 164). However, Cdi2 and Cdi3 have not been demonstrated to control expression of the phosphodiesterases. Nonetheless, phase variation by DNA inversion may be one understudied mechanism by which *C. difficile* generates phenotypic diversity.

Summary

Flagella contribute to *C. difficile* adherence and inflammation during infection (14). The majority of human-specific *C. difficile* strains contain genes for hierarchical assembly of flagella (106, 110). *C. difficile* produces flagella that aid in motility and adherence to the intestinal epithelium (114, 118, 119, 121). We and others demonstrated that SigD, the flagellar alternative sigma factor that activates late flagellar genes, indirectly activates the toxin genes, *tcdA* and *tcdB* (41, 109, 119). Both toxins inactivate substrates that regulate the actin cytoskeleton in epithelial cells to disrupt intestinal barrier function (10, 50, 69, 70). These events can lead to detection of

flagellin, the primary structural component of flagellum, by TLR5 on the basolateral epithelial surface (133). The synergy of *C. difficile* toxin activity and flagellin detection by TLR5 leads to the secretion of inflammatory cytokines and chemokines by epithelial cells and greater pathology during mouse infection (133, 136, 137). Yet clinical isolates of *C. difficile* exhibit phenotypic variability in flagellum production (114, 123, 165), and likely toxin production, that could contribute to evasion of host immune detection given their respective immunostimulatory profiles. What contributes to phenotypic heterogeneity in *C. difficile* flagella and toxin production, and the impact of this regulation on virulence are largely unknown.

We identified a new *cis*-acting regulator of flagellar gene expression in *C. difficile*. The leader RNA of the *flgB* operon is an unusual 498 nt compared to the average of 50 nt in *E. coli* (108). Previous work identified a c-di-GMP sensing riboswitch in the leader RNA, but this *cis*-acting regulatory element terminates transcription elongation in the 5' most 160 nt (108, 125). This begs the question, is there an additional *cis*-acting regulatory element present in the remaining 338 bp to control flagellar gene expression? Using bioinformatics tools, we identified a putative flagellar switch consisting of a 154 bp sequence, whose orientation varied among strains, flanked by 21 bp inverted repeats. We *hypothesized* that the flagellar switch enables phase variable production of flagella and toxins by DNA inversion in *C. difficile*. Phase variation involves the stochastic production of surface proteins in a bacterial population to generate phenotypic diversity for optimal fitness in diverse conditions (139). The *rationale* supporting our hypothesis is based on previous work showing phase variation of fimbriae in *E. coli* (153), flagella in *Salmonella enterica* serovar Typhimurium (156, 157), and a bacteriophage defense protein in *C. difficile* (159, 162) all by DNA inversion.

In our research, we have found that phase variation by DNA inversion controls production of flagella and toxins in *C. difficile* strain R20291 at a high frequency (Chapter 2). A single strain family designated as "ribotype 012" of *C. difficile* exhibits low frequency inversion of the flagellar switch in laboratory-adapted, environmental, and clinical isolates (Chapter 3). The gene regulatory mechanism by which the flagellar switch controls downstream flagellar gene expression involves the transcription terminator Rho factor (Chapter 4). And in the final section (Chapter 5), we further speculate on these published and unpublished studies in the context of gene regulation and virulence of *C. difficile* during host infection.
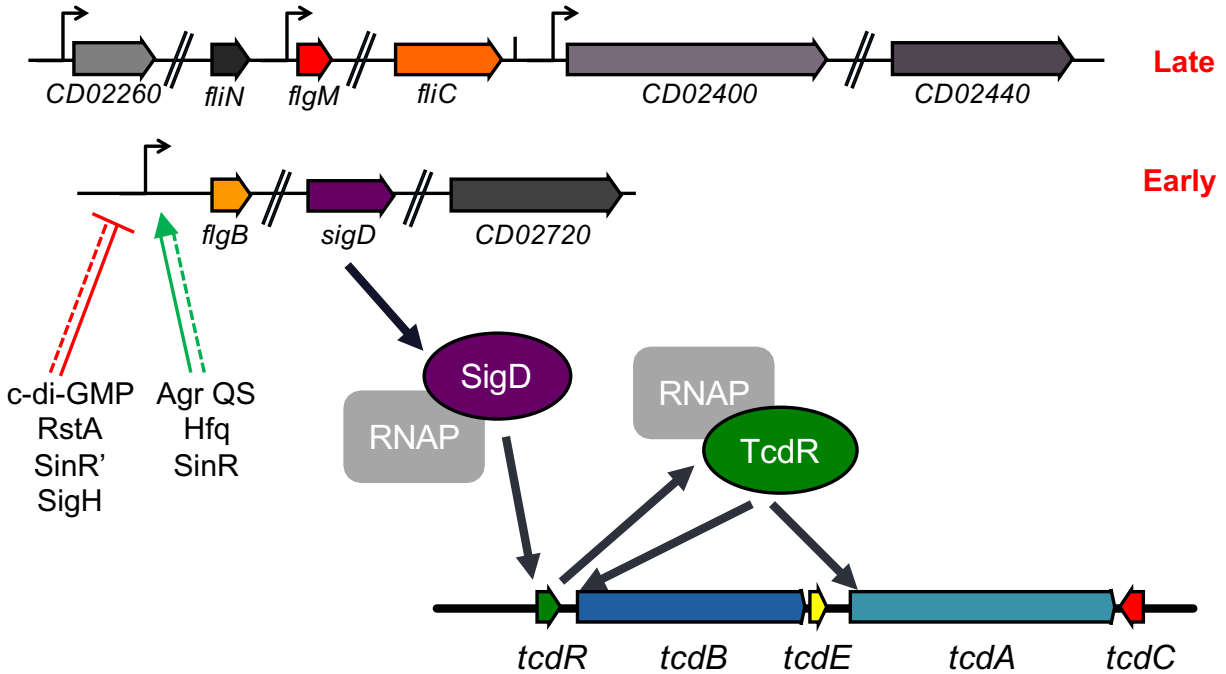
Figure 1.1. Positive and negative regulators of flagellar gene expression. A simplified schematic of the *flgB* operon with SigD regulation of the PaLoC. Negative regulators of flagellar gene expression include c-di-GMP, RstA, SinR', and SigH. Positive regulators of flagellar gene expression include Agr QS, Hfq, and SinR. The mechanisms by which many of these regulators control flagellar gene expression (promoter, leader, or post-translation) are largely unknown except for c-di-GMP through the Cd1 riboswitch (not shown). In addition, SigD can associate with RNAP to activate expression of late stage flagellar genes like *fliN, flgM,* and *fliC*. Lastly, SigD indirectly activates expression of *tcdA* and *tcdB* via *tcdR*.
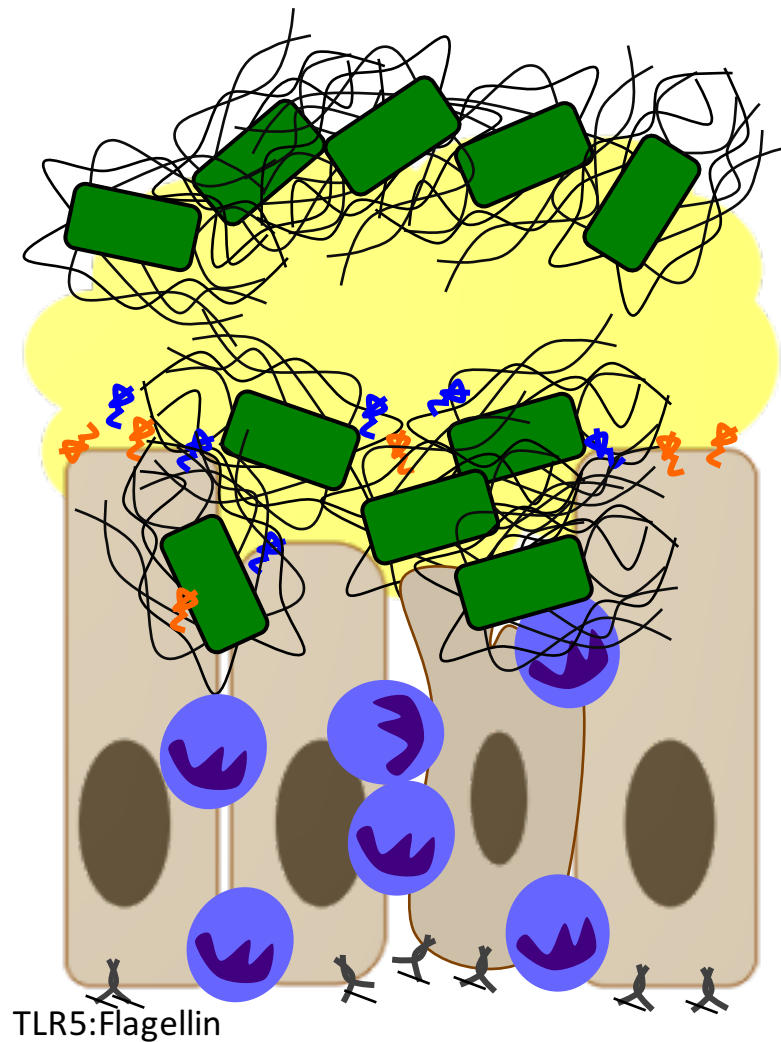
TLR5:Flagellin

Figure 1.2. Synergy of flagellin and toxin activity in the host intestine. *C. difficile* (depicted in green) engaging in swimming motility interacts with mucus (yellow) in the intestine and a subpopulation penetrates to adhere to intestinal epithelial cells (light brown). Under certain conditions*, C. difficile* produces TcdA and TcdB (orange and blue structures) that intoxicate host cells leading to disruption of the actin cytoskeleton and epithelial tight junctions. Flagellin monomers (black sticks) can interact with TLR5 on the basolateral surface after toxins have destroyed barrier function. TLR5 ligation stimulates signaling pathways that further potentiate inflammation and lead to the recruitment of neutrophils to the site of infection (purple circles). Model is based on previously published work (133, 136, 137, 166).
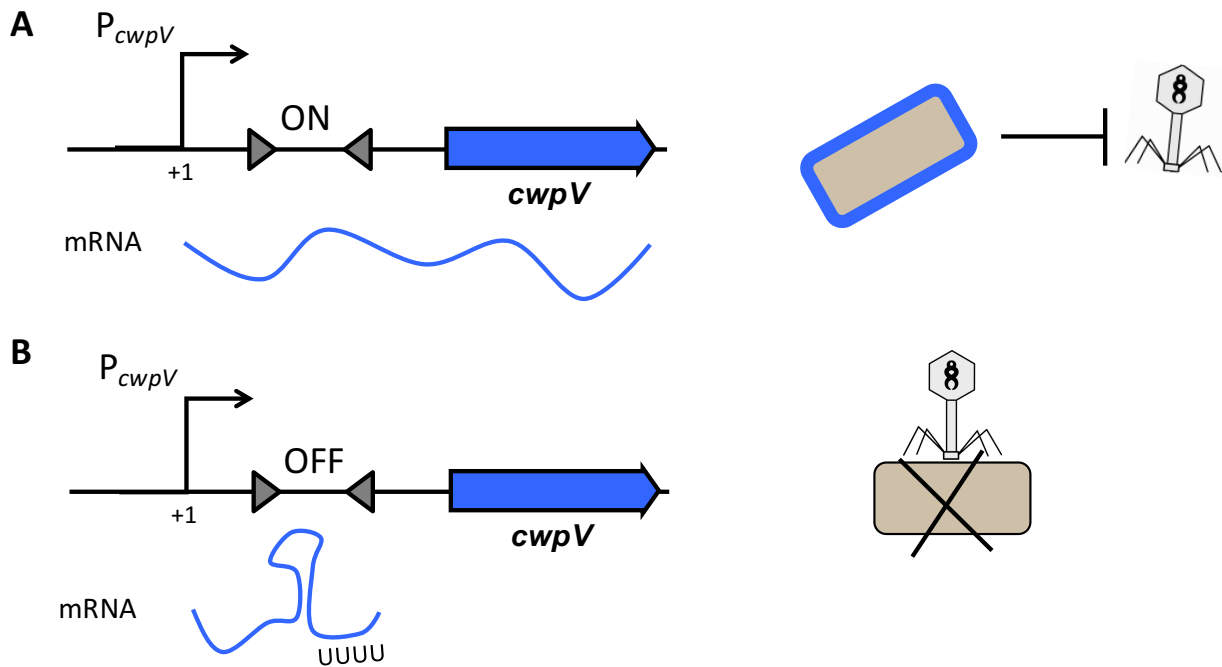
Figure 1.3. The *cwpV* switch affects *C. difficile* interaction with bacteriophage. (A) The *cwpV* switch is in the ON orientation and transcription elongation proceeds unhindered. CwpV is attached to the cell wall and resists predation by a bacteriophage. (B) The *cwpV* switch is in the OFF orientation and transcription elongation is interrupted by an intrinsic terminator resulting in premature termination and a failure to express *cwpV*. Lacking CwpV in the cell wall, *C. difficile* is susceptible to phage adsorption and tail spike DNA injection leading to lysis or death.

REFERENCES

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

2.   **Karen CC**, **John G B**. 2012. Biology of *Clostridium difficile*: implications for epidemiology and diagnosis. Annu Rev Microbiol **65**:501–521.

3.   **Smits WK**, **Lyras D**, **Lacy DB**, **Wilcox MH**, **Kuijper EJ**. 2016. *Clostridium difficile* infection. Nat Rev Dis Primers **2**:16020.

4.   **Lessa FC**, **Mu Y**, **Bamberg WM**, **Beldavs ZG**, **Dumyati GK**, **Dunn JR**, **Farley MM**, **Holzbauer SM**, **Meek JI**, **Phipps EC**, **Wilson LE**, **Winston LG**, **Cohen JA**, **Limbago BM**, **Fridkin SK**, **Gerding DN**, **McDonald LC**. 2015. Burden of *Clostridium difficile* infection in the United States. N Engl J Med **372**:825–834.

5.   **Freeman J**, **Bauer MP**, **Baines SD**, **Corver J**, **Fawley WN**, **Goorhuis B**, **Kuijper EJ**, **Wilcox MH**. 2010. The changing epidemiology of *Clostridium difficile* infections. Clin Microbiol Rev **23**:529–549.

6.   **Knetsch CW**, **Connor TR**, **Mutreja A**, **van Dorp SM**, **Sanders IM**, **Browne HP**, **Harris D**, **Lipman L**, **Keessen EC**, **Corver J**, **Kuijper EJ**, **Lawley TD**. 2014. Whole genome sequencing reveals potential spread of *Clostridium difficile* between humans and farm animals in the Netherlands, 2002 to 2011. Euro Surveill **19**:20954–262.

7.   **Theriot CM**, **Young VB**. 2015. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. Annu Rev Microbiol **69**:445–461.

8.   **McDonald LC**, **Gerding DN**, **Johnson S**, **Bakken JS**, **Carroll KC**, **Coffin SE**, **Dubberke ER**, **Garey KW**, **Gould CV**, **Kelly C**, **Loo V**, **Shaklee Sammons J**, **Sandora TJ**, **Wilcox MH**. 2018. Clinical practice guidelines for *Clostridium difficile* infection in adults and children: 2017 update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). Clin Infect Dis **66**:987–994.

9.   **Paredes-Sabja D**, **Shen A**, **Sorg JA**. 2014. *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends Microbiol **22**:406-416.

10.   **Chandrasekaran R**, **Lacy DB**. 2017. The role of toxins in *Clostridium difficile* infection. FEMS Microbiol Rev **41**:723–750.

11.   **Sorg JA**, **Sonenshein AL**. 2008. Bile salts and glycine as cogerminants for *Clostridium difficile* Spores. J Bacteriol **190**:2505–2512.

12.   **Kirk JA**, **Banerji O**, **Fagan RP**. 2017. Characteristics of the *Clostridium difficile* cell envelope and its importance in therapeutics. Microb Biotechnol **10**:76–90.

13. **Janoir C**. 2016. Virulence factors of *Clostridium difficile* and their role during infection. Anaerobe **37**:13–24.

14. **Stevenson E**, **Minton NP**, **Kuehne SA**. 2015. The role of flagella in *Clostridium difficile* pathogenicity. Trends Microbiol **23**:275–282.

15. **McKee RW**, **Aleksanyan N**, **Garrett EM**, **Tamayo R**. 2018. Type IV pili promote *Clostridium difficile* adherence and persistence in a mouse model of infection. Infect Immun 00943–17.

16. **Jenior ML**, **Leslie JL**, **Young VB**, **Schloss PD**, **Turnbaugh PJ**. 2017. *Clostridium difficile* colonizes alternative nutrient niches during infection across distinct murine gut microbiomes. mSystems **2**:e00063–17.

17. **Sun X**, **Hirota SA**. 2014. The roles of host and pathogen factors and the innate immune response in the pathogenesis of *Clostridium difficile* infection. Mol Immunol **63**:193-202.

18. **Fimlaid KA**, **Bond JP**, **Schutz KC**, **Putnam EE**, **Leung JM**, **Lawley TD**, **Shen A**. 2013. Global analysis of the sporulation pathway of *Clostridium difficile*. PLoS Genet **9**:e1003660.

19. **Pereira FC**, **Saujet L**, **Tomé AR**, **Serrano M**, **Monot M**, **Couture-Tosi E**, **Martin-Verstraete I**, **Dupuy B**, **Henriques AO**. 2013. The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. PLoS Genet **9**:e1003782.

20. **Saujet L**, **Pereira FC**, **Serrano M**, **Soutourina O**, **Monot M**, **Shelyakin PV**, **Gelfand MS**, **Dupuy B**, **Henriques AO**, **Martin-Verstraete I**. 2013. Genome-wide analysis of cell type-specific gene transcription during spore formation in *Clostridium difficile*. PLoS Genet **9**:e1003756.

21. **Cohen SH**, **Tang YJ**, **Silva J**. 2000. Analysis of the pathogenicity locus in *Clostridium difficile* strains. J Infect Dis **181**:659–663.

22. **Brouwer MSM**, **Roberts AP**, **Hussain H**, **Williams RJ**, **Allan E**, **Mullany P**. 2013. Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. Nat Commun **4**:2601.

23. **Mani N**, **Dupuy B**. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci USA **98**:5844–5849.

24. **Mani N**, **Lyras D**, **Barroso L**, **Howarth P**, **Wilkins T**, **Rood JI**, **Sonenshein AL**, **Dupuy B**. 2002. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J Bacteriol **184**:5971–5978.

25. **Girinathan BP**, **Monot M**, **Boyle D**, **McAllister KN**, **Sorg JA**, **Dupuy B**, **Govind R**. 2017. Effect of *tcdR* mutation on sporulation in the epidemic *Clostridium difficile* strain R20291. mSphere **2**: e00383-16.

26. **Matamouros S**, **England P**, **Dupuy B**. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol **64**:1274–1288.

27. **van Leeuwen HC**, **Bakker D**, **Steindel P**, **Kuijper EJ**, **Corver J**. 2013. *Clostridium difficile* TcdC protein binds four-stranded G-quadruplex structures. Nucleic Acids Res **41**:2382–2393.

28. **Carter GP**, **Douce GR**, **Govind R**, **Howarth PM**, **Mackin KE**, **Spencer J**, **Buckley AM**, **Antunes A**, **Kotsanas D**, **Jenkin GA**, **Dupuy B**, **Rood JI**, **Lyras D**. 2011. The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 Clinical isolate of *Clostridium difficile*. PLoS Pathog **7**:e1002317.

29. **Cartman ST**, **Kelly ML**, **Heeg D**, **Heap JT**, **Minton NP**. 2012. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdc genotype and toxin production. Appl Environ Microbiol **78**:4683–4690.

30. **Bakker D**, **Smits WK**, **Kuijper EJ**, **Corver J**. 2012. TcdC does not significantly repress toxin expression in *Clostridium difficile* 630Δ*erm*. PLoS ONE **7**:e43247.

31. **Govind R**, **Dupuy B**. 2012. Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. PLoS Pathog **8**:e1002727.

32. **Young R**. 2013. Phage lysis: do we have the hole story yet? Curr Opin Microbiol **16**:790–797.

33. **Lyerly DM**, **Lockwood DE**, **Richardson SH**, **Wilkins TD**. 1982. Biological activities of toxins A and B of *Clostridium difficile*. Infect Immun **35**:1147–1150.

34. **Lyerly DM**, **Saum KE**, **MacDonald DK**, **Wilkins TD**. 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun **47**:349–352.

35. **Wüst J**, **Sullivan NM**, **Hardegger U**, **Wilkins TD**. 1982. Investigation of an outbreak of antibiotic-associated colitis by various typing methods. J Clin Microbiol **16**:1096–1101.

36. **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

37. **Kuehne SA**, **Cartman ST**, **Heap JT**, **Kelly ML**, **Cockayne A**, **Minton NP**. 2010. The role of toxin A and toxin B in *Clostridium difficile* Infection. Nature **467**:711–713.

38. **O'Connor JR**, **Lyras D**, **Farrow KA**, **Adams V**, **Powell DR**, **Hinds J**, **Cheung JK**, **Rood JI**. 2006. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol **61**:1335–1351.

39. **Hussain HA**, **Roberts AP**, **Mullany P**. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* Strain 630 (630Δ*erm*) and demonstration that the

26. **Matamouros S**, **England P**, **Dupuy B**. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol **64**:1274–1288.

27. **van Leeuwen HC**, **Bakker D**, **Steindel P**, **Kuijper EJ**, **Corver J**. 2013. *Clostridium difficile* TcdC protein binds four-stranded G-quadruplex structures. Nucleic Acids Res **41**:2382–2393.

28. **Carter GP**, **Douce GR**, **Govind R**, **Howarth PM**, **Mackin KE**, **Spencer J**, **Buckley AM**, **Antunes A**, **Kotsanas D**, **Jenkin GA**, **Dupuy B**, **Rood JI**, **Lyras D**. 2011. The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 Clinical isolate of *Clostridium difficile*. PLoS Pathog **7**:e1002317.

29. **Cartman ST**, **Kelly ML**, **Heeg D**, **Heap JT**, **Minton NP**. 2012. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdc genotype and toxin production. Appl Environ Microbiol **78**:4683–4690.

30. **Bakker D**, **Smits WK**, **Kuijper EJ**, **Corver J**. 2012. TcdC does not significantly repress toxin expression in *Clostridium difficile* 630Δ*erm*. PLoS ONE **7**:e43247.

31. **Govind R**, **Dupuy B**. 2012. Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. PLoS Pathog **8**:e1002727.

32. **Young R**. 2013. Phage lysis: do we have the hole story yet? Curr Opin Microbiol **16**:790–797.

33. **Lyerly DM**, **Lockwood DE**, **Richardson SH**, **Wilkins TD**. 1982. Biological activities of toxins A and B of *Clostridium difficile*. Infect Immun **35**:1147–1150.

34. **Lyerly DM**, **Saum KE**, **MacDonald DK**, **Wilkins TD**. 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun **47**:349–352.

35. **Wüst J**, **Sullivan NM**, **Hardegger U**, **Wilkins TD**. 1982. Investigation of an outbreak of antibiotic-associated colitis by various typing methods. J Clin Microbiol **16**:1096–1101.

36. **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

37. **Kuehne SA**, **Cartman ST**, **Heap JT**, **Kelly ML**, **Cockayne A**, **Minton NP**. 2010. The role of toxin A and toxin B in *Clostridium difficile* Infection. Nature **467**:711–713.

38. **O'Connor JR**, **Lyras D**, **Farrow KA**, **Adams V**, **Powell DR**, **Hinds J**, **Cheung JK**, **Rood JI**. 2006. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol **61**:1335–1351.

39. **Hussain HA**, **Roberts AP**, **Mullany P**. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* Strain 630 (630Δ*erm*) and demonstration that the

conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. J Med Microbiol **54**:137–141.

40. **Collery MM**, **Kuehne SA**, **McBride SM**, **Kelly ML**, **Monot M**, **Cockayne A**, **Dupuy B**, **Minton NP**. 2016. What's a SNP between friends: the influence of single nucleotide polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and derivatives. Virulence 1–15.

41. **McKee RW**, **Mangalea MR**, **Purcell EB**, **Borchardt EK**, **Tamayo R**. 2013. The Second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of *sigD*. J Bacteriol **195**:5174–5185.

42. **Anjuwon-Foster BR**, **Maldonado-Vazquez N**, **Tamayo R**. 2018. Characterization of flagellar and toxin phase variation in *Clostridium difficile* ribotype 012 isolates. bioRxiv 256883.

43. **Kuehne SA**, **Collery MM**, **Kelly ML**, **Cartman ST**, **Cockayne A**, **Minton NP**. 2014. Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. J Infect Dis **209**:83–86.

44. **Carter GP**, **Chakravorty A**, **Pham Nguyen TA**, **Mileto S**, **Schreiber F**, **Li L**, **Howarth P**, **Clare S**, **Cunningham B**, **Sambol SP**, **Cheknis A**, **Figueroa I**, **Johnson S**, **Gerding D**, **Rood JI**, **Dougan G**, **Lawley TD**, **Lyras D**. 2015. Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during *Clostridium difficile* infections. MBio **6**:e00551.

45. **Monot M**, **Eckert C**, **Lemire A**, **Hamiot A**, **Dubois T**, **Tessier C**, **Dumoulard B**, **Hamel B**, **Petit A**, **Lalande V**, **Ma L**, **Bouchier C**, **Barbut F**, **Dupuy B**. 2015. *Clostridium difficile*: new insights into the evolution of the Pathogenicity Locus. Sci Rep **5**:15023.

46. **van den Berg RJ**, **Claas ECJ**, **Oyib DH**, **Klaassen CHW**, **Dijkshoorn L**, **Brazier JS**, **Kuijper EJ**. 2004. Characterization of toxin A-negative, toxin B-positive *Clostridium difficile* isolates from outbreaks in different countries by amplified fragment length polymorphism and PCR ribotyping. J Clin Microbiol **42**:1035–1041.

47. **Shin B-M**, **Kuak EY**, **Yoo SJ**, **Shin WC**, **Yoo HM**. 2008. Emerging toxin A-B+ variant strain of *Clostridium difficile* responsible for pseudomembranous colitis at a tertiary care hospital in Korea. Diagn Microbiol Infect Dis **60**:333–337.

48. **Cairns MD**, **Preston MD**, **Hall CL**, **Gerding DN**, **Hawkey PM**, **Kato H**, **Kim H**, **Kuijper EJ**, **Lawley TD**, **Pituch H**, **Reid S**, **Kullin B**, **Riley TV**, **Solomon K**, **Tsai PJ**, **Weese JS**, **Stabler RA**, **Wren BW**. 2017. Comparative genome analysis and global phylogeny of the toxin variant *Clostridium difficile* PCR ribotype 017 reveals the evolution of two independent sublineages. J Clin Microbiol **55**:865–876.

49. **Pruitt RN**, **Lacy DB**. 2012. Toward a structural understanding of *Clostridium difficile* toxins A and B. Front Cell Inf Microbio **2**:28.

50. **Aktories K**, **Schwan C**, **Jank T**. 2017. *Clostridium difficile* toxin biology. Annu Rev Microbiol **71**:281–307.

51. **Ho JGS**, **Greco A**, **Rupnik M**, **Ng KKS**. 2005. Crystal structure of receptor-binding C-terminal repeats from *Clostridium difficile* toxin A. Proc Natl Acad Sci U S A **102**:18373–18378.

52. **Pruitt RN**, **Chambers MG**, **Ng KKS**, **Ohi MD**, **Lacy DB**. 2010. Structural organization of the functional domains of *Clostridium difficile* toxins A and B. Proc Natl Acad Sci USA **107**:13467–13472.

53. **Dingle T**, **Wee S**, **Mulvey GL**, **Greco A**, **Kitova EN**, **Sun J**, **Lin S**, **Klassen JS**, **Palcic MM**, **Ng KKS**, **Armstrong GD**. 2008. Functional properties of the carboxy-terminal host cell-binding domains of the two toxins, TcdA and TcdB, expressed by *Clostridium difficile*. Glycobiology **18**:698–706.

54. **El-Hawiet A**, **Kitova EN**, **Kitov PI**, **Eugenio L**, **Ng KKS**, **Mulvey GL**, **Dingle TC**, **Szpacenko A**, **Armstrong GD**, **Klassen JS**. 2011. Binding of *Clostridium difficile* toxins to human milk oligosaccharides. Glycobiology **21**:1217–1227.

55. **Na X**, **Kim H**, **Moyer MP**, **Pothoulakis C**, **LaMont JT**. 2008. gp96 is a human colonocyte plasma membrane binding protein for *Clostridium difficile* toxin A. Infect Immun **76**:2862–2871.

56. **Olling A**, **Goy S**, **Hoffmann F**, **Tatge H**, **Just I**, **Gerhard R**. 2011. The repetitive oligopeptide sequences modulate cytopathic potency but are not crucial for cellular uptake of *Clostridium difficile* toxin A. PLoS ONE **6**:e17623.

57. **Orth P**, **Xiao L**, **Hernandez LD**, **Reichert P**, **Sheth PR**, **Beaumont M**, **Yang X**, **Murgolo N**, **Ermakov G**, **DiNunzio E**, **Racine F**, **Karczewski J**, **Secore S**, **Ingram RN**, **Mayhood T**, **Strickland C**, **Therien AG**. 2014. Mechanism of action and epitopes of *Clostridium difficile* toxin B-neutralizing antibody bezlotoxumab revealed by X-ray crystallography. J Biol Chem **289**:18008–18021.

58. **Yuan P**, **Zhang H**, **Cai C**, **Zhu S**, **Zhou Y**, **Yang X**, **He R**, **Li C**, **Guo S**, **Li S**, **Huang T**, **Perez-Cordon G**, **Feng H**, **Wei W**. 2014. Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B. Cell Res **25**:157-168.

59. **LaFrance ME**, **Farrow MA**, **Chandrasekaran R**, **Sheng J**, **Rubin DH**, **Lacy DB**. 2015. Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity. Proc Natl Acad Sci USA **112**:7073–7078.

60. **Tao L**, **Zhang J**, **Meraner P**, **Tovaglieri A**, **Wu X**, **Gerhard R**, **Zhang X**, **Stallcup WB**, **Miao J**, **He X**, **Hurdle JG**, **Breault DT**, **Brass AL**, **Dong M**. 2016. Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. Nature **538**:350-355.

61. **Papatheodorou P**, **Zamboglou C**, **Genisyuerek S**, **Guttenberg G**, **Aktories K**. 2010. Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. PLoS ONE

5:e10673.

62.   **Chandrasekaran R**, **Kenworthy AK**, **Lacy DB**. 2016. *Clostridium difficile* toxin A undergoes clathrin-independent, PACSIN2-dependent endocytosis. PLoS Pathog **12**:e1006070.

63.   **Qa'Dan M**, **Spyres LM**, **Ballard JD**. 2000. pH-induced conformational changes in *Clostridium difficile* toxin B. Infect Immun **68**:2470–2474.

64.   **Barth H**, **Pfeifer G**, **Hofmann F**, **Maier E**, **Benz R**, **Aktories K**. 2001. Low pH-induced formation of ion channels by *Clostridium difficile* toxin B in target cells. J Biol Chem **276**:10670–10676.

65.   **Reineke J**, **Tenzer S**, **Rupnik M**, **Koschinski A**, **Hasselmayer O**, **Schrattenholz A**, **Schild H**, **Eichel-Streiber von C**. 2007. Autocatalytic cleavage of *Clostridium difficile* toxin B. Nature **446**:415–419.

66.   **Egerer M**, **Giesemann T**, **Jank T**, **Satchell KJF**, **Aktories K**. 2007. Auto-catalytic cleavage of *Clostridium difficile* toxins A and B depends on cysteine protease activity. J Biol Chem **282**:25314–25321.

67.   **Pruitt RN**, **Chagot B**, **Cover M**, **Chazin WJ**, **Spiller B**, **Lacy DB**. 2009. Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in *Clostridium difficile* toxin A. J Biol Chem **284**:21934–21940.

68.   **Chumbler NM**, **Rutherford SA**, **Zhang Z**, **Farrow MA**, **Lisher JP**, **Farquhar E**, **Giedroc DP**, **Spiller BW**, **Melnyk RA**, **Lacy DB**. 2016. Crystal structure of *Clostridium difficile* toxin A. Nat Microbiol **1**:15002.

69.   **Just I**, **Wilm M**, **Selzer J**, **Rex G**, **Eichel-Streiber von C**, **Mann M**, **Aktories K**. 1995. The enterotoxin from *Clostridium difficile* (ToxA) monoglucosylates the Rho proteins. J Biol Chem **270**:13932–13936.

70.   **Just I**, **Selzer J**, **Wilm M**, **Eichel-Streiber von C**, **Mann M**, **Aktories K**. 1995. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature **375**:500–503.

71.   **Chen S**, **Sun C**, **Wang H**, **Wang J**. 2015. The role of Rho GTPases in toxicity of *Clostridium difficile* toxins. Toxins (Basel) **7**:5254–5267.

72.   **Nusrat A**, **Eichel-Streiber von C**, **Turner JR**, **Verkade P**, **Madara JL**, **Parkos CA**. 2001. *Clostridium difficile* toxins disrupt epithelial barrier function by altering membrane microdomain localization of tight junction proteins. Infect Immun **69**:1329–1336.

73.   **Kasendra M**, **Barrile R**, **Leuzzi R**, **Soriani M**. 2014. *Clostridium difficile* toxins facilitate bacterial colonization by modulating the fence and gate function of colonic epithelium. J Infect Dis **209**:1095–1104.

74. **Hecht G**, **Pothoulakis C**, **LaMont JT**, **Madara JL**. 1988. *Clostridium difficile* toxin A perturbs cytoskeletal structure and tight junction permeability of cultured human intestinal epithelial monolayers. J Clin Invest **82**:1516–1524.

75. **Hecht G**, **Koutsouris A**, **Pothoulakis C**, **LaMont JT**, **Madara JL**. 1992. *Clostridium difficile* toxin B disrupts the barrier function of T84 monolayers. Gastroenterology **102**:416–423.

76. **Fiorentini C**, **Fabbri A**, **Falzano L**, **Fattorossi A**, **Matarrese P**, **Rivabene R**, **Donelli G**. 1998. *Clostridium difficile* toxin B induces apoptosis in intestinal cultured cells. Infect Immun **66**:2660–2665.

77. **Gerhard R**, **Nottrott S**, **Schoentaube J**, **Tatge H**, **Olling A**, **Just I**. 2008. Glucosylation of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells. J Med Microbiol **57**:765–770.

78. **Xu H**, **Yang J**, **Gao W**, **Li L**, **Li P**, **Zhang L**, **Gong Y-N**, **Peng X**, **Xi JJ**, **Chen S**, **Wang F**, **Shao F**. 2014. Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. Nature **513**:237–241.

79. **Jorgensen I**, **Rayamajhi M**, **Miao EA**. 2017. Programmed cell death as a defence against infection. Nat Rev Immunol **17**:151–164.

80. **Chumbler NM**, **Farrow MA**, **Lapierre LA**, **Franklin JL**, **Haslam DB**, **Haslam D**, **Goldenring JR**, **Lacy DB**. 2012. *Clostridium difficile* toxin B causes epithelial cell necrosis through an autoprocessing-independent mechanism. PLoS Pathog **8**:e1003072.

81. **Farrow MA**, **Chumbler NM**, **Lapierre LA**, **Franklin JL**, **Rutherford SA**, **Goldenring JR**, **Lacy DB**. 2013. *Clostridium difficile* toxin B-induced necrosis is mediated by the host epithelial cell NADPH oxidase complex. Proc Natl Acad Sci USA **110**:18674–18679.

82. **Hundsberger T**, **Braun V**, **Weidmann M**, **Leukel P**, **Sauerborn M**, **Eichel-Streiber von C**. 1997. Transcription analysis of the genes *tcdA-E* of the Pathogenicity Locus of *Clostridium difficile*. Eur J Biochem **244**:735–742.

83. **Dupuy B**, **Sonenshein AL**. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Mol Microbiol **27**:107–120.

84. **Richardson AR**, **Somerville GA**, **Sonenshein AL**. 2015. Regulating the intersection of metabolism and pathogenesis in Gram-positive bacteria. Microbiol Spectr **3**.

85. **Antunes A**, **Martin-Verstraete I**, **Dupuy B**. 2010. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Mol Microbiol **79**:882–899.

86. **Sonenshein AL**. 2005. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr Opin Microbiol **8**:203–207.

87. **Dineen SS**, **McBride SM**, **Sonenshein AL**. 2010. Integration of metabolism and virulence by *Clostridium difficile* CodY. J Bacteriol **192**:5350–5362.

88. **Dineen SS**, **Villapakkam AC**, **Nordman JT**, **Sonenshein AL**. 2007. Repression of *Clostridium difficile* toxin gene expression by CodY. Mol Microbiol **66**:206–219.

89. **Dineen SS**, **McBride SM**, **Sonenshein AL**. 2010. Integration of metabolism and virulence by *Clostridium difficile* CodY. J Bacteriol **192**:5350–5362.

90. **Nawrocki KL**, **Edwards AN**, **Daou N**, **Bouillaut L**, **McBride SM**. 2016. CodY-dependent regulation of sporulation in *Clostridium difficile*. J Bacteriol **198**:2113–2130.

91. **Tan IS**, **Ramamurthi KS**. 2014. Spore formation in *Bacillus subtilis*. Environ Microbiol Rep **6**:212–225.

92. **Edwards AN**, **McBride SM**. 2014. Initiation of sporulation in *Clostridium difficile*: a twist on the classic model. FEMS Microbiol Lett **358**:110–118.

93. **Predich M**, **Nair G**, **Smith I**. 1992. *Bacillus subtilis* early sporulation genes *kinA*, *spo0F*, and *spo0A* are transcribed by the RNA polymerase containing sigma H. J Bacteriol **174**:2771–2778.

94. **Saujet L**, **Monot M**, **Dupuy B**, **Soutourina O**, **Martin-Verstraete I**. 2011. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol **193**:3186–3196.

95. **Rosenbusch KE**, **Bakker D**, **Kuijper EJ**, **Smits WK**. 2012. *C. difficile* 630Δ*erm* Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. PLoS ONE **7**:e48608.

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

97. **Pettit LJ**, **Browne HP**, **Yu L**, **Smits WK**, **Fagan RP**, **Barquist L**, **Martin MJ**, **Goulding D**, **Duncan SH**, **Flint HJ**, **Dougan G**, **Choudhary JS**, **Lawley TD**. 2014. Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics **15**:160.

98. **Edwards AN**, **Tamayo R**, **McBride SM**. 2016. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. Mol Microbiol **100**:954–971.

99. **Girinathan BP**, **Ou J**, **Dupuy B**, **Govind R**. 2018. Pleiotropic roles of *Clostridium difficile* sin locus. PLoS Pathog **14**:e1006940.

100. **Mandic-Mulec I**, **Doukhan L**, **Smith I**. 1995. The *Bacillus subtilis* SinR protein is a repressor of the key sporulation gene *spo0A*. J Bacteriol **177**:4619–4627.

101. **Rutherford ST**, **Bassler BL**. 2012. Bacterial quorum sensing: its role in virulence and possibilities for its control. Cold Spring Harb Perspect Med **2**:a012427.

102. **Martin MJ**, **Clare S**, **Goulding D**, **Faulds-Pain A**, **Barquist L**, **Browne HP**, **Pettit L**, **Dougan G**, **Lawley TD**, **Wren BW**. 2013. The *agr* locus regulates virulence and colonization genes in *Clostridium difficile* 027. J Bacteriol **195**:3672–3681.

103. **Darkoh C**, **Odo C**, **Dupont HL**. 2016. Accessory gene regulator-1 locus is essential for virulence and pathogenesis of *Clostridium difficile*. MBio **7**:e01237–16.

104. **Carter GP**, **Purdy D**, **Williams P**, **Minton NP**. 2005. Quorum sensing in *Clostridium difficile*: analysis of a luxS-type signalling system. J Med Microbiol **54**:119–127.

105. **Lee ASY**, **Song KP**. 2005. LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. Biochem Biophys Res Commun **335**:659–666.

106. **Stabler RA**, **Gerding DN**, **Songer JG**, **Drudy D**, **Brazier JS**, **Trinh HT**, **Witney AA**, **Hinds J**, **Wren BW**. 2006. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J Bacteriol **188**:7297–7305.

107. **Mukherjee S**, **Kearns DB**. 2014. The structure and regulation of flagella in Bacillus subtilis. Annu Rev Genet **48**:319–340.

108. **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**, **Severinov K**, **Le Bouguenec C**, **Coppée J-Y**, **Dupuy B**, **Martin-Verstraete I**. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

109. **Meouche El I**, **Peltier J**, **Monot M**, **Soutourina O**, **Pestel-Caron M**, **Dupuy B**, **Pons J-L**. 2013. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of *tcdR*. PLoS ONE **8**:e83748.

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

111. **Calvo RA**, **Kearns DB**. 2015. FlgM is secreted by the flagellar export apparatus in *Bacillus subtilis*. J Bacteriol **197**:81–91.

112. **Eichler J**, **Koomey M**. 2017. Sweet new roles for protein glycosylation in prokaryotes. Trends Microbiol **25**:662–672.

113. **Rossez Y**, **Wolfson EB**, **Holmes A**, **Gally DL**, **Holden NJ**. 2015. Bacterial flagella: twist and stick, or dodge across the kingdoms. PLoS Pathog **11**:e1004483.

114. **Twine SM**, **Reid CW**, **Aubry A**, **McMullin DR**, **Fulton KM**, **Austin J**, **Logan SM**. 2009. Motility and flagellar glycosylation in *Clostridium difficile*. J Bacteriol **191**:7050–7062.

115. **Faulds-Pain A**, **Twine SM**, **Vinogradov E**, **Strong PCR**, **Dell A**, **Buckley AM**, **Douce GR**, **Valiente E**, **Logan SM**, **Wren BW**. 2014. The post-translational modification of the *Clostridium difficile* flagellin affects motility, cell surface properties and virulence. Mol Microbiol **94**:272–289.

116. **Valiente E**, **Bouché L**, **Hitchen P**, **Faulds-Pain A**, **Songane M**, **Dawson LF**, **Donahue E**, **Stabler RA**, **Panico M**, **Morris HR**, **Bajaj-Elliott M**, **Logan SM**, **Dell A**, **Wren BW**. 2016. Role of glycosyltransferases modifying type b flagellin of emerging hypervirulent *Clostridium difficile* lineages and their impact on motility and biofilm formation. J Biol Chem **291**:25450–25461.

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

118. **Dingle TC**, **Mulvey GL**, **Armstrong GD**. 2011. Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect Immun **79**:4061–4067.

119. **Aubry A**, **Hussack G**, **Chen W**, **KuoLee R**, **Twine SM**, **Fulton KM**, **Foote S**, **Carrillo CD**, **Tanha J**, **Logan SM**. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun **80**:3521–3532.

120. **Twine SM**, **Reid CW**, **Aubry A**, **McMullin DR**, **Fulton KM**, **Austin J**, **Logan SM**. 2009. Motility and flagellar glycosylation in *Clostridium difficile*. J Bacteriol **191**:7050–7062.

121. **Baban ST**, **Kuehne SA**, **Barketi-Klai A**, **Cartman ST**, **Kelly ML**, **Hardie KR**, **Kansau I**, **Collignon A**, **Minton NP**. 2013. The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. PLoS ONE **8**:e73026.

122. **Hayashi F**, **Smith KD**, **Ozinsky A**, **Hawn TR**, **Yi EC**, **Goodlett DR**, **Eng JK**, **Akira S**, **Underhill DM**, **Aderem A**. 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. Nature **410**:1099–1103.

123. **Tasteyre A**, **Karjalainen T**, **Avesani V**, **Delmée M**, **Collignon A**, **Bourlioux P**, **Barc MC**. 2001. Molecular characterization of *fliD* gene encoding flagellar cap and its expression among *Clostridium difficile* isolates from different serogroups. J Clin Microbiol **39**:1178–1183.

124. **Purcell EB**, **Tamayo R**. 2016. Cyclic diguanylate signaling in Gram-positive bacteria. FEMS Microbiol Rev **40**:753–773.

125. **Sudarsan N**, **Lee ER**, **Weinberg Z**, **Moy RH**, **Kim JN**, **Link KH**, **Breaker RR**. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321**:411–413.

126. **Purcell EB**, **McKee RW**, **McBride SM**, **Waters CM**, **Tamayo R**. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol **194**:3307–3316.

127. **Boudry P**, **Gracia C**, **Monot M**, **Caillet J**, **Saujet L**, **Hajnsdorf E**, **Dupuy B**, **Martin-Verstraete I**, **Soutourina O**. 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. J Bacteriol **196**:3234–3248.

128. **Boudry P**, **Gracia C**, **Monot M**, **Caillet J**, **Saujet L**, **Hajnsdorf E**, **Dupuy B**, **Martin-Verstraete I**, **Soutourina O**. 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. J Bacteriol **196**:3234–3248.

129. **Kavita K**, **de Mets F**, **Gottesman S**. 2017. New aspects of RNA-based regulation by Hfq and its partner sRNAs. Curr Opin Microbiol **42**:53–61.

130. **Péchiné S**, **Janoir C**, **Collignon A**. 2005. Variability of *Clostridium difficile* surface proteins and specific serum antibody response in patients with *Clostridium difficile*-associated disease. J Clin Microbiol **43**:5018–5025.

131. **Andersen-Nissen E**, **Smith KD**, **Strobe KL**, **Barrett SLR**, **Cookson BT**, **Logan SM**, **Aderem A**. 2005. Evasion of Toll-like receptor 5 by flagellated bacteria. Proc Natl Acad Sci U S A **102**:9247–9252.

132. **Jarchum I**, **Liu M**, **Lipuma L**, **Pamer EG**. 2011. Toll-like receptor 5 stimulation protects mice from acute *Clostridium difficile* colitis. Infect Immun **79**:1498–1503.

133. **Yoshino Y**, **Kitazawa T**, **Ikeda M**, **Tatsuno K**, **Yanagimoto S**, **Okugawa S**, **Yotsuyanagi H**, **Ota Y**. 2013. *Clostridium difficile* flagellin stimulates toll-like receptor 5, and toxin B promotes flagellin-induced chemokine production via TLR5. Life Sci **92**:211–217.

134. **Gewirtz AT**, **Navas TA**, **Lyons S**, **Godowski PJ**, **Madara JL**. 2001. Cutting edge: bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression. J Immunol **167**:1882–1885.

135. **Rhee SH**, **Im E**, **Riegler M**, **Kokkotou E**, **O'brien M**, **Pothoulakis C**. 2005. Pathophysiological role of Toll-like receptor 5 engagement by bacterial flagellin in colonic inflammation. Proc Natl Acad Sci U S A **102**:13610–13615.

136. **Batah J**, **Denève-Larrazet C**, **Jolivot P-A**, **Kuehne S**, **Collignon A**, **Marvaud J-C**, **Kansau I**. 2016. *Clostridium difficile* flagella predominantly activate TLR5-linked NF-κB pathway in epithelial cells. Anaerobe **38**:116–124.

137. **Batah J**, **Kobeissy H**, **Bui Pham PT**, **Denève-Larrazet C**, **Kuehne S**, **Collignon A**,

Janoir-Jouveshomme C, **Marvaud J-C**, **Kansau I**. 2017. *Clostridium difficile* flagella induce a pro-inflammatory response in intestinal epithelium of mice in cooperation with toxins. Sci Rep **7**:3256.

138. **Merrigan M**, **Venugopal A**, **Mallozzi M**, **Roxas B**, **Viswanathan VK**, **Johnson S**, **Gerding DN**, **Vedantam G**. 2010. Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. J Bacteriol **192**:4904–4911.

139. **van der Woude MW**, **Bäumler AJ**. 2004. Phase and antigenic variation in bacteria. Clin Microbiol Rev **17**:581–611.

140. **van der Woude MW**. 2011. Phase variation: how to create and coordinate population diversity. Curr Opin Microbiol **14**:205–211.

141. **Bayliss CD**. 2009. Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. FEMS Microbiol Rev **33**:504–520.

142. **van der Woude MW**. 2006. Re-examining the role and random nature of phase variation. FEMS Microbiol Lett **254**:190–197.

143. **Zhou K**, **Aertsen A**, **Michiels CW**. 2014. The role of variable DNA tandem repeats in bacterial adaptation. FEMS Microbiol Rev **38**:119–141.

144. **Hendrixson DR**. 2006. A phase-variable mechanism controlling the *Campylobacter jejuni* FlgR response regulator influences commensalism. Mol Microbiol **61**:1646–1659.

145. **Hendrixson DR**. 2008. Restoration of flagellar biosynthesis by varied mutational events in *Campylobacter jejuni*. Mol Microbiol **70**:519–536.

146. **Hendrixson DR**, **DiRita VJ**. 2004. Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. Mol Microbiol **52**:471–484.

147. **Darmon E**, **Leach DRF**. 2014. Bacterial genome instability. Microbiol Mol Biol Rev **78**:1–39.

148. **Rotman E**, **Seifert HS**. 2014. The genetics of *Neisseria* species. Annu Rev Genet **48**:405–431.

149. **Quillin SJ**, **Seifert HS**. 2018. *Neisseria gonorrhoeae* host adaptation and pathogenesis. Nat Rev Micro **16**:226–240.

150. **Rajeev L**, **Malanowska K**, **Gardner JF**. 2009. Challenging a paradigm: the role of DNA homology in tyrosine recombinase reactions. Microbiol Mol Biol Rev **73**:300–309.

151. **Esposito D**, **Scocca JJ**. 1997. The integrase family of tyrosine recombinases: evolution

of a conserved active site domain. Nucleic Acids Res **25**:3605–3614.

152. **Smith MCM**, **Thorpe HM**. 2002. Diversity in the serine recombinases. Mol Microbiol **44**:299–307.

153. **Abraham JM**, **Freitag CS**, **Clements JR**, **Eisenstein BI**. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc Natl Acad Sci U S A **82**:5724–5727.

154. **McClain MS**, **Blomfield IC**, **Eisenstein BI**. 1991. Roles of fimB and fimE in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. J Bacteriol **173**:5308–5314.

155. **Gunther NW**, **Snyder JA**, **Lockatell V**, **Blomfield I**, **Johnson DE**, **Mobley HLT**. 2002. Assessment of virulence of uropathogenic *Escherichia coli* type 1 fimbrial mutants in which the invertible element is phase-locked on or off. Infect Immun **70**:3344–3354.

156. **Silverman M**, **Zieg J**, **Hilmen M**, **Simon M**. 1979. Phase variation in *Salmonella*: genetic analysis of a recombinational switch. Proc Natl Acad Sci U S A **76**:391–395.

157. **Ikeda JS**, **Schmitt CK**, **Darnell SC**, **Watson PR**, **Bispham J**, **Wallis TS**, **Weinstein DL**, **Metcalf ES**, **Adams P**, **O'Connor CD**, **O'Brien AD**. 2001. Flagellar phase variation of *Salmonella enterica* serovar Typhimurium contributes to virulence in the murine typhoid infection model but does not influence Salmonella-induced enteropathogenesis. Infect Immun **69**:3021–3030.

158. **Li J**, **Li J-W**, **Feng Z**, **Wang J**, **An H**, **Liu Y**, **Wang Y**, **Wang K**, **Zhang X**, **Miao Z**, **Liang W**, **Sebra R**, **Wang G**, **Wang W-C**, **Zhang J-R**. 2016. Epigenetic switch driven by dna inversions dictates phase variation in *Streptococcus pneumoniae*. PLoS Pathog **12**:e1005762.

159. **Emerson JE**, **Reynolds CB**, **Fagan RP**, **Shaw HA**, **Goulding D**, **Fairweather NF**. 2009. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol **74**:541–556.

160. **Reynolds CB**, **Emerson JE**, **la Riva de L**, **Fagan RP**, **Fairweather NF**. 2011. The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog **7**:e1002024.

161. **Sekulovic O**, **Fortier L-C**. 2015. Global transcriptional response of *Clostridium difficile* carrying the CD38 prophage. Appl Environ Microbiol **81**:1364–1374.

162. **Sekulovic O**, **Ospina Bedoya M**, **Fivian-Hughes AS**, **Fairweather NF**, **Fortier L-C**. 2015. The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. Mol Microbiol **98**:329–342.

163. **Bordeleau E**, **Fortier L-C**, **Malouin F**, **Burrus V**. 2011. c-di-GMP turn-over in

Clostridium difficile is controlled by a plethora of diguanylate cyclases and phosphodiesterases. PLoS Genet **7**:e1002039.

164. **Gao X**, **Dong X**, **Subramanian S**, **Matthews PM**, **Cooper CA**, **Kearns DB**, **Dann CE**. 2014. Engineering of *Bacillus subtilis* strains to allow rapid characterization of heterologous diguanylate cyclases and phosphodiesterases. Appl Environ Microbiol **80**:6167–6174.

165. **Tasteyre A**, **Karjalainen T**, **Avesani V**, **Delmée M**, **Collignon A**, **Bourlioux P**, **Barc MC**. 2000. Phenotypic and genotypic diversity of the flagellin gene (*fliC*) among *Clostridium difficile* isolates from different serogroups. J Clin Microbiol **38**:3179–3186.

166. **Tasteyre A**, **Barc M-C**, **Collignon A**, **Boureau H**, **Karjalainen T**. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect Immun **69**:7937–7940.

CHAPTER 2: A GENETIC SWITCH CONTROLS THE PRODUCTION OF FLAGELLA AND TOXINS IN *CLOSTRIDIUM DIFFICILE*[1]

Summary

Clostridium difficile is a bacterial pathogen that causes antibiotic associated diarrheal disease. Colonization of host tissues is a prerequisite step to disease development during bacterial infection. *C. difficile* produces flagella, which are proteinaceous structures on the bacterial surface that confer motility and participate in adherence to the host intestine. SigD, a regulator that coordinates flagellar gene expression, also activates expression of the toxin genes in *C. difficile*. Therefore, mechanisms controlling expression of flagellar genes, including *sigD*, will likely impact the severity of *C. difficile* infection given the impact on toxin production. In this body of work, we identified a genetic ON/OFF switch, which we term the "flagellar switch", encoded upstream of the flagellar genes. The orientation of the flagellar switch determines whether or not *C. difficile* produce flagella, engage in swimming motility, and secrete toxins. We identified the enzyme that catalyzes inversion of the flagellar switch. Inactivation of the

---

[1] This chapter contains a manuscript that previously appeared as an article in *PLOS Genetics*. I performed the experiments for Figures 2.1 to 2.20, except Dr. Rita Tamayo generated the R20291 *sigD* mutant in Figure 2.9 that was used in Figures 2.10, 2.11, 2.18, and 2.19. Rita and I wrote the original draft and revised version after peer review. This is the authors' version of the work. The full citation of the published version is: Anjuwon-Foster BR, Tamayo R. 2017. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet **13**:e1006701.

corresponding gene resulted in bacteria with the flagellar switch locked in the either the ON or OFF orientation, with concomitant effects on flagellum and toxin biosynthesis. The flagellar switch may represent a new regulatory strategy to coordinately control virulence determinants independent of previously described regulators.

Introduction

Clostridium difficile, a Gram-positive, spore-forming obligate anaerobe, is the leading cause of nosocomial disease in North America, Europe and Australia (1, 2). The Centers for Disease Control and Prevention list *C. difficile* infections (CDI) as an urgent threat related to the use of antibiotics (3). Antibiotic use perturbs the gastrointestinal microbiota that normally protects against CDI (1, 4). The rates of recurrence and mortality associated with CDI have increased in part due to the emergence of epidemic-associated strains with enhanced sporulation rates and toxin production (1, 5, 6). The *C. difficile* PCR ribotype 027 group is associated with greater odds of diarrheal disease severity, outcome, and death compared to many other PCR ribotypes (7). Therefore, an understanding of bacterial physiology and genetics in different strains of *C. difficile* ribotype 027 could reveal unique therapeutic or diagnostic targets to ameliorate severe CDI.

C. difficile is primarily transmitted as metabolically dormant spores, which germinate into actively growing vegetative cells in response to amino acids and bile salts such as glycine-conjugated taurocholate (8-10). CDI ranges in severity from mild self-limiting diarrhea to fulminant colitis characterized by neutrophil infiltration into the lamina propria, erosion of crypts and goblet cells, and extensive epithelial tissue damage. Diarrheal disease is associated with strains that produce the glucosylating toxins TcdA and/or TcdB (11, 12). Both TcdA and TcdB

glucosylate Rho and Rac GTPases in host cells to promote actin depolymerization, destruction of tight junctions at the mucosal barrier, and inflammation (13-16). The glucosylating toxins are necessary for diarrheal disease in both the mouse and hamster models of infection, with TcdB playing a more prominent role (11, 12, 17).

Colonization is a requisite step to diarrheal disease, and the bacterial surface structures, such as flagella, that participate in colonization are an active area of study. The flagellar apparatus confers motility and contributes to adherence, colonization, and disease in *C. difficile* (18-20). Flagellum biosynthesis genes are conserved in most sequenced *C. difficile* strains (19, 21). As in other bacterial species, flagellar gene transcription occurs in a hierarchical order to ensure proper protein assembly and to conserve energy. The early stage flagellar operon contains genes for assembly of the basal body. SigD (σ^D), the flagellar alternative sigma factor, is encoded in the early stage flagellar operon and activates the transcription of the late stage operons (22, 23). At least four operons contain late stage flagellar genes involved in assembly of the flagellar hook, filament, and cap, and for post-translational modification of the flagellar filament (18, 22, 24-26). Flagellar arrangement is peritrichous in most strains (18, 27-31).

The flagellar apparatus contributes to *C. difficile* fitness during gastrointestinal infection in a ribotype-dependent manner (20). In *C. difficile* 630Δ*erm* (ribotype 012), flagella are dispensable for adherence to cultured epithelial cells and colonization in the murine model (31, 32). In contrast, in the epidemic-associated *C. difficile* strain R20291 (ribotype 027), flagellar filaments promote adherence and colonization *in vitro* and during infection of mice (31). Flagellar motility is not required for adherence in this strain, as a MotB mutant that produces flagellar filaments with a nonfunctional motor displays wild type adherence (31). Interestingly, mice infected with the R20291 *fliC* mutant succumb to infection, whereas mice infected with the

parental strain do not (31). The R20291 *fliC* mutant has altered expression of genes involved in motility, membrane transport, sporulation, and metabolism *in vitro* (33), which may explain the enhanced virulence in mice. Post-translational modification of flagella also contributes to colonization kinetics in a mouse relapse model of infection in M68, an 017 ribotype strain (24). In the hamster model, *C. difficile* 630*Δerm* strains with a mutation in several early stage flagellar genes showed reduced virulence, whereas mutations in the late stage flagellar genes *fliC* and *fliD* generally showed increased virulence (22, 32). Several groups have observed that a 630*Δerm* *fliC* mutant produces more toxin, which may contribute to the increased virulence of the mutant (22, 31, 32). The animal model-dependent differences in virulence phenotypes of the flagellar gene mutants may be attributed to the greater sensitivity of hamsters to the glucosylating toxins (34), and to higher spore germination rates of *C. difficile* spores in hamsters than in mice (35).

Notably, the expression of the glucosylating toxin genes is linked to flagellar gene expression in *C. difficile* (20, 36). TcdA and TcdB are encoded on a Pathogenicity Locus (PaLoc) along with the TcdR sigma factor that positively regulates *tcdA* and *tcdB* (37), the anti-sigma factor TcdC suggested to inhibit TcdR function (38-41), and the TcdE holin-like protein suggested to be involved in toxin export (42, 43). Preliminary studies suggested that σ^D affects the expression of the glucosylating toxin genes (22). Subsequent work identified a σ^D consensus sequence in the *tcdR* promoter (23, 44). Overexpression of *sigD* in *C. difficile* increases toxin production in a TcdR-dependent manner (44), and recombinant σ^D with RNA polymerase directly binds the *tcdR* promoter (23). Taken together, these studies highlight a regulatory link between virulence factors critical to host colonization and to disease symptom development in *C. difficile*, similar to other diarrheal bacterial pathogens, such as *Vibrio cholerae* and *Campylobacter jejuni* (45, 46).

C. difficile flagellin and the glucosylating toxins stimulate pathogen recognition receptors (47-49), which promote pathogen clearance mechanisms, and therefore their production must be subject to precise regulation to avoid host recognition. In addition to σ^D, several transcriptional regulators modulate toxin production in *C. difficile*, potentially via regulation of flagellar gene expression. Spo0A, SigH, and RstA inhibit expression of the flagellar and toxin genes (50-53), while Agr quorum sensing and Hfq are positive regulators (29, 54). The exact mechanisms by which these regulators control flagellar and toxin gene expression are largely unknown.

Cyclic diguanylate (c-di-GMP) is a nucleotide second messenger that controls a multitude of bacterial processes, such as flagellar motility, biofilm formation, and virulence factor expression (55-58). In *C. difficile*, c-di-GMP regulates swimming motility, cytopathicity, Type IV-pilus dependent surface motility and biofilm formation (30, 44, 59, 60). Specifically, elevated c-di-GMP inhibits flagellar gene expression and swimming motility (59), and also negatively regulates toxin gene expression as a result of reduced *sigD* transcription (44). Flagellar gene regulation occurs via a c-di-GMP sensing riboswitch, Cd1, located in the 5' untranslated region of the early stage flagellar operon (61). c-di-GMP binding to Cd1 causes premature transcription termination, resulting in a truncated transcript of 160 nt of the 5' untranslated region (62).

Here, we describe the identification and characterization of an additional *cis*-acting regulatory element that mediates phase variable expression of flagellar genes and, consequently, the toxin genes. We identified a flagellar switch consisting of 154 bp flanked by 21 bp imperfect inverted repeats, between the Cd1 riboswitch and the first open reading frame in the early stage flagellar operon. The orientation of the flagellar switch controls downstream flagellar gene expression, including *sigD*, and therefore production of flagella and swimming motility.

Furthermore, the flagellar switch affects transcription of the toxin genes by way of σ^D, impacting

the production of the glucosylating toxins and the cytotoxicity of *C. difficile*. We provide

evidence that regulation through the flagellar switch occurs post-transcription initiation. Lastly,

we identified RecV, which also controls phase variation of the cell wall protein CwpV, as the

recombinase responsible for inversion at the flagellar switch in both directions. Together these

findings indicate that flagellar motility and toxin production are subject to phase variation in *C.*

difficile, thus identifying an additional level of regulation of these linked processes. Flagellar and

toxin phase variation in *C. difficile* may confer an advantage to the bacterium during infection of

the gastrointestinal tract as a bet hedging strategy to either promote colonization and

inflammation at certain tissue sites or persist and evade host immune stimulation.

Results

In silico identification of a flagellar switch upstream of the early flagellar biosynthesis operon

Soutourina, *et. al.*, identified a single transcriptional start site (TSS) for the early flagellar

operon in *C. difficile* strain 630Δ*erm*, located nearly 500 nucleotides (nt) upstream of the start

codon for *flgB*, the first gene in this operon (62). The *flgB* operon is also termed the "F3 locus"

elsewhere (21). The promoter for the *flgB* operon contains a σ^A consensus sequence (62). The

first 160 nt of the 5' untranslated region (UTR) of the *flgB* operon contains a class I (GEMM) c-

di-GMP riboswitch, previously named Cd1 (61). Binding of c-di-GMP to Cd1 causes premature

transcription termination, preventing flagellar gene transcription and inhibiting swimming

motility (59, 61, 62). Notably, an additional 338 nt lie between the Cd1 termination site and the

start codon of *flgB*. We speculated that the remaining 338 nt of the 5' UTR contains an additional

regulatory element that controls expression of the *flgB* operon. To explore this possibility, we

used Clustal Omega (63) to generate multiple sequence alignments and examine the 5' UTR of the *flgB* operon for regions of reduced sequence identity. Assuming the TSS is conserved in all strains, we found that all *C. difficile* genome sequences currently available through NCBI contain a 498 nt 5' UTR, except 630 which has a 496 nt 5' UTR. Strains of the 027 ribotype, which is associated with epidemics of CDI, were used for further analysis: CD196 (NCBI Accession No. FN538970.1), R20291 (FN545816.1), 2007855 (FN665654.1), and BI1 (FN668941.1) (19, 64, 65). The 5' UTRs of three strains exhibited 100% sequence identity, but BI1 had only 66.2% identity in a 154 bp region (Fig. 2.1A). Upon further scrutiny of the region, we identified near perfect 21 bp inverted repeats flanking the 154 bp region: 5'-AAGTT(A/G)CTAT(A/T)TTACAAAAAA-3' (Left inverted repeat). The presence of a sequence flanked by inverted repeats located upstream of genes encoding an immunostimulatory surface structure suggested a regulatory mechanism involving phase variation by site specific DNA recombination. These features reflect a regulatory mechanism employed by numerous mucosal bacterial pathogens to stochastically control the expression of cell surface structures that may be immunostimulatory, therefore promoting immune evasion and persistence in a host (66). We reasoned that the reduced sequence identity for BI1 at the 154 bp region could be due to DNA inversion between the inverted repeats. To test this, we repeated the multiple sequence alignment using the antiparallel, "inverse" sequence for the 154 bp sequence for BI1 and observed restoration of 100% identity to the other ribotype 027 sequences (Fig. 2.1A). Based on these *in silico* findings, we hypothesized that the 154 bp sequence undergoes inversion via site-specific recombination, and we term this the "flagellar switch" herein.

<u>The flagellar switch undergoes DNA inversion</u>

If the 154 bp sequence is capable of undergoing DNA inversion, we expect to detect the flagellar switch in both of the orientations in *C. difficile*, at least under some growth conditions. To test this, we used an orientation-specific PCR assay to detect and differentiate between the two orientations of the putative flagellar switch in multiple *C. difficile* strains grown in liquid medium (Fig. 2.1B) (67, 68). We use "published orientation" to refer to the sequence present in the indicated published genome for the given strain, and "inverse orientation" to refer to a sequence with an inversion between the inverted repeats. A common reverse primer, Rv, complementary to the *flgB* coding sequence was used in all PCRs (Fig. 2.1B, black). To detect the flagellar switch in the published orientation, we used a ribotype-specific forward primer F1 that anneals immediately 3' of the left inverted repeat (LIR) (Fig. 2.1B, green). To detect the flagellar switch in the inverse orientation, we used a forward primer F2 that anneals immediately 5' of the right inverted repeat (RIR) (Fig. 2.1B, red). Primers F1 and F2 are reverse-complementary. The primer pairs yield different products of different sizes depending on the orientation of the template sequence, and the sizes vary somewhat depending on the strain. We detected PCR products for both the published and inverse orientations for R20291 (FN545816.1) and ATCC 43598 (017 ribotype, sequence read archive SRX656590) (Fig. 2.1C). DNA sequencing of the PCR products confirmed the orientation of the template sequences. Furthermore, DNA sequencing suggests that DNA strand exchange for recombination would have to occur at or after position 12 of the IRs, because the 6th and 11th nucleotides that are not conserved do not change. We were unable to detect a PCR product for the inverse orientation for 630*Δerm* (Fig. 2.1C), a commonly used laboratory-adapted strain. Unlike the other published *C. difficile* genomes, 630Δ*erm* and its parental strain, 630, have shorter 20 bp IRs with an adenosine

absent from the 3' end of the LIR and a thymidine absent from the 5' end of the RIR. The length

of IRs can impact DNA inversion (68), which may explain the lack of an inverse product for

630Δ*erm*. Collectively, these data indicate that at least some *C. difficile* strains grow in broth

culture as a heterogeneous population of bacteria with the flagellar switch in the published or

inverse orientation, supporting our hypothesis that the flagellar switch undergoes site-specific

recombination. Based on the previously identified inversion sites in the *C. difficile* genomes (69),

we designate the flagellar switch site as Cdi4.

Quantifying the frequency of flagellar switch orientation in enriched flagellar phase variant

populations

We optimized an unbiased asymmetric PCR-digestion assay to distinguish the orientation

of the switch and determine the relative proportions of each in the population (Fig. 2.2A) (67, 68,

70). One pair of primers that flank the flagellar switch amplify a region encompassing the 5'

UTR *flgB* coding sequence of the *flgB* operon, yielding a 665 bp product that was subjected to

digestion with the restriction enzyme SwaI. The presence of a single SwaI restriction site in the

154 bp flagellar switch resulted in two fragments of different sizes depending on the orientation

of the sequence (Fig. 2.2A). If the flagellar switch is in the published orientation, we expect

fragments of 312 and 353 bp; if the switch is in the inverse orientation, we expect fragments of

418 and 247 bp. All four products are expected for a mixed population.

Spore preparations of *C. difficile* R20291, which contain a mix of bacteria with respect to

flagellar switch orientation (Fig. 2.3), were plated on BHIS agar supplemented with the

germinant taurocholate. Germinated colonies were collected and spotted onto standard BHIS

agar. We observed that over 96 hours, the R20291 colony morphology changed from smooth and

circular to a spreading, filamentous colony (Fig. 2.2B). Genomic DNA was isolated from

replicate colonies every 24 hours, and the asymmetric PCR-digestion assay was performed after

all samples were processed. Over time, the abundance of the SwaI digested products indicating

the inverse orientation increased (Fig. 2.2C). The proportions of bacteria with the switch in the

two orientations were quantified by measuring the pixel intensities of the bands, normalized to a

standard curve generated with titrated amounts of SwaI-digested DNA template for the published

and inverse orientations (71). The proportion of the bacteria with the flagellar switch in the

inverse orientation significantly increased after 48 hours (median of 33.8%, $P < 0.005$), 72 hours

(39.0%, $P < 0.05$), and 96 hours (67.4%, $P < 0.05$) compared to 24 hours (26.4%) (Fig. 2.2D).

These data provide evidence of DNA inversion and suggest that growth on a solid agar surface

favors the accumulation of *C. difficile* with the flagellar switch in the inverse orientation.

Single colonies derived from this growth (24-96 hours on BHIS agar) were subjected to

the asymmetric PCR-digestion assay to determine the orientation of the flagellar switch.

Individual colonies were enriched for one orientation of the flagellar switch. All colonies tested

yielded either 312 and 353 bp fragments indicating the switch in the published orientation (Fig.

2.4A) or 418 and 247 bp fragments indicating the switch in the inverse orientation (Fig. 2.4B).

However, the asymmetric PCR-digestion has limited sensitivity for detection of low abundance

targets. Thus, we used quantitative PCR (qPCR) of genomic DNA to quantify the frequency of

bacteria with the flagellar switch in each orientation. In isolated colonies that showed fragments

indicating the switch is in the published orientation by asymmetric PCR-digestion (Fig. 2.4A),

89.5% (+/- 6.6% SD) of the population had the switch in that orientation by qPCR (Fig. 2.4C).

Similarly, in colonies with the switch predominantly in the inverse orientation by asymmetric

PCR-digestion (Fig. 2.4B), 96.3% (+/- 2.2% SD) of the population has the switch in that

orientation by qPCR (Fig. 2.4C). These data indicate that asymmetric PCR-digestion successfully identifies colonies that are enriched for the flagellar switch in a single orientation. Furthermore, the orientation of the flagellar switch remains stable during growth on an agar surface for at least 24 hours and switch orientation did not affect growth (Fig. 2.5).

We note that during these experiments, we observed two distinct colony morphologies based on colony texture and edge– a smooth, circular (SC) morphotype, and a rough, filamentous (RF) morphotype (Fig. 2.6). The percentage of RF colonies increased, and the percentage of SC colonies decreased over time (Fig. 2.6). Bacteria from SC and RF colonies maintained their respective characteristic morphologies after passaging (Fig. 2.6). Colony morphology was not attributable to the orientation of the flagellar switch, as 72% of bacteria in the SC colonies had the flagellar switch in the published orientation, and 64% of bacteria in the RF colonies had the flagellar switch in the inverse orientation (Fig. 2.6).

The orientation of the flagellar switch controls downstream flagellar gene expression

The enrichment of two populations based on the orientation of the flagellar switch allowed us to determine the impact of switch orientation on downstream gene expression. We purified colonies with the flagellar switch in the published and inverse orientations (Fig. 2.7A) to assess gene expression. The abundances of early and late stage flagellar gene transcripts in the isolates were compared by quantitative reverse transcriptase PCR (qRT-PCR). Bacteria with the flagellar switch in the inverse orientation exhibited significantly reduced abundance of early stage flagellar gene transcripts, such as *flgB* and *sigD/fliA*, compared to bacteria with the flagellar switch in the published orientation (Fig. 2.7B) (23, 44). Accordingly, the abundances of the σ^D-dependent, late stage flagellar gene transcripts CDR20291_0227 (autolysin), *flgM*

(flagellar anti-sigma factor), and *fliC* (flagellin) (Fig. 2.7B) were also significantly decreased in bacteria with the flagellar switch in the inverse orientation. Consistent with this, transmission electron microscopy showed that the majority of bacteria from colonies with the flagellar switch in the published orientation displayed flagella (72% flagellated, n = 144), whereas bacteria from colonies with the switch in the inverse orientation were largely non-flagellated (98% non-flagellated, n = 163) (Fig. 2.7C). We also note that *C. difficile* strain R20291 was previously reported to produce a single flagellum (31), but we observed peritrichous flagella on our R20291 isolate.

In addition, we generated a transcriptional fusion of *mCherryOpt*, encoding a Red Fluorescent Protein-derivative (RFP) optimized for translation in *C. difficile* (72, 73), to the *flgM* promoter. The *flgM* gene was previously determined to be positively regulated by σ^D (23, 44). As σ^D is encoded in the *flgB* operon, it is indirectly phase variable and serves as an indicator of the status of the flagellar switch. This plasmid-borne reporter (P$_{flgM}$*::mCherryOpt*) was introduced into *C. difficile* R20291 isolates with the switch in the published or inverse orientation. In the isolate with the switch in the published orientation, the majority of bacteria were mCherryOpt positive: 76.7% (+/- 9.8% SD). In the isolate with the switch in the inverse orientation, few bacteria were mCherryOpt positive: 2.4% (+/- 2.2% SD) (Fig. 2.7D). These values are consistent with those obtained using qPCR to directly evaluate the switch orientation in these populations (Fig. 2.4C). Differences in the frequency of the switch orientation in enriched *flg* ON populations could be due to additional flagellar gene regulators that function independently of the *flg* switch, such as c-di-GMP (59, 62). Negative controls bearing promoterless *mCherryOpt (::mCherryOpt)* lacked fluorescence (Fig. 2.8); positive controls with inducible *mCherryOpt* (P$_{tet}$::*mCherryOpt*) showed more uniform red fluorescence (Fig. 2.8). We additionally constructed a *sigD* mutant in

61

R20291 by Targetron insertional mutagenesis as a negative control (Fig. 2.9). The *sigD* mutant

containing the P*flgM*::*mCherryOpt* reporter lacked mCherryOpt fluorescence, indicating that *flgM*

promoter activity is dependent on σ^D, and therefore on the flagellar switch (Fig. 2.7C).

Taken together, these data indicate that the orientation of the flagellar switch controls

flagellar gene expression and are consistent with phase variable gene expression. Thus, bacteria

with the flagellar switch in the published orientation express flagellar genes and produce flagella

and are thus flagellar phase ON ("*flg* ON" hereafter); bacteria with the switch in the inverse

orientation show decreased flagellar gene expression and flagellum biosynthesis and are

comparatively flagellar phase OFF ("*flg* OFF" hereafter).

Motility medium spatially segregates flagellar phase variant populations

Given that *flg* OFF bacteria were deficient in flagellum biosynthesis (Fig. 2.7), we

predicted that *flg* OFF bacteria would be non-motile compared to motile *flg* ON bacteria. To test

this, enriched *flg* ON and OFF isolates were examined for the ability to swim through BHIS-

0.3% agar (59). The R20291 *sigD* mutant was used as a non-motile control. The *flg* OFF bacteria

appeared motile (Fig. 2.10A), though they displayed a modest but statistically significant

reduction in swimming diameter at 24 and 48 hours post-inoculation compared to *flg* ON isolates

(Fig. 2.10B).

We reasoned that the enriched *flg* OFF inoculum would result in spatial restriction of

non-motile bacteria to the inoculation site. In contrast, the small fraction of *flg* ON bacteria

(detectable by qPCR and fluorescent reporters but below the limitation of detection by

asymmetric PCR digestion assay) would be capable of motility and chemotaxis and therefore

expand from the inoculation site, spatially segregating from *flg* OFF bacteria. To test this, for *flg*

ON and OFF isolates cultured in motility agar, we sampled bacteria from the inoculation site (center) and from the leading edge at 24, 48, and 72 hours and determined the orientation of the switch using the asymmetric PCR-digestion assay. As expected, *flg* ON maintained the flagellar switch in the ON orientation at all time points at both the center and edge (Fig. 2.10C, top). In contrast, *flg* OFF bacteria from the center contained the switch primarily in the OFF configuration, but bacteria at the leading edge had the switch to the ON orientation by 24 hours (Fig. 2.10C, bottom). The *flg* ON orientation was subsequently preserved along the leading edge at 48 and 72 hours (Fig. 2.10C, bottom). These results indicate that growth in motility medium introduced a selection barrier that restricted *flg* OFF bacteria to the inoculation site and allowed the expansion of a low frequency population of *flg* ON bacteria.

The orientation of the flagellar switch controls production of the glucosylating toxins

The flagellar alternative sigma factor, σ^D, controls transcription of genes within the Pathogenicity Locus (PaLoc) in addition to late stage flagellar genes in *C. difficile* (22, 23, 44). Therefore, we predicted that the flagellar switch, which impacts the expression of *sigD* (Fig. 2.7B), also indirectly controls the expression of the toxin genes *tcdA* and *tcdB* by activating the expression of the toxin sigma factor gene *tcdR*. We compared the expression of *tcdA*, *tcdB* and *tcdR* in *flgB* ON and OFF isolates using qRT-PCR. The abundances of the *tcdR*, *tcdA* and *tcdB* transcripts were significantly reduced in *flg* OFF bacteria compared to *flg* ON (Fig. 2.11A). Consistent with these results, TcdA protein level was also decreased in cell lysates of *flg* OFF bacteria compared to lysates of *flg* ON bacteria (Fig. 2.11B). TcdB levels were below the limit of detection by western blot.

TcdA and TcdB are exported protein toxins that inactivate host cell Rho and Rac GTPases through glucosylation, resulting in actin depolymerization and reduced host cell viability (13-16). To determine the effect of the flagellar switch on *C. difficile* cytotoxicity, filter-sterilized supernatants from *flg* ON and OFF isolates were tested in a cell viability assay with Vero cells. *Flg* ON isolates were significantly more cytotoxic compared to *flg* OFF isolates and a *sigD* control (Fig. 2.11C). Taken together, these data demonstrate that the flagellar switch orientation controls toxin production in addition to flagellum biosynthesis and swimming motility, indicating that these major virulence factors are coordinately phase variable in *C. difficile*.

The flagellar switch mediates regulation post-transcription initiation

Multiple mechanisms driving phase variable gene expression have been described: general homologous recombination, conservative site-specific recombination, short sequence repeat and slip strand mispairing, and DNA methylation (66, 74, 75). Classically, an invertible DNA element that undergoes site-specific recombination contains a promoter as the regulatory feature, and the orientation of the invertible DNA element (and promoter) determines whether the downstream genes are expressed. For example, in *E. coli* the orientation of the *fimS* switch, which contains a promoter, controls type I fimbrial gene expression (76). In contrast, in *C. difficile*, the genetic switch that regulates expression of *cwpV* operates after transcription initiation (68). In *cwpV* phase ON bacteria, transcriptional read-through of the 5' UTR and the *cwpV* coding sequence occurs. In *cwpV* phase OFF bacteria, the inversion of a DNA sequence in the 5' UTR leads to formation of a Rho-independent transcriptional terminator and results in premature transcription termination preventing expression of *cwpV*.

To begin to define the mechanism by which the flagellar switch controls gene expression, we generated a series of alkaline phosphatase (AP) transcriptional reporters (Fig. 2.12A) (77). First, the *phoZ* reporter gene was placed downstream of *flgB* (first gene of the *flgB* operon), under the control of the native *flgB* promoter and the full 498 bp 5' UTR with the flagellar switch (FS) in either the phase ON or OFF orientation (P*flgB*-5'UTR(FS$^{ON/OFF}$)-*flgB*::*phoZ*, Fig. 2.12A, #3 & 4). To determine if the flagellar switch contains a promoter, a truncated 307 bp 5' UTR retaining the flagellar switch (either phase ON or OFF orientation) but lacking the Cd1 riboswitch and native *flgB* promoter was placed upstream of *flgB::phoZ* (FS$^{ON/OFF}$-*flgB*::*phoZ*, Fig. 2.12A, #5 & 6). A transcriptional fusion of *phoZ* to the *flgB* promoter (P*flgB*::*phoZ*, Fig. 2.12A, #2) served as a positive control for the assay, and a promoterless construct (*::phoZ*, Fig. 2.12A, #1) was included as a negative control. All reporter fusions were integrated onto the chromosome of an R20291 *flg* ON isolate, and AP activity was assessed. In parallel with every AP assay, an asymmetric PCR-digestion assay was performed to ensure that the flagellar switches were in the expected orientation for both the native locus and the reporter fusion (Fig. 2.13). As anticipated, the promoterless control showed no AP activity, and the presence of the *flgB* promoter significantly increased activity (Fig. 2.12B). The P*flgB*-5'UTR(FS^{OFF})-*flgB*::*phoZ* reporter strain had significantly reduced activity compared to the P*flgB*-5'UTR(FS^{ON})-*flgB*::*phoZ* reporter strain (Fig. 2.12B), consistent with previous results (Fig. 2.7). No AP activity was detected in strains bearing the FS$^{ON/OFF}$-*flgB*::*phoZ* fusions (Fig. 2.12B). Similar results were obtained using the same six reporter fusions integrated onto the chromosome of a R20291 with the native flagellar switch in the OFF orientation, although AP activity was lower overall (Fig. 2.14). These results suggest that the flagellar switch does not contain a promoter to initiate transcription independently of the *flgB* promoter.

We considered alternative mechanisms by which expression could be inhibited in bacteria with the switch in the OFF orientation: (1) the formation of a Rho-independent transcriptional terminator, (2) the requirement for a *trans*-acting element such as an RNA binding protein or small non-coding RNA that selectively inhibits transcription, or (3) a Rho-dependent terminator. To examine these possibilities, we evaluated the activity of the *phoZ* reporters in a heterologous bacterium, *B. subtilis*, in which *C. difficile*-specific factors will be absent, but features inherent to the flagellar switch will be preserved. The reporter fusions (Fig. 2.12A) were integrated into the *B. subtilis* BS49 chromosome, and the orientation of the flagellar switch was monitored by orientation-specific PCR (Fig. 2.13). As in *C. difficile*, the positive control strain with the P$_{flgB}$*::phoZ* fusion showed high AP activity, but no activity was observed for the *::phoZ* negative control (Fig. 2.12B). Also consistent with *C. difficile*, *B. subtilis* with the FS$^{ON/OFF}$-*flgB*::*phoZ* fusions produced negligible AP activity, indicating the absence of a promoter in the flagellar switch (Fig. 2.12C). In contrast to the *C. difficile* results, we observed comparable AP activity in *B. subtilis* with P$_{flgB}$-5'UTR(FS$^{ON/OFF}$)-*flgB*::*phoZ* reporters (Fig. 2.12C).Thus, reduced gene expression as a result of the flagellar switch in the OFF orientation is specific to *C. difficile*, indicating that the regulatory feature is not inherent to the sequence, instead supporting the role of a *trans*-acting factor that is specific to *C. difficile*.

When considered along with the qRT-PCR data from Fig. 2.7B and 2.11A, these results suggest regulation post-transcription initiation from the *flgB* operon promoter and implicates changes in RNA stability or premature termination. Therefore, we used northern blot analysis to determine if transcription terminates prematurely or if the transcript is destabilized when the flagellar switch is in the OFF orientation. We were unable to detect the full-length transcript for the *flgB* operon (~ 23 kb) using a probe specific to the Cd1 region of the 5' UTR, so we used

strains with the P_{flgB}-5'UTR(FS$^{ON/OFF}$)-$flgB$::$phoZ$ reporters in the R20291 flg ON background to evaluate premature transcription termination. We observed an RNA corresponding to a full-length transcript of ~2400 nt for the P_{flgB}-5'UTR(FS^{ON})-$flgB$::$phoZ$ reporter but reduced transcript for the P_{flgB}-5'UTR(FS^{OFF})-$flgB$::$phoZ$ reporter (Fig. 2.12D). Furthermore, we did not detect a smaller transcript species other than termination through the Cd1 riboswitch, which eliminates the possibility of Rho-dependent or -independent termination unless the transcript is degraded faster than we can detect by northern blot. Collectively, the reporter assays in *C. difficile flg* ON and OFF and in *B. subtilis* with northern blot analysis suggest that regulation via the flagellar switch occurs post-transcription initiation and involves an unidentified *trans*-acting element that destabilizes the mRNA to reduce gene expression in *flg* OFF bacteria.

RecV, a tyrosine recombinase, catalyzes recombination at the flagellar switch in both orientations

Serine and tyrosine recombinases catalyze site-specific recombination to mediate DNA inversion at phase variable genetic switches in a RecA-independent manner (78, 79). In *E. coli*, two recombinases catalyze inversion at the fimbrial switch; FimB can catalyze both orientations, whereas FimE is restricted to ON to OFF inversion events (80, 81). A conserved tyrosine recombinase called RecV catalyzes strand exchange in both orientations at the *cwpV* switch in *C. difficile* (68, 82-84). We postulated that the recombinase(s), which catalyze inversion at the *flg* switch, would be present in all published *C. difficile* genomes with intact flagellum biosynthesis genes. We identified eight conserved serine or tyrosine recombinases, including RecV (CDR20291_1004), and used a two-plasmid system in a heterologous bacterium to identify the flagellar switch recombinase(s) (68). One plasmid contains one of the eight recombinase genes

cloned downstream of an anhydrotetracycline (ATc) inducible promoter (85), and the second

plasmid contains one of the P$_{flgB}$-5'UTR(FS$^{ON/OFF}$)-*flgB*::*phoZ* reporters (Fig. 2.12A, #3 & 4).

The plasmids were transformed into *E. coli*, induced with ATc, and orientation-specific PCR was

used on purified plasmids to determine if inversion occurred. We found that RecV was sufficient

to catalyze recombination of the flagellar switch from the ON to OFF (Fig. 2.15A) and OFF to

ON (Fig. 2.15B) orientations in *E. coli*; the other seven were unable to promote recombination.

These data indicate that the RecV recombinase is sufficient to catalyze inversion in both

orientations at both the flagellar and *cwpV* switches in a heterologous organism. These data do

not rule out the possibility of an additional recombinase that requires a *trans*-acting element,

such as a recombination directionality factor (86), to catalyze inversion at the flagellar switch.

     To determine whether RecV is involved in inversion of the flagellar switch in *C. difficile*,

we first overexpressed *recV* in the backgrounds of R20291 *flg* ON and OFF using the

anhydrotetracycline (ATc)-inducible expression vector to determine whether RecV would

promote inversion to the opposite orientation (87). We found that overexpression of *recV* in the

flg ON and OFF backgrounds leads to a mixed population of both orientations compared to

empty vector, regardless of the presence of ATc (Fig. 2.16, panel A). Transcription of *recV* in the

ATc-inducible expression vector was leaky in the absence of inducer (Fig. 2.16, panel B), which

is in line with previous work (87, 88).

RecV mutants are phase-locked for flagellum and toxin production

     The *C. difficile* R20291 *recV* mutants were previously shown to be phase-locked for

CwpV production (83, 84). The identification of RecV as the recombinase mediating flagellar

switch inversion suggests that mutation of *recV* would similarly result in phase-locked

phenotypes with respect to flagellum and toxin production. To evaluate the requirement of RecV for flagellar phase variation in *C. difficile*, we obtained two *C. difficile* R20291 mutants with ClosTron insertions in *recV* (kind gift from Dr. Louis-Charles Fortier). One mutant contains the *cwpV* switch locked in the ON orientation, and the other, the OFF orientation (82, 83). Both mutants contain the flagellar switch in the OFF orientation (herein, "*flg* OFF*"), as determined using orientation-specific PCR (Fig. 2.1B, Fig. 2.15C).

To obtain a *recV* mutant in which the flagellar switch is locked in the ON orientation, *recV* was expressed from a plasmid to allow inversion of the flagellar switch (Fig. 2.16C). This complemented strain was then passaged in the absence of antibiotic selection to allow loss of the plasmid. Thiamphenicol-sensitive clones were screened for flagellar switch orientation to identify *recV* mutants that are *flg* ON (herein, "*flg* ON*") (Fig. 2.17).

The *recV flg* ON* and OFF* mutants, each bearing vector or *recV* under the control of the P_{tet} ATc-inducible promoter, were assayed for swimming motility in BHIS-0.3% agar. A non-motile *sigD* mutant and enriched *flg* ON and OFF isolates were included as controls. Unlike the enriched *flg* OFF isolate that appeared motile due to a small frequency of *flg* ON bacteria, the *recV flg* OFF* mutant was non-motile at 24 hours incubation (Fig. 2.18A,B). This phase-locked phenotype was dependent on RecV, as complementation with *recV in trans* restored motility. In contrast, the *recV flg* ON* mutant showed motility, and providing *recV in trans* had no measurable effect in this assay. We note that extended incubation (greater than 24 hours) of the *recV flg* OFF* mutant results in motile progeny (Fig. 2.18). These motile bacteria retain the flagellar switch in the OFF orientation based on orientation-specific PCR and DNA sequencing results (Fig. 2.19), so the motile phenotype is due to suppressor mutations.

We next evaluated toxin production by the *recV flg* ON* and OFF* mutants by measuring TcdA production by western blot. As shown above, the enriched *flg* OFF isolate produced less TcdA than the enriched *flg* ON isolate (Fig. 2.18C, lanes 1,2). TcdA levels were greater in the *recV flg* ON* mutant compared to the enriched *flg* ON isolate. The difference is likely attributable to the presence of ~10.5% or 23.2% of bacteria with the flagellar switch in the OFF orientation remaining in the enriched *flg* ON isolate based on qPCR and fluorescence microscopy estimates, respectively (Fig. 2.4C, 2.7D). TcdA abundance in the *recV flg* OFF* mutant was comparable to that in the *sigD* mutant, which was below the limit of detection. Provision of *recV in trans* in either the *recV flg* ON* or OFF* mutant, which results in heterogeneity in the orientation of the flagellar switch (Fig. 2.15D), resulted in intermediate TcdA levels (Fig. 2.18C, lanes 5,7). The observed differences in TcdA levels among the strains appear to depend on the proportion of *flg* ON and OFF bacteria in the population. These results indicate that RecV mediates flagellum and toxin phase variation, in addition to CwpV phase variation, in *C. difficile*.

Discussion

Colonization is a prerequisite step to diarrheal disease development in *C. difficile* infection. The identification of colonization factors and the mechanisms controlling their production is an active area of study. Transcriptional regulators (51-53), a peptide-based quorum sensing system (29), and an RNA chaperone (54) activate or repress flagellar gene expression, albeit the exact regulatory mechanisms remain undetermined. Prior to this study, Cd1, a class I c-di-GMP sensing riboswitch, was the only identified *cis*-acting regulatory element to control

flagellar gene expression (59, 61, 62). Here, we report an additional *cis*-acting regulatory element, a "flagellar switch," present upstream of the early stage flagellar operon.

The flagellar switch is the fourth identified *C. difficile* inversion site (Cdi4) and consists of a 154 bp sequence flanked by 21 bp imperfect inverted repeats (69). The flagellar switch is functional, undergoing site-specific recombination in at least two different *C. difficile* ribotypes, 027 and 017. The flagellar switch inverted repeats are conserved in all sequenced *C. difficile* genomes that contain the flagellum biosynthesis genes, except the 630 genome has shortened 20 bp inverted repeats (Fig. 2.20). The flagellar switch appears locked ON in *C. difficile* 630, which may be due to the shortened inverted repeats. Alternatively, the 630 genome (012 ribotype) may differ in RecV transcription, production, or activity compared to the 027 and 017 ribotype strains, although there are no obvious differences in the *recV* promoter or open reading frame sequences to indicate that this is the case. Genetic evidence suggests that reduced inverted repeat length can impair recombination at the *cwpV* switch in *C. difficile* (68). The *cwpV* switch (Cdi1) controls the production of a large cell wall protein that promotes bacterial aggregation *in vitro* and is postulated to promote intestinal colonization (68, 82). CwpV may also contribute to *C. difficile* survival in the host by conferring resistance to predation by bacteriophages (84). Deposition of CwpV on the surface of *C. difficile* reduces phage adsorption and specifically prevents the injection of phage DNA into the bacterial cell (84). The other two identified inversion sites in *C. difficile*, although not yet demonstrated to be functional switches, are under active investigation.

Unlike the *cwpV* switch that controls the expression of a single structural gene, the flagellar switch controls the expression of both structural and regulatory genes, including *sigD*. Flagellar phase ON bacteria, as well as a *recV* mutant with the flagellar switch locked in the ON

orientation, display peritrichous flagella, engage in swimming motility, and produce the glucosylating toxins. Flagellar phase OFF bacteria and *recV* mutants with the flagellar switch locked in the OFF orientation are attenuated for transcription of flagellar genes, and therefore grow as aflagellate, non-motile bacteria also attenuated for toxin production. Thus, the flagellar switch coordinately controls the production of a colonization factor and essential virulence determinants by impacting *sigD* expression. Moreover, the flagellar switch orientation is expected to affect additional genes in the σ^D regulon, such as metabolic pathways, multiple cell wall proteins, metabolic transporters for amino acids and divalent cations, and several transcriptional regulators (23).

Classically, genetic switches that toggle between two orientations and facilitate phase variable expression of downstream genes contain a promoter to alter transcription initiation (66). Transcriptional reporter data excluded the presence of a promoter within the flagellar switch, indicating that the switch controls downstream gene expression after transcription initiation from the *flgB* operon promoter has occurred. Indeed, the switch is within the previously defined 5' UTR of the *flgB* (F3) operon, downstream of the c-di-GMP riboswitch (62). This arrangement of regulatory elements suggests that flagellar phase variation is only relevant when c-di-GMP levels are sufficiently low to permit transcriptional read-through beyond the riboswitch and into the flagellar switch. Regulation post-transcription initiation by genetic switches has been described previously. The *cwpV* switch in *C. difficile* also lacks promoter activity and instead regulates *cwpV* expression post-transcriptionally (68). When the switch is in the *cwpV* phase OFF orientation, the mRNA adopts a structure containing a Rho-independent terminator resulting in premature transcription termination (68). In *E. coli*, the *fimE* mRNA is subject to regulation post transcription initiation. FimE is encoded immediately 5' of *fimS*, and when *fimS* is in the OFF

orientation, the *fimE* mRNA forms a Rho-dependent terminator that reduces transcript stability (89). If the flagellar switch in *C. difficile* contained a Rho-independent terminator, we would expect no AP activity in *B. subtilis* when the flagellar switch was in the OFF orientation. However, AP activity was indistinguishable between the *flg* ON and OFF reporters in *B. subtilis,* suggesting a *trans*-acting regulatory element that is expressed only in *C. difficile*. We note that AP activity from the reporters was substantially higher in *B. subtilis* compared to *C. difficile*; it is possible that a regulatory factor in *B. subtilis* masked the unidentified *trans*-acting regulatory factor. Moreover, northern blot analysis suggests a mechanism independent of both Rho-dependent and independent termination since we were unable to detect a truncated transcript from the phase OFF condition. Collectively, our results suggest that the early stage flagellar operon mRNA is destabilized or degraded when the flagellar switch DNA is in the OFF orientation. We speculate that a constitutively-expressed *C. difficile*-specific regulatory factor, either an RNA binding protein or small, non-coding RNA, binds to the leader sequence of the mRNA to directly or indirectly degrade or destabilize the mRNA. Thus, the flagellar switch in *C. difficile* represents a novel system that requires an unknown additional regulatory factor to function as an ON/OFF switch.

In many mucosal pathogens, phase variation is a regulatory feature that modulates the expression of an immunostimulatory cell surface structure (66). Flagellar phase variation occurs in *Clostridium chauvoei* (90), *Salmonella enterica* serovar *enterica* Typhimurium (91, 92), *Helicobacter pylori* (93), and *Campylobacter jejuni* (94). Wild strains of *B. subtilis* grown under planktonic conditions bifurcate into two distinct populations: flagellated, motile single bacteria and aflagellate, non-motile chains of bacteria (95). Expression and activity of σ^D determines whether these populations arise in *B. subtilis*. Expression of *swrA*, which encodes a master

regulator of swarming motility, is subject to phase variation by slip strand mispairing in the coding sequence (95-97). SwrA controls transcription of the flagellar and chemotaxis (*fla/che*) operon by regulating the activity of the transcriptional regulator DegU (98). SigD, encoded within the *fla/che* operon, also positively regulates *swrA* transcription (98). Thus, a feedforward loop resulting from *swrA* expression introduces a bias for σ^D ON bacteria that are flagellated and motile. These and other data present a model in which flagellar gene expression in *B. subtilis* is bistable (97). Although additional work is needed to determine if *C. difficile* flagellar gene expression meets the criteria of bistability (99-101), inversion of the *C. difficile* flagellar switch similarly results in phenotypic heterogeneity that would confer an advantage in an environment that selects for one phenotype over the other (99-101). Thus, *B. subtilis* and *C. difficile* may use distinct mechanisms to accomplish the same outcome.

We identified RecV as necessary and sufficient for inversion of the flagellar switch in both orientations, by over-expression of *recV* in *C. difficile* and in *E. coli* bearing the flagellar switch sequence, and by mutation of *recV* in *C. difficile* R20291. Notably, RecV was previously reported to control inversion of the *cwpV* switch in *C. difficile* (68, 82). The use of a single recombinase to control inversion of multiple genetic switches has been described previously. In *Bacteroides fragilis*, the Mpi recombinase controls inversion of at least 13 genetic switches, including seven switches that control expression of different capsular polysaccharide loci (102). These switches are scattered through the genome and contain inverted repeats that harbor a 10 bp core consensus sequence (102). In contrast, the inverted repeats and the flanking half sites for the *cwpV* and flagellar switches lack recognizable sequence identity, although the identification of additional RecV-controlled genetic switches might reveal a core consensus sequence. RecV may bind the recognition sequences with different affinities, allowing hierarchical inversion of

switches bearing disparate inverted repeats. Regulation of *recV* transcription and/or RecV activity could therefore differentially affect inversion of the *cwpV* and flagellar switches. Hierarchical binding could be achieved solely by RecV, or via additional proteins that direct RecV to specific target sequences. The need for these binding partners may have been bypassed in *E. coli* due to over-expression of *recV*. It is also possible that *trans*-acting DNA binding proteins, such as a histone-like protein or recombination directionality factor (RDF) (86, 103), influence DNA bending to affect the recombination reaction at RecV-controlled genetic switches. Such DNA binding proteins can integrate environmental signals to bias switch orientation and generate genetic variants that have a fitness advantage in that environment. For example, leucine-responsive regulatory protein and H-NS affect the recombination reaction at *fimS* in *E. coli* (104, 105). The existence of factors influencing RecV activity in *C. difficile* is supported by the observation that *C. difficile* R20291 carrying the prophage ϕCD38-2 has the *cwpV* switch biased to the ON orientation; *recV* transcription is unaltered (83). Prophage infection of a *recV* mutant with the *cwpV* switch in the OFF orientation remains locked, suggesting that the phage does not encode a recombinase to promote inversion, but that instead a phage gene product modulates RecV activity (83). ϕCD38-2 did not affect flagellum or toxin gene transcription in R20291, but *C. difficile* may produce another factor that similarly influences recombination within the flagellar (and *cwpV*) switch. Recent evidence demonstrates a role for RDFs in *C. difficile* physiology: an RDF pairs with a serine recombinase to excise a prophage-like element from a sporulation-specific sigma factor gene (106). Future studies will explore how RecV can control multiple genetic switches with divergent inverted repeat sequences.

Ultimately, iterative combinations of *cwpV* and *flg* ON and OFF phenotypes may influence *C. difficile* fitness in the host intestinal environment. A dual *cwpV* and *flg* ON phenotype could facilitate penetration of the bacteriophage-rich colonic mucus (107), and/or increase bacterial attachment to and colonization of the intestinal mucosa. Conversion of one or both phenotypes to phase OFF might reduce the probability of the host simultaneously developing an antibody response to these immunostimulatory surface structures. The role of flagellum and toxin phase variation in *C. difficile* pathogenesis remains to be determined, though it stands to reason that both the phase ON and OFF phenotypes confer advantages during the course of an infection. The *flg* ON bacteria would be competent for efficient intestinal colonization for R20291 (31). Yet purified flagellin from multiple *C. difficile* strains can activate host Toll-like receptor 5 (TLR5) and stimulate p38 MAPK activation and IL-8 secretion *in vitro* (47, 108, 109). Although TLR5 is dispensable in a mouse model of CDI (49), recombinant *C. difficile* FliC is immunogenic and protective in a mouse model of CDI (110). In addition, the *C. difficile* glucosylating toxins have been implicated in activation of the NLRP3 and Pyrin inflammasomes, which could promote pathogen clearance (48, 111). *C. difficile flg* OFF bacteria could thus evade TLR5 recognition and inflammasome stimulation, enhancing bacterial colonization and persistence within the host. Future studies will determine the contribution of flagellum and toxin phase variation to *C. difficile* virulence.

Materials and Methods

Growth and Maintenance of Bacterial Strains

The strains and plasmids used in this study are listed in Table 2.1, and details on their construction are described later in the methods. *C. difficile* was maintained in an anaerobic

chamber (Coy Laboratories) at an atmosphere of 90% N_2, 5% CO_2, and 5% H_2 and grown in

Brain Heart Infusion medium (Becton Dickinson) supplemented with 5% yeast extract (Becton

Dickinson) (BHIS) at 37°C. Bacteria were also cultured in Tryptone Yeast (TY) broth media

where indicated. All *C. difficile* broth cultures were grown statically. *Escherichia coli* DH5α,

BL21, and HB101(pRK24) were cultured at 37°C in Luria-Bertani (LB) medium with the

indicated antibiotics for plasmid selection. *Bacillus subtilis* strain BS49 was grown in BHIS with

the appropriate antibiotics. Antibiotics were used at the following concentrations:

chloramphenicol, 10 μg/ml; thiamphenicol, 10 μg/ml; kanamycin, 100 μg/ml; ampicillin, 100

μg/ml; erythromycin, 5.0 μg/ml, lincomycin 20.0 μg/ml.

Orientation-Specific PCR Assay

Spores of *C. difficile* strains 630*Δerm*, R20291, and ATCC 43598 were plated on BHIS

supplemented with 0.1% sodium taurocholate (Sigma) (BHIS+TA). After 24 hours, individual

colonies were grown in BHIS at 37°C. Overnight cultures were diluted 1:50 in fresh BHIS and

grown to early stationary phase. An aliquot of each culture was diluted 1:5 into 10 mM Tris-HCl,

pH 7.5, 1 mM EDTA (TE) buffer. Boiled lysates served as templates in PCR using primers

designed for each strain based the genome sequences for strains 630 (Accession No.

AM180355.1) and R20291 (Accession No. FN545816.1). The ribotype 017 strain ATCC 43598

has not been sequenced, so we used M68 (Accession No. FN688375.1) as a representative

ribotype 017 strain to design primers for the flagellar switch. Primer sequences are listed in

Table 2.2. Primers R1614 (published/ON orientation) and R1615 (inverse/OFF orientation) were

used for R20291. Primers R1622 (published/ON orientation) and R1623 (inverse/OFF

orientation) were used for ATCC 43598. Primers R1751 (published/ON orientation) and R1752

(inverse/OFF orientation) were used for 630Δerm. All reactions used R857 as the reverse primer given the sequence identity between strains for the first gene in the early flagellar operon, *flgB*. The results shown are representative of three independent experiments each with at least two biological replicates for each strain.

Asymmetric PCR-Digestion Assay

To amplify the flagellar switch, chromosomal DNA purified from *C. difficile* R20291, as previously described (112), served as the template in PCRs with primers R591 and R857, yielding a product of 665bp. The PCR products (500 ng) were digested with SwaI (NEB), which results in different products depending on the orientation of the flagellar switch: 312 bp and 353 bp for the published orientation, and 418 bp and 247 bp for the inverse orientation (Fig. 2.2). A mixture of bacteria with the flagellar switch in the published and inverse orientation results in four bands at varying intensities. SwaI reaction products were separated in 2.5% agarose gels, which were stained with ethidium bromide (EtBr) for imaging with a G:BOX Chemi Imaging system. For all experiments with purified phase variant populations, three or four biological replicates of R20291 *flg* ON and OFF were assessed by the asymmetric PCR-digest assay to ensure a homogeneous population. We used ImageJ to perform a densitometry analysis of the digested bands to determine the relative proportions of bacteria with flagellar switch in the published/ON (312, 353 bp) and inverse/OFF (418, 247 bp) orientations. The pixel intensities of published/ON or inverse/OFF bands were divided by the total pixel intensity for that sample. These values were normalized to a standard curve generated by mixing known quantities of *flg* ON and OFF plasmid template and subjecting them to the asymmetric PCR-digestion assay, giving the percentage of bacteria in a sample with the switch in a given orientation.

To evaluate the effect of surface growth on the flagellar switch, *C. difficile* R20291 spores were plated on BHIS-TA to induce germination. After 24 hours, individual colonies were suspended and spotted onto four BHIS plates. Every 24 hours for four days, one plate was used for recovery of bacteria from colony for chromosomal DNA extraction. Once chromosomal DNA was collected, the orientation of the flagellar switch was determined using the asymmetric PCR-digestion assay. Densitometry analysis was performed using ImageJ software as described above. Values for bands corresponding to the published/ON and inverse/OFF orientations were compared to a standard curve. The data were combined from three independent experiments, each with two to four biological replicates.

Purification of Flagellar Phase Variant Populations

C. difficile R20291 spores were plated on BHIS+TA. After 24 hours, six to ten individual colonies were suspended in BHIS, spotted onto individual BHIS plates, and grown 24-96 hours as indicated. Single colonies were derived from this growth (24-96 hours) by passaging onto BHIS plates and individually screened for the flagellar switch orientation by the asymmetric PCR digestion assay. A higher frequency of colonies with the flagellar switch in the OFF orientation was observed at later time points (96 hours). For all experiments using enriched flagellar phase variant populations, except the *phoZ* reporters, *mCherryOpt* reporters, and cell viability assay, glycerol stocks were *not* made to ensure robust reproducibility of the phenotypes between independent experiments and reproducibility of the enrichment protocol. At least two to four biological replicates of *flg* phase ON and OFF populations were used for all experiments.

Quantitative PCR Analysis of the Flagellar Switch Orientation

In all reactions, SYBR Green Real-Time qPCR reagents (Thermo Fisher) were used, with primers at a final concentration of 500 nM and an annealing temperature of 55°C. Titrated amounts of genomic DNA from *recV flg* ON* and *flg* OFF* were used to determine the amount of template necessary for the reactions and the primer efficiency. Primers for detection of the ON orientation (R2175 and R2177) had a PCR efficiency of 88.5% for the *recV flg* ON* DNA, and the primers for detection of the OFF orientation (R2176 and R2177) had a PCR efficiency of 84.5% for the *recV flg* OFF* DNA. To determine the frequency of the *flg* ON and OFF orientations in populations of enriched *C. difficile* R20291 *flg* ON and OFF isolates, three biological replicates of each were grown in BHIS medium to an OD_{600} of ~1.0, and chromosomal DNA was extracted as previously described (112). Quantitative PCR was done using 4 ng of DNA from *flg* ON and OFF isolates. DNA copy number was calculated using the *ΔΔCt* method, with the *rpoC* gene as the indicated reference gene for DNA copy number.

RNA Extraction, cDNA Synthesis, and Quantitative Reverse Transcriptase PCR

Isolated *C. difficile flg* ON and OFF variants were grown overnight in BHIS medium, 1:50 in 3mL of BHIS medium and grown to stationary phase (OD_{600} of 1.8 - 2.0). RNA was isolated as described previously (59, 112). Briefly, cells were collected by centrifugation and lysed by bead beating in cold TriSURE (Bioline). Nucleic acid was extracted with chloroform (Sigma), precipitated from the aqueous phase with isopropanol, washed with ethanol, and suspended in RNase-free water. To remove contaminating genomic DNA, RNA was treated with TURBO DNase (Thermo Fisher) according to the manufacturer's protocol. Synthesis of cDNA was done using a Tetro cDNA Synthesis kit (Bioline) and random hexamers according to the

manufacturer's instructions and including a no-reverse transcriptase control. Real-time PCRs were done using 2 ng of cDNA and SYBR Green Real-Time qPCR reagents (Thermo Fisher). Transcript abundance was calculated using the ΔΔCt method, with *rpoC* as the control gene and the indicated reference condition/strain (59).

Swimming Motility Assay

C. difficile R20291, *flg* ON and OFF isolates, and *recV* mutants were assayed for flagellar motility as previously described (59). An R20291 *sigD* mutant was included as a non-motile control in all experiments. Details regarding the generation of the *sigD* mutant are in S1 Methods and S5 Fig. Autoclaved 0.5 X BHIS with 0.3% agar (30 mL) was poured into 100 mm Petri dishes allowed to solidify overnight. These soft agar plates were kept in the anaerobic chamber for at least 4 hours prior to the experiment. Bacteria were grown in BHIS medium overnight, diluted 1:50 in fresh BHIS broth the next day, and grown to an OD_{600} of 1.0. Two microliters of *flg* ON, *flg* OFF, and *sigD* were inoculated into the agar, then incubated at 37°C. The diameter of growth was measured after 24, 48, and 72 hours; two perpendicular measurements were made for each swim site and averaged for each replicate. Three to four biological replicates were evaluated, each in technical duplicate (values averaged), in two independent experiments. Images were taken using the G:BOX Chemi imaging system with the Upper White Light illuminator. A Student's t-test was used to determine statistical significance.

Alkaline Phosphatase Assay

C. difficile phoZ reporter strains were grown from glycerol stocks on BHIS plates and incubated at 37°C. After 24 hours, 2-3 colonies of each reporter strains were grown overnight in

TY medium and diluted 1:50 into BHIS medium. Late exponential phase cells ($OD_{600} \approx 1.5$, 1.5 mL) were collected by centrifugation, the supernatant was discarded, and pellets were stored at -20°C overnight. Frozen pellets were thawed on ice and the alkaline phosphatase (AP) assay was performed as previously described (77). *Bacillus subtilis* BS49 *phoZ* reporter strains were grown from glycerol stocks on BHIS-Erm plates under aerobic conditions, and the assay was done as above. Construction of AP reporters is described in S1 Methods.

Detection of RNA by Northern Blot

    C. difficile phoZ reporter strains were grown in 4mL of BHIS medium to OD_{600} 1.5. Bacteria were collected by centrifugation, and RNA was extracted as described above with the following exceptions. Four rounds of bead beating were done, and the purified RNA was rigorously treated with DNase I to ensure removal of contaminating DNA. Digoxigenin (DIG)-labeled DNA probes for the northern blot were generated by PCR using genomic DNA from R20291 as template for the Cd1 probe and a 5S rRNA (CDR20291_ r03) probe, a DIG High Prime kit was used (Roche) (Table 2.2). Reagents from the NorthernMax-Gly Kit (Thermo Fisher) were used for electrophoresis and hybridization. RNA samples (15ug for mRNA probe and 500ng for rRNA probe) were briefly thawed on ice, mixed 1:1 with Glyoxal Load Dye, incubated at 50°C for 30 minutes, and electrophoresed in a 1% agarose gel made with 1X Gel Prep/Running Buffer. The agarose gel was imaged using the G:BOX Chemi Imaging system to confirm RNA integrity. The gel was briefly soaked in 20X SSC (3M NaCl and 300mM sodium citrate), then RNA was transferred via capillary action onto a nylon membrane (Amersham Hybond-N+) overnight in 20X SSC for 16 hours. RNA was crosslinked to the membrane using a UV Stratalinker 1800 (Stratagene). After prehybridization with ULTRAhyb buffer at 42°C for 1

hr, 4 μL of a DIG-labeled DNA probe specific to the Cd1 riboswitch or the 5S rRNA loading

control gene was added. After 16 hours, the membrane washed in low and high stringency

buffers sequentially (NorthernMax-Gly Kit). To detect the DIG-labeled probes on the membrane,

we used the buffers and followed the manufacturer's instructions in the DIG High Prime DNA

Labeling and Detection Starter Kit II (Roche) and the chemiluminescent substrate CDP-Star

(Roche). Membranes were exposed to film, which was then imaged using a developer.

Detection of TcdA by Western Blot

    C. difficile was grown as patches on BHIS agar for 24 hours. The patches were

thoroughly suspended in BHIS, and cell densities were normalized to an OD_{600} $1.0 - 1.5$. The

cells were again pelleted by centrifugation and then suspended in 1X SDS-PAGE sample buffer

(113) and heated at 95°C for 10 minutes. The lysates were separated on an 8% SDS-

polyacrylamide gel then transferred to a nitrocellulose membrane (Bio-Rad). TcdA was detected

using a mouse α-TcdA primary antibody (Novus Biologicals) and goat anti-mouse IgG

conjugated with IR800 (Thermo Fisher) as described previously (44). Detection was done using

a LI-COR Odyssey Imager (LI-COR Biosciences). Three independent experiments were done,

each with three to four biological replicates of *flg* ON and OFF isolates.

Cell Viability Assay

    Vero cells were seeded in 96 well plates at 8,000 cells per well in DMEM (Life

Technologies) supplemented with 10% Fetal Bovine Serum (Sigma) (DMEM-FBS) and 1X

Penicillin/Streptomycin (Life Technologies) for 16 hours. *C. difficile* R20291 *flg* ON and OFF

isolates, and a *sigD* mutant control, were grown on BHIS plates for 16-24 hours. Four colonies

of each isolate/strain were grown in TY medium overnight, then diluted 1:50 into fresh TY and grown to stationary phase ($OD_{600} \approx 2.0$). Cultures were normalized to the same OD_{600}. Bacteria were removed by centrifugation, and supernatants were collected, filter sterilized (0.45 micron), and serially diluted in TY medium. For each well with Vero cells, the cell culture medium was replaced with fresh DMEM-FBS. Supernatants were added in the following dilutions with cell culture media: 1:4, 1:40, 1:80, 1:400, and 1:800. As a negative control, Vero cells were treated with TY medium at a 1:4 dilution in DMEM-FBS. After 24 hours at 37°C, the medium was carefully removed by aspiration, and 100 μL each of cell culture medium and Promega CellTiter Glo reagent were added to each well. After 10 minutes at room temperature on a shaker, Vero cell lysates were transferred to opaque-walled 96 well plates. Luminescence was measured using a Synergy H1 Hybrid plate reader with Gen5 software (BioTek). Data were combined from two independent experiments each with four biological replicates of each isolate/strain.

Identification of the Recombinase That Catalyzes Inversion of the Flagellar Switch

Relevant *E. coli* strains are listed in Table 2.1. Bacteria were passaged onto LB plates with ampicillin 50 μg/mL and kanamycin 50 μg/mL and grown overnight for single colonies. Bacteria were inoculated into LB with the above-mentioned antibiotics to grow overnight at 30°C, diluted 1:50 into fresh medium in duplicates, and grown to an OD_{600} 0.4. Anhydrotetracycline (ATc) was added to a biological replicate for each flagellar switch/recombinase pair at a final concentration of 200 ng/mL, grown until OD_{600} 1.0, and plasmids were purified using the GeneJET Plasmid Miniprep Kit (Thermo Fisher). Purified plasmids were used as template in an orientation-specific PCR assay to identify the recombinase that catalyzes inversion from ON to OFF and OFF to ON.

Microscopy

For transmission electron microscopy, *C. difficile flg* ON and OFF populations were

isolated on BHIS agar plates as described above and briefly washed in Dulbecco's PBS (Sigma)

prior to suspension in PBS-4 % paraformaldehyde for fixation for 1 hour in the anaerobic

chamber. Cell suspensions were adsorbed onto Formvar/copper grids, washed in water, and

stained for 30 seconds in two sequential drops of 1.5% aqueous uranyl acetate. Cells were

observed on a LEO EM 910 Transmission Electron Microscope (Carl Zeiss Microscopy, LLC,

Thornword, NY) and recorded with a Gatan Orius SC1000 digital camera with Digital

Micrograph 3.11.0 software (Gatan, Inc., Pleaston, CA). To image bacterial colony morphology,

we used a Zeiss Stereo Discovery V8 dissecting microscope with a glass stage for illumination of

bacterial colonies with light from the top and bottom.

For fluorescence microscopy, *C. difficile* enriched *flg* ON and OFF isolates and the *sigD*

mutant bearing *mCherryOpt* reporter fusions (Table 2.1) were cultured overnight in BHIS

medium, diluted 1:50 in fresh BHIS medium, and grown to an OD_{600} of ~0.5. For strains with the

ATc-inducible *mCherryOpt* reporter (pDSW1728), 15 ng/mL of ATc was added to the culture at

OD_{600} ~0.3 to induce transcription of the *mCherryOpt* gene. One milliliter of each culture was

briefly pelleted in the anaerobic chamber, washed and suspended in phosphate buffered saline

(PBS), and fixed according to published methods in PBS and paraformaldehyde (73). After 3

hours in fixative in the dark at 4°C, bacterial pellets were washed three times in PBS and

suspended in 0.5 mL PBS with 3 µg/mL 4′,6-diamidino-2-phenylindole (DAPI) to label DNA.

After overnight incubation at 4°C in the dark, samples were pelleted and suspended in 1 mL of

PBS. Bacteria were immobilized on agar pads as previously described (73) and covered with 1.5

thickness glass cover slips (ThermoFisher) for fluorescence microscopy. The Olympus BX61

Upright Wide Field Microscope with a 60X/1.42 Oil PlanApo N objective lens was used for

imaging samples, and Volocity 6.3 software was used for image acquisition. Multiple fields were

taken for each sample in a coordinated fashion to ensure no repeated sections. Three images were

automatically taken for each field in the differential contrast (DIC) channel, DAPI channel

(Excitation: 377 nm (+/- 25 nm), Emission: 447 nm (+/- 30nm)), and Texas Red/RFP channel

(Excitation: 562 nm (+/- 20 nm), Emission: 642 nm (+/- 20 nm)) with consistent settings for

side-by-side comparison. Images were processed using the FIJI version of ImageJ (114), and

bacteria were visually enumerated for +/- mCherryOpt fluorescence, with DAPI staining and

DIC channels allowing imaging of all bacteria.

Generation of standard curve using known quantities of *flg* ON and OFF DNA

The region spanning the promoter of the *flgB* operon, the 5' UTR, and the *flgB* coding

sequence was amplified by PCR from genomic DNA of *C. difficile* R20291 *flg* ON and OFF

isolates using primers R1512 and R1611. The resulting 1063 bp products, P$_{flgB}$-

5'UTR(FS$^{ON/OFF}$)-*flgB*, were digested with BamHI and EcoRI. The *phoZ* gene was PCR

amplified from pMC358 template using primers R1609 and R1610. The 1473 bp product was

digested with EcoRI and SphI. The digested P$_{flgB}$-5'UTR(FS$^{ON/OFF}$)-*flgB* and *phoZ* fragments

were ligated into BamHI- and SphI-digested pMC123 to generate *flg* ON (pRT1323) and OFF

(pRT1324) reporter plasmids for the standard curve. Preparations of these plasmids were mixed

in known ratios of OFF/ON plasmids (0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30,

80:20, 90:10, 100:0) for the asymmetric PCR-digestion assay. The mixtures were used as

templates in PCR reactions with R591 and R857 to amplify the flagellar switch region. The PCR

products were column purified, digested with the restriction enzyme SwaI, and electrophoresed

in a 2.5% agarose gel. Images of the EtBr-stained gels were collected under UV using the

G:BOX Chemi Imaging system, and band intensities in each lane were determined using Image

J, as previously described (71).

Construction of alkaline phosphatase reporters for use in *B. subtilis* and *C. difficile*

The *slpA* transcriptional terminator sequence was amplified from pRPF185 with R1848

and R1849, resulting in a 161 bp product that was digested with BamHI and HindIII and ligated

into similarly digested pSMB47 to generate pRT1532. Alkaline phosphatase reporter fusions

were subsequently cloned into this vector, 5' of the transcriptional terminator. The full length

reporters (P$_{flgB}$-5'UTR(FS$^{ON/OFF}$)-*flgB::phoZ*) were amplified from plasmids RT1323 and

RT1324 with primers R1512 and R1610. The resulting 2530 bp product was digested with SphI

and BamHI, and ligated into similarly digested pRT1532. The promoterless *phoZ* construct was

made by PCR amplifying *phoZ* from pMC358 with primers R1632 and R1610, digesting the

1479 bp product with SphI and BamHI, and ligating into similarly digested pRT1532. The

P$_{flgB}$*::phoZ* reporter was made by PCR amplifying the *flgB* operon promoter region from *C.*

difficile R20291 genomic DNA with primers R1512 and R1608 (250 bp product digested with

SphI and EcoRI) and *phoZ* from pMC358 with primers R1609 and R1610 (1476 bp product

digested with EcoRI and BamHI). Digested products for P$_{flgB}$ and *phoZ* were ligated into BamHI-

and SphI-digested pRT1532. To generate fusions of the flagellar switch only (FS$^{ON/OFF}$) to *phoZ*,

the region downstream of the Cd1 riboswitch was amplified by PCR with primers R1673 and

R1610 using pRT1323 and pRT1324 as templates for the *flg* ON and OFF orientations,

respectively. The 2109 bp PCR products were digested with SphI and BamHI and ligated into

87

similarly digested pRT1532. Purified plasmids for each reporter were transformed into *Bacillus subtilis* BS49, as previously described (71, 112), and are listed in Table 2.1. Transformants with a flagellar switch were screened by orientation-specific PCR with R1614/1615 and R1706 to ensure that the expected orientation of ON or OFF was present and maintained. BS49 strains were conjugated with *C. difficile* R20291, and transformants were selected for lincomycin and tetracycline resistance, as previously described (112). *C. difficile* transformants were screened for the Tn916 AP reporter by PCR and for the orientation of the native flagellar switch and the Tn916 AP reporter flagellar switch using the asymmetric PCR-digestion assay.

Cloning of the putative recombinase genes for expression in *E. coli* and *C. difficile*

The genes for the eight recombinases conserved in the *C. difficile* genomes are listed with the R20291 locus tag: CDR20291_1004 (*recV*), CDR20291_1060, CDR20291_1068, CDR20291_1174, CDR20291_1826, CDR20291_1855, CDR20291_1973, and CDR20291_3416. The genes were amplified from *C. difficile* R20291 genomic DNA with primers named according to the pattern "locus tag F" and "locus tag R" (Table 2.2). PCR products were digested with KpnI and HindIII, except PCR products for CDR20291_1826 and CDR20291_1973 were digested with KpnI and BamHI. All PCR products were ligated into similarly digested pMWO-074 and transformed into *E. coli* DH5α. Sequence integrity was confirmed by sequencing the insertion in each resulting plasmid. Plasmids were purified from DH5α and co-transformed with either P_{flgB}-5'UTR(FS^{ON})-*flgB::phoZ* (RT1323) or P_{flgB}-5'UTR(FS^{OFF})-*flgB::phoZ* (RT1324) into *E. coli* DH5α to generate strains with a plasmid encoding one conserved recombinase and with a plasmid containing the flagellar switch in the ON or OFF orientation (Table 2.1).

To express *recV* in *C. difficile*, the *C. difficile* R20291 *recV* gene was PCR amplified

from pRT1164 using primers R1853 and R1854. The 628 bp PCR product was digested with

SacI and BamHI, ligated into similarly digested pRPF185, and transformed into *E. coli* DH5α,

yielding strain pRT1529. The purified plasmid was transformed into *E. coli* HB101(pRK24) to

allow transfer by conjugation of the *recV* expression plasmid into *C. difficile* R20291 *flg* ON and

OFF isolates. Thiamphenicol and kanamycin resistant *C. difficile* transconjugants were selected

for evaluation (112). A control plasmid of the ATc-inducible expression vector was generated by

removing the *gusA* gene from pRPF185 by digestion with SacI and BamHI and religating the

vector backbone following treatment with Klenow. The vector was introduced into *C. difficile*

R20291 *flg* ON and OFF isolates via conjugation with *E. coli* HB101(pRK24), and

thiamphenicol and kanamycin resistance colonies were selected (112). The presence of plasmids

in *C. difficile* was confirmed by PCR with plasmid-specific primers flanking the multiple cloning

site.

Generation of the *C. difficile* R20291 *sigD* mutant

To generate the *sigD* mutant in R20291, we used the previously generated retargeting

plasmid pBL100::*sigD* (pRT1073), which inserts the Group II intron targeting sequence into

sigD after nucleotide position 228 (60). pRT1073 was transformed into *E. coli* HB101(pRK24)

to allow its transfer into *C. difficile* R20291 via conjugation. Thiamphenicol- and kanamycin-

resistant transconjugants were selected and passaged on BHIS-agar with lincomycin to identify

clones with an insertion of the Group II intron. Candidate mutants were screened by PCR with

primers for the *sigD* gene and for the Group II intron as described in Fig 2.9.

Construction of *mCherryOpt* reporter fusions and introduction into *C. difficile*

The pDSW1728 plasmid with the *mCherryOpt* gene under the transcriptional control of an anhydrotetracycline-inducible promoter (gifted by Dr. Craig Ellermeier) served as the base for reporter construction (72, 73). To generate pP$_{flgM}$*::mCherryOpt*, P$_{flgM}$ was PCR amplified from R20291 genomic DNA using primers R2117 and R2046 and digested with NheI and SacI to generate a 51 bp product. The NheI and SacI digested P$_{flgM}$ was ligated into similarly digested pRPF144. The *mCherryOpt* gene was digested from pDSW1728 with SacI and BamHI. The 739 bp product was ligated into similarly digested pRPF144 with P$_{flgM}$, yielding pRT1676. To generate the promoterless vector control plasmid (p*::mCherryOpt*), pP$_{flgM}$*::mCherryOpt* was digested with NheI and SacI to remove the P$_{flgM}$, and the vector backbone was treated with Klenow before being religated to yield pRT1685. The plasmids were confirmed by PCR and sequencing of the insertions, then transformed into *E. coli* HB101(pRK24) to allow conjugation with *C. difficile* R20291 enriched *flg* ON and *flg* OFF, and *sigD::ermB* as appropriate. Thiamphenicol- and kanamycin-resistant colonies were selected, and plasmids were detected by PCR with vector specific primers.

Generation of *C. difficile recV* mutant derivatives

The *C. difficile* R20291 *recV::ermB* (*recV*) *cwpV* ON* and OFF* (* locked) were obtained from Dr. Louis-Charles Fortier (84). Orientation-specific PCR (Fig 1A) was performed to determine the orientation of the flagellar switch using primers R1614 and R857 for the ON orientation (375 bp) and R1615 and R857 for the OFF orientation (281 bp). Both *recV* isolates were found to be *flg* OFF* (Fig 2.15). To obtain a *recV flg* ON* isolate, we took advantage of the facts that pRPF185*::recV* (pRT1529) has leaky transcription in the absence of inducer, and that

low levels of RecV are sufficient to promote a mixed population of *flg* ON and OFF bacteria (Fig 2.17). The *C. difficile* R20291 *recV cwpV* OFF* and *flg* OFF* isolate (RT1693) was conjugated with *E. coli* HB101(pRK24) transformed with pRT1529 to obtain thiamphenicol and kanamycin resistant transconjugant RT1697. RT1697 was passaged in TY broth medium, diluting every 24 hours for 4 days in the absence of thiamphenicol selection to allow plasmid loss and to obtain a mixture of genetic switch orientations. Broth cultures were sampled on the 4th day and cultured on plain BHIS agar to obtain single colonies. Eight single colonies were replica plated on BHIS +/- thiamphenicol plates. Colonies that were sensitive to thiamphenicol (indicating plasmid loss) were evaluated for the orientations of the flagellar and *cwpV* switches by orientation-specific PCR. The *cwpV* switch orientation was determined to identify a *cwpV* OFF* isolate, to prevent implication of *cwpV* in any phenotypes attributed to the flagellar switch. Primers R1920 and R1050 were used to detect the *cwpV* switch in the OFF orientation (493 bp product), while primers R1921 and R1050 were used to detect the *cwpV* ON orientation (346 bp product). Two *recV flg* ON*/*cwpV* OFF* clones were obtained, RT1702 and RT1703. Complementation of the *recV flg* ON* mutant was achieved by conjugation with *E. coli* HB101(pRK24) transformed with pRPF185::*recV* (pRT1529) and isolateion of thiamphenicol and kanamycin resistant transconjugants.

Construction of *C. difficile* strains for swimming motility assay and western blot

E. coli* (pRK24) transformed with pRT1611 (reporterless derivative of pRPF185) was conjugated with the following *C. difficile* R20291 *strains*: enriched *flg* ON, enriched *flg* OFF, *sigD*, *recV flg* ON* (*cwpV* OFF*), and *recV flg* OFF* (*cwpV* OFF*). Thiamphenicol and kanamycin resistant *C. difficile* transconjugants were selected for evaluation. These strains are

listed in the Table 2.1 and were used in the swimming motility assay and western blot for TcdA in Figure 2.18 with *recV* mutants complemented with *recV* on as plasmid.

Evaluation of the *flgB* operon regulatory region from *recV flg* OFF* motile suppressors

The region upstream of the *flgB* operon consisting of the promoter and 5'UTR was amplified by PCR using primers R1512 and R1611 to generate a 1066 bp product from the following *C. difficile* R20291 *recV* strains: *flg* ON*, *flg* OFF*, and the eight *flg* OFF* motile suppressor mutants. The 1066 bp product was cloned into the pCR4-TOPO vector (ThermoFisher), as described by the manufacturer's instructions, transformed into *E. coli* DH5α, and selected on agar medium for ampicillin resistance and loss of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) activity. White colonies were screened by colony lysis PCR for insertion of the 1066 bp into the vector using vector specific primers. Plasmid was purified from positive transformants and used as template in orientation-specific PCR to evaluate the orientation of the flagellar switch. PCR products were resolved in a 2.5% agarose gel and stained with EtBr for visualization. The region of interest was also sequenced using plasmid specific primers that flank the 1066 bp regulatory region and *flgB* gene. Sequencing results for the 1066 bp region from the eight *recV flg* OFF* suppressor mutants showed neither single nucleotide mutations nor inversion of the flagellar switch to the ON orientation compared to *recV flg* OFF*.
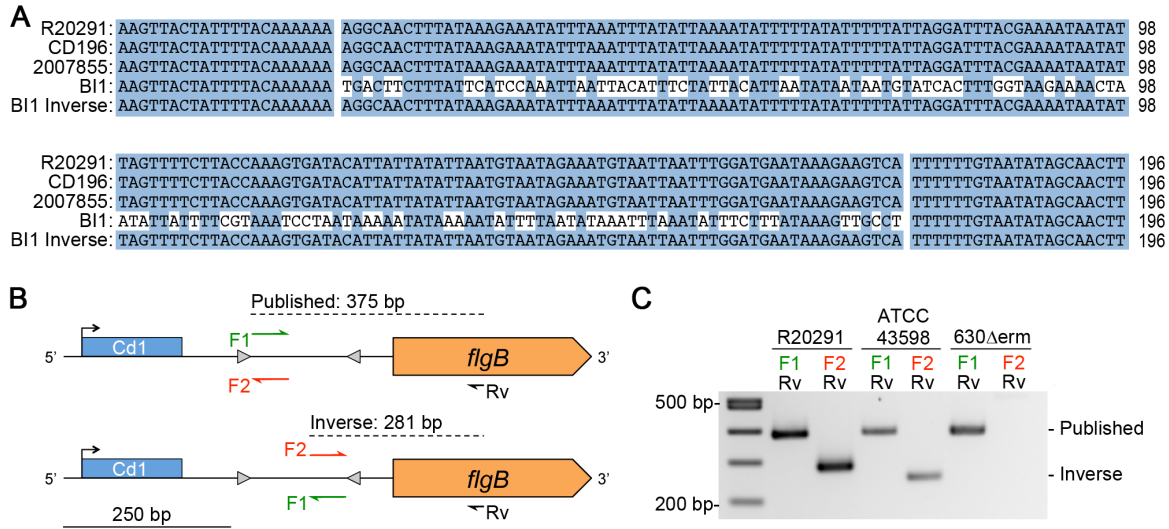
Figure 2.1. Evidence for DNA inversion at the flagellar switch. (A) Nucleotide sequences corresponding to the 5' UTR of the *flgB* operon from genome sequences available for PCR ribotype 027 strains were aligned using Clustal Omega. Shown are the regions corresponding to the putative flagellar switch and flanking imperfect inverted repeats. For strain BI1 "inverse", the alignment was repeated after replacing the putative switch with its reverse complement. Identical nucleotides are indicated with blue shading. (B) Diagram of the PCR strategy used to detect the putative flagellar switch orientation. The primer names and sequences used for each strain are listed in the Table 2.2. The predicted product sizes are based on R20291 sequence. (C) Orientation-specific PCR products for the flagellar switch from three *C. difficile* strains representing three ribotypes (R20291, 027, NCBI Accession No FN545816.1; ATCC43598, 017, NCBI sequence read archive SRX656590 (115); 630Δ*erm*, 012, NCBI Accession No. EMBL:LN614756 (116)).
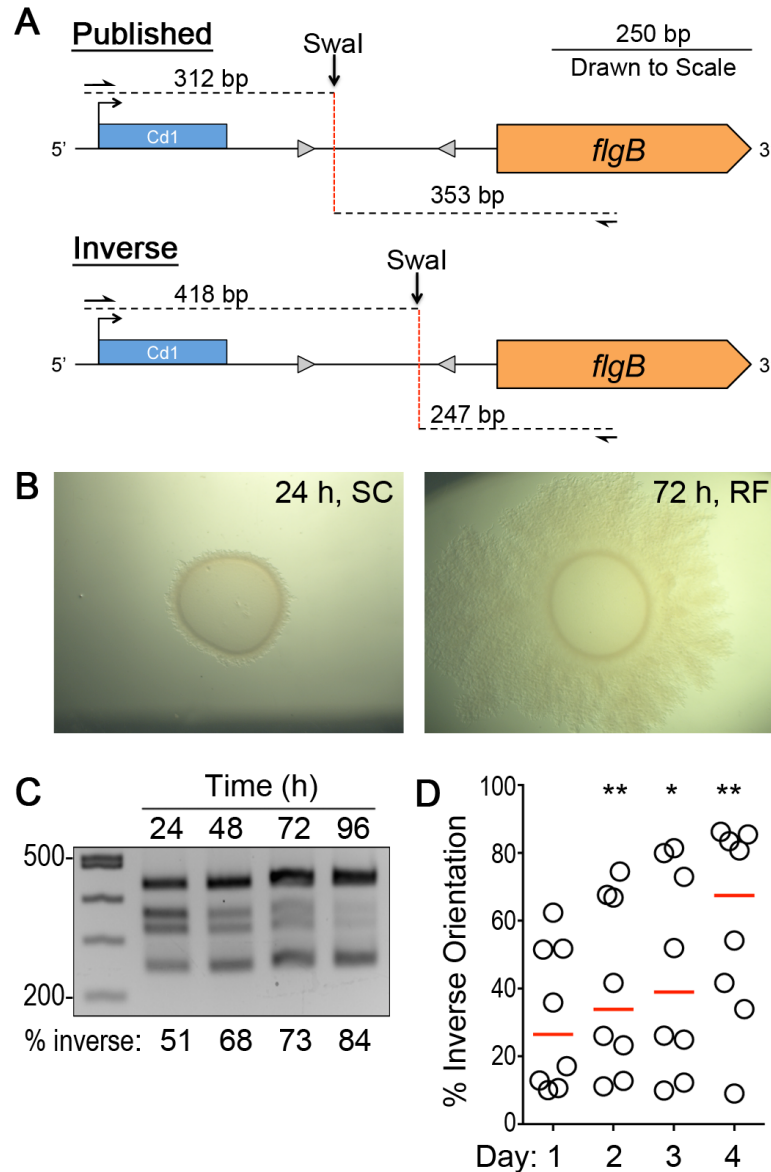
Figure 2.2. Enrichment for flagellar phase variant populations. (A) Diagram of the asymmetric PCR-digestion method used to determine the proportions of bacteria in the population with the putative flagellar switch in each orientation. (B) Representative images of *C. difficile* R20291 colonies after 24 hours of growth, which are circular with smooth edges (left), and after 72 hours of growth, which are rough with filamentous edges (right). (C) Asymmetric PCR-digestion assay showing the proportions of bacteria with the flagellar switch in each orientation within colonies collected every 24 hours for four days. Below each lane is the percentage of the colony population in which the flagellar switch is in the inverse orientation, calculated by comparing relevant band intensities to the total normalized to a standard curve. (D) Quantitative measurements of the percentage of bacteria in each colony with the switch in the inverse orientation over the course of four days, with relevant band intensities normalized to a standard curve. Each circle represents one of eight biological replicates monitored over time. * $p < 0.05$, ** $p < 0.01$ by one-way ANOVA and Dunnett's multiple comparisons test.

94
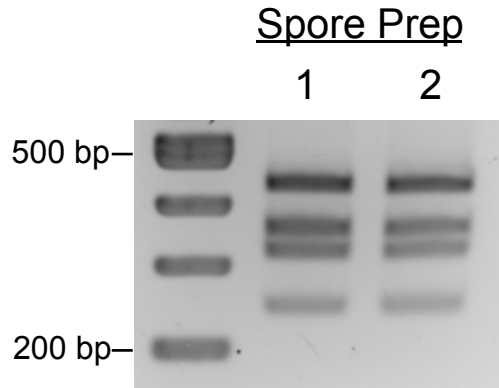
Figure 2.3. Spore stocks of *C. difficile* R20291 contain both *flg* ON and OFF bacteria.
Boiled lysates of *C. difficile* R20291 spores served as the templates in an asymmetric PCR-digestion assay with primers R591 and R857 and the restriction enzyme SwaI. Shown are the results for two independent spore preparations.
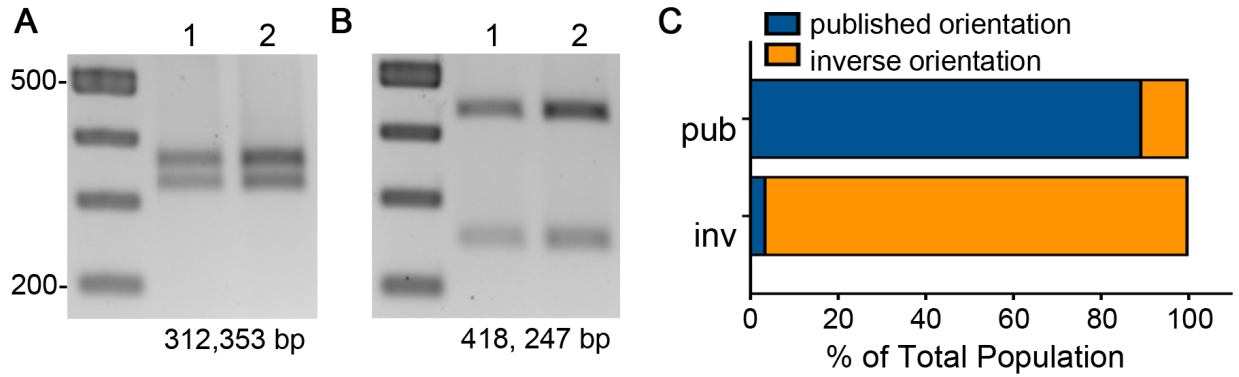
Figure 2.4. Assessing the purity of enriched populations. (A, B) Asymmetric PCR-digestion assay of genomic DNA to determine the orientation of the flagellar switch in individual colonies. Bacteria from individual colonies contain the switch predominantly in the published orientation (A) or the inverse orientation (B). Two representatives of each are shown. (C) Quantitative PCR results of the flagellar switch orientation in enriched phase variant populations (published, n=9; inverse, n=10). The *ΔΔCt* method was used to determine the relative DNA copies of the flagellar switch orientation in enriched populations relative to the *rpoC* gene.
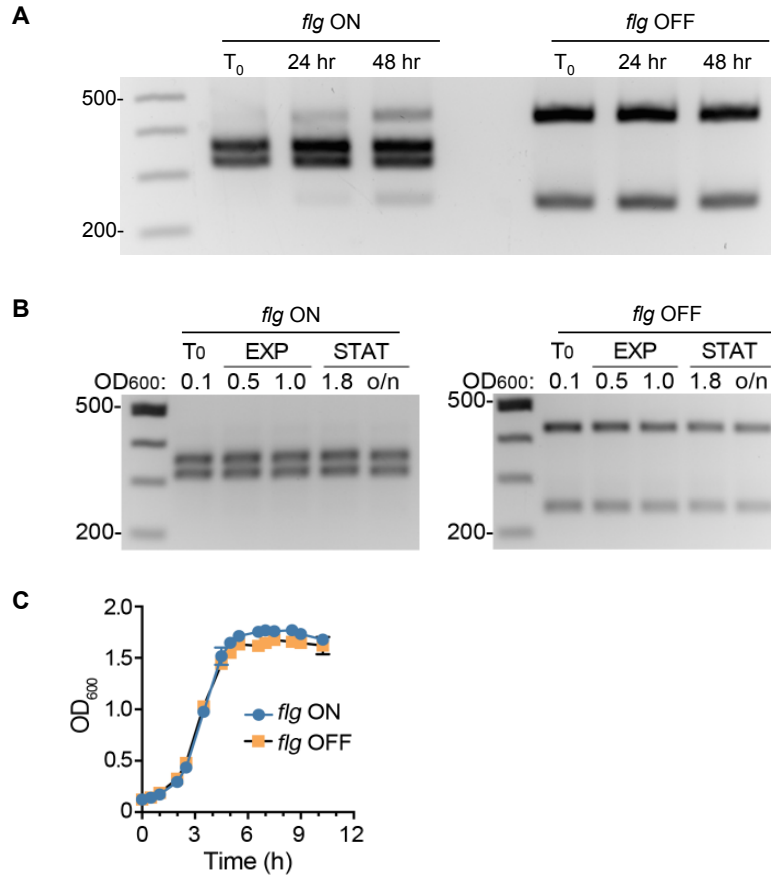
Figure 2.5. Stability of the flagellar switch during growth in liquid and solid media. (A) Asymmetric PCR-digestion assay with template from enriched *C. difficile* R20291 *flg* ON and OFF isolates grown on an agar surface for 24 hours and 48 hours. T_0 represents the cultures used to inoculate the agar plates. (B) Asymmetric PCR-digestion assay of *C. difficile* R20291 *flg* ON and OFF isolates grown in BHIS medium collected at T_0, two exponential phase time points (EXP, OD_{600} 0.5 and 1.0), and two stationary phase time points (STAT, OD_{600} 1.8 and overnight, O/N). Images are representative from two independent experiments with four replicates of each *flg* phase. (C) Growth curve of *C. difficile* R20291 *flg* ON and OFF isolates in BHIS medium. Data are combined from two independent experiments each with two replicates of each *flg* phase, and means and standard deviations are shown.
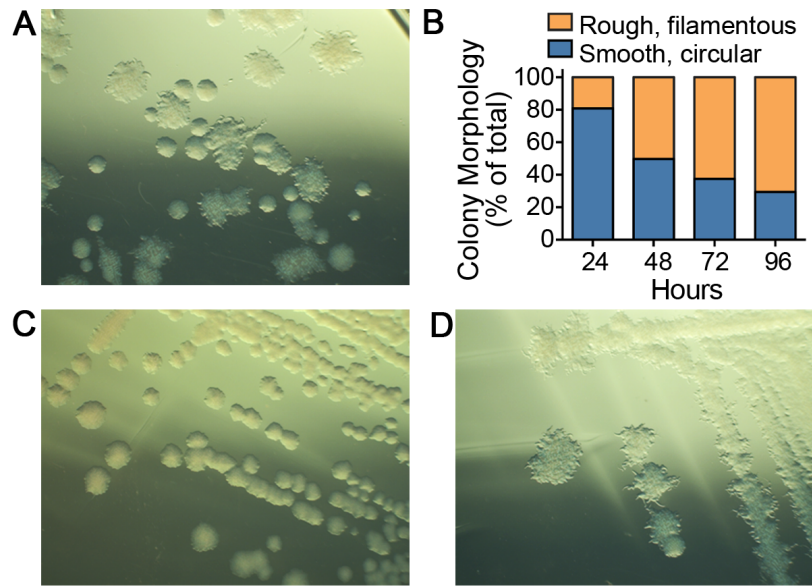
Figure 2.6. Appearance of smooth, circular (SC) and rough, filamentous (RF) colony morphotypes. *C. difficile* R20291 spores were plated on BHIS agar supplemented with taurocholate to obtain single colonies. After 24 hours, individual colonies were suspended in media and spotted onto fresh BHIS plates. Every 24 hours for four days, colony spots were suspended, serially diluted, and plated on BHIS to enumerate colonies based on their morphologies. (A) We observed two distinct colony morphologies based on colony texture and edge– a smooth, circular (SC) morphotype, and a rough, filamentous (RF) morphotype. (B) The percentage of RF colonies increased, and the percentage of SC colonies decreased over time. Using the asymmetric PCR-digestion assay to determine the orientation of the flagellar switch in SC and RF colonies after 96 hours, colony morphology and switch orientation were determined to be unlinked. 72% of bacteria in the SC colonies had the flagellar switch in the published orientation, and 64% in the RF colonies had the switch in the inverse orientation. (C,D) Bacteria from SC and RF colony morphology maintained their respective characteristic morphologies after passaging.
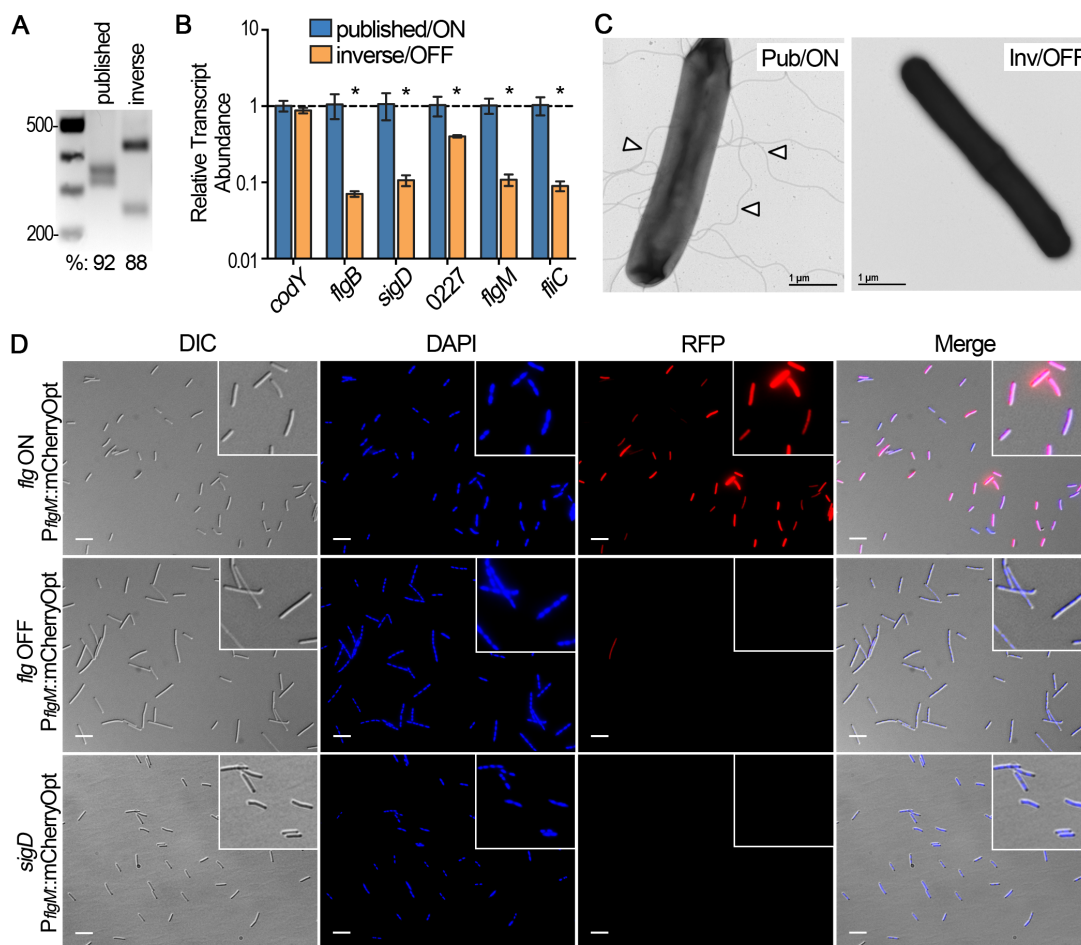
Figure 2.7. The orientation of the flagellar switch impacts the expression of the downstream flagellar genes. (A) Asymmetric PCR-digestion assay performed on *C. difficile* R20291 isolates with the flagellar switch in the published and inverse orientations respectively. (B) qRT-PCR was used to determine the abundance of representative flagellar gene transcripts in isolates with the flagellar switch in the published and inverse orientation. Four independent isolates were tested, and *Ct* values for each flagellar gene and the *codY* gene (non-regulated control) were normalized to those of the housekeeping gene *rpoC*; the published orientation samples were arbitrarily chosen as the reference condition. Shown are the means and standard deviations. * $p < 0.05$ by t-tests comparing mean transcript abundances between published and inverse samples, n = 4. (C) Visualization of flagella by transmission electron microscopy at 25,000X magnification. Size bars = 1 micron. Representative images of bacterial flagellar switch isolates are shown. Arrowheads indicate flagella. (D) Micrographs of enriched *flg* ON, *flg* OFF, and a *sigD* mutant transformed with the pP*flgM::mCherryOpt* reporter. Channels used are indicated for each column; the fourth column images are a merge of the DIC, DAPI, and RFP. RFP positive and negative bacteria were visually enumerated relative to the DIC and DAPI channels, and quantifications are shown in Fig 2.8. White bars = 10 microns.
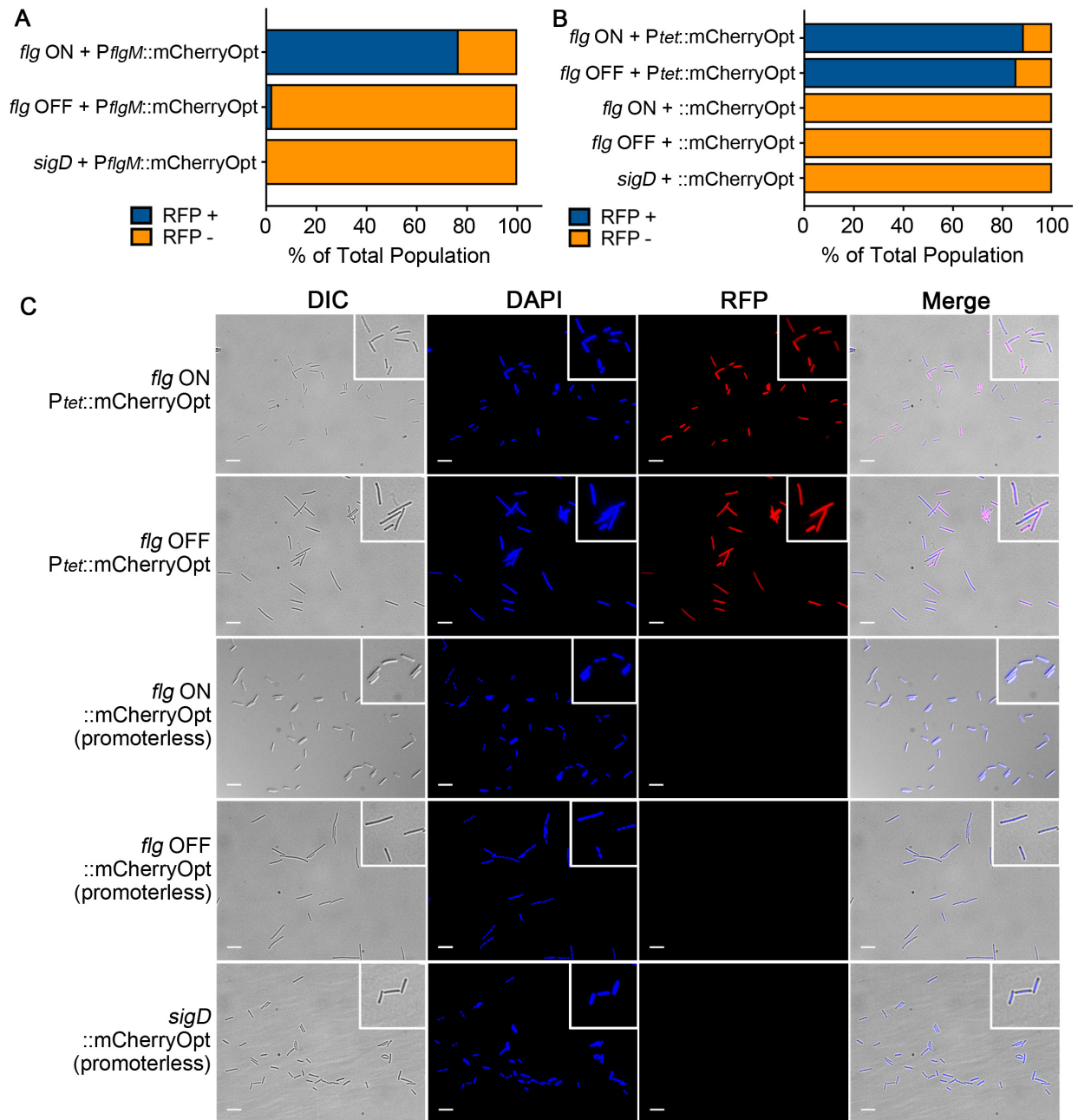
Figure 2.8. Quantification and controls for fluorescence microscopy studies in Fig 2.7. (A) Frequency of RFP (mCherryOpt)-positive and -negative bacteria for strains evaluated in Fig. 2.7D. Number of bacteria counted per group transformed with P*flgM*::*mCherryOpt*: *flg* ON (n=3321 bacteria; 7 biological replicates), *flg* OFF (n=3720 bacteria; 6 biological replicates), and *sigD* (n=1348, 3 biological replicates). (B) Frequency of RFP (mCherryOpt)-positive and -negative bacteria of strains in panel C. Number of bacteria counted per group transformed with P$_{Tet}$::*mCherryOpt*: *flg* ON (n=662 bacteria; 3 biological replicates) and *flg* OFF (n=401 bacteria; 3 biological replicates). Number of bacteria counted per group transformed with promoterless ::*mCherryOpt*: *flg* ON (n=964 bacteria; 3 biological replicates), *flg* OFF (n=772 bacteria; 3

biological replicates), and *sigD* (n=670, 3 biological replicates). (C) Representative fluorescence micrographs for the indicated strains containing fusions of *mCherryOpt* to the P*tet* promoter (grown with ATc, positive control) or promoterless *mCherryOpt* (negative controls). White bar = 10 microns. The channels used are indicated at the top.
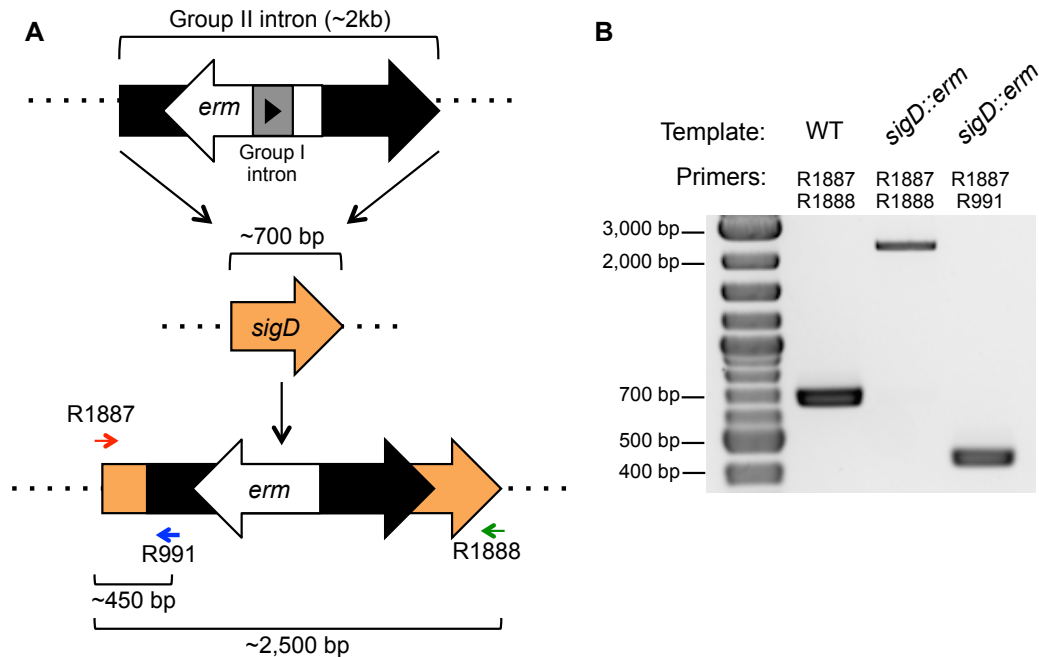
Figure 2.9. Construction and confirmation of the *sigD* mutation in *C. difficile* R20291. (A).
Schematic diagram of the Group II intron disruption of *sigD* in *C. difficile* R20291
(CDR20291_0270). The Targetron construct was designed previously to insert at nucleotide
position 228 of the *sigD* gene in the sense orientation (60). The *sigD* gene is 702 bp. The forward
and reverse primers, R1887 (red arrow) and R1888 (green arrow), partially flank *sigD* and
produce a PCR product of 716 bp. Insertion of the Group II intron into *sigD* (*sigD::ermB*) results
in a PCR product of ~2500 bp. A second PCR reaction was used to confirm *sigD* by using R991,
a Group II intron specific primer called EBS Universal. A PCR reaction with the R1887 and
R991 yields a product of ~450 bp if the Group II intron is in *sigD*. (B)**.** Image of an EtBr stained
agarose gel with PCR products for the different reactions detailed in (A). Lane 1: Wildtype
R20291 used as template with primers R1887 and R1888. Lane 2: The putative R20291 *sigD*
mutant used as template with primers R1887 and R1888. Lane 3: The putative R20291 *sigD*
mutant used as template with primers R1887 and R991.
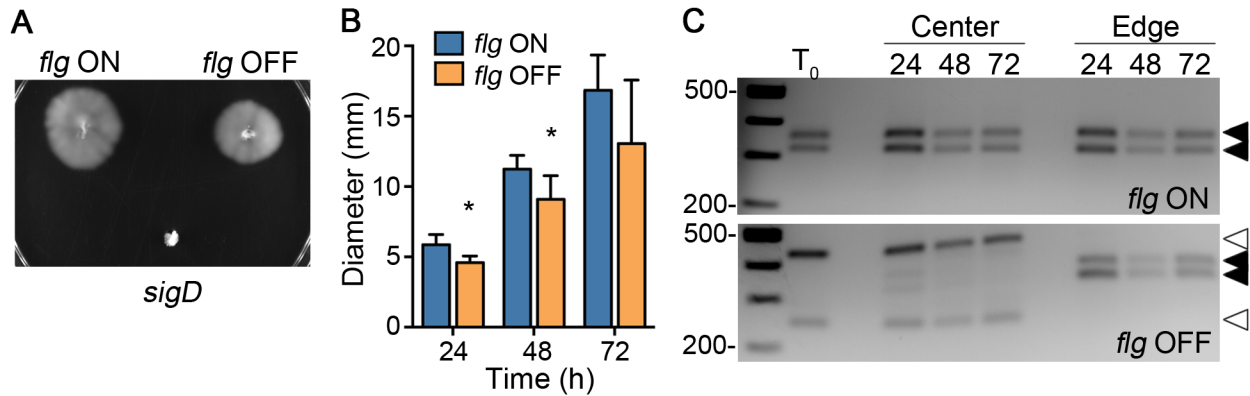
Figure 2.10. Motility medium spatially segregates flagellar phase variant populations. (A) The motility of *C. difficile* R20291 *flg* ON and OFF isolates was evaluated in soft agar medium. A non-motile R20291 *sigD* mutant was included as a control. (B) The motility of *flg* ON and OFF isolates was quantified by measuring the diameters every 24 hours for 3 days, and means and standard deviations are shown. * $p < 0.05$ by t-tests comparing values at each time point. The data are combined from two independent experiments with four biological replicates of each *flg* phase. (C) At 24, 48 and 72 hours, bacteria were sampled from the center and edges of the growth in the motility assays for *flg* ON and OFF isolates (top and bottom panels, respectively) and subjected to asymmetric PCR-digestion assays. Black arrows indicate products for the ON orientation; open arrows, the OFF orientation. Shown are representative images from 2 independent experiments, each with two biological replicates of each *flg* phase.
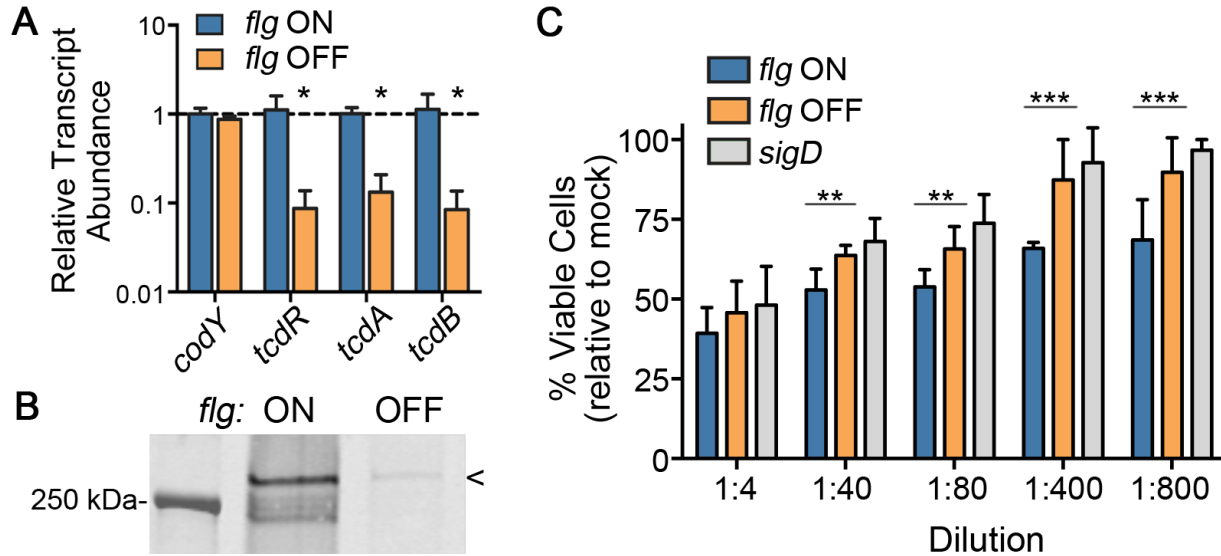
Figure 2.11. The orientation of the flagellar switch impacts toxin production. (A) qRT-PCR was used to determine the abundance of the indicated transcripts in *flg* ON and OFF isolates of *C. difficile* R20291. Four independent isolates were tested, and C*t* values for each gene were normalized to those of the housekeeping gene *rpoC*; the *flg* ON samples were arbitrarily chosen as the reference condition. Shown are means and standard deviations. * $p < 0.05$ by t-tests comparing mean transcript abundances between *flg* ON and OFF samples, n = 4. (B) TcdA protein levels in cell lysates of *flg* ON and OFF isolates were evaluated by western blot. Shown is a representative image for three independent experiments each with at least three replicates of each *flg* phase. (C) The *flg* ON and OFF isolates, as well as the *sigD* mutant control, were grown to stationary phase in TY medium. The supernatants were serially diluted and applied to Vero cells for 24 hours. Cell viability was assessed using the CellTiter Glo assay. Data are combined from two independent experiments each with four replicates of *sigD* mutant and *flg* phase variants, and means and standard deviations are shown. ** $p < 0.01$, *** $p < 0.001$ by one-way ANOVA comparing the means for each dilution.
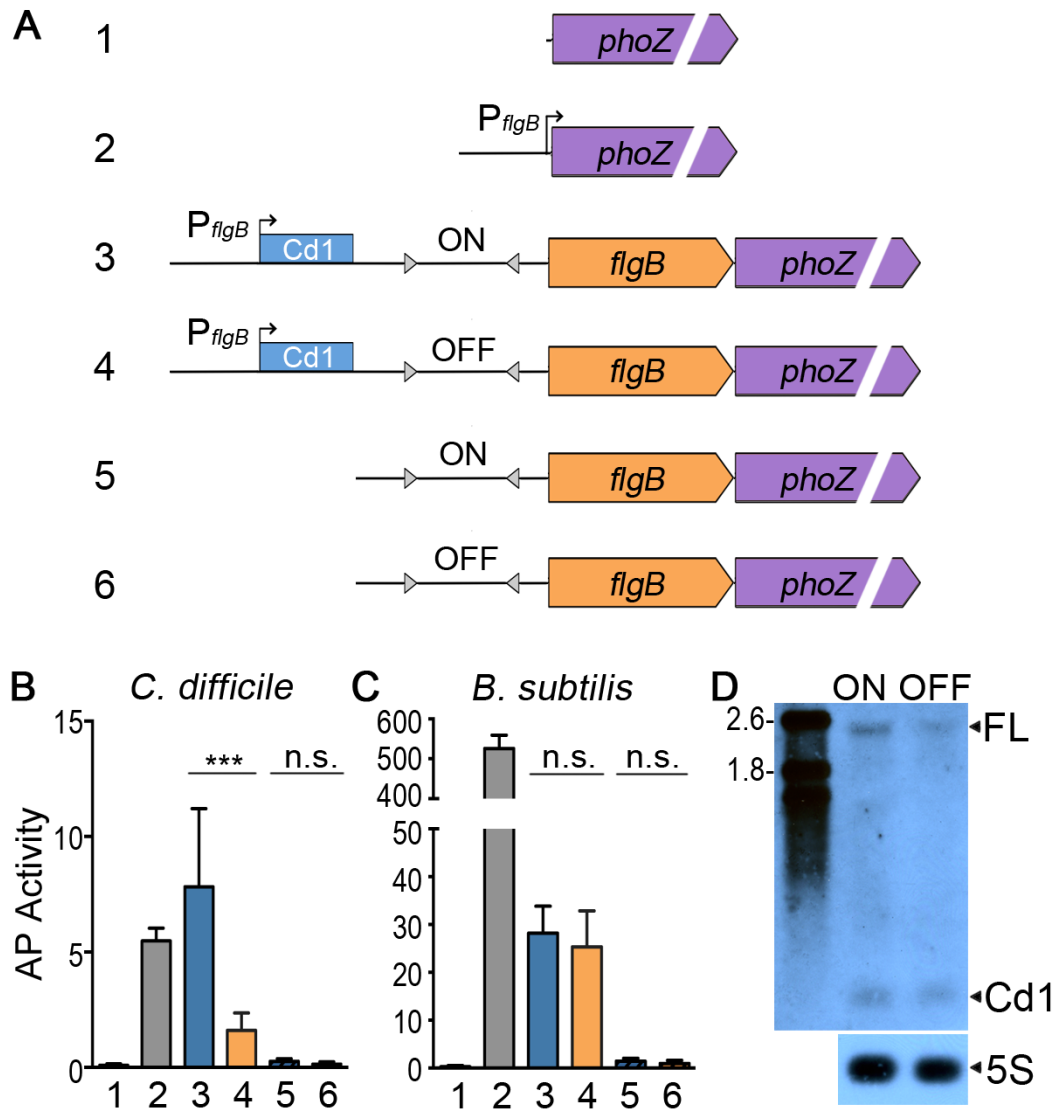
Figure 2.12. The orientation of the flagellar switch controls flagellar gene expression post-transcription initiation. (A) Diagram of the reporter gene fusions. The known *flgB* operon promoter, the Cd1 c-di-GMP riboswitch, and the orientation of the flagellar switch are indicated, if present. (B,C) The fusions in A were integrated into the *C. difficile* R20291 chromosome (left) or the *B. subtilis* BS49 chromosome via Tn916 (112). Alkaline phosphatase (AP) activity was measured as described previously (77). Means and standard deviations are shown. *** $p < 0.001$ by one-way ANOVA and Bonferroni's multiple comparisons test, n = 8. n.s. = not significant. (D) Northern blot detection of the *phoZ*-containing transcripts from *C. difficile* R20291 bearing fusions 3 or 4. The full length (FL) transcript of ~2400 nt is indicated. 5S RNA served as the loading control. The image is representative of three biological replicates for each strain.

105

A <u>Native flagellar switch</u>: ON: 579 bp, 100bp; OFF: 473 bp, 206 bp

Reporter: No Pr (1) Pr only (2) Full ON (3) Full OFF (4) FS ON (5) FS OFF (6)

Native *flg*: ON OFF ON OFF ON OFF ON OFF ON OFF ON OFF

600 bp-

100 bp-

B <u>Reporter flagellar switch</u>: ON: 589 bp, 100bp; OFF: 483 bp, 206 bp

Reporter: Full ON (3) Full OFF (4) FS ON (5) FS OFF (6)

Reporter *flg*: ON ON OFF OFF ON ON OFF OFF

Native *flg*: ON OFF ON OFF ON OFF ON OFF

600 bp-

100 bp-

C *B. subtilis*: <u>Reporter flagellar switch</u>: ON: 610 bp, OFF: 516 bp

Reporter: Full ON (3) Full OFF (4) FS ON (5) FS OFF (6)

Primer: ON OFF ON OFF ON OFF ON OFF

700 bp-

400 bp-

Figure 2.13. Confirmation of flagellar switch orientation accompanying alkaline phosphatase assays in *C. difficile* and *B. subtilis*. (A, B) Asymmetric PCR-digestion assay to determine the orientation of the flagellar switch. For the native locus, we used primers R1705 and R1704 in PCR reactions. R1705 (forward primer) anneals 3' of the Cd1 riboswitch DNA sequence, and R1704 (reverse primer) anneals to the second gene in the *flgB* operon, *flgC* (CDR20291_0249). For the reporter locus (B), we used R1705 and R1706, a reverse primer that anneals to the alkaline phosphatase gene, *phoZ*, for PCR amplification. Biological replicates for each AP reporter in *C. difficile* R20291 were combined and genomic DNA was extracted, based on published methods, to simultaneously determine the orientation of the native flagellar switch in each. The numerical designations for each reporter are shown in parentheses and correspond to Fig. 2.12A. We confirmed that both the native and reporter flagellar switches for each AP reporter strain were in their expected orientations. (C) Orientation-specific PCR assay of AP reporters #3-6 in *B. subtilis* showing that the AP reporter flagellar switches maintained their orientations.
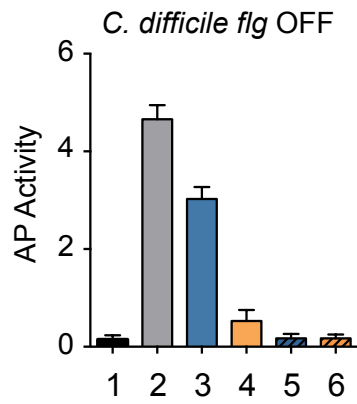
Figure 2.14. Alkaline phosphatase activity of the *phoZ* gene reporters in *C. difficile flg* OFF. Activity was assessed as described in the main text in tandem with the same reporters in the R20291 *flg* ON background.
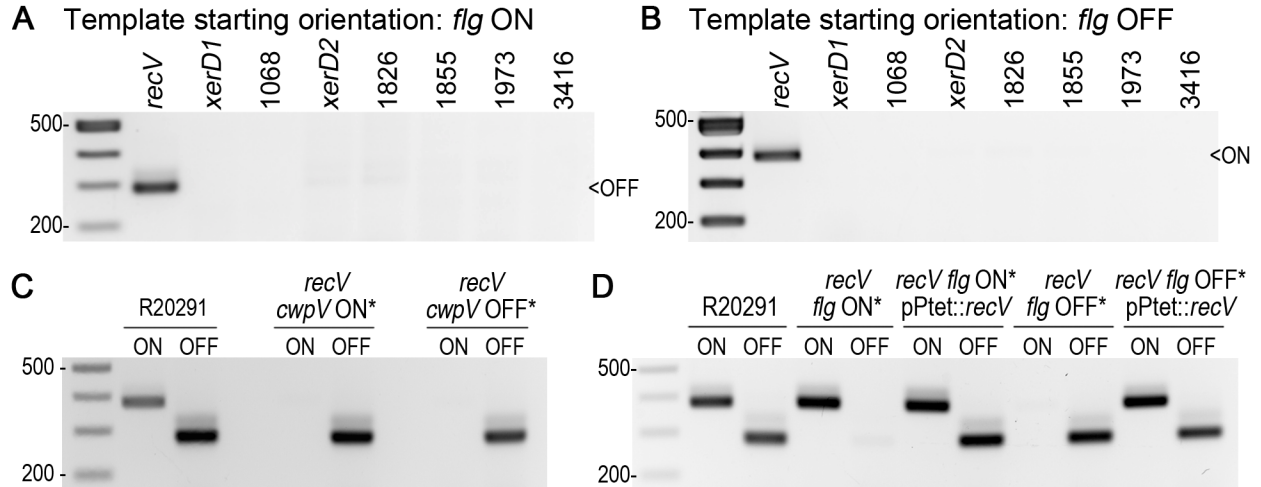
Figure 2.15. Identification of the recombinase that mediates inversion of the flagellar switch. (A,B) Orientation-specific PCR (Fig 1B) assays to identify the conserved recombinase that catalyzes inversion at the flagellar switch from ON to OFF (A) and OFF to ON (B). The gene name or R20291 locus tag numbers for the eight conserved recombinases are shown. A 375 bp product indicates the ON orientation; a 281 bp product indicates the OFF orientation. (C) Orientation-specific PCR for the flagellar switch to determine whether the *recV cwpV* locked ON and OFF mutants were locked for the flagellar switch. (D) Orientation-specific PCR for the flagellar switch in complemented *recV* mutants (pP*tet*::*recV*) and controls.
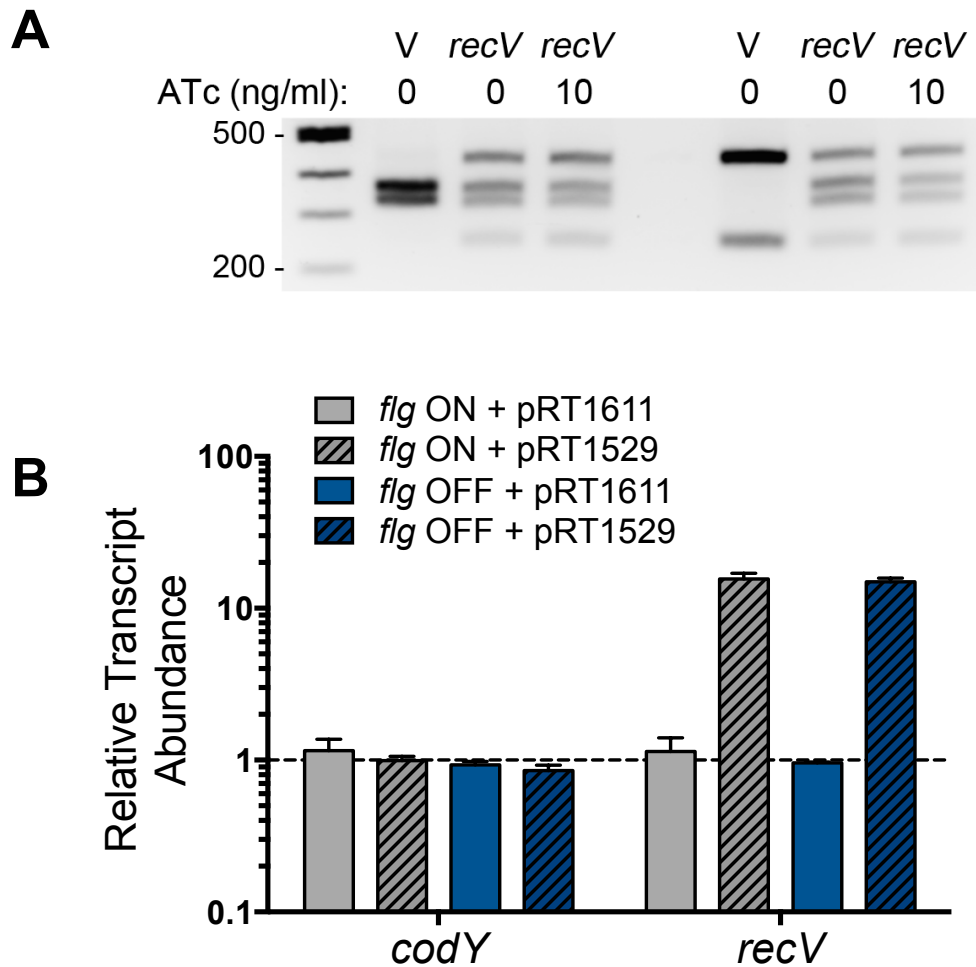
Figure 2.16. Expression of *recV* in *C. difficile flg* ON and *flg* OFF isolates stimulates inversion of the flagellar switch. (A) Asymmetric PCR-digestion assay of products from *C. difficile* R20291 *flg* ON and OFF isolates transformed with pRT1611 (V, vector), or pRT1529 (pRPF185::*recV*), grown with or without 10 ng/ml ATc (Table 2.1). (B) Increased transcription of *recV* occurs in *C. difficile* with pRT1529 in the absence of ATc induction. qRT-PCR measuring transcript abundance of *recV* and *codY* control gene in R20291 *flg* ON and OFF, each with pRT1611 (vector) and RT1529 (*recV*). Data were normalized to the *rpoC* gene and expressed relative to the respective parental strain with pRT1611. Shown are means and standard deviations.
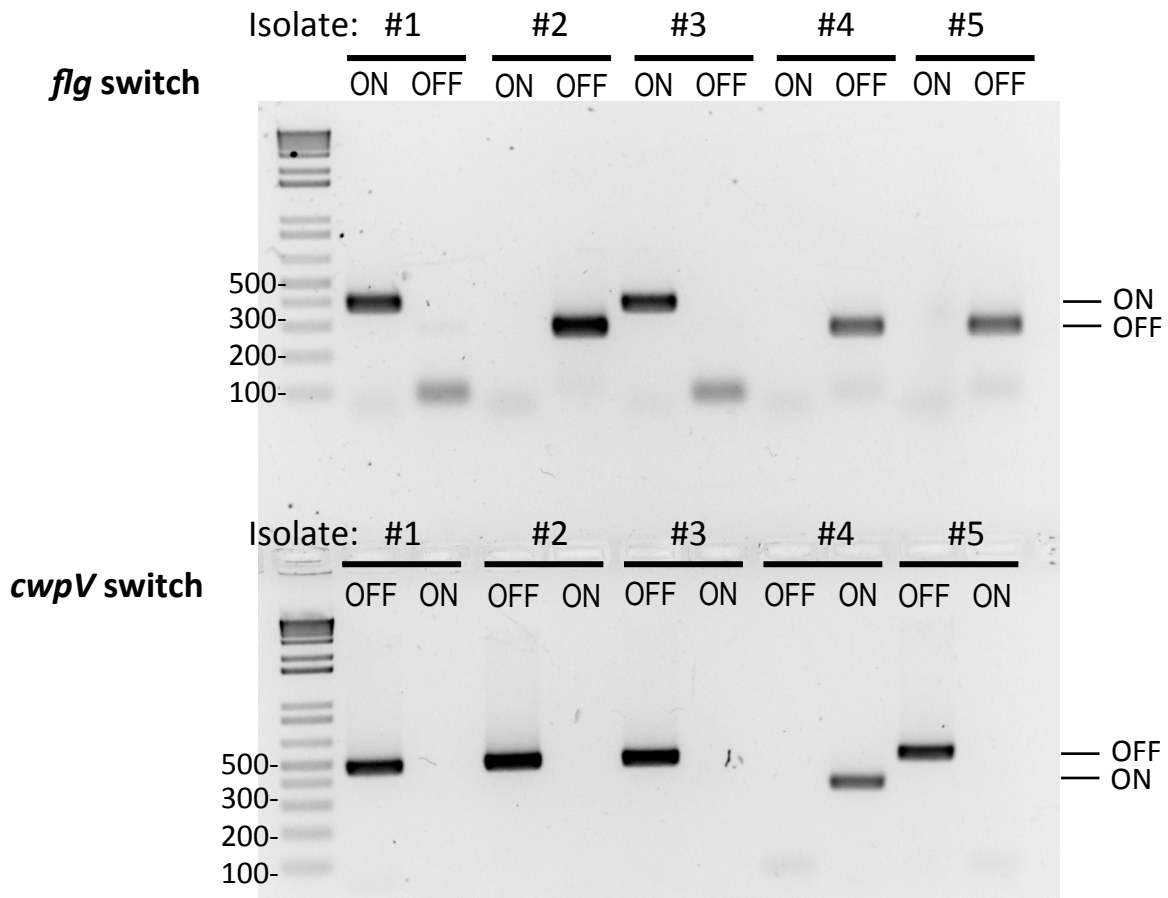
Figure 2.17. Identification of *C. difficile* R20291 *recV flg* ON* mutants. *C. difficile* R20291 *recV flg* OFF* strain was transformed with a plasmid for expression of *recV* to allow flagellar switch inversion. The strain was then passaged without antibiotics to allow plasmid loss. Five thiamphenicol-sensitive colonies were identified and screened by PCR as in Fig. 2.1B for the orientation of the *cwpV* and flagellar switches. Shown: orientation-specific PCR assay of five isolates, two of which (#1, #3) have the flagellar switch in the ON orientation; both have the *cwpV* switch in the OFF orientation.
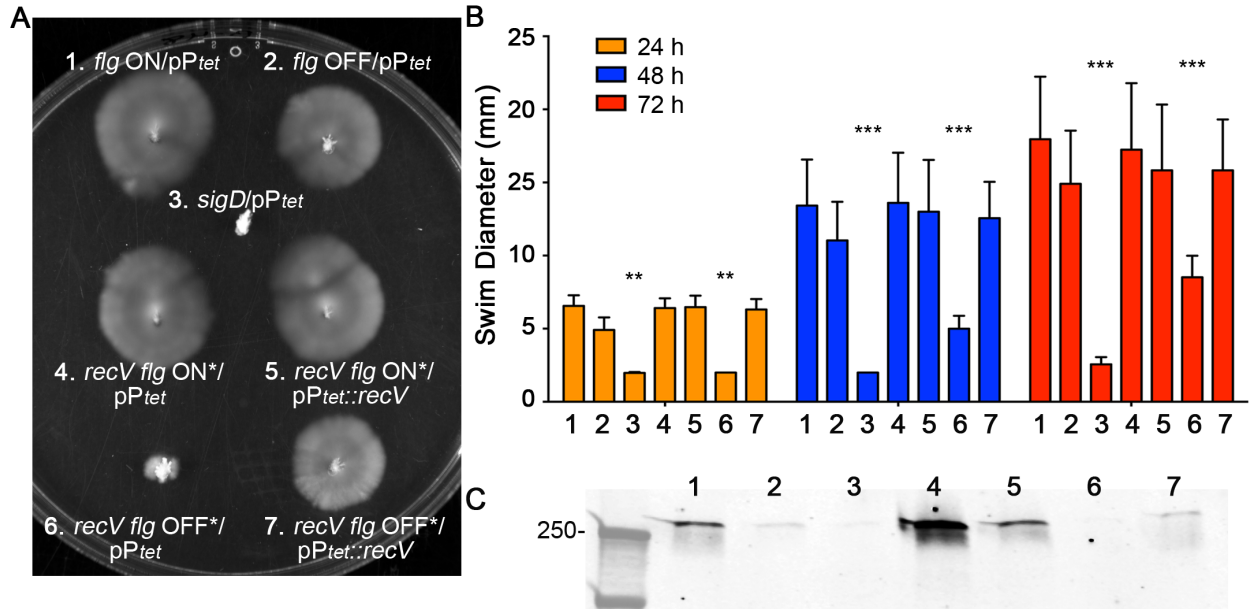
Figure 2.18. Mutation of *recV* results in phase-locked motility and toxin production. (A,B)
C. difficile strains were assayed for motility in BHIS-0.3% agar. (A) Photographs were taken
every 24 hours, and a representative image from 48 hours is shown. (B) Measurements of the
diameters of motility were taken at 24, 48 and 72 hours. Shown are means and standard
deviations from six biological replicates. ** $p < 0.01$, *** $p < 0.001$, by two-way ANOVA and
Tukey's multiple comparison test, comparing to *flg* ON/pP$_{tet}$. (C) TcdA protein levels in cell
lysates were evaluated by western blot. Shown is a representative image of three independent
assays. (B, C) Numbers indicate the strain, as indicated in panel A.
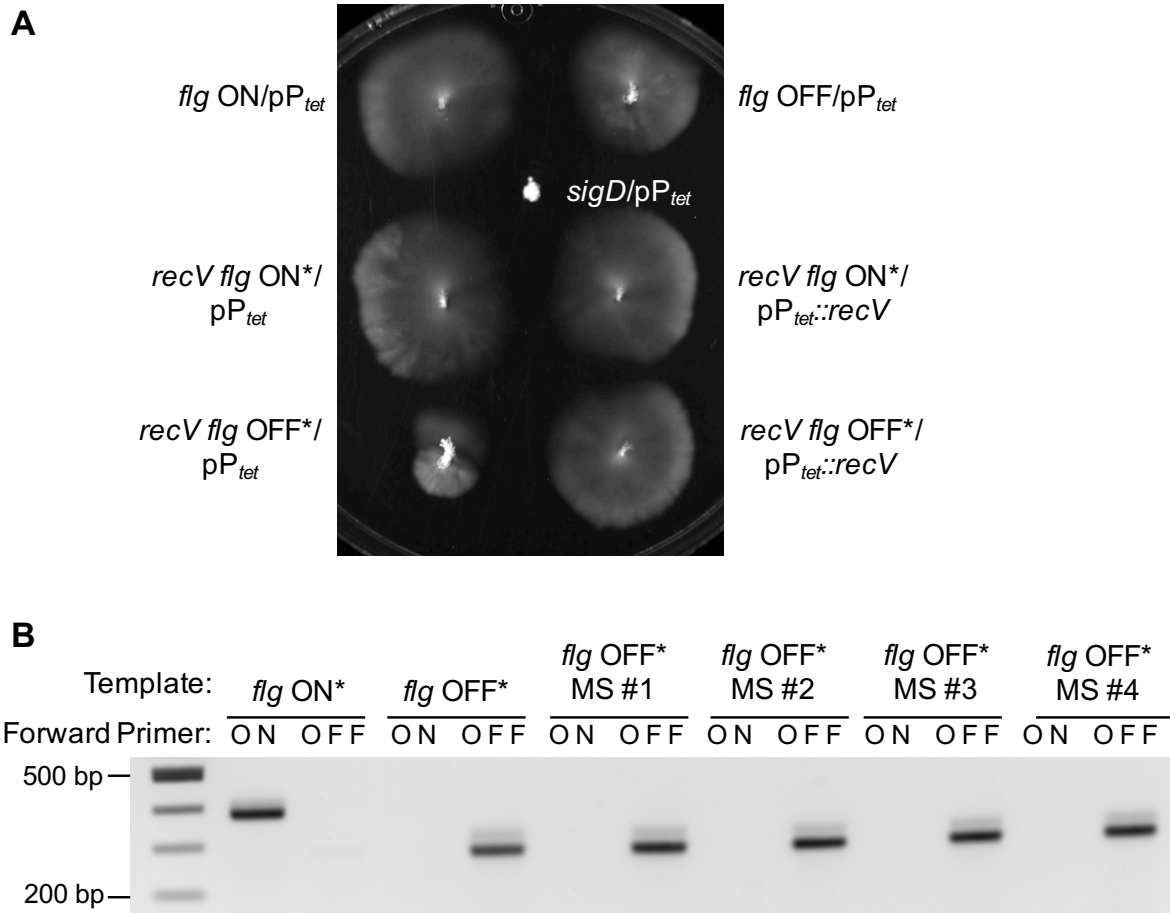
Figure 2.19. Isolation of motile *recV flg* OFF* suppressor mutants. (A) Representative image of a motility assay after 72 hours, with *flg* ON/pP*tet* (RT1615), *flg* OFF/pP*tet* (RT1617), *sigD*/pP*tet* (RT1690), *recV flg* ON*/pP*tet* (RT1715), *recV flg* ON*/pP*tet::recV* (RT1716), *recV flg* OFF*/pP*tet* (RT1691), and *recV flg* OFF*/pP*tet::recV* (RT1697). Strain numbers are listed in parentheses. The *recV flg* OFF*/pP*tet::recV* (RT1691) showed motility upon this prolonged incubation. (B) Orientation-specific PCR assay of the flagellar switch from *recV flg* ON*, *recV flg* OFF*, and three motile suppressor (MS) mutants of *recV flg* OFF* (pRT1719 – RT1724). Image representative of two independent experiments with eight biological replicates of motile suppressor mutants of *recV flg* OFF*.

```
630          AAGTTACTATTTTACAAAAA- GAGCAACTTTTCGAAGAAATATTTAAATACATATTAAAATTTTT  64
BI9          AAGTTACTATTTTACAAAAAA GAGCAACTTTTCGAAGAAATATTTAAATACATATTAAAATTTTT  65
M68          AAGTTACTATTTTACAAAAAA GAGCAACTTTTTGAAGAAATATTTAAATACATATTAAAATTTTT  65
CF5 Inverse  AAGTTACTATTTTACAAAAAA GAGCAACTTTTTGAAGAAATATTTAAATACATATTAAAATTTTT  65
R20291       AAGTTACTATTTTACAAAAAA AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTT  65
CD196        AAGTTACTATTTTACAAAAAA AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTT  65
BI1 Inverse  AAGTTACTATTTTACAAAAAA AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTT  65
2007855      AAGTTACTATTTTACAAAAAA AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTT  65

630          TTATATTTTCATGGGACTTCATTAAAAAAATATTAGTTTTCTTACTAAATTGATACATTATTGTATT  131
BI9          TTATATTTTCATGGGACTTCATTGAAAAAATATTAGTTTTCTTACTAAAATGATACATTATTGTATT  132
M68          TTATATTTTCATGGGACTTCATTGAAAAAATATTAGTTTTCTTACTAAAATGATACATTATTGTATT  132
CF5 Inverse  TTATATTTTCATGGGACTTCATTGAAAAAATATTAGTTTTCTTACTAAAATGATACATTATTGTATT  132
R20291       TTATATTTTTATTAGGATTTACGAAAATAATATTAGTTTTCTTACCAAAGTGATACATTATTATATT  132
CD196        TTATATTTTTATTAGGATTTACGAAAATAATATTAGTTTTCTTACCAAAGTGATACATTATTATATT  132
BI1 Inverse  TTATATTTTTATTAGGATTTACGAAAATAATATTAGTTTTCTTACCAAAGTGATACATTATTATATT  132
2007855      TTATATTTTTATTAGGATTTACGAAAATAATATTAGTTTTCTTACCAAAGTGATACATTATTATATT  132

630          AATGTAATGGAAATGCGATTAATTTAAATAAATAAAGAAGTCA TT-TTTGTAATATAGCAACTT  195
BI9          AATGTAATGGAAATGCGATTAATTTAAATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
M68          AATGTAATGGAAATGCGATTAATTTAAATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
CF5 Inverse  AATGTAATGGAAATGCGATTAATTTAAATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
R20291       AATGTAATAGAAATGTAATTAATTTGGATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
CD196        AATGTAATAGAAATGTAATTAATTTGGATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
BI1 Inverse  AATGTAATAGAAATGTAATTAATTTGGATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
2007855      AATGTAATAGAAATGTAATTAATTTGGATGAATAAAGAAGTCA TTTTTTGTAATATAGCAACTT  196
```

Figure 2.20. Sequence alignment of the flagellar switch and inverted repeat sequences from NCBI accessible genomes of *C. difficile.* The following *C. difficile* genomes were used in a sequence alignment for the flagellar switch: 630, BI9, M68, CF5, R20291, CD196, BI1, and 2007855.

Table 2.1: Strains and plasmids used in this study.

Strain/Plasmid	Description/Purpose	Reference
Plasmids		
pMC123	*E. coli* – *C. difficile* shuttle vector; Amp^R, Cm^R/Tm^R	(117)
pMC358	pMC123 with *Enterococcus faecalis phoZ,* including native *phoZ* promoter	(77)
pRT1323	pMC123::P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ* pMC123 with *flgB* operon promoter, 5' UTR with the flagellar switch in the ON orientation, and *flgB,* fused to *phoZ*	This work
pRT1324	pMC123::P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ* pMC123 with *flgB* operon promoter, 5' UTR with the flagellar switch in the OFF orientation, and *flgB,* fused to *phoZ*	This work
pSMB47	Tn*916* integrational vector Erm^R/CmR	(118)
pRT1532	pSMB47::*slpATT;* pSMB47 with the *C. difficile* 630 *slpA* Rho-independent terminator	This work
pRT1346	Promoterless *phoZ* in pRT1532 (::*phoZ*)	This work
pRT1345	*C. difficile* R20291 *flgB* operon promoter fused to *phoZ* in pRT1532 (P$_{flgB}$::*phoZ*)	This work
pRT1326	*C. difficile* R20291 *flgB* operon promoter, 5' UTR with the flagellar switch in the ON orientation, and *flgB,* fused to *phoZ* in pRT1532 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This work
pRT1344	*C. difficile* R20291 *flgB* operon promoter, 5' UTR with the flagellar switch in the OFF orientation, and *flgB,* fused to *phoZ* in pRT1532 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This work
pRT1254	*C. difficile* R20291 flagellar switch in the ON orientation and *flgB,* fused to *phoZ* in pRT1532 (FS^{ON}-*flgB::phoZ*)	This work
pRT1286	*C. difficile* R20291 flagellar switch in the OFF orientation and *flgB* gene, fused to *phoZ* in pRT1532 (FS^{OFF}-*flgB::phoZ*)	This work
pMWO-074	Low copy anhydrotetracycline inducible expression vector, Kan^R	(85)
pRT1164	pMWO-074::*recV* (CDR20291_1004)	This work
pRT1165	pMWO-074::*xerD1* (CDR20291_1060)	This work
pRT1166	pMWO-074::CDR20291_1068	This work
pRT1167	pMWO-074::*xerD2* (CDR20291_1174)	This work
pRT1221	pMWO-074::CDR20291_1826	This work
pRT1222	pMWO-074::CDR20291_1855	This work
pRT1223	pMWO-074::CDR20291_1973	This work
pRT1224	pMWO-074::CDR20291_3416	This work
pRPF185	Anhydrotetracycline (ATc) inducible expression vector; Cm^R/Tm^{10}	(87)
pRT1611	β-glucuronidase gene*, gusA,* removed to generate reporterless derivative of pRPF185	This work
pRT1529	pRPF185::*recV* (CDR20291_1004)	This work
pBL100	Targetron vector containing untargeted group II intron; Amp^R	(119)
pRT1073	pBL100::*sigD,* for Targetron integration at nucleotide position 228 of *sigD* (CDR20291_0270), sense orientation	(23, 60)
pDSW1728	*mCherryOpt*; gene in MCS downstream of anhydrotetracycline-inducible promoter system derivative of pRPF185 (RT1685)	(72)
pRPF144	*E. coli-C. difficile* shuttle vector with P$_{cwp2}$::*gusA* cassette; Cm^R/Tm^R	(87)
pRT1676	pRPF144::P$_{flgM}$::*mCherryOpt* (replaces P$_{cwp2}$::*gusA* in pRPF144 with P$_{flgM}$)	This work
pRT1483	pRPF144::*mCherryOpt* (promoterless)	This work
pRT1719	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{ON})-*flgB* cloned from *C. difficile* R20291 RT1702	This work
pRT1720	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{OFF})-*flgB* cloned from *C.*	This work

	difficile R20291 RT1693	
pRT1721	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{OFF})-*flgB* cloned from *C. difficile* R20291 RT1704	This work
pRT1722	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{OFF})-*flgB* cloned from *C. difficile* R20291 RT1705	This work
pRT1723	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{OFF})-*flgB* cloned from *C. difficile* R20291 RT1706	This work
pRT1724	*E. coli* DH5α pCR4-TOPO: P$_{flgB}$-5'UTR(FS^{OFF})-*flgB* cloned from *C. difficile* R20291 RT1707	This work
***Escherichia coli* Strains**		
DH5α	F- Φ80*lacZ* Δ M15 (*lacZYA-argF)*U169 *recA1 endA1 hsdR*17(rκ -, mκ +) *phoA supE44 thi*-1 *gyrA*96 *relA1 λ- tonA*	Invitrogen, (120)
HB101(pRK24)	F- *mcrB mrr hsdS20(r$_B$-m$_B$)recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20* (pRK24)	(117)
RT1246	*E. coli* DH5α co-transformed with pRT1323 and pRT1164	This work
RT1247	*E. coli* DH5α co-transformed with pRT1323 and pRT1165	This work
RT1248	*E. coli* DH5α co-transformed with pRT1323 and pRT1166	This work
RT1249	*E. coli* DH5α co-transformed with pRT1323 and pRT1167	This work
RT1250	*E. coli* DH5α co-transformed with pRT1323 and pRT1221	This work
RT1251	*E. coli* DH5α co-transformed with pRT1323 and pRT1222	This work
RT1252	*E. coli* DH5α co-transformed with pRT1323 and pRT1223	This work
RT1253	*E. coli* DH5α co-transformed with pRT1323 and pRT1224	This work
RT1310	*E. coli* DH5α co-transformed with pRT1324 and pRT1164	This work
RT1311	*E. coli* DH5α co-transformed with pRT1324 and pRT1165	This work
RT1312	*E. coli* DH5α co-transformed with pRT1324 and pRT1166	This work
RT1313	*E. coli* DH5α co-transformed with pRT1324 and pRT1167	This work
RT1314	*E. coli* DH5α co-transformed with pRT1324 and pRT1221	This work
RT1315	*E. coli* DH5α co-transformed with pRT1324 and pRT1222	This work
RT1316	*E. coli* DH5α co-transformed with pRT1324 and pRT1223	This work
RT1317	*E. coli* DH5α co-transformed with pRT1324 and pRT1224	This work
***Bacillus subtilis* Strains**		
BS49	CU2189::Tn916 for conjugation into *C. difficile*; Tet^R	(121) (122)
RT1392	BS49 transformed with pRT1346 (promoterless *phoZ*)	This work
RT1393	BS49 transformed with pRT1345 (P$_{flgB}$::*phoZ*)	This work
RT1394	BS49 transformed with pRT1326 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This work
RT1395	BS49 transformed with pRT1344 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This work
RT1331	BS49 transformed with pRT1254 (FS^{ON}-*flgB::phoZ*)	This work
RT1332	BS49 transformed with pRT1286 (FS^{OFF}-*flgB::phoZ*)	This work
***Clostridium difficile* Strains**		
630*Δerm*	Ribotype 012, erythromycin susceptible derivative of *C. difficile* 630	(123)
ATCC 43598	Ribotype 017, gift from Shonna McBride	ATCC
R20291	Ribotype 027, epidemic isolate	(19)
RT1566	*C. difficile* R20291 with a Targetron insertion at nucleotide position 228 of *sigD*, sense orientation (*sigD::erm*)	This work
RT1536	*C. difficile* R20291 *flg* ON isolate with promoterless *phoZ* fusion	This work

	integrated on the chromosome via Tn916 (RT1392)	
RT1537	*C. difficile* R20291 *flg* OFF isolate with promoterless *phoZ* fusion integrated on the chromosome via Tn916 (RT1392)	This work
RT1538	*C. difficile* R20291 *flg* ON isolate with P*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1393)	This work
RT1539	*C. difficile* R20291 *flg* OFF isolate with P*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1393)	This work
RT1540	*C. difficile* R20291 *flg* ON isolate with P*flgB*-5'UTR(FS^{ON})-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1394)	This work
RT1541	*C. difficile* R20291 *flg* OFF isolate with P*flgB*-5'UTR(FS^{ON})-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1394)	This work
RT1542	*C. difficile* R20291 *flg* ON isolate with P*flgB*-5'UTR(FS^{OFF})-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1395)	This work
RT1543	*C. difficile* R20291 *flg* OFF isolate with P*flgB*-5'UTR(FS^{OFF})-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1395)	This work
RT1544	*C. difficile* R20291 *flg* ON isolate with FS^{ON}-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1331)	This work
RT1545	*C. difficile* R20291 *flg* OFF isolate with FS^{ON}-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1331)	This work
RT1546	*C. difficile* R20291 *flg* ON isolate with FS^{OFF}-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1332)	This work
RT1547	*C. difficile* R20291 *flg* OFF isolate with FS^{OFF}-*flgB::phoZ* fusion integrated on the chromosome via Tn916 (RT1332)	This work
RT1615	*C. difficile* R20291 *flg* ON isolate with pRT1611	This work
RT1616	*C. difficile* R20291 *flg* ON isolate with pRT1529	This work
RT1617	*C. difficile* R20291 *flg* OFF isolate with pRT1611	This work
RT1618	*C. difficile* R20291 *flg* OFF isolate with pRT1529	This work
RT1690	*C. difficile* R20291 *sigD::ermB* with pRT1611	This work
RT1689	*C. difficile* R20291 *sigD::ermB* with pRT1676	This work
RT1695	*C. difficile* R20291 *flg* ON isolate with pRT1676	This work
RT1696	*C. difficile* R20291 *flg* OFF isolate with pRT1676	This work
RT1698	*C. difficile* R20291 *flg* ON isolate with pDSW1728	This work
RT1699	*C. difficile* R20291 *flg* OFF isolate with pDSW1728	This work
RT1712	*C. difficile* R20291 *flg* ON isolate with pRT1483	This work
RT1713	*C. difficile* R20291 *flg* OFF isolate with pRT1483	This work
RT1714	*C. difficile* R20291 *sigD::ermB* with pRT1483	This work
RT1693	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* OFF)[a]	(84)
RT1694	*C. difficile* R20291 *recV::ermB* (*cwpV* ON; *flg* OFF)[a]	(84)
RT1691	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* OFF)[a] with pRT1611	This work
RT1697	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* OFF)[a] with pRT1529	This work
RT1702	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* ON)[a]	This work
RT1715	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* ON)[a] with pRT1611	This work
RT1716	*C. difficile* R20291 *recV::ermB* (*cwpV* OFF; *flg* ON)[a] with pRT1529	This work
RT1705	*C. difficile* R20291 *recV::ermB flg* OFF motile suppressor mutant #1	This work
RT1706	*C. difficile* R20291 *recV::ermB flg* OFF motile suppressor mutant #2	This work
RT1707	*C. difficile* R20291 *recV::ermB flg* OFF motile suppressor mutant #3	This work
RT1708	*C. difficile* R20291 *recV::ermB flg* OFF motile suppressor mutant #4	This work

[a]Genotype in parentheses reflects starting orientations of the indicated switches.

Table 2.2. Primers used in this study.

Primer Name[a]	Sequence (5' → 3')[b]	Lab Annotation
*flg*Ribo012_Pub	GAGCAACTTTTCGAAGAAATATTTAAATAC	R1751
*flg*Ribo012_Inv	GTATTTAAATATTTCTTCGAAAAGTTGCTC	R1752
*flg*Ribo017_Pub	GAGCAACTTTTTGAAGAAATATTTAAATAC	R1622
*flg*Ribo017_Inv	GTATTTAAATATTTCTTCAAAAAGTTGCTC	R1623
*flg*Ribo027_Pub	AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTTTTATATTTTTATTAGG	R1614
*flg*Ribo027_Inv	CCTAATAAAAATATAAAAATATTTTAATATAAATTTAAATATTTCTTTATAAAGTTGCCT	R1615
*flgB*Conserv_Rev	AGGCATAGCATCATTTAGTGTTTCTTC	R857
*flg*UTR5'end_For	C**GAATTC**GTATACTTAAGTTAAAACTAAATAGGCAAATC (**EcoRI**)	R591
EF2973_phoZRBSF	GTT**GAATTC**AGGAGGAAACAAGGAAATGAAG (**EcoRI**)	R1609
EF2973_phoZCDSR	CAAGTT**GGATCC**CGTTCTGCTTTTTCTTCATTTTG (**BamHI**)	R1610
CDR202_PflgBnew	GTTCAA**GCATGC**GATATATTGTACAAATAAAATTGAAATATATGG (**SphI**)	R1512
CDR20291_0248CDSR	CAA**GAATTC**TTACCTCCCACTTATTATTGA	R1611
CDR20291_asyPCRF	CACCTATAATAACATATTATACCAAATATAATTTAAAATAAAG	R1705
*slpA*Term_F	GCGCACC**GGATCC**TATAAGT (**BamHI**)	R1848
*slpA*Term_R	TGACTC**AAGCTT**CATCTTTTTATTTAGG (**HindIII**)	R1849
EF2973_phoZRBSF	GTTCAA**GCATGC**AGGAGGAAACAAGGAAATGAAG (**SphI**)	R1632
CDR20291_0248PR	GTT**GAATTC**AACTTAAGTATACAATAAATAAC (**EcoRI**)	R1608
CDR20291_5'flgLIR	GTCAA**GCATGC**CACCTATAATAACATATTATACC (**SphI**)	R1673
phoZ_AsymR	GAATGTTAATAAGGTAACCCCTAGCAAAGCTCTTTTCTTC	R1706
CDR20291_0249R	GCTGTTAATCCACTAGCAGATATTCTC	R1704
CDR20291_1004F (*recV*)	GTCAA**GGTACC**GCAACAAGACCTATAGAAATAG (**KpnI**)	R1675
CDR20291_1004R (*recV*)	CAGTT**AAGCTT**TTAACCAATAAAGAAATTTTCAC (**HindIII**)	R1676
CDR20291_1060F	GTCAA**GGTACC**GATATTATAGAGGGATATATAGATTAC (**KpnI**)	R1677
CDR20291_1060R	CAGTT**AAGCTT**CTATTTAGCTTCTGCTTTTTGTAATTCTAC (**HindIII**)	R1678
CDR20291_1068F	GTCAA**GGTACC**AAAGACCAAGGTATCATATTAGAAAC (**KpnI**)	R1679
CDR20291_1068R	CAGTT**AAGCTT**TTACTGGTCTTCATGTAATTTCAAATGAAC (**HindIII**)	R1680
CDR20291_1174F	GTCAA**GGTACC**AAAAGGAAAGGAGTTATTATG (**KpnI**)	R1681
CDR20291_1174R	CAGTT**AAGCTT**CTATTTTTTATTTATATCCAAATTATC (**HindIII**)	R1682
CDR20291_1826F	GTCAA**GGTACC**AAAGCTGCAATTTATTCAAG (**KpnI**)	R1683
CDR20291_1826R	CAGTT**GGATCC**CTAATTGCTTAAAGAAACAGAATTATCCA (**BamHI**)	R1684
CDR20291_1855F	GTCAA**GGTACC**AATAATGATAAGAAAATTATAAAAGTAC (**KpnI**)	R1685
CDR20291_1855R	CAGTT**AAGCTT**TTAATCGTCCTCAATCCATTCAAAATTC (**HindIII**)	R1686
CDR20291_1973F	GTCAA**GGTACC**TATAAAATTTTGATTGGTGGGTTATATATG (**KpnI**)	R1687
CDR20291_1973R	CAGTT**GGATCC**TTAATTAGCATTAAATAAACTTTCAG (**BamHI**)	R1688
CDR20291_3416F	GTCAA**GGTACC**ATTATTATATATGGATATTGC (**KpnI**)	R1689
CDR20291_3416R	CAGTT**AAGCTT**CTAAATTGATTTTTTAGCACGAATAAGTG (**HindIII**)	R1690
CDR20291_1115q F (*codY*)	ATTAGGAACATTGGTACTTTCAAGAT	R910
CDR20291_1115q	TTGAACTACAGCTTTCTTTCTCATTT	R911

R (*codY*)		
CDR20291_0248q F (*flgB*)	GCAACTAATCTAAGAAGTCAGACAATAGC	R856
CDR20291_0248q R (*flgB*)	AGGCATAGCATCATTTAGTGTTTCTTC	R857
CDR20291_0270q F (*sigD*)	GAATATGCCTCTTGTAAAGAGTATAGCA	R860
CDR20291_0270q R (*sigD*)	TGCATCAATCAATCCAATGACTCC	R861
CDR20291_0227q F (autolysin)	CATCTGGATTTGATATGATTATG	R1671
CDR20291_0227q R (autolysin)	AATCTACTAGCTGTATTATTTACT	R1672
CDR20291_0230q F (*flgM*)	AAGGAAATGGCAAGTGTG	R1669
CDR20291_0230q R (*flgM*)	TTATCCTCGCATATCCTCT	R1670
CDR20291_0240q F (*fliC*)	CAAAGTAAGTCTATGGAGAA	R1584
CDR20291_0240q R (*fliC*)	ACAGATATACCATCTTGAAC	R1585
CDR20291_0581q F (*tcdR*)	AGCAAGAAATAACTCAGTAGATGATT	R908
CDR20291_0581q R (*tcdR*)	TTATTAAATCTGTTTCTCCCTCTTCA	R909
CDR20291_0584q F (*tcdA*)	GGAGAAGTCAGTGATATTGCTCTTG	R852
CDR20291_0584q R (*tcdA*)	CAGTGGTAGAAGATTCAACTATAGCC	R853
CDR20291_0582q F (*tcdB*)	AAGGAATATCTAGTTACAGAAGTATTAGAGC	R854
CDR20291_0582q R (*tcdB*)	GCAGTGTCATTTATTTGACCTCCA	R855
CDR20291_1004q F (*recV*)	TTACATACTGGCTTTAC	R1799
CDR20291_1004q R (*recV*)	GACCTAAATTAGCTTCT	R1800
CDR20291_r03No rthF	GCAAAGAGGATACACCTGT	R1795
CDR20291_r03No rthR	GGCTACGTCCTACTCTCCCA	R1796
Cd1Ribo_NorthF	GCAAATCTAGAGAAATCTAGTGACG	R1791
Cd1Ribo_NorthR	CGTATACAATATACCTAGAACTCTT	R1792
recV_RT709F	GAGTCA**GAGCTC**CAAATCAAACTAAGAGGAGTGGTTGAAA (**SacI**)	R1853
recV_RT709R	TGACTC**GGATCC**ACGTTATAATTAACCAATAAAG (**BamHI**)	R1854
CDR20291_0270F	GTAGTTAATGAATAGAGAAGAATTAAT	R1887
CDR20291_0270R	TCACCATCTATATAGAATATTTAAG	R1888
EBSuniv	CGAAATTAGAAACTTGCGTTCAGTAAAC	R991
PflgM_F_NheI	CAAT**GCTAGC**TATGCGTGAATTATAT (**NheI**)	R2117
R202flgMpromo_ R	CTAAC**GAGCTC**TAATTTAATTATCGCTC (**SacI**)	R2046
CwpVIS_R202F	CAAAACCATGTTTTTTATAACAATTCATTAAC	R1920
CwpVIS_R202R	GTTAATGAATTGTTATAAAAAACATGGTTTTG	R1921
CwpVqR	AGCATCTGCTATAGATGAGTCGTTT	R1050
qPCR_FlgSwit-ON	GTTTTCTTACCAAAGTGATACATTATTATATTAATG	R2175

qPCR_FlgSwit-OFF	CATTAATATAATAATGTATCACTTTGGTAAGAAAAC	R2176
qPCR_FlgSwit-REV	GCTATTGTCTGACTTCTTAAATTAGTTGCAT	R2177
pCR4-TOPO T3_F	ATTAACCCTCACTAAAGGGA	
pCR4-TOPO T3_R	TAATACGACTCACTATAGGG	

[a] Where locus tags are used in the primer name, the corresponding gene name is noted in parentheses

[b] Restriction site sequences are highlighted with bold text; the corresponding enzyme is listed in parentheses.

REFERENCES

1. **Cartman ST**, **Heap JT**, **Kuehne SA**, **Cockayne A**, **Minton NP**. 2010. The emergence of "hypervirulence" in *Clostridium difficile*. Int J Med Microbiol **300**:387–395.

2. **Smits WK**, **Lyras D**, **Lacy DB**, **Wilcox MH**, **Kuijper EJ**. 2016. *Clostridium difficile* infection. Nat Rev Dis Primers **2**:16020.

3. **Lessa FC**, **Mu Y**, **Bamberg WM**, **Beldavs ZG**, **Dumyati GK**, **Dunn JR**, **Farley MM**, **Holzbauer SM**, **Meek JI**, **Phipps EC**, **Wilson LE**, **Winston LG**, **Cohen JA**, **Limbago BM**, **Fridkin SK**, **Gerding DN**, **McDonald LC**. 2015. Burden of *Clostridium difficile* infection in the United States. N Engl J Med **372**:825–834.

4. **Theriot CM**, **Young VB**. 2015. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. Annu Rev Microbiol **69**:445–461.

5. **Merrigan M**, **Venugopal A**, **Mallozzi M**, **Roxas B**, **Viswanathan VK**, **Johnson S**, **Gerding DN**, **Vedantam G**. 2010. Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. J Bacteriol **192**:4904–4911.

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

7. **See I**, **Mu Y**, **Cohen J**, **Beldavs ZG**, **Winston LG**, **Dumyati G**, **Holzbauer S**, **Dunn J**, **Farley MM**, **Lyons C**, **Johnston H**, **Phipps E**, **Perlmutter R**, **Anderson L**, **Gerding DN**, **Lessa FC**. 2014. NAP1 strain type predicts outcomes from *Clostridium difficile* infection. Clin. Infect. Dis. **58**:1394–1400.

8. **Sorg JA**, **Sonenshein AL**. 2008. Bile salts and glycine as cogerminants for *Clostridium difficile* spores. J Bacteriol **190**:2505–2512.

9. **Francis MB**, **Allen CA**, **Shrestha R**, **Sorg JA**. 2013. Bile acid recognition by the *Clostridium difficile* germinant receptor, CspC, is important for establishing infection. PLoS Pathog **9**:e1003356.

10. **Paredes-Sabja D**, **Shen A**, **Sorg JA**. 2014. *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends Microbiol **22**:406-416.

11. **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

12. **Kuehne SA**, **Cartman ST**, **Heap JT**, **Kelly ML**, **Cockayne A**, **Minton NP**. 2010. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature **467**:711–713.

13. **Just I**, **Wilm M**, **Selzer J**, **Rex G**, **Eichel-Streiber von C**, **Mann M**, **Aktories K**. 1995.

The enterotoxin from *Clostridium difficile* (ToxA) monoglucosylates the Rho proteins. J Biol Chem **270**:13932–13936.

14.   **Just I**, **Selzer J**, **Wilm M**, **Eichel-Streiber von C**, **Mann M**, **Aktories K**. 1995. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature **375**:500–503.

15.   **Pruitt RN**, **Lacy DB**. 2012. Toward a structural understanding of *Clostridium difficile* toxins A and B. Front Cell Inf Microbio **2**:28.

16.   **Shen A**. 2012. *Clostridium difficile* toxins: mediators of inflammation. J Innate Immun **4**:149–158.

17.   **Carter GP**, **Chakravorty A**, **Pham Nguyen TA**, **Mileto S**, **Schreiber F**, **Li L**, **Howarth P**, **Clare S**, **Cunningham B**, **Sambol SP**, **Cheknis A**, **Figueroa I**, **Johnson S**, **Gerding D**, **Rood JI**, **Dougan G**, **Lawley TD**, **Lyras D**. 2015. Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during *Clostridium difficile* infections. MBio **6**:e00551.

18.   **Twine SM**, **Reid CW**, **Aubry A**, **McMullin DR**, **Fulton KM**, **Austin J**, **Logan SM**. 2009. Motility and flagellar glycosylation in *Clostridium difficile*. J Bacteriol **191**:7050–7062.

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

20.   **Stevenson E**, **Minton NP**, **Kuehne SA**. 2015. The role of flagella in *Clostridium difficile* pathogenicity. Trends Microbiol **23**:275–282.

21.   **Stabler RA**, **Gerding DN**, **Songer JG**, **Drudy D**, **Brazier JS**, **Trinh HT**, **Witney AA**, **Hinds J**, **Wren BW**. 2006. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J Bacteriol **188**:7297–7305.

22.   **Aubry A**, **Hussack G**, **Chen W**, **KuoLee R**, **Twine SM**, **Fulton KM**, **Foote S**, **Carrillo CD**, **Tanha J**, **Logan SM**. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun **80**:3521–3532.

23.   **Meouche El I**, **Peltier J**, **Monot M**, **Soutourina O**, **Pestel-Caron M**, **Dupuy B**, **Pons J-L**. 2013. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of *tcdR*. PLoS ONE **8**:e83748.

24.   **Faulds-Pain A**, **Twine SM**, **Vinogradov E**, **Strong PCR**, **Dell A**, **Buckley AM**, **Douce GR**, **Valiente E**, **Logan SM**, **Wren BW**. 2014. The post-translational modification of the *Clostridium difficile* flagellin affects motility, cell surface properties and virulence. Mol Microbiol **94**:272–289.

25. **Valiente E**, **Bouché L**, **Hitchen P**, **Faulds-Pain A**, **Songane M**, **Dawson LF**, **Donahue E**, **Stabler RA**, **Panico M**, **Morris HR**, **Bajaj-Elliott M**, **Logan SM**, **Dell A**, **Wren BW**. 2016. Role of glycosyltransferases modifying type b flagellin of emerging hypervirulent *Clostridium difficile* lineages and their impact on motility and biofilm formation. J Biol Chem **291**:25450–25461.

26. **Bouché L**, **Panico M**, **Hitchen P**, **Binet D**, **Sastre F**, **Faulds-Pain A**, **Valiente E**, **Vinogradov E**, **Aubry A**, **Fulton K**, **Twine S**, **Logan SM**, **Wren BW**, **Dell A**, **Morris HR**. 2016. The type b flagellin of hypervirulent *Clostridium difficile* is modified with novel sulfonated peptidylamido-glycans. J Biol Chem **291**:25439–25449.

27. **Delmée M**, **Avesani V**, **Delferriere N**, **Burtonboy G**. 1990. Characterization of flagella of *Clostridium difficile* and their role in serogrouping reactions. J Clin Microbiol **28**:2210–2214.

28. **Tasteyre A**, **Barc MC**, **Karjalainen T**, **Dodson P**, **Hyde S**, **Bourlioux P**, **Borriello P**. 2000. A *Clostridium difficile* gene encoding flagellin. Microbiology (Reading, Engl) **146 (Pt 4)**:957–966.

29. **Martin MJ**, **Clare S**, **Goulding D**, **Faulds-Pain A**, **Barquist L**, **Browne HP**, **Pettit L**, **Dougan G**, **Lawley TD**, **Wren BW**. 2013. The *agr* locus regulates virulence and colonization genes in *Clostridium difficile* 027. J Bacteriol **195**:3672–3681.

30. **Purcell EB**, **McKee RW**, **Bordeleau E**, **Burrus V**, **Tamayo R**. 2015. Regulation of Type IV pili contributes to surface behaviors of historical and epidemic strains of *Clostridium difficile*. J Bacteriol **198**:565–577.

31. **Baban ST**, **Kuehne SA**, **Barketi-Klai A**, **Cartman ST**, **Kelly ML**, **Hardie KR**, **Kansau I**, **Collignon A**, **Minton NP**. 2013. The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. PLoS ONE **8**:e73026.

32. **Dingle TC**, **Mulvey GL**, **Armstrong GD**. 2011. Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect Immun **79**:4061–4067.

33. **Barketi-Klai A**, **Monot M**, **Hoys S**, **Lambert-Bordes S**, **Kuehne SA**, **Minton N**, **Collignon A**, **Dupuy B**, **Kansau I**. 2014. The flagellin FliC of *Clostridium difficile* is responsible for pleiotropic gene regulation during in vivo infection. PLoS ONE **9**:e96876.

34. **Lyerly DM**, **Saum KE**, **MacDonald DK**, **Wilkins TD**. 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun **47**:349–352.

35. **Francis MB**, **Allen CA**, **Sorg JA**. 2013. Muricholic acids inhibit *Clostridium difficile* spore germination and growth. PLoS ONE **8**:e73653.

36. **Martin-Verstraete I**, **Peltier J**, **Dupuy B**. 2016. The regulatory networks that control

Clostridium difficile toxin synthesis. Toxins (Basel) **8**:153.

37.   **Mani N**, **Dupuy B**. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci USA **98**:5844–5849.

38.   **Matamouros S**, **England P**, **Dupuy B**. 2007. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol **64**:1274–1288.

39.   **Carter GP**, **Douce GR**, **Govind R**, **Howarth PM**, **Mackin KE**, **Spencer J**, **Buckley AM**, **Antunes A**, **Kotsanas D**, **Jenkin GA**, **Dupuy B**, **Rood JI**, **Lyras D**. 2011. The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of *Clostridium difficile*. PLoS Pathog **7**:e1002317.

40.   **Bakker D**, **Smits WK**, **Kuijper EJ**, **Corver J**. 2012. TcdC does not significantly repress toxin expression in *Clostridium difficile* 630Δ*erm*. PLoS ONE **7**:e43247.

41.   **Cartman ST**, **Kelly ML**, **Heeg D**, **Heap JT**, **Minton NP**. 2012. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the TcdC genotype and toxin production. Appl Environ Microbiol **78**:4683–4690.

42.   **Govind R**, **Dupuy B**. 2012. Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. PLoS Pathog **8**:e1002727.

43.   **Olling A**, **Seehase S**, **Minton NP**, **Tatge H**, **Schröter S**, **Kohlscheen S**, **Pich A**, **Just I**, **Gerhard R**. 2012. Release of TcdA and TcdB from *Clostridium difficile* cdi 630 is not affected by functional inactivation of the *tcdE* gene. Microb Pathog **52**:92–100.

44.   **McKee RW**, **Mangalea MR**, **Purcell EB**, **Borchardt EK**, **Tamayo R**. 2013. The second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of *sigD*. J Bacteriol **195**:5174–5185.

45.   **Syed KA**, **Beyhan S**, **Correa N**, **Queen J**, **Liu J**, **Peng F**, **Satchell KJF**, **Yildiz F**, **Klose KE**. 2009. The *Vibrio cholerae* flagellar regulatory hierarchy controls expression of virulence factors. J Bacteriol **191**:6555–6570.

46.   **Barrero-Tobon AM**, **Hendrixson DR**. 2012. Identification and analysis of flagellar coexpressed determinants (Feds) of *Campylobacter jejuni* involved in colonization. Mol Microbiol **84**:352–369.

47.   **Yoshino Y**, **Kitazawa T**, **Ikeda M**, **Tatsuno K**, **Yanagimoto S**, **Okugawa S**, **Yotsuyanagi H**, **Ota Y**. 2013. *Clostridium difficile* flagellin stimulates toll-like receptor 5, and toxin B promotes flagellin-induced chemokine production via TLR5. Life Sci **92**:211–217.

48.   **Xu H**, **Yang J**, **Gao W**, **Li L**, **Li P**, **Zhang L**, **Gong Y-N**, **Peng X**, **Xi JJ**, **Chen S**, **Wang F**, **Shao F**. 2014. Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. Nature **513**:237–241.

49.   **Jarchum I**, **Liu M**, **Lipuma L**, **Pamer EG**. 2011. Toll-like receptor 5 stimulation protects mice from acute *Clostridium difficile* colitis. Infect Immun **79**:1498–1503.

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

51.   **Pettit LJ**, **Browne HP**, **Yu L**, **Smits WK**, **Fagan RP**, **Barquist L**, **Martin MJ**, **Goulding D**, **Duncan SH**, **Flint HJ**, **Dougan G**, **Choudhary JS**, **Lawley TD**. 2014. Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics **15**:160.

52.   **Saujet L**, **Monot M**, **Dupuy B**, **Soutourina O**, **Martin-Verstraete I**. 2011. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol **193**:3186–3196.

53.   **Edwards AN**, **Tamayo R**, **McBride SM**. 2016. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. Mol Microbiol **100**:954–971.

54.   **Boudry P**, **Gracia C**, **Monot M**, **Caillet J**, **Saujet L**, **Hajnsdorf E**, **Dupuy B**, **Martin-Verstraete I**, **Soutourina O**. 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. J Bacteriol **196**:3234–3248.

55.   **Tamayo R**, **Pratt JT**, **Camilli A**. 2012. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. Annu Rev Microbiol **61**:131–148.

56.   **Römling U**, **Galperin MY**, **Gomelsky M**. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev **77**:1–52.

57.   **Purcell EB**, **Tamayo R**. 2016. Cyclic diguanylate signaling in Gram-positive bacteria. FEMS Microbiol Rev **40**:753-773.

58.   **Bordeleau E**, **Burrus V**. 2015. Cyclic-di-GMP signaling in the Gram-positive pathogen *Clostridium difficile*. Curr Genet **61**:497–502.

59.   **Purcell EB**, **McKee RW**, **McBride SM**, **Waters CM**, **Tamayo R**. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol **194**:3307–3316.

60.   **Bordeleau E**, **Purcell EB**, **Lafontaine DA**, **Fortier L-C**, **Tamayo R**, **Burrus V**. 2015. Cyclic di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. J Bacteriol **197**:819–832.

61.   **Sudarsan N**, **Lee ER**, **Weinberg Z**, **Moy RH**, **Kim JN**, **Link KH**, **Breaker RR**. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321**:411–413.

62.   **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**,

Severinov K, Le Bouguenec C, Coppée J-Y, Dupuy B, Martin-Verstraete I. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

63. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol **7**:539.

64. Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, Frost E, McDonald LC. 2005. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet **366**:1079–1084.

65. McDonald LC, Killgore GE, Thompson A, Owens RC, Kazakova SV, Sambol SP, Johnson S, Gerding DN. 2005. An epidemic, toxin gene-variant strain of *Clostridium difficile*. N Engl J Med **353**:2433–2441.

66. van der Woude MW, Bäumler AJ. 2004. Phase and antigenic variation in bacteria. Clin Microbiol Rev **17**:581–611.

67. Zhao H, Li X, Johnson DE, Blomfield I, Mobley HL. 1997. In vivo phase variation of MR/P fimbrial gene expression in *Proteus mirabilis* infecting the urinary tract. Mol Microbiol **23**:1009–1019.

68. Emerson JE, Reynolds CB, Fagan RP, Shaw HA, Goulding D, Fairweather NF. 2009. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol **74**:541–556.

69. Stabler RA, Valiente E, Dawson LF, He M, Parkhill J, Wren BW. 2010. In-depth genetic analysis of *Clostridium difficile* PCR-ribotype 027 strains reveals high genome fluidity including point mutations and inversions. Gut Microbes **1**:269–276.

70. Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE. 2001. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature **414**:555–558.

71. Lim JK, Gunther NW, Zhao H, Johnson DE, Keay SK, Mobley HL. 1998. In vivo phase variation of *Escherichia coli* type 1 fimbrial genes in women with urinary tract infection. Infect Immun **66**:3303–3310.

72. Ransom EM, Ellermeier CD, Weiss DS. 2015. Use of mCherry red fluorescent protein for studies of protein localization and gene expression in *Clostridium difficile*. Appl Environ Microbiol **81**:1652–1660.

73. Ransom EM, Weiss DS, Ellermeier CD. 2016. Use of mCherryOpt fluorescent protein in *Clostridium difficile*. Methods Mol Biol **1476**:53–67.

74. **van der Woude MW**. 2011. Phase variation: how to create and coordinate population diversity. Curr Opin Microbiol **14**:205–211.

75. **Srikhanta YN**, **Fox KL**, **Jennings MP**. 2010. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nat Rev Micro **8**:196–206.

76. **Abraham JM**, **Freitag CS**, **Clements JR**, **Eisenstein BI**. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc Natl Acad Sci U S A **82**:5724–5727.

77. **Edwards AN**, **Pascual RA**, **Childress KO**, **Nawrocki KL**, **Woods EC**, **McBride SM**. 2015. An alkaline phosphatase reporter for use in *Clostridium difficile*. Anaerobe **32C**:98–104.

78. **Smith MCM**, **Thorpe HM**. 2002. Diversity in the serine recombinases. Molecular Microbiology **44**:299–307.

79. **Esposito D**, **Scocca JJ**. 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucleic Acids Res **25**:3605–3614.

80. **Klemm P**. 1986. Two regulatory fim genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. EMBO J **5**:1389–1393.

81. **Gally DL**, **Leathart J**, **Blomfield IC**. 1996. Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. Mol Microbiol **21**:725–738.

82. **Reynolds CB**, **Emerson JE**, **la Riva de L**, **Fagan RP**, **Fairweather NF**. 2011. The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog **7**:e1002024.

83. **Sekulovic O**, **Fortier L-C**. 2015. Global transcriptional response of *Clostridium difficile* carrying the CD38 prophage. Appl Environ Microbiol **81**:1364–1374.

84. **Sekulovic O**, **Ospina Bedoya M**, **Fivian-Hughes AS**, **Fairweather NF**, **Fortier L-C**. 2015. The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. Mol Microbiol **98**:329–342.

85. **Obrist MW**, **Miller VL**. 2012. Low copy expression vectors for use in *Yersinia* sp. and related organisms. Plasmid **68**:33–42.

86. **Lewis JA**, **Hatfull GF**. 2001. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. Nucleic Acids Res **29**:2205–2216.

87. **Fagan RP**, **Fairweather NF**. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286**:27483–27493.

88.   **Oliveira Paiva AM**, **Friggen AH**, **Hossein-Javaheri S**, **Smits WK**. 2016. The signal sequence of the abundant extracellular metalloprotease PPEP-1 can be used to secrete synthetic reporter proteins in *Clostridium difficile*. ACS Synth Biol **5**:1376-1382.

89.   **Joyce SA**, **Dorman CJ**. 2002. A Rho-dependent phase-variable transcription terminator controls expression of the FimE recombinase in *Escherichia coli*. Mol Microbiol **45**:1107–1117.

90.   **Tamura Y**, **Kijima-Tanaka M**, **Aoki A**, **Ogikubo Y**, **Takahashi T**. 1995. Reversible expression of motility and flagella in *Clostridium chauvoei* and their relationship to virulence. Microbiology (Reading, Engl) **141 (Pt 3)**:605–610.

91.   **Silverman M**, **Zieg J**, **Hilmen M**, **Simon M**. 1979. Phase variation in *Salmonella*: genetic analysis of a recombinational switch. Proc Natl Acad Sci U S A **76**:391–395.

92.   **Ikeda JS**, **Schmitt CK**, **Darnell SC**, **Watson PR**, **Bispham J**, **Wallis TS**, **Weinstein DL**, **Metcalf ES**, **Adams P**, **O'Connor CD**, **O'Brien AD**. 2001. Flagellar phase variation of *Salmonella enterica* serovar Typhimurium contributes to virulence in the murine typhoid infection model but does not influence Salmonella-induced enteropathogenesis. Infect Immun **69**:3021–3030.

93.   **Josenhans C**, **Eaton KA**, **Thevenot T**, **Suerbaum S**. 2000. Switching of flagellar motility in *Helicobacter pylori* by reversible length variation of a short homopolymeric sequence repeat in *fliP*, a gene encoding a basal body protein. Infect Immun **68**:4598–4603.

94.   **Hendrixson DR**. 2006. A phase-variable mechanism controlling the *Campylobacter jejuni* FlgR response regulator influences commensalism. Mol Microbiol **61**:1646–1659.

95.   **Kearns DB**, **Losick R**. 2005. Cell population heterogeneity during growth of *Bacillus subtilis*. Genes Dev **19**:3083–3094.

96.   **Kearns DB**, **Chu F**, **Rudner R**, **Losick R**. 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol Microbiol **52**:357–369.

97.   **Mukherjee S**, **Kearns DB**. 2014. The structure and regulation of flagella in *Bacillus subtilis*. Annu Rev Genet **48**:319–340.

98.   **Calvio C**, **Osera C**, **Amati G**, **Galizzi A**. 2008. Autoregulation of swrAA and motility in *Bacillus subtilis*. J Bacteriol **190**:5720–5728.

99.   **Dubnau D**, **Losick R**. 2006. Bistability in bacteria. Mol Microbiol **61**:564–572.

100.   **Veening J-W**, **Smits WK**, **Kuipers OP**. 2008. Bistability, epigenetics, and bet-hedging in bacteria. Annu Rev Microbiol **62**:193–210.

101.   **Norman TM**, **Lord ND**, **Paulsson J**, **Losick R**. 2015. Stochastic switching of cell fate

in microbes. Annu Rev Microbiol **69**:381–403.

102. **Coyne MJ**, **Weinacht KG**, **Krinos CM**, **Comstock LE**. 2003. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proc Natl Acad Sci U S A **100**:10446–10451.

103. **Dorman CJ**. 2004. H-NS: a universal regulator for a dynamic genome. Nat Rev Micro **2**:391–400.

104. **Blomfield IC**, **Calie PJ**, **Eberhardt KJ**, **McClain MS**, **Eisenstein BI**. 1993. Lrp stimulates phase variation of type 1 fimbriation in *Escherichia coli* K-12. J Bacteriol **175**:27–36.

105. **Eisenstein BI**, **Sweet DS**, **Vaughn V**, **Friedman DI**. 1987. Integration host factor is required for the DNA inversion that controls phase variation in *Escherichia coli*. Proc Natl Acad Sci U S A **84**:6506–6510.

106. **Serrano M**, **Kint N**, **Pereira FC**, **Saujet L**, **Boudry P**, **Dupuy B**, **Henriques AO**, **Martin-Verstraete I**. 2016. A recombination directionality factor controls the cell type-specific activation of σK and the fidelity of spore development in *Clostridium difficile*. PLoS Genet **12**:e1006312.

107. **Barr JJ**, **Auro R**, **Furlan M**, **Whiteson KL**, **Erb ML**, **Pogliano J**, **Stotland A**, **Wolkowicz R**, **Cutting AS**, **Doran KS**, **Salamon P**, **Youle M**, **Rohwer F**. 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc Natl Acad Sci USA **110**:10771–10776.

108. **Batah J**, **Denève-Larrazet C**, **Jolivot P-A**, **Kuehne S**, **Collignon A**, **Marvaud J-C**, **Kansau I**. 2016. *Clostridium difficile* flagella predominantly activate TLR5-linked NF-κB pathway in epithelial cells. Anaerobe **38**:116–124.

109. **Andersen-Nissen E**, **Smith KD**, **Strobe KL**, **Barrett SLR**, **Cookson BT**, **Logan SM**, **Aderem A**. 2005. Evasion of Toll-like receptor 5 by flagellated bacteria. Proc Natl Acad Sci U S A **102**:9247–9252.

110. **Ghose C**, **Eugenis I**, **Sun X**, **Edwards AN**, **McBride SM**, **Pride DT**, **Kelly CP**, **Ho DD**. 2016. Immunogenicity and protective efficacy of recombinant *Clostridium difficile* flagellar protein FliC. Emerg Microbes Infect **5**:e8.

111. **Ng J**, **Hirota SA**, **Gross O**, **Li Y**, **Ulke-Lemee A**, **Potentier MS**, **Schenck LP**, **Vilaysane A**, **Seamone ME**, **Feng H**, **Armstrong GD**, **Tschopp J**, **MacDonald JA**, **Muruve DA**, **Beck PL**. 2010. *Clostridium difficile* toxin-induced inflammation and intestinal injury are mediated by the inflammasome. Gastroenterology **139**:542–52–552.e1–3.

112. **Bouillaut L**, **McBride SM**, **Sorg JA**. 2011. Genetic Manipulation of *Clostridium difficile*. Curr Protoc Microbiol **20**:9A2.1-9A2.17.

113. **Laemmli UK**. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680–685.

114. **Schindelin J**, **Arganda-Carreras I**, **Frise E**, **Kaynig V**, **Longair M**, **Pietzsch T**, **Preibisch S**, **Rueden C**, **Saalfeld S**, **Schmid B**, **Tinevez J-Y**, **White DJ**, **Hartenstein V**, **Eliceiri K**, **Tomancak P**, **Cardona A**. 2012. Fiji: an open-source platform for biological-image analysis. Nat Meth **9**:676–682.

115. **Knetsch CW**, **Terveer EM**, **Lauber C**, **Gorbalenya AE**, **Harmanus C**, **Kuijper EJ**, **Corver J**, **van Leeuwen HC**. 2012. Comparative analysis of an expanded *Clostridium difficile* reference strain collection reveals genetic diversity and evolution through six lineages. Infect Genet Evol **12**:1577–1585.

116. **van Eijk E**, **Anvar SY**, **Browne HP**, **Leung WY**, **Frank J**, **Schmitz AM**, **Roberts AP**, **Smits WK**. 2015. Complete genome sequence of the *Clostridium difficile* laboratory strain 630Δ*erm* reveals differences from strain 630, including translocation of the mobile element CTn5. BMC Genomics **16**:31.

117. **McBride SM**, **Sonenshein AL**. 2011. Identification of a genetic locus responsible for antimicrobial peptide resistance in *Clostridium difficile*. Infect Immun **79**:167–176.

118. **Manganelli R**, **Provvedi R**, **Berneri C**, **Oggioni MR**, **Pozzi G**. 1998. Insertion vectors for construction of recombinant conjugative transposons in *Bacillus subtilis* and *Enterococcus faecalis*. FEMS Microbiol Lett **168**:259–268.

119. **Bouillaut L**, **Self WT**, **Sonenshein AL**. 2013. Proline-dependent regulation of *Clostridium difficile* Stickland metabolism. J Bacteriol **195**:844–854.

120. **Hanahan D**. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.

121. **Christie PJ**, **Korman RZ**, **Zahler SA**, **Adsit JC**, **Dunny GM**. 1987. Two conjugation systems associated with *Streptococcus faecalis* plasmid pCF10: identification of a conjugative transposon that transfers between *S. faecalis* and *Bacillus subtilis*. J Bacteriol **169**:2529–2536.

122. **Browne HP**, **Anvar SY**, **Frank J**, **Lawley TD**, **Roberts AP**, **Smits WK**. 2015. Complete genome sequence of BS49 and draft genome sequence of BS34A, Bacillus subtilis strains carrying Tn916. FEMS Microbiol Lett **362**:1–4.

123. **Hussain HA**, **Roberts AP**, **Mullany P**. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δ*erm*) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. J Med Microbiol **54**:137–141.

CHAPTER 3: CHARACTERIZATION OF FLAGELLAR AND TOXIN PHASE

VARIATION IN *CLOSTRIDIUM DIFFICILE* RIBOTYPE 012 ISOLATES[2]

Summary

Clostridium difficile produces flagella that enhance bacterial motility, and secretes toxins that promote diarrheal disease symptoms. Previously, we found that production of flagella and toxins are co-regulated via a flippable DNA element termed the "flagellar switch", which mediates the phase variable production of these factors. Here we evaluate multiple isolates of *C. difficile* ribotype 012 strains and find them to be primarily *flg* OFF. Some, but not all, of these isolates showed the ability to switch between *flg* ON and OFF states. These findings suggest heterogeneity in the ability of *C. difficile* ribotype 012 strains to phase vary flagellum and toxin production, which may broadly apply to pathogenic *C. difficile*.

---

[2] This chapter contains a manuscript reviewed at *Journal of Bacteriology*, for which I am addressing reviewer criticism with new experiments, and is available on the preprint server *bioRxiv*. Natalia Maldonado-Vazquez, the co-first author, performed the experiments for Figures 3.1-3.5 under my supervision. I isolated the motile derivatives of *C. difficile* JIR8094 (MD #1-3) used throughout these figures, and I performed the western blot for TcdA in Figure 3.5B. I performed all the experiments for Figures 3.6-3.11. Natalia and Dr. Tamayo wrote an initial draft of the Introduction, Materials & Methods, and part of the Results. I rewrote the aforementioned sections to incorporate results generated from Figures 3.6-3.11 and wrote the Discussion. Dr. Tamayo and I edited the version submitted for peer review.

Introduction

The obligate anaerobe *Clostridium difficile* is a leading cause of nosocomial intestinal infections. *C. difficile* associated infections (CDI) are most common in individuals who have undergone antibiotic therapy, which disrupts the usually protective microbiota and creates a niche for *C. difficile* outgrowth (1). Virulence of *C. difficile* is largely mediated by two glucosylating toxins, TcdA and TcdB, which target and inactivate Rho and Rac GTPases in the intestinal epithelium, resulting in depolymerization of the actin cytoskeleton and eventually host cell death (2). Toxin-mediated damage to the epithelium leads to disruption of the intestinal barrier, diarrheal symptoms, and a robust inflammatory response (2).

C. difficile produces peritrichous flagella that are essential for swimming motility and contribute to host cell adherence (3-6). As in other bacterial species, the expression of flagellar genes occurs in a hierarchical manner (7). In *C. difficile,* at least four different operons encode flagellar genes. The early stage flagellar genes (*flgB* operon) are transcribed first, and they encode the basal body, motor, and rod of the flagella. The *flgB* operon also encodes the alternative sigma factor SigD (σ^D, also known as FliA or σ^{28}). In addition to activating late stage flagellar gene expression, SigD also positively regulates the expression of *tcdR*, which encodes a sigma factor that activates transcription of *tcdA* and *tcdB* (7-10). Therefore, the regulation of flagellar genes also impacts virulence by affecting toxin gene expression.

The expression of flagellum and toxin genes is subject to complex regulation (6, 7, 11-13). Recently, we demonstrated that flagellum and toxin biosynthesis is phase variable via site-specific recombination that inverts a DNA element termed the "flagellar switch" (14). The flagellar switch consists of a 154 bp invertible DNA sequence flanked by 21 bp inverted repeats (left inverted repeat (LIR) and right inverted repeat (RIR)) and lies upstream of the *flgB* operon.

The orientation of the flagellar switch controls expression of the *flgB* operon, and therefore *sigD* and the toxin genes, through an unidentified mechanism occurring post-transcription initiation (14). Bacteria with the flagellar switch in an orientation resulting in flagellum production, swimming motility, and high toxin production were termed flagellar phase ON (*flg* ON). In contrast, bacteria with the flagellar switch in the opposite orientation resulting in the absence of flagella, sessility, and reduced toxin production were termed flagellar phase OFF (*flg* OFF). RecV, a site-specific tyrosine recombinase previously shown to mediate inversion of the *cwpV* switch (15, 16), also mediates inversion of the flagellar switch (14). The phase variable production of flagella and toxins was proposed to allow *C. difficile* to balance the benefits of swimming motility and toxinogenesis with the cost of producing these immunogenic factors (17-20).

Flagellum and toxin phase variation in *Clostridium difficile* appears to vary across ribotypes (14). For *C. difficile* strains R20291, a ribotype 027 strain, and ATCC 43598, a ribotype 017 strain, both the *flg* ON and OFF orientations were apparent under multiple conditions tested. However, only the *flg* ON orientation was detectable in 630, a ribotype 012 strain originally isolated from a patient with *C. difficile* infection (21). However, recent work from Collery *et al.* suggests the flagellar switch is capable of inversion in 630 (22), suggesting this strain is not phase-locked. JIR8094 (also named 630E) was isolated through serial passage of 630 to obtain erythromycin-sensitive isolates amenable to genetic manipulation with tools relying on an erythromycin resistance cassette (23). Comparison of the genome sequence of JIR8094 to 630 revealed numerous secondary mutations that arose during and since isolation (22). One polymorphism identified was an inversion of the flagellar switch to the OFF orientation in JIR8094 compared to 630, which is *flg* ON. The JIR8094 genotype explains the

non-motile phenotype and reduced flagellum and toxin gene expression previously reported for JIR8094 (7), and indicates an ability to invert the flagellar switch in this strain lineage. In contrast, the 630Δ*erm* strain was similarly derived from 630 through serial passaging, but it remains motile and toxinogenic (24).

In this study, we examined multiple ribotype 012 strains, including laboratory-adapted, clinical, and environmental isolates, to assess their ability to invert the flagellar switch and to phase vary flagellum and toxin production. Analysis of these strains and their motile derivatives indicates that flagellum and toxin phase variation is conserved in ribotype 012 *C. difficile*, and that there is considerable variation in the frequency of phase variation in this and perhaps other ribotypes.

Results

The laboratory-adapted *C. difficile* strain JIR8094 is *flg* OFF

The *C. difficile* strain JIR8094 is a non-motile, toxin-attenuated derivative of the ribotype 012 strain 630 (7, 22, 23). In *C. difficile*, expression of the *flgB* operon is controlled by multiple *cis*-acting mechanisms, including a σ^A-dependent promoter (31), a c-di-GMP riboswitch in the 5'UTR (27, 31, 32), and the flagellar switch (14). Sequencing of the *flgB* operon regulatory region from *C. difficile* 630 and JIR8094 revealed no differences in the promoter region or the c-di-GMP riboswitch sequences. However, while the LIR and RIR were identical in the two strains, the intervening 154 bp was inverted in JIR8094 relative to 630 (Figure 3.1A). Alignment of the flagellar switch sequence using the reverse-complement of the sequence obtained for JIR8094 restored sequence identity to 100% with 630 (Figure 3.1A). Inversion of the flagellar

switch in JIR8094 relative to 630 was confirmed by orientation-specific PCR (Figure 3.1B). For this assay, a common reverse primer (Rv) is used in combination with a primer (Fw1) that leads to amplification when the template is in the ON orientation, or with a primer (Fw2) that amplifies when the template is in the OFF orientation (Figure 3.1B) (14). As shown previously, *C. difficile* R20291 cultures contain a mixture of *flg* ON and OFF bacteria (Figure 3.1C) (14). In contrast, only the *flg* ON orientation was detected for strain 630, and only the *flg* OFF orientation was detected for JIR8094 (Figure 3.1C). In addition, strain 630 demonstrated swimming through motility medium, while JIR8094 was non-motile after 72 hours (Figure 3.1D). These observations are consistent with a recent report identifying the sequence inversion in JIR8094 (22), and with phenotypes corresponding to the *flg* OFF state (7).

The flagellar switch undergoes inversion at low frequency in JIR8094

Throughout our experimentation, we did not observe motility or motile flares from JIR8094 in standard assays (data not shown), and Collery et al. similarly reported an inability to recover motile bacteria from this strain (22). These findings suggested that JIR8094 is phase-locked OFF or undergoes inversion of the flagellar switch at a low rate. To distinguish between these possibilities, we modified the motility assay to accommodate higher density inoculums. Rather than inoculating motility agar with 2 µL of saturated broth culture, we first concentrated the culture 30-fold, reasoning that this would increase the odds of detecting rare *flg* ON bacteria. Over 48-72 hours, we were able to detect motility from a subset of JIR8094 inoculums (Figure 3.2A).

To determine whether motility was restored as a result of flagellar switch inversion to ON or due to extragenic suppressor mutations, we used orientation-specific PCR (Figure 3.1B) to

assess the orientation of the flagellar switch. Three motile derivatives of JIR8094 (MD #1-3) were arbitrarily chosen after recovery from the edge of the motile colony for further analysis. The term "motile derivatives" was chosen over "*flg* ON" for JIR8094 derivatives since, as described later in the text, these strains remain attenuated for *flg* ON phenotypes compared to strain 630. All three MDs yielded PCR products indicative of the *flg* ON orientation, but not the *flg* OFF orientation, as seen in the parental strain 630 (Figure 3.2B). Sequencing of the flagellar switch confirmed the *flg* ON orientation of the switch in JIR8094 MD #1-3 (Figure 3.2C). These results indicate that the flagellar switch remains reversible in JIR8094, though the frequency of inversion appears to be lower than previously observed for R20291 (14).

To estimate the frequency of *flg* ON bacteria in JIR8094, a range of bacterial densities were assayed for swimming motility. As controls, we used the motile 630 strain, the non-motile 630 Δ*erm sigD* mutant (33), and an isolate of *C. difficile* R20291 *flg* OFF which contains ~3 % *flg* ON bacteria (14). This small population of *flg* ON bacteria spatially segregates from non-motile *flg* OFF bacteria in motility medium, leading to a motile phenotype that is readily apparent within 24 hours (14). Whereas R20291 appeared motile at all inoculums tested, motility of JIR8094 was only observed at the highest density, suggesting that flagellar switch inversion occurs at a low frequency in this strain (Figure 3.3).

Characterization of JIR8094 motile derivatives

We predicted that restoration of the flagellar switch to the ON orientation would restore flagellar motility and toxin production of the JIR8094 MDs to levels observed in the original 630 parental strain. To test this, we assayed swimming motility using standard (not concentrated) inoculums, with *C. difficile* 630 and JIR8094 strains included as positive and negative controls,

135

respectively. The JIR8094 MD #1-3 were motile at all time points tested, unlike JIR8094 (Figure 3.4A), but motility was not fully restored to the level seen for 630 (Figure 3.4B). We also observed that JIR8094 MD #1-3 produced fewer flagella than motile and peritrichous 630, while JIR8094 remained aflagellate (Figure 3.4C).

The intermediate motility of JIR8094 MD #1-3 could be due to either partial recovery of flagellar gene transcription and/or of flagellum biosynthesis, both of which could be due to additional mutations in JIR8094. Using quantitative reverse transcriptase PCR (qRT-PCR), we measured the transcript abundance for three flagellar genes, *flgB*, *sigD,* and *fliC.* Both *flgB* and *sigD* are in the operon directly controlled by the flagellar switch (early stage flagellar genes), and *fliC* serves as a representative *sigD*-regulated late stage flagellar gene. As previously reported, the *flgB*, *sigD*, and *fliC* transcripts were significantly lower in JIR8094 compared to 630 (Figure 3.4D) (7). In MD #1-3, *flgB*, *sigD,* and *fliC* transcript levels were higher than in JIR8094, but were not restored to the levels in 630 (Figure 3.4D). Nonetheless, the intermediate level of flagellar gene expression and number of flagella on the JIR8094 MDs were sufficient to confer some motility.

Similar intermediate phenotypes were observed with regards to toxin production. Analysis of *tcdR*, *tcdA*, and *tcdB* transcripts in stationary phase cultures (favoring toxin gene expression) by qRT-PCR showed that all three transcripts were less abundant in JIR8094 than in 630, as seen previously (7), although only *tcdA* was significantly lower (Figure 3.5A). The transcripts trended higher in MD #1-3 than in their JIR8094 parent, but the differences were not statistically significant (Figure 3.5A). Consistent with this, the amounts of TcdA protein produced by MD #1-3 appeared more similar to the JIR8094 parent than 630 (Figure 3.5B). Together, these data indicate that JIR8094 MD #1-3 have restored the ability to transcribe

flagellar and toxin genes, and thus motility and toxinogenesis, but suggest that other mutations in JIR8094 also affect flagellum and toxin gene expression in this strain.

Topoisomerase activity in JIR8094 motile derivatives impacts motility

Comparison of the genome sequences of *C. difficile* 630 and its 630Δ*erm* and JIR8094 derivatives revealed multiple single nucleotide polymorphisms, including 11 found in JIR8094 but not the other two strains (one of which is the flagellar switch inversion) (22). We speculated that one or more of these polymorphisms is responsible for the difference in flagellar gene expression between 630 and the JIR8094 MD #1-3. Of interest was a nonsense mutation in a *topA* orthologue that introduced a stop codon at amino acid 386, while the full-length protein is 695 amino acids. Because *topA* encodes DNA topoisomerase I, which controls DNA negative supercoiling, we postulated that the mutation in *topA* could result in changes to the *flgB* operon promoter and/or coding sequence that impair gene transcription in JIR8094 and its derivatives, compared to 630, which encodes the full length TopA. If so, expressing *topA* in the JIR8094 MD #1-3 is expected to restore motility and toxin production to the levels in 630.

To test this, we ectopically expressed *topA* from an ATc-inducible promoter in 630, JIR8094, and the MD #1-3 and compared the motility of these strains to vector controls. A 630Δ*erm sigD* mutant bearing vector or the *topA* expression plasmid (pTopA) was included as a non-motile control. After 72 hours, expression of *topA* did not significantly alter the motility of 630 or the *sigD* mutant as expected (Figure 3.6A and 3.6B). Expression of *topA* alone is not sufficient to augment motility in *flg* OFF bacteria, as JIR8094 + pTopA remained non-motile (Figure 3.6A and 3.6B). However, expression of *topA* in JIR8094 MD #1-3 significantly enhanced motility, restoring it to 630 levels (Figures 3.6A and 3.6B). Presumably TcdA levels would concomitantly

increase upon restoration of TopA activity, but western blot analysis indicated TcdA levels were unchanged in the JIR8094 MD #1-3 + pTopA derivatives compared to parent JIR8094 + pTopA (Figure 3.6C). These results suggest that the flagellar switch must be in the ON orientation to allow flagellar gene expression and motility, and DNA supercoiling affects flagellar gene expression. Moreover, additional mutations in JIR8094 may interfere with the link between toxin and flagellar gene expression.

Clinical and environmental Isolates of *C. difficile* ribotype 012 are *flg* OFF

Both flagellar switch orientations were readily detected in *C. difficile* strains R20291, a ribotype 027, and ATCC 43598, a ribotype 017, but only the *flg* ON was detected in 630 (14). The 630 strain was isolated in 1982 (21), so passaging within and between labs could have led to accumulation of mutations that reduced the frequency of flagellar phase variation in 630. We thus asked whether the lack of (or very low frequency of) flagellar switch inversion in 630 is unique to this strain and its derivatives, or a broader attribute of currently circulating ribotype 012 *C. difficile* isolates. To test this, we evaluated flagellar phase variation in six riboype 012 isolates: SS235, a community environmental isolate; SE838, a hospital environmental isolate; and MAM30, MT4768, MT5065, and MT5066, isolates from patients with CDI (kindly gifted by Dr. Kevin Garey). Sequencing of the flagellar switches revealed that all six isolates contain the flagellar switch in the same OFF orientation as JIR8094, and opposite that of strain 630 (Figure 3.7) (22). The flagellar switches of two isolates also contained polymorphisms within the inverted repeats (Figure 3.7). MT4768 had one nucleotide (nt) substitution in the LIR and an additional nucleotide in the RIR. SE838 contained an additional 3 nt in the LIR and 2 nt in the RIR. In addition, MT4768 possessed three single nucleotide substitutions within the flagellar

switch (Figure 3.7). SS235, MAM30, MT5065, and MT5066 had flagellar switches identical to that in JIR8094.

Sequencing results indicate the conservation of the regulatory flagellar switch element, but they did not reveal if it is functional. It is also possible that the sequencing results reflect the most abundant switch orientation in a population, masking the presence of a lower abundance orientation. Therefore, we used PCR with orientation-specific primer sets that discriminate between the two DNA orientations (Figure 3.8A) (14). As a control we included 630, for which only the *flg* ON orientation is detectable (Figure 3.8A) (14). For all six isolates only the *flg* OFF orientation of the flagellar switch was detectable (Figure 3.8A).

Given that *flg* OFF bacteria in R20291 are phenotypically aflagellate and non-motile (14), we predicted that the six isolates with the flagellar switch in the OFF orientation would be similarly non-motile. The six riboype 012 isolates were examined for the ability to swim through motility medium. Strain 630 and a 630Δ*erm sigD* mutant were used as motile and non-motile controls, respectively. After 24 hours incubation, all six ribotype 012 isolates showed a non-motile phenotype in comparison to the motile 630, which migrated outward from the original inoculation site (Figure 3.8B). Collectively, these data indicate that the flagellar switch is conserved in ribotype 012 clinical and environmental isolates and appear to be "phase-locked" in an *flg* OFF state or undergo flagellar switch inversion at a low frequency *in vitro*.

Isolation and phenotypic characterization of *flg* ON derivatives of clinical and environmental ribotype 012 isolates

While examining the motility of the ribotype 012 isolates, we observed that a subset of replicates showed motile flares extending from the stab sites for some isolates (Figure 3.9A).

This is in contrast with R20291, for which we previously found that clonal *flg* OFF populations yielded a motile phenotype within 24 hours in 100% of replicates (14). To determine the frequency at which motile flares arise in the ribotype 012 isolates, we increased the number of biological replicates (n = 32 over 2 experiments) and monitored motility every 24 hours for 3 days. A small proportion of colonies of three of the isolates developed motile flares by 24 hours (MT5066, 6.3%; SE838, 9.4%; SS235, 3.1%), and a fourth isolate showed motile flares by 48 hours (MT5065, 18.9 %) (Figure 3.9B). The number of colonies with motile flares for these four isolates increased and then remained relatively stable between 48 and 72 hours (48 hours: MT5066, 71.9%; SE838, 100%; SS235, 59.4%; MT5065, 18.8%) (Figure 3.9B). The remaining two isolates, MAM30 and MT5065, failed to develop motile flares throughout the course of the experiments.

To confirm that the bacteria from the edge of the flares regained motility through inversion of the flagellar switch to the ON orientation, we arbitrarily chose and isolated bacteria from three motile flares each for MT5066, SE838, and SS235 and subjected them to orientation-specific PCR and sequenced the flagellar switch (Figure 3.1B). All nine motile isolates yielded a product for the *flg* ON orientation, indicating that DNA inversion occurred in these strains, yet a *flg* OFF product was also observed (Figure 3.10) (Figure 3.9C). This result is in contrast to the JIR8094 MD #1-3, for which only the *flg* ON orientation was detected (Figure 3.2B). In almost all cases, the intensity of the *flg* ON band was greater than the *flg* OFF, suggesting the majority of the population was *flg* ON. Sequencing results confirmed flagellar switch inversion to the ON orientation for all nine *flg* ON derivatives and 100% sequence identity with the strain 630 flagellar switch. We noted that in all three *flg* ON derivatives of parent SE828, the RIR length increased from 22 bp to 23 bp (Figure 3.10).

To determine whether flagellar switch inversion to the ON orientation was sufficient to restore *flg* ON phenotypes in these strains, we assayed the original ribotype 012 isolates and their *flg* ON derivatives for swimming motility and toxin production (14). As expected, all three *flg* ON derivatives of each strain showed motility compared to their respective non-motile parent strains (MT5066, SE838, and SS235) (Figure 3.11A). TcdA protein levels were visibly higher in cell lysates of the *flg* ON derivatives relative to the *flg* ON MT5066, SE838, and SS235 strains (Figure 3.11B). These results demonstrate that the flagellar switch is capable of inversion in some but not all clinical and environmental ribotype 012 isolates, and reiterate the regulatory link between toxin and flagellar gene expression via SigD. In addition, the ribotype 012 strains exhibit heterogeneity in their ability to phase vary, which may be attributed to differences in the sequences of the inverted repeats flanking the flagellar switch.

Discussion

Our prior study characterizing flagellum and toxin phase variation in *C. difficile* focused on the epidemic-associated 027 isolate R20291. In this strain, phase variation occurred readily *in vitro*, and the strain consisted of a mixture of *flg* ON and OFF bacteria. However, we noted that the ribotype 012 strain 630, which was originally isolated from a patient with *C. difficile* infection, existed primarily as *flg* ON bacteria. Here, we asked whether the 630 lineage is unusual, or whether reduced phase variation is a broader attribute of ribotype 012 strains. Strain 630's non-motile JIR8094 derivative primarily exists as *flg* OFF bacteria, and phase variation to *flg* ON occurs with low frequency. Our survey of six clinical and environmental ribotype 012 isolates of *C. difficile* indicates that within the 012 ribotype there exists a range of abilities to phase vary flagellum and toxin production, at least from the *flg* OFF to ON states. This work

141

represents the first examination of flagellar and toxin phase variation within a ribotype family, and our results indicate potentially diverse patterns of phase variation across *C. difficile* strains.

We found that the laboratory-adapted strain JIR8094 contains the flagellar switch in the OFF orientation, which is consistent with its non-motile and toxin-attenuated phenotypes. We were able to isolate rare motile derivatives of JIR8094. These were determined to have inverted the flagellar switch to the ON orientation, suggesting the recovery of motility was attributable to flagellar switch inversion. However, phenotypic characterization of these motile derivatives of JIR8094 indicated that motility and toxin production were only partially recovered compared to the original 630 strain (JIR8094 parent). However, JIR8094 contains 11 polymorphisms compared to 630 (22), and we found that restoring expression of one of the affected genes, *topA*, was sufficient to fully restore motility. This finding indicates that both loss of *topA* and inversion of the flagellar switch from ON to OFF negatively affected flagellar gene expression in JIR8094, and underscores previously stated cautions in using and interpreting results from research done with JIR8094. For example, one study using strain JIR8094 indicated that TcdA is dispensable for virulence in the hamster model (34), while another using 630*Δerm* showed that both TcdA and TcdB are sufficient for eliciting disease in hamsters (35). The initial studies suggesting TcdA was dispensable for infection may have been due to the *flg* OFF status of JIR8094. A *tcdA* mutation in the background of a motile derivative of JIR8094 could be attenuated for virulence in the hamster model supporting a role for TcdA in diarrheal disease development.

The conditions that favor inversion from *flg* ON to OFF in 630 may be gleaned from how JIR8094 was generated, specifically through serial passage on solid medium (23). In our studies with R20291, we observed a bias toward the *flg* OFF orientation during growth on nutrient rich solid medium over several days, while switch orientation appeared stable during growth in the

same liquid medium (14). In contrast to JIR8094, strain 630Δerm was generated by passaging

630 in non-selective liquid medium until a spontaneous erythromycin sensitive isolate was

identified, and the flagellar switch remained in the *flg* ON orientation, unlike 630E (24). The

ability to bias a phase variable genetic switch in one orientation over another is common in other

mucosal bacterial pathogens. For example, incubation in human urine biases the Type I fimbrial

switch to the OFF orientation in uropathogenic *Escherichia coli* (36). Anoxic conditions bias the

fimbrial switch to the ON orientation in both uropathogenic *Proteus mirabilis* and *E. coli* (37).

For *C. difficile*, factors that bias the flagellar switch orientation could be host factors sensed by

gene products that influence recombination.

All six of the original ribotype 012 isolates, both clinical and environmental, existed

primarily, and perhaps exclusively, in a *flg* OFF state, both genotypically and phenotypically.

Two strains, MAM30 and MT4768, appeared to be phase-locked, as we were unable to recover

flg ON bacteria upon growth in motility medium, which selects for outgrowth of motile bacteria.

The remaining four strains varied in their propensity to switch to *flg* ON. One of these, SE838,

readily switched from *flg* OFF to ON, similar to R20291. A limitation of our study is that it is

uncertain whether the isolates were *flg* OFF at the time of collection, or whether the strains

switched during outgrowth in the laboratory. However, the fact that all six isolates from different

patients and environments were *flg* OFF argues that our findings are not an artifact of *in vitro*

work, but a common feature in the *C. difficile* ribotype 012 family. Recent work found that in a

patient simultaneously infected with three *C. difficile* isolates, the ribotype isolate 012 was *flg*

ON, providing evidence that not all ribotype 012 isolates are biased to *flg* OFF, or the isolate

may have arisen from a stage of infection favoring *flg* ON bacteria (38). It is also possible that

most ribotypes show a strong *flg* OFF bias and lower inversion frequency, similar to the ribotype

012, and the ribotype 027 and 017 strains examined previously are outliers (14). A comprehensive analysis of strains representing a variety of ribotypes both *in vitro* and during infection will help determine the population dynamics of flagellum and toxin phase variation in *C. difficile*.

Inverted repeat length may affect recombination of the flagellar switch in the 012 ribotype strains. Our previous analysis of the flagellar switch with publically available sequenced genomes of *C. difficile* indicated 21 bp inverted repeats in all strains. Interestingly, all of the clinical and environmental 012 ribotype isolates had both inverted repeats less than or greater than, but not equal to 21 bp (Figure 3.3). In addition, the SE838 isolate had flagellar switch IRs of 23 and 22 bp, but its *flg* ON derivatives had a RIR of 23 bp, suggesting some flexibility in the ability of the recombinase RecV to interact with these inverted repeats. It is also possible that a recombination directionality factor (RDF) is required for flagellar switch inversion in *C. difficile* strains that harbor non-ideal inverted repeats, and that the RDF is only produced under certain environmental conditions (39). The involvement of an RDF may allow precise control over when to promote recombination at the flagellar switch, similar to the contribution of an RDF that contributes to the excision of a DNA element and promotes sporulation (40). Alternatively, changes to repeat length could reduce affinity for RecV, and perhaps allow recombination by another recombinase. Current work seeks to identify a role for inverted repeat sequence and length in affecting flagellar switch inversion frequency in both directions in all ribotypes.

Flagellar and toxin phase variation in *C. difficile* creates population heterogeneity that has the potential to influence diagnosis, disease progression, and transmission. Characterization of flagellar and toxin phase variation in multiple ribotypes of *C. difficile* could provide insight into factors that influence flagellar switch inversion and find correlations between diarrheal disease

symptoms and switch orientation. Identifying surface and exported proteins in *C. difficile* that are consistently produced and not subject to phase variable expression could serve as better diagnostic and therapeutic targets to ensure targeting of the entire population.

Materials and Methods

Bacterial strains and growth conditions

All bacterial strains used in this study are listed in Table 3.1. *C. difficile* strains were cultivated statically at 37°C in brain heart infusion medium (BD, Beckton Dickinson) supplemented with 5% yeast extract (BHIS) or tryptone yeast (TY) medium, as specified. All *C. difficile* growth was done anaerobically in a Coy anaerobic chamber with an atmosphere of (90% N_2, 5% CO_2, 5% H_2). Unless otherwise indicated, *E. coli* was cultured at 37°C under aerobic conditions in Luria Bertani (LB) medium. For selection of plasmids in *E. coli*, 100 µg/mL ampicillin (Amp) and/or 10 µg/mL chloramphenicol (Cm) was used, as indicated. Kanamycin (Kan) 100 µg/ml was used to select against *E. coli* in conjugations. For maintenance of plasmids in *C. difficile*, 10 µg/mL thiamphenicol (Tm) was used.

Characterization of *C. difficile* ribotype 012 clinical and environmental isolates

Clinical and environmental isolates of *C. difficile* used are part of the F ribotype database (25). The strains were ribotyped by fluorescent PCR and sequencing of ribosomal RNA genes as previously described (25). The flagellar switch orientation was determined by PCR amplification of the region and sequencing of the products using primers R591 and R857.

Orientation-specific PCR

The flagellar switch orientation was determined by PCR using two primer sets that distinguish the orientation of the switch (14, 26). Primers were designed using the published genomes sequences for strains 630 (Accession No. FN688375.1) and R20291 (Accession No. FN545816.1). Primer sequences are listed in Table 3.2. For the 630 lineage and other ribotype 012 isolates, R1751 and R857 were used to detect the ON orientation, and R1752 and R857 were used to detect the OFF. For R20291, R1614 and R857 were used to detect the ON orientation, and R1615 and R857 were used to detect the OFF. PCR products were separated on a 2% agarose gel and stained with ethidium bromide (EtBr) for imaging using the G:Box Chemi imaging system.

Motility assays

Swimming motility was assayed as described previously (14, 27). Briefly, individual colonies of the indicated strains were grown in TY broth overnight, then diluted 1:50 in fresh BHIS and grown to an $OD_{600} \sim 1.0$. In standard assays, 2 µL of culture were inoculated into anaerobic motility medium (0.5X BHIS-0.3% agar, autoclaved for 30 minutes). In experiments aimed at identifying JIR8094 motile revertants, a range of volumes of overnight bacterial cultures were centrifuged to collect and concentrate bacterial cells. The cells were suspended in 100 µL of phosphate buffered saline (PBS), and 2 µL were stabbed into soft agar plates. Swimming motility was assayed by measuring the swim diameter after 24, 48, and 72 hours incubation as indicated. Two perpendicular measurements were averaged for each colony. Photographs shown were taken with the G:Box imaging system (Syngene). To measure the frequency of motile recovery in the ribotype 012 isolates, motility assays were performed as

described above with 32 independent colonies in two independent experiments. A stab site positive for motility was identified by the presence of a motile flare.

Recovery of motile derivatives of parent JIR8094 and select ribotype 012 isolates

After incubation in motility agar as described above, motile bacteria were collected from the edge of an expanded colony using a sterile loop and passaged onto BHIS agar for recovery. All isolates were confirmed to be motile by independently assaying swimming through motility medium, and the orientation of the flagellar switch was determined by PCR as described.

Microscopy

For transmission electron microscopy, colonies of *C. difficile* were grown in BHIS medium overnight, diluted into fresh BHIS medium, and grown to an OD600 of ~1.0. Bacteria were washed in Dulbecco's PBS (Sigma) prior to suspension in PBS-4 % paraformaldehyde for fixation for 1 hour in the anaerobic chamber. Cell suspensions were adsorbed onto Formvar/copper grids, washed with water, and stained for 30 seconds in two sequential drops of aqueous uranyl acetate (1.5%). Cells were viewed on a LEO EM 910 Transmission Electron Microscope (Carl Zeiss Microscopy, LLC, Thornword, NY). Documentation was done with a Gatan Orius SC1000 digital camera and Digital Micrograph 3.11.0 software (Gatan, Inc., Pleaston, CA).

Detection of TcdA production by western blot

Overnight cultures were grown in TY broth and then diluted 1:50 in fresh medium. Strains were grown until late stationary phase (OD_{600} 1.8-2.0) for 24 hours, normalized to the

same density, and then cells were collected by centrifugation. Bacteria were lysed by suspension in SDS-PAGE buffer and heating to 100ºC for 10 minutes. Samples were electrophoresed on 4-15 % TGX gels (Bio-Rad), then transferred to nitrocellulose membranes. Membranes were stained with Ponceau S (Sigma) to determine equal loading and imaged using the G:Box Chemi imaging system. TcdA was detected using mouse anti-TcdA primary antibodies (Novus Biologicals) followed by goat anti-mouse IgG secondary antibody conjugated to DyLight 800 4X PEG (Invitrogen). Blots were imaged using the Odyssey imaging system (LI-COR). All strains were assayed in at least 3 independent experiments.

RNA extraction, cDNA synthesis and quantitative reverse transcriptase–PCR

C. difficile was grown in TY medium overnight, diluted in fresh medium, and grown until mid-exponential phase (OD600 ~1.0) or stationary phase (OD600 1.8-2.0) as indicated. Cells were collected by centrifugation, and RNA was isolated as described previously (14, 27). RNA was treated with DNase I (Ambion), and cDNA was then synthesized using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher). qRT-PCR was done using SensiMix SYBR and Fluorescein Kit reagents (Bioline). The primers used were named following the convention gene-qF and gene-qR (Table 3.2). Data were normalized using the $\Delta\Delta Ct$ method, with *rpoC* serving as the reference gene. At least four independent samples were included.

Generation of *topA* expression strains

The *topA* gene was amplified from 630 genomic DNA using primers R2302 and R2303, which introduced SacI and BamHI restriction sites, respectively. After digestion with these enzymes, the *topA* fragment was ligated into similarly digested pRPF185, which allows for ATc-

inducible expression of the cloned gene (28). Ligation products were transformed into *E. coli*

DH5α. Chloramphenicol-resistant clones were collected and screened by PCR with R2302 and

R2303. Clones were confirmed by sequencing of the insert. The resulting pRPF185::*topA* was

introduced into *C. difficile* 630, JIR8094, and the motile derivatives of JIR8094 by conjugation

via HB101(pRK24) as previously described (27, 29, 30). Transconjugants were selected on agar

media with Tm and Kan and confirmed by PCR with plasmid-specific primers R1832 and

R1833. To induce expression of *topA*, anhydrotetracycline was added at a final concentration of

20 ng/mL to motility medium to assay swimming motility and broth medium for assaying TcdA

production by western blot.

A

JIR8094	AAGTTACTATTTTACAAAAA	TGACTTCTTTATTTATTTAAATTAATCGCATTTCCATTACATTAA
JIR8094-INV	AAGTTACTATTTTACAAAAA	GAGCAACTTTTCGAAGAAATATTTAAATACATATTAAAATTTTTT
630	AAGTTACTATTTTACAAAAA	GAGCAACTTTTCGAAGAAATATTTAAATACATATTAAAATTTTTT

JIR8094	TACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGAAGTCCCATGAAAATATAAA
JIR8094-INV	TATATTTTCATGGGACTTCATTAAAAAAATATTAGTTTTCTTACTAAATTGATACATTATTGTATT
630	TATATTTTCATGGGACTTCATTAAAAAAATATTAGTTTTCTTACTAAATTGATACATTATTGTATT

JIR8094	AAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTTGCTC	TTTTTGTAAAATAGCAACTT
JIR8094-INV	AATGTAATGGAAATGCGATTAATTTAAATAAATAAAGAAGTCA	TTTTTGTAAAATAGCAACTT
630	AATGTAATGGAAATGCGATTAATTTAAATAAATAAAGAAGTCA	TTTTTGTAATATAGCAACTT

Figure 3.1. *C. difficile* **JIR8094 has the flagellar switch in the OFF orientation.** (A) Multiple sequence alignment of the flagellar switch sequences from *C. difficile* 630, its JIR8094 derivative, and JIR8094-inv in which the 154 bp sequence between the imperfect inverted repeats was replaced with the reverse-complement. Blue shading indicates identical nucleotides. (B) Schematic for the orientation specific PCR assay used to determine the orientation of the flagellar switch. The Rv primer (R857) is used with either Fw1/ON (R1751) or Fw2/OFF (R1752) to detect the ON or OFF orientation of the flagellar switch, respectively. In the 630 lineage, Fw1/ON + Rv yield a 374 bp product; Fw2/OFF + Rv yield a 250 bp product. Diagram not drawn to scale. (C) Orientation-specific PCR products using template DNA from *C. difficile* R20291, 630 and JIR8094. (D) Motility of 630 and JIR8094 in motility medium (0.5X BHIS-0.3% agar) after 72 hours growth.
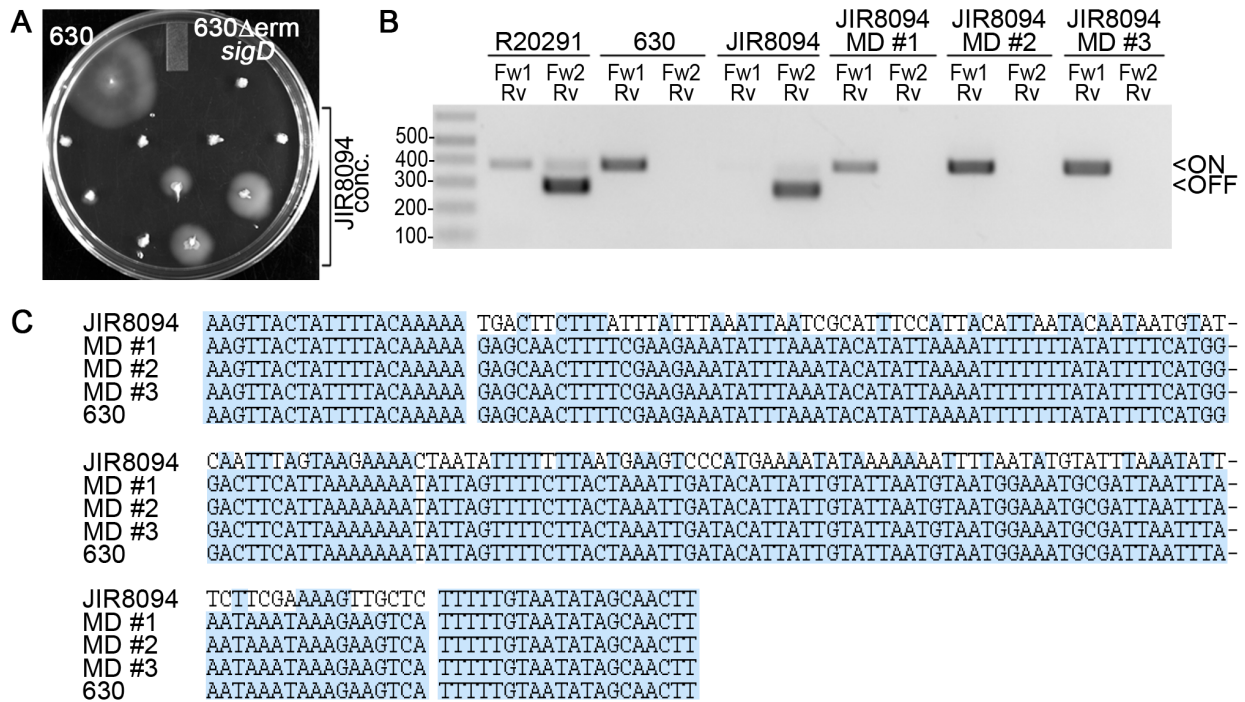
Figure 3.2. The flagellar switch undergoes inversion at a low frequency in JIR8094. (A) Concentrated liquid cultures of JIR8094 were inoculated into motility medium along with standard (i.e., not concentrated) inoculums of 630 (motile) and 630Δ*erm sigD* (non-motile) as controls. Shown is a representative experiment in which a JIR8094 culture was concentrated 30-fold (9 replicates). After 48 hours growth at 37°C, motility or motile flares was detected for three concentrated inoculums with JIR8094. Data is representative of 3 independent experiments. (B) Orientation-specific PCR products for three motile *flg* ON derivatives of JIR8094 (MD #1-3). *C. difficile* R20291, 630, and JIR8094 were included as controls for both PCR primer sets. (C) Multiple sequence alignment of the flagellar switch of JIR8094, 630, and three *flg* ON derivatives of JIR8094. Blue shading indicates identical nucleotides.
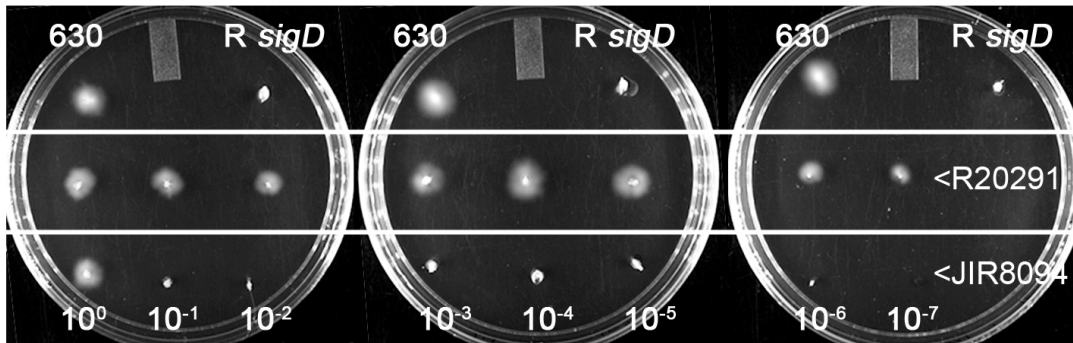
Figure 3.3. Motile derivatives of JIR8094 arise at a low frequency. JIR8094 and R20291 cultures were concentrated, and then serial dilutions were assayed for motility in motility medium, with 630 and an R20291 *sigD* mutant as motile and non-motile controls, respectively.
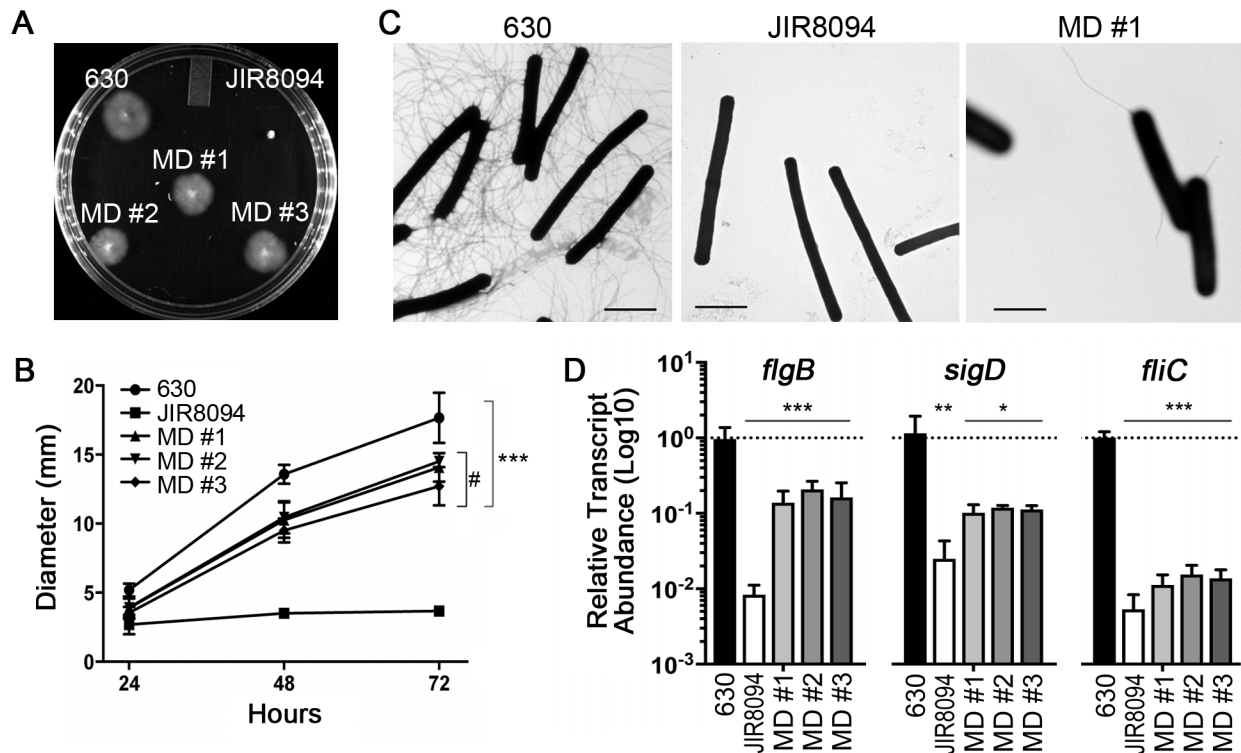
Figure 3.4. JIR8094 motile derivatives have partial recovery of flagellar gene expression, flagellum biosynthesis, and motility. (A) Motility assay for 630, JIR8094, and JIR8094 MD #1-3. Image is representative of 3 biological replicates. (B) Quantification of motility after 24, 48, and 72 hours. Shown are the means and standard deviations of the swim diameters from 3 independent experiments. At both 48 and 72 hours, *** $p < 0.001$ compared to JIR8094, and [#] $p < 0.05$ compared to 630, by one way ANOVA and Tukey's multiple comparisons test. The motilities of the JIR8094 MD #1-3 were not significantly different from each other. (C) Transmission electron microscopy (TEM) images for 630, JIR8094, and JIR8094 MD #1 showing the presence or absence of flagella. (D) The abundance of flagellar gene transcripts *flgB*, *sigD*, and *fliC* in logarithmic phase BHIS cultures of 630, JIR8094, and JIR8094 MD #1-3 isolates. The data were analyzed using the $\Delta\Delta Ct$ method with *rpoC* as the control gene and 630 as the reference strain. Shown are the means and standard deviations. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ compared to the 630 reference by one-way ANOVA and Tukey's multiple comparisons test. None of the differences between JIR8094 and MD #1-3 values met $p < 0.05$.
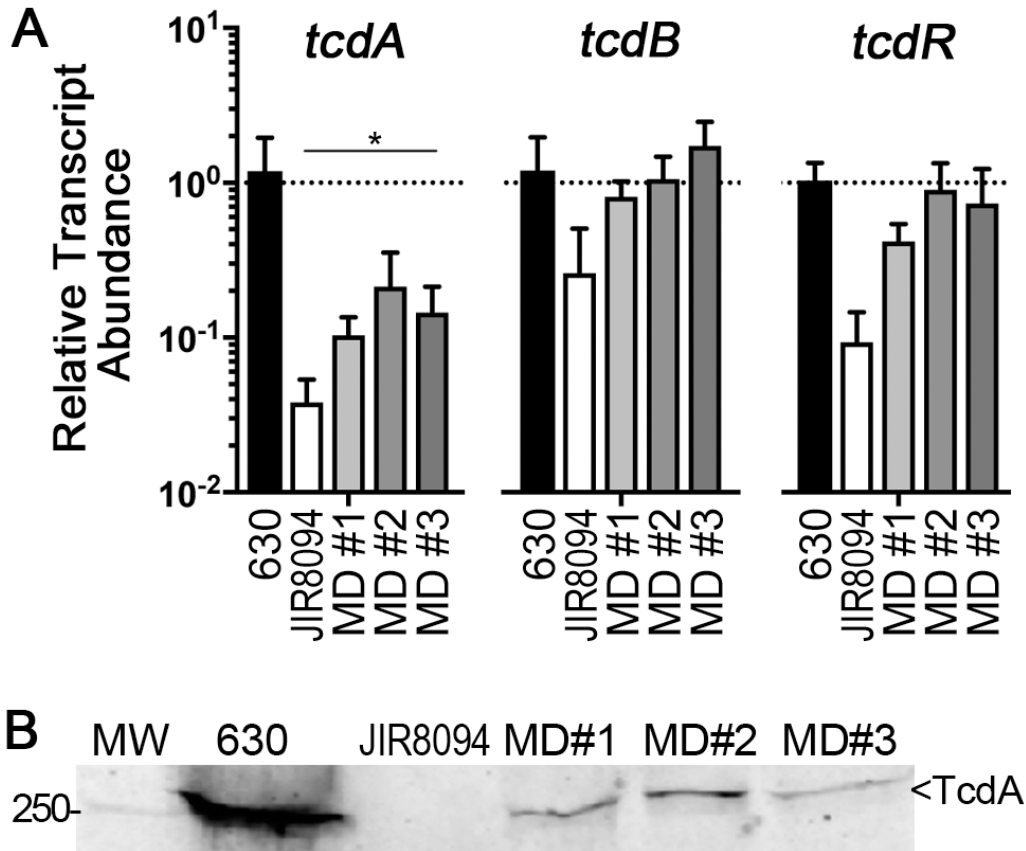
Figure 3.5. JIR8094 motile derivatives show intermediate toxin gene expression and toxin production. (A) qRT-PCR analysis of *tcdR*, *tcdA,* and *tcdB* transcript abundance in stationary phase cultures of 630, JIR8094, and JIR8094 MD #1-3. The data were analyzed using the $\Delta\Delta Ct$ method with *rpoC* as the control gene and 630 as the reference strain. Shown are the means and standard deviations. * $p < 0.05$ compared to 630 by one-way ANOVA and Tukey's multiple comparisons test. None of the differences between JIR8094 and MD #1-3 values met $p < 0.05$. (B) Western blot detection of TcdA production by 630, JIR8094, and JIR8094 MD #1-3. A representative image from 3 independent experiments is shown.
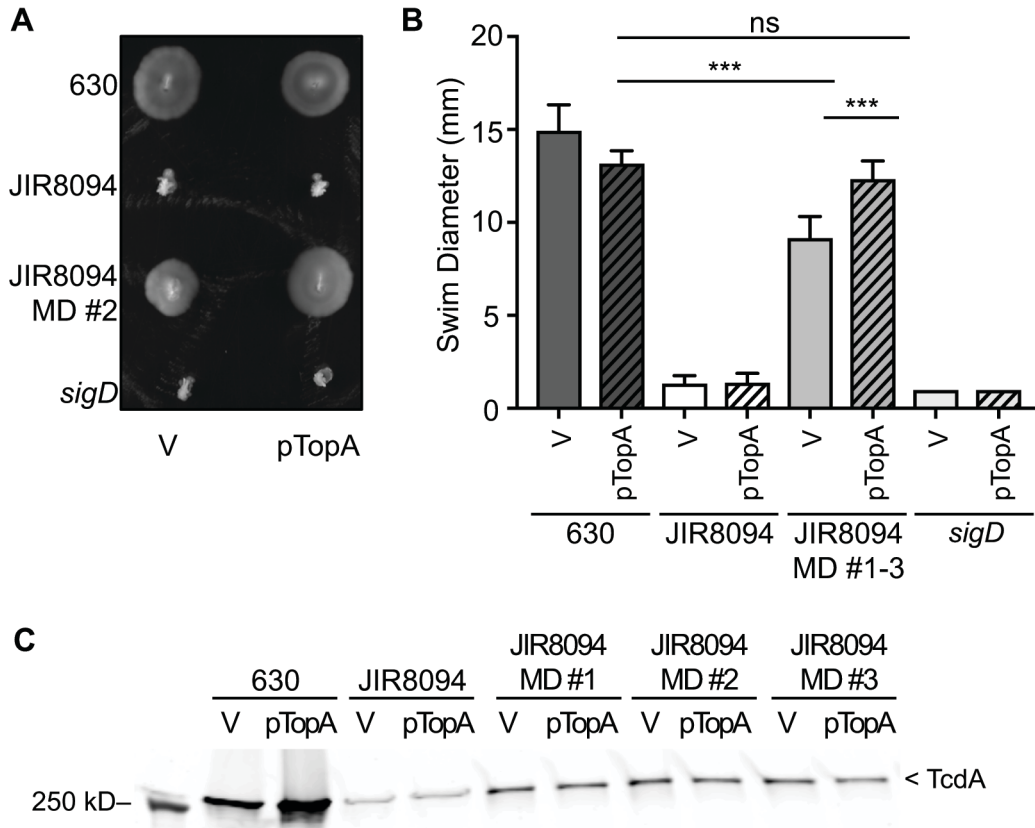
Figure 3.6. Ectopic expression of *topA* restores motility of JIR8094 motile derivatives to 630 levels. (A,B) Motility of 630, JIR8094, JIR8094 MD #2, and a *sigD* mutant, each bearing vector (V) or pTopA, after 72 hours growth. (A) Image is representative of two independent experiments with six biological replicates of each strain. Equivalent results were obtained for JIR8094 MD #1 and MD #2. (B) Quantification of motility in the above strains after 72 hours, with the data for JIR8094 MD #1-3 combined. Two independent experiments with six biological replicates were done. *** $p < 0.001$ by one-way ANOVA and Tukey's multiple comparisons test. (C) Detection of TcdA in the indicated strains by western blot. Data are representative of two independent experiments with each strain.

```
     630  AAGTTACTATTTT-ACAAA--AA GAGCAACTTTTCGAAGAAATATTTAAATACATAT  54
  JIR8094  AAGTTACTATTTT-ACAAA--AA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  54
    MAM30  AAGTTACTATTTT-ACAAA--AA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  54
   MT4768  AAGTTACTATTT--ACAAA-AA  TGACTTCTTTATTCATTTAAATTAATCGCATTTC  54
   MT5065  AAGTTACTATTTT-ACAAA--AA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  54
   MT5066  AAGTTACTATTTT-ACAAA--AA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  54
    SE838  AAGTTACTATTTTTACAAAAAAA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  57
    SS235  AAGTTACTATTTT-ACAAA--AA TGACTTCTTTATTTATTTAAATTAATCGCATTTC  54

     630  TAAAATTTTTTTATATTTTCATGGGACTTCATTAAAAAAATATTAGTTTTCTTACTAA  112
  JIR8094  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  112
    MAM30  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  112
   MT4768  CATTACATTAATACAATAATGTATCATTTTAGTAAGAAAACTAATATTTTTTCAATGA  112
   MT5065  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  112
   MT5066  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  112
    SE838  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  115
    SS235  CATTACATTAATACAATAATGTATCAATTTAGTAAGAAAACTAATATTTTTTTAATGA  112

     630  ATTGATACATTATTGTATTAATGTAATGGAAATGCGATTAATTTAAATAAATAAAGAA  170
  JIR8094  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170
    MAM30  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170
   MT4768  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170
   MT5065  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170
   MT5066  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170
    SE838  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  173
    SS235  AGTCCCATGAAAATATAAAAAAATTTTAATATGTATTTAAATATTTCTTCGAAAAGTT  170

     630  GTCA TT--TTTGTAATATAGCAACTT  194
  JIR8094  GCTC TT--TTTGTAATATAGCAACTT  194
    MAM30  GCTC TT--TTTGTAATATAGCAACTT  194
   MT4768  GCTC TTT-TTTGTAATATAGCAACTT  195
   MT5065  GCTC TT--TTTGTAATATAGCAACTT  194
   MT5066  GCTC TT--TTTGTAATATAGCAACTT  194
    SE838  GCTC TTTTTTTGTAATATAGCAACTT  199
    SS235  GCTC TT--TTTGTAATATAGCAACTT  194
```

Figure 3.7. Sequence alignment of flagellar switch and inverted repeat sequences from the clinical and environmental ribotype 012 *C. difficile* isolates. A multiple sequence alignment generated using Clustal Omega with the flagellar switch and inverted repeats from the *C. difficile* strains 630, JIR8094, MAM30, MT4768, MT5065, SE838, and SS235.
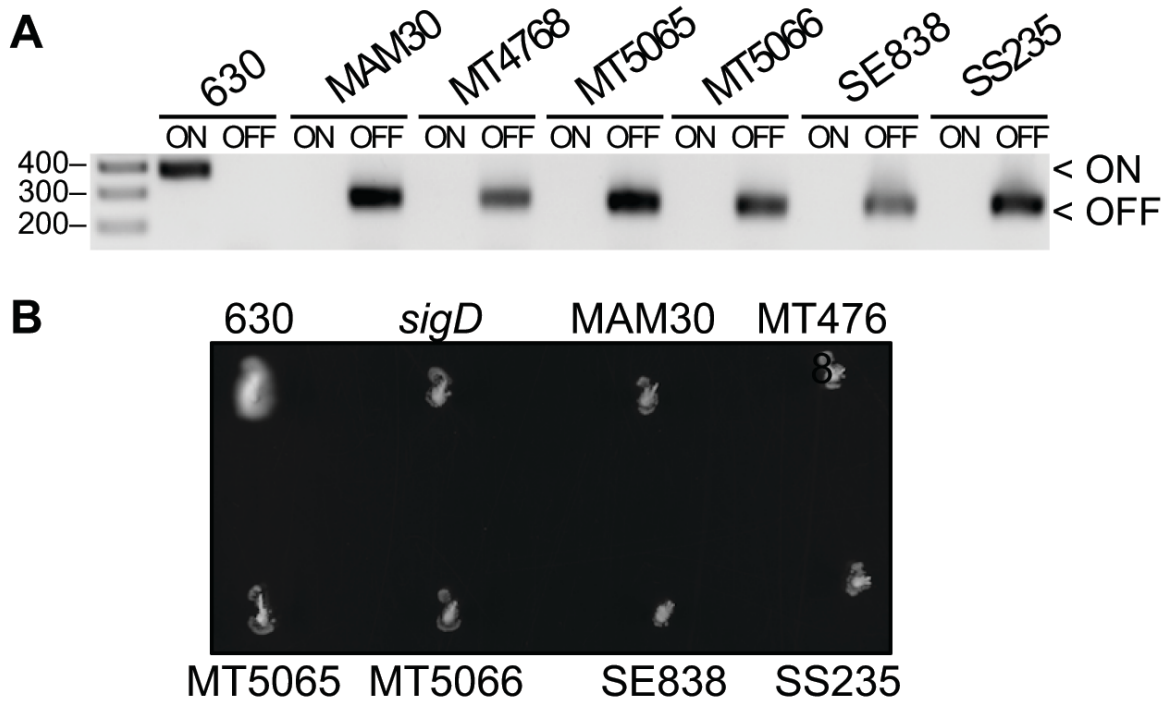
Figure 3.8. Clinical and environmental ribotype 012 isolates are *flg* OFF. (A) Orientation-specific PCR products using the indicated strains as template. Template from 630 is included as a control to detect the *flg* ON orientation. Image is representative of four biological replicates. (B) Swimming motility after 24 hours. Data are representative of three independent experiments with eight biological replicates of each strain.
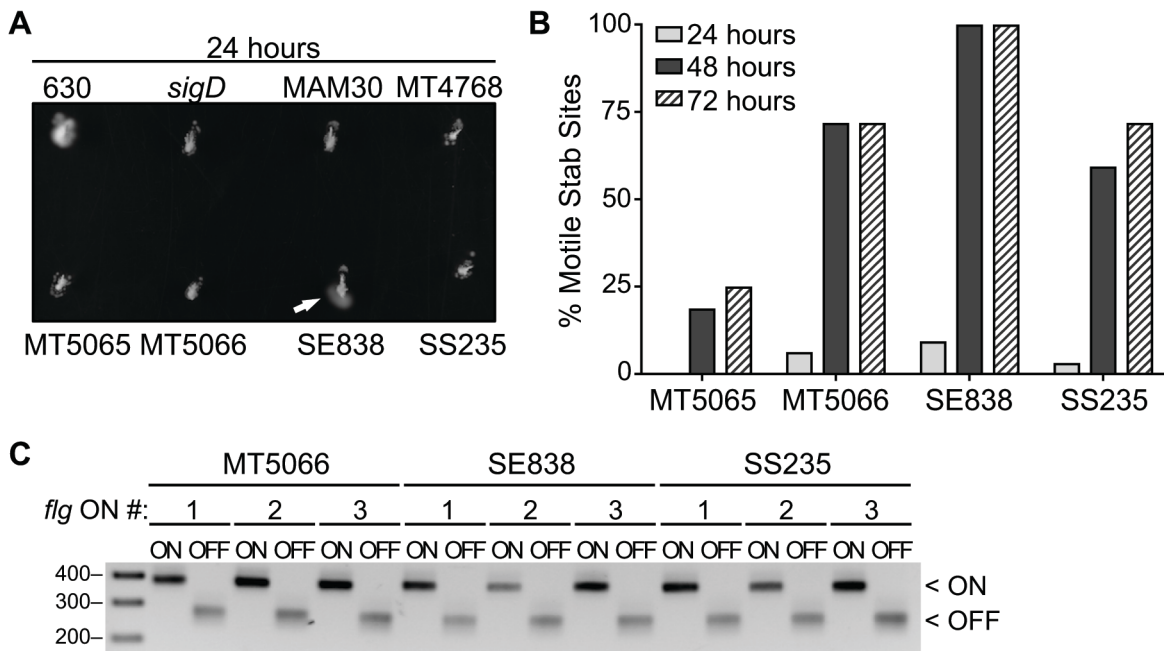
Figure 3.9. Low frequency recovery of *flg* ON bacteria from select ribotype 012 isolates. (A) Swimming motility after 24 hours. A representative image with a motile flare (white arrow) is shown. (B) The isolates differ in the frequency of recovery of motile revertants. Shown are pooled data from two independent experiments, each examining 16 biological replicates (n = 32). Frequency of motile flares cumulative for each time point, with those detected at 72 hours including those from 48 hours. (C) Orientation-specific PCR products using the indicated strains as template. All nine recovered *flg* ON isolates derived from *flg* OFF MT5066, SE838, and SS235 yielded PCR products indicating both *flg* ON and *flg* OFF populations. Data are representative of three independent experiments.

```
          630  AAGTTACTATTTT-ACAA--AAA GAGCAACTTTTCGAAGAAATATTTAAAT  48
MT5066 flg ON  AAGTTACTATTTT-ACAA--AAA GAGCAACTTTTCGAAGAAATATTTAAAT  48
 SE838 flg ON  AAGTTACTATTTTTACAAAAAAA GAGCAACTTTTCGAAGAAATATTTAAAT  51
 SS235 flg ON  AAGTTACTATTTT-ACAA--AAA GAGCAACTTTTCGAAGAAATATTTAAAT  48

          630  ACATATTAAAATTTTTTTATATTTTCATGGGACTTCATTAAAAAAATATTAG  100
MT5066 flg ON  ACATATTAAAATTTTTTTATATTTTCATGGGACTTCATTAAAAAAATATTAG  100
 SE838 flg ON  ACATATTAAAATTTTTTTATATTTTCATGGGACTTCATTAAAAAAATATTAG  103
 SS235 flg ON  ACATATTAAAATTTTTTTATATTTTCATGGGACTTCATTAAAAAAATATTAG  100

          630  TTTTCTTACTAAATTGATACATTATTGTATTAATGTAATGGAAATGCGATTA  152
MT5066 flg ON  TTTTCTTACTAAATTGATACATTATTGTATTAATGTAATGGAAATGCGATTA  152
 SE838 flg ON  TTTTCTTACTAAATTGATACATTATTGTATTAATGTAATGGAAATGCGATTA  155
 SS235 flg ON  TTTTCTTACTAAATTGATACATTATTGTATTAATGTAATGGAAATGCGATTA  152

          630  ATTTAAATAAATAAAGAAGTCA TT--TTTGTAA-TATAGCAACTT        194
MT5066 flg ON  ATTTAAATAAATAAAGAAGTCA TT--TTTGTAA-TATAGCAACTT        194
 SS235 flg ON  ATTTAAATAAATAAAGAAGTCA TTTTTTTGTAAATATAGCAACTT        200
 SE838 flg ON  ATTTAAATAAATAAAGAAGTCA TT--TTTGTAA-TATAGCAACTT        194
```

Figure 3.10. Sequence alignment of flagellar switch and inverted repeat sequences from the motile derivatives of clinical and environmental ribotype 012 *C. difficile* isolates. A multiple sequence alignment generated using Clustal Omega with the flagellar switch and inverted repeats from the *C. difficile* strains 630 and *flg* ON isolates of MT5066, SE838 and SS235.
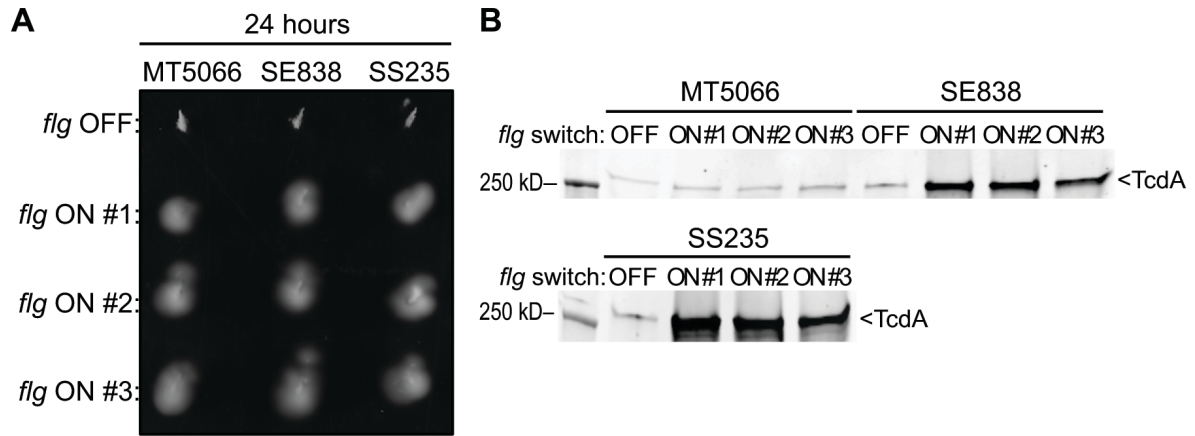
Figure 3.11. Characterization of select *flg* ON ribotype 012 isolates. (A) Swimming motility assay after 24 hours. All nine recovered *flg* ON isolates were motile at 24 hours compared to their non-motile *flg* OFF MT5066, SE838, and SS235 parents. Data are representative of three independent experiments with five biological replicates for each strain. (B) Detection of TcdA in the indicated strains by western blot. Data are representative of three independent experiments.

Table 3.1. Strains and plasmids used in this study.

Strain/Plasmid	Description/Purpose	Reference
Plasmids		
pRPF185	Anhydrotetracycline (ATc) inducible expression vector; Cm^R/Tm^{10}	(28)
pRT1611	pRPF185 with *gusA* removed (vector control)	(14)
pRT1830	pRPF185::*topA* (CD630_12740)	This work
***Escherichia coli* Strains**		
DH5α	F- Φ80*lacZ* Δ M15 (*lacZYA-argF*)U169 *recA1 endA1 hsdR*17(rκ -, mκ +) *phoA supE44 thi*-1 *gyrA*96 *relA1 λ- tonA*	Invitrogen, (41)
HB101(pRK24)	F- *mcrB mrr hsdS20*(r${_B}^-$m${_B}$)*recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20* (pRK24)	(30)
***Clostridium difficile* Strains**		
630	Ribotype 012	(21, 42)
630*Δerm*	Ribotype 012, erythromycin susceptible derivative of *C. difficile* 630	(24)
JIR8094 (630E)	Ribotype 012, erythromycin susceptible derivative of *C. difficile* 630	(23)
R20291	Ribotype 027, epidemic isolate	(43)
MAM30	Ribotype 012, clinical isolate from hospitalized patient	Gift from KW Garey
MT4768	Ribotype 012, clinical isolate from hospitalized patient	Gift from KW Garey
MT5065	Ribotype 012, clinical isolate from hospitalized patient	Gift from KW Garey
MT5066	Ribotype 012, clinical isolate from hospitalized patient	Gift from KW Garey
SE838	Ribotype 012, hospital environmental isolate	Gift from KW Garey
SS235	Ribotype 012, community environmental isolate	Gift from KW Garey
RT1931	MT5066 *flg* ON #1	This work
RT1946	MT5066 *flg* ON #2	This work
RT1947	MT5066 *flg* ON #3	This work
RT1932	SE838 *flg* ON #1	This work
RT1933	SE838 *flg* ON #2	This work
RT1948	SE838 *flg* ON #3	This work
RT1934	SS235 *flg* ON #1	This work
RT1935	SS235 *flg* ON #2	This work
RT1949	SS235 *flg* ON #3	This work
RT1566	R20291 *sigD::erm* (Targetron insertion at nucleotide position 228 of *sigD*, sense orientation)	(14)
RT782	JIR8094 motile derivative #1	This work
RT783	JIR8094 motile derivative #2	This work
RT784	JIR8094 motile derivative #3	This work
RT1609	JIR8094 motile derivative #1 + pRT1611	This work
RT1860	JIR8094 motile r derivative #1 + pRT1830	This work
RT1613	JIR8094 motile derivative #2 + pRT1611	This work
RT1861	JIR8094 motile derivative #2 + pRT1830	This work
RT1631	JIR8094 motile derivative #3 + pRT1611	This work
RT1862	JIR8094 motile derivative #3 + pRT1830	This work
RT1863	630 Δ*erm sigD::erm* + pRT1611	This work
RT1864	630 Δ*erm sigD::erm* + pRT1830	This work
RT1576	630 + pRT1611	This work

RT1858	630 + pRT1830	This work
RT1578	JIR8094 + pRT1611	This work
RT1859	JIR8094 + pRT1830	This work

Table 3.2. Oligonucleotides used in this study.

Primer Name[a]	Sequence (5' → 3')[b]	Lab Annotation
*flg*Ribo012_Pub	GAGCAACTTTTCGAAGAAATATTTAAATAC	R1751
*flg*Ribo012_Inv	GTATTTAAATATTTCTTCGAAAAGTTGCTC	R1752
*flg*Ribo027_Pub	AGGCAACTTTATAAAGAAATATTTAAATTTATATTAAAATATTTTTATATTTTTATTAGG	R1614
*flg*Ribo027_Inv	CCTAATAAAAATATAAAAATATTTTAATATAAATTTAAATATTTCTTTATAAAGTTGCCT	R1615
*flg*UTR5'end_For	**CGAATTC**GTATACTTAAGTTAAAACTAAATAGGCAAATC (**EcoRI**)	R591
CDR20291_0248 qF (*flgB*)	GCAACTAATCTAAGAAGTCAGACAATAGC	R856
CDR20291_0248 qR (*flgB*)	AGGCATAGCATCATTTAGTGTTTCTTC	R857
CDR20291_0270 qF (*sigD*)	GAATATGCCTCTTGTAAAGAGTATAGCA	R860
CDR20291_0270 qR (*sigD*)	TGCATCAATCAATCCAATGACTCC	R861
CDR20291_0240 qF (*fliC*)	CAAAGTAAGTCTATGGAGAA	R1584
CDR20291_0240 qR (*fliC*)	ACAGATATACCATCTTGAAC	R1585
CDR20291_0581 qF (*tcdR*)	AGCAAGAAATAACTCAGTAGATGATT	R908
CDR20291_0581 qR (*tcdR*)	TTATTAAATCTGTTTCTCCCTCTTCA	R909
CDR20291_0584 qF (*tcdA*)	GGAGAAGTCAGTGATATTGCTCTTG	R852
CDR20291_0584 qR (*tcdA*)	CAGTGGTAGAAGATTCAACTATAGCC	R853
CDR20291_0582 qF (*tcdB*)	AAGGAATATCTAGTTACAGAAGTATTAGAGC	R854
CDR20291_0582 qR (*tcdB*)	GCAGTGTCATTTATTTGACCTCCA	R855
CD630_topA_F	CAA**GAGCTC**GATTAGGTAGTGGGGGTGTAT (**SacI**)	R2302
CD630_topA_R	CAA**GGATCC**CAAGAGAATTATGTGATTTTTTTCAATTT (**BamHI**)	R2303
pRPF185MCSflank_F	TATTTCGATGCCCTGGACT	R1832
pRPF185MCSflank_R	ATCCCCTACTACTGACAGCTTC	R1833

[a] Where locus tags are used in the primer name, the corresponding gene name is noted in parentheses

[b] Restriction site sequences are highlighted with bold text; the corresponding enzyme is listed in parentheses.

REFERENCES

1. **Theriot CM**, **Young VB**. 2015. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. Annu Rev Microbiol **69**:445–461.

2. **Chandrasekaran R**, **Lacy DB**. 2017. The role of toxins in *Clostridium difficile* infection. FEMS Microbiol Rev **41**:723–750.

3. **Tasteyre A**, **Barc M-C**, **Collignon A**, **Boureau H**, **Karjalainen T**. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect Immun **69**:7937–7940.

4. **Dingle TC**, **Mulvey GL**, **Armstrong GD**. 2011. Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect Immun **79**:4061–4067.

5. **Baban ST**, **Kuehne SA**, **Barketi-Klai A**, **Cartman ST**, **Kelly ML**, **Hardie KR**, **Kansau I**, **Collignon A**, **Minton NP**. 2013. The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. PLoS ONE **8**:e73026.

6. **Stevenson E**, **Minton NP**, **Kuehne SA**. 2015. The role of flagella in *Clostridium difficile* pathogenicity. Trends Microbiol **23**:275–282.

7. **McKee RW**, **Mangalea MR**, **Purcell EB**, **Borchardt EK**, **Tamayo R**. 2013. The second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of *sigD*. J Bacteriol **195**:5174–5185.

8. **Aubry A**, **Hussack G**, **Chen W**, **KuoLee R**, **Twine SM**, **Fulton KM**, **Foote S**, **Carrillo CD**, **Tanha J**, **Logan SM**. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun **80**:3521–3532.

9. **Meouche El I**, **Peltier J**, **Monot M**, **Soutourina O**, **Pestel-Caron M**, **Dupuy B**, **Pons J-L**. 2013. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of *tcdR*. PLoS ONE **8**:e83748.

10. **Mani N**, **Dupuy B**. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci USA **98**:5844–5849.

11. **Martin-Verstraete I**, **Peltier J**, **Dupuy B**. 2016. The regulatory networks that control *Clostridium difficile* toxin synthesis. Toxins (Basel) **8**:153.

12. **Martin MJ**, **Clare S**, **Goulding D**, **Faulds-Pain A**, **Barquist L**, **Browne HP**, **Pettit L**, **Dougan G**, **Lawley TD**, **Wren BW**. 2013. The *agr* locus regulates virulence and colonization genes in *Clostridium difficile* 027. J Bacteriol **195**:3672–3681.

13. **Girinathan BP**, **Ou J**, **Dupuy B**, **Govind R**. 2018. Pleiotropic roles of *Clostridium difficile sin* locus. PLoS Pathog **14**:e1006940.

14.   **Anjuwon-Foster BR**, **Tamayo R**. 2017. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet **13**:e1006701.

15.   **Emerson JE**, **Reynolds CB**, **Fagan RP**, **Shaw HA**, **Goulding D**, **Fairweather NF**. 2009. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol **74**:541–556.

16.   **Reynolds CB**, **Emerson JE**, **la Riva de L**, **Fagan RP**, **Fairweather NF**. 2011. The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog **7**:e1002024.

17.   **Anjuwon-Foster BR**, **Tamayo R**. 2017. Phase variation of *Clostridium difficile* virulence factors. Gut Microbes 1–8.

18.   **Shen A**. 2012. *Clostridium difficile* toxins: mediators of inflammation. J Innate Immun **4**:149–158.

19.   **Yoshino Y**, **Kitazawa T**, **Ikeda M**, **Tatsuno K**, **Yanagimoto S**, **Okugawa S**, **Yotsuyanagi H**, **Ota Y**. 2013. *Clostridium difficile* flagellin stimulates toll-like receptor 5, and toxin B promotes flagellin-induced chemokine production via TLR5. Life Sci **92**:211–217.

20.   **Batah J**, **Kobeissy H**, **Bui Pham PT**, **Denève-Larrazet C**, **Kuehne S**, **Collignon A**, **Janoir-Jouveshomme C**, **Marvaud J-C**, **Kansau I**. 2017. *Clostridium difficile* flagella induce a pro-inflammatory response in intestinal epithelium of mice in cooperation with toxins. Sci Rep **7**:3256.

21.   **Wüst J**, **Sullivan NM**, **Hardegger U**, **Wilkins TD**. 1982. Investigation of an outbreak of antibiotic-associated colitis by various typing methods. J Clin Microbiol **16**:1096–1101.

22.   **Collery MM**, **Kuehne SA**, **McBride SM**, **Kelly ML**, **Monot M**, **Cockayne A**, **Dupuy B**, **Minton NP**. 2016. What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and derivatives. Virulence 1–15.

23.   **O'Connor JR**, **Lyras D**, **Farrow KA**, **Adams V**, **Powell DR**, **Hinds J**, **Cheung JK**, **Rood JI**. 2006. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol **61**:1335–1351.

24.   **Hussain HA**, **Roberts AP**, **Mullany P**. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δ*erm*) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. J Med Microbiol **54**:137–141.

25.   **Martinson JNV**, **Broadaway S**, **Lohman E**, **Johnson C**, **Alam MJ**, **Khaleduzzaman M**, **Garey KW**, **Schlackman J**, **Young VB**, **Santhosh K**, **Rao K**, **Lyons RH**, **Walk ST**. 2015. Evaluation of portability and cost of a fluorescent PCR ribotyping protocol for *Clostridium difficile* epidemiology. J Clin Microbiol **53**:1192–1197.

26. **Zhao H**, **Li X**, **Johnson DE**, **Blomfield I**, **Mobley HL**. 1997. In vivo phase variation of MR/P fimbrial gene expression in *Proteus mirabilis* infecting the urinary tract. Mol Microbiol **23**:1009–1019.

27. **Purcell EB**, **McKee RW**, **McBride SM**, **Waters CM**, **Tamayo R**. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol **194**:3307–3316.

28. **Fagan RP**, **Fairweather NF**. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286**:27483–27493.

29. **Bouillaut L**, **McBride SM**, **Sorg JA**. 2005. Genetic Manipulation of *Clostridium difficile*. Curr Protoc Microbiol **20**:9A2.1-9A2.17.

30. **McBride SM**, **Sonenshein AL**. 2011. Identification of a genetic locus responsible for antimicrobial peptide resistance in *Clostridium difficile*. Infect Immun **79**:167–176.

31. **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**, **Severinov K**, **Le Bouguenec C**, **Coppée J-Y**, **Dupuy B**, **Martin-Verstraete I**. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

32. **Sudarsan N**, **Lee ER**, **Weinberg Z**, **Moy RH**, **Kim JN**, **Link KH**, **Breaker RR**. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321**:411–413.

33. **Bordeleau E**, **Purcell EB**, **Lafontaine DA**, **Fortier L-C**, **Tamayo R**, **Burrus V**. 2015. Cyclic di-GMP riboswitch-regulated Type IV pili contribute to aggregation of *Clostridium difficile*. J Bacteriol **197**:819–832.

34. **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

35. **Kuehne SA**, **Cartman ST**, **Heap JT**, **Kelly ML**, **Cockayne A**, **Minton NP**. 2010. The role of Toxin A and Toxin B in *Clostridium difficile* infection. Nature **467**:711–713.

36. **Greene SE**, **Hibbing ME**, **Janetka J**, **Chen SL**, **Hultgren SJ**. 2015. Human urine decreases function and expression of Type 1 pili in uropathogenic *Escherichia coli*. MBio **6**:e00820–15.

37. **Lane MC**, **Li X**, **Pearson MM**, **Simms AN**, **Mobley HLT**. 2009. Oxygen-limiting conditions enrich for fimbriate cells of uropathogenic *Proteus mirabilis* and *Escherichia coli*. J Bacteriol **191**:1382–1392.

38. **Groß U**, **Brzuszkiewicz E**, **Gunka K**, **Starke J**, **Riedel T**, **Bunk B**, **Spröer C**, **Wetzel D**, **Poehlein A**, **Chibani C**, **Bohne W**, **Overmann J**, **Zimmermann O**, **Daniel R**, **Liesegang H**. 2018. Comparative genome and phenotypic analysis of three *Clostridioides*

difficile strains isolated from a single patient provide insight into multiple infection of *C. difficile*. BMC Genomics **19**:1.

39. **Lewis JA**, **Hatfull GF**. 2001. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. Nucleic Acids Res **29**:2205–2216.

40. **Serrano M**, **Kint N**, **Pereira FC**, **Saujet L**, **Boudry P**, **Dupuy B**, **Henriques AO**, **Martin-Verstraete I**. 2016. A recombination directionality factor controls the cell type-specific activation of σK and the fidelity of spore development in *Clostridium difficile*. PLoS Genet **12**:e1006312.

41. **Hanahan D**. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.

42. **Sebaihia M**, **Wren BW**, **Mullany P**, **Fairweather NF**, **Minton N**, **Stabler R**, **Thomson NR**, **Roberts AP**, **Cerdeño-Tárraga AM**, **Wang H**, **Holden MT**, **Wright A**, **Churcher C**, **Quail MA**, **Baker S**, **Bason N**, **Brooks K**, **Chillingworth T**, **Cronin A**, **Davis P**, **Dowd L**, **Fraser A**, **Feltwell T**, **Hance Z**, **Holroyd S**, **Jagels K**, **Moule S**, **Mungall K**, **Price C**, **Rabbinowitsch E**, **Sharp S**, **Simmonds M**, **Stevens K**, **Unwin L**, **Whithead S**, **Dupuy B**, **Dougan G**, **Barrell B**, **Parkhill J**. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet **38**:779–786.

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

CHAPTER 4: RHO FACTOR IS REQUIRED FOR PHASE VARIATION OF

FLAGELLA IN *CLOSTRIDIUM DIFFICILE*[3]

Summary

Clostridium difficile is a bacterial pathogen that causes an antibiotic-associated diarrheal disease. *C. difficile* produces flagella that aid in bacterial motility and secretes toxins that are necessary for diarrheal disease symptoms. Both flagella and toxins are co-regulated in *C. difficile* via a genetic ON/OFF switch in a process called phase variation. Motile and toxigenic bacteria are phase ON, whereas non-motile and non-toxigenic bacteria are phase OFF. Herein, we identified Rho factor as a requirement for the ON/OFF production of flagella and toxins in *C. difficile*. Mutations in Rho factor that are predicted to inactivate essential features of Rho activity restore motility to phase OFF bacteria. Normally, Rho factor terminates messages for production of proteins near the end. We predict that Rho factor directly terminates the message for production of flagella and toxins in phase OFF bacteria early. Our findings add to an emerging paradigm that Rho factor can terminate messages early and late in their expression.

Introduction

Clostridium difficile is a Gram-positive, spore-forming anaerobe and the leading cause of antibiotic-associated diarrheal disease. Antibiotic treatment for a preexisting infection increases

---

[3] I performed the experiments for Figures 4.1-4.6 and Tables 4.1-4.2, except Dr. Rita Tamayo isolated seven motile suppressors of *recV flg* OFF* listed in Table 4.3 (RT1939-RT1945). I wrote an original draft of Chapter 4 and Dr. Tamayo and I revised the final version herein.

susceptibility to *C. difficile* colonization (1). Antibiotic use disrupts the gastrointestinal microbial community that normally affords resistance to *C. difficile* infection. Diarrheal disease symptoms are most attributable to the glucosylating toxins TcdA and TcdB (2-5). In host cells, TcdA and TcdB glucosylate the regulatory regions of Rho family GTPases to prevent interaction with downstream effectors (6, 7). As a consequence, Rho GTPases can no longer interact with cellular factors controlling the actin cytoskeleton. In cell culture, phenotypically this results in a disruption of tight junctions and loss of epithelial cell integrity (8, 9), and downstream cell death through diverse mechanisms. Toxin activity during infection results in the disruption of the intestinal barrier and immune cell recruit with inflammation.

The expression of both toxin genes is subject to diverse regulatory mechanisms to precisely control production of these energetically costly toxins (10). For example, in stationary phase conditions where amino acids and carbohydrates are low, the nutrient sensing transcriptional regulators CodY and CcpA derepress transcription of the genes encoding the toxin genes and their positive regulator TcdR (11-14). However, the toxins are dispensable for initial colonization in small animal models of disease (2-5). *C. difficile* encodes various surface bound appendages that enable bacterial colonization in the intestine, albeit with some redundancy to ensure survival (15, 16). Flagella in *C. difficile* contribute not only to swimming motility (17, 18), but also adherence to intestinal epithelial cells, colonization, and virulence in animal models of infection (19-21). Similar to *Bacillus subtilis*, the flagellar gene operons are arranged in discrete loci and expressed in a hierarchical manner to ensure assembly of the right complexes at the appropriate time (22). The early stage flagellar operon (*flgB*) encodes an alternative sigma factor, SigD, that is essential for expression of late stage flagellar genes (23, 24). In addition, SigD contributes to expression of the toxin genes by activating transcription of

tcdR (20, 23, 24). Thus, factors that regulate expression of the *flgB* operon can also impact toxin production via SigD. Numerous regulators, such as Agr quorum sensing (25), SinR (26), RstA (27), and cyclic diguanylate (23) have been shown to affect both flagellar and toxin production in *C. difficile*. However, the exact regulatory mechanisms for many remain undefined.

Recently we demonstrated that *C. difficile* flagella and toxin production is phase variable in both ribotypes 027 and 012 (28, 29). Phase variation is a process in bacteria allowing stochastic control of the production of cell surface or secreted proteins to generate phenotypic diversity (30). Diverse genetic mechanisms are at play in phase variation (30). In *C. difficile*, conservative site-specific recombination of a genetic switch termed the "flagellar switch" upstream of the *flgB* operon is responsible for phase variation of flagella and toxins (28). The flagellar switch is an invertible 154 bp element and flanked by 21 bp imperfect inverted repeats. In flagellar phase ON (*flg* ON) bacteria, the switch is in an orientation permissive for downstream gene expression, and therefore *C. difficile* produces flagella, is motile, and secretes the glucosylating toxins. In flagellar phase OFF (*flg* OFF) bacteria, the switch restricts gene expression, and *C. difficile* is aflagellate, non-motile, and attenuated for toxin production. The canonical regulatory mechanism inherent to genetic switches relies on an invertible promoter element. In one orientation, the promoter is positioned to control transcription of a downstream gene by RNA polymerase. However, inversion of the switch disrupts the orientation-specific nature of the -10 and -35 elements such that RNA polymerase can no longer transcribe the target gene. The fimbrial switch in *Escherichia coli* contains a promoter element for *fimA* operon expression (31). In *Salmonella*, a flagellar switch encodes the Hin recombinase necessary for inversion as well as a promoter to control expression of *fljBA* operon (32).

In *C. difficile*, the flagellar switch lies between the previously identified, upstream σ^A-dependent promoter and the *flgB* coding sequence. Thus, the σ^A-dependent promoter drives transcription through the 5' untranslated region (UTR) including the flagellar switch (33). We showed the mechanism of regulation via the flagellar switch occurs post-transcription initiation and involves an unidentified *trans*-acting posttranscriptional regulator specific to *C. difficile* (28). The only other genetic switch functionally characterized in *C. difficile* is 5' of the *cwpV* gene (34-36). The *cwpV* switch results in the formation of an intrinsic terminator in the mRNA of *cwpV* OFF bacteria, but allows transcriptional readthrough in *cwpV* ON bacteria (34). The flagellar switch in *C. difficile* does not contain an intrinsic terminator (28). Therefore, we sought to identify the regulatory mechanism by which the flagellar switch controls downstream *flgB* operon expression.

The tyrosine recombinase RecV is both necessary and sufficient to mediate inversion of the flagellar switch in both direction in *C. difficile* strain R20291 (28). Mutation of *recV* in R20291 locks the flagellar switch in either orientation. We denote the phase-locked mutants with an asterisk to differentiate them from phase variable states: phase-locked *recV::erm flg* ON (*recV flg* ON*) or phase-locked *recV::erm flg* OFF (*recV flg* OFF*) (28). The genotype of the flagellar switch predicts the phenotype in these mutants: *recV flg* ON* produce flagella, are motile, and secrete toxins, whereas *recV flg* OFF* are aflagellate, non-motile, and attenuated for toxin production (28). We previously observed that extended incubation of the *recV flg* OFF* mutant in motility medium can infrequently give rise to "motile flares" in which motile bacteria migrate from the stab site of non-motile bacteria (Figure 2.19A). We hypothesized that the motile *recV flg* OFF* bacteria were suppressors with mutations in a gene(s) responsible for inhibiting gene expression in *flg* OFF bacteria (28, 37). In this body of work, we sequenced the suppressors,

identified the mutations, and characterized the suppressor mutants in motility and toxin assays. Mutations were identified in the *rho* gene, suggesting that Rho factor is responsible for termination of gene expression in *flg* OFF bacteria. Rho factor is an ATP-dependent RNA helicase that mediates transcription termination at the 3' end of a gene or operon (38). Emerging evidence shows Rho factor can mediate termination in the leader RNA as a regulatory mechanism due to changes in RNA structure (39, 40). Based on our results, we propose models in which Rho factor either directly or indirectly mediates transcription termination in the *flgB* operon leader RNA of *flg* OFF bacteria.

Results

Frequency in detection of motile suppressors

We previously described four motile suppressors (MS) from a *recV flg* OFF* mutant. Despite recovering swimming motility, the MS strains retained the flagellar switch in the OFF orientation. These data suggest that under the conditions tested RecV is the sole recombinase responsible for inversion of the flagellar switch. No mutations were present in the *flgB* operon promoter and leader sequence in the *recV flg* OFF* MS compared to the parental strain (data not shown). We hypothesized that the *recV flg* OFF* MS have either (1) extragenic gain-of-function mutations in positive regulators of *flgB* operon expression to override the *flg* OFF regulatory mechanism, or (2) extragenic loss-of-function mutations in an unidentified *C. difficile*-specific, *trans*-acting regulatory factor responsible for transcription termination in *flg* OFF bacteria. Motility assays were performed with 64 biological replicates of *recV flg* OFF* over two independent experiments to calculate the frequency of MS by counting motile flares. The *recV*

172

flg ON* mutant served as a motile control, and a *sigD::erm* (*sigD*) mutant served as a non-motile control. Motility was monitored every 24 hours for three days. Only 3.1% of *recV flg* OFF* colonies stabbed into the motility medium developed motile flares at 24 hours (Figure 4.1). The percentage of colonies showing motility increased from 10.9% at 48 hours to 21.9% at 72 hours (Figure 4.1). Based on these data, we conclude that recovery of motility in the non-motile *recV flg* OFF* mutant is an infrequent event.

The motile suppressors display restored motility

Based on the presence of motile flares, we predicted the MS would have restored swimming motility. We isolated fourteen MS from motility medium and assayed for swimming motility compared to motile *recV flg* ON*, non-motile *recV flg* OFF*, and a non-motile *sigD* mutant control strains. All 14 MS were motile at 24 and 48 hours compared to the non-motile *recV flg* OFF* and *sigD* mutants (Figure 4.2A & 4.2B). In addition, MS #8-14 had a significant increase in motility compared to *recV flg* ON* at 48 hours *(p < 0.05 *)* (Figure 4.2B). From these data, we concluded that the infrequent motile flares detected in the motility medium (Figure 4.1) from the parental *recV flg* OFF* mutant represent a motile suppressor population with recovered swimming motility.

Identification of the mutations in motile suppressors

To identify the genes involved in recovery of swimming motility in the *recV flg* OFF* motile suppressors, whole genome sequencing and variant analysis was performed on seven independent MS and the parental *recV flg* OFF* mutant. A minimum threshold of 50% of sequencing reads with an allelic variant relative to the reference genome was arbitrarily set to

identify mutations present in the MS, but absent from the parental strain (Materials & Methods).

Only one gene, CDR20291_33240, which encodes an orthologue of Rho factor from *E. coli* (38),

was altered in all seven MS; MS #1-7 contained a frameshift, missense, or nonsense mutation in

rho (Table 4.1). Additional polymorphisms were identified in a subset of these strains. In 5/7

strains (MS #2-6), we identified a single nucleotide polymorphism (SNP) in the intergenic region

between two convergent genes: a gene coding a prophage site-specific recombinase and the *ilvB*

gene, coding for the acetolactate synthase large subunit. In the same 5 MS, we identified two

additional mutations upstream of a gene annotated to encode a manganese-containing catalase. In

MS #7, we identified a SNP in a gene coding for a predicted transposase-like protein B element.

Seven additional MS of the *recV flg* OFF* mutant (MS #8-14) were isolated from motile

flares for sequencing of the *rho* gene, and 6/7 contained a mutation in *rho* (Table 4.1). MS #11

did not contain a mutation in *rho*, but was included for comparison in most assays since it had a

mild growth defect unlike most of the suppressors (discussed later). Based on these results, we

focused on evaluating Rho factor as the lead candidate gene product required for termination of

transcription in the *flgB* operon leader RNA of *flg* OFF bacteria

Predicted functional consequences of *rho* mutations

Rho factor is an homohexamer that functions as an ATP-dependent RNA helicase that

mediates transcription termination at the 3' end of a gene or operon (38). Domains and amino

acid residues important for Rho factor function have been delineated in great detail in the model

organism *E. coli*. In this section, we briefly review the functional domains in *E. coli* Rho

compared to *C. difficile* Rho, and we predict the functional consequence of most mutations

(Figure 4.3A and 4.3B). *C. difficile* Rho monomer can be divided into two domains, C-terminal

and N-terminal, each with distinct subdomains that contribute to function (Figure 4.1). The N-terminal domain (NTD) has three subdomains: an N-terminal Helix Bundle (NHB), an Insertion Domain, and a Primary Binding Site (PBS). The NHB consists of amino acids 11-51 and helps recruit nascent RNA to the PBS. The PBS consists of amino acids 171 to 223 and, in *E. coli*, lies in the outer leaflet of the hexamer, and binds RNA rich in pyrimidine bases, specifically cytosines. The sequence recognized by the PBS is called the *rho* utilization (*rut*) site and lies 60-70 nucleotides or bases upstream of the termination zone. Rho factor remains tethered to the *rut* site via the PBS, and nascent RNA is thread through the interior channel of the hexamer. Mutations in the PBS affect termination efficiency. The Insertion Domain varies in length and amino acid composition and is present in some bacteria and contributes to various functions. The CTD consists of several features that contribute to ATP binding, hydrolysis, and RNA binding. The CTD has Walker A (AA 288-295) and B (AA 371-375) motifs that are present in many proteins that use ATP binding and hydrolysis for energy. ATPase activity is required for both RNA-DNA helicase function of Rho in separating the nascent mRNA from the template DNA out of the transcription elongation complex of RNAP, and the termination activity. The Q-loop (AA 390-402) and R-loop (AA 431-436) consist of the Secondary Binding Site (SBS), which contributes to RNA binding along the interior of the hexamer during translocation. Mutations in the SBS affect the termination site. RNA binding by SBS induces conformation changes to allow for ATP binding and hydrolysis and translocation of Rho along the transcript.

Although our suppressor screen was not saturating, we identified several mutants presumably defective in key functions of Rho. Several mutants have missense mutations in domains that contribute to RNA binding in either the PBS and the Q-loop of the SBS – MS #7 (R201M) and MS #9 (L404P) (Figure 4.3B). The proline likely introduces conformational

rigidity due to the cyclic structure and prevents conformational changes in the Q-loop. Several

mutants have missense mutations in the CTD likely affecting interdomain interactions, ATP

binding and hydrolysis, and structural transitions: MS #2 (G474V), MS#3 (G279E), MS #8

(S473F), MS #12 (R322I), and MS #13 (D320A) (Figure 4.3B). The MS #12 (R322I) is the

critical arginine valve residue necessary for ATP → ADP transitions. Several mutants have

nonsense mutations resulting in truncated Rho proteins of various sizes and lacking critical

domains: MS #4 (E444Stop), MS #5 (AA 70 frameshift), MS #6 (E255Stop), MS #10

(E108Stop), and MS #14 (E247Stop) (Figure 4.3B). The MS #1 (N528S) remains to be

determined since it is the second to last amino acid. The mutation in MS #11 has not been

identified and may be upstream of *rho* affecting expression or in another gene that contributes to

termination in *flg* OFF bacteria. Collectively, these mutants can be used to probe specific

domains in *C. difficile*, specifically the ATP binding and hydrolysis, interdomain interactions,

and RNA binding domains PBS and SBS.

Putative inactivating *rho* mutations attenuate planktonic growth

      Inactivation of *rho* in *E. coli* is lethal (41), but *rho* mutants in *B. subtilis* and

Staphylococcus aureus are viable with negligible growth defects (42, 43). Some of the MS

developed smaller colonies on nutrient rich agar medium, and as detailed above, some of the

observed mutations are predicted to abolish critical features for Rho activity, such as ATP

binding and hydrolysis. Therefore, we evaluated the growth kinetics of the 14 MS relative to the

recV flg ON* mutant and the parental *recV flg* OFF* mutant. We found that 4/14 MS (i.e., MS

#1, MS #7, MS #8, and MS #11) had reproducible growth kinetics comparable to the *recV flg*

OFF* and *recV flg* ON* mutants in rich medium (Figure 4.4A). Growth rates were calculated

using exponential phase time points at 4 hours and 5.5 hours. All motile suppressors exhibited variable growth rates compared to the parent *recV flg* OFF* mutant in exponential phase (Figure 4.4B). That all the MS grew *in vitro* is consistent with previous work showing that the *rho* gene is not essential for growth in *C. difficile* strain R20291 (44). Targeted mutation of the *rho* gene in R20291 is currently underway.

Rho factor is dispensable for regulation of toxin production

The flagellar switch orientation affects expression of *sigD* in the *flgB* operon (28, 29). Moreover, a *recV flg* ON* mutant produces high concentrations of TcdA, whereas the *recV flg* OFF* mutant is attenuated for TcdA production (28). SigD, the flagellar alternative sigma factor, is required for expression of late stage flagellar genes and contributes to expression of the toxin genes. In a hierarchy, SigD activates *tcdR* expression, and TcdR activates *tcdA* and *tcdB* expression (20, 23, 24). Therefore, we predicted that TcdA levels would increase in the MS relative to the parental strain. TcdA levels were assessed by western blot in strains grown for 12 hours in TY and BHIS medium. MS #1, #7, #8, and #11 were used since their growth curves were most similar to the *recV flg* OFF* mutant (Figure 4.4A and 4.4B). The *recV flg* ON* mutant served as a positive control, and both the *recV flg* OFF* and *sigD* mutants served as negative controls. In TY medium, we found that TcdA levels were indistinguishable among the *recV flg* OFF*, *recV flg* ON*, MS #1, MS #7, MS #8, MS #11, and *sigD* mutants (Figure 4.5A). Although peptide-rich TY medium is ideal for assessing toxin production (45), SigD activity may be higher in glucose-rich media, such as BHIS, which is used in the motility medium. In BHIS, we found that TcdA levels, as previously published (28), were significantly reduced in the *recV flg* OFF* mutant compared to the *recV flg* ON* mutant, and the MS #1, #7, #8, and #11 had

variable recovery in TcdA production between independent experiments (Figure 4.5B). Loss of

function mutations in Rho factor may have pleotropic effects on the expression and production of

other positive and negative regulators of toxin production that confound the role of the flagellar

switch and SigD.

Rho mutant alleles exhibit dominant negative effects

To load onto a nascent RNA and mediate transcription termination, Rho factor must

assemble into a hexamer to function (38). Mutant monomers expressed *in trans* to a wildtype

allele could be antagonistic to the wildtype multimeric protein rendering the protein non-

functional. To determine if the mutations in the *rho* gene of the motile suppressors are dominant

negative, we cloned each of the 13 *rho* mutant alleles downstream of the anhydrotetracycline

(ATc)-inducible promoter (46). The wildtype allele of *rho* from the *recV flg* OFF* mutant served

as a negative control for the assay. The plasmids were transformed into the *recV flg* OFF*

mutant, and resulting strains were assayed for the frequency of motile flares in motility medium

every 24 hours for two days to determine whether the *rho* mutant alleles could exert a dominant

negative effect (47). Frequency of motile reversion was determined with and without ATc

inducer since the P_{tet} promoter is known to be leaky (47). As expected, expression of the

wildtype *rho* allele resulted in the lowest frequency of motile stab sites regardless of ATc

concentration. Less than 25% of stab sites had motile flares for strains expressing the G279E,

AA70→FS→Stop79, L404P, or D320A *rho* allele at 24 hours; by 48 hours the strains expressing

the G279E, L40P, and D320A alleles yielded motile flares in 100% of stab sites without ATc

(Table 4.2). Leaky expression of the remaining mutant *rho* alleles exerted a dominant negative

effect with greater than 25% of stab sites positive for motile flares by 24 hours and 100% positive by 48 hours (Table 4.2).

Dominant negative mutants should exhibit a dose-dependent effect over the wildtype gene product. Interestingly, we found that the addition of ATc to motility medium reduced the frequency of motile conversion for all the Rho mutants at 24 hours compared to uninduced, except the R322I allele increased from 75% to 100% (Table 4.2). By 48 hours, motility medium with ATc had at least 75% of stab sites positive for motile flares with all the Rho mutants (Table 4.2). These data demonstrate that Rho mutant alleles can exert a dominant negative effect over the wildtype Rho homohexamer. We propose that this occurs through generation of heterohexamers of wildtype and mutant Rho, resulting in loss of Rho-mediated inhibition of expression in *flg* OFF* bacteria. The reciprocal experiments with Rho mutant overexpression in the *recV flg* ON* mutant had no effect on motility with or without ATc inducer (Table 4.3). Motile flares detected for expression of the wildtype Rho in the *recV* flg OFF* background were isolated for future characterization.

Rho-dependent regulation occurs after transcription initiation of *flgB* operon

Rho inactivation can affect gene expression by direct, indirect *cis*, or indirect *trans* effects (48). Direct regulation is the canonical mechanism involving Rho factor binding, translocating, and terminating transcription by paused RNAP at specific sites. Indirect regulation *in cis* can be due to continued transcription into a neighboring gene or operon if an upstream terminator is Rho dependent. Indirect regulation *in trans* can be due to the production of a factor encoded elsewhere that impacts gene expression at the promoter or post-transcription initiation. To determine whether Rho factor affects *flgB* operon expression at the promoter level (indirect

trans) or post-transcription initiation (direct or indirect *trans*), we used alkaline phosphatase (AP) gene reporters. The same AP reporter constructs described previously were used (28), except cloned into a low-copy plasmid for higher overall AP activity. Briefly, the native *flgB* promoter (σ^A-dependent), the full 498 bp 5' UTR with the flagellar switch (FS) in either the phase ON or OFF orientation, and *flgB* (first gene of the *flgB* operon) were placed upstream of the *phoZ* reporter gene (P$_{flgB}$-5'UTR(FS$^{ON/OFF}$)-*flgB*::*phoZ*; ON or OFF). A transcriptional fusion of *phoZ* to the *flgB* promoter (P$_{flgB}$::*phoZ*; Promoter) served as a positive control for the assay, and a promoterless construct (::*phoZ*; Promoterless) was included as a negative control. The four reporters were transformed into the following mutants: *recV flg* OFF*, *recV flg* ON*, MS #1 (N528S), MS #7 (R201M), MS #8 (S473F), and MS #11 (suppressor mutation undetermined). These MS were chosen because they have growth kinetics most similar to the *recV flg* ON* and OFF* strains. As anticipated, the Promoterless reporter control showed no AP activity, and the presence of the Promoter reporter significantly increased activity in all strains tested (Figure 4.6). AP activity was detectable and comparable for the ON reporter in the *recV flg* OFF* and *recV flg* ON* mutants, indicating that the *flg* ON leader RNA is resistant to Rho-dependent termination (Figure 4.6). However, AP activity was higher for the ON reporter in all four MS compared to *recV flg* ON*, which suggests that either Rho factor affects gene expression in the *flg* ON background to a minor degree, or a regulator of transcription elongation or RNA stability is highly expressed or more active in *rho* mutants. These possibilities are not mutually exclusive. The OFF reporter had negligible AP activity and was indiscernible between the *recV flgB* ON* and *recV flgB* OFF* mutants, but was significantly higher in all four MS (Figure 4.6). Proficient gene expression with the OFF reporter in MS is consistent with the phenotypic recovery of motility observed in these strains. Collectively, these data eliminate the possibility that Rho

factor exerts an indirect *trans* effect at the promoter of the *flgB* operon. Moreover, we predict Rho factor mediates transcription termination at the *flgB* operon leader RNA either through a direct or indirect *trans* mechanism.

Discussion

Based on these studies, our data suggest that flagellar phase variation in *Clostridium difficile* strain R20291 requires the activity of Rho factor. In summary, a flagellar switch upstream of the *flgB* operon mediates the ON/OFF expression of flagellar and toxin genes (28, 29). Phase-locked non-motile bacteria infrequently give rise to motile suppressors with mutations in *rho*, implicating Rho factor in flagellar and toxin phase variation. The identified Rho mutant alleles are predicted to be non-functional since their overexpression restores motility to phase-locked non-motile bacteria.

We isolated a constellation of missense Rho mutants predicted to affect mRNA binding, translocation, and termination – all essential steps in Rho-dependent transcription termination. These results, particularly the isolation of multiple isolates with nonsense mutations in Rho, suggest that *rho* is not essential for growth. The motile suppressors have growth defects in nutrient rich medium suggesting that *rho* inactivation does attenuate expression or activity of important processes in *C. difficile*. In contrast, *rho* inactivation in *B. subtilis* and *S. aureus* have negligible effects on growth (42, 43). In a saturating transposon screen in *C. difficile* R20291 (44), Rho was absent from the list of essential genes, which supports our findings. Interestingly, *nusA* and *nusG* mutants were not recovered in that screen. NusA associates with RNAP to help fold and stabilize hairpin-dependent pauses and terminators in bacteria (49), but also antagonizes Rho factor activity through direct competition for specific sites on transcripts (38, 50). NusG

interacts with RNAP to stimulate transcription elongation (51), and also stabilizes the Rho secondary binding site interaction with RNA to promote termination (38, 52).

Rho factor is required for phase variation of fimbriae in *E. coli* and capsule in *Neisseria meningitidis* (53, 54). In *E. coli*, the FimE recombinase controls inversion of the fimbrial switch (*fimS*), which contains a promoter, in the ON to OFF direction to prevent transcription of the *fimA* structural gene. The *fimE* gene is located immediately 5' of *fimS*. Joyce *et al.* demonstrated that *fimE* mRNA half-life is differentially regulated based on *fimS* orientation. In the *fimS* ON orientation, the transcript for *fimE* extends across the switch into *fimA*, whereas the OFF transcript terminates in *fimS*. Rho inactivation restored *fimE* mRNA stability in *fimS* OFF bacteria. The *fimS* mRNA contains a *boxA* sequence, a sequence associated with Rho-mediated termination (55), exclusively when in the OFF orientation to facilitate Rho-dependent transcription termination. In *N. meningitidis*, slip-strand mispairing of a repeat sequence within the coding sequence of a polysialyltransferase gene reveals a cryptic Rho-dependent transcription termination site in capsule-OFF bacteria (54). We propose that in *C. difficile*, the change in mRNA sequence due to DNA switch inversion reveals either a *rut* sequence or a strong RNAP pause site in *flg* OFF bacteria. The strand specific determinant of Rho factor binding or RNAP pausing determines the outcome of Rho-dependent flagellum phase variation in *C. difficile*.

Rho inactivation in *B. subtilis* results in hypersporulation due to loss of an internal *kinB* gene Rho-dependent terminator (48). KinB is a histidine kinase that acts in a phosphorelay to activate Spo0A, the master regulator of sporulation initiation in *B. subtilis* (56). A mechanism involving an indirect effect of Rho via inappropriate readthrough of upstream genes can be eliminated since the genes 5' to the *flgB* operon are σ^D-dependent (24). Interestingly, the

flagellar glycosylation genes 5' of the *flgB* operon may require Rho-factor for transcription termination since 7/14 motile suppressors were hypermotile relative to *recV flg* ON*. Loss of Rho-dependent termination in the hypothetical gene 5' of the *flgB* operon would lead to transcriptional read-through into the *flgB* operon promoting a positive feedback loop. However, overexpression of Rho mutants in the *recV flg* ON* had negligible effects on motility. Our working model is that Rho factor mediates direct transcription termination in the leader RNA of the *flgB* operon in *flg* OFF bacteria. The OFF reporter had significantly higher activity in the MS compared to either phase locked mutant. These data are supported by the recovery of swimming motility observed in the MS. The high activity of the ON reporter in MS strains could be due to increased expression and/or activity of positive regulators of flagellar gene expression. Specifically, we speculate that AgrA and/or SinR, which regulate expression of c-di-GMP diguanylate cyclases and phosphodiesterases (25, 26), indirectly control flagellar gene expression in *flg* ON and OFF bacteria.

Multiple *in vitro* and *in vivo* approaches are underway to distinguish between direct and indirect mechanisms of flagellar phase variation via Rho factor. Targeted mutagenesis of Rho is needed to confirm that Rho relieves inhibition of motility in *flg* OFF bacteria. Treatment with bicyclomycin, a Rho-specific antibiotic (57), or heterologous expression of Psu, a Phage-Derived Rho Inhibitor (58, 59), can resolve any doubts of the contribution of Rho to flagellar phase variation. Transcription termination assays using the leader sequence for the *flgB* operon with the switch in the ON and OFF orientation will help discern a direct role for Rho factor in termination. Should such studies indicate that Rho factor does not directly terminate transcription in the leader RNA of the *flgB* operon, then whole transcriptome studies could identify gene networks dysregulated in a *rho* mutant due to global loss of transcription termination. In

addition, one motile suppressor lacking mutations in *rho* had growth kinetics comparable to either phase-locked mutant. Identifying the gene(s) mutated in this motile suppressor is of high priority, because it may reveal the factor that mediates indirect control of flagellar gene expression by Rho.

Materials and Methods

Bacterial strains and growth conditions

All bacterial strains used in this study are listed in Table 4.3. *C. difficile* was cultured at 37ºC in anaerobic conditions with a gas atmosphere of 90% N_2, 5% CO_2, and 5% H_2 in a Coy chamber. Brain Heart Infusion with 5% yeast extract (BHIS) or Tryptone Yeast (TY) media were used for cultivation of *C. difficile* where indicated. Unless otherwise indicated, *E. coli* was cultured at 37°C under aerobic conditions in Luria Bertani (LB) medium. For selection of plasmids in *E. coli*, 100 μg/mL ampicillin (Amp) and/or 10 μg/mL chloramphenicol (Cm) was used, as indicated. Kanamycin (Kan) 100 μg/ml was used to select against *E. coli* in conjugations with *C. difficile*. For maintenance of plasmids in *C. difficile*, 10 μg/mL thiamphenicol (Tm) was used.

Motility Assays

The *recV flg* OFF* mutant was used to isolate motile suppressors. Motility assays were set-up as previously described (28, 29, 60). Briefly, bacteria were cultured overnight in TY media and diluted 1:50 into fresh BHIS media for growth until reaching an Optical Density (OD) of ~1.0 at 600nm. A 1.5 μL volume of bacterial culture was used to stab inoculate anaerobic

0.3% 0.5X BHIS agar medium. Motility plates were incubated at 37℃ and motile flares were visually detected and enumerated, for frequency estimates, every 24 hours for three days. An asymmetric motile flare coming from the non-motile stab site was indicative of a potential motile suppressor. Using inoculating loops, motile suppressors were recovered from the lead the edge of the swim site and passaged on BHIS agar to isolate single colonies. Individual colonies were used to confirm motility in soft agar plates. For characterization of motile suppressors, the same conditions mentioned previously were used. For strains with plasmids, motility medium contained thiamphenicol at a final concentration of 10 μg/mLSwim diameters were measured every 24 hours as indicated, averaging the diameters from the narrowest and widest edges of the colony. For experiments determining whether a specific rho mutant allele could exert a dominant negative effect in *recV flg* OFF* and *recV flg* ON*, motile flares were enumerated every 24 hours for two days. Single colonies were suspended in 100 μL of 1X PBS and 1.5 μL was used to stab inoculate motility medium.

Whole Genome Sequencing

Genomic DNA was extracted from seven motile suppressors (RT1705 to RT1711 in Table 4.3) and the parental strain *recV flg* OFF* (RT1693) as previously described (61). Genomic DNA was prepared using the KAPA HyperPrep Kit (Roche) and sequenced using an Illumina HiSeq 2500 Rapid Run platform with paired ends and 100X coverage by the UNC-CH High Throughput Genomic Sequencing Facility. Data was processed and sequencing reads were mapped to the reference *C. difficile* R20291 genome (Accession No. FN545816.1) using CLC Genomics Workbench software with default settings. Mutations present exclusively in suppressors, but absent from the reference *recV flg* OFF* strain are presented in Table 4.1.

Western Blots for TcdA

Overnight cultures were grown in TY broth and then diluted 1:50 in fresh TY or BHIS, where indicated. Strains were grown until late stationary phase (OD_{600} 1.8-2.0) for 12 hours, normalized to the same density, and then cells were collected by centrifugation. Bacteria were lysed by suspension in SDS-PAGE buffer and heating to 100ºC for 10 minutes. Samples were electrophoresed on 4-15 % TGX gels (Bio-Rad), then transferred to nitrocellulose membranes. Membranes were stained with Ponceau S (Sigma) to determine equal loading and imaged using the G:Box Chemi imaging system. TcdA was detected using mouse anti-TcdA primary antibodies (Novus Biologicals) followed by goat anti-mouse IgG secondary antibody conjugated to DyLight 800 4X PEG (Invitrogen). Blots were imaged using the Odyssey imaging system (LI-COR). All strains were assayed in at least 2 independent experiments in TY and BHIS media.

Alkaline phosphatase reporter assays

Strains were passaged from glycerol stock and single colonies were used to inoculate TY medium. Overnight cultures were diluted 1:50 into BHIS medium and grown until OD600 ~1.0-1.5. One milliliter of culture was collected by centrifugation, supernatant discarded, and the pellet was stored in -20 ℃ freezer until ready to perform the assay. The AP assay was performed as previously described (28, 62).

Generation of strains

Alkaline phosphatase reporters were cloned from previously described templates using R2282 and R1706 described in Table 4.3 (28). Products for promoterless *phoZ*, P*flgB*::*phoZ*, P*flgB*-5'UTR(FS^{ON})-*flgB*::*phoZ*, and P*flgB*-5'UTR(FS^{OFF})-*flgB*::*phoZ* were digested with NheI and BamHI and ligated into similarly digested pRPF144 (46). *E. coli* transformants were screened with primers R2282 and R1706, and positive cloned were sequenced with the same primers to confirmation. Plasmids were transformed into *E. coli* HB101, which was used in conjugations with the indicated *C. difficile* strains in Table 4.3 after brief heat shock (63). *C. difficile* transconjugants were selected on BHIS with thiamphenicol and kanamycin.

For dominant negative motility assays, constructs bearing a mutated *rho* allele were cloned as described. Briefly, genomic DNA from the 13 motile suppressors with mutations in the *rho* was used as template in a PCR using primers R2308 and R2307 to amplify the *rho* gene (CDR20291_3324) (61). PCR products were digested with EcoRV and BamHI and ligated into pRPF185. pRPF185 was digested with SacI, Klenow treated to generate blunt ends, and digested with BamHI prior to ligation with insert. *E. coli* transformants were grown at 30℃ to hinder additional mutations in *rho* and were screened with R2308 and R2307. Sequence integrity was confirmed using primers R2308, R2307, R2366, and R2367. Plasmids were transformed into HB101 for conjugation with the *C. difficile* strains listed in Table 4.3 (63). *C. difficile* transconjugants were selected on BHIS agar medium with thiamphenicol and kanamycin.
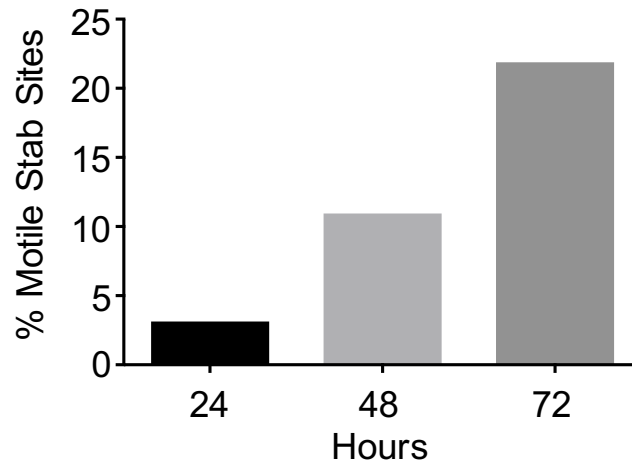
Figure 4.1: Low frequency recovery of motile suppressors from the *recV flg* ON* mutant. The *recV flg* OFF* mutant was inoculated into motility medium, and the presence of motile flares was recorded every 24 hours for three days. Data are expressed as a percentage of motile flares divided by the total number of replicates. Data are combined from two independent experiments with 64 total biological replicates.
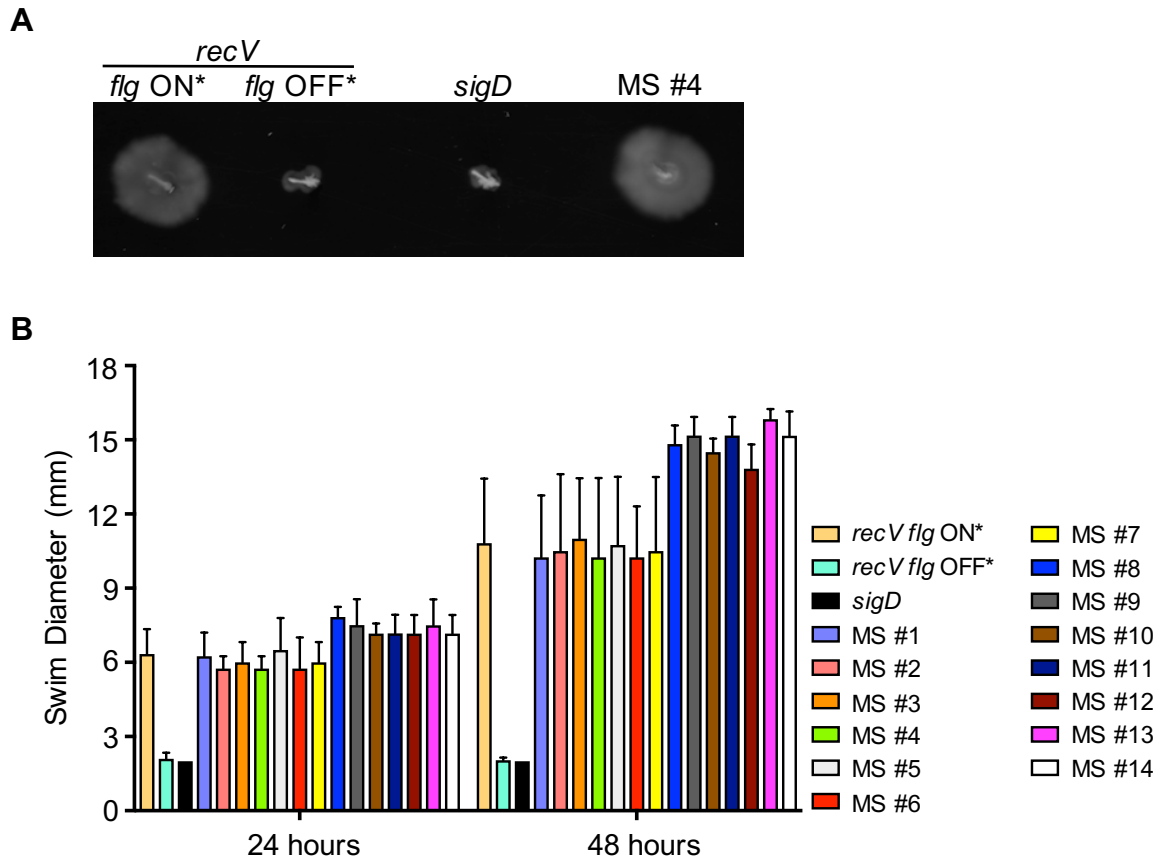
Figure 4.2.: The motile suppressors display restored motility. (A) Representative image of a swimming motility plate at 48 hours with the *recV flg* ON*, *recV flg* OFF*, *sigD*, and MS #4 mutants. (B) Quantification of swim diameter measured at 24 and 48 hours with the *recV flg* ON*, *recV flg* OFF*, *sigD*, and MS #1-14 mutants. Data were pooled from 2 independent experiments with a total of 4-7 biological replicates of each strain.

A

NHB

Insertion Domain

PBS

Linker

Walker motifs
A   B

SBS
Q-loop
R-loop

11 – 51   52 – 167   171 – 223   266 – 514

N-terminal domain (NTD)   C-terminal domain (CTD)

100 AA
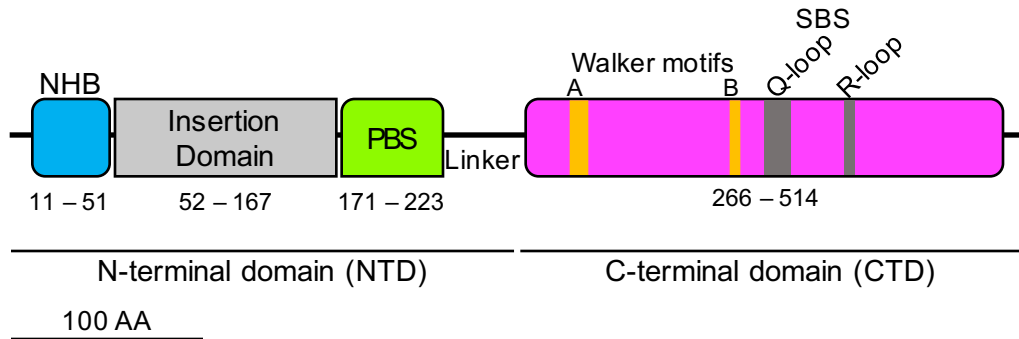
B

```
LILDDKVLEG LGSKTLVELR EIAKELKIKS ITTYKKNELI EIINSKSKNE  50
EKVNEETKEK NEDIKEDIKN RKENIERESI EKEVEYKSPT KSYNKKEIDT 100
QNISQNQENR NQRNDYNKTT SVNDSNKSSN ISNRMVRNNN KNYYMPKQVD 150
ESKIVDEFNT SKEDEVVGVL EILPDGFGFL RGSNYLSTEG DVYVSPSQIR 200
RFNMKTGDKI KGITRHPKSG EKFRALLYVQ KINDENPDTA IQRNAFETLT 250
PIFPEERLTL ETNRNEIATR IIDLISPIGK GQRGLIVAPP KAGKTVLLKS 300
VANSIAKNHP NVELIVLLID ERPEEVTDMK ESIEGDVIYS TFDQVSSHHV 350
KVAEMVLNRA QRLVEHGKDV VILLDSITRL ARAYNLTISP TGRTLSGGID 400
PGALHGPKKF FGAARNIRQG GSLTILGTAL VETGSRMDDV IFEEFKGTGN 450
MELHLDRKLA EKRIFPAIDI YKSGTRRDDL LLDDEEKTAL WRLRREMSNN 500
SVMEITDKVI ELIKRTKDNK EFVKSIKNL
```

Figure 4.3. Schematic and amino acid alignment with features of *C. difficile* Rho factor. (A, B) Schematic of Rho factor with annotated and color-coded domains important for function. N-terminal Helix Bundle (NHB) domain is colored blue and consists of amino acids (AA) 11-51. The Insertion Domain is colored grey and is AA 52-167. The Primary Binding Site (PBS) domain is colored green and is AA 171-223. The C-terminal domain (CTD) has a smaller region of AA 266-514. Walker motif A domain is orange (A) and bolded AA 288-295 in (B). Walker motif B domain is orange (A) and underlined and bolded AA 371-375 in (B). The Q loop domain is colored grey in (A) and is wavy underlined AA 390-402 in (B). The R loop domain is grey in (A) and is AA 431-436 in (B).
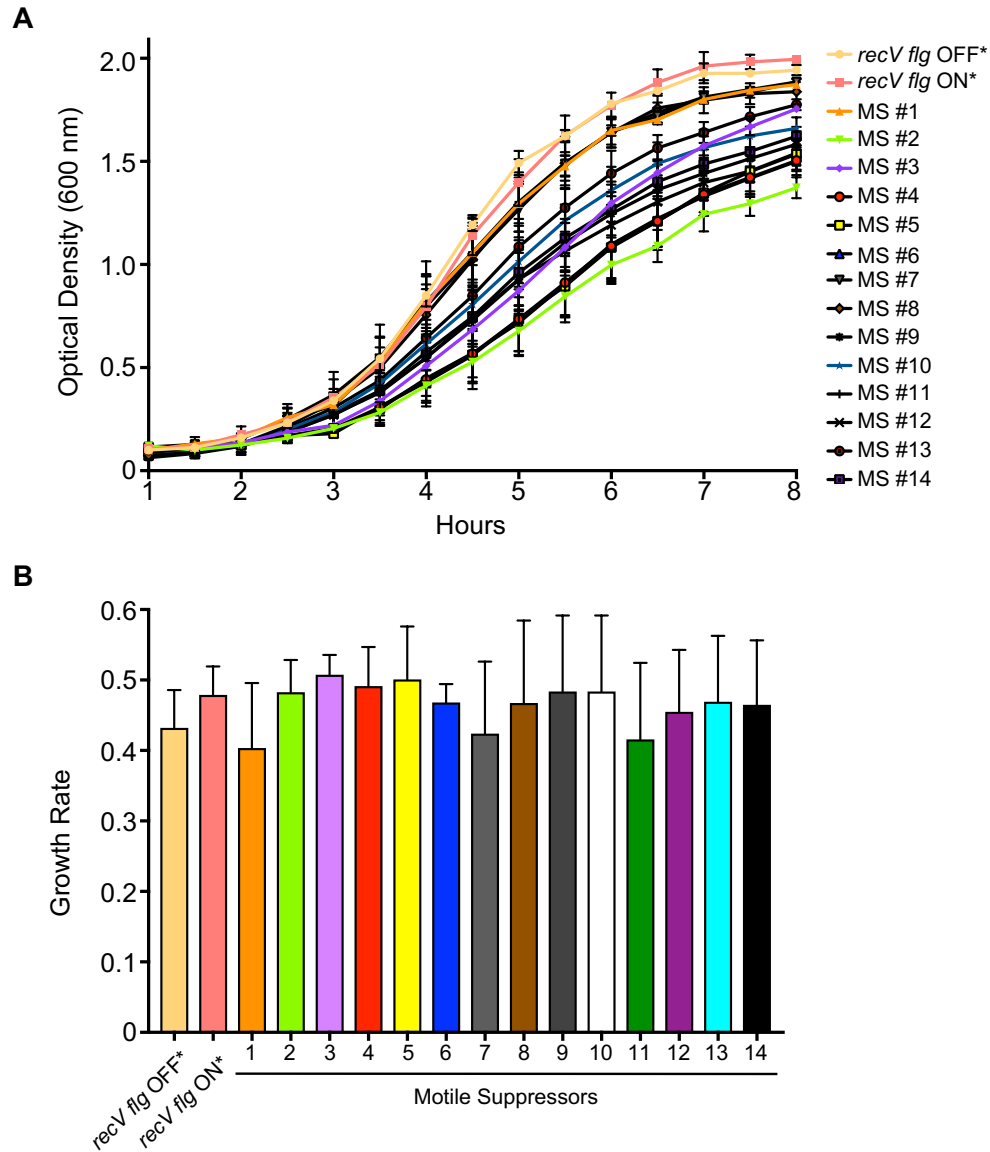
Figure 4.4. Growth curves and rates of motile suppressors. (A) Growth curves of the *recV flg* ON*, *recV flg* OFF*, and MS #1-14 strains in BHIS medium. Data are pooled from two independent experiments with 4-6 biological replicates of each mutant. (B) Growth rates of the *recV flg* ON*, *recV flg* OFF*, and MS #1-14 strains calculated from (A) using time points at 4 and 5.5 with the following equation: $\Delta\ln(OD_{600})/\Delta time$.

A

B
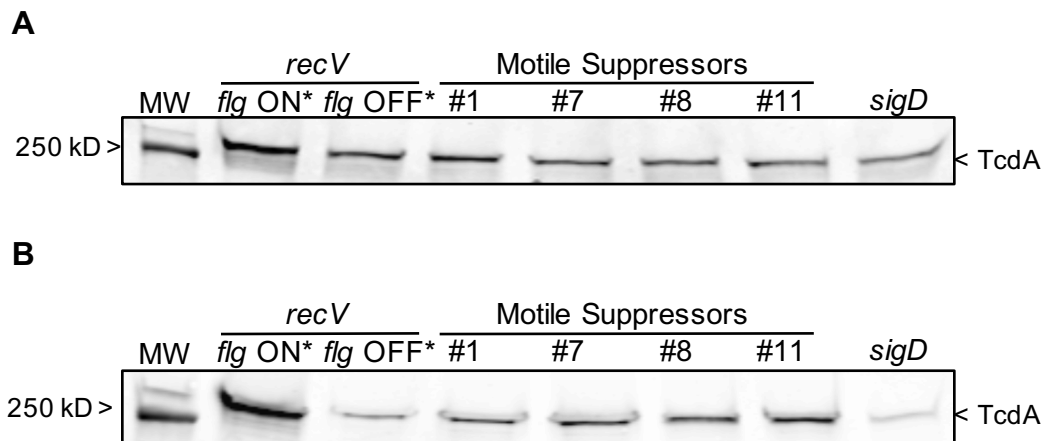
Figure 4.5. Motile suppressors produce TcdA in variable levels depending on growth medium. (A, B) Western blots for TcdA from the cell lysates of the *recV flg* ON*, *recV flg* OFF*, MS #1, MS #7, MS #8, MS #11, and *sigD* mutants in (A) TY medium and (B) BHIS medium after 12 hours of growth. Data in both panels are representative of two independent experiments.
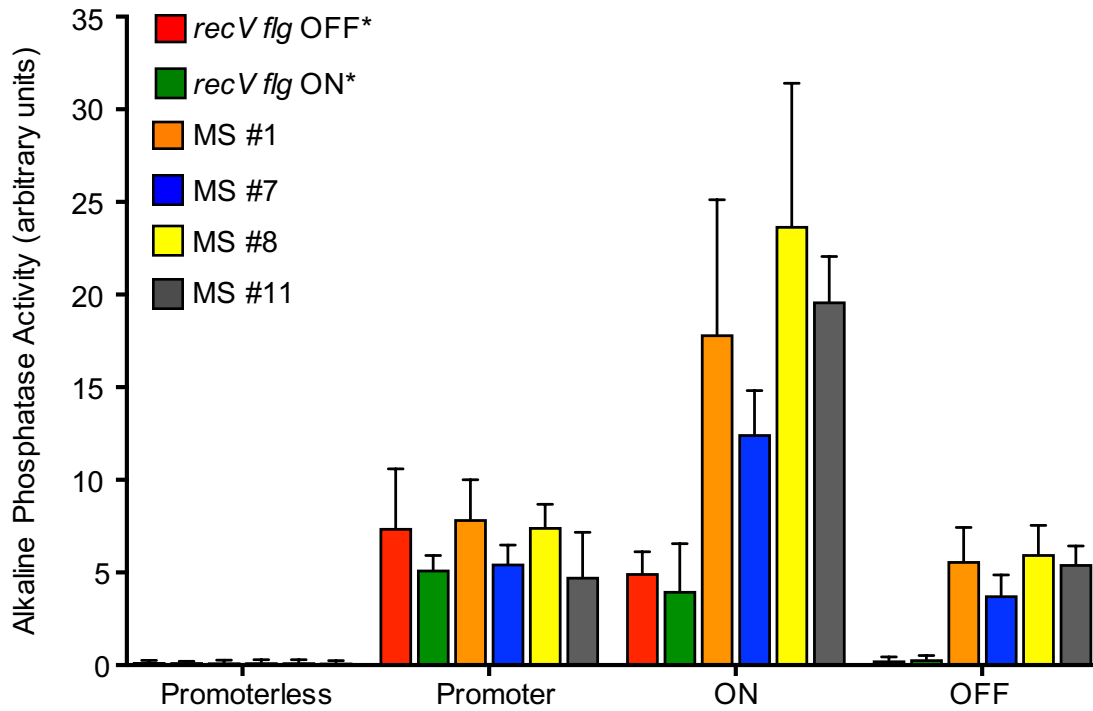
Figure 4.6. Evaluation of alkaline phosphatase reporters in the motile suppressors. Alkaline phosphatase activity for promoterless *phoZ* (Promoterless), $P_{flgB}::phoZ$ (Promoter), P_{flgB}-5'UTR(FS^{ON})-*flgB::phoZ* (ON), and P_{flgB}-5'UTR(FS^{OFF})-*flgB::phoZ* (OFF) reporters in the *recV flg* OFF*, *recV flg* ON*, MS #1, MS #7, MS #8, and MS #11 mutants. Data are pooled from two independent experiments with four biological replicates of each strain.

Table 4.1. Mutations identified from whole genome sequencing of motile suppressors.

Suppressor Strain	Position on R20291 Chromosome	Nucleotide			Locus Tag in CDR20291	Gene Annotation	Notes
		Reference	Alteration	Frequency			
recV flg ON MS #1	3955138	T	C	99.28	3324	rho	N528S
recV flg OFF* MS #2	3955300	C	A	99.13	3324	rho	G474V
	1670842	G	A	74.76	1414 – 1415		
	1726838	A	T	79.58	5' of 1465		
	1726842	T	G	77.68	5' of 1465		
recV flg OFF* MS #3	3955885	C	T	97.66	3324	rho	G279E
	1670842	G	A	57.32	1414 – 1415		
	1726838	A	T	55.11	5' of 1465		
	1726842	T	G	51.76	5' of 1465		
recV flg OFF* MS #4	3955391	C	A	99.05	3324	rho	E444Stop
	1670842	G	A	68.91	1414 – 1415		
	1726838	A	T	74.24	5' of 1465		
	1726842	T	G	71.44	5' of 1465		
recV flg OFF* MS #5	3956539	T	/	89.83	3324	rho	Frameshift
	1670842	G	A	70.43	1414 – 1415		
	1726838	A	T	73.19	5' of 1465		
	1726842	T	G	71.26	5' of 1465		
recV flg OFF* MS #6	3955955	C	A	99.22	3324	rho	E255Stop
	1670842	G	A	74.76	1414 – 1415		
	1726838	A	T	79.58	5' of 1465		
	1726842	T	G	77.68	5' of 1465		
recV flg OFF* MS #7	3956119	T	C	99.02	3324	rho	R201M
	4128329	A	G	90.91	5' of 3478		
recV flg OFF* MS #8	1418	C	T	N/A	3324	rho	S473F
recV flg OFF* MS #9	1212	T	C	N/A	3324	rho	L404P
recV flg OFF* MS #10	322	G	T	N/A	3324	rho	E108Stop
recV flg OFF* MS #11	N/A	N/A	N/A	N/A		rho	N/A
recV flg OFF* MS #12	965	G	T	N/A	3324	rho	R322I
recV flg OFF* MS #13	957	G	A	N/A	3324	rho	D320A
recV flg OFF* MS #14	739	G	T	N/A	3324	rho	E247Stop

Table 4.2. Frequency estimates of dominant negative effects in a motility assay.

rho alleles	Frequency at 24 hours[a]		Frequency at 48 hours[a]	
	0 ATc[b,c]	15 ATc[b,d]	0 ATc[b,c]	15 ATc[b,d]
Wildtype	0.00	0.00	29.17	0.00
MS #1: N528S	91.67	0.00	100.00	100.00
MS #2: G474V	45.83	0.00	100.00	75.00
MS #3: G279E	16.67	0.00	100.00	100.00
MS #4: E444Stop	95.83	16.67	100.00	100.00
MS #5: Frameshift	8.33	0.00	70.83	75.00
MS #6: E255Stop	66.67	0.00	100.00	100.00
MS #7: S473F	87.50	25.00	100.00	100.00
MS #8: L404P	41.67	0.00	100.00	100.00
MS #9: E108Stop	8.33	0.00	100.00	100.00
MS #10	41.67	0.00	100.00	91.67
MS #11: R322I	75.00	100.00	100.00	100.00
MS #12: D320A	16.67	0.00	100.00	100.00
MS #13: E247Stop	91.67	0.00	100.00	100.00
MS #14: S473F	16.67	0.00	100.00	100.00

[a]Frequency is expressed as a percentage of number the motile flares divided by the number of biological replicates tested.
[b]Concentration of ATc added in ng/mL.
[c]For uninduced conditions, wildtype and MS #1-7 frequency estimates were calculated from two independent experiments with 24 biological replicates whereas MS #8-14 was from a single experiment with 12 biological replicates.
[d]For ATc induced, data were from a single experiment with 12 biological replicates for each allele.

Table 4.3. Strains and plasmids used in this study.

Strain/Plasmid	Description/Purpose	Reference
Plasmids		
pRPF185	Anhydrotetracycline (ATc) inducible expression vector; Cm^R/Tm^{10}	(46)
pRT1611	pRPF185 with *gusA* removed (vector control)	(28)
pRPF144	pMTL960 with *gusA* under the transcriptional control of the promoter for *cwp2*	(46)
pRT1824	pMTL960: *::phoZ* (promoterless)	This study
pRT1825	pMTL960: P*flgB::phoZ*	This study
pRT1826	pMTL960: P*flgB*-5'UTR(FS^{ON})-*flgB::phoZ*	This study
pRT1827	pMTL960: P*flgB*-5'UTR(FS^{OFF})-*flgB::phoZ*	This study
pRT1882	pRPF185: *rho* wildtype from RT1693	This study
pRT1883	pRPF185: *rho* allele from RT1705 MS #1	This study
pRT1884	pRPF185: *rho* allele from RT1706 MS #2	This study
pRT1885	pRPF185: *rho* allele from RT1707 MS #3	This study
pRT1886	pRPF185: *rho* allele from RT1708 MS #4	This study
pRT1887	pRPF185: *rho* allele from RT1709 MS #5	This study
pRT1888	pRPF185: *rho* allele from RT1710 MS #6	This study
pRT1889	pRPF185: *rho* allele from RT1711 MS #7	This study
pRT2099	pRPF185: *rho* allele from RT1939 MS #8	This study
pRT2100	pRPF185: *rho* allele from RT1940 MS #9	This study
pRT2101	pRPF185: *rho* allele from RT1941 MS #10	This study
pRT2102	pRPF185: *rho* allele from RT1942 MS #11	This study
pRT2103	pRPF185: *rho* allele from RT1943 MS #12	This study
pRT2104	pRPF185: *rho* allele from RT1944 MS #13	This study
pRT2105	pRPF185: *rho* allele from RT1945 MS #14	This study
***Escherichia coli* Strains**		
DH5α	F- Φ80*lacZ* Δ M15 (*lacZYA-argF*)U169 *recA1 endA1 hsdR*17(rκ -, mκ +) *phoA supE44 thi*-1 *gyrA*96 *relA1* λ⁻ *tonA*	Invitrogen, (64)
HB101(pRK24)	F⁻ *mcrB mrr hsdS20(r_B⁻m_B)recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20* (pRK24)	(65)
***Clostridium difficile* Strains**		
RT1693	R20291 *recV::erm cwpV* OFF*; *flg* OFF*	(36)
RT1694	R20291 *recV::erm cwpV* ON*; *flg* OFF*	(36)
RT1702	R20291 *recV::erm cwpV* OFF*; *flg* ON*	(28)
RT1705	MS #1 from RT1693	This study
RT1706	MS #2 from RT1693	This study
RT1707	MS #3 from RT1693	This study
RT1708	MS #4 from RT1693	This study
RT1709	MS #5 from RT1693	This study
RT1710	MS #6 from RT1693	This study
RT1711	MS #7 from RT1693	This study
RT1939	MS #8 from RT1694	This study
RT1940	MS #9 from RT1694	This study
RT1941	MS #10 from RT1694	This study
RT1942	MS #11 from RT1694	This study
RT1943	MS #12 from RT1694	This study

RT1944	MS #13 from RT1694	This study
RT1945	MS #14 from RT1694	This study
RT1838	RT1693 + pRT1824 (*::phoZ*)	This study
RT1839	RT1693 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1840	RT1693 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1841	RT1693 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1842	RT1702 + pRT1824 (*::phoZ*)	This study
RT1843	RT1702 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1844	RT1702 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1845	RT1702 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1846	RT1705 + pRT1824 (*::phoZ*)	This study
RT1847	RT1705 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1848	RT1705 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1849	RT1705 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1850	RT1711 + pRT1824 (*::phoZ*)	This study
RT1851	RT1711 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1852	RT1711 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1853	RT1711 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1981	RT1939 + pRT1824 (*::phoZ*)	This study
RT1982	RT1939 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1983	RT1939 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1984	RT1939 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1985	RT1942 + pRT1824 (*::phoZ*)	This study
RT1986	RT1942 + pRT1825 (P$_{flgB}$*::phoZ*)	This study
RT1987	RT1942 + pRT1826 (P$_{flgB}$-5'UTR(FS^{ON})-*flgB::phoZ*)	This study
RT1988	RT1942 + pRT1827 (P$_{flgB}$-5'UTR(FS^{OFF})-*flgB::phoZ*)	This study
RT1904	RT1693 + pRT1882	This study
RT1905	RT1693 + pRT1883	This study
RT1906	RT1693 + pRT1884	This study
RT1907	RT1693 + pRT1885	This study
RT1908	RT1693 + pRT1886	This study
RT1909	RT1693 + pRT1887	This study
RT1910	RT1693 + pRT1888	This study
RT1911	RT1693 + pRT1889	This study
RT1912	RT1702 + pRT1882	This study
RT1913	RT1702 + pRT1883	This study
RT1914	RT1702 + pRT1884	This study
RT1915	RT1702 + pRT1885	This study
RT1916	RT1702 + pRT1886	This study
RT1917	RT1702 + pRT1887	This study
RT1918	RT1702 + pRT1888	This study
RT1919	RT1702 + pRT1889	This study

Table 4.4. Primers used in this study.

Primer Name[a]	Sequence (5' → 3')[b]	Lab Annotation
PhoZ_RPF144_F	GAATGCTAGCGAGCTGACTGGGTTGAAG	R2282
*phoZ*_AsymR	GAATGTTAATAAGGTAACCCCTAGCAAAGCTCTTTTCTTC	R1706
R202_rho_F2	CAAGATATCCATGCCTTCAATTTCCAAAATAAG	R2308
R202_rho_R	CAAGGATCCTCCATATTTGCATTTGTATTATTGATT	R2307
CDR20291_3324intF	GGAAACTAATAGAAATGAAATAGCT	R2366
CDR20291_3324intR	AGCTATTTCATTTCTATTAGTTTCC	R2367

REFERENCES

1.   **Theriot CM**, **Young VB**. 2015. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. Annu Rev Microbiol **69**:445–461.

2.   **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

3.   **Kuehne SA**, **Cartman ST**, **Minton NP**. 2011. Both, Toxin A and Toxin B, are important in *Clostridium difficile* infection. Gut Microbes **2**:252–255.

4.   **Kuehne SA**, **Collery MM**, **Kelly ML**, **Cartman ST**, **Cockayne A**, **Minton NP**. 2014. Importance of Toxin A, Toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. J Infect Dis **209**:83–86.

5.   **Carter GP**, **Chakravorty A**, **Pham Nguyen TA**, **Mileto S**, **Schreiber F**, **Li L**, **Howarth P**, **Clare S**, **Cunningham B**, **Sambol SP**, **Cheknis A**, **Figueroa I**, **Johnson S**, **Gerding D**, **Rood JI**, **Dougan G**, **Lawley TD**, **Lyras D**. 2015. Defining the roles of TcdA and TcdB in localized gastrointestinal disease, systemic organ damage, and the host response during *Clostridium difficile* infections. MBio **6**:e00551.

6.   **Aktories K**, **Schwan C**, **Jank T**. 2017. *Clostridium difficile* toxin biology. Annu Rev Microbiol **71**:281–307.

7.   **Chandrasekaran R**, **Lacy DB**. 2017. The role of toxins in *Clostridium difficile* infection. FEMS Microbiol Rev **41**:723–750.

8.   **Nusrat A**, **Eichel-Streiber von C**, **Turner JR**, **Verkade P**, **Madara JL**, **Parkos CA**. 2001. *Clostridium difficile* toxins disrupt epithelial barrier function by altering membrane microdomain localization of tight junction proteins. Infect Immun **69**:1329–1336.

9.   **Kasendra M**, **Barrile R**, **Leuzzi R**, **Soriani M**. 2014. *Clostridium difficile* toxins facilitate bacterial colonization by modulating the fence and gate function of colonic epithelium. J Infect Dis **209**:1095–1104.

10.   **Martin-Verstraete I**, **Peltier J**, **Dupuy B**. 2016. The regulatory networks that control *Clostridium difficile* toxin synthesis. Toxins (Basel) **8**:153.

11.   **Dineen SS**, **Villapakkam AC**, **Nordman JT**, **Sonenshein AL**. 2007. Repression of *Clostridium difficile* toxin gene expression by CodY. Mol Microbiol **66**:206–219.

12.   **Antunes A**, **Martin-Verstraete I**, **Dupuy B**. 2010. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Mol Microbiol **79**:882–899.

13.   **Mani N**, **Dupuy B**. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci USA **98**:5844–5849.

14.   **Girinathan BP**, **Monot M**, **Boyle D**, **McAllister KN**, **Sorg JA**, **Dupuy B**, **Govind R**. 2017. Effect of *tcdR* mutation on sporulation in the epidemic *Clostridium difficile* strain R20291. mSphere **2**:e00383-16.

15.   **Janoir C**. 2016. Virulence factors of *Clostridium difficile* and their role during infection. Anaerobe **37**:13–24.

16.   **Kirk JA**, **Banerji O**, **Fagan RP**. 2017. Characteristics of the *Clostridium difficile* cell envelope and its importance in therapeutics. Microb Biotechnol **10**:76–90.

17.   **Twine SM**, **Reid CW**, **Aubry A**, **McMullin DR**, **Fulton KM**, **Austin J**, **Logan SM**. 2009. Motility and flagellar glycosylation in *Clostridium difficile*. J Bacteriol **191**:7050–7062.

18.   **Stevenson E**, **Minton NP**, **Kuehne SA**. 2015. The role of flagella in *Clostridium difficile* pathogenicity. Trends Microbiol **23**:275–282.

19.   **Dingle TC**, **Mulvey GL**, **Armstrong GD**. 2011. Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect Immun **79**:4061–4067.

20.   **Aubry A**, **Hussack G**, **Chen W**, **KuoLee R**, **Twine SM**, **Fulton KM**, **Foote S**, **Carrillo CD**, **Tanha J**, **Logan SM**. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun **80**:3521–3532.

21.   **Baban ST**, **Kuehne SA**, **Barketi-Klai A**, **Cartman ST**, **Kelly ML**, **Hardie KR**, **Kansau I**, **Collignon A**, **Minton NP**. 2013. The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. PLoS ONE **8**:e73026.

22.   **Mukherjee S**, **Kearns DB**. 2014. The structure and regulation of flagella in *Bacillus subtilis*. Annu Rev Genet **48**:319–340.

23.   **McKee RW**, **Mangalea MR**, **Purcell EB**, **Borchardt EK**, **Tamayo R**. 2013. The second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of *sigD*. J Bacteriol **195**:5174–5185.

24.   **Meouche El I**, **Peltier J**, **Monot M**, **Soutourina O**, **Pestel-Caron M**, **Dupuy B**, **Pons J-L**. 2013. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of *tcdR*. PLoS ONE **8**:e83748.

25.   **Martin MJ**, **Clare S**, **Goulding D**, **Faulds-Pain A**, **Barquist L**, **Browne HP**, **Pettit L**, **Dougan G**, **Lawley TD**, **Wren BW**. 2013. The *agr* locus regulates virulence and colonization genes in *Clostridium difficile* 027. J Bacteriol **195**:3672–3681.

26.   **Girinathan BP**, **Ou J**, **Dupuy B**, **Govind R**. 2018. Pleiotropic roles of *Clostridium difficile* sin locus. PLoS Pathog **14**:e1006940.

27. **Edwards AN**, **Tamayo R**, **McBride SM**. 2016. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. Mol Microbiol **100**:954–971.

28. **Anjuwon-Foster BR**, **Tamayo R**. 2017. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet **13**:e1006701.

29. **Anjuwon-Foster BR**, **Maldonado-Vazquez N**, **Tamayo R**. 2018. Characterization of flagellar and toxin phase variation in *Clostridium difficile* ribotype 012 isolates. bioRxiv 256883.

30. **van der Woude MW**, **Bäumler AJ**. 2004. Phase and antigenic variation in bacteria. Clin Microbiol Rev **17**:581–611.

31. **Abraham JM**, **Freitag CS**, **Clements JR**, **Eisenstein BI**. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc Natl Acad Sci U S A **82**:5724–5727.

32. **Silverman M**, **Zieg J**, **Hilmen M**, **Simon M**. 1979. Phase variation in *Salmonella*: genetic analysis of a recombinational switch. Proc Natl Acad Sci U S A **76**:391–395.

33. **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**, **Severinov K**, **Le Bouguenec C**, **Coppée J-Y**, **Dupuy B**, **Martin-Verstraete I**. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

34. **Emerson JE**, **Reynolds CB**, **Fagan RP**, **Shaw HA**, **Goulding D**, **Fairweather NF**. 2009. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol **74**:541–556.

35. **Reynolds CB**, **Emerson JE**, **la Riva de L**, **Fagan RP**, **Fairweather NF**. 2011. The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog **7**:e1002024.

36. **Sekulovic O**, **Ospina Bedoya M**, **Fivian-Hughes AS**, **Fairweather NF**, **Fortier L-C**. 2015. The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. Mol Microbiol **98**:329–342.

37. **Anjuwon-Foster BR**, **Tamayo R**. 2017. Phase variation of *Clostridium difficile* virulence factors. Gut Microbes 1–8.

38. **Mitra P**, **Ghosh G**, **Hafeezunnisa M**, **Sen R**. 2017. Rho protein: roles and mechanisms. Annu Rev Microbiol **71**:687–709.

39. **Kriner MA**, **Groisman EA**. 2017. RNA secondary structures regulate three steps of Rho-dependent transcription termination within a bacterial mRNA leader. Nucleic Acids Res **45**:631–642.

40.   **Sedlyarova N**, **Shamovsky I**, **Bharati BK**, **Epshtein V**, **Chen J**, **Gottesman S**, **Schroeder R**, **Nudler E**. 2016. sRNA-mediated control of transcription termination in *E. coli*. Cell **167**:111–121.e13.

41.   **Freddolino PL**, **Goodarzi H**, **Tavazoie S**. 2012. Fitness landscape transformation through a single amino acid change in the rho terminator. PLoS Genet **8**:e1002744.

42.   **Quirk PG**, **Dunkley EA**, **Lee P**, **Krulwich TA**. 1993. Identification of a putative *Bacillus subtilis rho* gene. J Bacteriol **175**:647–654.

43.   **Washburn RS**, **Marra A**, **Bryant AP**, **Rosenberg M**, **Gentry DR**. 2001. Rho is not essential for viability or virulence in *Staphylococcus aureus*. Antimicrob Agents Chemother **45**:1099–1103.

44.   **Dembek M**, **Barquist L**, **Boinett CJ**, **Cain AK**, **Mayho M**, **Lawley TD**, **Fairweather NF**, **Fagan RP**. 2015. High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. MBio **6**:e02383.

45.   **Dupuy B**, **Sonenshein AL**. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Mol Microbiol **27**:107–120.

46.   **Fagan RP**, **Fairweather NF**. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286**:27483–27493.

47.   **Oliveira Paiva AM**, **Friggen AH**, **Hossein-Javaheri S**, **Smits WK**. 2016. The signal sequence of the abundant extracellular metalloprotease PPEP-1 can be used to secrete synthetic reporter proteins in *Clostridium difficile*. ACS Synth Biol **5**:1376–1382.

48.   **Bidnenko V**, **Nicolas P**, **Grylak-Mielnicka A**, **Delumeau O**, **Auger S**, **Aucouturier A**, **Guérin C**, **Repoila F**, **Bardowski J**, **Aymerich S**, **Bidnenko E**. 2017. Termination factor Rho: from the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. PLoS Genet **13**:e1006909.

49.   **Yang X**, **Lewis PJ**. 2010. The interaction between RNA polymerase and the elongation factor NusA. RNA Biol **7**:272-275.

50.   **Qayyum MZ**, **Dey D**, **Sen R**. 2016. Transcription elongation factor NusA is a general antagonist of Rho-dependent termination in *Escherichia coli*. J Biol Chem **291**:8090–8108.

51.   **Yakhnin AV**, **Babitzke P**. 2014. NusG/Spt5: are there common functions of this ubiquitous transcription elongation factor? Curr Opin Microbiol **18**:68–71.

52.   **Valabhoju V**, **Agrawal S**, **Sen R**. 2016. Molecular basis of NusG-mediated regulation of rho-dependent transcription termination in bacteria. J Biol Chem **291**:22386–22403.

53.   **Joyce SA**, **Dorman CJ**. 2002. A Rho-dependent phase-variable transcription terminator controls expression of the FimE recombinase in *Escherichia coli*. Mol

Microbiol **45**:1107–1117.

54. **Lavitola A**, **Bucci C**, **Salvatore P**, **Maresca G**, **Bruni CB**, **Alifano P**. 1999. Intracistronic transcription termination in polysialyltransferase gene (*siaD*) affects phase variation in *Neisseria meningitidis*. Mol Microbiol **33**:119–127.

55. **Konan KV**, **Yanofsky C**. 2000. Rho-dependent transcription termination in the tna operon of *Escherichia coli*: roles of the *boxA* sequence and the *rut* site. J Bacteriol **182**:3981–3988.

56. **Tan IS**, **Ramamurthi KS**. 2014. Spore formation in *Bacillus subtilis*. Environ Microbiol Rep **6**:212–225.

57. **Zwiefka A**, **Kohn H**, **Biochemistry WW**. 1993. Transcription termination factor Rho: the site of Bicyclomycin inhibition in *Escherichia coli*. Biochemistry **32**:3564-2570.

58. **Sauer B, Ow O, Ling L, Calendar R.**1981. Mutants of satellite bacteriophage P4 that are defective in the suppression of transcriptional polarity. J. Mol. Biol. **145**:29–46.

59. **Ranjan A**, **Banerjee R**, **Pani B**, **Sen U**, **Sen R**. 2013. The moonlighting function of bacteriophage P4 capsid protein, Psu, as a transcription antiterminator. Bacteriophage **3**:e25657.

60. **Purcell EB**, **McKee RW**, **McBride SM**, **Waters CM**, **Tamayo R**. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol **194**:3307–3316.

61. **Bouillaut L**, **McBride SM**, **Sorg JA**. 2005. Genetic Manipulation of *Clostridium difficile*. Curr Protoc Microbiol **20**:9A2.1-9A2.17.

62. **Edwards AN**, **Pascual RA**, **Childress KO**, **Nawrocki KL**, **Woods EC**, **McBride SM**. 2015. An alkaline phosphatase reporter for use in *Clostridium difficile*. Anaerobe **32C**:98–104.

63. **Kirk JA**, **Fagan RP**. 2016. Heat shock increases conjugation efficiency in *Clostridium difficile*. Anaerobe **42**:1–5.

64. **Hanahan D**. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.

65. **McBride SM**, **Sonenshein AL**. 2011. Identification of a genetic locus responsible for antimicrobial peptide resistance in *Clostridium difficile*. Infect Immun **79**:167–176.

CHAPTER 5: DISCUSSION[4]

Discovery of flagellum and toxin phase variation in *C. difficile*

Clostridium difficile infections (CDI) are a cause of significant morbidity and mortality in industrialized countries. Diarrheal disease caused by CDI ranges from mild inflammatory diarrhea to pseudomembranous colitis, a severe condition characterized by pathologic lesions on the mucosal surface of colonic tissue. Antibiotic treatment is the leading risk factor for CDI, because the antibiotics disrupt the intestinal microbial community that is usually protective against CDI (1). In response to certain bile salts in the intestine, *C. difficile* spores germinate into actively growing vegetative cells (2, 3). The vegetative bacteria secrete the protein toxins TcdA and TcdB, which are largely responsible for the inflammation, intestinal pathology, and diarrheal disease symptoms seen in CDI (4).

C. difficile surface proteins mediate adherence to other microbial species, mucus, and intestinal cells within the colon for colonization (5). The peritrichous flagella produced by *C. difficile* aid in motility and modulate colonization in an animal model of infection (6). In addition, TcdA and TcdB production is linked to flagellum biosynthesis (7-10). SigD (σD), an alternative sigma factor encoded within the early stage flagellar gene operon (*flgB* operon), is

---

[4] Parts of Chapter 5 were published in a manuscript that previously appeared as an autocommentary in *Gut Microbes*. I co-wrote the manuscript with Dr. Tamayo. This is the authors' version of the work. The full citation of the definitive version is: Anjuwon-Foster BR, Tamayo R. 2017. Phase variation of *Clostridium difficile* virulence factors. *Gut Microbes* 1–8, https://doi.org/10.1080/19490976.2017.1362526

essential for transcription of late stage flagellar genes, and also positively regulates transcription of the toxin genes (Fig. 5.1) (11, 12). Therefore, factors that regulate expression of the *flgB* operon not only control bacterial motility, they are also likely to impact toxin production and therefore virulence of *C. difficile*. The 5' untranslated region (UTR) of the *flgB* operon mRNA contains a riboswitch (Cd1) specific to the nucleotide second messenger cyclic diguanylate (c-di-GMP) (13, 14). C-di-GMP binding to Cd1 causes premature transcription termination within the first 160 nucleotides of the 5' UTR of the *flgB* operon mRNA, inhibiting flagellar gene expression, motility, and toxin production. Because the *flgB* 5' UTR is 498 nucleotides, we postulated that an additional *cis*-acting regulatory element downstream of Cd1 could control flagellar and toxin gene expression.

In Chapter 2, we determined that flagellum and toxin production is subject to phase variation via DNA inversion in *C. difficile* R20291, a representative 027 strain isolated in 2006 from an epidemic of CDI in the U.K (15, 16). Between Cd1 and the first open reading frame of the *flgB* operon lies a 154 bp invertible DNA element flanked by 21 bp imperfect inverted repeats (Fig. 5.1). We termed the invertible element the "flagellar switch." Based on genomes currently available on public databases, the flagellar switch sequence and flanking inverted repeats are conserved in all *C. difficile* genomes with flagellar biosynthesis genes. The flagellar switch is capable of inversion in at least three ribotypes (i.e., 027, 017, and 012), under the conditions tested.

Our study in Chapter 2 showed that *C. difficile* with the flagellar switch oriented according to the R20291 published genome (FN545816.1) expresses and produces peritrichous flagella, engages in swimming motility, and secretes the glucosylating toxins, and is thus flagellar phase ON ("*flg* ON"). Conversely, bacteria with the sequence oriented in the inverse

orientation are non-flagellated, non-motile, and attenuated for toxin secretion, and are thus flagellar phase OFF ("*flg* OFF"). The tyrosine recombinase RecV was determined to catalyze inversion of the flagellar switch, and mutation of *recV* in *C. difficile* showed that RecV is required for flagellar switch inversion. RecV was previously shown to control inversion of a genetic switch upstream of *cwpV*, which encodes a cell wall protein involved in autoaggregation and resistance to certain bacteriophage (17-19). Interestingly, the flagellar and *cwpV* switches have seemingly disparate sequences in the inverted repeats and surrounding DNA.

Strain Dependent Differences in Flagellar Phase Variation

The *C. difficile* strain 630 (ribotype 012) was isolated in 1982 from a patient with severe CDI in Switzerland (20). Erythromycin-sensitive derivatives of 630, 630Δ*erm* and 630E, are more amenable to currently available genetic tools and used most often to study *C. difficile* physiology and virulence (21, 22). However, unlike the R20291 strain in which we could detect the flagellar switch in both ON and OFF orientations, only the ON orientation was detected in 630Δ*erm* and in the 630 parent (15). These results suggest that the flagellar switch is locked in these strains. The inability of the flagellar switch to invert to the *flg* OFF orientation in 630 and 630Δ*erm* could be due to the shorter inverted repeats flanking the flagellar switches in these strains, with 20 bp instead of the 21 bp evident in all of the other available published genomes of *C. difficile* strains with a flagellar switch. Inverted repeat length has been demonstrated to affect recombination at the *cwpV* switch in several *C. difficile* strains (17). Reduced inverted repeat length could similarly prevent inversion of the flagellar switch in 630 and 630*Δerm*.

Recent work suggests that the flagellar switch in 630 is capable of inversion from *flg* ON to OFF in some condition(s). Collery *et al.* found the flagellar switch in the OFF orientation in

JIR8094 (also referred to as 630E), another derivative of 630 (23). To obtain JIR8094, *C. difficile* 630 was serially passaged an undisclosed number of times on non-selective agar medium (22). In the process, JIR8094 acquired a non-motile and less toxigenic phenotype. We previously reported that ectopic expression of *sigD* in JIR8094 was sufficient to rescue toxin production, consistent with mutations that affect expression of *sigD* in the *flgB* operon (11). Collery *et al.* inverted the flagellar switch in JIR8094 to that in 630 and 630*Δerm*, which would presumably restore flagellar motility and toxin production (23, 24). However, the JIR8094 mutant strain remained non-motile and attenuated for production of both toxins (23). The lack of restored motility and toxinogenesis may be due to inversion of only 150 bp of the flagellar switch, whereas we experimentally determined that 154 bp comprise the flagellar switch (15). Alternatively, several other genetic differences between 630E and its 630 parent have been identified and may account for abrogated motility and toxin production in 630E strain.

In Chapter 3, we assessed the frequency and phenotypic outcomes of flagellar switch inversion in multiple *C. difficile* ribotype 012 isolates (16). The laboratory-adapted strain JIR8094 and six clinical and environmental isolates were all found to be *flg* OFF: non-motile and attenuated for toxin production. We isolated low frequency motile derivatives of JIR8094 with partial recovery of motility and toxin production. We speculated that an additional accumulated mutation in JIR8094 accounted for the partial restoration. Indeed, overexpression of a topoisomerase gene that was mutated in JIR8094 fully restored motility to 630 levels in the JIR8094 motile derivatives. However, TcdA levels were unchanged in JIR8094 motile derivatives overexpressing the topoisomerase relative to a vector control strain. Repairing the additional identified genetic polymorphisms in JIR8094 (23), such as those in genes encoding an RNA helicase or oligopeptide transporter, may fully restore toxin biosynthesis.

The clinical and environmental isolates varied considerably in the frequency in which *flg* ON derivatives arose, and these derivatives showed full restoration of motility and toxin production. We observed heterogeneity in the size of inverted repeats flanking the flagellar switch among clinical and environmental isolates. However, there was no correlation between inverted repeat size and frequency of *flg* ON derivatives, suggesting that inverted repeat size does not influence recombination. Ultimately, studying the differences between 630 and JIR8094 may help reveal the mechanism by which the orientation of the flagellar switch modulates downstream gene expression. In subsequent sections we discuss additional implications of our findings based on this study.

Wildtype strains: Fake News or False Alarms?

An overarching approach used in basic research of bacterial pathogens is linking genotype to phenotype. Does a bacterial gene product contribute to fitness under these conditions, and will genetic complementation of a mutant *in trans* restore fitness? Mucosal bacterial pathogens that use diverse mechanism to mediate the ON/OFF production of surface or secreted virulence factors can hamper research progress. As presented in Chapter 3, laboratory passage of *C. difficile* strain 630 lead to the isolation of not only a genetically amenable derivative named JIR8094, but a flagellar and toxin phase OFF isolate. Given its amenability to mutagenesis, JIR8094 has been used in many studies to identify and characterize virulence gene regulators (25, 26), sporulation genes (27, 28), and cell wall modifications genes (29). Early work using JIR8094 suggested that TcdA was dispensable for diarrheal disease in hamsters (30), but this outcome was most likely because the flagellar switch orientation (*flg* OFF) masked the contribution of TcdA. However, not all studies that used JIR8094 revealed phenotypes that were

significantly different than when studied in 630Δ*erm*. For example, two major studies co-released in *PLOS Genetics* assessed the contribution of predicted sporulation-specific sigma factors to each stage in sporulation (27, 28). Both independent studies yielded overwhelmingly similar results with one paper using JIR8094 (27) and the other 630Δ*erm* (28).

Misattribution of a phenotype or over/underestimation of a virulence factor has been demonstrated previously in other mucosal pathogens. Anderson *et al.* initially attributed NGO0322, a hydrogen peroxide induced gene based on microarray data, to gonococcal resistance to neutrophil phagocytosis and killing through mutation and complementation (31). However, 11 genes encoding Opa proteins in *Neisseria gonorrhea* are subject to stochastic expression by phase variation and contribute to gonococcal survival in neutrophils (32). Immunoblots for Opa proteins revealed distinct differences in Opa profiles between the *NGO0322* mutant and the parental strain. Ultimately, the Opa profile present in the *NGO0322* mutant, and not the function of *NGO0322* itself, was responsible for gonococcal survival in neutrophils. Phenotypic diversity poses challenges for basic research in linking genotype to phenotype in bacteria that undergo high frequency phase variation of surface proteins that influence host-pathogen interactions.

Genetically phase locked mutants in *Streptococcus pneumoniae* have helped clarify phenotypes, to a degree. Oliver *et al.* generated genetically colony phase-locked opaque and transparent mutants of *Streptococcus pneumoniae* (33). Opaque colonies are associated with higher capsule production, invasive disease, whereas transparent colonies are associated with asymptomatic colonization of the nasal cavity and less invasive disease. *S. pneumoniae* generates six different alleles of a methylase through site-specific recombination and Oliver *et al.* generated genetically locked mutants for all six. Four were phase locked transparent and two

phase locked opaque and exhibited differences in mortality and biofilm formation, but not adherence to epithelial cells (33). Alarmingly, their phase locked alleles generated colony types that were contradictory to previous work using the same strain background (34). The differences could be due to strain passage and SNPs throughout the genome between independent labs. Nonetheless, these studies shed light on the challenges poised when working with chameleon-like pathogens. Research is underway to generate genetically "phase-locked" strains of *C. difficile* to assess the contribution of individual phase variable gene products to virulence. Combinatorial mutants of *recV* with the various genetic switches locked would benefit the *C. difficile* community by clarifying the contribution of individual phase variable gene products to colonization, diarrheal disease symptoms, and transmission during host infection.

Topoisomerase: A Flip & Twist for Phenotypic Diversity

In Chapter 3, we found that inversion of the flagellar switch to *flg* ON in JIR8094 only partially restored motility and toxin production, suggesting that other genetic variations in JIR8094 negatively impact these processes. JIR8094 and derivative strains carry a non-sense mutation in the *topA* gene that truncates the protein (23). Overexpression of *topA*, which encodes a type II topoisomerase, fully restored the motility of the JIR8094 motile derivatives. However, the JIR8094 parent remained non-motile upon overexpression of TopA, suggesting that DNA relaxation is not sufficient to recover motility in *flg* OFF bacteria. Although, under the conditions tested, DNA supercoiling and topoisomerase activity did not influence flagellar switch inversion, previous studies have specified roles for DNA topology and supercoiling in affecting *fimS* inversion in *E. coli*. In laboratory-adapted strains of *E. coli*, Insertion Sequence 2 elements inactivate *fimE*, a recombinase that favors the ON to OFF bias of *fimS*, which results in equal

efficiency of *fimS* inversion by FimB (35). Treatment of FimB-only strains with Novobiocin, a

DNA gyrase-specific antibiotic that relaxes DNA (positive supercoiling), biases cells to be *fimS*

ON (36). Leucine-responsive regulatory protein (Lrp) and Integration Host Factor (IHF) can now

introduce additional bends and kinks into the *fimS* DNA making the switch an unfavorable

substrate for FimB-mediated inversion, so cells are "trapped" in an ON state (37, 38). Mutations

in Lrp and IHF binding sites in *fimS* restore FimB inversion, and cells become biased to OFF

(37, 38). In the *fimS* OFF state, Histone-like nucleoid-structuring protein (H-NS) bends the DNA

to restrict FimB-mediated inversion and additional mutations in H-NS binding sites restore

inversion from OFF-to-ON and ON-to-OFF (38). Collectively, Lrp, IHF, and H-NS function as

recombination directionality factors (RDF) to bias *fimS* in a given orientation for exclusively

fimbrial phase ON and phase OFF bacteria. In the absence of antibiotics, DNA topology in

stationary phase is relaxed and the aforementioned RDFs can affect *fimS* inversion frequency.

Does DNA topology and supercoiling influence flagellar switch inversion in the ribotype 027

and 017 strains? Treatment of cells with a DNA gyrase inhibitor such as a fluoroquinolone

antibiotic, which would increase positive supercoiling as a result of DNA topoisomerase activity,

would allow an assessment of the impact of topology on flagellar switch orientation.

Mechanism of Phase Variable Gene Regulation

Phase variation occurs by multiple mechanisms in diverse mucosal bacterial pathogens,

including site-specific recombination, general homologous recombination, slip strand mispairing,

and DNA methylation (39). For site-specific recombination, a recombinase recognizes a specific

DNA sequence and catalyzes strand exchange and DNA inversion. The orientation of the DNA

sequence dictates whether a nearby gene is expressed. Classically, the invertible DNA element

contains a promoter to transcriptionally control an adjacent gene or operon. The best-studied example of this mechanism is the Type I fimbrial biosynthetic operon in *E. coli*. Here, the invertible element *fimS* lies upstream of the Type I fimbrial operon. Within *fimS* is a promoter that, when properly oriented, drives transcription of *fimA* (40). Using a series of transcriptional reporters and growth conditions known to be permissive for flagellar gene expression, we found that the flagellar switch in *C. difficile* does not contain a promoter, and expression of the *flgB* operon relies on the upstream σ^A-dependent promoter (15).

Phase variable expression of *cwpV* in *C. difficile* occurs through a post-transcriptional mechanism (17). In one orientation (ON) of the *cwpV* switch, the leader region of the *cwpV* mRNA adopts a structure allowing transcriptional elongation into the *cwpV* coding sequence. In the opposite orientation (OFF), the leader region forms an intrinsic terminator that causes premature termination of transcription, precluding transcription of the *cwpV* coding sequence and CwpV biosynthesis. To determine whether a similar *cis*-acting mechanism occurs via the flagellar switch, we evaluated the same transcriptional reporters in *Bacillus subtilis,* a spore-forming obligate aerobe, where we postulated that *C. difficile*-specific *trans*-acting regulatory factors would be absent. In contrast with results from *C. difficile,* we found that the *flg* ON and OFF reporters yielded equivalent activity in *B. subtilis*, indicating that no intrinsic transcription terminator is present in the flagellar switch. Northern blotting failed to detect evidence of premature transcript termination, suggesting that factor-dependent termination also does not occur (15). However, the low sensitivity of northern blotting using a Digoxigenin-labeled probe does not allow us to eliminate the possibility of factor-dependent termination (15).

Interstate 5': The Rho(d) Less Travelled

Data presented in Chapter 4 implicate Rho factor in mediating strand-specific transcription termination in *flg* OFF bacteria. Motile suppressor mutants were isolated from phase-locked *flg* OFF mutants in motility medium. All sequenced motile suppressors harbored missense or nonsense mutations in the *rho* gene resulting in non-functional, unstable, or severely truncated variants of Rho factor. Growth curves and calculated growth rates demonstrated that *rho* is not essential for growth, but does attenuate growth kinetics in nutrient rich media. Why is the Rho mutant attenuated? As demonstrated in other organisms, Rho factor could help resolve R-loops, terminate the timely expression of cryptic prophage genes, and prevent pervasive antisense transcription that is toxic (41). RNA sequencing and CHIP-seq experiments could illuminate a role for Rho factor in these processes in *C. difficile*. All motile suppressors had motility comparable to the phase locked flagellar OFF mutant. The results of gene reporter experiments suggest that Rho factor is constitutively expressed under the conditions tested, since a *flg* OFF reporter was inactive in a phase-locked *flg* ON* mutant. Gene reporter data suggest that Rho factor directly mediates strand-specific termination in the leader RNA of the *flgB* operon of OFF bacteria and/or controls the production of another factor to indirectly control flagellar gene expression in a phase-dependent manner. Experiments are ongoing to distinguish between these possibilities.

Because of the propensity of the phase-locked off strain to restore motility via suppressor mutations in Rho, we favor a model in which Rho factor directly impacts *flgB* operon expression. Based on the literature, we present three working models that coincide with the steps of Rho factor function: binding, translocation, and termination. In the first model, a cytosine-rich *rut* sequence is present exclusively in the *flg* OFF RNA and in the 5' to 3' orientation to recruit Rho

factor via the primary binding site (PBS). A limitation of this model is that the flagellar switch sequence is limited in cytosines and AT-rich. However, Rho function has not been studied in AT-rich species, so it is formally possible that *rut* sites have different sequence features in *C. difficile*. In the second model, the *rut* sequence is present in the cytosine-rich sequence 5' of the flagellar switch, in the Cd1 riboswitch region. This model is predicated on the ability of Insertion Domain present in the NTD of Rho factor to unwind highly structured RNA. To allow phase-specific control of *flg* gene expression, a Rho Antagonizing RNA Element (RARE) sequence present in the *flg* ON strand prevents Rho translocation by preventing closed complex formation. Sevostayanova *et al.* found that a RARE antagonizes Rho-dependent termination in the leader RNA of the *mgtCBR* operon in *Salmonella enterica serovar* Typhimurium (42). Alternative stem loop formation would dictate whether RARE is accessible to restrict Rho activity downstream of binding the *rut* sequence. Furthermore, in this model RARE would be absent from *flg* OFF RNA because the factor is strand-specific or inaccessible by alternate stem loop formation. In the third model, neither availability of the *rut* sequence nor Rho translocation are compromised, but instead a strand-specific pause site is present exclusively in the *flg* OFF state (and a weak or no pause site is present in the *flg* ON state). Either the 5' to 3' orientation of the sequence or the secondary structure would dictate RNAP polymerase accessibility to the pause site. In the leader RNA of *corA* in *S.* Typhimurium, one of two mutually exclusive stem loop conformations not only makes a *rut* sequence accessible, but stimulates a strong RNAP pause site normally repressed by the alternative stem loop conformation (43). These three models coincide with an emerging paradigm that Rho factor can serve not just as a constitutive terminator of transcription at the 3' end of genes and operons, Rho can also mediate regulated termination within leader sequences of mRNAs to control gene expression in *cis* (44).

Numerous other questions remain. What features of the leader RNA of the *flgB* operon are required for Rho function? Does transcription termination require NusG at a weak pause site? Finally, is there a therapeutic benefit of targeting Rho factor to force all bacteria to produce immunostimulatory flagellin and toxins for clearance by immune system? The role of the insertion domain, the bases required for Rho interaction, and the usefulness of Rho as a therapeutic target are areas of future investigation

Rho(llin') on Highly Structured RNA: A Unique Insertion Domain of *C. difficile* Rho factor

Rho factor in *C. difficile* contains the canonical domains present in *E. coli* Rho that contribute to transcription termination in the N- and C-terminal domains, plus one unique feature. In *C. difficile*, Rho factor contains an additional domain, the Insertional Domain, between the N-terminal Helix Bundle (NHB) and Primary Binding Site (PBS) (Figure 4.3). The better characterized Rho proteins from *E. coli* and *B. subtilis* lack an insertional domain, but a few Rho proteins with insertion domains from other species have been characterized (45-48). For example, in *Clostridium botulinum*, a Rho insertion domain forms prion-like proteins when overproduced in bacteria and permissive fungi (47). The function of the insertional domain in *C. difficile* Rho is unknown. Interestingly, an insertion domain in *Micrococcus leutus* Rho factor aids in unwinding of highly structure RNA immediately downstream of a promoter to mediate termination (46). We speculate that the insertion domain of *C. difficile* Rho similarly unwinds the secondary structure of the c-di-GMP sensing riboswitch to access a *rut* sequence. This hypothesis is consistent with our inability to detect termination or loss of gene expression of the *flg* OFF reporter in *B. subtilis* (15). *B. subtilis* Rho does not have the N-terminal insertion domain. Chimeras of *B. subtilis* Rho with the *C. difficile* insertion domain added could elucidate

if this domain is sufficient for *rut* sequence access both *in vitro* and *in vivo*. Mechanistic insight

into the function of the Rho insertion domain in *C. difficile* could inform us of the functional

contribution of this feature in many genetically intractable bacteria (48).

Two Are Better Than One: c-di-GMP and DNA Inversion Regulate Flagellum Biosynthesis

Cyclic diguanylate and phase variation by DNA inversion control flagellum biosynthesis

and toxin production in *C. difficile* (49, 50). C-di-GMP binds to the Cd1 riboswitch in the leader

RNA of the *flgB* operon to cause premature transcription termination at the 160 nt. Mutation of a

single base in the P2 domain of the Cd1 riboswitch renders the riboswitch unable to bind and

respond to c-di-GMP resulting in robust reporter activity, as determined using a reporter fusion

in *B. subtilis* (51). Low c-di-GMP concentrations are permissive for transcription through to the

downstream flagellar switch and flagellar genes, but elevated c-di-GMP leads to termination 89

nt before reaching the LIR of the flagellar switch. Notably, these findings indicate that c-di-GMP

is epistatic to the phase variation mechanism, i.e., phase variation is only relevant under low

intracellular c-di-GMP conditions. Thus, the proposed mechanism for phase variable gene

regulation via Rho is similarly only relevant when c-di-GMP levels are sufficiently low to allow

readthrough of the full leader sequence. Cooperation between a riboswitch and Rho factor within

a leader sequence has been previously demonstrated in *S.* Typhimurium. Holland *et al.*

demonstrated that under high Mg^{2+} concentrations, the Mg^{2+}-sensing riboswitch in the leader

sequence of the *mgtA* mRNA requires Rho factor for termination (52). Specifically, low Mg^{2+}

promotes a secondary structure that favors transcription elongation into the *mgtA* coding

sequence, which functions in Mg^{2+} import. However, in high Mg^{2+} the RNA adopts a relaxed

conformation that exposes a *rut* sequence, allowing Rho to bind and terminate transcription

prematurely, and preventing the import of toxic levels of Mg^{2+}. Treatment with the Rho-specific inhibitor Bicyclomycin abrogates Rho-dependent termination in permissive conditions (high Mg^{2+}). The independent functions of a riboswitch and Rho factor in the same leader mRNA, as seen in the case of the *C. difficile flgB* leader, is new territory.

Environmental signals that elevate intracellular c-di-GMP in *C. difficile* would repress the SigD regulon, which includes genes involved in flagellar motility, toxin production, various adhesins, and other factors (12). Yet increasing c-di-GMP is expected to have further pleotropic effects, such as increased Type IV pilus production and related phenotypes (13, 49, 53). McKee *et al.* found that mutation of the major pilin gene *pilA1* or the pilus ATPase gene *pilB1* reduces *C. difficile* persistence in a mouse infection model (54). Interestingly, *pilA1* expression is positively regulated by a c-di-GMP "ON" riboswitch. These studies suggest that the intracellular concentration of c-di-GMP is elevated in bacterial cells at least transiently during infection. Accordingly, termination through the Cd1 riboswitch is also likely to occur during infection. Mutation of the Cd1 riboswitch and the flagellar switch individually and in combination would inform the individual contribution of these *cis*-acting regulators to *flgB* operon expression. These mutagenesis techniques remain challenging in *C. difficile*, but are feasible using newly available tools (55).

RecV-Dependent Changes in *C. difficile* Colony Morphology

Differences in the expression of cell surface structures can affect how bacteria assemble into a colony, and changes to gross colony morphology can provide insight into microbial physiology and virulence (56). Reynolds *et al.* reported that the phase variable expression of *cwpV* influences colony morphology (18). *C. difficile recV* mutants characterized as CwpV phase

217

ON ("*cwpV* ON") colonies exhibited a densely packed morphology with smooth edges, whereas

a *recV* mutant noted as *cwpV* phase OFF ("*cwpV* OFF") bacteria yield dispersed, ruffled colonies

(18). Other genes have also been implicated in colony morphology development (57, 58). In

pursuit of obtaining enriched flagellar phase variant populations, we observed smooth, circular

(SC) colonies and rough, filamentous (RF) colonies, but colony morphology was not strongly

associated with flagellar switch orientation (15). To identify the genes responsible for the SC and

RF colony morphologies, we evaluated a panel of *C. difficile* mutants. We noted that a *C.*

difficile R20291 *sigD* mutant, which is aflagellate and non-motile, can form both SC and RF

colonies, indicating that colony morphology is independent of flagellum biosynthesis and

motility (Fig. 5.2A). Mutating *recV* resulted in SC colonies in both flagellar phase locked ON

and OFF strains, indicating that the RF colony morphology is independent of flagellar switch

orientation but dependent on RecV (Fig. 5.2B). In our work, in contradiction with the prior

report, the RF colony morphology is independent of the *cwpV* switch orientation because *recV*

mutants with the *cwpV* switch locked ON and OFF both yield SC colonies.

Unpublished data indicates that RecV controls DNA inversion of another switch(es)

controlling phase variable expression of genes mediating SC and RF colony morphology

(personal communication, Ognjen Sekulovic, *et al.*). If so, generating independent *recV* mutants

could give rise to exclusively SC colonies (as observed) or RF colonies depending on the

orientation of such a switch at the time of *recV* inactivation. To determine if RecV modulates

colony morphology, we transformed *C. difficile* from SC and RF colonies with a plasmid

allowing expression of *recV* or the vector control (Figure 5.2C & 5.2D) (15, 59). Bacteria from

SC and RF colonies bearing vector retained their respective colony morphologies, and bacteria

from SC colonies expressing *recV* retained the SC colony morphology. In contrast, bacteria from

RF colonies expressing *recV* yielded SC colonies. From these data, we surmise that RecV mediates recombination at a genetic switch upstream of genes responsible for the RF colony morphology. RecV inactivation in a previous study suggested the *cwpV* ON is responsible for the RF colonies (18), but based on our findings we predict another RecV-dependent switch was locked in their strain and *cwpV* ON was misattributed to RF colonies.

Few studies have evaluated the ultrastructure of *C. difficile* colonies (60). Our data suggest that the production of a surface protein or polysaccharides is subject to RecV-dependent phase variation and impacts colony morphology. There are three known extracellular polysaccharides that could influence colony morphology, named PSI-PSIII, and the export machinery and cell wall proteins that anchor these polysaccharides to the surface could be involved as well (61, 62). Alternatively, RecV may indirectly affect colony morphology by mediating phase variation of a protein controlling the transcription, translation, or localization of a surface protein or polysaccharide. As noted above, two other invertible DNA sequences have been predicted based on the comparison of genomes of multiple *C. difficile* strains (63). These are located upstream of CDR20291_15140 and CDR20291_06850, which encode functional c-di-GMP phosphodiesterases (64, 65). The nucleotide second messenger c-di-GMP has been shown to affect colony morphology in multiple bacteria, and preliminary data suggests c-di-GMP may do so in *C. difficile* (unpublished data, Garrett and Tamayo) (66, 67). The function and virulence contribution of the RecV-dependent phase variable surface proteins or polysaccharides are under investigation.

Virulence contribution of RecV-dependent phase variation during host infection

The role of phase variation of flagella and toxins during *C. difficile* infection remains to be determined. Ideally, such studies would use phase-locked mutants, so that the *inability* to switch between *flg* ON and OFF states can be assessed. A *recV* mutant is phase-locked, but we refrained from evaluating this mutant in animal models of *C. difficile* disease because of the likely pleiotropic effects of the mutation. RecV controls inversion of both the flagellar and *cwpV* switches (15, 17, 18), and impacts at least two other loci (personal communication, Ognjen Sekulovic, *et al.*). Thus, combinatorial genetic switch mutants through inactivation of *recV* would be required to assess the effect of each individual switch on *C. difficile* virulence. An alternative approach for generating phase-locked mutants is site-directed mutagenesis of the inverted repeats, a method successful for phase locking the fimbrial switch in *E. coli* (68), which would allow us to lock the flagellar switch without affecting inversion of other switches.

Gunther *et al.* found that the ability of uropathogenic *Escherichia coli* (UPEC) to phase vary type I fimbriae biosynthesis affects colonization (68). A *fimS* phase-locked OFF mutant was attenuated for colonization in a mouse model of urinary tract infection compared to wild type and *fimS* phase-locked ON strains, consistent with a known role for these fimbriae in UPEC virulence (68). Flagellar filaments contribute to *C. difficile* R20291 colonization during infection of mice (9), so flagellar phase locked *C. difficile* similarly may display colonization phenotypes consistent with the ON/OFF status of the flagellar switch and flagellum biosynthesis. However, the fitness benefit of adherence comes at a cost to *flg* ON bacteria: both flagellin and the glucosylating toxins are immunostimulatory. *C. difficile* flagellin stimulates host Toll-like receptor 5 (TLR5), which activates signaling pathways leading to the secretion of IL-8, a neutrophil chemokine, in epithelial cells (69, 70). The link between flagellum and toxin

220

biosynthesis is also likely to impact *C. difficile* fitness in the intestine. The glucosylating toxins stimulate apoptosis, necrosis, or pyroptosis depending on the host cell type and model of infection (71-74). Batah *et al.* found that flagellin and the toxins synergistically elicit a robust inflammatory response from the intestinal epithelium during infection of mice with R20291, whereas the individual antigens alone were not sufficient for eliciting such a response (75). Thus, while the *flg* ON state may be advantageous for establishing a *C. difficile* infection and disease, the *flg* OFF state may be selected at later stages of infection because it allows avoidance of immune clearance.

The role for CwpV in *C. difficile* in the context of the anaerobic host intestine must also be considered. CwpV promotes bacterial autoaggregation *in vitro*, which was proposed to suggest a role for CwpV in host colonization (18). CwpV also promotes resistance to bacteriophage predation by reducing phage adsorption and phage tail spike DNA injection (19). Thus, *cwpV* ON bacteria could resist phage attack in the intestine. Given the functional contributions of CwpV to *C. difficile in vitro*, the potential contribution of *cwpV* OFF bacteria to CDI is puzzling. Bacteria that are *cwpV* OFF may be less likely to generate an antibody response, and/or they could disperse from an unfavorable colonization site during infection.

Because RecV mediates inversion of multiple sequences, it is tempting to speculate that RecV allows coordination of their production, resulting in synergism or antagonism between those factors. Considering flagella and CwpV together, these may have complementary roles in colonization, with a *cwpV* ON phenotype compensating for a lack of flagella in a *flg* OFF bacterium, and vice versa. One of the antigens may be more immunodominant such that patient infection with another strain could allow for growth of a phase OFF for that antigen. Although antibodies against CwpV can be recovered after infection, CwpV might still be less

221

immunostimulatory compared to flagellin and the toxins (76). In bacteria that are phase ON for both surface structures, the adherence of these bacteria could be greater than individually if they engage distinct receptors. However, in each case, the cost of immunogenicity will play a role in the overall survival of the respective bacteria.

Several other issues complicate the ability to predict the virulence outcome during infection. CwpV could alter flagellar function, although results in *C. difficile* 630 contradict this assertion: motility and flagellin production are unchanged in bacteria overexpressing CwpV (18). The SigD regulon (under the control of the flagellar switch) includes several predicted and functionally characterized adhesins that may synergize with CwpV and/or flagella for colonization (12). Finally, additional RecV-dependent switch(es) may influence *C. difficile* virulence. Importantly, there is no evidence to date that RecV preferentially recognizes one switch orientation over the other for any of its targets, so presumably the abundance of phase ON and OFF of each target in a population relies on external selective pressures.

Translational Impact: Diagnostics & Treatment

Our findings have significant implications for CDI diagnosis and treatment. The methods used for the diagnosis of CDI are a point of contentious debate given the diversity in disease symptoms and tests (77, 78). *C. difficile* can be diagnosed using Nucleic Acid Amplification Tests (NAAT) to detect *C. difficile* rRNA, toxin, or metabolic genes. NAAT have high sensitivity with low specificity. Another approach relies on detection of a specific *C. difficile* product, namely glutamate dehydrogenase (GDH), an intracellular and secreted enzyme that contributes to virulence (79, 80), or the glucosylating toxins. Testing for specific products has relatively low sensitivity, but high specificity. Exclusively using NAAT for diagnosis can lead to

a false positive result for CDI: the patient can be asymptomatically colonized, or remnants of DNA from past exposure may confound the test results. Tests for GDH provide the advantage of stating the patient has an actively growing vegetative population of *C. difficile,* and the toxin immunoassay determines the presence of toxin producing strains. Patients presenting with diarrheal disease symptoms can be subjected to a testing algorithm for *C. difficile* involving detection GDH to denote active replication of the organism, and subsequent NAAT or toxin immunoassay to indicate the presence of virulent *C. difficile*. If negative by GDH, then no other tests are needed. A positive result for the GDH test is followed with a confirmatory toxin immunoassay or NAAT. In many cases, a negative result for the toxin immunoassay arises for a patient who is not currently experiencing diarrheal disease symptoms. We propose that such patients may be colonized with a *flg* OFF isolate. We obtained six ribotype 012 isolates and found that all were *flg* OFF and therefore poor producers of the glucosylating toxin TcdA. TcdA and TcdB levels in such strains could be below the level of sensitivity in a predominately *flg* OFF colonizing *C. difficile*. We suggest labs consider a NAAT evaluating the orientation of the flagellar switch as a prognostic marker of CDI. Conditions that bias the population to *flg* ON would result in increased toxin production, which would become readily detectable by toxin immunoassays and require immediate treatment with metronidazole or vancomycin. Patients with recurrent infection could be dealing with *C. difficile* undergoing switching between *flg* ON with fulminant disease and *flg* OFF with no disease.

For CDI treatment, the new monoclonal antibodies targeting both toxins developed by Merck may be ineffective since the target is subject to phase variable production. Treatment with monoclonal antibodies to the toxins will improve diarrheal disease symptoms, but have no effects on colonization. Carter *et al*. demonstrated that toxin mutants of ribotype 027 strains

maintained colonization in mice at levels similar to parental toxin positive strains for at least 48 hours post-infection. These data suggest that colonization will remain, and these patients could serve as asymptomatic carriers of *C. difficile*. Identifying surface and exported proteins in *C. difficile* that are consistently produced and not subject to phase variable production could serve as better diagnostic and therapeutic targets to ensure targeting of the entire population.

Concluding Remarks

The discovery and characterization of flagellum and toxin phase variation in *C. difficile* has provided new knowledge of a *cis*-acting regulatory feature controlling the production of these virulence factors, but has also left many unanswered questions.

1. How does Rho factor mediate transcription termination in the leader RNA of the *flgB* operon? What features of the leader RNA impart Rho factor recognition? What is the functional contribution of the insertion domain of Rho in comparison to previously characterized Rho proteins? Does Rho factor contribute to regulation at other genetic switches?

2. RecV appears to control multiple switches. What is the phase variable surface protein or polysaccharide controlled by RecV and affecting colony morphology? What is the sequence that RecV recognizes among the switches? Is a single nucleotide deletion in the inverted repeats, as seen in 630E, sufficient to inhibit DNA inversion? Could an accessory factor(s) help RecV discriminate between its targets, leading to different rates of inversion? If so, what is the functional outcome to *C. difficile* physiology?

3. Lastly, the contribution of RecV-mediated phase variation of flagella, toxins, autoaggregation, and phage resistance to virulence in a host has yet to be examined.

Potential interactions between these phenotypes could substantially alter the course of

CDI and transmission to a new host.

Defining the mechanism controlling phase variation of virulence factors in *C. difficile* could

expose new targets for the development of therapeutic agents.
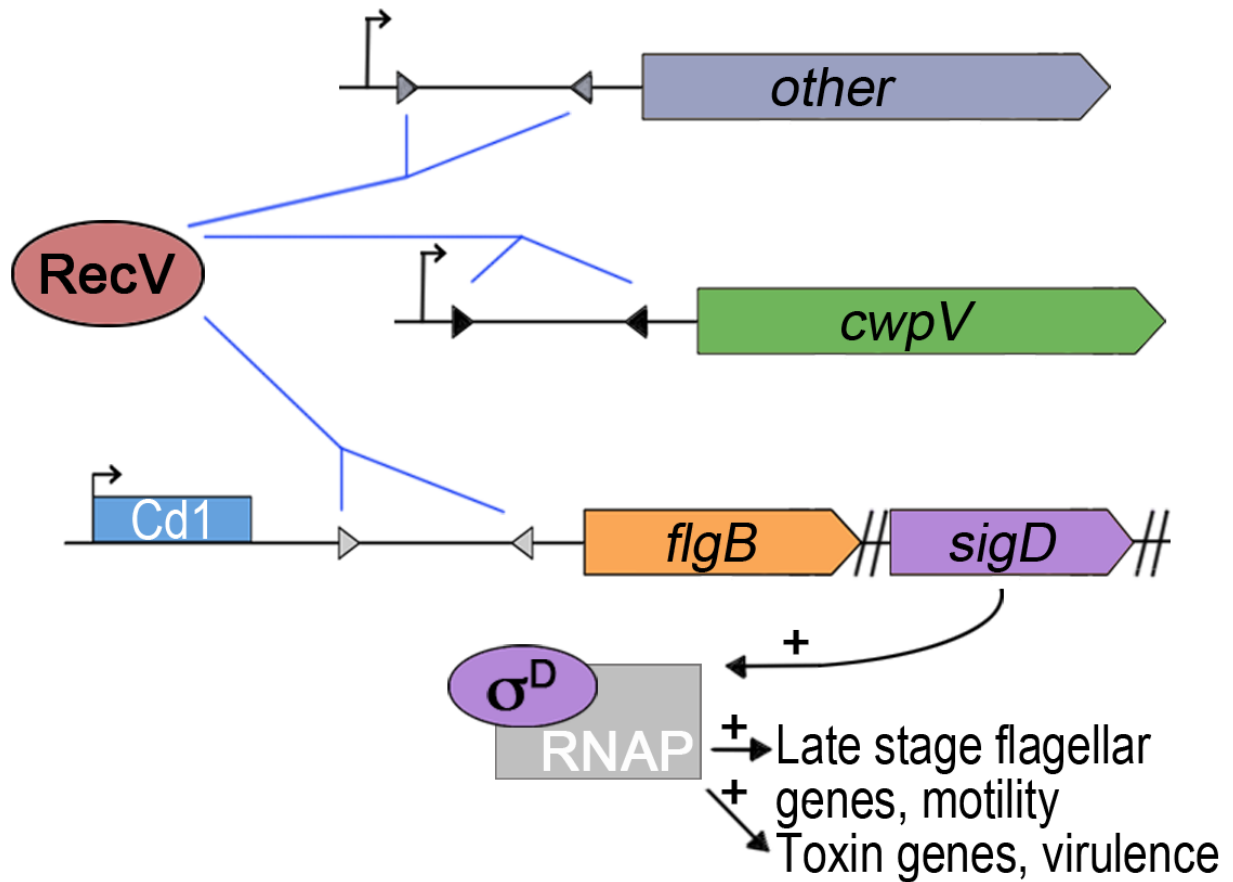
Figure 5.1. Diagram of the regulatory scheme for flagellar and toxin phase variation in *C. difficile*. A DNA invertible element, which we termed the "flagellar switch," is located in the 5' UTR of the *flgB* operon and controls transcription of structural and regulatory genes necessary for flagellum biosynthesis and motility. One orientation of the flagellar switch, but not the other, is permissive for downstream gene expression. In addition, the flagellar switch regulates toxin production by controlling the expression of *sigD*, which located in the *flgB* operon and encodes a sigma factor, σ^D, that promotes toxin gene transcription. RecV, a tyrosine recombinase, catalyzes DNA inversion in both directions at the flagellar and *cwpV* switches, and also impacts one or more unidentified genetic switches that affect colony morphology.
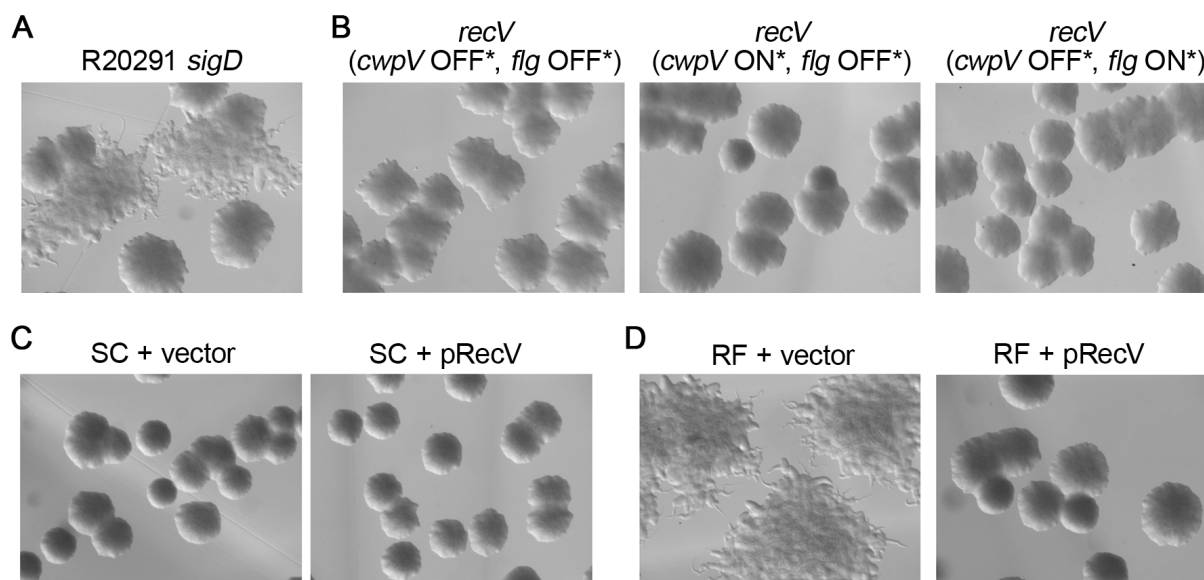
Figure 5.2. RecV controls a genetic switch responsible for the RF morphology in a σ^D-independent manner. Light microscopy images of *C. difficile* R20291 colonies grown on BHIS agar. (A) A R20291 *sigD* mutant develops smooth circular (SC) and rough filamentous (RF) colonies (B) All *recV* mutants yield SC colonies, regardless of flagellar and *cwpV* switch orientations. Switch genotypes are indicated in parentheses, with asterisks denoting phase-locked orientations. (C) Bacteria from SC colonies transformed with a plasmid for ectopic *recV* expression (pRecV, pRT1529) or the vector control (pRT1611) retain the SC colony morphology of the parent isolate. (D) Bacteria from RF colonies transformed with the vector control develop the parental RF colony morphology, while overexpression of *recV* (pRecV) results in conversion to the SC morphology.

REFERENCES

1. **Theriot CM**, **Young VB**. 2015. Interactions between the gastrointestinal microbiome and *Clostridium difficile*. Annu Rev Microbiol **69**:445–461.

2. **Sorg JA**, **Sonenshein AL**. 2008. Bile salts and glycine as cogerminants for *Clostridium difficile* spores. J Bacteriol **190**:2505–2512.

3. **Paredes-Sabja D**, **Shen A**, **Sorg JA**. 2014. *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. Trends Microbiol **22**:406-416.

4. **Pruitt RN**, **Lacy DB**. 2012. Toward a structural understanding of *Clostridium difficile* toxins A and B. Front Cell Inf Microbio **2**:28.

5. **Janoir C**. 2016. Virulence factors of *Clostridium difficile* and their role during infection. Anaerobe **37**:13–24.

6. **Stevenson E**, **Minton NP**, **Kuehne SA**. 2015. The role of flagella in *Clostridium difficile* pathogenicity. Trends Microbiol **23**:275–282.

7. **Dingle TC**, **Mulvey GL**, **Armstrong GD**. 2011. Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect Immun **79**:4061–4067.

8. **Aubry A**, **Hussack G**, **Chen W**, **KuoLee R**, **Twine SM**, **Fulton KM**, **Foote S**, **Carrillo CD**, **Tanha J**, **Logan SM**. 2012. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun **80**:3521–3532.

9. **Baban ST**, **Kuehne SA**, **Barketi-Klai A**, **Cartman ST**, **Kelly ML**, **Hardie KR**, **Kansau I**, **Collignon A**, **Minton NP**. 2013. The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. PLoS ONE **8**:e73026.

10. **Barketi-Klai A**, **Monot M**, **Hoys S**, **Lambert-Bordes S**, **Kuehne SA**, **Minton N**, **Collignon A**, **Dupuy B**, **Kansau I**. 2014. The flagellin FliC of *Clostridium difficile* is responsible for pleiotropic gene regulation during *in vivo* infection. PLoS ONE **9**:e96876.

11. **McKee RW**, **Mangalea MR**, **Purcell EB**, **Borchardt EK**, **Tamayo R**. 2013. The second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of *sigD*. J Bacteriol **195**:5174–5185.

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

13. **Bordeleau E**, **Purcell EB**, **Lafontaine DA**, **Fortier L-C**, **Tamayo R**, **Burrus V**. 2015. Cyclic di-GMP riboswitch-regulated Type IV pili contribute to aggregation of *Clostridium difficile*. J Bacteriol **197**:819–832.

14. **Purcell EB**, **Tamayo R**. 2016. Cyclic diguanylate signaling in Gram-positive bacteria. FEMS Microbiol Rev **40**:753–773.

15. **Anjuwon-Foster BR**, **Tamayo R**. 2017. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet **13**:e1006701.

16. **Anjuwon-Foster BR**, **Maldonado-Vazquez N**, **Tamayo R**. 2018. Characterization of flagellar and toxin phase variation in *Clostridium difficile* ribotype 012 isolates. bioRxiv 256883.

17. **Emerson JE**, **Reynolds CB**, **Fagan RP**, **Shaw HA**, **Goulding D**, **Fairweather NF**. 2009. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol **74**:541–556.

18. **Reynolds CB**, **Emerson JE**, **la Riva de L**, **Fagan RP**, **Fairweather NF**. 2011. The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. PLoS Pathog **7**:e1002024.

19. **Sekulovic O**, **Ospina Bedoya M**, **Fivian-Hughes AS**, **Fairweather NF**, **Fortier L-C**. 2015. The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. Mol Microbiol **98**:329–342.

20. **Sebaihia M**, **Wren BW**, **Mullany P**, **Fairweather NF**, **Minton N**, **Stabler R**, **Thomson NR**, **Roberts AP**, **Cerdeño-Tárraga AM**, **Wang H**, **Holden MT**, **Wright A**, **Churcher C**, **Quail MA**, **Baker S**, **Bason N**, **Brooks K**, **Chillingworth T**, **Cronin A**, **Davis P**, **Dowd L**, **Fraser A**, **Feltwell T**, **Hance Z**, **Holroyd S**, **Jagels K**, **Moule S**, **Mungall K**, **Price C**, **Rabbinowitsch E**, **Sharp S**, **Simmonds M**, **Stevens K**, **Unwin L**, **Whithead S**, **Dupuy B**, **Dougan G**, **Barrell B**, **Parkhill J**. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet **38**:779–786.

21. **Hussain HA**, **Roberts AP**, **Mullany P**. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δ*erm*) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. J Med Microbiol **54**:137–141.

22. **O'Connor JR**, **Lyras D**, **Farrow KA**, **Adams V**, **Powell DR**, **Hinds J**, **Cheung JK**, **Rood JI**. 2006. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol **61**:1335–1351.

23. **Collery MM**, **Kuehne SA**, **McBride SM**, **Kelly ML**, **Monot M**, **Cockayne A**, **Dupuy B**, **Minton NP**. 2016. What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and derivatives. Virulence 1–15.

24. **Ng YK**, **Ehsaan M**, **Philip S**, **Collery MM**, **Janoir C**, **Collignon A**, **Cartman ST**, **Minton NP**. 2013. Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using *pyrE* alleles. PLoS ONE **8**:e56051.

25. **Antunes A**, **Martin-Verstraete I**, **Dupuy B**. 2010. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Mol Microbiol **79**:882–899.

26. **Saujet L**, **Monot M**, **Dupuy B**, **Soutourina O**, **Martin-Verstraete I**. 2011. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol **193**:3186–3196.

27. **Fimlaid KA**, **Bond JP**, **Schutz KC**, **Putnam EE**, **Leung JM**, **Lawley TD**, **Shen A**. 2013. Global analysis of the sporulation pathway of *Clostridium difficile*. PLoS Genet **9**:e1003660.

28. **Saujet L**, **Pereira FC**, **Serrano M**, **Soutourina O**, **Monot M**, **Shelyakin PV**, **Gelfand MS**, **Dupuy B**, **Henriques AO**, **Martin-Verstraete I**. 2013. Genome-wide analysis of cell type-specific gene transcription during spore formation in *Clostridium difficile*. PLoS Genet **9**:e1003756.

29. **Ho TD**, **Williams KB**, **Chen Y**, **Helm RF**, **Popham DL**, **Ellermeier CD**. 2014. *Clostridium difficile* Extracytoplasmic function σ factor σV regulates lysozyme resistance and is necessary for pathogenesis in the hamster model of infection. Infect Immun **82**:2345–2355.

30. **Lyras D**, **O'Connor JR**, **Howarth PM**, **Sambol SP**, **Carter GP**, **Phumoonna T**, **Poon R**, **Adams V**, **Vedantam G**, **Johnson S**, **Gerding DN**, **Rood JI**. 2009. Toxin B is essential for virulence of *Clostridium difficile*. Nature **458**:1176–1179.

31. **Anderson MT**, **Seifert HS**. 2013. Phase variation leads to the misidentification of a *Neisseria gonorrhoeae* virulence gene. PLoS ONE **8**:e72183.

32. **Quillin SJ**, **Seifert HS**. 2018. *Neisseria gonorrhoeae* host adaptation and pathogenesis. Nat Rev Micro **16**:226–240.

33. **Oliver MB**, **Basu Roy A**, **Kumar R**, **Lefkowitz EJ**, **Swords WE**. 2017. *Streptococcus pneumoniae* TIGR4 phase-locked opacity variants differ in virulence phenotypes. mSphere **2**:e00386-17.

34. **Li J**, **Li J-W**, **Feng Z**, **Wang J**, **An H**, **Liu Y**, **Wang Y**, **Wang K**, **Zhang X**, **Miao Z**, **Liang W**, **Sebra R**, **Wang G**, **Wang W-C**, **Zhang J-R**. 2016. Epigenetic switch driven by DNA inversions dictates phase variation in *Streptococcus pneumoniae*. PLoS Pathog **12**:e1005762.

35. **Blomfield IC**, **McClain MS**, **Princ JA**, **Calie PJ**, **Eisenstein BI**. 1991. Type 1 fimbriation and *fimE* mutants of *Escherichia coli* K-12. J Bacteriol **173**:5298–5307.

36. **Dove SL**, **Dorman CJ**. 1994. The site-specific recombination system regulating expression of the Type 1 fimbrial subunit gene of *Escherichia coli* is sensitive to changes in DNA supercoiling. Mol Microbiol **14**:975–988.

37. **Kelly A**, **Conway C**, **O Cróinín T**, **Smith SGJ**, **Dorman CJ**. 2006. DNA supercoiling

and the Lrp protein determine the directionality of fim switch DNA inversion in *Escherichia coli* K-12. J Bacteriol **188**:5356–5363.

38. **Corcoran CP**, **Dorman CJ**. 2009. DNA relaxation-dependent phase biasing of the fim genetic switch in *Escherichia coli* depends on the interplay of H-NS, IHF and LRP. Mol Microbiol **74**:1071–1082.

39. **van der Woude MW**, **Bäumler AJ**. 2004. Phase and antigenic variation in bacteria. Clin Microbiol Rev **17**:581–611.

40. **Abraham JM**, **Freitag CS**, **Clements JR**, **Eisenstein BI**. 1985. An invertible element of DNA controls phase variation of Type 1 fimbriae of *Escherichia coli*. Proc Natl Acad Sci U S A **82**:5724–5727.

41. **Mitra P**, **Ghosh G**, **Hafeezunnisa M**, **Sen R**. 2017. Rho protein: roles and mechanisms. Annu Rev Microbiol **71**:687–709.

42. **Sevostyanova A**, **Groisman EA**. 2015. An RNA motif advances transcription by preventing Rho-dependent termination. Proc Natl Acad Sci USA **112**:E6835–43.

43. **Kriner MA**, **Groisman EA**. 2017. RNA secondary structures regulate three steps of Rho-dependent transcription termination within a bacterial mRNA leader. Nucleic Acids Res **45**:631–642.

44. **Kriner MA**, **Sevostyanova A**, **Groisman EA**. 2016. Learning from the leaders: gene regulation by the transcription termination factor Rho. Trends in Biochem Sci **41**:690–699.

45. **Mitra A**, **Misquitta R**, **Nagaraja V**. 2014. *Mycobacterium tuberculosis* Rho is an NTPase with distinct kinetic properties and a novel RNA-binding subdomain. PLoS ONE **9**:e107474.

46. **Nowatzke WL**, **Burns CM**, **Richardson JP**. 1997. Function of the novel subdomain in the RNA binding domain of transcription termination factor Rho from *Micrococcus luteus*. J Biol Chem **272**:2207–2211.

47. **Yuan AH**, **Hochschild A**. 2017. A bacterial global regulator forms a prion. Science **355**:198–201.

48. **D'Heygère F**, **Rabhi M**, **Boudvillain M**. 2013. Phyletic distribution and conservation of the bacterial transcription termination factor Rho. Microbiology (Reading, Engl) **159**:1423–1436.

49. **Purcell EB**, **McKee RW**, **McBride SM**, **Waters CM**, **Tamayo R**. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol **194**:3307–3316.

50. **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**,

Severinov K, **Le Bouguenec C**, **Coppée J-Y**, **Dupuy B**, **Martin-Verstraete I**. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

51. **Sudarsan N**, **Lee ER**, **Weinberg Z**, **Moy RH**, **Kim JN**, **Link KH**, **Breaker RR**. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321**:411–413.

52. **Hollands K**, **Proshkin S**, **Sklyarova S**, **Epshtein V**, **Mironov A**, **Nudler E**, **Groisman EA**. 2012. Riboswitch control of Rho-dependent transcription termination. Proc Natl Acad Sci USA **109**:5376–5381.

53. **Purcell EB**, **McKee RW**, **Bordeleau E**, **Burrus V**, **Tamayo R**. 2015. Regulation of Type IV pili contributes to surface behaviors of historical and epidemic strains of *Clostridium difficile*. J Bacteriol **198**:565–577.

54. **McKee RW**, **Aleksanyan N**, **Garrett EM**, **Tamayo R**. 2018. Type IV pili promote *Clostridium difficile* adherence and persistence in a mouse model of infection. Infect Immun 00943–17.

55. **McAllister KN**, **Bouillaut L**, **Kahn JN**, **Self WT**, **Sorg JA**. 2017. Using CRISPR-Cas9-mediated genome editing to generate *C. difficile* mutants defective in selenoproteins synthesis. Sci Rep **7**:14672.

56. **Weiser JN**, **Austrian R**, **Sreenivasan PK**, **Masure HR**. 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect Immun **62**:2582–2589.

57. **la Riva de L**, **Willing SE**, **Tate EW**, **Fairweather NF**. 2011. Roles of cysteine proteases Cwp84 and Cwp13 in biogenesis of the cell wall of *Clostridium difficile*. J Bacteriol **193**:3276–3285.

58. **Kirby JM**, **Ahern H**, **Roberts AK**, **Kumar V**, **Freeman Z**, **Acharya KR**, **Shone CC**. 2009. Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of *Clostridium difficile*. J Biol Chem **284**:34666–34673.

59. **Fagan RP**, **Fairweather NF**. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem **286**:27483–27493.

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

61. **Willing SE**, **Candela T**, **Shaw HA**, **Seager Z**, **Mesnage S**, **Fagan RP**, **Fairweather NF**. 2015. *Clostridium difficile* surface proteins are anchored to the cell wall using CWB2 motifs that recognise the anionic polymer PSII. Mol Microbiol **96**:596-608.

62. **Chu M**, **Mallozzi MJG**, **Roxas BP**, **Bertolo L**, **Monteiro MA**, **Agellon A**, **Viswanathan VK**, **Vedantam G**. 2016. A *Clostridium difficile* cell wall glycopolymer locus influences

bacterial shape, polysaccharide production and virulence. PLoS Pathog **12**:e1005946.

63. **Stabler RA**, **Valiente E**, **Dawson LF**, **He M**, **Parkhill J**, **Wren BW**. 2010. In-depth genetic analysis of *Clostridium difficile* PCR-ribotype 027 strains reveals high genome fluidity including point mutations and inversions. Gut Microbes **1**:269–276.

64. **Bordeleau E**, **Fortier L-C**, **Malouin F**, **Burrus V**. 2011. c-di-GMP turn-over in *Clostridium difficile* is controlled by a plethora of diguanylate cyclases and phosphodiesterases. PLoS Genet **7**:e1002039.

65. **Gao X**, **Dong X**, **Subramanian S**, **Matthews PM**, **Cooper CA**, **Kearns DB**, **Dann CE**. 2014. Engineering of *Bacillus subtilis* strains to allow rapid characterization of heterologous diguanylate cyclases and phosphodiesterases. Appl Environ Microbiol **80**:6167–6174.

66. **Römling U**, **Galperin MY**, **Gomelsky M**. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev **77**:1–52.

67. **Beyhan S**, **Odell LS**, **Yildiz FH**. 2008. Identification and characterization of cyclic diguanylate signaling systems controlling rugosity in Vibrio cholerae. J Bacteriol **190**:7392–7405.

68. **Gunther NW**, **Snyder JA**, **Lockatell V**, **Blomfield I**, **Johnson DE**, **Mobley HLT**. 2002. Assessment of virulence of uropathogenic *Escherichia coli* Type 1 fimbrial mutants in which the invertible element is phase-locked on or off. Infect Immun **70**:3344–3354.

69. **Batah J**, **Denève-Larrazet C**, **Jolivot P-A**, **Kuehne S**, **Collignon A**, **Marvaud J-C**, **Kansau I**. 2016. *Clostridium difficile* flagella predominantly activate TLR5-linked NF-κB pathway in epithelial cells. Anaerobe **38**:116–124.

70. **Yoshino Y**, **Kitazawa T**, **Ikeda M**, **Tatsuno K**, **Yanagimoto S**, **Okugawa S**, **Yotsuyanagi H**, **Ota Y**. 2013. *Clostridium difficile* flagellin stimulates toll-like receptor 5, and toxin B promotes flagellin-induced chemokine production via TLR5. Life Sci **92**:211–217.

71. **Ng J**, **Hirota SA**, **Gross O**, **Li Y**, **Ulke-Lemee A**, **Potentier MS**, **Schenck LP**, **Vilaysane A**, **Seamone ME**, **Feng H**, **Armstrong GD**, **Tschopp J**, **MacDonald JA**, **Muruve DA**, **Beck PL**. 2010. *Clostridium difficile* toxin-induced inflammation and intestinal injury are mediated by the inflammasome. Gastroenterology **139**:542–52–552.e1–3.

72. **Xu H**, **Yang J**, **Gao W**, **Li L**, **Li P**, **Zhang L**, **Gong Y-N**, **Peng X**, **Xi JJ**, **Chen S**, **Wang F**, **Shao F**. 2014. Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. Nature **513**:237–241.

73. **Warny M**, **Keates AC**, **Keates S**, **Castagliuolo I**, **Zacks JK**, **Aboudola S**, **Qamar A**, **Pothoulakis C**, **LaMont JT**, **Kelly CP**. 2000. p38 MAP kinase activation by *Clostridium difficile* Toxin A mediates monocyte necrosis, IL-8 production, and enteritis. J Clin Invest

105:1147–1156.

74.   **Farrow MA**, **Chumbler NM**, **Lapierre LA**, **Franklin JL**, **Rutherford SA**, **Goldenring JR**, **Lacy DB**. 2013. *Clostridium difficile* Toxin B-induced necrosis is mediated by the host epithelial cell NADPH oxidase complex. Proc Natl Acad Sci USA **110**:18674–18679.

75.   **Batah J**, **Kobeissy H**, **Bui Pham PT**, **Denève-Larrazet C**, **Kuehne S**, **Collignon A**, **Janoir-Jouveshomme C**, **Marvaud J-C**, **Kansau I**. 2017. *Clostridium difficile* flagella induce a pro-inflammatory response in intestinal epithelium of mice in cooperation with toxins. Sci Rep **7**:3256.

76.   **Wright A**, **Drudy D**, **Kyne L**, **Brown K**, **Fairweather NF**. 2008. Immunoreactive cell wall proteins of *Clostridium difficile* identified by human sera. J Med Microbiol **57**:750–756.

77.   **Burnham C-AD**, **Carroll KC**. 2013. Diagnosis of *Clostridium difficile* infection: an ongoing conundrum for clinicians and for clinical laboratories. Clin Microbiol Rev **26**:604–630.

78.   **Fang FC**, **Polage CR**, **Wilcox MH**. 2017. Point-Counterpoint: what is the optimal approach for detection of *Clostridium difficile* infection? J Clin Microbiol **55**:670–680.

79.   **Girinathan BP**, **Braun S**, **Sirigireddy AR**, **Lopez JE**, **Govind R**. 2016. Importance of glutamate dehydrogenase (GDH) in *Clostridium difficile* colonization *in vivo*. PLoS ONE **11**:e0160107.

80.   **Girinathan BP**, **Braun SE**, **Govind R**. 2014. *Clostridium difficile* glutamate dehydrogenase is a secreted enzyme that confers resistance to H_2O_2. Microbiology (Reading, Engl) **160**:47–55.