Integrating sporulation, toxin production, and motility by redefining the role of TcdR and characterizing the sin regulon in *Clostridium difficile*

by

Brintha Parasumanna Girinathan

B.S., Bharathiyar University, 2000
M.S., Bharathiyar University, 2005

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

Clostridium difficile is a gram-positive anaerobic, motile, spore-forming opportunistic bacterium. It is a nosocomial pathogen, and the symptoms of *C. difficile* infection (CDI) range from mild diarrhea to life-threatening pseudomembranous colitis and toxic megacolon. Antibiotic use is the primary risk factor for the development of CDI as it disrupts the healthy protective gut flora which enables *C. difficile* to colonize and establish in the colon.

C. difficile damages the host tissue by secreting toxins and disseminates in the environment by forming spores. The two-major toxin-encoding genes, *tcdA*, and *tcdB* are located within a 19.6 kb pathogenicity locus (PaLoc), which also includes the gene encoding an RNA polymerase sigma factor TcdR, that is essential for toxin gene expression. We created a site-directed mutation in *tcdR* in the epidemic-type *C. difficile* R20291 strain and found that disruption of *tcdR* affected sporulation in addition to toxin production. Spores of the *tcdR* mutant were more heat- sensitive and required nearly three-fold higher taurocholate to germinate when compared to the wild-type (WT). Transmission Electron Microscopic analysis of the *tcdR* mutant spores also revealed a weakly assembled exosporium. Consistent with our phenotypic assays, our comparative transcriptome analysis also showed significant downregulation of sporulation genes in the *tcdR* mutant when compared to the WT strain. Our findings on *tcdR* suggest that the regulatory networks of toxin production and sporulation in *C. difficile* R20291 strain are interlinked with each other.

Transcriptome analysis revealed the *sin* operon to be significantly downregulated in the *tcdR* mutant which made us hypothesize the link between *sin* operon regulation and sporulation. The sin locus coding SinR (113 aa) and SinI (57 aa) is responsible for sporulation inhibition in *B. subtilis*. SinR in *B. subtilis* mainly acts as a repressor of its target genes to control sporulation,

biofilm formation, and autolysis. SinI is an inhibitor of SinR, and SinI/SinR interaction determines whether or not the SinR can inhibit target gene expression.

The *C. difficile* genome carries two *sinR* homologs in the operon, and we named it as *sinR* and *sinR*', coding for SinR (112 aa) and SinR' (105 aa), respectively. To identify the regulation of *sin* on sporulation, we created a site-directed mutation in the *sin* locus in two different *C. difficile* strains R20291 and JIR8094. Comparative transcriptome analysis of the *sinRR'* mutants revealed their pleiotropic roles in controlling several essential pathways including sporulation, toxin production, and motility (STM) in *C. difficile*.

We performed several genetic and biochemical experiments, to prove that SinR regulates transcription of crucial regulators in STM pathways, which includes *sigD*, *spo0A*, and *codY*. Unlike *B. subtilis*, SinR' acts as an antagonist of SinR and SinR'/SinR determines SinR activity. Our *in vivo* experiment using hamster model also demonstrated the importance of *sin* locus for successful *C. difficile* colonization. Our findings above reveal that *sin* locus acts as a central link that regulates essential pathways including sporulation, toxin production, and motility, which are critical for *C. difficile* pathogenesis.

The final section of this dissertation analyzes a variant *codY* gene in the epidemic *C. difficile* R20291 strain. In this strain the CodY, a global nutrient sensor-regulator carry a missense mutation where the 146th tyrosine residue is replaced with asparagine (CodY^{Y146N}). Our preliminary study with the mutated CodY^{Y146N} suggested its differential role in its regulatory activity. Further analysis of CodY^{Y146N} might give some possible clues behind the hypervirulent nature of epidemic R20291 strain.

Taken together, studies performed on both *tcdR* and *sinR* mutants reveal a significant amount of crosstalk occurring between the powerful regulators of STM pathways under the

directionality of TcdR and SinR in determining their ultimate cell fate. Our findings on CodY^{Y146N} suggest how the bacteria could switch to a hypervirulence mode by manipulating one of its vital regulators like CodY.

Integrating sporulation, toxin production, and motility by redefining the role of TcdR and characterizing the sin regulon in *Clostridium difficile*

by

Brintha Parasumanna Girinathan

B.S., Bharathiyar University, 2000
M.S., Bharathiyar University, 2005

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Interdepartmental Genetics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Revathi Govind

Copyright

Abstract

Clostridium difficile is a gram-positive anaerobic, motile, spore-forming opportunistic bacterium. It is a nosocomial pathogen, and the symptoms of *C. difficile* infection (CDI) range from mild diarrhea to life-threatening pseudomembranous colitis and toxic megacolon. Antibiotic use is the primary risk factor for the development of CDI as it disrupts the healthy protective gut flora which enables *C. difficile* to colonize and establish in the colon.

C. difficile damages the host tissue by secreting toxins and disseminates in the environment by forming spores. The two-major toxin-encoding genes, *tcdA*, and *tcdB* are located within a 19.6 kb pathogenicity locus (PaLoc), which also includes the gene encoding an RNA polymerase sigma factor TcdR, that is essential for toxin gene expression. We created a site-directed mutation in *tcdR* in the epidemic-type *C. difficile* R20291 strain and found that disruption of *tcdR* affected sporulation in addition to toxin production. Spores of the *tcdR* mutant were more heat- sensitive and required nearly three-fold higher taurocholate to germinate when compared to the wild-type (WT). Transmission Electron Microscopic analysis of the *tcdR* mutant spores also revealed a weakly assembled exosporium. Consistent with our phenotypic assays, our comparative transcriptome analysis also showed significant downregulation of sporulation genes in the *tcdR* mutant when compared to the WT strain. Our findings on *tcdR* suggest that the regulatory networks of toxin production and sporulation in *C. difficile* R20291 strain are interlinked with each other.

Transcriptome analysis revealed the *sin* operon to be significantly downregulated in the *tcdR* mutant which made us hypothesize the link between *sin* operon regulation and sporulation. The sin locus coding SinR (113 aa) and SinI (57 aa) is responsible for sporulation inhibition in *B. subtilis*. SinR in *B. subtilis* mainly acts as a repressor of its target genes to control sporulation,

biofilm formation, and autolysis. SinI is an inhibitor of SinR, and SinI/SinR interaction determines whether or not the SinR can inhibit target gene expression.

The *C. difficile* genome carries two *sinR* homologs in the operon, and we named it as *sinR* and *sinR*', coding for SinR (112 aa) and SinR' (105 aa), respectively. To identify the regulation of *sin* on sporulation, we created a site-directed mutation in the *sin* locus in two different *C. difficile* strains R20291 and JIR8094. Comparative transcriptome analysis of the *sinRR'* mutants revealed their pleiotropic roles in controlling several essential pathways including sporulation, toxin production, and motility (STM) in *C. difficile*.

We performed several genetic and biochemical experiments, to prove that SinR regulates transcription of crucial regulators in STM pathways, which includes *sigD*, *spo0A*, and *codY*. Unlike *B. subtilis*, SinR' acts as an antagonist of SinR and SinR'/SinR determines SinR activity. Our *in vivo* experiment using hamster model also demonstrated the importance of *sin* locus for successful *C. difficile* colonization. Our findings above reveal that *sin* locus acts as a central link that regulates essential pathways including sporulation, toxin production, and motility, which are critical for *C. difficile* pathogenesis.

The final section of this dissertation analyzes a variant *codY* gene the epidemic *C. difficile* R20291 strain. In this strain the CodY, a global nutrient sensor-regulator carry a missense mutation where the 146th tyrosine residue is replaced with asparagine (CodY^{Y146N}). Our preliminary study with the mutated CodY^{Y146N} suggested its differential role in its regulatory activity. Further analysis of CodY^{Y146N} might give some possible clues behind the hypervirulent nature of epidemic R20291 strain.

Taken together, studies performed on both *tcdR* and *sinR* mutants reveal a significant amount of crosstalk occurring between the powerful regulators of STM pathways under the

directionality of TcdR and SinR in determining their ultimate cell fate. Our findings on CodY^{Y146N} suggest how the bacteria could switch to a hypervirulence mode by manipulating one of its vital regulators like CodY.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to extend my inaugural sincere gratitude to my supervisor cum mentor, Dr. Revathi Govind for providing me this excellent opportunity to pursue Ph.D. in her laboratory under her expert supervision and guidance. She stood by me, supported and encouraged me and is the key player for my achievements.

Next, I want to thank my committee members Dr. Govindsamy Vediyappan, Dr. Jereon Roelofs and Dr. Sanjeev Narayanan for their valuable guidance and suggestions throughout my graduate program.

Additionally, I would like to acknowledge all the funding sources especially Johnson Cancer research funding for their constant support which helped me pursue my research throughout summer.

I would also like to make a special mention to all those who helped me during parts of my Ph.D. Dr. Tracy Miesner and her colleagues in LACS for providing me sufficient training to handle hamsters and Dr. Dan Boyle for his incredible EM images.

A special thanks to all the other people who encouraged, supported and advised me during my Ph.D., I cannot list you all, but thank you all for everything. A massive thanks to all my lab mates for all their support throughout my stay in the lab.

Finally, I would like to express my big thanks to my family members and friends for their unconditional and ongoing support. My sincere thanks to my wonderful, caring husband Dr. Chandrasekar Raman who stood by me during all my hard times, overwhelming me with his constant support and encouragement.

Dedication

I dedicate this thesis to my parents, Girinathan Parasumanna Sokkier and Kalavathy Maraimuthu for their never-ending trust in my ability, especially my father, who is in heaven, keeping a track of my progress. If not for him, I would have never dreamt of pursuing a Ph.D. program in the U.S.

Chapter 1 - Introduction

1.1 Clostridium (Clostridioides) difficile

Clostridium difficile is a Gram-positive anaerobic endospore-forming bacterium belonging to the phylum of Firmicutes and has re-emerged as one of the predominant healthcare-associated diseases worldwide over the past couple of decades due to the frequency, severity, and recurrence of this infection [1-3] . *C. difficile* is the leading cause of antibiotic-associated diarrhea and is the apparent cause of morbidity and mortality in hospitalized patients. Nearly 25% to 60% of *C. difficile* treated patients suffer from recurrent infection posing a severe threat to the healthcare industry [4-7].

In this chapter, I will discuss how *C. difficile* evolved to be a vital healthcare-associated pathogen over the years, as well as the epidemiology, bacterial pathogenesis and therapeutics with more emphasize on its molecular regulation in STM (sporulation, toxin production and motility) pathway.

1.1.1 Etymologia and history of *C. difficile* infection

The discovery of *C. difficile* dates back the 19th century in the pre-antibiotic era. John Finney first reported it and Sir William Osler at Johns Hopkins Hospital on July 28th, 1892, when a 22-year-old female showed typical CDI symptoms ranging from mild diarrhea to frequent bloody stools. She succumbed to infection within five days of severe CDI [8], and her autopsy records indicated “diphtheritic colitis” in the small bowel, which was attributed to the boric acid stomach irrigation, as a means of local antiseptic before the surgery. Later in 1935, Hale and O’ Toole isolated the bacteria from the stools of healthy neonates and named it as *Bacillus difficillis*, the

species name difficile is a form of Latin adjective difficilis because of the difficulties they encountered in its isolation and culturing [9]. Since the bacteria appeared spindle under the microscope, eventually it was reassigned to genus *Clostridium* which comes from Greek word kloster meaning spindle. *Clostridium difficile* is also called as *C. difficile*/C.diff/CDF/cdf, and initially, it was considered as part of a normal intestinal flora rather than an opportunistic pathogen since even today asymptomatic infants are reported to be colonized with *C. difficile* during the first half year of their life [10]. *C. difficile* infections (CDI) were sporadic for several decades, and there were only occasional reported cases of diarrhea in hospitalized patients making it inconspicuous for several decades since its initial discovery. The first active relationship between clindamycin and Pseudomembranous colitis (PMC), a hallmark of CDI was reported by Tedesco *et al*. in 1974. He named it as "clindamycin-colitis" [11], which was reversed by administration of vancomycin [12], suggesting the role of gut bacteria in the pathogenesis of PMC. Based on the initial hypothesis that the use of broad-spectrum antibiotics leads to overgrowth of a pathogenic organism by altering the gut microbiota set the strong foundation for our current understanding of CDI [13]. Toxin-producing *C. difficile* was attributed to be the causative agent for PMC by Barlett et al., in 1978 [14]. By the late 1980s, based on several studies relating to the epidemiology, clinical features, diagnostic tests and effective treatment strategy devised for CDI, *C. difficile* was assumed to be well understood and easily manageable bacteria [15]. All the assumptions about *C. difficile* were turned on its head in the early 21st century when the bacteria reemerged with an increase in incidence and severity and established itself as a significant public health threat ever since [2, 16].

1.1.2 Genome and ribotypes used in present study

C. difficile has a highly mosaic genome prone to genetic exchange [17]. To date, several PCR ribotypes have been identified with diverse population structure that spreads across six phylogenetic clades [18, 19]. The widely used *C. difficile* strain 630 belonged to PCR ribotype 012 and was isolated in 1985. The genome of *C. difficile* 630 is 4,290,252 bp and has low G+C content of 29.06%, and it was first sequenced by Sebaihia, M. *et al*., in 2006. It encodes for 3,776 coding sequence (CDS) with more than 80% of CDSs encoded from the forward strand [17, 20, 21].

1.1.2.1 *C. difficile* strain JIR8094

Rood lab (Monash, Australia) and Mullany lab (UCL, London, UK) isolated spontaneous erythromycin sensitive derivative of the reference strain 630 by 30 repeated serial passaging in an antibiotic-free media and named it as 630Δerm and 630Estrain, also called JIR8094 respectively [22-25]

1.1.2.2 *C. difficile* strain R20291

The PCR ribotype 027 (also called NAP1/REA type B1- North American Pulsotype 1/ restriction endonuclease analysis type B1) had emerged in North America around 2000 due to the frequent use of commonly prescribed antibiotic classes fluoroquinolone [26-29]. Fluoroquinolone resistance in these isolates was due to two identical independent mutation (Thr82Ile) in the *gyrA* gene of two distinct epidemic lineages of ribotype 027. These epidemic lineages were responsible for dissemination of 027 strains globally and led to the outbreak of epidemic *C. difficile* strain R20291 (PCR ribotype 027) in continental Europe that was isolated in 2004-2005 [19, 30]. *C.*

difficile R20291, a hypervirulent strain shares 3,247 CDS with the reference strain 630, including all essential virulent genes and some genes unique to itself [30, 31]

1.2 Epidemiology of *C. difficile*

C. difficile is a nosocomial (hospital-acquired) pathogen and virtually most of the hospital environment including medical personnel, floors, bedrails, hospital surfaces are heavily contaminated with *C. difficile* along with the infected person [2, 32]. Nevertheless, *C. difficile* infected patients are quarantined in hospitals, the risk of CDI is directly proportional to the duration of hospital stay thereby making hospitalization a compelling requirement for acquiring CDI [33]. In both asymptomatic and symptomatic patients, *C. difficile* is transmitted from the patients by the contaminated hands of healthcare workers [34]. The diarrheal symptom of CDI is the primary explosive weapon for massive *C. difficile* shedding to the local environment. The physical proximity to infected patients was hypothesized to be the prime risk factor in the horizontal transmission of *C. difficile* which was strengthened by the observational study performed by McFarland who showed an increased incidence of *C. difficile* patients following an infected patient in a commode [35, 36].

Nowadays, increased incidence of CDI is seen in population like children and young healthy women, who were earlier considered to be at low risk [3, 35, 37]. These include communities of individuals with no known history of the hospital or antimicrobial exposure, however, is linked to the frequent use of proton pump inhibitors and closer proximity of livestock. Severe community-acquired CDI is noted in healthy women following child birth, and it requires hospitalization to be treated [3]. *C. difficile* ribotype 078 and other unique ribotypes are associated

with community-acquired CDI [3, 38, 39]. The exact reason for such incidence requires a thorough groundwork.

1.2.1 Economic burden and infection control

C. difficile infects roughly 500,000 people each year in the U.S., sending more than 347,000 to the hospital for treatment. An estimated *C. diff*-associated death range from 14,000 to 30,000 annually and it costs health care system several billion/year to treat CDI [40] rendering this pathogen a severe threat by the Centre for Disease Control and Prevention (CDC).

To control CDIs and to reduce the associated burden on health care providers, infection control guidelines based on scientific awareness is critical. As per the literature survey, no single approach has been successful in controlling the spread of CDI in hospitals. Instead, a multidisciplinary approach is fruitful taking into account all the principal reservoirs of infection including the patients, their environment, the medical personnel involved and the equipment's used [41]. The appropriate way to control the spread of CDI is keeping it contained. Hence the Infectious Diseases Society of America (IDSA) and the Society of Healthcare Epidemiology of America (SHEA) recommends specific guidelines for CDI [42]. It includes quarantining *C. difficile* patients in isolated rooms with self-contained toilets or particular *C. difficile* unit with chlorine cleaned environment and medical equipment dedicated to each patient. Employ barrier nurses using disposable gloves and aprons who are well trained in treating CDI, antibiotic stewardship (vancomycin or fidaxomicin replacing metronidazole as first-line therapy in adults) and most importantly hand hygiene [42-49]. Recommendations involving testing and isolating CDI is also mentioned in the updated guidelines [42]. It is noteworthy that *C. difficile* spores easily bypass the alcohol-based hand gels which are used in hospitals [49].

1.3 Symptoms and treatment of *C. difficile* Infections (CDI)

The prevalent symptom of CDI is mild diarrhea which requires no treatment other than cutting off the provoking antibiotics which result in replenishing of the gut flora to keep *C. difficile* under control [50]. However, the more severe form of CDI leads to the life-threatening condition called pseudomembranous colitis (PMC) and toxic megacolon [51] The colonoscopy findings of PMC relative to the healthy colon is shown in (Figure 1.1). The characteristic feature of PMC are yellow-white plaques over the colonic mucosa filled with host immune cells like leukocytes, neutrophils, fibrin, mucous and inflammatory debris, which can be defined by histopathology [34, 52]. Another lethal complication of CDI is toxic megacolon, characterized by dilated and perforated colon accompanied by constant abdominal pain and reduced bowel movement. It is the most severe form of CDI and if left untreated, can lead to septicemia, peritonitis and ultimate death [51].

The main culprit of CDI is antibiotics, and ironically antibiotics are used to treat this infection. In mild cases, CDI can be treated conservatively by withdrawing the provoking antibiotics and rescuing the colonic microbiota alongside supportive therapy like electrolyte and fluid replacement [53]. However, in cases, where antibiotic treatment is a prerequisite for underlying infection, then a narrow spectrum antibiotic could be used which does not correlate with CDI [54, 55]. In treating severe cases of CDI, fecal microbiota therapy (FMT) is used widely, and treatment with immunotherapy is underway [56].

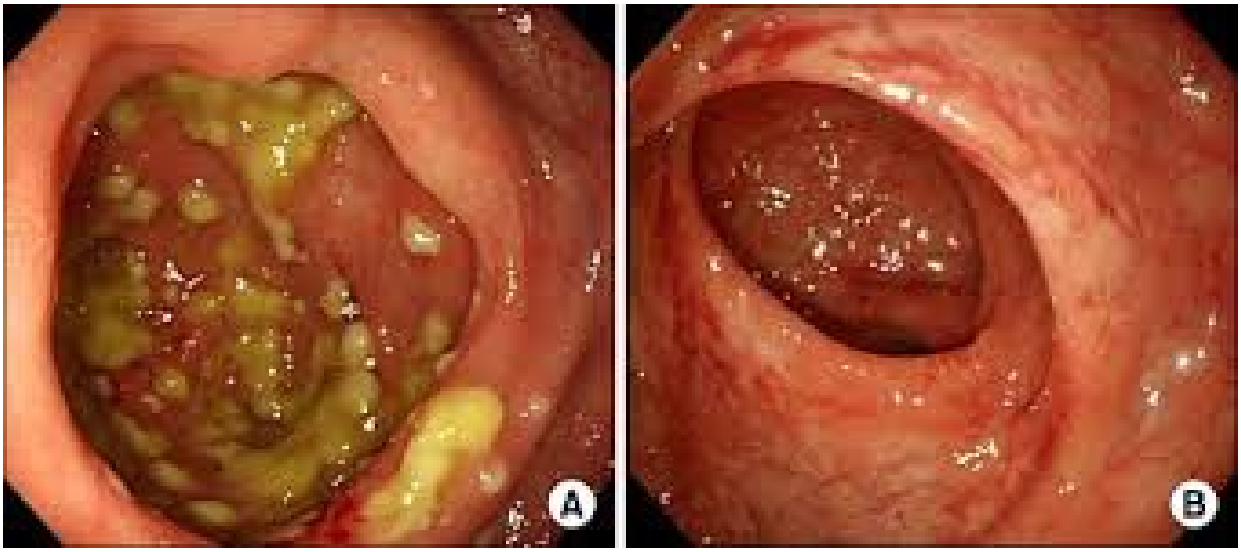


Figure 1.1 Colonoscopy view of PMC relative to normal colon

To the left is the image of PMC colon with raised yellowish plaques over colonic mucosa and to the right is normal colon (Image obtained from Dr. Sung Yeun Yang, South Korea)

1.4 Pathogenic mechanisms of *C. difficile*

1.4.1 Procurement and Colonization

C. difficile produces highly dormant and infective spores [57]. Spores are excreted in the feces and the most importantly, they are resistant to most of the disinfectants [57-59]. The spores serve as an infectious vehicle, responsible for transmission of the disease and persistence of the organism in the environment and the infected patients [34, 60]. The spore biology and its role in *C. difficile* transmission are reviewed in greater detail in Chapters 1.4.2.

Infection begins when *C. difficile* gets ingested by the host. Most of the vegetative *C. difficile* die in the stomach due to the harsh acidic environment [36]. However, the spores that are resistant to stomach acid survive and pass through the stomach and become established in the small intestine [57] and they remain under constant surveillance by the host enteric microbiota. When

this control is lost, due to antibiotic treatment, then the spores rapidly germinate in response to bile derivatives like taurocholic acids [61, 62] and migrate to the colon [63].

C. difficile establishes in the colon by expressing several virulence factors that are associated with *C. difficile* pathogenesis which include (i) Flagella (FliC – flagellin, FliD flagellar cap proteins) and fimbriae – mediate movement and adherence to the gut mucosal layer, (ii) Proteolytic and hydrolytic enzymes – to rupture the gut mucosal barrier (iii) Adhesins and surface proteins - Cwp66 (cell wall proteins) – facilitate binding to enterocytes and (iv) anti-phagocytic capsule to prevent opsonization and engulfment by neutrophils (PMN - polymorphonuclear leukocytes) [64, 65]. On successful colonization and establishment in the colon, *C. difficile* produces two primary virulence factors, toxin A and toxin B.

1.4.2 **Disease progression via toxins**

C. difficile was first defined as a toxin-mediated pathogen by Bartlett et al. in 1978 [14], and since then only the toxin-producing strains have been reported to be pathogenic, and the strains which do not produce toxins are non-pathogenic – e.g., PaLoc-ve [18, 66, 67]. *C. difficile* produce two enterotoxins (toxins that target the intestine), named toxin A and toxin B which are the primary virulence factors of the pathogen [68]. However, the relative contribution of toxin A and toxin B in CDI has been on debate for several years. Toxin A was initially thought to be the primary virulence determinant of CDI, and the virulence was found to be augmented by toxin B in the presence of toxin A [67], thereby both toxins worked in harmony to establish CDI. This theory was turned on its head by Janezic 2015, who showed that toxin B caused CDI and its activity was independent of toxin A which was supported by the isolation of naturally occurring pathogenic A-B+ variant of *C. difficile* strain [69, 70]. These paradoxical findings were finally challenged by

Kuehne and his colleagues, who showed that isogenic *C. difficile* mutants producing either toxin A or toxin B were cytotoxic *in vitro* which translated directly into virulence *in vivo*, suggesting that both toxins have the pathogenic potential [71].

Toxins A & B have three compartments as shown in (Figure 1.2)– A N terminal catalytic domain (1-543 aa) with enzymatic (glucosyltransferase) activity which reaches the host cell cytoplasm. A central translocation domain which contains cysteine protease domain (543-801aa) and a hydrophobic domain with the pore-forming region (801-1132 aa) for membrane insertion and translocation of toxins into the host cell cytoplasm. The C terminal receptor binding domain (1832-2710 aa) - facilitates the host cell receptor binding [72, 73].

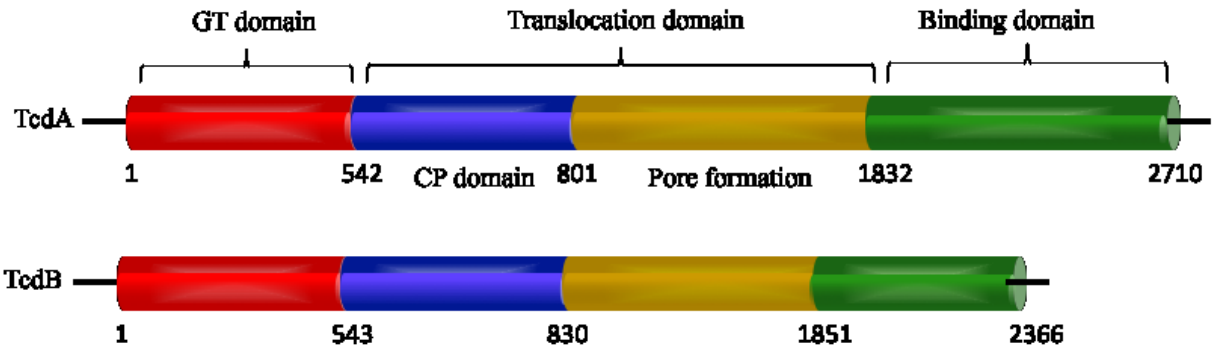


Figure 1.2 Structural organization of toxin genes

Both toxins have glycosyltransferase domain in the N terminus, Translocation domain containing cysteine protease and delivery or pore forming domain in the middle and binding domain in the C terminus which is slightly modified in toxin B. Cellular intoxication of the host by *C. difficile.*

Following colonization, the intoxication of the host cells with *C. difficile* toxins occur via an ABCD model (A-biological activity, B – binding, C -cutting, and D- delivery) which is shown in (Figure 1.3) [74]. The toxins initially bind to its receptor through the C terminal receptor binding domain followed by endocytosis of the toxins and trafficking into the cellular endosomes through clathrin and dynamin-dependent or independent pathway [75]. The toxin in the endosome

undergoes a conformational change due to the acidification of the endosomal compartment resulting in the formation of protrusions and insertion of toxins into the endosomal membrane and subsequent formation of the pore within the endosomal membrane [72]. The toxin then binds to InsP6 (inositol hexakisphosphate - a host cell-derived molecule) [76], which induces a conformational change in the cysteine protease domain mediating autocatalytic cleavage and facilitating the release of the active N terminal glucosyltransferase domain of the toxin into the cell cytosol [76, 77]. It, in turn, results in glucosylation and inactivation of Rho family of small GTPase (such as Rho, Ras, and Rac) [78, 79] by covalently transferring cellular glucose moiety from UDP-glucose to the threonine amino acid of the target proteins. Thereby it forms an O-glycosidic bond and makes them inactive [74]. This inactivation of Rho GTPases affects many downstream cellular pathways like loss of structural integrity due to lack of actin depolymerization and the associated drop in cellular levels of F-actin resulting in a characteristic cell rounding phenotype following caspase 3 and 9 mediated cellular apoptosis [80]. The disruption of cellular tight junctions follows within the intestinal epithelium due to lack of colonocyte death and eventual loss of epithelial barrier function. Collectively, intestinal permeability and fluid accumulation increases resulting diarrhea, the hallmark symptom of CDI.

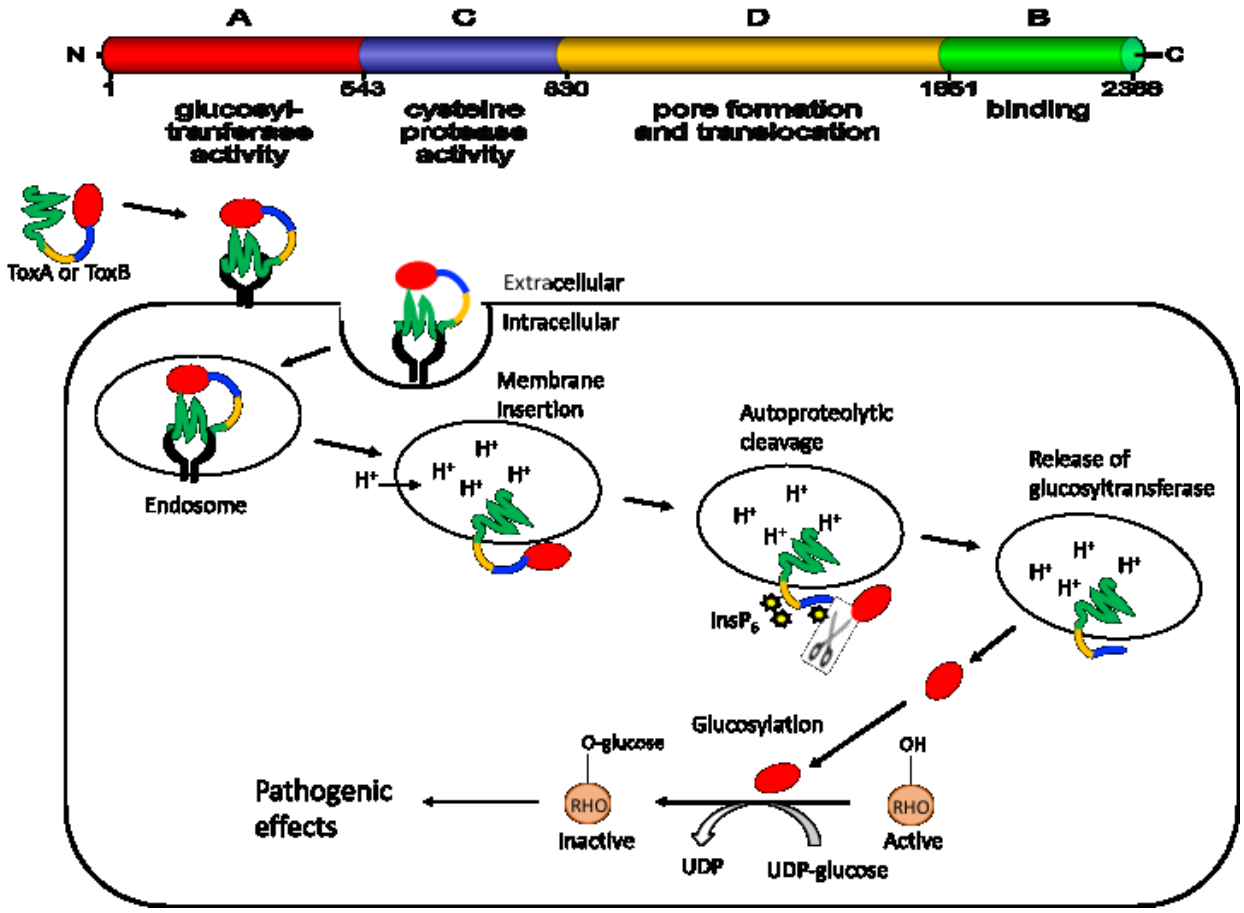


Figure 1.3 Intoxication of host cell

The cellular intoxication begins when toxins bind to the receptors followed by endocytosis into the endosomes and integration into the endosomal membrane and translocate into the cytosol at low pH, due to acidification of the endosome. It results in autoproteolytic cleavage of the toxin and release of an active glycosyltransferase domain into the cytosol which causes pathogenic effects by inactivating the Rho GTPases resulting in inhibition of Rho-dependent signaling. Red - N terminal catalytic domain, Blue - cysteine protease domain, Yellow - central hydrophobic domain and Green - C terminal domain (Adapted and modified from [81]

1.5 Regulation of toxin production in *C. difficile*

1.5.1 Toxin gene regulation

The toxin genes *tcdA* (toxin A) and *tcdB* (toxin B) are located within a 19.6 Kb region of the chromosome known as Pathogenicity Locus (PaLoc) which is spatially present in all

toxigenic *C. difficile* strains making it less likely to be a mobile genetic element [82]. Other three accessory genes present in this locus are positive regulator *tcdR*/earlier known as *tcdD* [83], negative regulator *tcdC* [84] and *tcdE* (gene encoding for a putative holin like protein) [85].

1.5.1.1 Role of *tcdR* on toxin production

Toxins gene expressions in *C. difficile* is responsive to cellular growth phase and constituents of the medium. Toxins levels are low during the exponential period, and the pathogenicity genes (*tcdR, tcdB, tcdE,* and *tcdE*) are transcribed only when the cells enter the stationary phase when there is some form of nutrient limitation or accumulation of growth inhibiting substance [86]. The positive regulatory roles of TcdR on toxin production was identified when the *tcdR* gene was found to activate *tcdA* and *tcdB* reporter fusions when expressed *in trans* in *E. coli* [87]. *In vitro* transcription experiments performed using core RNA polymerase suggested that TcdR was able to interact with the *tcdR, tcdA, tcdB* promoter regions and stimulate their transcription specifically and activate its transcription. Protein-protein interaction experiments indicated that TcdR was able to bind directly with the core enzyme in the absence of DNA [87-90] thereby concluding TcdR as alternative RNA polymerase sigma factor required for the transcription of self and toxin gene expression in *C. difficile*. Due to its distinct nature, TcdR is classified as a member of a discrete group V of sigma factors within the σ70 family.

Environmental factors that found to inhibit toxin production includes, specific nutrients like glucose, a mixture of amino acids (like glycine, methionine, tryptophan, threonine, isoleucine, leucine, valine) proline, cysteine, butanol, and biotin. Availability of butyrate in the medium and the growth temperature at 37^0C are known to activate toxin production [91].

1.5.1.2 TcdC, a postulated negative regulator of toxin production

TcdC is an acidic, membrane-associated protein with the ability to form dimers in its purified form due to the presence of coiled-coiled motifs in the middle of the protein. Based on the the timing of its relative expression and emergence of high toxin-producing epidemic strains with deletions or frameshift mutations in the *tcdC* gene, TcdC was suggested to be the negative regulator of toxin production [84, 86]. TcdC was later found to inactivate or destabilize the open complex formation before transcription initiation by interfering with TcdR and TcdR containing RNAP holoenzyme, but how it does explicitly, that is poorly understood [84, 92]. However, TcdC was found to be insensitive once the TcdR containing RNAP holoenzyme formed a stable open complex with toxin promoters like *tcdA* or *tcdB*. The toxin production in *C. difficile* strain M7404 (PCR ribotype of NAP1/027), which carries a non-functional *tcdC* gene was found to be decreased by introduction of a functional *tcdC* gene. However, chromosomal complementation in *C. difficile* strain R20291 (PCR ribotype – NAP1/027) with an inactive *tcdC* was found not to affect toxin production [92]. Hence a lot of uncertainty prevails about the role of TcdC in *C. difficile* pathogenesis and its influence on TcdR.

1.5.1.3 Role of other regulators in Toxin production

Given the complex nature of toxin gene expression, other global regulators like CodY, CcpA, σD, Spo0A, σH are also found to influence toxin production in *C. difficile*. Metabolite sensing regulators like CodY and CcpA regulate toxin gene expression by directly repressing the expression of *tcdR* [93, 94]. CcpA, a global regulator of carbon metabolism pathway gets activated in the presence of glucose metabolism intermediate, fructose 1,6 biphosphate (FBP) and repress the expression of toxin genes. CodY, a nutrient sensing global regulator binds tightly to the

promoter region of *tcdR* in the presence of cofactors like GTP and BCAA and repress it [95]. Hence both CcpA and CodY respond to three metabolites namely FBP, GTP, and BCAA.

σD, an alternative sigma factor of the flagellar operon positively regulates the toxin gene expression by directing the RNA polymerase to recognize the promoter region of only *tcdR* (not *tcdA* or *tcdB*) [96]. These findings were strengthened when relatively increased mRNA levels of *tcdA, tcdB, tcdR* and unaltered levels of *tcdC* were noticed in *C. difficile* overexpressing *sigD* confirming the positive regulatory role of σD on toxin gene expression [86, 97]. The σD-mediated toxin gene expression in *C. difficile* is in response to intracellular c-di-GMP concentration [97].

Spo0A, the master regulator of sporulation is found to repress toxin gene expression. However, it is very likely that this effect is indirect, due to the lack of Spo0A binding sequence on the upstream of any PaLoc genes [98-101]. Transition phase sigma factor, σH also influences *C. difficile* toxin production. A null mutation in *sigH* leads to overexpression of *tcdR*, *tcdA*, and *tcdB* implying that σH is an inhibitor of toxin gene expression. σH modulates the expression of *spo0A*, the butyrate biosynthesis pathway, where butyrate turn out to be an activator of toxin production [102].

1.6 Sporulation in *C. difficile*

The persistence of *C. difficile* in the environment is mainly due to the presence of spores as they are retained in the surroundings for a more extended period thereby favoring its transmission.

The structure of a mature spore is shown in (Figure 1.4). Spore is made up of a central core, which is made up of cellular DNA and ribosomes. The DNA in the center is saturated by small acid-soluble proteins (SASPs), and the semi-dehydrated state of the core is maintained by high

concentrations of dipicolinic acid (DPA) [103, 104]. The spore DNA and DPA contribute to the resistance of spore against heat, desiccation and genotoxic invasions [103].

The spore core is lined by an inner membrane, which serves as a putative permeability barrier, the germ cell wall, which during germination becomes the cell wall of vegetative cells and a thick cortex, which is a well-hydrated peptidoglycan rich matrix, in contrast to the core. The hydrated nature of the cortex maintains the dehydrated quality of the center via physical and osmotic pressure. Surrounding the cortex is the outer membrane, coat, and the exosporium. The coat is the hardened proteinaceous carbohydrate-rich layer which protects the spore layers from enzymatic and chemical agents as spore defective in the coat are no longer resistant to gastric juice and the outer most layer of the spore; the exosporium is a defined shell rich in glycoproteins. Exosporium is required for cell adhesion and colonization and is unique to *C. difficile* [105, 106].

In a laboratory setting, the spores are germinated in the presence of bile acid-like taurocholate [62]. The germinated vegetative cells which are grown in rich medium undergo repeated cycles of vegetative replication by binary fission, representing a metabolically active stage of growth. At the end of the exponential phase, *C. difficile* senses and integrates various environmental and physiological cues through regulatory molecules like sigma factors, positive and negative transcription factors and sensors of nutritional and growth status [106]. In response to the signals, *C. difficile* switches on a series of reaction, to prevent cells from an inappropriate commitment to dormancy and to overcome hostile environmental conditions [107].

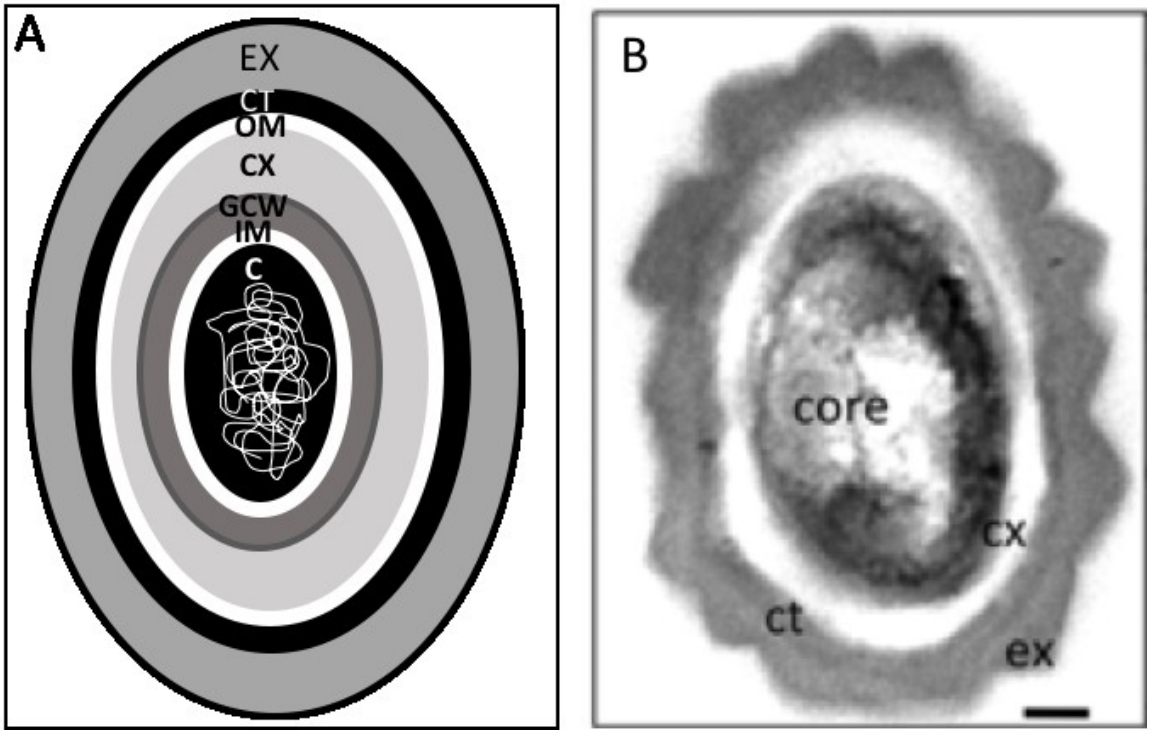


Figure 1.4 Ultrastructure of bacterial endospore

(A) Schematic representation of different layers of endospore. (B) TEM image of mature *C. difficile* endospore showing innermost core (C) containing DNA and ribosomes is surrounded by continuous inner membrane (IM) and germ cell wall (GCW). Over this, is the thick peptidoglycan rich cortex (CX) encircled by outer membrane (OM), spore coat (CT) containing distinct striated layers and exosporium, the outermost (EX). TEM image was taken by Dr. Dan Boyle (K-State).

During the switch from logarithmic to stationary phase, *C. difficile* become motile to explore their environment for nutrients [96]. It activates the transcription of virulence genes [86], secrete extracellular protease - to degrade polymeric nutrients (the minor activity) [108], produce antibiotics to eliminate competitors [109, 110]. Finally, if no other option is available, then the stationary phase cells enter sporulation [98, 99, 102].

1.6.1 General process of spore formation

The spore formation begins with the cell's commitment to the release of spore from the mother cell through a complicated sequential development process involving several biochemical and physiological changes [111]. The entire process is classified into seven stages, and an overview of it is shown in (Figure 1.5)

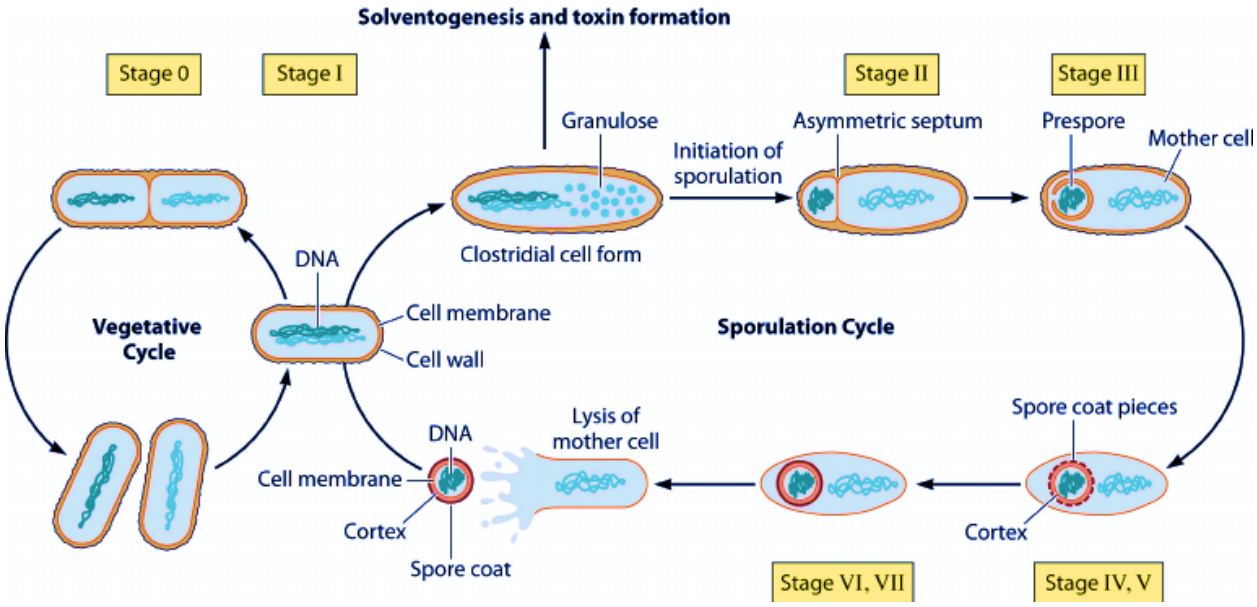


Figure 1.5 Schematic overview of Spore formation and Germination

The spore formation cycle is determined by 7 cytological stages. In response to molecular cues, the spores shed off their protective layers, becomes metabolically active and outgrows back to vegetative cell which is termed as germination. (Adapted and modified from [111]

Stage 0 and Stage I is the preparatory phase for the commitment of the vegetative cells to sporulation. Stage II is a cell fate-determining step during which, the first morphologically visible change occurs. The committed cells favor a polar division site to initiate septum formation and

undergo asymmetric cell division instead of the regular central division. This asymmetric division generates a large mother cell and smaller pre-spore compartments with distinct developmental fates [112]. In stage III, the engulfment process occurs when the mother cell creates a cell-within-a-cell state by engulfing the pre-spore via phagocytosis. The engulfed pre-spore contains double membrane, and the peptidoglycan rich cortex is deposited between the layers. Both stage IV and V are morphogenesis stages, and the cell begins to form its refractive properties in stage IV, followed by stage V during which outer pre-spore membrane gets multi-layered with spore coat. Stage VI-VII have programmed cell death stages, and final phase of spore maturation occurs in stage VI which involves DPA synthesis which is synthesized and transported from the mother to the daughter cell [113, 114]. Stage VII is the final stage when the mother cell lysis to release the mature spore [57].

1.7 Regulation of sporulation

Endospores are formed in bacteria generally in response to nutritional stress. The entire process of sporulation requires a lot of time and energy to be completed and this process is irreversible once when cells are committed to sporulate by forming the asymmetric polar septum. Hence the bacterial population tries to block sporulation to the fullest extent and sporulation is initiated only as an extreme response to adversity. Once committed, various extracellular (surfactin) and intracellular signals (nutritional stress like the drop in amino acids and GTP levels) initiate sporulation with several hundred of genes explicitly devoted to sporulation and their regulatory pathways involving key sigma factors and numerous transcriptional regulators. The signals are sensed by phosphorelay system which eventually phosphorylates the Stage 0

sporulation protein A Spo0A, which is the transcription factor and master response regulator of sporulation pathway [115].

Sporulation pathway is controlled by a regulatory cascade of kinases, phosphatases, finely tuned activation-by-phosphorylation of Spo0A, key transition phase sigma factor, σH and various compartmentalized sigma factors in the bacterial cell [98, 102, 116, 117]. The different stages of spore formation are explained in Chapter 1.4.2.1. The compartment-specific sigma factors, σG and σK in the mother cell and σF and σE in prespore are required for the successful formation, maturation and release of matured spores [112, 117].

In *B. subtilis*, major regulators that influence sporulation are SinR, CodY, σH, AbrB [[118-121]. A comparison of *B. subtilis* sporulation specific gene, regulators and its orthologues in different *Clostridia* species is shown in (Figure 1.6) [122].

1.7.1 SinR - Sporulation Inhibition Regulator

SinR role in transition phase regulation is extensively studied in *B. subtilis* and for convenience sake, the *B. subtilis* sin locus genes are labeled as BsSinR and BsSinI and *C. difficile* sin locus genes are marked as CdSinR and CdSinR' respectively.

The *B. subtilis* sin locus codes for BsSinR (113aa), a 14kDa protein, and BsSinI (57aa), a 6kDa protein that acts as an antagonist of BsSinR [123]. BsSinR is a multimeric DNA binding protein with HTH DNA binding domain (1-69 aa) in the N-terminus and multimerization domain (74-111 aa) in its C-terminus. BsSinI, on the other hand, resembles a truncated form of BsSinR with only multimerization domain [124].

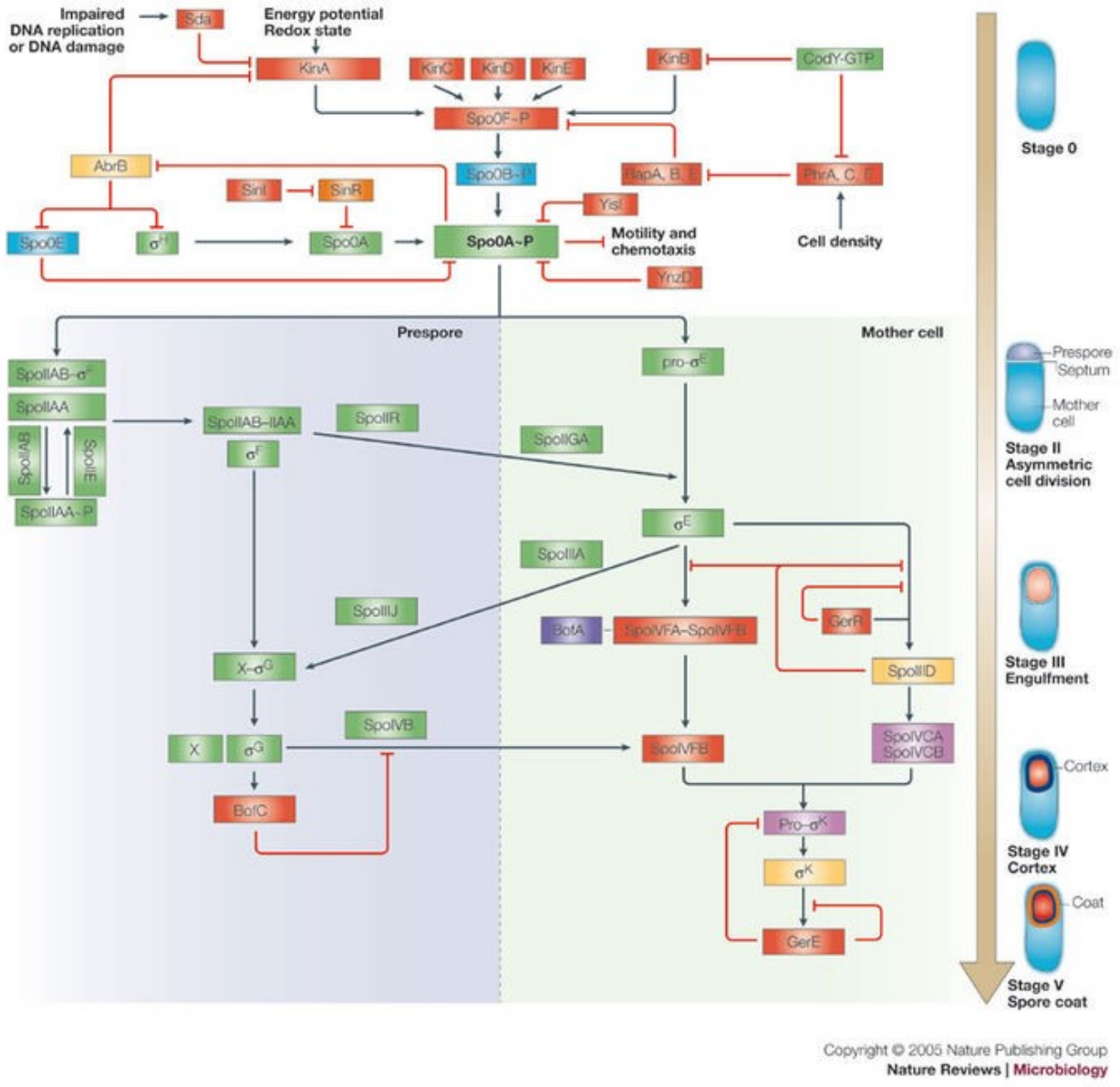


Figure 1.6 Canonical overview of genes regulating spore formation

Schematic representation of sporulation cascade in *B. subtilis* with its putative orthologues in Clostridial species. The orthologues are colored coded according to their presence or absence is as follows: Green- present in all Clostridia with >30% identity, Red – Not identified in any Clostridia, Yellow – present in all Clostridia except *Clostridium tetani* (*C. tetani*), Blue – possible presence only in *C. tetani*, Orange – present only in *Clostridium acetobutylicum* (*C. acetobutylicum*) and *C. tetani*, Pink - Post translational processing is required to form a functional protein in *B. subtilis* and *C. difficile* alone. (Adapted from [122]).

B. subtilis SinR, a primary transition phase regulator, responds to the nutrient depletion by acting as a developmental switch and controlling alternate developmental pathways. It is a pleiotropic dual function regulatory protein having both positive and negative effects on gene expression during late growth [118, 125-127]. Some of the major pathways regulated by BsSinR are discussed below.

1.7.1.1 Role of SinR in sporulation

In *Bacillus subtilis*, the proteins of the *sin* locus (sporulation inhibition) form a component of a distinct molecular circuitry that governs the cells commitment to sporulation [118, 123]. SinR (Sporulation initiation Repressor), a tetrameric protein functions as a transcriptional repressor of the genes essential for entry into sporulation at different levels by directly binding to their DNA [118]. It acts as a repressor of *kinB* by binding to SinR binding consensus sequence (GTTTYT, Y=T/C) in an inverted orientation between nucleotide-57-42 bps [128]. Next primary target for SinR is *sigE* (early-acting sigma factor in the mother cell), *spoIIA, spoIIG* and *spo0A* [118, 129, 130]. The master regulator of sporulation, *spo0A* is expressed from two promoters driven by σA and σH [131]. SinR binds as a tetramer directly to the promoter region that is recognized by σH RNAP through specific guanine residues that are crucial for SinR recognition [130] and prevents *spo0A* transcription. SinR mediated repression is overcome through the activity of SinI, which is activated by Spo0A-P. The presence of SinI disrupts the SinR tetramer through the formation of a SinI-SinR heterodimer complex [123, 124]. Function of SinI/SinR is highly sensitive to the gene dosage and chromosome positioning. During DNA replication, genes present near the origin of replication are present in high copy number than those near the terminus. The *sin* locus is located near the terminus, and its location in the chromosome is crucial because once when the

concentration of SinR is doubled in the cells, then its antagonist SinI is less effective in preventing SinR from binding to its targets. It suggests the requirement of spatiotemporal gene expression of *sin* locus for fine-tuning its regulation during adverse conditions [132-134]. Similar to *B. subtilis*, in *C. difficile*, the *sin* locus contains two genes *cdsinR* encoding 339bp SinR and *cdsinI* encoding 318 bp SinI. Both CdSinR and CdSinI has an N terminal HTH DNA binding domain and C terminal multimerization domain and are 43% and 35% identical to *Bs*SinR, suggesting that *Cd*SinR and *Cd*SinI might have arisen through a gene duplication event [135]. Hence *Cd*SinI was renamed as *Cd*SinR'. Previous studies have shown that inactivation of *spo0A* and *σH* resulted in an asporogenous phenotype and upregulation of *sin* operon [98, 102]. However, the null mutation in *opp* and *app* permeases resulted in a hyper sporulation phenotype with upregulation of *sin* locus suggesting differential regulation of *sin* locus in asporogenic and hyper sporulation phenotype [136]. These controversial findings led us to characterize the pleiotropic role of *C. difficile sin* locus which is the focus of second half of this study.

1.7.1.2 Role of SinR in motility

In *C. difficile* and *B. subtilis*, σD - an alternative RNA polymerase sigma factor positively regulates the genes needed for the formation of flagellar assembly and motility [96, 137]. In *B. subtilis*, cells high in σD positively regulate the expression of genes coding for the flagellar filament protein (hag) which eventually lead to motile cells which are free to explore the nutrients from the environment [138]. FlgM is an anti-sigma factor of σD that gets sequestered intracellularly in both *C. difficile* and *B. subtilis* [96, 139]. A functionally assembled hook-basal body (HBB) is required to actively transport FlgM outside the cells thereby freeing σD to direct the transcription of late flagellar genes [137, 140]. Initial studies performed by Rashid *et al.* in *B.*

subtilis, showed that mutation in *sinR* affected SigD dependent function at multiple points suggesting a direct correlation between role of SinR on σD and its requirement in motility, however, this requirement was nullified when *flgM* was mutated in cells lacking *sinR* [127, 139] which suggest BsSinR positively regulates motility by repressing flgM. However, the function of SinR is altered in the presence of another SinR like protein called SlrR. SlrR is a SinR like protein with N terminal HTH DNA binding domain and C terminal multimerization domain. The amino-terminal domain of SlrR resembles SinR repressor, and its carboxy-terminal domain resembles SinI antirepressor [141]. Biochemical experiments showed that SlrR could directly interact with SinR and this interaction seems to affect the motility. The SlrR influences SinR activity by forming a heterodimer, and the SinR/SlrR represses the promoters of *hag* genes which encodes flagellin [132, 133, 142].

Both motile (*C. difficile* strain R202921) and non-motile strains (*C. difficile* JIR8094) are present in *C. difficile* (refer to Section 1.1.2.1 and 1.1.2.2). The motile strains have an intact flagellar operon, and in the non-motile JIR8094 strain, the flagellar operon is known to carry some uncharacterized deletions that hampering its ability to form motile cells [22]. Other than SigD, motility in *C. difficile* is found to be regulated by other regulators like CodY (Global nutrient sensing regulator) [95], σH (Transition phase sigma factor) [102], Spo0A (Master regulator of sporulation) [99], RstA (Regulator for sporulation and toxin) [143], and LexA (SOS master regulator) [144]. However, the cross-talk between these global regulators in regulating motility is not well studied. The transcription of *sigD* gene is directly controlled by the increased intracellular levels of secondary messenger molecules called c-di-GMP in both *C. difficile* and *B. subtilis* [97, 145]. Within the cells, the c-di-GMP is synthesized from two molecules of GTP by diguanylate cyclase (DGCs) and hydrolyzed by phosphodiesterase (PDEs) [146], and the intracellular levels

of c-di-GMP are based on the levels of DGCs and PDEs in the cell. In *C. difficile*, the c-di-GMP blocks the flagellar driven motility by binding to the Cd1 riboswitch located upstream of the *flgB* operon and promote premature termination of transcription [97, 147, 148]. Determining the role of SinR in motility and its effect on σD and c-di-GMP is part of this study.

1.7.1.3 Role of SinR in autolysis

In *B. subtilis*, the σD is a positive regulator of the cell separating autolysin genes *lytC* and *lytF* [125]. Since BsSinR positively regulates σD indirectly by inhibiting *flgM*, SinR is found to control σD dependent genes like autolysins [127]. Like motility, SinR regulation on autolysis is altered in the presence of SlrR. In the presence of SlrR, SlrR/SinR functions as an epigenetic switch and repress the genes involved in cell separation by binding specifically to promoter regions of autolysin genes (*lytABC* & *lytF*) [132, 133, 142]. Major autolysin operon *lytABC* is absent is *C. difficile*, and blast searches revealed no SlrR homologs in *C. difficile*. However, the effect of SinR on autolysis is explained in Chapter 3.

1.7.1.4 Role of SinR in biofilm formation

In *B. subtilis*, biofilms consist of a long chain of tightly associated bacterial cells that are held together in bundles encased within a slimy self-produced extracellular matrix which is mainly composed of exopolysaccharide (EPS) encoded by *epsA-O* operon and protein TasA encoded by *tapA/yqM-sipW-tasA* operon [149-151]. The genetically identical cells within ubiquitous communities of biofilm can further differentiate and express a different set of genes with distinct functions.

In *B. subtilis*, cells are proposed to exist in two states -SlrR low (corresponding to single, motile cells) and SlrR high (relevant to chains of matrix-producing cells). When SlrR levels are low in the cells, SinR inhibits the *slrR* transcription, thereby keeps the level of SlrR protein low, which in turn results in no biofilm formation. SinR also inhibits biofilm formation by acting as a potent negative transcriptional regulator of *epsA-O* (EPS component of the matrix) and *yqM-sipW-tasA* (Protein component of the matrix) [152, 153]. SinR binds directly to the conserved DNA sequence upstream (5' –TTTGTTCTCTAAAGAGAACTTA-3') as a homotetramer and inhibits the transcription of these genes, which are essential for biofilm formation [150]. Spo0A-P on the other hand positively regulate matrix gene expression by indirectly influencing its activity on SinR through SinI [154]. It is achieved by Spo0A-P binding to the operators of *sinI* and activating its expression [155]. SinI when expressed, blocks SinR mediated repression on biofilm by forming SinR/SinI complex thereby preventing SinR from binding to its target DNA [123, 124, 153]. The formation of SinR/SinI also switches the cells to SlrR high state. Once, when the SlrR levels are high, they tend to stay in the high state for several generations. In this state, SlrR binds to SinR and traps it in the heteromeric SinR/SlrR complex. As a result, the derepression of matrix genes (*eps A-O* and *tapA-sipW-tasA*) occurs in the presence of SlrR/SinR complex [132, 133, 142]. Another regulator which is very similar to SinI that has been identified in *B. subtilis* is SlrA which act as a positive regulator of biofilm formation by directly inhibiting SinR [142].

Including motility, the c-di-GMP concentrations are also known to influence biofilm formation in *B. subtilis* and *C. difficile*. Studies show that overexpression of DGCs in *C. difficile* resulted in an increased intracellular concentration of c-di-GMP and robust pilus dependent biofilm formation [145, 147]. SinR is the principal regulator of biofilm formation in *B. subtilis*, and its role has been extensively studied. However, the function of SinR in *C. difficile* biofilm

formation remains unknown. Homologous genes regulating the EPS component and protein component of the matrix is absent in *C. difficile*, and blast search also does not reveal sequence similar to SlrR. However, it is very likely that SinR' in *C. difficile* could replace the role of SlrR and based on the domain homology, we hypothesize that SinR/SinR' interaction could be similar to SinR/SlrR interaction in *B. subtilis*.

1.8 Aims of the study

C. difficile toxin production, motility, sporulation all occur around the same transition phase. Hence it is essential to understand the links between them, and there seems to be a lot of crosstalk happening between metabolism and virulence gene regulators involved in the three major pathways. Since all these pathways are found to have a direct influence on pathogenicity, it is essential to understand how it is regulated and analyze the inter-regulation between transition phase regulatory pathways in *C. difficile*, which is the main objective of this thesis. The integration between STM pathway is shown by

(i) Redefining the role of *tcdR* on other transition phase regulatory pathways like sporulation.

(ii) Characterizing the pleiotropic role of *sin* locus genes on *C. difficile* pathogenesis.

Analyzing the variant *codY* gene in epidemic strain R20291

1.9 References

1. Chapman, B.C., et al., Fecal microbiota transplant in patients with Clostridium difficile infection: A systematic review. J Trauma Acute Care Surg, 2016. **81**(4): p. 756-64.

2. Moore, S.C., Clostridium difficile: More Challenging than Ever. Crit Care Nurs Clin North Am, 2018. **30**(1): p. 41-53.

3. Ofori, E., et al., Community-acquired Clostridium difficile: Epidemiology, Ribotype, Risk Factors, Hospital and Intensive Care Unit Outcomes, and Current and Emerging Therapies. J Hosp Infect, 2018.

4. Wilcox, M.H., et al., Recurrence of symptoms in Clostridium difficile infection--relapse or reinfection? J Hosp Infect, 1998. **38**(2): p. 93-100.

5. Wistrom, J., et al., Frequency of antibiotic-associated diarrhoea in 2462 antibiotic-treated hospitalized patients: a prospective study. J Antimicrob Chemother, 2001. **47**(1): p. 43-50.

6. Ramos Martinez, A., et al., [Clostridium difficile enteritis]. Gastroenterol Hepatol, 2011. **34**(8): p. 539-45.

7. Hopkins, R.J. and R.B. Wilson, Treatment of recurrent Clostridium difficile colitis: a narrative review. Gastroenterol Rep (Oxf), 2018. **6**(1): p. 21-28.

8. Finney-JMT, Gastro-enterostomy for cicatrizing ulcer of the pylorus. Bull Johns Hopkins Hosp. 1893;4:53.

9. Hall, I.C. and E. O'Toole, Intestinal flora in new-born infants: With a description of a new pathogenic anaerobe, bacillus difficilis. American Journal of Diseases of Children, 1935. **49**(2): p. 390-402.

10. Rousseau, C., et al., Clostridium difficile carriage in healthy infants in the community: a potential reservoir for pathogenic strains. Clin Infect Dis, 2012. **55**(9): p. 1209-15.

11. Tedesco, F.J., R.W. Barton, and D.H. Alpers, Clindamycin-associated colitis. A prospective study. Ann Intern Med, 1974. **81**(4): p. 429-33.

12. Bartlett, J.G., A.B. Onderdonk, and R.L. Cisneros, Clindamycin-associated colitis in hamsters: protection with vancomycin. Gastroenterology, 1977. **73**(4 Pt 1): p. 772-6.

13. Depestel, D.D. and D.M. Aronoff, Epidemiology of Clostridium difficile infection. J Pharm Pract, 2013. **26**(5): p. 464-75.

14. Bartlett, J.G., et al., Antibiotic-associated pseudomembranous colitis due to toxin-producing clostridia. N Engl J Med, 1978. **298**(10): p. 531-4.

15. Gerding, D.N., Clostridium difficile 30 years on: what has, or has not, changed and why? Int J Antimicrob Agents, 2009. **33 Suppl 1**: p. S2-8.

16. Ghose, C., Clostridium difficile infection in the twenty-first century. Emerg Microbes Infect, 2013. **2**(9): p. e62.

17. Sebaihia, M., et al., The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. Nat Genet, 2006. **38**(7): p. 779-86.

18. Dingle, K.E., et al., Clinical Clostridium difficile: clonality and pathogenicity locus diversity. PLoS One, 2011. **6**(5): p. e19993.

19. Knight, D.R., et al., Diversity and Evolution in the Genome of Clostridium difficile. Clin Microbiol Rev, 2015. **28**(3): p. 721-41.

20. Monot, M., et al., Reannotation of the genome sequence of Clostridium difficile strain 630. J Med Microbiol, 2011. **60**(Pt 8): p. 1193-9.

21. Riedel, T., et al., Genome Resequencing of the Virulent and Multidrug-Resistant Reference Strain Clostridium difficile 630. Genome Announc, 2015. **3**(2).

22. Collery, M.M., et al., What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of Clostridium difficile strain 630 and derivatives. Virulence, 2017. **8**(6): p. 767-781.

23. Dannheim, H., et al., Manual curation and reannotation of the genomes of Clostridium difficile 630Deltaerm and C. difficile 630. J Med Microbiol, 2017. **66**(3): p. 286-293.

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

25. Smits, W.K., SNP-ing out the differences: Investigating differences between Clostridium difficile lab strains. Virulence, 2017. **8**(6): p. 613-617.

26. O'Connor, J.R., S. Johnson, and D.N. Gerding, Clostridium difficile infection caused by the epidemic BI/NAP1/027 strain. Gastroenterology, 2009. **136**(6): p. 1913-24.

27. Almalki, Z.S., et al., Off-label use of oral fluoroquinolone antibiotics in outpatient settings in the United States, 2006 to 2012. Pharmacoepidemiol Drug Saf, 2016. **25**(9): p. 1042-51.

28. Linder, J.A., et al., Fluoroquinolone prescribing in the United States: 1995 to 2002. Am J Med, 2005. **118**(3): p. 259-68.

29. Werner, N.L., et al., Unnecessary use of fluoroquinolone antibiotics in hospitalized patients. BMC Infect Dis, 2011. **11**: p. 187.

30. Valiente, E., M.D. Cairns, and B.W. Wren, The Clostridium difficile PCR ribotype 027 lineage: a pathogen on the move. Clin Microbiol Infect, 2014. **20**(5): p. 396-404.

31. Stabler, R.A., et al., Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol, 2009. **10**(9): p. R102.

32. Monegro, A.F. and H. Regunath, Hospital Acquired Infections, in StatPearls. 2018, StatPearls Publishing

StatPearls Publishing LLC.: Treasure Island (FL).

33. Andrejak, M., J.L. Schmit, and A. Tondriaux, The clinical significance of antibiotic-associated pseudomembranous colitis in the 1990s. Drug Saf, 1991. **6**(5): p. 339-49.

34. Durovic, A., A.F. Widmer, and S. Tschudin-Sutter, New insights into transmission of Clostridium difficile infection-narrative review. Clin Microbiol Infect, 2018.

35. McFarland, L.V., et al., Implications of the changing face of Clostridium difficile disease for health care practitioners. Am J Infect Control, 2007. **35**(4): p. 237-53.

36. Crobach, M.J.T., et al., Understanding Clostridium difficile Colonization. Clin Microbiol Rev, 2018. **31**(2).

37. Cherkasskaia, R.S., et al., [Clostridium difficile and diarrhea in infants in the first half-year of life]. Pediatriia, 1992(7-9): p. 15-20.

38. Munoz, M., et al., Community-acquired infection with hypervirulent Clostridium difficile isolates that carry different toxin and antibiotic resistance loci: a case report. Gut Pathog, 2017. **9**: p. 63.

39. Gross, U., et al., 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, 2018. **19**(1): p. 1.

40. Kwon, J.H., M.A. Olsen, and E.R. Dubberke, The morbidity, mortality, and costs associated with Clostridium difficile infection. Infect Dis Clin North Am, 2015. **29**(1): p. 123-34.

41. Gil, F., et al., Clostridioides (Clostridium) difficile infection: current and alternative therapeutic strategies. Future Microbiol, 2018. **13**: p. 469-482.

42. McDonald LC, G.D., Johnson S, et al. , 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 Disease., 2018.

43. Bartlett, J.G., Clostridium difficile Infection. Infect Dis Clin North Am, 2017. **31**(3): p. 489-495.

44. Martin, M., et al., National European guidelines for the prevention of Clostridium difficile infection: a systematic qualitative review. J Hosp Infect, 2014. **87**(4): p. 212-9.

45. Bui, C., et al., Antimicrobial stewardship programs that target only high-cost, broad-spectrum antimicrobials miss opportunities to reduce Clostridium difficile infections. Am J Infect Control, 2016. **44**(12): p. 1684-1686.

46. Stites, S.D., et al., The tipping point: patients predisposed to Clostridium difficile infection and a hospital antimicrobial stewardship programme. J Hosp Infect, 2016. **94**(3): p. 242-248.

47. Sartelli, M., et al., WSES guidelines for management of Clostridium difficile infection in surgical patients. World J Emerg Surg, 2015. **10**: p. 38.

48. Spellberg, B., A. Srinivasan, and H.F. Chambers, New Societal Approaches to Empowering Antibiotic Stewardship. Jama, 2016. **315**(12): p. 1229-30.

49. Knight, N., et al., Clostridium difficile colitis: a retrospective study of incidence and severity before and after institution of an alcohol-based hand rub policy. Am J Infect Control, 2010. **38**(7): p. 523-8.

50. Ayyagari, A., J. Agarwal, and A. Garg, Antibiotic associated diarrhoea: infectious causes. Indian J Med Microbiol, 2003. **21**(1): p. 6-11.

51. Eaton, S.R. and J.E. Mazuski, Overview of severe Clostridium difficile infection. Crit Care Clin, 2013. **29**(4): p. 827-39.

52. Potter, V.A. and A. Aravinthan, Identifying patients at risk of severe Clostridium difficile-associated disease. Br J Hosp Med (Lond), 2012. **73**(5): p. 265-70.

53. Napolitano, L.M. and C.E. Edmiston, Jr., Clostridium difficile disease: Diagnosis, pathogenesis, and treatment update. Surgery, 2017. **162**(2): p. 325-348.

54. Vaishnavi, C., Clostridium difficile infection: clinical spectrum and approach to management. Indian J Gastroenterol, 2011. **30**(6): p. 245-54.

55. Sartelli, M., et al., Antimicrobials: a global alliance for optimizing their rational use in intra-abdominal infections (AGORA). World J Emerg Surg, 2016. **11**: p. 33.

56. Vaishnavi, C., Fecal microbiota transplantation for management of Clostridium difficile infection. Indian J Gastroenterol, 2014. **33**(4): p. 301-7.

57. Gil, F., et al., Updates on Clostridium difficile spore biology. Anaerobe, 2017. **45**: p. 3-9.

58. Doona, C.J., et al., Effects of High-Pressure Treatment on Spores of Clostridium Species. Appl Environ Microbiol, 2016. **82**(17): p. 5287-97.

59. Edwards, A.N., et al., Chemical and Stress Resistances of Clostridium difficile Spores and Vegetative Cells. Front Microbiol, 2016. **7**: p. 1698.

60. Swick, M.C., T.M. Koehler, and A. Driks, Surviving Between Hosts: Sporulation and Transmission. Microbiol Spectr, 2016. **4**(4).

61. Bhattacharjee, D., K.N. McAllister, and J.A. Sorg, Germinants and Their Receptors in Clostridia. J Bacteriol, 2016. **198**(20): p. 2767-75.

62. Sorg, J.A. and A.L. Sonenshein, Bile salts and glycine as cogerminants for Clostridium difficile spores. J Bacteriol, 2008. **190**(7): p. 2505-12.

63. Blanchi, J., J. Goret, and F. Megraud, Clostridium difficile Infection: A Model for Disruption of the Gut Microbiota Equilibrium. Dig Dis, 2016. **34**(3): p. 217-20.

64. Solomon, K., et al., Mortality in patients with Clostridium difficile infection correlates with host pro-inflammatory and humoral immune responses. J Med Microbiol, 2013. **62**(Pt 9): p. 1453-60.

65. Yu, H., et al., Cytokines Are Markers of the Clostridium difficile-Induced Inflammatory Response and Predict Disease Severity. Clin Vaccine Immunol, 2017. **24**(8).

66. Roy Chowdhury, P., et al., Comparative genomic analysis of toxin-negative strains of Clostridium difficile from humans and animals with symptoms of gastrointestinal disease. BMC Microbiol, 2016. **16**: p. 41.

67. Monot, M., et al., Clostridium difficile: New Insights into the Evolution of the Pathogenicity Locus. Sci Rep, 2015. **5**: p. 15023.

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

69. Janezic, S., et al., A new type of toxin A-negative, toxin B-positive Clostridium difficile strain lacking a complete tcdA gene. J Clin Microbiol, 2015. **53**(2): p. 692-5.

70. Rupnik, M., et al., New types of toxin A-negative, toxin B-positive strains among Clostridium difficile isolates from Asia. J Clin Microbiol, 2003. **41**(3): p. 1118-25.

71. Kuehne, S.A., S.T. Cartman, and N.P. Minton, Both, toxin A and toxin B, are important in Clostridium difficile infection. Gut Microbes, 2011. **2**(4): p. 252-5.

72. Pruitt, R.N., et al., Structural organization of the functional domains of Clostridium difficile toxins A and B. Proc Natl Acad Sci U S A, 2010. **107**(30): p. 13467-72.

73. Pruitt, R.N. and D.B. Lacy, Toward a structural understanding of Clostridium difficile toxins A and B. Front Cell Infect Microbiol, 2012. **2**: p. 28.

74. Jank, T. and K. Aktories, Structure and mode of action of clostridial glucosylating toxins: the ABCD model. Trends Microbiol, 2008. **16**(5): p. 222-9.

75. Chandrasekaran, R. and D.B. Lacy, The role of toxins in Clostridium difficile infection. FEMS Microbiol Rev, 2017. **41**(6): p. 723-750.

76. Egerer, M., et al., Autocatalytic processing of Clostridium difficile toxin B. Binding of inositol hexakisphosphate. J Biol Chem, 2009. **284**(6): p. 3389-95.

77. Pruitt, R.N., et al., Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in Clostridium difficile toxin A. J Biol Chem, 2009. **284**(33): p. 21934-40.

78. Aktories, K., Bacterial protein toxins that modify host regulatory GTPases. Nat Rev Microbiol, 2011. **9**(7): p. 487-98.

79. Jank, T., T. Giesemann, and K. Aktories, Rho-glucosylating Clostridium difficile toxins A and B: new insights into structure and function. Glycobiology, 2007. **17**(4): p. 15r-22r.

80. Chumbler, N.M., et al., Clostridium difficile Toxins TcdA and TcdB Cause Colonic Tissue Damage by Distinct Mechanisms. Infect Immun, 2016. **84**(10): p. 2871-7.

81. Awad, M.M., et al., Clostridium difficile virulence factors: Insights into an anaerobic spore-forming pathogen. Gut Microbes, 2014. **5**(5): p. 579-93.

82. Elliott, B., et al., The complexity and diversity of the Pathogenicity Locus in Clostridium difficile clade 5. Genome Biol Evol, 2014. **6**(12): p. 3159-70.

83. Moncrief, J.S., L.A. Barroso, and T.D. Wilkins, Positive regulation of Clostridium difficile toxins. Infect Immun, 1997. **65**(3): p. 1105-8.

84. Matamouros, S., P. England, and B. Dupuy, Clostridium difficile toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol, 2007. **64**(5): p. 1274-88.

85. Govind, R. and B. Dupuy, Secretion of Clostridium difficile toxins A and B requires the holin-like protein TcdE. PLoS Pathog, 2012. **8**(6): p. e1002727.

86. Hundsberger, T., et al., Transcription analysis of the genes tcdA-E of the pathogenicity locus of Clostridium difficile. Eur J Biochem, 1997. **244**(3): p. 735-42.

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

88. Mani, N., et al., Environmental Response and Autoregulation of Clostridium difficile TxeR, a Sigma Factor for Toxin Gene Expression. Journal of Bacteriology, 2002. **184**(21): p. 5971-5978.

89. Mani, N., B. Dupuy, and A.L. Sonenshein, Isolation of RNA polymerase from Clostridium difficile and characterization of glutamate dehydrogenase and rRNA gene promoters in vitro and in vivo. J Bacteriol, 2006. **188**(1): p. 96-102.

90. Dupuy, B. and S. Matamouros, Regulation of toxin and bacteriocin synthesis in Clostridium species by a new subgroup of RNA polymerase sigma-factors. Res Microbiol, 2006. **157**(3): p. 201-5.

91. Karlsson, S., et al., Expression of Clostridium difficile Toxins A and B and Their Sigma Factor TcdD Is Controlled by Temperature. Infection and Immunity, 2003. **71**(4): p. 1784-1793.

92. Carter, G.P., et al., The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of Clostridium difficile. PLoS Pathog, 2011. **7**(10): p. e1002317.

93. Antunes, A., et al., Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. Nucleic Acids Res, 2012. **40**(21): p. 10701-18.

94. Antunes, A., I. Martin-Verstraete, and B. Dupuy, CcpA-mediated repression of Clostridium difficile toxin gene expression. Mol Microbiol, 2011. **79**(4): p. 882-99.

95. Dineen, S.S., S.M. McBride, and A.L. Sonenshein, Integration of metabolism and virulence by Clostridium difficile CodY. J Bacteriol, 2010. **192**(20): p. 5350-62.

96. El Meouche, I., et al., Characterization of the SigD regulon of C. difficile and its positive control of toxin production through the regulation of tcdR. PLoS One, 2013. **8**(12): p. e83748.

97. McKee, R.W., et al., The second messenger cyclic Di-GMP regulates Clostridium difficile toxin production by controlling expression of sigD. J Bacteriol, 2013. **195**(22): p. 5174-85.

98. Deakin, L.J., et al., The Clostridium difficile spo0A gene is a persistence and transmission factor. Infect Immun, 2012. **80**(8): p. 2704-11.

99. Pettit, L.J., et al., Functional genomics reveals that Clostridium difficile Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics, 2014. **15**: p. 160.

100. Rosenbusch, K.E., et al., C. difficile 630Deltaerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. PLoS One, 2012. **7**(10): p. e48608.

101. Mackin, K.E., et al., Spo0A differentially regulates toxin production in evolutionarily diverse strains of Clostridium difficile. PLoS One, 2013. **8**(11): p. e79666.

102. Saujet, L., et al., The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in Clostridium difficile. J Bacteriol, 2011. **193**(13): p. 3186-96.

103. Francis, M.B., C.A. Allen, and J.A. Sorg, Spore Cortex Hydrolysis Precedes Dipicolinic Acid Release during Clostridium difficile Spore Germination. J Bacteriol, 2015. **197**(14): p. 2276-83.

104. Francis, M.B. and J.A. Sorg, Dipicolinic Acid Release by Germinating Clostridium difficile Spores Occurs through a Mechanosensing Mechanism. mSphere, 2016. **1**(6).

105. Paredes-Sabja, D., A. Shen, and J.A. Sorg, Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. Trends Microbiol, 2014. **22**(7): p. 406-16.

106. Zhu, D., J.A. Sorg, and X. Sun, Clostridioides difficile Biology: Sporulation, Germination, and Corresponding Therapies for C. difficile Infection. Front Cell Infect Microbiol, 2018. **8**: p. 29.

107. Janoir, C., et al., Adaptive strategies and pathogenesis of Clostridium difficile from in vivo transcriptomics. Infect Immun, 2013. **81**(10): p. 3757-69.

108. Bakker, D., et al., The HtrA-like protease CD3284 modulates virulence of Clostridium difficile. Infect Immun, 2014. **82**(10): p. 4222-32.

109. McBride, S.M. and A.L. Sonenshein, The dlt operon confers resistance to cationic antimicrobial peptides in Clostridium difficile. Microbiology, 2011. **157**(Pt 5): p. 1457-65.

110. Nawrocki, K.L., E.K. Crispell, and S.M. McBride, Antimicrobial Peptide Resistance Mechanisms of Gram-Positive Bacteria. Antibiotics (Basel), 2014. **3**(4): p. 461-92.

111. Al-Hinai, M.A., S.W. Jones, and E.T. Papoutsakis, The Clostridium sporulation programs: diversity and preservation of endospore differentiation. Microbiol Mol Biol Rev, 2015. **79**(1): p. 19-37.

112. Fimlaid, K.A., et al., Global analysis of the sporulation pathway of Clostridium difficile. PLoS Genet, 2013. **9**(8): p. e1003660.

113. Donnelly, M.L., K.A. Fimlaid, and A. Shen, Characterization of Clostridium difficile Spores Lacking Either SpoVAC or Dipicolinic Acid Synthetase. J Bacteriol, 2016. **198**(11): p. 1694-707.

114. De Santis, S., et al., Nutritional Keys for Intestinal Barrier Modulation. Front Immunol, 2015. **6**: p. 612.

115. Strauch, M.A. and J.A. Hoch, Signal transduction in Bacillus subtilis sporulation. Curr Opin Genet Dev, 1993. **3**(2): p. 203-12.

116. Underwood, S., et al., Characterization of the sporulation initiation pathway of Clostridium difficile and its role in toxin production. J Bacteriol, 2009. **191**(23): p. 7296-305.

117. Fimlaid, K.A. and A. Shen, Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. Curr Opin Microbiol, 2015. **24**: p. 88-95.

118. Cervin, M.A., et al., The Bacillus subtilis regulator SinR inhibits spoIIG promoter transcription in vitro without displacing RNA polymerase. Nucleic Acids Res, 1998. **26**(16): p. 3806-12.

119. Ababneh, Q.O. and J.K. Herman, CodY Regulates SigD Levels and Activity by Binding to Three Sites in the fla/che Operon. J Bacteriol, 2015. **197**(18): p. 2999-3006.

120. Belitsky, B.R., et al., Interactive regulation by the Bacillus subtilis global regulators CodY and ScoC. Mol Microbiol, 2015. **97**(4): p. 698-716.

121. Phillips, Z.E. and M.A. Strauch, Bacillus subtilis sporulation and stationary phase gene expression. Cell Mol Life Sci, 2002. **59**(3): p. 392-402.

122. Paredes, C.J., K.V. Alsaker, and E.T. Papoutsakis, A comparative genomic view of clostridial sporulation and physiology. Nat Rev Microbiol, 2005. **3**(12): p. 969-78.

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

124. Lewis, R.J., et al., Crystallisation of the Bacillus subtilis sporulation inhibitor SinR, complexed with its antagonist, SinI. FEBS Lett, 1996. **378**(1): p. 98-100.

125. Margot, P., V. Lazarevic, and D. Karamata, Effect of the SinR protein on the expression of the Bacillus subtilis 168 lytABC operon. Microb Drug Resist, 1996. **2**(1): p. 119-21.

126. Newman, J.A., C. Rodrigues, and R.J. Lewis, Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in Bacillus subtilis. J Biol Chem, 2013. **288**(15): p. 10766-78.

127. Rashid, M.H. and J. Sekiguchi, flaD (sinR) mutations affect SigD-dependent functions at multiple points in Bacillus subtilis. J Bacteriol, 1996. **178**(22): p. 6640-3.

128. Fujita, Y., et al., Dual Regulation of Bacillus subtilis kinB Gene Encoding a Sporulation Trigger by SinR through Transcription Repression and Positive Stringent Transcription Control. Front Microbiol, 2017. **8**: p. 2502.

129. Zhang, B., P. Struffi, and L. Kroos, sigmaK can negatively regulate sigE expression by two different mechanisms during sporulation of Bacillus subtilis. J Bacteriol, 1999. **181**(13): p. 4081-8.

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

131. Dixon, L.G., et al., Developmental gene expression in Bacillus subtilis crsA47 mutants reveals glucose-activated control of the gene for the minor sigma factor sigma(H). J Bacteriol, 2001. **183**(16): p. 4814-22.

132. Chai, Y., R. Kolter, and R. Losick, Reversal of an epigenetic switch governing cell chaining in Bacillus subtilis by protein instability. Mol Microbiol, 2010. **78**(1): p. 218-29.

133. Chai, Y., et al., An epigenetic switch governing daughter cell separation in Bacillus subtilis. Genes Dev, 2010. **24**(8): p. 754-65.

134. Chai, Y., et al., Evidence that metabolism and chromosome copy number control mutually exclusive cell fates in Bacillus subtilis. EMBO J, 2011. **30**(7): p. 1402-13.

135. Girinathan, B.P., et al., Pleiotropic roles of Clostridium difficile sin locus. PLoS Pathog, 2018. **14**(3): p. e1006940.

136. Edwards, A.N., K.L. Nawrocki, and S.M. McBride, Conserved oligopeptide permeases modulate sporulation initiation in Clostridium difficile. Infect Immun, 2014. **82**(10): p. 4276-91.

137. Marquez-Magana, L.M. and M.J. Chamberlin, Characterization of the sigD transcription unit of Bacillus subtilis. J Bacteriol, 1994. **176**(8): p. 2427-34.

138. Estacio, W., et al., Dual promoters are responsible for transcription initiation of the fla/che operon in Bacillus subtilis. J Bacteriol, 1998. **180**(14): p. 3548-55.

139. Fredrick, K. and J.D. Helmann, FlgM is a primary regulator of sigmaD activity, and its absence restores motility to a sinR mutant. J Bacteriol, 1996. **178**(23): p. 7010-3.

140. West, J.T., W. Estacio, and L. Marquez-Magana, Relative roles of the fla/che P(A), P(D-3), and P(sigD) promoters in regulating motility and sigD expression in Bacillus subtilis. J Bacteriol, 2000. **182**(17): p. 4841-8.

141. Chu, F., et al., A novel regulatory protein governing biofilm formation in Bacillus subtilis. Mol Microbiol, 2008. **68**(5): p. 1117-27.

142. Chai, Y., R. Kolter, and R. Losick, Paralogous antirepressors acting on the master regulator for biofilm formation in Bacillus subtilis. Mol Microbiol, 2009. **74**(4): p. 876-87.

143. Edwards, A.N., R. Tamayo, and S.M. McBride, A novel regulator controls Clostridium difficile sporulation, motility and toxin production. Mol Microbiol, 2016. **100**(6): p. 954-71.

144. Walter, B.M., et al., The SOS Response Master Regulator LexA Is Associated with Sporulation, Motility and Biofilm Formation in Clostridium difficile. PLoS One, 2015. **10**(12): p. e0144763.

145. Chen, Y., et al., Evidence for cyclic Di-GMP-mediated signaling in Bacillus subtilis. J Bacteriol, 2012. **194**(18): p. 5080-90.

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

147. Purcell, E.B., et al., A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls Clostridium difficile Biofilm and Toxin Production during Stationary Phase. Infect Immun, 2017. **85**(9).

148. Purcell, E.B., et al., Cyclic diguanylate inversely regulates motility and aggregation in Clostridium difficile. J Bacteriol, 2012. **194**(13): p. 3307-16.

149. Branda, S.S., et al., A major protein component of the Bacillus subtilis biofilm matrix. Mol Microbiol, 2006. **59**(4): p. 1229-38.

150. Chu, F., et al., Targets of the master regulator of biofilm formation in Bacillus subtilis. Mol Microbiol, 2006. **59**(4): p. 1216-28.

151. Kearns, D.B., et al., A master regulator for biofilm formation by Bacillus subtilis. Mol Microbiol, 2005. **55**(3): p. 739-49.

152. Vlamakis, H., et al., Sticking together: building a biofilm the Bacillus subtilis way. Nat Rev Microbiol, 2013. **11**(3): p. 157-68.

153. Colledge, V.L., et al., Structure and organisation of SinR, the master regulator of biofilm formation in Bacillus subtilis. J Mol Biol, 2011. **411**(3): p. 597-613.

154. Dubnau, E.J., et al., A protein complex supports the production of Spo0A-P and plays additional roles for biofilms and the K-state in Bacillus subtilis. Mol Microbiol, 2016. **101**(4): p. 606-24.

155. Chai, Y., et al., Bistability and biofilm formation in Bacillus subtilis. Mol Microbiol, 2008. **67**(2): p. 254-63.

Chapter 2 - Effect of *tcdR* mutation on sporulation in the epidemic *Clostridium difficile* R20291 strain

2.1 Publication arising from this chapter

The key findings from this chapter has resulted in the following publication (see Appendix A)

Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, Govind R. Effect of *tcdR* mutation on sporulation in the epidemic *Clostridium difficile* strain R20291. mSphere, 2017 .2(1):e00383-16.

2.2 Abstract

Clostridium difficile is an important nosocomial pathogen and the leading cause of hospital-acquired diarrhea. Antibiotic use is the primary risk factor for the development of *C. difficile*-associated disease because it disrupts normally protective gut flora and enables *C. difficile* to colonize the colon. *C. difficile* damages host tissue by secreting toxins and disseminates by forming spores. The toxin-encoding genes, *tcdA* and *tcdB*, are part of a pathogenicity locus, which also includes the gene *tcdR* that codes for TcdR, an alternate sigma factor that initiates transcription of *tcdA* and *tcdB* genes. We created a *tcdR* mutant in the epidemic-type *C. difficile* R20291 strain in an attempt to identify the global role of *tcdR*. A site-directed mutation in *tcdR* affected both toxin production and sporulation in *C. difficile* R20291. Spores of the *tcdR* mutant were more heat-

sensitive than the wild type. Nearly three-fold more taurocholate was needed to germinate spores from the *tcdR* mutant than the spores prepared from the WT strain. Transmission Electron Microscopic analysis of the spores also revealed a weakly assembled exosporium on the *tcdR* mutant spores. Accordingly, comparative transcriptome analysis showed many differentially expressed sporulation genes in the *tcdR* mutant when compared to the WT strain. These data suggests that regulatory networks of toxin production and sporulation in *C. difficile* R20291 strain are linked with each other.

2.3 Importance

C. difficile infects thousands of hospitalized patients every year causing significant morbidity and mortality. *C. difficile* spores play a pivotal role in the transmission of the pathogen in the hospital environment. During infection, the spores germinate, and the vegetative bacterial cells produce toxins that damage host tissue. Thus, sporulation and toxin production are two important traits of *C. difficile*. In this study, we show that a mutation in *tcdR*, the toxin gene regulator affects both toxin production and sporulation in the epidemic-type *C. difficile* R20291 strain.

2.4 Introduction

Clostridium difficile is a Gram-positive, spore-forming, anaerobic bacillus, and is the leading cause of hospital-acquired diarrheal diseases (1, 2). Nearly 50% of all patients carry *C. difficile* asymptomatically after hospitalization (2, 3). Nearly 10% of all *C. difficile* infected patients develop pseudomembranous colitis, and 3% develop severe, life-threatening

complications such as fulminant colitis and toxic megacolon (4). *C. difficile* infection (CDI) is commonly acquired from *C. difficile* spores present in the hospital environment, and individuals become infected when the normal colonic microbiota is suppressed by antibiotic therapy (5). In the gut, *C. difficile* spores germinate to the toxin-producing vegetative form in response to certain bile acids, *e.g.*, taurocholic acid, and amino acids. *C. difficile* toxins A (TcdA) and B (TcdB) are then secreted from the vegetative cell and cause tissue damage, necrosis, and inflammation, and are the main reason for this disease outcome (6).

In *C. difficile*, the toxin genes, *tcdA* and *tcdB*, are encoded within a 19-kb pathogenicity locus (PaLoc) and the *tcdR* gene, encoded upstream of *tcdB,* is required for expression of the toxin genes. TcdR is an alternate sigma factor that directs transcription by recruiting RNA polymerase to the toxin gene promoters and its own promoter (7, 8). Previous studies have reported that other proteins can regulate toxin gene expression in response to different environmental stimuli by controlling the transcription of *tcdR*. The sigma factor SigD positively regulates toxin production by controlling the transcription of *tcdR* (9). CodY, a global transcriptional regulator, represses the toxin gene expression by binding with high affinity to the *tcdR* promoter region (10, 11). Finally, in response to sugar availability, CcpA, a major regulator of carbon catabolite repression binds to the promoter region or the 5' ends of several PaLoc genes, with the strongest affinity to the promoter region of *tcdR* (12, 13).

TcdR was the first described member of the group V family of alternative sigma factors (14). We recently determined that TcsR, a toxin gene regulator in *Clostridium sordellii*, is also a member of this family of sigma factors (15). Most of these alternative sigma factors are auto-regulated (7, 16) and are induced by environmental stresses, such as nutritional limitation, DNA

damage, or non-optimal temperatures (8, 14, 17) suggesting that these sigma factors function during these sub-optimal growth conditions.

In this study, we created and characterized a mutation in *tcdR* in the epidemic-type *C. difficile* R20291 strain to determine whether TcdR influenced cellular processes other than toxin production. We found that the *tcdR* mutant sporulated less efficiently than the WT strain. Moreover, spores prepared from the *tcdR* mutant were more heat-sensitive and had lower germination efficiency than the wildtype, parental strain. Electron microscopic analysis of the *tcdR* mutant spores also revealed a weakly assembled exosporium. In agreement with these findings, comparative RNA-seq analyses of the WT and the *tcdR* mutant strains revealed several sporulation genes to be affected by the *tcdR* mutation. These results suggested that a mutation in *tcdR* not only affects toxin production but also influences the sporulation pathway in the *C. difficile* R20291 strain. Interestingly, mutating *tcdR* in the *C. difficile* 630Δ*erm* strain, however, did not result in this phenotype suggesting that the TcdR regulon may be strain-specific.

2.5 Aim of the work described in this chapter

In this chapter, I have used genetic, biochemical and computational tools to decipher the effect of *tcdR* on other transition phase regulatory pathways like sporulation.

2.6 Materials and methods

2.6.1 Bacterial strains and growth conditions

Clostridium difficile strains (Table 2.1)were grown anaerobically in TY (Tryptose, Yeast extract) or 70:30 medium (18) as described previously (15, 19). Cefoxitin (Cef; 25 µg/mL), thiamphenicol

(Thio; 15 μg/mL) and lincomycin (Lin; 15 μg/mL) were added to *C. difficile* cultures whenever necessary. *Escherichia coli* strains were grown in (LB) broth. *E. coli* strain S17-1 (20), used for conjugation, was supplemented with ampicillin (100 μg/mL) or chloramphenicol (25 μg/mL) when indicated and cultured aerobically in LB broth.

Table 2.1 Bacterial Strains and Plasmids used in this study

Strains/ Plasmids used	Description	Reference
C. difficile R20291	NAP1/027 ribotype	(40)
C. difficile R20291::*tcdR*	R20291with intron insertion in *tcdR* gene	This study
*C. difficile 630*Δ*erm*	Erm^s derivative of strain 630	(63)
*C. difficile 630*Δ*erm*::*tcdR*	*630*Δ*erm* with intron insertion in *tcdR* gene	This study
E. coli DH5α	*endA1 recA1 deoR hsdR17* (r_K^- m_K^+)	NEB labs
E. coli S17-1	Strain with integrated RP4 conjugation transfer function for conjugation between *E. coli* and *C. difficile*	(20)
E. coli GM241(DE3)	*gusA* mutant lysogenized with DE3 phage and host for *gusA* reporter plasmids	(19)
pMTL007-CE5	ClosTron plasmid	(21)
pMTL007-CE5::*tcdR*-141	pMTL007-CE5 carrying *tcdR*-specific intron	This study
pRPF185	*C. difficile* shuttle vector	(64)
pRGL294	pRPF185 with *tcdR* expressed from its own promoter	This study
pACYC184	*E. coli* cloning vector; compatible with pET16B	NEB
pACYC515	pACYC184 vector carrying *gusA* gene under *tcdR promoter*	(19)
pET16b	*E. coli* expression vector	Novagen
pRGL312	pET16B with *tcdR*	This study
pRGL320	pACYC184 vector carrying *gusA* gene under *bclA2 promoter*	This study
pRGL321	pACYC184 vector carrying *gusA* gene under *bclA3 promoter*	This study
C. difficile R20291::*tcdR* +pRGL294	R20291::*tcdR complemented with tcdR*	This study
C. difficile R20291::*tcdR* +pRPF185	R20291::*tcdR with vector control*	This study

2.6.2 Construction of a *tcdR* mutant

A *tcdR* mutation was constructed in *C. difficile* strain using the ClosTron gene knockout system (21). The group II intron insertion site in the antisense orientation between nucleotides 141

and 142 of the *tcdR* ORF was selected using the Perutka algorithm, a Web-based design tool available at http://www.clostron.com. The designed retargeted intron was cloned into pMTL007-CE5 and the resulting plasmid pMTL007-CE5::Cdi-*tcdR*-141a was transferred into R20291 by conjugation as described previously (15, 22). The selection of thiamphenicol-resistant transconjugants in 15 µg ml^{-1} lincomycin plates confers potential Ll.ltrB insertions within the target *tcdR* gene in the chromosome of R20291. The presence of a putative *tcdR* mutant was identified by PCR using *tcdR*-specific primers (Table S1) in combination with the EBSu universal and ERM primers. Specific single integration of the group II intron into the genome was verified by Southern blot using ([^{32}P]dATP) radiolabeled probe specific for the *tcdR* gene as described previously (15, 22). Complementation of the *C. difficile* R20291::*tcdR* mutant is described in the supplementary methods.

2.6.3 Toxin Assays

Cultures of R20291 and R20291::*tcdR* mutant were centrifuged after 10 hours in TY medium, and toxin ELISAs were performed as described previously (15). Details are presented in supplementary methods.

2.6.4 Sporulation Assay (Microscopic analysis)

C. difficile cultures were grown overnight in TY medium supplemented with 0.1% taurocholate to induce germination of any spores that were present. Cells were then diluted in TY medium to an OD_{600} nm of 0.5 and then 100 µl was spread on 70:30 sporulation agar (18). Plates were incubated at 37°C and monitored for the production of spores. Cells were harvested from the plates after 24 hours and were suspended in TY medium for phase-contrast microscopy as described previously (18). At least four fields per strain were obtained, and the number of spores

and vegetative cells were counted to calculate the percentage of spores based on the total number of spores and vegetative cells. Experiments were performed at least three independent times.

2.6.5 Sporulation Assay (Ethanol-resistance method)

C. difficile strains were inoculated and grown on 70:30 sporulation agar as described above. After 24 hr growth, cells were scraped from the plates and suspended in 70:30 sporulation liquid medium to an OD_{600} = 1.0. Cells were immediately serially diluted and plated onto TY agar + 0.1% taurocholate to enumerate viable vegetative cells and spores. A 0.5 ml aliquot of the culture was removed from the chamber, mixed with 0.5 ml of 95% ethanol, vortexed and incubated at room temperature for 15 min. Ethanol-treated cells were serially diluted in 1X PBS, returned to the anaerobic chamber and plated onto TY + 0.1% taurocholate plates to enumerate spores. After 24 hours of growth, CFU were enumerated, and the percentage sporulation was calculated as the number of ethanol-resistant spores divided by the total number of viable cells (vegetative cells and spores).

2.6.6 Spore preparation

Spores were generated and purified as previously described (23, 24). Details are presented in the supplementary methods section.

2.6.7 RNA-seq Analysis and Quantitative reverse transcription PCR (qRT-PCR)

RNA-seq analysis was performed at the DNA Core facility at the University of Missouri and the data was analyzed using methods described previously (25-27). Details of RNA seq and the qRT-PCR (28, 29) are detailed in the supplementary methods.

2.6.8 Germination

Purified *C. difficile* spores were heat activated at 65 °C for 30 min and then placed on ice. Ten microliters of the heat-activated spores were added to a final optical density at 600 nm (OD_{600}) of 0.5 in 990 µl of BHIS medium alone or medium supplemented with 2, 5, 10, 20, or 50 mM of taurocholic acid (TA). Germination was monitored at 600 nm for 30 minutes in a Perkin Elmer (Waltham, MA) Lambda25 UV/Vis spectrophotometer. The data points at OD_{600} (T_x) were normalized to the starting OD_{600} value (T_0). The germination rates and the Effective Concentration (EC_{50}) were calculated using the slopes of the linear portions of the germination plots as described previously (23, 30). EC_{50} is the concentration of germinant needed to reach 50% of the maximum germination rate. EC_{50} values were individually calculated from each germination experiment and are reported as averages with the standard error of the mean.

2.6.9 Spores heat resistance

Purified spores (nearly 1 x 10^5) prepared as above were resuspended in 500 µL of water and incubated at 70 °C. Samples were removed at 0.5 hours, 4 hours and 8 hours, serially diluted in PBS, plated onto TY agar plates with 0.1% taurocholate and grown anaerobically for 48 hours before counting (31, 32). As a control for non-heat-treated spores, an aliquot was plated onto TY+ 0.1% taurocholate agar plates prior to the experiment and colonies were counted as described above.

2.6.10 Transmission Electron Microscopy

All steps in sample preparation were performed at room temperature using pelleted spores in 1.5 microcentifuge tube and solutions were prepared in 1X PBS unless indicated otherwise. For transmission electron microscopy, spores (10^{10}) were fixed for 2 hrs in a solution of 2% glutaraldehyde and 2% paraformaldehyde. The spores were thoroughly rinsed in 1X PBS (5 minutes each) and post fixed with 1% osmium tetroxide with constant rotation for 1-2 hours. The samples were then washed thrice with 1X PBS (5 minutes each) and enblock stained with 2% aqueous Uranyl acetate for 1hr light protected and then washed three times (5 min each) with distilled water. The spores were further dehydrated in a graded 50% -95% acetone series (vol/vol) for 5 minutes and left in 100% acetone overnight. Infiltration was carried out in graded acetone/EMBED 812/Araldite resin (Electron Microscopy Sciences) ratio (1:1, 1:2) for 10min each at RT with constant rotation and incubated in 100% resin overnight. Resin was cured at 60^0C for 24-48h and thin sections (silver to gold color) cut and absorbed onto on 200 mesh copper grids. Sections were examined with a transmission electron microscope (CM100, FEI Company) at 100 kV and images were captured using a side mounted Hamamatsu digital camera model C8484 with AMT image capture software version 602.591n.

2.7 Results

2.7.1 Mutation in *tcdR* affects both toxin production and sporulation in *C. difficile* R20291 strain

To analyze the global role of *tcdR* in *C. difficile* R20291 strain, we used the ClosTron system (21) to inactivate the *tcdR* gene. Insertion of the group II intron into the target gene (Figure

S 2.1A) was verified by PCR using intron-specific primers and *tcdR* gene-specific primers (Figure S 2.1B, Table S 2.1). A Southern blot confirmed the single chromosomal insertion of the intron in the *tcdR* gene (Figure S 2.1C). Growth kinetics were performed and indicated that the inactivation of the *tcdR* gene did not affect the normal growth of the bacterium (Figure 2.1A). A Toxin ELISA was performed with the cytosolic protein extracts of *tcdR* mutant and WT strain. We observed a dramatic reduction in toxin production (Figure 2.1B) in the mutant compared to WT supporting the previously known function of TcdR as a positive regulator of the toxin genes (7, 8, 16). Further, we measured the sporulation efficiency of the *tcdR* mutant at a 24-hour time point. Nearly three-fold reduction in ethanol-resistant spores was observed in the *tcdR* mutant when compared to the WT strain (Figure 2.2A). A similarly reduced sporulation rate (~2.6 fold) was observed when the number of sporulation cells in the population was counted microscopically (Figure 2.2B). We then complemented the *tcdR* mutant by cloning and expressing *tcdR* from its own promoter. Toxin production in the complemented strain was fully recovered (Figure 1B), whereas the effect on sporulation could be only partially restored (Figure 2.2 A and Figure 2.2B). Unlike toxin gene regulation (where TcdR directly regulates *tcdA* and *tcdB* transcription), sporulation is regulated by multiple transcription factors and alternative RNA polymerase sigma factors (29, 33, 34). . Sporulation also involves finely tuned spatially and temporally regulated gene expression programs and may not be mimicked exactly in the complemented. All of these regulatory mechanisms could result in partial complementation of the sporulation. Another explanation could be that, when TcdR sigma factor is over-expressed, the availability of RNA core polymerase for other sigma factors needed for sporulation could be limited and could result in partial complementation of the sporulation phenotype.

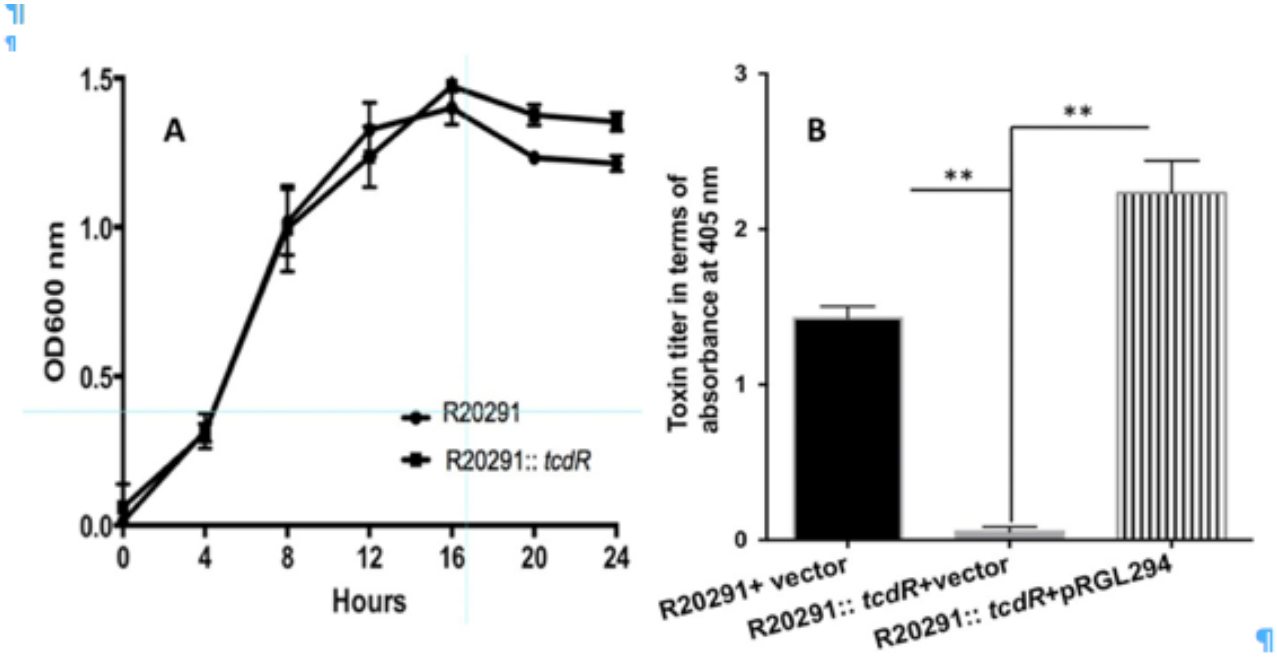


Figure 2.1. Effect of *tcdR* inactivation on bacterial growth kinetics and toxin production.
(A) Growth curve of R20291 and R20291::*tcdR* in TY medium. (B) TcdA and TcdB levels in cytosolic fractions after 10h growth. *C. difficile* strains were grown in TY medium and toxins were quantified using ELISA. The data represent the averages of the results of three independent assays. Error bars both panel A and B correspond to the standard error of the means. The asterisks (**) in panel B indicates statistical difference at $p < 0.005$.

2.7.2 Transcriptome analysis of *tcdR* mutant

Global regulators, *ccpA* and *codY,* are known to influence both sporulation and toxin production in *C. difficile* (10, 11). We performed qRT-PCR analysis and found no significant change in their transcript levels in the *tcdR* mutant when compared to the WT strain (Figure S 2.2). Since this initial analysis failed to explain the reasons behind the unexpected phenotype of the *tcdR* mutant, we decided to perform a transcriptome study using RNA-seq analysis. RNAs were prepared from stationary phase cultures of the *tcdR* mutant (R20291::*tcdR*) and the WT (R20291) and were subjected to RNA-seq analysis. The data observed for selected genes were confirmed by performing qRT-PCR analysis (Figure S 2.2 and Figure S 2.3). RNA-seq analysis of *tcdR* mutant showed that most of the virulence genes are under-expressed and revealed that two major classes

of genes were particularly affected, *i.e.* PaLoc genes and sporulation-associated genes (see in NCBI-GEO accession number GSE85395). However, few genes were upregulated in the *tcdR* mutant. Among those that were over-expressed, we found the *srlR* gene encoding the regulator of Glucitol/sorbitol-specific PTS system (CDR20291_0690 to 0696). PaLoc genes (*tcdA*, *tcdB*, *tcdR*, and *tcdE*) were down-regulated (33-fold, 12-fold, 5-fold and 3-fold respectively) in the *tcdR* mutant, as expected. Auto-regulation of TcdR and its need for toxin gene transcription were already well characterized (7, 8, 16). However, no report was available on the TcdR-mediated transcription of *tcdE* in the . TcdE is a holin-like protein and was found to mediate toxin release from *C. difficile* cells (22, 35). Our data suggest that TcdR is also needed to initiate *tcdE* transcription in *C. difficile*.

2.7.3 Many sporulation associated genes were significantly repressed in the *tcdR* mutant

In addition, many genes in the sporulation pathways were repressed in the *tcdR* mutant when compared to the WT (Table 2.2). Sporulation is a highly complex cellular process regulated by a cascade of events (33, 34, 36). Spo0A is the master regulator of sporulation, and its transcript levels were unchanged in the *tcdR* mutant as observed in both RNA seq and qRT-PCR analysis (Table 2.2and Figure 2.3). However, we saw that transcripts of the specific sporulation sigma factor genes *sigE, sigF, sigG,* and *sigK* were under-expressed in the *tcdR* mutant (Table 2.2and Figure 2.3). Even though the levels of transcription of these genes were moderately under-expressed in the *tcdR* mutant in the RNA-seq analysis when compared to the WT strain (1.5 to 2-fold), we observed through qRT-PCR analyses that their transcription levels in the *tcdR* mutant were significantly reduced along the time growth (Figure 2.3).

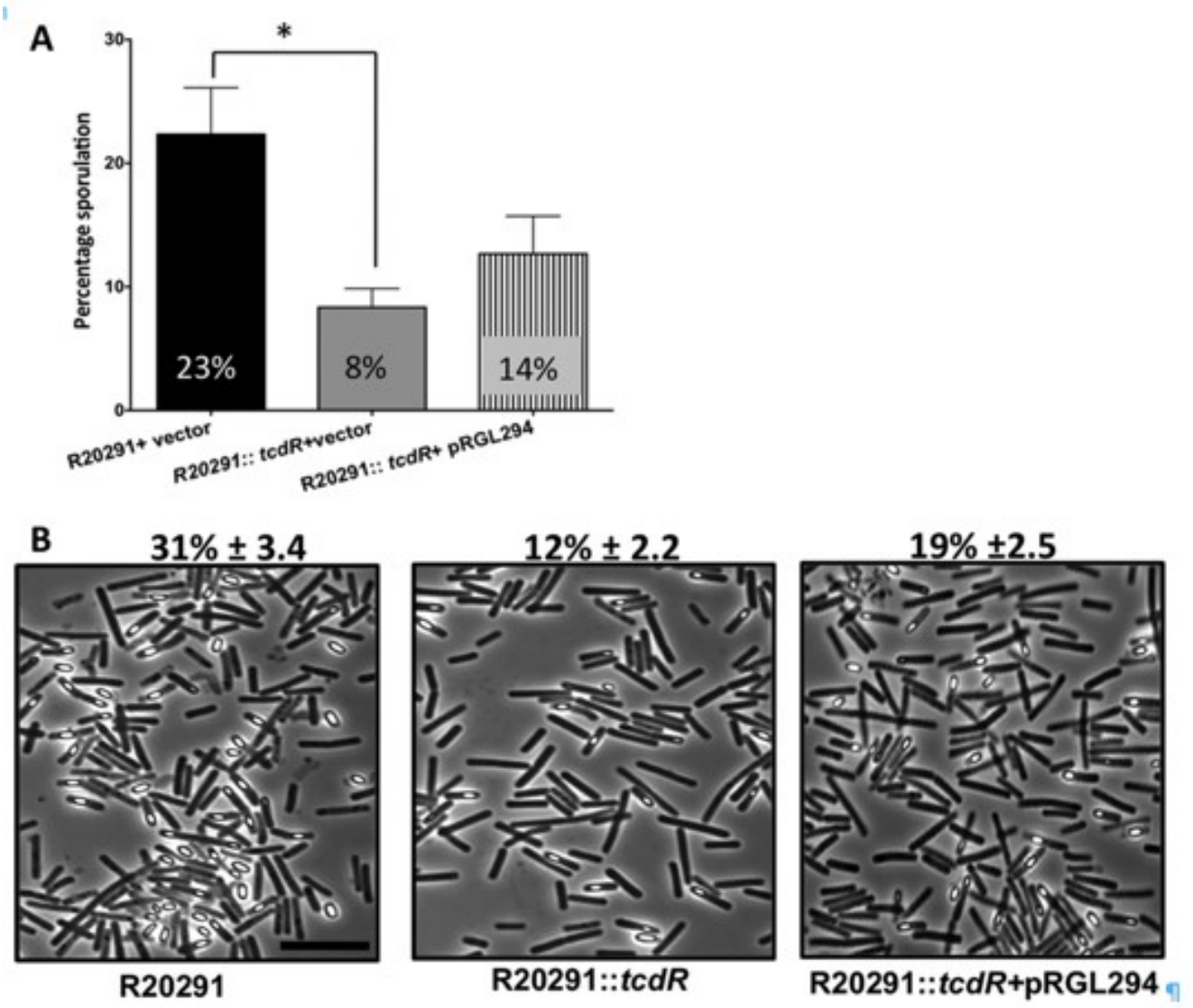


Figure 2.2 Mutation in *tcdR* affects the sporulation efficiency in the R20291 strain.

A. Sporulation frequency (CFU ml^{-1} ethanol-resistant spores) of R20291 + pRPF185, R20291::*tcdR*+pRPF185 and R20291::*tcdR*+pRGL294 (pRPF185 derivative plasmid expressing *tcdR*) strains grown for 24h in 70:30 sporulation medium. The error bars correspond to standard error mean from 3 biological replicates. * $p < 0.05$ by two-tailed Student's t-test. At least three independent experiments were performed. B. Phase contrast microscopy of paraformaldehyde-fixed R20291, R20291::*tcdR*+pRPF185 and R20291::*tcdR* +pRGL294 strains grown for 24h in 70:30 sporulation plate. Percentage sporulation (± standard deviation) was calculated (the number of spores divided by the total number of spores and vegetative cells) from at least three independent experiments. Scale bar = 10µm.

RNA-seq analysis also revealed several sporulation genes controlled by *sigE*, *sigG,* and *sigK* to be significantly affected in the *tcdR* mutant (Table 2.2) (29, 33, 34). SigE is a mother cell-specific sigma factor responsible for the transcription of early sporulation-specific genes, and the SigE-regulated genes identified to be affected in *tcdR* mutant included: *spoIVA* (stage IV sporulation protein A), *spmBA* (spore maturation proteins B & A) and *sigK*, the second mother cell-specific sigma factor. SigG is the forespore-specific factor that controls final stages in sporulation. The SigG-regulated genes found to be repressed in *tcdR* mutant included: *pdaA* (spore specific deacetylase), *sspA* (small acid soluble protein) and *spoVAC/AD* (stage V sporulation proteins). SigG and SigE activities were found previously to be required for the production of heat-resistant spores (34). The *sigK C. difficile* mutant was able to make heat- resistant spores, however at 3-log lower level than the parent strain (34). SigK regulates many genes encoding spore structure proteins that participate to the synthesis of the spore coat and spore exosporium. In fact, we found that many of the SigK-regulated genes such as *cotJBD, cotA, cotB, cotE, bclA3, and bclA2* as well as *sleC* and *cdeC* genes were significantly under-expressed in *tcdR* mutant when compared to the WT strain. The down-regulation of these genes was confirmed by qRT-PCR analysis (Figure S 2.3 and Figure S 2.4).

Table 2.2 Differentially expressed sporulation genes in R20291::*tcdR* mutant.

Genes were considered differentially expressed if the fold change was ≥ 2.0 and their adjusted *p* value is ≤0.05. Expression levels of genes marked with (*) were not statistically significant.

Gene ID	**Gene name if assigned, known/predicted function**	**Fold down regulation in mutant (WT/*tcdR* mutant)**	**Known or predicted Sigma Factor needed for expression**
CDR20291_0124	cell wall endopeptidase	**3.844**	Sig F
CDR20291_2145	hypothetical protein	**5.993**	Sig F
CDR20291_2363	*gpr*, germination protease	**4.008**	Sig F
CDR20291_3400	spore cortex-lytic enzyme	**5.652**	Sig F
CDR20291_3401	*spoIIR*, stage II sporulation protein	**4.228**	Sig F

Gene ID	Gene name if assigned, known/predicted function	Fold down regulation in mutant (WT/*tcdR* mutant)	Known or predicted Sigma Factor needed for expression
CDR20291_2530	*sigG*	**2.14**	Sig F
CDR20291_0125	*spoIIID*, stage III sporulation protein D	**5.323**	Sig E
CDR20291_0714	stage IV sporulation protein	**12.140**	Sig E
CDR20291_1031	*spoIIIAB*, stage III sporulation protein AB	**3.600**	Sig E
CDR20291_1032	*spoIIIAC*, stage III sporulation protein AC	**4.031**	Sig E
CDR20291_1033	*spoIIIAD*, stage III sporulation protein AD	**4.458**	Sig E
CDR20291_1034	*spoIIIAE*, stage III sporulation-related protein	**3.733**	Sig E
CDR20291_2147	*cspBA*, germination-specific protease	**4.346**	Sig E
CDR20291_2513	*spoIVA*, stage IV sporulation protein A	**4.773**	Sig E
CDR20291_3376	*spmB*, spore maturation protein B	**4.333**	Sig E
CDR20291_3377	*spmA*, spore maturation protein A	**5.447**	Sig E
CDR20291_1073	hypothetical protein	**4.563**	Sig E
CDR20291_0702	*spoVAC*, stage V sporulation protein AC	**5.524**	Sig G
CDR20291_0703	*spoVAD*, stage V sporulation protein AD	**5.682**	Sig G
CDR20291_1130	small acid-soluble spore protein	**4.816**	Sig G
CDR20291_1131	*dacF*, D-alanyl-D-alanine carboxypeptidase	**5.891**	Sig G
CDR20291_1529	*sodA*, superoxide dismutase	**5.714**	Sig G
CDR20291_2576	*sspA*, small acid-soluble spore protein A	**4.500**	Sig G
CDR20291_2802	*spoVFB*, dipicolinate synthase subunit B	**3.914**	Sig G
CDR20291_3080	small acid-soluble spore protein	**4.107**	Sig G
CDR20291_3107	*sspB*, small acid-soluble spore protein B	**4.690**	Sig G
CDR20291_0212	spore coat protein	**6.600**	Sig K
CDR20291_0316	spore coat assembly asparagine-rich protein	**6.101**	Sig K
CDR20291_0337	Fragment of putative exosporium glycoprotein	**12.666**	Sig K
CDR20291_0522	*cotJB1*, spore-coat protein	**8.666**	Sig K
CDR20291_0523	*cotJC1*, spore-coat protein	**6.842**	Sig K
CDR20291_2290	*cotJB2*, spore-coat protein	**5.679**	Sig K
CDR20291_2291	*cotJC2*, spore-coat protein	**5.165**	Sig K
CDR20291_2803	*dpaA*, dipicolinate synthase subunit A	**4.291**	Sig K
CDR20291_3090	*bclA2*, exosporium glycoprotein	**6.302**	Sig K
CDR20291_3193	*bclA3*, exosporium glycoprotein	**12.612**	Sig K
CDR20291_3466	cell wall hydrolase	**4.631**	Sig K
CDR20291_0476	*sleC*, spore peptidoglycan hydrolase	**5.502**	Partly by Sig F, SigK
CDR20291_2121	*sinR*	**20.5**	Unknown
CDR20291_2122	*sinR* like DNA binding protein	**27.25**	Unknown
CDR20291_0701	*sigF**	**1.23**	Sig H
CDR20291_2531	*sigE**	**1.56**	Sig H

Gene ID	Gene name if assigned, known/predicted function	Fold down regulation in mutant (WT/*tcdR* mutant)	Known or predicted Sigma Factor needed for expression
CDR20291_1052	*spo0A**	**1.56**	Sig H
CDR20291_1067B	*sigK**	**1.78**	Sig E

2.7.4 CDR20291_2121 and CDR20291_2122 (sin operon) were repressed in *tcdR* mutant

Other than genes involved in sporulation morphology, we also found some regulatory genes potentially involved in sporulation to be affected in the *tcdR* mutant. In fact, the transcript levels of *CDR20291_2121* (coding for a SinR-like protein of *Bacillus subtilis*) and *CDR20291_2122* (coding for a DNA binding protein) genes were nearly 20-fold lower in the *tcdR* mutant than the WT strain (Table 2.2). This result was confirmed by qRT-PCR (Figure 2.3). In *B. subtilis*, SinR is encoded within the *sin* locus carrying both *sinI* and *sinR* genes. In *B. subtilis*,

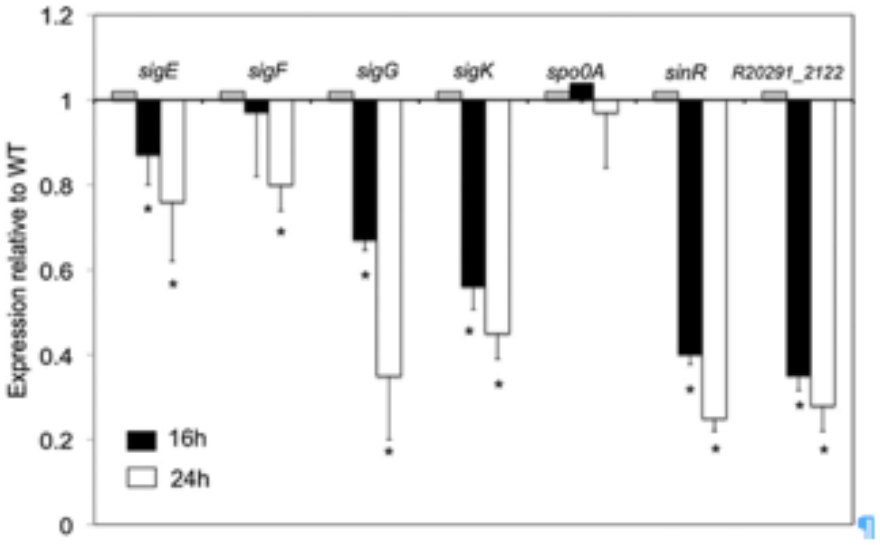


Figure 2.3 Decreased expression of key sporulation genes in the *tcdR* mutant.

qRT-PCR analysis of *sigE, sigF, sigG, sigK, spo0A, sinR*, and *R20291_2122* expression after 16 and 24 h of *C. difficile* growth in 70:30 sporulation medium. Error bars correspond to the standard error mean from at least three biological replicates. * $p < 0.05$ by two-tailed Student's t-test.

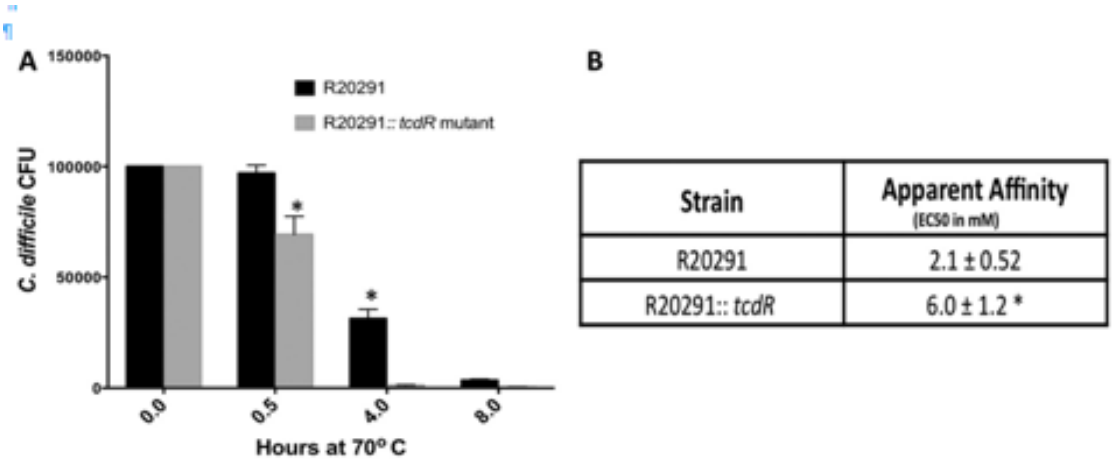
SinR forms tetramers, which repress *spo0A* transcription, although SinI is an inhibitor of SinR (37). If SinR functions similarly in *C. difficile*, a decrease in SinR activity should lead to an increase of sporulation. However, we observed decreased sporulation in the *tcdR* mutant (Figure. 2), suggesting that the products of the *sin* locus must function differently in *C. difficile*.

2.7.5 Spores derived from the *tcdR* mutant have increased heat sensitivity.

To compare the heat sensitivity of spores between the WT and the *tcdR* mutant strains, we incubated purified spores at 70ºC for 0.5 h, 4 h, and 8 h. When we monitored cell viability from the heat-treated spores, we found that spores from the *tcdR* mutant lost most of their viability upon 4 hours of heat treatment, and they were nearly ten-fold more sensitive to heat than the WT spores (Figure 2.4A). This could be due to the decreased expression of both *sigG* and *sigE* in the *tcdR* mutant as observed in our transcriptional analysis, whose activities are known to be involved in the formation of heat resistance of spores (34). In addition, the lower expression of many of the spore structure proteins (including *cdeC*) in the *tcdR* mutant can also explain the heat sensitivity of these spores.

2.7.6 Increased taurocholate was required by *tcdR* mutant spores to germinate

To test if the lower transcription of sporulation-associated genes observed in the *tcdR* mutant (Table 2.2) affects the ability of *C. difficile* spores to germinate, we determined the apparent interaction of spores with taurocholic acid (Figure 2.4B). *C. difficile* spores were suspended in rich medium alone or supplemented with increasing concentrations of the germinant, taurocholate.



Strain	Apparent Affinity (EC50 in mM)
R20291	2.1 ± 0.52
R20291:: *tcdR*	6.0 ± 1.2 *

Figure 2.4 The *tcdR* mutant affects spore germination.

Heat resistance of spores of *C. difficile* strain R20291 and its derivatives *tcdR* mutant was measured by heat-treating aliquots at 70°C for 0.5 hours, 4 hours and 8 hours. Surviving spores were enumerated as described in the Materials and Methods. The data represent the averages of the results of three independent experiments, and error bars represent standard errors of the means. Asterisks (*) indicate statistical difference at $p < 0.05$. B. Apparent affinity of taurocholate for *C. difficile* spores. EC_{50} values were individually calculated from three independent germination experiments and are reported as averages with the standard error of the mean. A Student *t* test was performed and asterisk indicates the calculated *p* value is <0.05.

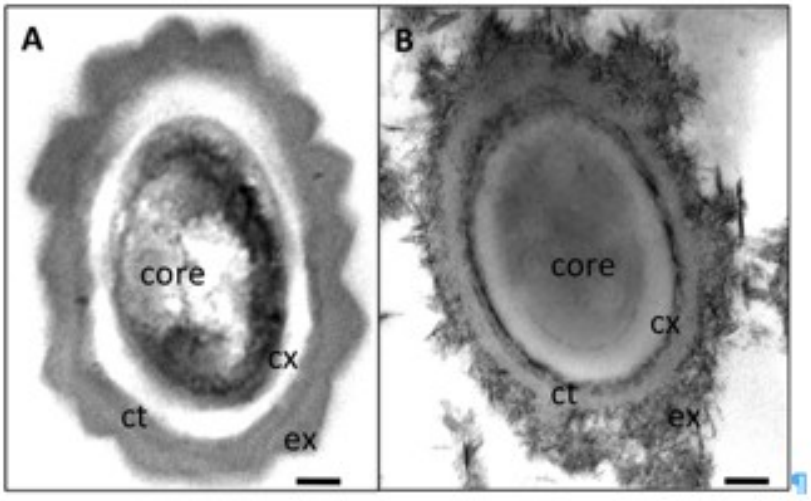


Figure 2.5 Transmission Electron Microscopic analysis of *C. difficile* spores.

Thin section of spores from (A) WT R20291 strain and (B) R20291::*tcdR* mutant. Abbreviations: ex, exosporium; ct, coat; co, core; cx, cortex. Bar size equals to 100nm.

The kinetics of spore germination were followed by measuring the rate of the decrease in OD_{600} as the spores germinate (see Materials and Methods). Though not traditional enzyme kinetics, this assay allows us to understand how spores interact with the taurocholate germinant. *C. difficile* R20291 spores display an EC_{50} value of 2.1 mM (similar to what has been previously reported for other strains) (23, 24, 30). However, the *tcdR* mutant spores display an EC_{50} value of 6.0 mM corresponding to a 3-fold reduction (<0.05 *p* value) in the TA affinity. These results support the overall observation that spore-associated functions were affected when *tcdR* was inactivated in R20291 strain.

2.7.7 Exosporium assembly was affected in the *tcdR* mutant

Spores of the R20291::*tcdR* mutant were compared to those of the WT using electron microscopy to assess any effect on gross spore morphology. Samples were viewed as embedded thin sections and the analysis revealed that *tcdR* mutant spores had a defect in their exosporium assembly (Figure 2.5B and Figure S 2.5). The spore core of the *tcdR* mutant was stained weakly when compared to the WT spores, and darkly stained particulate materials were present over the spore coat and throughout these preparations. Weaker exosporium in the *tcdR* mutant spores could have made them susceptible to structural changes during chemical fixation procedures resulting in these darker particles around the spores. In contrast, most of the R20291 WT spore has an intact exosporium that fully encloses the spore coat (Figure 2.5A) and is devoid of the darker debris observed in the *tcdR* mutant spores. This observation suggests that the *tcdR* mutation in R20291 affects the spore structure with a profound effect on its exosporium assembly.

2.7.8 Effect of *tcdR* on sporulation is strain-specific

Previous studies have reported that mutations in a specific gene can result in different phenotypes in different *C. difficile* strain backgrounds (38, 39). To understand whether the effect of TcdR is strain-dependent, we created a *tcdR* mutant in the 630Δ*erm* strain using the ClosTron system. Toxin production in 630Δ*erm::tcdR* was severely down-regulated as observed in R20291*::tcdR* strain (Figure S 2.6). But unlike in the R20291*::tcdR* strain, the sporulation efficiency of 630Δ*erm::tcdR* was nearly two-fold more than its WT strain (Figure 2.6A and Figure 2.6B). The similar opposing phenotype was already reported for the *spo0A* mutants of R20291 vs. 630Δ*erm*, which also affects the toxin production (38, 39). Though the *spo0A* mutation resulted in increased toxin production in the R20291 strain, it resulted in reduced toxin production in the 630Δ*erm* background. Even though the R20291 and the 630 strains share 3,247 core genes, their genomes are significantly different from one another (40). Whereas there are 47 coding sequences unique in R20291 compared to 630 strain, and 505 coding sequences unique to 630 compared to the R20291 strain (40). Therefore, the difference we observed in these two strains concerning the impact of the *tcdR* mutation on sporulation might be related to the presence or absence of any of these unique genes. Even though we do not know the exact reason for these differences, these observations suggest that the *C. difficile* genome is dynamic and its regulatory networks are fluid in nature.

2.8 Discussion

TcdR-mediated toxin gene regulation is well studied in *C. difficile* (7, 8, 16). The aim of this study was to understand whether TcdR could influence cellular processes other than toxin production. To do so, we created a *tcdR* mutant in the R20291 strain and performed several

phenotypic assays. As expected, the *tcdR* mutant strain produced no or below detectable levels of toxins. Surprisingly, we also observed that the level of spores produced by the mutant is significantly reduced compared to the wild type.

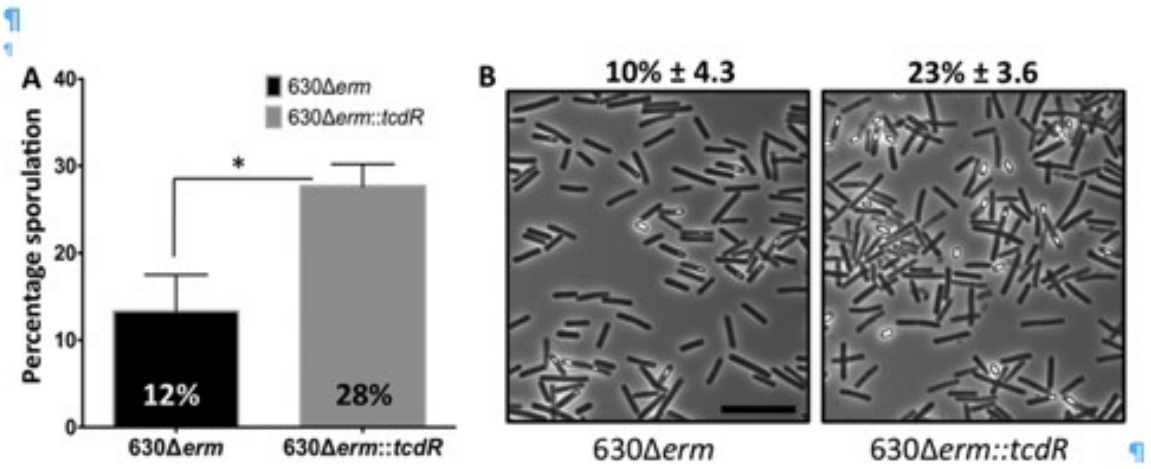


Figure 2.6 Effect of *tcdR* on sporulation is strain-specific.

A. Sporulation frequency (CFU ml^{-1} ethanol-resistant spores) of 630Δ*erm* and 630Δ*erm*::*tcdR* strains grown for 24hr in 70:30 sporulation medium. The error bars correspond to the standard deviation from at least three biological replicates. The asterisk (*) indicates $P< 0.05$ by two-tailed Student's t test. B. Phase contrast microscopy of paraformaldehyde-fixed 630Δ*erm* and 630Δ*erm*::*tcdR* strains grown for 48hr in 70:30 sporulation plate. At least three independent experiments were performed to calculate the percentage sporulation (± standard error mean).

The link between toxin production and sporulation in *C. difficile* has always been suggested but not well studied. For example, in *C. difficile* R20291, a mutation in *spo0A*, the master regulator of sporulation, resulted in changes in toxin production (41). More recently, Edwards *et al.* reported that inactivation of CD3688 (*rstA*) in *C. difficile* strain 630 affects sporulation, toxin production and motility (42). Moreover, it has been shown that the global regulators CodY and CcpA regulate toxin production along with sporulation (10, 12, 43). Thus, if the *tcdR* mutation affects *codY*, *ccpA* or *spo0A* expression, both toxin production and sporulation could be influenced. When we

measured transcript levels of these genes by qRT-PCR, we found no change in their levels in the *tcdR* mutant compared to the WT (Figure S 2.2). However, the genome-wide transcriptome analyses of the *tcdR* mutant confirmed that many sporulation genes were affected.

Nearly 50% of the sporulation genes downregulated in the *tcdR* mutant are known (or predicted) to be under the control of SigE and SigK for their transcription (Table 2.2) (29, 33, 34, 44, 45). Among the down-regulated SigE-dependent genes, we found *sigK*, which could explain the transcriptional decrease of several SigK target genes in the *tcdR* mutant. The RNA-seq analyses of the *tcdR* mutant showed that transcription of *spoIIR* and *spoIIID* genes is reduced (Table 2.2). SpoIIR is essential for the activation of SigE (29, 44) and *spoIIID*, encoding a transcriptional regulator, is involved in the transcription of *sigK* (45). In *C. difficile*, like in *B. subtilis*, SigE is activated by proteolytic cleavage of the SigE precursor form (pro-SigE) (33). In *B. subtilis*, the enzyme SpoIIGA, which is responsible for pro-SigE processing, is co-expressed with *sigE* and is only activated when the mother cell and forespore compartments are formed (46-48). The trigger for SpoIIGA activation is the signal protein SpoIIR that is synthesized in the newly formed forespore and whose presence is communicated to the mother cell (49, 50). In *B. subtilis*, *spoIIR* is regulated by SigF, whereas in *C. difficile*, partial SigE processing is observed in *sigF* mutants, suggesting a lower level expression of *spoIIR* in the absence of SigF (29, 33). If the expression of *spoIIR* in a *sigF* mutant is influenced by TcdR, this could explain the partial processing of SigE in *sigF* mutants. Thus, a reduced abundance of *spoIIR* in the *tcdR* mutant could lead to low levels of activated SpoIIGA and part of pro-SigE would remain unprocessed and inactive. If so, this would also result in a decrease of *spoIIID* as observed in the transcriptome (Table 2.2) and therefore little or no transcription of *sigK* would occur resulting in poor spore maturation.

Most of the genes identified as affected in the *tcdR* mutant code for proteins that are part of the spore proteome (33, 41, 51) and are involved in spore structure and germination. To determine whether *tcdR* mutant spores properties are different from those of the WT spores, we performed heat sensitivity and germination assays using purified spores. TcdR mutant spores were ten times more heat-sensitive than WT spores (Figure 2.4A). Accordingly, transcriptome analysis showed that several exosporium and coat proteins coding genes were under-expressed in the *tcdR* mutant. A recent study on the *C. difficile* exosporium protein BclA3 demonstrated its role in spore heat resistance (52). The authors found that BclA3 is glycosylated by a glycosyltransferase encoded by the adjacent gene (CD3350) within the same operon whose mutation resulted in un-glycosylated BclA3. They showed that spores from this mutant were highly susceptible to heat treatment when compared to the WT spores (52). The same heat susceptibility was observed with the exosporium protein CdeC, which present only in *C. difficile* and is needed for the assembly of exosporium (32). Also, *C. difficile spoVAC* and *dpaAB* mutants also produced heat-sensitive spores (31). The dipicolinate synthase enzyme subunits (SpoVFB and DpaA) are responsible for the production of dipicolinic acid (DPA) that protects spores during heat treatment (53-55). Moreover, proteins encoded in the *spoVA* operon are responsible for transporting DPA from mother cell to fore-spore during spore development (56).

All these results are consistent with the transcriptome analysis of the *tcdR* mutant, which showed decreased expression of *bclA3*, *cdeC*, *spoVAC* and DPA synthase coding operon. This probably results in the production of spores with weaker exosporium that are must be more sensitive to heat treatment than the WT strain (Figure 2.4A). TEM analysis of the *tcdR* mutant

spores confirmed this speculation, where the exosporium was found to be defective and weakly assembled (Figure 2.5).

Germination of bacterial spores is induced when the germinant receptors (GR) sense germinants and subsequently trigger the release of spore core DPA (53). The release of DPA from the spore core leads to the activation of cortex hydrolases that degrade the peptidoglycan (PG) cortex layer, which then allows core hydration. In *C. difficile*, CspC is the bile salt-sensing germinant receptor and is necessary for the release of DPA from spores (23). SleC is the spore cortex lytic enzyme, and its activation depends on CspC (through CspB-mediated cleavage of the pro-domain to generate active SleC) (23, 57-59). A mutation in *sleC* was previously reported to affect germination in *C. difficile* (58, 60). Thus, lower transcription of *sleC* in *tcdR* mutant (Table 2.2), suggested that *tcdR* mutant spores could have inefficient germination. In agreement, we have shown that TA affinity of *C. difficile tcdR* spores is low when compared to the WT spores (Figure 2.4B), indicating that germination is significantly reduced.

Several studies have previously identified sequence "TTTACA" as the -35 region of the TcdR-dependent promoters (7, 8). To test whether some of the downregulated sporulation genes in the *tcdR* mutant can be directly controlled by TcdR, we looked for the presence of this consensus sequence in the promoter regions of these genes (Table 2.2). In fact, we found 8 genes carrying the sequence in the -35 region of the TcdR-dependent promoters. These genes include *bclA2*, *bclA3*, *cotJBD*, *spoVFB*, *cotA*, *cotB*, *cotE*, *dpaA* and *sin*. To test if any of these genes are directly controlled by TcdR, we constructed transcriptional fusions between the promoter of the *bclA2* and *bclA3* genes and the *E. coli* □-glucuronidase (*gusA*) gene that we introduced in a *gus* negative *E. coli* strain expressing or not expressing TcdR as we did previously (19). Compared to the control strains, we did not see any TcdR mediated transcription of *bclA2* or *bclA3* promoters, indicating

that TcdR is not a direct regulator of these genes (Figure S 2.7). However, we cannot exclude for these genes that TcdR may act together with a specific regulator present in the R20291 strain.

Finally, the effect of TcdR on sporulation could be indirect. TcdR is an alternate sigma factor, and its presence or absence could influence the availability of the RNA polymerase core enzyme for other sigma factors in the cell, which in turn can influence the gene expression pattern. Thus, the absence of TcdR in the R20291::*tcdR* strain, increasing the availability of RNA polymerase core enzyme to other sigma factors, could indirectly affect those involved in the sporulation process. On the other hand, there may be common regulators that connect toxin gene regulation with the sporulation pathway in *C. difficile* that could be affected by the *tcdR* mutation. Previous studies have identified several regulators in *C. difficile* regulating toxin production along with sporulation, which strongly suggested that these two pathways were linked (10, 12, 28, 42, 43).

In the past decade, large *C. difficile* outbreaks, with higher relapse rates and increased mortality were reported throughout the world and were attributed to *C. difficile* strains belonging to ribotype 027. The strain R20291 used in this study is a 027 ribotype isolate (61). Genetic and phenotypic features of this ribotype hint that the strains grouped under 027 ribotype are different from other *C. difficile* strains (40). Recently, Lyon *et al.* reported that CdtR, a regulator in the binary toxin locus CdtLoc could regulate toxin production only in 027 ribotype and not in others (62). The authors of this study proposed that CdtR could be regulating toxin production by regulating the TcdR through a yet-to-be identified intermediary regulator in the 027 ribotype. It has been previously proposed that the ability to regulate toxin production in response to various environmental cues with various regulatory responses may be different for 027 ribotype in comparison to other *C. difficile* ribotypes (40). Results from subsequent studies are in agreement

with this proposal. For example, a mutation in the highly-conserved *codY* gene results in different phenotypes from 027 ribotype and other ribotypes. The *codY* mutation results in hyper sporulation phenotype in 027 ribotype (UK1 strain) and produce only a moderate effect on the sporulation in the 012 ribotype (630 strain) (43). It is also worth noting that *sin* locus expression was different in *codY* mutants in these two different *C. difficile* backgrounds (43). Similarly, a mutation in *spo0A* resulted in increased toxin production only in 027 ribotype and not in 012 ribotype (38, 39). In this current study, we observed the positive influence of TcdR on sporulation only in R20291 of 027 ribotype and not in 630 strain of ribotype 012. Even though these studies along with our observation suggest that ribotype 027 has unique gene regulatory networks that differ from other *C. difficile* strains, variations may present strains within 027 ribotype. Detailed study is needed to check whether the gene regulatory networks of toxins synthesis and sporulation pathway are connected in all known 027 ribotype strains. In such case, the ability to synchronize the toxin production, and the sporulation can provide the selective advantage to 027 ribotype isolates to be more successful with increased virulence and high transmission ability. Deciphering the connections between toxin and sporulation regulatory network could lead to the discovery of other novel regulators and pathways that can be targeted for the development of new therapeutics to manage *C. difficile* infections. Any treatment that leads to inhibition of toxin production and spore formation in patients with *C. difficile* infection can potentially lower the severity of the disease in addition to the transmission and recurrence of infection through the spores dissemination.

2.9 Acknowledgement

We thank Nigel Minton, University of Nottingham for sharing the plasmid pMTL007C-E5; and Robert Fagan for providing plasmid pRPF185. We also thank Jose E. Lpoez for technical

assistance throughout the study. RG is supported by 1R15AI122173 from NIAID. Funds from the Johnson Cancer Center- KSU and a pilot project to R.G. from CBID-KU (1P20GM113117-01) also supported this work. J.A.S is supported by Award Number 5R01AI116895 from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies or the National Institutes of Health.

2.10 References

1. CDC. Antibiotic resistance threats in the United States, 2013. 2013; Available from: http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf.

2. McFarland LV. 1998. Epidemiology, risk factors and treatments for antibiotic-associated diarrhea. Digestive diseases 16:292-307.

3. Cohen SH, Gerding DN, Johnson S, Kelly CP, Loo VG, McDonald LC, Pepin J, Wilcox MH. 2010. Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). Infection control and hospital epidemiology 31:431-55.

4. Kyne L, Hamel MB, Polavaram R, Kelly CP. 2002. Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 34:346-53.

5. Goudarzi M, Seyedjavadi SS, Goudarzi H, Mehdizadeh Aghdam E, Nazeri S. 2014. *Clostridium difficile* Infection: Epidemiology, Pathogenesis, Risk Factors, and Therapeutic Options. Scientifica 2014:916826.

6. Popoff MR, Bouvet P. 2009. Clostridial toxins. Future microbiology 4:1021-64.

7. Mani N, Dupuy B. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proceedings of the National Academy of Sciences of the United States of America 98:5844-9.

8. 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. Journal of bacteriology 184:5971-8.

9. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, Pons JL. 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. Dineen SS, McBride SM, Sonenshein AL. 2010. Integration of metabolism and virulence by *Clostridium difficile* CodY. Journal of bacteriology 192:5350-62.

11. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL. 2007. Repression of *Clostridium difficile* toxin gene expression by CodY. Molecular microbiology 66:206-19.

12. Antunes A, Camiade E, Monot M, Courtois E, Barbut F, Sernova NV, Rodionov DA, Martin-Verstraete I, Dupuy B. 2012. Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. Nucleic acids research 40:10701-18.

13. Antunes A, Martin-Verstraete I, Dupuy B. 2011. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Molecular microbiology 79:882-99.

14. Dupuy B, Matamouros S. 2006. Regulation of toxin and bacteriocin synthesis in *Clostridium* species by a new subgroup of RNA polymerase sigma-factors. Research in microbiology 157:201-5.

15. Sirigi Reddy AR, Girinathan BP, Zapotocny R, Govind R. 2013. Identification and characterization of *Clostridium sordellii* toxin gene regulator. Journal of bacteriology 195:4246-54.

16. Dupuy B, Sonenshein AL. 1998. Regulated transcription of *Clostridium difficile* toxin genes. Molecular microbiology 27:107-20.

17. Dupuy B, Mani N, Katayama S, Sonenshein AL. 2005. Transcription activation of a UV-inducible *Clostridium perfringens* bacteriocin gene by a novel sigma factor. Molecular microbiology 55:1196-206.

18. Putnam EE, Nock AM, Lawley TD, Shen A. 2013. SpoIVA and SipL are *Clostridium difficile* spore morphogenetic proteins. Journal of bacteriology 195:1214-25.

19. Govind R, Vediyappan G, Rolfe RD, Dupuy B, Fralick JA. 2009. Bacteriophage-mediated toxin gene regulation in *Clostridium difficile*. Journal of virology 83:12037-45.

20. Teng F, Murray BE, Weinstock GM. 1998. Conjugal transfer of plasmid DNA from *Escherichia coli* to enterococci: a method to make insertion mutations. Plasmid 39:182-6.

21. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP. 2010. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. Journal of microbiological methods 80:49-55.

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

23. 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 pathogens 9:e1003356.

24. Sorg JA, Sonenshein AL. 2010. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. Journal of bacteriology 192:4983-90.

25. Criscuolo A, Brisse S. 2013. AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. Genomics 102:500-6.

26. Monot M, Orgeur M, Camiade E, Brehier C, Dupuy B. 2014. COV2HTML: a visualization and analysis tool of bacterial next generation sequencing (NGS) data for postgenomics life scientists. Omics : a journal of integrative biology 18:184-95.

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

28. 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*. Journal of bacteriology 193:3186-96.

29. 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 genetics 9:e1003756.

30. Bhattacharjee D, Francis MB, Ding X, McAllister KN, Shrestha R, Sorg JA. 2015. Reexamining the Germination Phenotypes of Several *Clostridium difficile* Strains Suggests Another Role for the CspC Germinant Receptor. Journal of bacteriology 198:777-86.

31. Donnelly ML, Fimlaid KA, Shen A. 2016. Characterization of *Clostridium difficile* Spores Lacking Either SpoVAC or Dipicolinic Acid Synthetase. Journal of bacteriology 198:1694-707.

32. Barra-Carrasco J, Olguin-Araneda V, Plaza-Garrido A, Miranda-Cardenas C, Cofre-Araneda G, Pizarro-Guajardo M, Sarker MR, Paredes-Sabja D. 2013. The *Clostridium difficile* exosporium cysteine (CdeC)-rich protein is required for exosporium morphogenesis and coat assembly. Journal of bacteriology 195:3863-75.

33. 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 genetics 9:e1003660.

34. Pereira FC, Saujet L, Tome 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 genetics 9:e1003782.

35. Govind R, Fitzwater L, Nichols R. 2015. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. Journal of bacteriology 197:2600-9.

36. Losick R, Stragier P. 1992. Crisscross regulation of cell-type-specific gene expression during development in *B. subtilis*. Nature 355:601-4.

37. Bai U, Mandic-Mulec I, Smith I. 1993. SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction. Genes & development 7:139-48.

38. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. 2013. Spo0A differentially regulates toxin production in evolutionarily diverse strains of *Clostridium difficile*. PloS one 8:e79666.

39. 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. Infection and immunity 80:2704-11.

40. 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 biology 10:R102.

41. 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.

42. Edwards AN, Tamayo R, McBride SM. 2016. A Novel Regulator Controls *Clostridium difficile* Sporulation, Motility and Toxin Production. Molecular microbiology.

43. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. 2016. CodY-dependent Regulation of Sporulation in *Clostridium difficile*. Journal of bacteriology.

44. Janoir C, Deneve C, Bouttier S, Barbut F, Hoys S, Caleechum L, Chapeton-Montes D, Pereira FC, Henriques AO, Collignon A, Monot M, Dupuy B. 2013. Adaptive strategies and pathogenesis of *Clostridium difficile* from in vivo transcriptomics. Infection and immunity 81:3757-69.

45. Pishdadian K, Fimlaid KA, Shen A. 2015. SpoIIID-mediated regulation of sigmaK function during Clostridium difficile sporulation. Molecular microbiology 95:189-208.

46. Fujita M, Losick R. 2002. An investigation into the compartmentalization of the sporulation transcription factor sigmaE in *Bacillus subtilis*. Molecular microbiology 43:27-38.

47. Jonas RM, Weaver EA, Kenney TJ, Moran CP, Jr., Haldenwang WG. 1988. The *Bacillus subtilis* spoIIG operon encodes both sigma E and a gene necessary for sigma E activation. Journal of bacteriology 170:507-11.

48. LaBell TL, Trempy JE, Haldenwang WG. 1987. Sporulation-specific sigma factor sigma 29 of *Bacillus subtilis* is synthesized from a precursor protein, P31. Proceedings of the National Academy of Sciences of the United States of America 84:1784-8.

49. Hofmeister AE, Londono-Vallejo A, Harry E, Stragier P, Losick R. 1995. Extracellular signal protein triggering the proteolytic activation of a developmental transcription factor in *B. subtilis*. Cell 83:219-26.

50. Karow ML, Glaser P, Piggot PJ. 1995. Identification of a gene, spoIIR, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in *Bacillus subtilis*. Proceedings of the National Academy of Sciences of the United States of America 92:2012-6.

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

52. Strong PC, Fulton KM, Aubry A, Foote S, Twine SM, Logan SM. 2014. Identification and characterization of glycoproteins on the spore surface of *Clostridium difficile*. Journal of bacteriology 196:2627-37.

53. Paredes-Sabja D, Setlow P, Sarker MR. 2011. Germination of spores of Bacillales and Clostridiales species: mechanisms and proteins involved. Trends in microbiology 19:85-94.

54. Orsburn B, Melville SB, Popham DL. 2008. Factors contributing to heat resistance of *Clostridium perfringens* endospores. Applied and environmental microbiology 74:3328-35.

55. Chen NY, Jiang SQ, Klein DA, Paulus H. 1993. Organization and nucleotide sequence of the *Bacillus subtilis* diaminopimelate operon, a cluster of genes encoding the first three enzymes of diaminopimelate synthesis and dipicolinate synthase. The Journal of biological chemistry 268:9448-65.

56. Vepachedu VR, Setlow P. 2007. Role of SpoVA proteins in release of dipicolinic acid during germination of *Bacillus subtilis* spores triggered by dodecylamine or lysozyme. Journal of bacteriology 189:1565-72.

57. Adams CM, Eckenroth BE, Putnam EE, Doublie S, Shen A. 2013. Structural and functional analysis of the CspB protease required for *Clostridium* spore germination. PLoS pathogens 9:e1003165.

58. Burns DA, Heap JT, Minton NP. 2010. SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. Journal of bacteriology 192:657-64.

59. Gutelius D, Hokeness K, Logan SM, Reid CW. 2014. Functional analysis of SleC from *Clostridium difficile*: an essential lytic transglycosylase involved in spore germination. Microbiology 160:209-16.

60. Francis MB, Allen CA, Sorg JA. 2015. Spore Cortex Hydrolysis Precedes Dipicolinic Acid Release during *Clostridium difficile* Spore Germination. Journal of bacteriology 197:2276-83.

61. Elliott B, Dingle KE, Didelot X, Crook DW, Riley TV. 2014. The complexity and diversity of the Pathogenicity Locus in *Clostridium difficile* clade 5. Genome biology and evolution 6:3159-70.

62. Lyon SA, Hutton ML, Rood JI, Cheung JK, Lyras D. 2016. CdtR Regulates TcdA and TcdB Production in *Clostridium difficile*. PLoS pathogens 12:e1005758.

63. Hussain HA, Roberts AP, Mullany P. 2005. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. Journal of medical microbiology 54:137-41.

64. Fagan RP, Fairweather NF. 2011. *Clostridium difficile* has two parallel and essential Sec secretion systems. The Journal of biological chemistry 286:27483-93.

2.11 Supplemental data included in this file:

Supplementary methods: I. Spore preparation; II. Toxin ELISA, III; Complementation of *tcdR* mutant; IV. RNA seq analysis; V. Quantitative reverse transcriptase PCR.
Table S1. Oligonucleotides used in the study

Figure S1. Construction and characterization of *tcdR* mutant in *C. difficile*.
Figure S2. Expression analysis of known toxin gene regulators in *tcdR* mutant by q-RT PCR.
Figure S3. Expression analysis of selected sporulation genes during growth in 70:30 medium.
Figure S4. Expression analysis of selected sporulation genes during growth in TY medium.
Figure S5. TEM analysis of spores from the R20291 and R20291::*tcdR* strains
Figure S6. Toxin ELISA for 630Δ*erm* and 630Δ*erm*::*tcdR* strains.
Figure S7. Beta-Glucuronidase activity of *bclA2* promoter-*gusA* and *bclA3* promoter-*gusA* fusions in the presence or absence of TcdR

2.11.1 Supplementary methods

2.11.1.1 Spore preparation from *C. difficile* cultures

C. difficile strains were streaked onto 20 plates of BHIS agar medium. After 4 days, the growth from each plate was scraped into 1 ml of sterile water and incubated overnight at 4 °C. The following day, the growth suspension was washed five times with sterile water. In between each wash step, the layer of white cell debris was removed. After washing, the spores/cell debris suspension in water was layered on top of a 60% (w / v) sucrose solution. The gradient was centrifuged in a swinging-bucket rotor at 3,200 X g for 20 minutes. During the centrifugation, the dense spores travel through the sucrose and form a pellet on the bottom while the cell debris is caught in the upper layer. After centrifugation, the solution was removed and the pellet containing the spores was washed five times with sterile water to remove any sucrose. The spores were then suspended in sterile water up to 1 ml. Purified spores were examined under phase contrast microscopy and determined to be >99.9% pure and phase bright.

2.11.1.2 Toxin ELISA

Bacterial cells were harvested and were resuspended in 10mM Tris buffer (pH 8.0) containing a protease inhibitor cocktail (Roche, Mannheim, Germany) and sonicated to release the

cytosolic contents. A Bio-Rad protein assay reagent was used to determine the total protein concentration and equal amounts of cytosolic protein (20 µg) were assayed for their relative toxin level using *C. difficile* Tox A/B II ELISA kit from Tech Lab (VA, USA).

2.11.1.3 Complementation of the R20291::tcdR mutant

R20291 chromosomal DNA was used to PCR amplify the *tcdR* ORF along with its upstream DNA using primers ORG403 and ORG209 (Table S 2.1), which carried restriction sites *KpnI* and *BamHI* respectively. The resulting PCR product and the vector pRPF185 were digested with *KpnI*, *BamHI* and ligated to yield pRGL294 (Table 2.1). The empty vector or the pRGL294 was introduced into *C. difficile* R20291::*tcdR* by conjugation. Transconjugants were grown overnight in TY medium supplemented with thiamphenicol (15 µg/ml). One hundred microliters of overnight culture was used to initiate fresh 10 ml cultures. The 10 h old culture was used to measure the toxins.

2.11.1.4 RNA seq analysis

Total RNA was isolated from three biological replicates of each strain belonging to late stationary phase (18 hours after inoculation) and quality was checked using Agilent 2100 Bioanalyzer. The rRNA content in the selected samples was depleted using Epicenter Bacterial Ribo-Zero kit. The depleted rRNA fraction was used to construct strand specific single end cDNA libraries according to manufacturers' instructions (using Truseq Small Stranded Total RNA sample prep kit, Illumina). Illumina HiSeq2000 sequencer (multiplexing 3 samples per lane) was used to sequence libraries. Sequences were cleaned (AlienTrimmer [26]) of adapter sequences, low quality sequences and only sequences with a minimum of 30 nucleotide in length was considered for

further analysis. Out of three biological replicates, based on the quality of the data, only two were qualified for statistical analysis. Cleaned genes were aligned to reference genome (FN545816.1) using Bowtie (version 1.0.1) [25]. DESeq2 version 1.8.3 was used to perform normalization and differential analysis. Genes were considered differentially expressed if the fold change was ≥ 2.0 and their adjusted *p* value is ≤0.05. The mapped reads were formatted and visualized usingCOV2HTML [27] (https://mmonot.eu/COV2HTML/visualisation.php?str_id=-40). RNA-Seq data have been deposited in NCBI-GEO with accession n° GSE85395.

2.11.1.5 Quantitative reverse transcription PCR (qRT-PCR)

C. difficile R20291 and R20291*::tcdR* cultures were grown in TY or in 70:30 medium and cells were harvested at 16 h and 24 h by centrifugation at 4 °C for 2 min. Total RNA was extracted from the harvested cells following the protocol described previously [9, 25] and treated with DNase (Turbo; Ambion) for 30 min at 37 °C. 30 µL of final reaction volume comprising of 5 µg of template RNA, 4 µL of deoxynucleoside triphosphates (dNTP; 10mM each), 1 µg of hexamer oligonucleotide primer (5 µg/µL pdN_6; Roche), and 6 µL of reverse transcription (RT) buffer was heated at 80 °C for 5 min and cDNA was synthesized at 42 °C for 2 hours using avian myeloblastosis virus (AMV) reverse transcriptase (Promega). Final 20 µL reaction volume containing 10 ng or 10 pg (for 16S rRNA) of cDNA, 400 nM gene-specific primers, and 12.75 µL of SYBR PCR master mix (BioRad) was used to perform Real-time quantitative PCR using iQPCR real-time PCR instrument (BioRad). Amplification and detection was performed as described previously [9]. Quantity of cDNA of a gene in each sample was normalized to the quantity of *C. difficile* 16S rRNA gene and the ratio of normalized target concentrations (threshold cycle [$2^{-\Delta\Delta Ct}$] method) [9, 28] gives the relative change in gene expression.

2.11.2 Supplementary tables

Table S 2.1 Oligonucleotides used in this study

OLIGO Name	Sequence	Description
EBS universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Intron specific
tcdR-IBS	AAAAAAGCTTATAATTATCCTTAGAAACCGATTTAGTGCGCCCAGATAGGGTG	*tcdR* intron retarget
tcdR-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGATTTAATTAACTTACCTTTCTTTGT	*tcdR* intron retarget
tcdR-EBS2	TGAACGCAAGTTTCTAATTTCGATTGTTTCTCGATAGAGGAAAGTGTCT	*tcdR* intron retarget
ORG403	GGTACCCATTTGATTAAAAATAACAAAATATTAAATAATTC	*tcdR* upstream- forward with KpnI
ORG81	ATGCAAAAGTCTTTTTATGAATTAATTGTT	*tcdR coding-Forward*
ORG82	GTTAAAATAATTTTCATAGTCTTTTTTTA	*tcdR* coding-Reverse
ORG208	GAGCTCATATAAGAGAGGATGATTTTATGC	*tcdR coding-Forward*
ORG209	GGATCCTTAATGATGATGATGATGATGCAAGTTAAAATAATTTTC	*tcdR* coding-Reverse with BamHI
RG-RT37	TGACTTTACACTTTCATCTGTTTCTAGC	*sigE* qPCR-Forward
RG-RT38	GGGCAAATATACTTCCTCCTCCAT	*sigE* qPCR-Reverse
RG-RT41	CGCTCCTAACTAGACCTAAATTGC	*sigF* qPCR-Forward
RG-RT42	GGAAGTAACTGTTGCCAGAGAAGA	*sigF* qPCR -Reverse
RG-RT49	CATATGTTGCTAATCGAGTTCCTTTAT	*sigK* qPCR-Forward
RG-RT50	TCAACGGAAGATCAGGATGATTTA	*sigK* qPCR-Reverse
RG-RT45	CAAACTGTTGTCTGGCTTCTTC	*sigG* qPCR-Forward
RG-RT46	GTGGTGTTAATACATCAGAACTTCC	*sigG* qPCR-Reverse
RG-RT33	CATGAAATAGGAGTACCAGCTCA	*spo0A* qPCR-Forward
RG-RT34	CTCCATGCAACCTCTATTGC	*spo0A* qPCR-Reverse
RG-RT23	GAGGAGAGTGGAATTCCTAGTGTAG	*16srRNA* qPCR-Forward
RG-RT24	GGACTACCAGGGTATCTAATCCTGT	*16srRNA* qPCR-Reverse
RG-RT25	AGGCAGGTTTACATCCAACATA	*sinR* qPCR
RG-RT26	AGTGGTATGTCTAAAGCAGTAGC	*sinR* qPCR
RG-RT27	AAAGACTTAAAGAAGAACGGAAAA	CDR20291_2122-Forward
RG-RT28	TTGGATTCTTTTTACCACTTTCG	CDR20291_2122-Reverse
sleB-RT	GATATTGTAGAGAACCCCTAATCC	QRT-Forward
sleB-RT	GCAAATCCTAAAGCTAAAAATAC	QRT-Reverse
Gpr-RT	GGTGTTACTATTAAGTTCTTGTCAT	QRT-Forward
Gpr-RT	CTGGTGGAGGTGTTGGCAATACTAG	QRT-Reverse
sspA-RT	CTATCTGTTGCTTTTTCCAGCC	QRT-Forward
sspA-RT	GTATGAGTAATTATCAACAAGTTG	QRT-Reverse
spoVAC-RT	GTAGACCAAATAAGCCCAAAACC	QRT-Forward
spoVAC-RT	CAGAACTAGCACCTAGTTTATC	QRT-Reverse
spoVAD-RT	GTGGCGATTTAATAAATCAAATAG	QRT-Forward
spoVAD-RT	CACTTCCTGCTCCTGTAACTGTCC	QRT-Reverse
bclA3-RT	CTGCTGCGTTTGTAAGGTCTATTAC	QRT-Forward
bclA3-RT	GAGCAACAGGTCCAACAGGAGCAAC	QRT-Reverse

bclA2-RT	GAGTTACTCCATTAAAGTTAG	QRT-Forward
bclA2-RT	GGAGTAGCAGGAGCGATAGGACC	QRT-Reverse
cdeC-RT	GATGAAATAAATTCAGAAGACATGA	QRT-Forward
cdeC-RT	GGCACTGCATTTGATACAGAGAAG	QRT-Reverse
sleC-RT	CTGTTCCATAGATACCATCTTC	QRT-Forward
sleC-RT	GGGCAGTAAAGACTTAGGTGACC	QRT-Reverse
cotCB-RT	GGTACAGAGGAAATGGCTCATGTTG	QRT-Forward
cotCB-RT	CTTGTAGTAAAGTTTACTCCATTAG	QRT-Reverse
cotE-RT	GAATATTGATAAAGCATCATCATATG	QRT-Forward
cotE-RT	GCCATAAGAGATGTTATAGGGGATG	QRT-Reverse
cotB-RT	GATTTTATCTTACACTGTTCTATTCC	QRT-Forward
cotB-RT	GGACCATATTATGATGGAACATGCTC	QRT-Reverse
cotA-RT	CTTACCTAGAACTTCAACACCAGTTA	QRT-Forward
cotA-RT	CATTGTGTAATCTTAAAGCTGTTGC	QRT-Reverse
pdaA-RT	CATCTAATATCTCAGTATTTGTTG	QRT-Forward
pdaA-RT	GCGAACAATCTTTAAAATATACACAA	QRT-Reverse
pbclA2-F	GGTACCAGATAAGCAATTATATAATTTTGTGGATGCCTTA	*bclA2* promoter-Forward
pbclA2-R	TCTAGATAATTAATCCTCCTTTTTTAAAGTTAGAGTATTAC	*bclA2* promoter-Reverse
pbclA3-F	GGTACCTAATGAATAGGAATGAATAGGAATGAATAAAGTA	*bclA3* promoter-Forward
pbclA3-R	TCTAGAACAACTCCTTTGTCTTCTATATATTGCAGACATAA	*bclA3* promoter-Reverse

2.11.3 Supplementary figures

Figure S 2.1 Construction and characterization of *tcdR* mutant in *C. difficile*.

A. Schematic representation of insertional inactivation of tcdR by group II intron. B. The intron insertion in tcdR coding region was verified by PCR using the intron-specific primer EBS along with gene-specific primers ORG81 or ORG82 in the parent (R20291), tcdR mutant (R20291::tcdR) and in tcdR complemented strain (R20291:: tcdR+pRG294). The same strategy was followed to verify tcdR mutation in the 630Δerm strain. C. Southern blot analysis of genomic DNA from the WT and tcdR mutant strains with a tcdR specific probe. The shift in the hybridization band indicates the integration of intron within tcdR coding region.

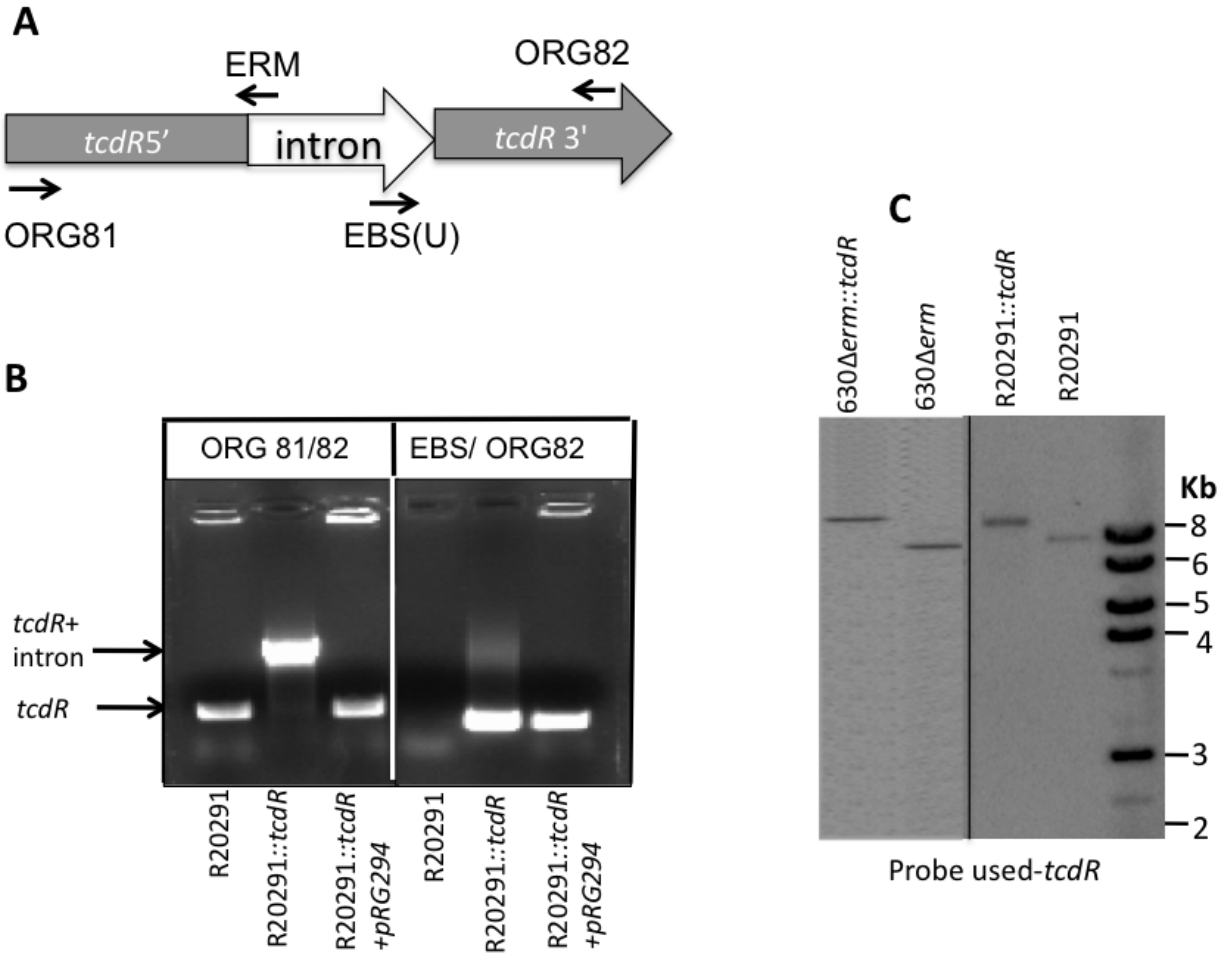


Figure S 2.2 Expression analysis of toxin genes and known toxin gene regulators in *tcdR* mutant.

RNA was prepared from R202091 and R20291*:: tcdR* strains that grown in TY medium for 16h. qPCR analysis was performed for selected regulators coding genes. Statistical analysis by *t*-test and the error bars indicate the standard error of the mean (s.e.m.). ** *p* value < 0.01

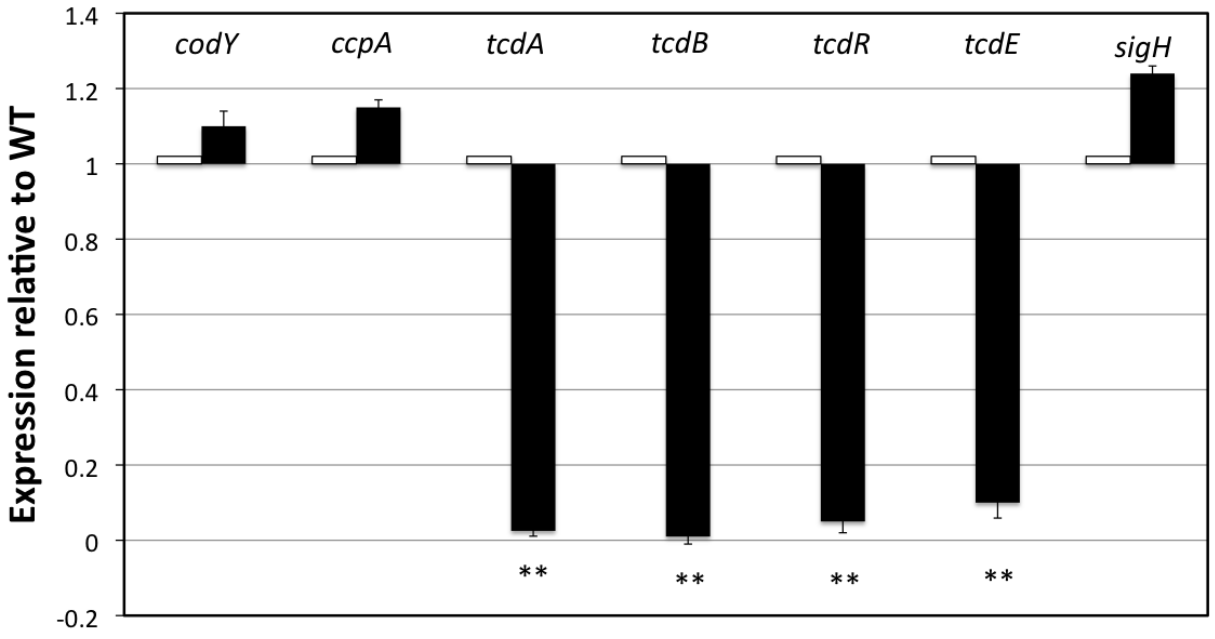


Figure S 2.3 Expression analysis of selected sporulation genes during growth in 70:30 medium.

RNA was prepared from R202091 and R20291*:: tcdR* strains that grown in 70:30 medium for 24h. qRT-PCR analysis was performed for selected sporulation genes. Statistical analysis by *t*-test and the error bars indicate the standard error of the mean (s.e.m.). * *p* value < 0.05

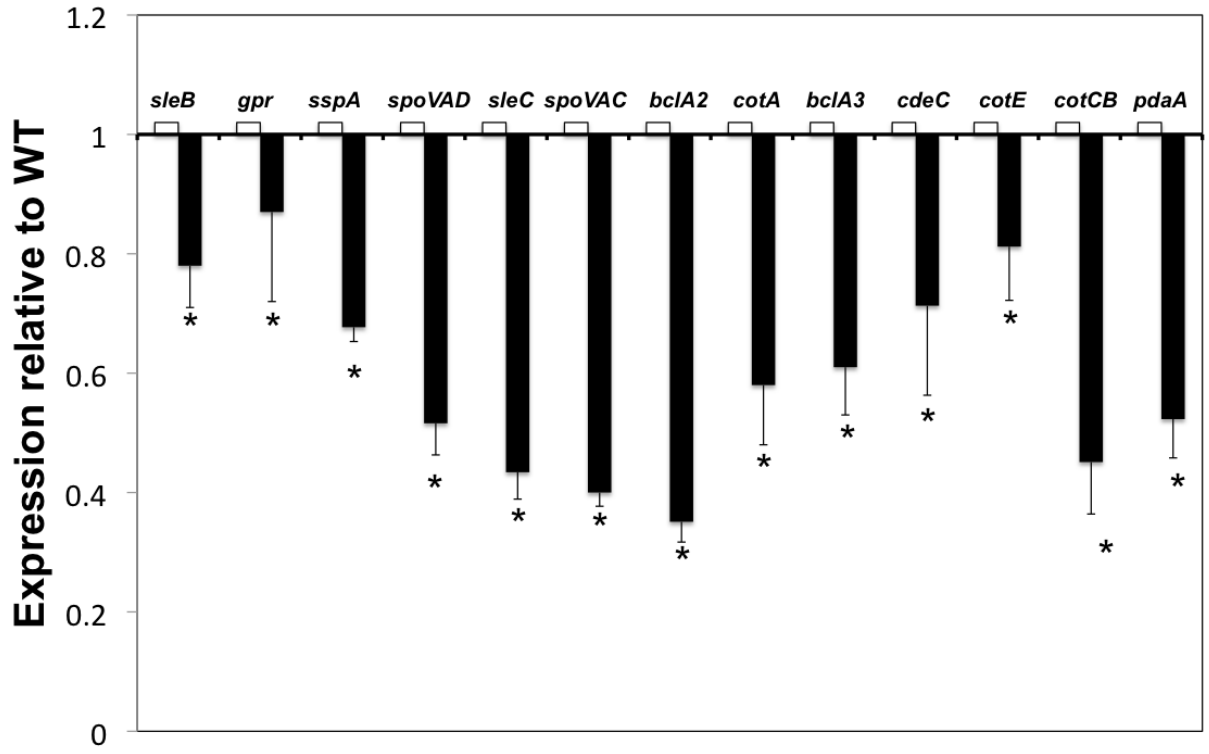


Figure S 2.4 Expression analysis of selected sporulation genes during growth in TY medium.

RNA was prepared from R202091 and R20291*:: tcdR* strains that grown in TY medium for 24h. qRT-PCR analysis was performed for selected sporulation genes. Statistical analysis by *t*-test and the error bars indicate the standard error of the mean (s.e.m.). * *p* value < 0.05

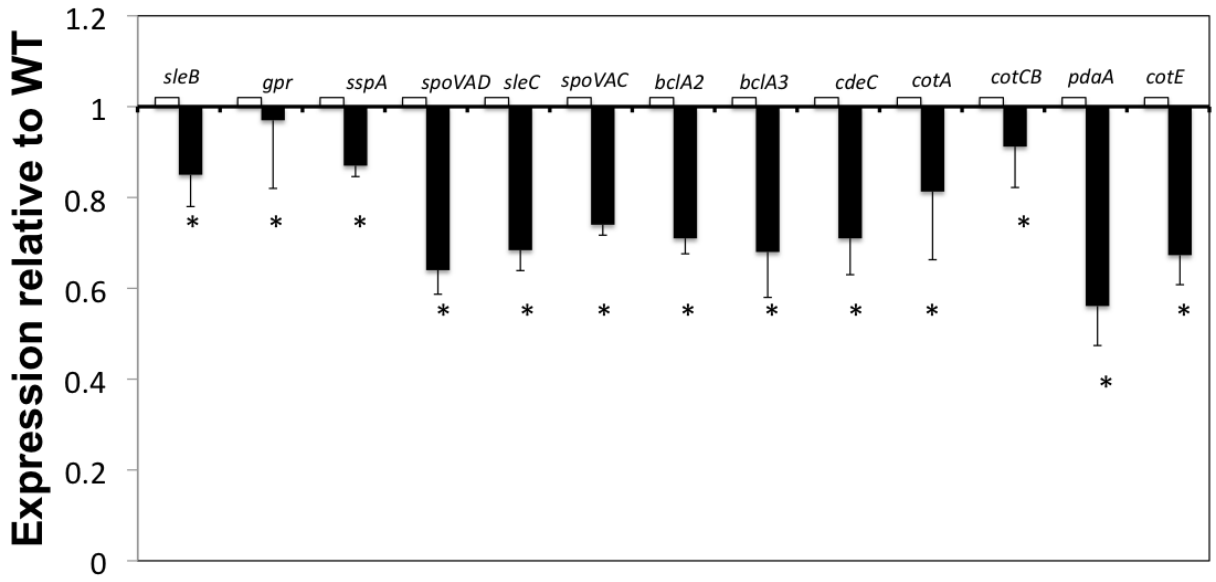


Figure S 2.5 TEM analysis of spores from the R20291 and R20291::tcdR strains.

The *tcdR* mutant spores (n=60) were scored 100% for the presence of ruffled defective exosporium (marked with black arrow) and 98% for the presence of weakly stained core (marked as open triangle). Bar, 100nm.

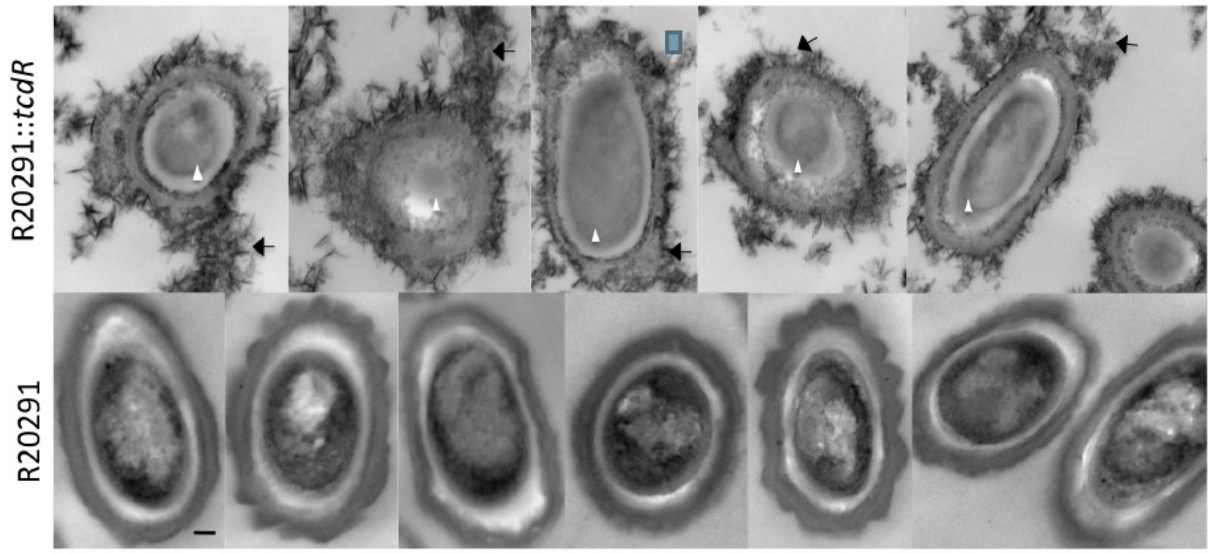


Figure S 2.6 Toxin ELISA for 630Δerm and 630Δerm:: tcdR strains.

TcdA and TcdB expression levels in cytosolic fractions of 10h grown *C. difficile* strains in TY medium were quantified using ELISA. Shown is a representative experiment of three independent assays. Error bars correspond to the standard error mean from at three biological replicates. The asterisks (***) in panel B indicates statistical difference at $p < 0.001$, estimated by student *t* test.

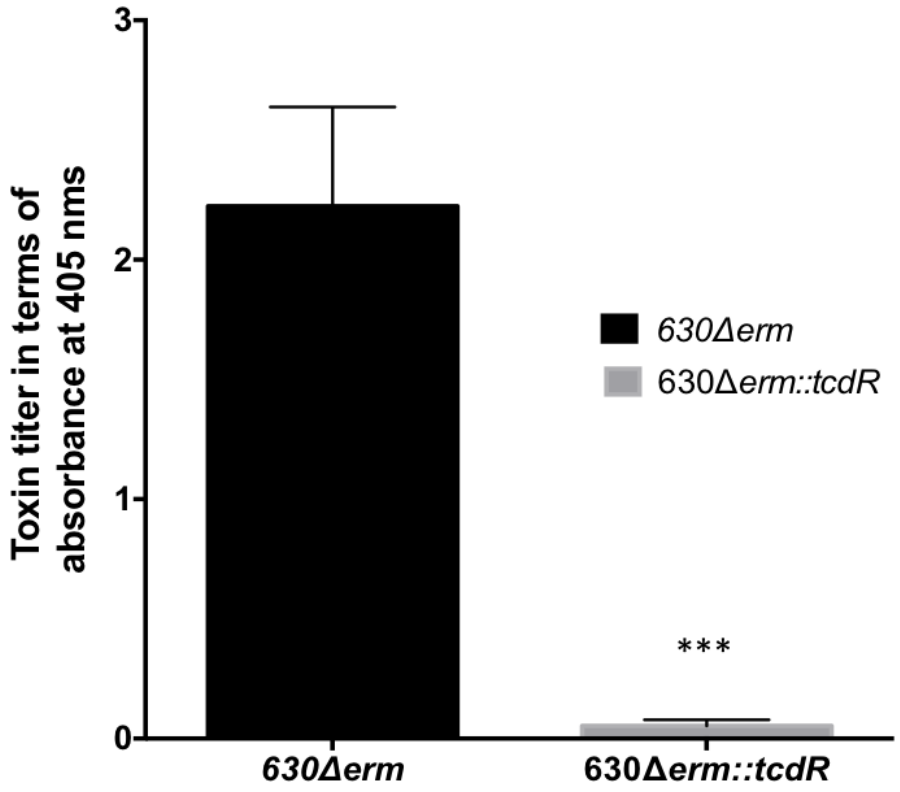
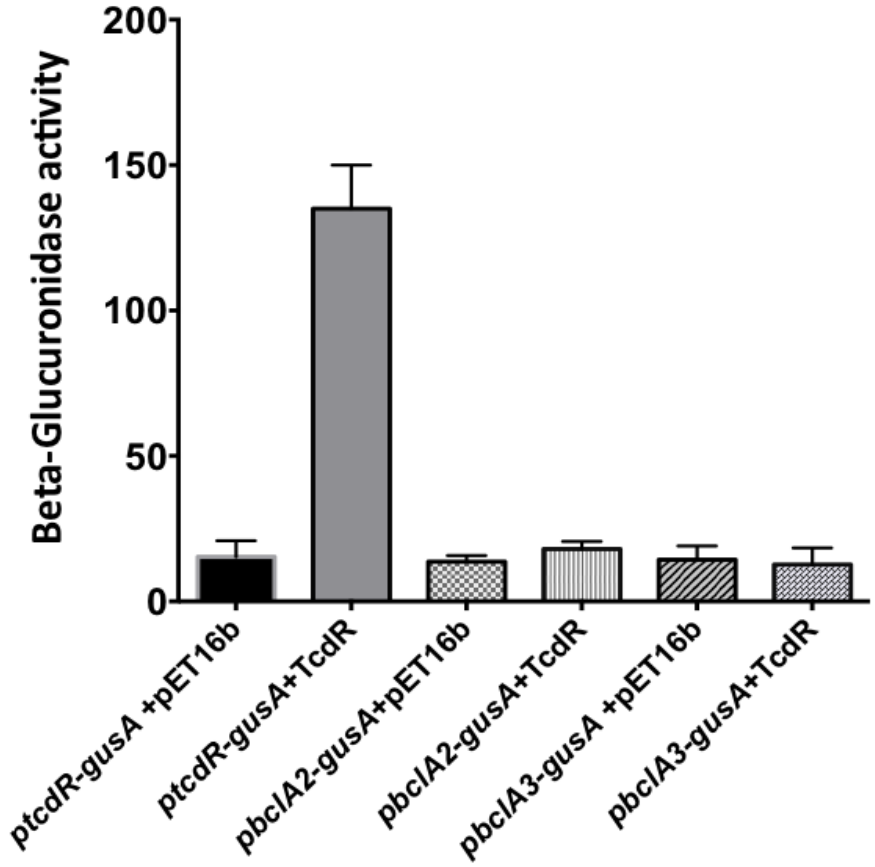


Figure S 2.7 Beta-Glucuronidase activity of bclA2 promoter-gusA and bclA3 promoter-gusA fusions in the presence or absence of TcdR.

E. coli strains carrying promoter fusion plasmids (*pbclA2-gusA* and *pbclA3-gusA*) along with a TcdR-expressing plasmid (pRGL312) or the vector (pET16b) were grown for 6 hours and the TcdR expression was induced with 1mM IPTG for four hours at 37°C. Bacterial cultures were harvested and were assayed for beta-glucuronidase activity (in miller units) as described previously [18]. The values represent the means of three independent experiments. Statistical analysis by *t*-test and the error bars indicates standard error of mean.



Chapter 3 - Pleiotropic roles of *Clostridium difficile* sin locus

3.1 Publication arising from this chapter

The key findings from this chapter has resulted in the following publication (see Appendix B)

Girinathan, B.P., Junjun Ou, Bruno Dupuy, Revathi Govind Pleiotropic roles of *Clostridium difficile* sin locus. PLoS Pathog, 2018. **14**(3): p. e1006940.

3.2 Abstract

Clostridium difficile is the primary cause of nosocomial diarrhea and pseudomembranous colitis. It produces dormant spores, which serve as an infectious vehicle responsible for transmission of the disease and persistence of the organism in the environment. In *Bacillus subtilis*, the *sin* locus coding SinR (113 aa) and SinI (57 aa) is responsible for sporulation inhibition. In *B. subtilis*, SinR mainly acts as a repressor of its target genes to control sporulation, biofilm formation, and autolysis. SinI is an inhibitor of SinR, so their interaction determines whether SinR can inhibits target gene expression. The *C. difficile* genome carries two *sinR* homologs in the operon that we named *sinR* and *sinR'*, coding for SinR (112 aa) and SinR' (105 aa), respectively. In this study, we constructed and characterized *sin* locus mutants in two different *C. difficile* strains R20291 and

JIR8094, to decipher the locus's role in *C. difficile* physiology. Transcriptome analysis of the *sinRR'* mutants revealed their pleiotropic roles in controlling several pathways including sporulation, toxin production, and motility in *C. difficile*. Through various genetic and biochemical experiments, we have shown that SinR can regulate transcription of key regulators in these pathways, which includes *sigD*, *spo0A*, and *codY*. We have found that SinR' acts as an antagonist to SinR by blocking its repressor activity. Using a hamster model, we have also demonstrated that the *sin* locus is needed for successful *C. difficile* infection. This study reveals the *sin* locus as a central link that connects the gene regulatory networks of sporulation, toxin production, and motility; three key pathways that are important for *C. difficile* pathogenesis.

3.3 Introduction

Clostridium difficile, a major nosocomial pathogen, is the causative agent of antibiotic-associated diarrhea and pseudomembranous colitis [1, 2]. Every year, nearly half a million cases of *C. difficile* infections (CDI) occur in the United States and result in approximately 14,000 deaths [3]. *C. difficile* toxins damage the colonic epithelium, which results in moderate to severe diarrhea [4]. Recent studies have shown that these toxins are essential for *C. difficile* pathogenesis [4–7]. Due to the strictly anaerobic nature of the vegetative cell, *C. difficile* survives outside the host in the form of dormant spores, which are highly resilient and resistant to most disinfectants. Thus, *C. difficile* spores are critical for its host to host transmission and persistence in the hospital environment [8].

C. difficile Toxins A and B are encoded by the *tcdA* and *tcdB* genes respectively, and their expression is dependent on TcdR, an alternative RNA polymerase sigma factor [9–11].

Environmental stresses, such as alteration of the redox potential, high temperature, or limitation of nutrients like glucose, and biotin, modulate toxin production by influencing the expression of *tcdR* [9–12]. Similar to toxin production, the sporulation pathway in *C. difficile* is also known to be influenced by nutrient availability and uptake [13, 14]. The regulators involved in con- trolling toxin synthesis in response to nutrients are the global regulatory proteins CcpA and CodY [14–18]. Among them, CcpA mediates glucose-dependent toxin gene repression [15, 16], and CodY blocks the transcription of toxin genes during the exponential growth phase of the bacterial culture [17, 18]. Other than affecting toxin production, mutations in *codY* and *ccpA* were also found to affect sporulation [13, 16]. Other genes that are known to influence both toxin production and sporulation include *spo0A*, *sigH*, and *rstA* [19–22]. New evidence suggests that the toxin, motility, and sporulation regulatory networks are linked together in *C. difficile* [19, 23, 24]. The sigma factor SigD needed for transcription of the flagellar operon was identified to regulate *tcdR* transcription to influence toxin production [25, 26] positively. Mutations in *spo0A*, *rstA*, and *sigH* also influenced motility along with toxin production and sporulation [19–22]. This study identified that mutation of the *sin* locus in *C. difficile* could affect toxin production and sporulation along with motility and thus reports a new regulatory element of this network.

In *Bacillus subtilis*, the *sin* (sporulation inhibitor) locus codes for two proteins SinR and SinI and regulates several genes involved in sporulation, motility, competency, proteolysis, and biofilm formation [27–31]. In this study, we have created *C. difficile sin* locus mutants in two different strains. Using RNA-Seq analysis, we compared the transcriptome of the mutants with respective parent strains to identify and assess the transcriptional regulation of *sin* locus coded regulators. Follow up phenotypic analyses and complementation experiments showed that the Sin

regulators in *C. difficile* are also pleiotropic as in *B*. *subtilis*. Here, their regulatory roles in toxin production, sporulation, and motility were further investigated and discussed.

3.4 Aim of the work described in this chapter

In this chapter, I have used genetic, biochemical and computational tools to characterize the pleiotropic role of sin locus genes on *C. difficile* pathogenesis both in vitro and in vivo

3.5 Materials and methods

3.5.1 Ethics statement

All animal procedures were performed with prior approval from the KSU Institutional Animal Care and Use Committee (protocol #3657). Animals showing signs of disease were euthanize by CO2 asphyxia followed by thoracotomy as a secondary means of death, in accordance with Panel on Euthanasia of the American Veterinary Medical Association. Kansas State University is accredited by AAALAC International (Unit #000667) and files an Assurance Statement with the NIH Office of Laboratory Animal Welfare (OLAW). KSU Animal Welfare Assurance Number is D16-00369 (A3609-01), and USDA Certificate Number is 48-R-0001. Kansas State University utilizes the United States Government Principles for the utilization and care of vertebrate animals used in testing, research and training guidelines for appropriate animal use in a research and teaching setting.

3.5.2 Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study are listed in (Table S 3.1) and cloning strategies used are listed in S1 Text. *Clostridium difficile* strains were grown anaerobically (10% H_2, 10% $C0_2$ and 80% N_2) in TY (Tryptose and Yeast extract) agar or broth as described previously [60, 66]. Erythromycin (Erm; 2.5 μg ml^{-1}), Lincomycin (Linc 20ug/ml), Cefoxitin (Cef; 25 μg/ml), thiamphenicol (Thio; 15 μg ml^{-1}) were added to culture medium whenever necessary. Sporulation was induced in respective *C. difficile* strains by growing them in 70:30 sporulation medium (63 g Bacto-Peptone, 3.5 g Protease-Peptone, 11.1 g BHI, 1.5 g Yeast-Extract, 1.06 g Tris base, 0.7 g NH_4SO_4, 15 g agar per liter) [67]. *Escherichia coli* strain S17-1 [68] was used for conjugation and cultured aerobically in Luria-Bertani (LB) broth and supplemented with chloramphenicol (25μg ml^{-1}) or ampicillin (100μg ml^{-1}) as indicated.

3.5.3 Construction of *C. difficile* mutant strains

ClosTron gene knockout system [69] was used to construct *sinRR'* and *sinR'* mutants. For *sinRR'* disruption, the group II intron insertion site between nucleotides 141 and 142 in *sinR* gene in the antisense orientation was selected using the Perutka algorithm, a Web-based design tool available at http://www.Clostron.com. For *sinR'* mutant construction, the group II intron insertion site between nucleotides 129 and 130 in the sense direction was selected. The designed retargeted intron was cloned into pMTL007-CE5 as described previously [59, 70]. The resulting plasmids pMTL007-CE5::Cdi-*sin*R-141s or pMTL007-CE5::Cdi-*sin*R'-129s was transferred into *C. difficile* cells by conjugation as described earlier [59, 70]. The potential Ll. ltrB insertions within the target genes in the *C. difficile* chromosome was conferred by the selection of erythromycin or lincomycin resistant transconjugants in 5 μg ml^{-1}erythromycin or 20 μg ml^{-1} lincomycin plates.

PCR using gene-specific primers (Table S 3.2) in combination with the EBS-U universal and ERM primers was performed to identify putative *C. difficile* mutants.

3.5.4 General DNA techniques

DNeasy Blood and Tissue Kit (Qiagen) was used to extract chromosomal DNA from the *C. dif- ficile* cultures. Primers used throughout the study are listed (Table S 3.2 and Table S 3.3). Gene- clean Kit (mpbio) was used to gel extract the PCR products, and QIAprep Spin Miniprep Kit (Qiagen) was used to extract plasmid DNA. Standard procedures were used to perform routine cloning.

3.5.5 Sporulation efficiency assays

Sporulation assays were performed in 70:30 sporulation medium as described previously [60]. *difficile* strains were grown on 70:30 sporulation agar. After 30 h of growth, cells were scraped from the plates and suspended in 70:30 sporulation liquid medium to an OD_{600} of 1.0. Cells were immediately serially diluted and plated onto TY agar with 0.1% taurocholate to enumerate viable vegetative cells and spores. To determine the number of spores present, 500µl of the samples from each culture were mixed 1:1 with 95% ethanol and incubated for 1hour to kill all the vegetative cells. The ethanol-treated samples were then serially diluted, plated on TY agar with 0.1% taurocholate and incubated at 37°C for 24 to 48 hours to enumerate the number of spores. Dividing the number of spores by the total number of CFU and multiplying the value by 100 determined the percentage of ethanol-resistant spores. The results were based on a minimum of three biological replicates.

3.5.6 Phase-contrast microscopy

C. difficile strains were grown in 70:30 medium as described above. At indicated time points, 1 ml of culture was removed from the anaerobic chamber, centrifuged at 17,000g for 1min and suspended in 30µl of sterile PBS. A thin layer of 0.7% agarose was applied to the surface of slide and 2µl of concentrated culture was placed on it. Phase contrast microscopy was per- formed using 100x oil immersion objective on OLYMPUS BX41 microscope. The PixeLINK camera was used to acquire the view of at least three fields for each strain.

3.5.7 Transmission electron microscopy

All steps in sample preparation were performed at room temperature and solutions were pre- pared in 1X PBS (phosphate-buffered saline) unless indicated otherwise. For transmission electron microscopy, cells (10^{10}) were fixed overnight in a solution of 2% glutaraldehyde and 2% paraformaldehyde. The cells were thoroughly rinsed with 1X PBS (5 minutes each) and post-fixed with 1% osmium tetroxide with constant rotation for 1–2 hours. The samples were then washed thrice with 1X PBS (5 minutes each), enblock stained with 2% Uranyl acetate in water for 1hr with light protection, and finally washed three times (5 min each) with distilled water. The cells were further dehydrated in a graded 50% -100% acetone series (vol/vol) for 5 minutes and infiltrated in graded EMBED 812/Araldite resin (Electron Microscopy Sciences) at RT with constant rotation. Thin sections of polymerized resin were placed on copper grids and stained with 2% alcoholic uranyl acetate and Reynolds' lead citrate respectively. Sections were examined with a transmission electron microscope (Philips CM100) and regions containing the cross-section of the cells were photographed at 80 kV for image analysis.

To visualize the flagella, whole bacterial cells harvested from overnight cultures were processed as above and were negatively stained with 2% uranyl acetate before transmission electron microscopy analysis

3.5.8 RNA-Seq analysis

We isolated total RNA from three biological replicates of each strain belonging to early-stationary phase (12 hours after inoculation) and quality was checked using Agilent 2100 Bioanalyzer. The RNA-Seq was performed as previously described [60]. Briefly, we depleted the rRNA content in the selected samples using Epicenter Bacterial Ribo-Zero kit. Strand-specific single end cDNA libraries were prepared using Truseq Small Stranded Total RNA sample prep kit Illumina as per the manufacturers' instructions. Illumina HiSeq2000 sequencer (multiplexing three samples per lane) was used to sequence libraries. Sequences were cleaned with AlienTrimmer [71] of adapter sequences. Only high-quality sequences with a minimum of 30 nucleotides in length were considered for further analysis. Cleaned genes were aligned to reference genomes (FN545816.1 and AM180355.1) using Bowtie (version 1.0.1) [25, 60, 72].

DESeq2 version 1.8.3 was used to perform normalization and differential analysis. Genes were considered differentially expressed if the fold change was 2: $\log_2$ 1.5 and their adjusted *p*-value was :::0.05.

3.5.9 Cloning, expression, and purification of SinR-6His, SinR'-6His, and CodY-6His proteins in *E. coli*

SinR, SinR' and CodY proteins were overexpressed in Rosetta *E. coli* DE3 cells using pET16B expression system. The ORFs for cloning were PCR amplified from JIR8094

chromosome using gene-specific primers (listed in S2 Table), and the amplified gene fragments were then digested with *Xho1* and *BamH1* to clone into pET16B digested with the same enzymes. The resulting plasmids were then transformed into *E.coli* Rosetta DE3 (Novagen) competent cells to obtain recombinant strains. To overexpress SinR-6His, and SinR'-6His, the *E. coli* recombinant strains were grown at 37°C in LB medium containing chloramphenicol (25μg ml^{-1}) and ampicillin (100ug ml^{-1}). Protein expression was achieved by inducing with 1mM IPTG at 17°C overnight. Cells were harvested by centrifugation, and the 6His-tagged proteins were purified by affinity chromatography on Ni^{++} agarose (Sigma-Aldrich) beads following the manufacturer's recommendations.

3.5.10 Antibody production

The anti-SinR used in this study was raised against SinR-His_6 in rabbits by Lampire Biologicals (Everett, PA). The anti-SinR' was raised against SinR'-His_6 in mice by Lampire Biologicals (Everett, PA).

3.5.11 Western blot analysis

C. difficile cells for western blot analysis were harvested and washed in 1x PBS solution before suspending in sample buffer (Tris 80mM; SDS 2%; and Glycerol 10%) for sonication. Whole cell extracts were then heated at 100°C for 7 min and centrifuged at 17,000 g for 1 min, and the proteins were separated by SDS-PAGE and electro-blotted onto PVDF membrane. Immobilized proteins in the membranes were then probed with specific antibodies at a dilution of 1:10,000. The blot was subsequently probed with HRP-conjugated secondary antibodies at a dilution of 1:10000. Immuno-detection of proteins was performed with ECL Kit (Thermo Sci-

entific) following the manufacturer's recommendations and were developed using Typhoon 9100 scanner.

3.5.12 Toxin ELISA

Cytosolic toxins from 16h old *C. difficile* cultures grown in TY medium were measured as described previously [70, 73]. In brief, one ml of *C. difficile* cultures were harvested and suspended in 200 µl of sterile PBS, sonicated and centrifuged to harvest the cytosolic protein. One hundred µg of cytosolic proteins was used to measure the relative toxin levels using *C. difficile* premier Toxin A &B ELISA kit from Meridian Diagnostics Inc. (Cincinnati, OH).

3.5.13 Motility assay

C. difficile cultures were grown until mid-exponential phase at 37°C. After adjusting their OD@600 to 0.5, 3µl of each strain was inoculated by stabbing or spotting into BHI medium with 0.3% w/v agar in tubes and plates respectively. After incubation at 37°C, the motility was quantified by measuring the radius of the cultures at different time points. Motility assay was performed in 4 replicates and independently repeated at least three times

3.5.14 SinR-6His; SinI-GST pull-down experiment

To express SinR'-GST protein we cloned the *sinR'* gene in the pGST-parallel2 expression system [74]. First, the *sinR'* gene was PCR amplified using primers ORG619 and ORG620 (Table S 3.2) and R20291 chromosomal DNA as a template. The PCR fragments were then cloned in between *NcoI* and *SalI* sites of the pGST-parallel2 vector. The resulting plasmid was then transformed into *E.coli* Rosetta DE3 competent cells to obtain recombinant strain. To overexpress SinR'-GST, *E. coli*

recombinant strains were grown at 37°C in LB medium containing chloramphenicol (25µg ml^{-1}). Protein expression was achieved by inducing with 1mM IPTG at 17°C overnight with mild agitation. To perform the pull-down experiment, 200 µgs of whole cell lysate proteins from the *E. coli* cells expressing SinR'-GST was mixed with ~20 µgs of purified SinR-6His protein and incubated at 4°C for 1hr. The mixture was then passed through the Ni^{++} affinity column (Sigma-Aldrich) to trap and elute SinR-6His protein. Whole lysates from *E. coli* cells expressing GST alone was also mixed with purified SinR-6His protein, and this control mixture was processed in the same way as the test sample. The elutes from Ni++ columns were then separated by SDS-PAGE and were electro blotted onto PVDF membrane. Membranes with immobilized proteins were then probed with either Anti-6His antibodies at 1:10,000 dilution or with anti-GST antibodies at the dilution of 1:5000. Immunodetection of proteins was performed with Pierce ECL 2 Western blotting Substrate Kit (Thermo Scientific) and the Typhoon 9100 scanner.

3.5.15 Electrophoretic mobility shift assay (EMSAs)

SinR and SinR' binding was performed with radioactively labeled DNA probes. The *codY* up- stream and the *gluD* upstream regions were amplified using primer pairs ORG629- ORG630 and ORG72-ORG73, respectively and the products were cloned into a pGEMT cloning vector. The region was then excised from the plasmid construct using *EcoRI* and was radiolabeled using Klenow fragment of DNA polymerase I (NEB. labs) and [α- ^{32} P]dATP-6000 Ci/mmol (PerkinElmer Life Sciences). Binding experiments with radioactively labelled *codY* upstream DNA with SinR-6His or SinR'-6His was performed using reaction buffer containing 10 mM Tris–HCl (pH 8.0), 0.1 mM DTT, 150 mM KCl, 0.5mM EDTA, 0.1% Triton X-100 and 12.5% glycerol. For binding experiments containing both SinR and SinR', proteins were mixed in the reaction buffer at a specified concentration and were incubated at room temperature for 30 minutes

before adding the DNA probe. Reactions were loaded onto a 6% native polyacrylamide gel in 1XTBE (Tris/Borate/EDTA) and subjected to electrophoresis at 100 V for 45 minutes. Gels were then dried, and the autoradiography was performed with Molecular Dynamics Phosphor-Imager technology.

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

3.5.16 Hamster model for *C. difficile* pathogenesis

Syrian golden hamsters (100–120 g) were used for *C. difficile* infection. Upon their arrival, fecal pellets were collected from all hamsters, homogenized in 1 ml saline, and examined

for *C. difficile* by plating on CCFA-TA (Cycloserine Cefoxitin Fructose Agar- 0.1% Taurocholate) to ensure that the animals did not harbor indigenous *C. difficile*. After this initial screen, they were housed individually in sterile cages with *ad libitum* access to food and water for the duration of the study. Hamsters were first gavaged with 30 mg/kg clindamycin [59, 75]. *C. difficile* infection was initiated five days after clindamycin administration by gavage with vegetative cells. We used vegetative *C. difficile* cells because of the test strain R20291::*sinRR*' is asporogenic and do not produce any spores. Bacterial inoculums were standardized and prepared immediately before challenge as described in our earlier study [59]. They were transported in independent 1.5 ml Eppendorf tubes to the vivarium using the Remel AnaeroPack system (one box for each strain) to maintain viability. Immediately before and after infecting the animal, a 10 µL sample of the inoculum was plated onto TY agar with cefoxitin to confirm the bacterial count and viability. There were five groups of animals, including the uninfected control group. Ten animals per group were used for the infection. Approximately, 2000 *C. difficile* vegetative cells of R20291 strain and R20291::*sinRR*' were used for the animal challenge. In the uninfected control (group 5) only five animals were used, and they received only antibiotics and sterile PBS. Animals were monitored for signs of disease (lethargy, poor fur coat, sunken eyes, hunched posture, and wet tail) every four hours (six times per day) throughout the study period. Hamsters were scored from 1 to 5 for the signs mentioned above (1-normal and 5-severe). Fresh fecal pellets were collected daily from every animal to monitor *C. difficile* colonization until they began developing diarrheal symptoms. Hamsters showing signs of severe disease (a cumulative score of 12 or above) were euthanized by CO_2 asphyxiation. Surviving hamsters were euthanized 15 days after *C. difficile* infection. Thoracotomy was performed as a secondary mean of death. The cecal contents from these hamsters were collected in 15ml Nalgene tubes, secured air tight and

were transported to the lab using Remel AnaeroPack system. They were then immediately subjected to CFU enumeration. For CFU enumeration, the daily fecal samples or the cecal contents collected post-mortem were resuspended in 1X PBS, serially diluted and plated onto CCFA agar with 0.1% Taurocholate (CCFA-TA). The CFU were counted after 48 h of incubation. The survival data of the challenged animals were graphed as Kaplan-Meier survival analyses and compared for statistical significance using the log-rank test using GraphPad Prism 6 software (GraphPad Software, San Diego, CA).

3.6 Results

3.6.1 Comparison of *C. difficile* and *B. subtilis sin* loci

In *B. subtilis*, the *sin* locus carries two small ORFs, *sinI* and *sinR* [32, 33] (Figure 3.1A). *B. subtilis* SinR (BsSinR) is a DNA-binding protein that binds to a conserved DNA sequence upstream of the translational start site of target genes to negatively control their transcription. SinI, encoded by a gene adjacent to *sinR*, has an antagonistic relationship with SinR and binds directly to the SinR protein to inhibit its activity. This causes the pathways that were repressed by SinR to switch on. In *B. subtilis*, SinR contains 113 aa, and the DNA binding domain is located at the N-terminus part, which spans from residues 5–61 [32, 33] (Figure 3.1A). The C-terminal part ofSinR forms alpha-helices and is responsible for multimerization and SinI interaction. The SinI protein, on the other hand, resembles a truncated SinR without the DNA binding region and carries only the alpha-helical structure to drive the hetero-dimerization of SinR-SinI complex [32–34]. In *C. difficile* the *sin* locus contains two ORFs CDR20291_2121 and CDR20291 _2122 (in *C. difficile* R20291 reference genome), which codes for proteins that are 43% and 35% identical to *B. subtilis* SinR, respectively (Figure 3.1B). Both these proteins are predicted to be DNA-

binding since they carry HTH (Helix-Turn-Helix) domains in their N-terminal regions. Hence we named CDR20291_2121 as *sinR* and CDR20291_ CD2122 as *sinR'*.

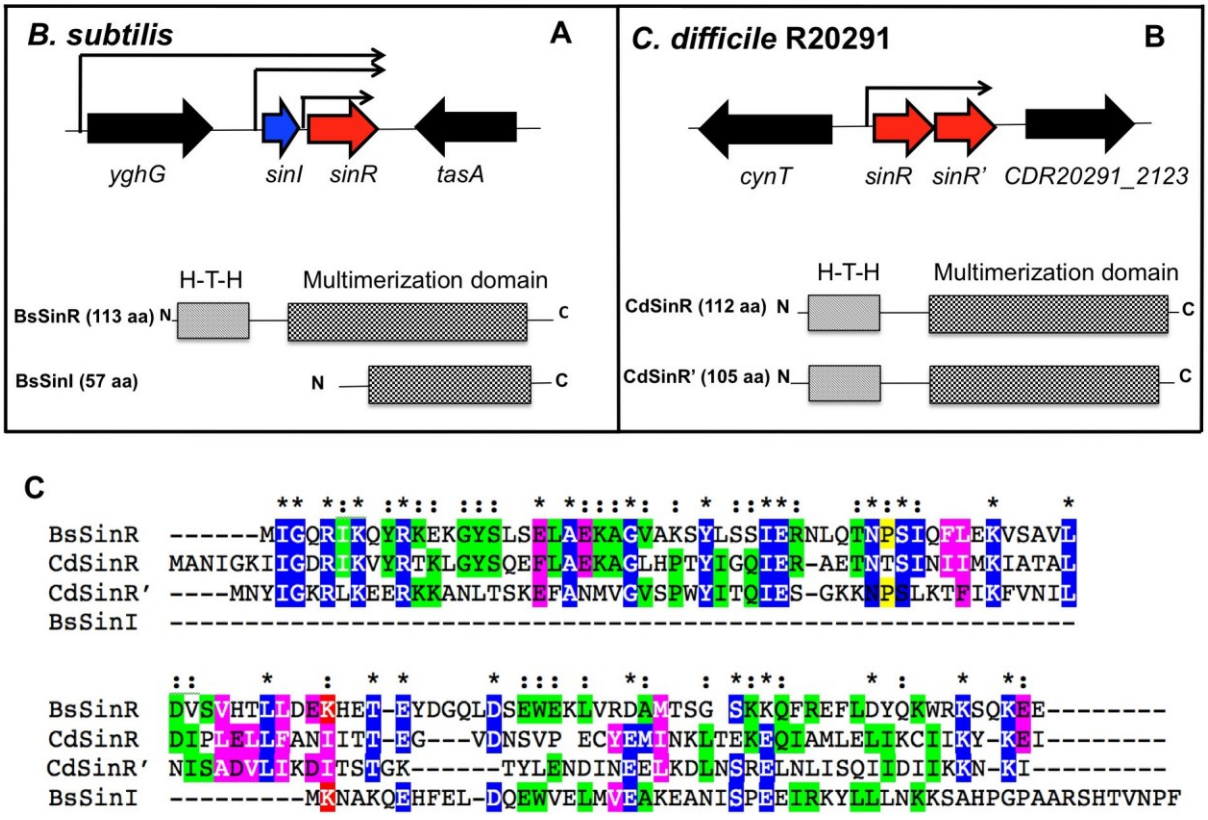


Figure 3.1 Genetic organization of genes in the sin locus

B. subtilis (A) and *C. difficile* R20291 strain (B). The different domains within Sin proteins are presented below. (C) Sequence alignment of the *C. difficile* SinR (CdSinR) and SinR' (CdSinR') with *B. subtilis* SinR (BsSinR) and SinI (BsSinI) using ClustalW.

The *C. difficile* SinR (CdSinR) contains 112 amino acids, and its predicted HTH domain spans residues 11 to 66. The SinR' (CdSinR') protein carries 105 aa, and its predicted HTH domain spans from residues 7 to 62 (Figure 3.1B). Both CdSinR and CdSinR' shows the highest homology to BsSinR in this DNA-binding domain, where within the 50 residues of HTH domain, 13 of them are identical and 19 of them represent conservative substitutions (Figure 3.1C). CdSinR and CdSinR' shows similarity with each other (33% identity) only in their N terminal DNA binding

domain. The C terminus multimerization domains of these proteins show variations, and there is less similarity of CdSinR and CdSinR' to BsSinR and each other in this region.

In various *Bacillus sp*. SinR homologs are known to control the expression of the genes adjacent to the *sin* loci. Thus, identifying genes adjacent to the *sin* loci were helpful in predicting at least a few functions of the Sin regulators in these bacterial species. For example, in *B*. *subtilis*, the *sin* locus is adjacent to the *tapA-sipW-tasA* operon, and SinR represses the expression of this operon whose products are involved in the production of the biofilm matrix [31]. In *Bacillus anthracis*, the *sin* locus is next to *calY* that codes for camelysin, a cell surface associated pro- tease, and SinR in this species is known to repress the *calY* expression [35]. In *C. difficile*, the *sin* locus is located in between *cynT* (codes for carbonic anhydrase) and CDR20291_2123 (unknown function) (Figure 3.1B) and is not close to any other genes that are known to be essential for virulence in this pathogen. Thus, the location of the *sin* locus in *C. difficile* chromosome did not provide us any clues about its possible functions. To get more information about the locus and its role in *C. difficile* physiology we decided to construct and characterize mutants in *sin* locus.

3.6.2 Construction and verification of sinRR' mutants in *C. difficile* strains JIR8094 and R20291

An erythromycin resistant marker was introduced in the *sinR* at nucleotide 141 using Clos- tron, a TargeTron-based group II intron in *C. difficile* JIR8094 [36] and R20291 strains [37]. The presence of the retargeted intron in the correct gene in both mutant strains was confirmed by PCR (Figure S 3.1). In *B*. *subtilis*, three different promoters drive the transcription of the *sin* genes [33]. In *B*. *subtilis*, the polycistronic *sinIR* transcript is produced from two different promoters, and the *sinR* transcript is driven from an independent promoter immediately downstream of *sinI* (Figure

3.1A) [33]. In *C. difficile*, the operon upstream of *sin* locus transcribes in the opposite direction, and no read-through transcription of *sin* locus is possible from its promoter (Figure 3.1B). Using cDNA prepared from the JIR8094 and the mutant strain, we performed RT-PCR analysis and checked for the presence of *sinR*, *sinR'* and *sinRR'* transcripts (Figure S 3.2)

We could detect *sinR*, *sinR'* and also the read through *sinRR'* transcripts, which confirmed that the *sinR* and *sinR'* are transcribed as a single transcript (Figure S 3.2). When the same analysis was performed using the mutant strain cDNA both the *sinR*, *sinR'* and *sinRR*' transcripts were absent (Figure S 3.2). The QRT-PCR analysis of the *sinR* mutant showed significant reduction of both *sinR* and *sinR'* transcript levels (Figure S 3.2). It also revealed that similar to *B*. *subtilis*, the *C. difficile sin* locus is expressed between late-exponential and early-stationary growth phase (10 to 12 h) (Figure S 3.2). Similar results were obtained in RT-PCR analyses of cDNA from the R20291 strain (Figure S 3.2). When we performed the western blot analysis using the SinR and SinR' specific antibodies (see M&M), both SinR and SinR' were found to be absent in the mutant (Figure S 3.3)

Our western blot and the RT-PCR results together suggest that *sinR and sinR'* are part of an operon. However, there is a possibility that *sinR'* could have an independent promoter coded within the *sinR* coding region, which was not expressed in the growth conditions tested. Since the insertion of the intron in *sinR* (first gene in the operon) disrupted both *sinR*, *sinR'* transcripts, and SinR, SinR' production in the growth conditions tested, we named the mutant strains with the disrupted *sinR* gene as JIR8094::*sinRR'* and R20291::*sinRR'*.

3.6.3 Impact of *sinRR'* inactivation in *C. difficile*

We first analyzed the impact of *sin* locus inactivation on the growth of *C. difficile* in TY medium. During the exponential phase of the growth, both parents and mutants grew at a similar

rate. However, when they entered the stationary phase, we observed a decrease in the turbidity of the mutant cultures as measured as OD@600 nm (Figure S 3.3). We performed the Triton X-100 autolysis assay to check the influence of SinRR' on global autolysis of *C. difficile* [25].

We used the 16h old stationary phase culture to perform this assay, where the R20291::*sinRR'* lysed at a faster rate compared to the parent (Figure S 3.4). These results suggested that inactivation of *sinRR'* induced autolysis in *C. difficile*. In *B. subtilis*, SinR along with another regulatory protein SlrR represses the expression of *lytA-lytB-lytC* and *lytF* autolysins [38]. Our initial observation of lysis phenotype in the *sinRR'* mutants suggested that like *B. subtilis* SinR, *C. difficile* SinR might also be controlling the autolysin genes. In *B. subtilis* the SinR is a pleiotropic regulator and controls various pathways including autolysis [29–31, 33, 38, 39]. We suspected that SinR and SinR' in *C. difficile* might also regulate several targets to control multiple functions. Hence, to identify the *sinRR*' regulated pathways in *C. difficile*, we performed the transcriptome analysis of the *sinRR'* mutants in comparison with their respective parents.

3.6.4 Assessment of the *sinRR'* regulon in *C. difficile*

Based on the growth pattern of the *sinRR'* mutants (Figure S 3.4) and the expression kinetics of *sinRR'* in the parent strains (Figure S 3.2) we decided to compare the transcriptomes of mutant strains with their respective parent strains during the early stationary phase (i.e., 12 h of growth) in TY medium. We used three biological replicates and genes were considered differentially expressed if the fold change was ::: $\log_2$ 1.5 and their adjusted *p*-value was :::0.05.

In the RNA seq analysis, it was observed that 437 and 425 genes were over-expressed in R20291::*sinRR'* and in JIR8094::*sinRR'* mutant strains, respectively, while 668 and 208 genes were under-expressed in R20291::*sinRR'* and JIR8094::*sinRR'* mutant strains, respectively.

Results from the transcriptome analysis confirm that as in *B. subtilis*, SinRR' in *C. difficile* also regulates a wide range of genes involved in several pathways including sporulation, motility, metabolism, membrane transport, stress response and toxin synthesis (Figure 3.2A) (Figure 2A). A list of genes identified to be differentially regulated in mutants R20291::*sinRR' and* JIR8094::*sinRR'* compared to their parent strains are listed in Table S 3.4, Table S 3.5, Table S 3.6 and Table S 3.7 respectively. To test and validate the transcriptome profiles, we performed relevant phenotypic assays and functional analysis with parent and mutant strains for major pathways (sporulation, motility, toxin production and autolysis) that were suggested to be regulated by SinR and SinR'.

We have included following strains in the phenotypic analysis: parent strain, *sinRR'* mutant, *sinRR'* mutant with pRGL311 (plasmid with *sinRR'* under its native promoter), and *sinRR'* mutant with pRG334 (plasmid with *sinRR'* under the inducible promoter). To determine the independent role of SinR and SinR' in the phenotypes, the *sinRR'* mutant with plasmids: pRG300 (*sinR* gene alone with its promoter region); pRG310 (*sinR* under the inducible promoter); and pRG306 (*sinR'* alone under the inducible promoter) were used. Western blot analysis with SinR and SinR' specific antibodies were performed to confirm their expressions from the constructs, and the *sinRR'* mutant with vector alone was used as negative controls (Figure 3.2B). Growth curve analysis showed when *sinRR'* was expressed from its promoter or the inducible promoter in the *sinRR'* mutant, no autolysis was observed, and they grew similar as the wild type (Figure 3.2C and (Figure S 3.4). In the Triton X-100 autolysis assay, a partial recovery from autolysis was observed when either SinR or SinR' alone was expressed in the mutant (Figure S 3.4).

3.6.5 *C. difficile sinRR'* mutants are asporogenic

To determine the role of *sinRR'* on sporulation, we grew the test strains on 70:30 sporulation agar for 30h. Initial analysis through phase contrast microscopy detected no spores in R20291:: *sinRR'* (Figure 3.3A). Transmission electron microscopy (TEM) further confirmed this observation (Figure 3.3B). Fully mature spores could be detected in R20291, whereas the *sinRR'* mutant cells were devoid of any spores. Similar results were obtained for JIR8094::*sinRR'* mutant as well (Figure S 3.5).

We performed ethanol treatment-based sporulation efficiency assay where the ability of the bacteria to produce viable spores were analyzed by counting the total number of CFU (Colony Forming Units) following ethanol treatment. The mean sporulation efficiency of the parental strain R20291 was 18.7% (Figure 3.3C). The *sinRR'* mutant strain did not produce any spores, and the percentage of sporulation was near zero.

We were surprised by the observation that expression of either *sinRR'* or *sinR*/*sinR'* alone also did not revive the sporulation in the *sinRR'* mutants (Figure 3.3C). Sporulation in *C. difficile* is initiated with the activation of Spo0A, which in turn triggers early sporulation gene transcription [22, 40]. Transcripts of *spo0A* were 3.5-fold and 2.9-fold under-expressed in JIR8094::*sinRR'* and in R20291::*sinRR'* strains respectively, when compared to parent strains. We performed western blot analysis with the Spo0A specific antibodies [41]. We detected GDH (glutamate dehydrogenase) for loading control since its production was found to be unaffected in the *sinRR'* mutants. Western blot analysis showed that in R20291:: *sinRR'* the Spo0A was absent or below the detectable level (Figure 3.3C, Figure S 3.5). Lower production of Spo0A can result in down-regulation of all sporulation genes under its control. Our transcriptomic data indeed found many sporulation-associated genes to be affected (Table 3.1, Table S 3.4 and Table S 3.6) in the *sinRR'*

mutant. The QRT-PCR analysis performed on selected sporulation genes confirmed their down-regulation in the *sinRR'* mutants (Table S 3.8).

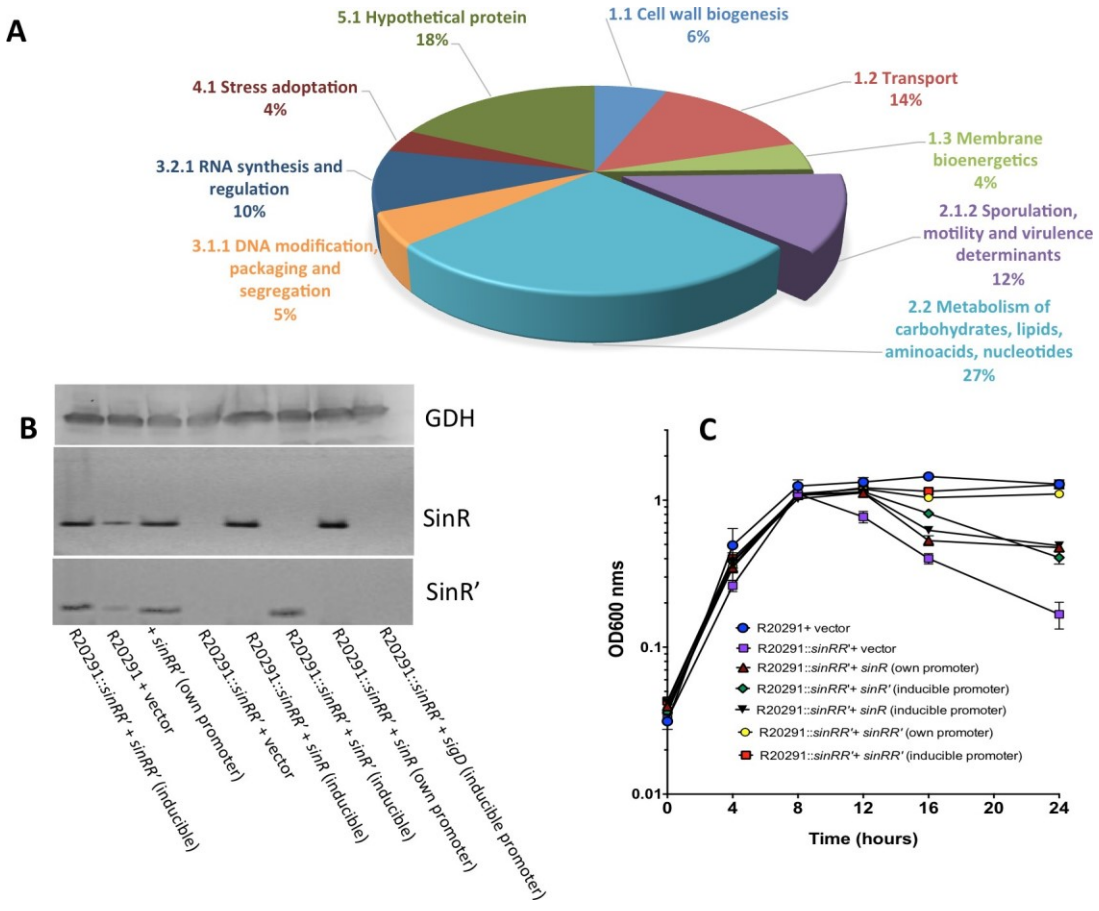


Figure 3.2 Characterization of sin locus (sinRR') mutant in *C. difficile*.

(A) Functional categorization of genes affected by sin locus mutation in R20291 strains based on RNA seq data. (B) Western blot analysis with SinR and SinR' specific antibodies demonstrating the absence of both SinR and SinR' in the *sinRR'* mutants and their presence after the complementation. GDH detection using anti-GDH antibodies was used as loading control. (C) Growth curve of the parent (R20291), *sinRR'* mutant and the *sinRR'* mutant complemented strains in TY medium. The data shown are means ± standard errors of three replicates.

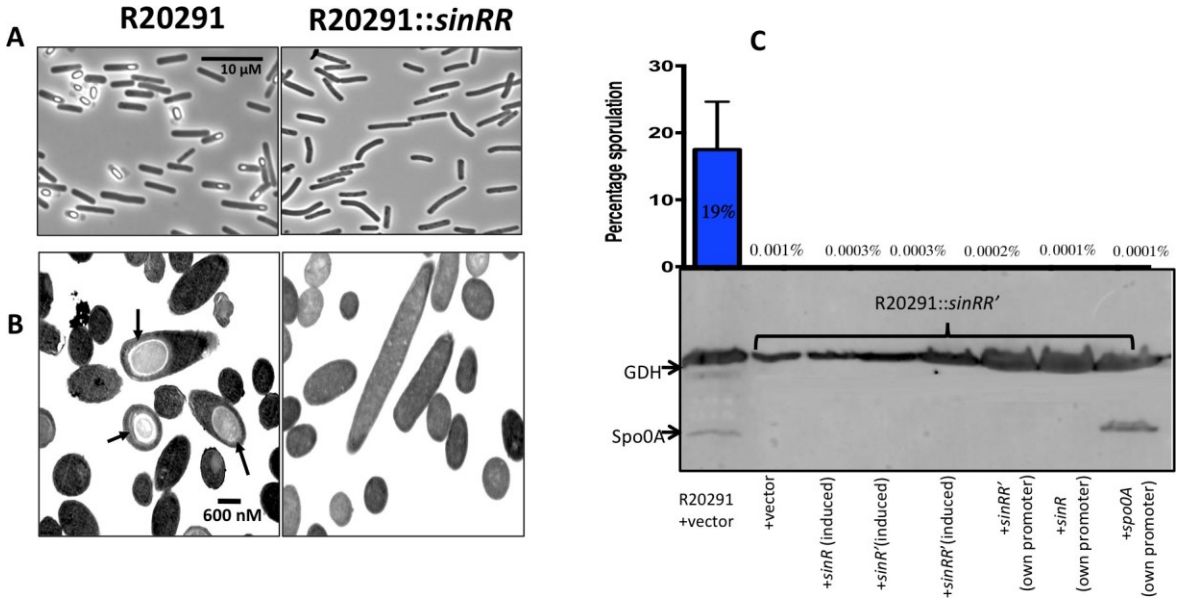


Figure 3.3 Sporulation in sinRR' mutant.

(A) Phase contrast microscopy of paraformaldehyde-fixed R20291::*sinRR'* strains revealed no spores. (B) R20291::*sinRR'* was asporogenic as shown in representative TEM images in comparison with the parent strain. Black arrows indicate mature spores in parent strains. C. Asporulation phenotype of *sinRR'* mutant could not be complemented. Sporulation frequency (CFU/ml of ethanol resistant spores) of R20291, *sinRR'* mutant and mutant complemented with different constructs were determined. The *sinRR'* mutant strain expressing spo0A from its own promoter was also included in this analysis. Below the sporulation frequency graph is the multiplex-western blot analysis of *sinRR'* mutant complemented strain proteins using Spo0A and GDH specific antibodies.

Since our transcriptome analysis and western blot analysis revealed a lower Spo0A in R20291::*sinRR'*, we decided to test whether the asporogenic phenotype of the *sinRR'* mutants is due to the lower production of Spo0A. We expressed *spo0A* from its native promoter (pRGL312) in the R20291::*sinRR'* and production of Spo0A in *sinRR'* mutants was verified through the western blot analysis using Spo0A specific antibodies (Figure 3.3C) [41]. To our surprise, production of Spo0A in the *sinRR'* mutants did not induce the sporulation in the R20291::*sinRR'* strain (Figure 3.3C). For sporulation to proceed normally, the Spo0A protein should get activated by phosphorylation [42]. Spo0A~P then acts as a transcriptional activator for many downstream genes in the sporulation pathway that includes sigma factors, the forespore specific *sigF*, and the

mother cell-specific *sigE* [22, 40, 42]. We performed QRT-PCR to detect the transcripts of Spo0A~P activated *sigF* and *sigE* genes. We did not observe increases in *sigF* and *sigE* transcript levels in the *spo0A* expressing *sinRR'* mutant when compared to the *sinRR'* mutant with vector alone control. This result suggests that activation of Spo0A to Spo0A~P is affected in the *sinRR'* mutant.

In *Bacillus sp*., the pathway that controls Spo0A phosphorylation is well characterized [43– 47]. In *Clostridia*, the components of this phosphorelay are absent, and it has been hypothesized that sporulation-associated sensor kinases may directly phosphorylate the Spo0A for its activation. In *C. difficile*, four orphan kinases (CD630_01352, CD630_2492, CD630_01579, and CD630_1949) are present, among which, the CD630_1579 kinase was shown to phosphorylate Spo0A *in vitro*, and the CD630_2492 mutant was found to be less efficient in sporulation [48]. In the transcriptome data, the CD630_1579 and the CD630_ 2492 kinases were to be under-expressed ~1.5-fold and ~3-fold, respectively, in the JIR8094::*sinRR'* mutant. However, their homologs CDR20291_1476 and CDR20291_2385 in the R20291::*sinRR'* were not affected suggesting that these kinases might not be the main reason for Spo0A inactivation in the *sinRR'* mutants. Since the regulatory network of Spo0A activation is largely unknown, there is a possibility that unknown kinases could have been affected in *sinRR'* mutants.

Table 3.1 Under-expressed sporulation genes in R20291::*sinRR'*.

Locus Tag	Gene	Protein Name	Fold-Change: R20291/*sinRR'* mutant	log2 ratio	Known/predicted sigma factor needed for expression	Adjusted *p* value
CDR20291_0104	*cwlD*	Germination N-acetylmuramoyl-L-alanine amidase, Autolysin	67.7	6.1	SigE	1.01E-09
CDR20291_0125	*spoIIID*	Stage III sporulation protein D	68.0	6.1	SigE	6.24E-05
CDR20291_0128		putative sporulation protein yyac, DUF1256 family	50.2	5.6	SigE	0.005
CDR20291_0213		hypothetical protein	3041.9	11.6	SigE	6.05E-14
CDR20291_0316		spore coat assembly asparagine rich protein	81.8	6.4	SigE	0.00057
CDR20291_0713		Putative sporulation protein YunB	63.0	6.0	SigE	0.00666
CDR20291_1005		Putative membrane protein, BDBH YlbJ involved in spore cortex formation	5.7	2.5	SigE	5.93E-06
CDR20291_1030	*spoIIIAA*	Stage III sporulation protein AA	18464.8	14.2	SigE	0.00005
CDR20291_1031	*spoIIIAB*	Stage III sporulation protein AB	90.8	6.5	SigE	8.33E-06
CDR20291_1033	*spoIIIAD*	Stage III sporulation protein AD	3571.7	11.8	SigE	1.32E-08
CDR20291_1034	*spoIIIAE*	Stage III sporulation protein AE	13.9	3.8	SigE	1.48E-09
CDR20291_1035	*spoiIIIAF*	Stage III sporulation protein AF	119.5	6.9	SigE	0.00825
CDR20291_1036	*spoIIIAG*	Stage III sporulation protein AG	794.0	9.6	SigE	2.08E-09
CDR20291_1051	*spoIVB*	Stage IV sporulation protein AB	6.2	2.6	SigE, SigG	6.61E-06
CDR20291_1073		Putative phage protein, skin element	98.9	6.6	SigE	0.00821
CDR20291_1282	*cotE*	Spore coat protein CotE peroxiredoxin/chitinase	18.7	4.2	SigE	0.00029
CDR20291_1360	*cotB*	Spore outer coat layer protein CotB	332.0	8.4	SigE	1.03E-07
CDR20291_2146	*cspC*	Subtilisin-like serine germination related protease-CspC	16.3	4.0	SigE	0.00921

CDR20291_2147	*cspBA*	Subtilisin like serine germination related protease, CspB	19.5	4.3	SigE	7.88E-07
CDR20291_2289	*cotJA*	Putative spore coat protein	103	7.0	SigE	0.00074
CDR20291_2291	*cotD*	Spore coat protein CotD manganese catalase	139.6	7.1	SigE	0.056
CDR20291_2334	*spoIV*	Stage IV sporulation protein	4.1	2.0	SigE	9.981E-06
CDR20291_2335		putative sporulation protein yyac	1593.4	10.6	SigE	0.00097
CDR20291_2513	*spoIVA*	Stage IV sporulation protein AA	309.1	8.3	SigE	2.35E-06
CDR20291_2573	*spoIIE*	Stage II sporulation protein E	36.13938	5.17	SigE	7.04E-05
CDR20291_3331		Putative spore protein	3.1	1.6	SigE	6.75E-09
CDR20291_3376	*spmB*	Spore maturation protein B	41.4	5.4	SigE	0.00081

3.6.6 *C. difficile sinRR'* mutants are non-motile

The JIR8094 strain was intrinsically non-motile due to mutations within the flagellar operon [49]. Hence, we choose only R20291 and R20291::*sinRR'* to perform motility-related experiments. The R20291::*sigD* mutant and the R20291::*sinRR'* strains with vector alone (pRPF185) were used as the controls. Exponentially growing bacterial cultures were spotted on BHI with 0.3% agar and was incubated at 37˚C for 36h to monitor motility. The bacterial cultures expressing *sinRR'*, or *sinR* or *sinR'* from the tet-inducible promoters were spotted on BHI with 50 ng/ml of ATc and 0.3% agar. In the motility assays, the R20291::*sinRR'* strain was defective in motility (Figure 3.4C and Figure S 3.6). The transcriptome analysis supported our observation, where *sigD*, the sigma factor needed for the transcription of the flagellar operons, was found to be 14-fold under-expressed in the R20291::*sinRR'* (Figure 3.4A, Table S 3.4) along with other motility-related genes. Electron microscopic analysis followed by negative staining failed to detect flagellar structures in the R20291::*sinRR'* (Figure 3.4B).

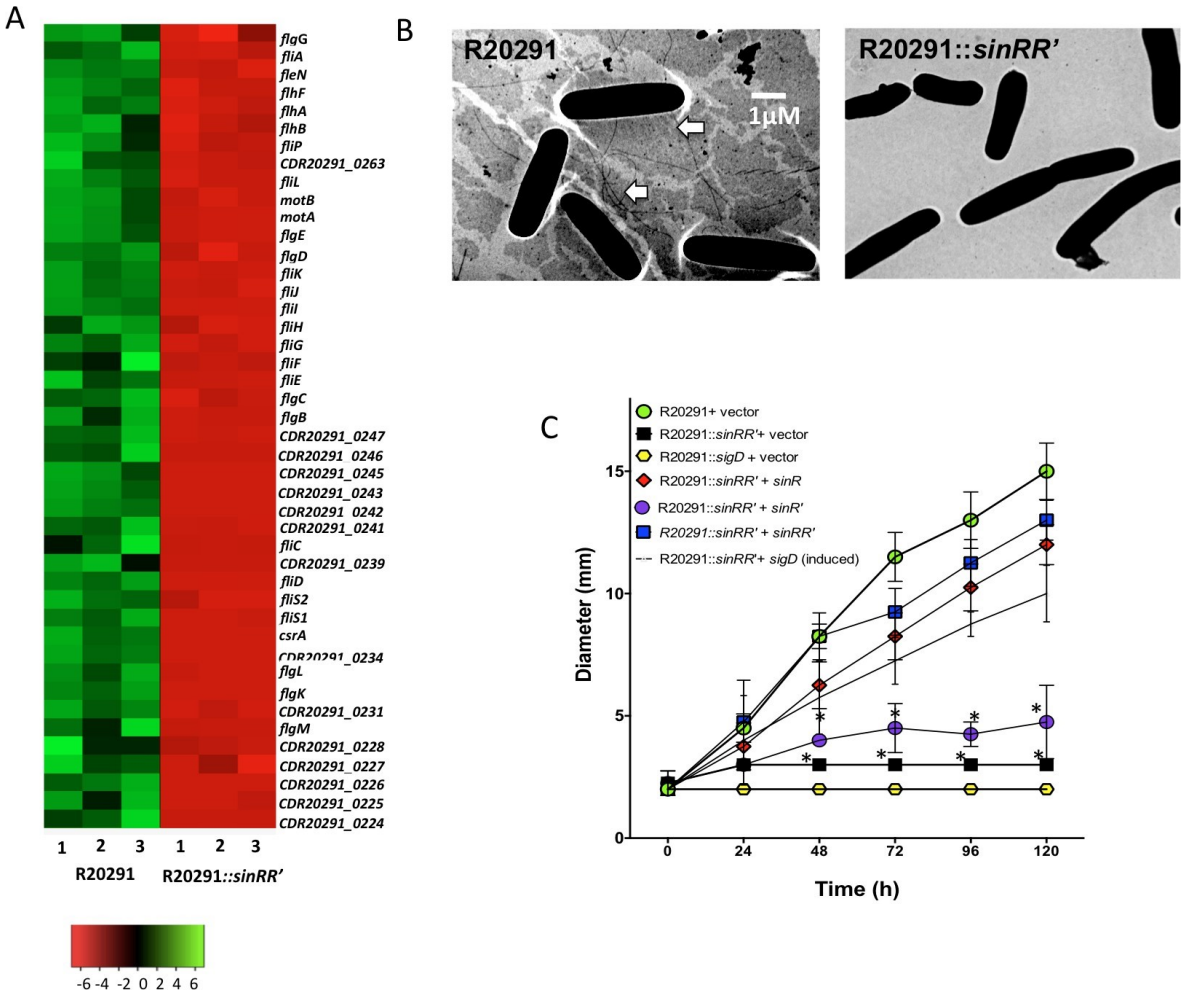


Figure 3.4 Mutation in the sin locus affects *C. difficile* flagellar synthesis.

(A) Heat map showing the lower expression of flagellar and motility-related genes in the R20291::*sinRR'* mutant compared to the parent. Color intensity in each cell represents corresponding Log2 expression values in the color scale bar. (B) Transmission electron micrographs of negatively stained *C. difficile* cells. White arrows point to flagella. (C) Motility of R20291, *sinRR'* mutant and the *sinRR'* mutant complemented strains in BHIS with 0.3% agar. The *sigD* mutant and the *sinRR'* mutant expressing s*igD* from an inducible promoter were included in this analysis. The swim diameters (mm) was measured every 24 h for a total of 120 h is shown and the data shown are means ± standard errors of three biological replicates. The experiments were repeated at least three times independently (*, p:::0.05 by a two-tailed Student's t-test).

A dot blot analysis with FliC (the flagellar structural protein) specific antibodies also confirmed the absence of flagella in the R20291::*sinRR'* strain (Figure S 3.6). Expression of *sinRR'*

from its promoter or the inducible pro- moter revived the motility (Figure 3.4C). Interestingly, expression of SinR alone was sufficient to revive the motility in the R20291::*sinRR'* strain, whereas the SinR' expression alone did not have any effect (Figure 3.4C).

SigD is needed for the transcription of the flagellar operon in *C. difficile* [25, 26]. To determine whether the non-motile phenotype of *sinRR'* mutant is due to the reduced levels of *sigD* in the *sinRR'* mutants, we expressed *sigD* from the tetracycline-inducible promoter by introducing the construct pRGL291 into the R20291::*sinRR'* strain (Figure S 3.1). We observed motility was partially restored in the R20291::*sinRR'* when the *sigD* expression was induced (Figure 3.4C), suggesting that *sinRR'* controls motility by controlling the expression of *sigD* in *C. difficile*.

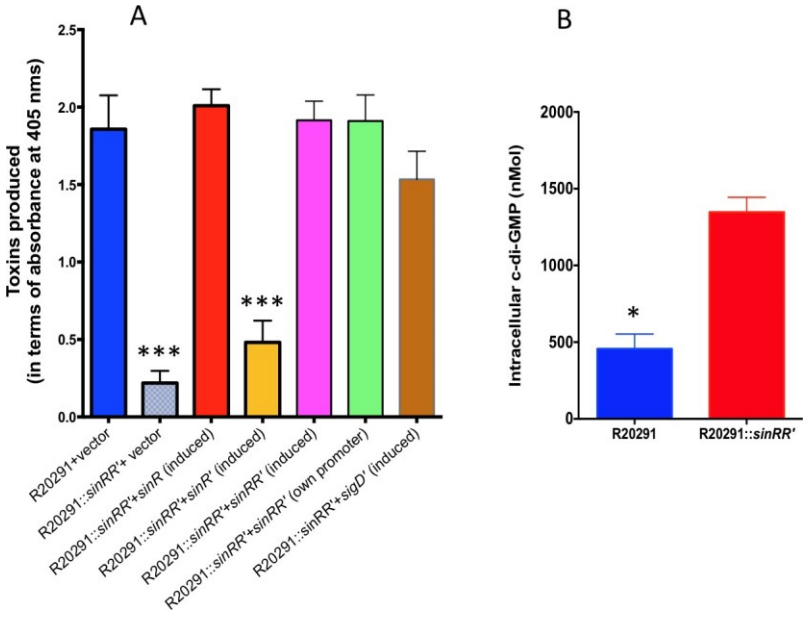


Figure 3.5 SinRR' positively influences the expression of PaLoc genes.

(A) Quantification of toxins in parent R20291 and the *sinRR'* mutant complemented strains using toxins specific ELISA. The data shown are means ± standard errors of three replicates. Statistical significance was tested using one-way ANOVA, followed by Dunnett's multiple comparisons test comparing values to the average of the parent with vector control (*** <0.0005 p-value). (B) Increased intracellular levels of c-di-GMP in the *sinRR'* mutant. Statistical analysis was performed using two-tailed t- test (* <0.05 p-value).

3.6.7 *C. difficile sinRR'* mutants produce less toxins than their parent strains

The transcriptome analysis and the follow-up QRT-PCR (Figure 3.5A, Table 3.2, Table S 3.4, Table S 3.6 and Table S 3.8) result suggested *sin* locus's role in toxin gene regulation. Toxin ELISA was performed with the cytosolic protein extracts of *sinRR'* mutants and their respective parent strains.

Bacterial cultures expressing either *sinRR'* or *sinR/sinR'* alone from the tetracycline-inducible promoter were grown for 6h in TY medium and were induced with 50ng/ml of ATc for 5 hours. Cytosolic proteins harvested from these induced cultures were used for toxin ELISA. We observed a six-fold reduction in toxin production (Figure 3.5A) in the R20291::*sinRR'* when compared to the R20291 strain. In JIR8094::*sinRR'* however, a moderate two-fold reduction in toxin level was recorded when compared to the parent strain (Figure S 3.7).

Table 3.2 Expression levels of PaLoc genes and their regulators in R20291::*sinRR'*.

Gene	Known or predicted function	RNA-Seq Analysis Fold-change WT/mutant			Q-RT-PCR Analysis Fold-change WT/mutant			Expression in *sinRR'* mutant
		Actual	Log_2 ratio	Adj.*p* value	Actual	Log_2 ratio	Adj.*p* value	
tcdR	Sigma factor for toxin genes	32.9	5.0	1.11E-16	6.09	2.06	5.35E-03	Under-expressed
tcdB	Toxin B	88.4	6.5	5.80E-12	8.02	3.00	1.94E-11	Under-expressed
tcdE	Holin like protein	44.2	5.5	6.57E-05	6.29	2.65	3.32E-04	Under-expressed
tcdA	Toxin A	13.1	3.7	3.04E-04	7.89	2.98	0.00452	Under-expressed
sigD	Sigma factor for flagellar operon	14.4	3.8	4.13E-08	24.76	4.63	6.39E-03	Under-expressed
dccA	Diguanylate cyclase	0.10	-3.3	2.05E-24	0.008	-6.93	9.67E-04	Over-expressed
codY	GTP sensing transcriptional regulator	0.35	-1.5	7.94E-18	0.12	-3.03	2.23E-05	Over-expressed
ccpA	transcriptional regulator	1.29	0.4	3.45E-03	1.95	-0.97	8.45E-08	No significant change

Expression of *sinRR'* in the mutants brought the toxin production back to the level comparable to the parent strains. As we observed in the motility assay, expression of *sinR* alone was sufficient to bring back the toxin production in the *sinRR'* mutant, while expression of *sinR'* did not show any effect. In *C. difficile*, SigD positively regulates *tcdR*, the sigma factor needed for toxin gene transcription [25, 26]. Interestingly, the expression of *sigD* from an inducible promoter

revived the toxin production in *sinRR'* mutants, suggesting that *sinRR'* controls both toxin production and motility by regulating *sigD* in *C. difficile*.

3.6.8 Elevated c-di-GMP levels are present in *sinRR'* mutant

We observed that SigD expression in the *sinRR'* mutants partially recovered both the motility and the toxin production in that strain (Figure 3.4C and Figure 3.5A). The main question that arises from this observation is how SinR controls *sigD* expression. The *sigD* gene is part of the flagellar operon, whose transcription is directly controlled by the intracellular cyclic di-GMP (c-di- GMP) concentration [26, 50]. Within the cells, the c-di-GMP is synthesized from two molecules of GTP by diguanylate cyclases (DGCs) and is hydrolyzed by phosphodiesterases (PDEs) [50, 51]. The functionality of several of these *C. difficile* DGCs and PDEs has been confirmed by expressing them heterologously in *Vibrio cholerae*, where they resulted in phenotypes (bio- film formation and motility) that correspond to elevated or lowered levels of intracellular c-di- GMP [51]. In *C. difficile* when CD630_1420 (*dccA*) was expressed from an inducible promoter, it resulted in elevated levels of intracellular c-di-GMP and reduced bacterial motility [50]. In R20291::*sinRR'*, ten-fold more (-3.3 Log_2 fold) *dccA* (CDR2029_1267) transcript was observed (Figure S 3.5) compared to parent. We measured the intracellular concentration of c-di-GMP (Figure S 3.8) and observed a nearly three-fold increase in the c-di-GMP concentration in the *sinRR'* mutant compared to the parent R20291 strain (Figure 3.5B). This elevated intracellular level of c-di- GMP in *sinRR'* mutants can block the *sigD* expression, which in turn will result in reduced motility and toxin production (Figure 3.4C and Figure 3.5B). Hence, when *sigD* was expressed from the tetracycline-inducible promoter (which is not affected by c-di-GMP concentration), motility and toxin production in the *sinRR'* mutant could be revived. These two

findings corroborate our conclusion that elevated levels of c-di-GMP in *sinRR'* mutant plays a major role in controlling its toxin production and motility. We are currently performing experiments to test whether SinR can directly regulate *dccA* in *C. difficile*.

3.6.9 Inactivation of SinR' results in hyper-sporulation, higher toxin production, and motility than the parent strain

Results from the *sinR* and *sinR'* complementation experiments showed that expression of SinR alone could revive the toxin production and the motility in the R20291::*sinRR'* strain, whereas SinR' expression alone did not have any effect on the toxin production or the motility (Figure 3.4C and Figure 3.5A). These results suggested that among SinR and SinR', only SinR can directly influence the toxin production and the motility, which raised the question on the role of SinR' in these pathways. To find the answer, we created a *sinR'* mutant which expressed SinR in the absence of SinR'(Figure S 3.9). Our repeated attempts to create a *sinR'* mutant using the similar technique in the JIR8094 background failed for unknown reasons. Mutation in *sinR'* was confirmed by PCR (Figure S 3.9) and western blot analysis using SinR' specific antibodies.

As expected the SinR' mutant produced SinR protein, but not the SinR' (Figure S 3.9). The R20291::*sinR'* grew almost similar to the parent strain and did not show any profound autolysis phenotype as the R20291:: *sinRR'* (Figure S 3.6). We performed the assays to measure sporulation, motility and toxin production in the R20291::*sinR'*. In the sporulation assay, it was found that R20291::*sinR'* produced nearly three-fold more spores than the parent R20291 strain (Figure 3.6A). The R20291::*sinR'* was more motile than the R20291 strain (Figure 3.6B). Similarly, a 2.5-fold increase in the toxin produc- tion was observed in the R20291::*sinR'* when compared to the parent strain (Figure 3.6C). These initial results revealed that SinR' can negatively influence

sporulation, toxin production, and motility. In our complementation of R20291::*sinRR'* we showed that presence of SinR' alone in the *C. difficile* cells in the absence of SinR could not influence either toxin production or the motility (Figure 3.4C and Figure 3.5A)

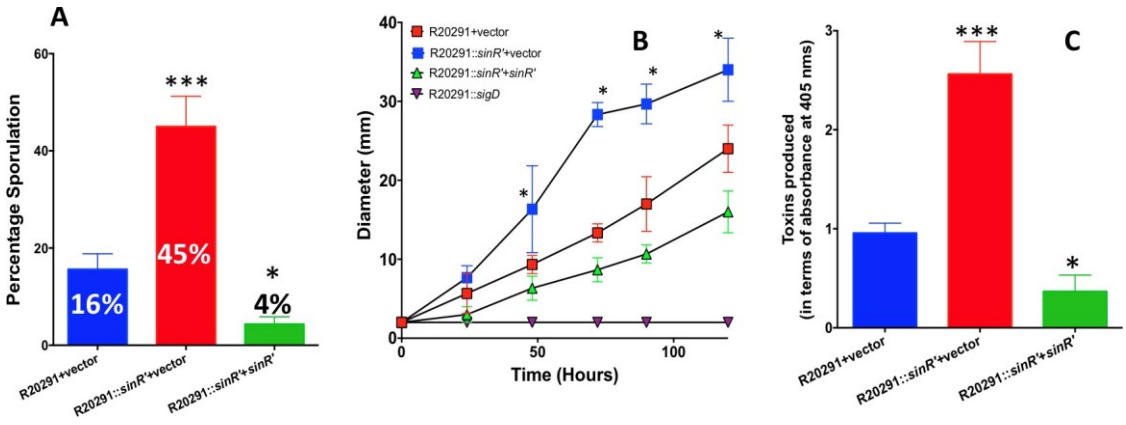


Figure 3.6 Characterization of *C. difficile* R20291::*sinR'*.

(A) *C. difficile* cultures were grown in 70:30 medium for 30 h under anaerobic conditions and Sporulation frequency (CFU/ml of ethanol resistant spores) of R20291, *sinR'* mutant was determined. The data shown are means ± standard errors of three biological replicates. (B) Motility assays of the *C. difficile* R20291, sinR' mutant and complemented *sinR'* mutant. The experiments were repeated at least three times independently (*, P 0.05 by a two-tailed Student's t-test). (C) Toxin production measured by ELISA. Statistical analysis was performed using one way-ANOVA with Dunnett's multiple comparisons test comparing values to the average of the parent with vector control (***<0.0005, *< 0.05 p-value).

Hence, SinR' must be influencing these pathways through its action on SinR. For example, if SinR' is an inhibitor of SinR then the absence of SinR' in the R20291::*sinR'* would result in increased SinR activity, which in turn may result in increased sporulation, toxin production and motility in this strain. To test this hypothesis, we performed two experiments. First, tested the effect of over-expressed SinR in the wild-type strain; Second, we checked for physical interaction of SinR with SinR' proteins by performing pull-down experiments.

3.6.10 Overexpression of SinR in the wildtype strain R20291 results in hyper-sporulation and increased the toxin production and motility

The plasmid construct with either *sinR* (pRG300) or *sinR'* (pRG306) under tetracycline-inducible promoter were introduced into R20291 parent strain and were tested for their toxin production, sporulation, and motility upon induction with ATc. The R20291 strain with the vector alone was used as the control in these assays. To perform the sporulation assay, we used bacterial cultures grown in 70:30 medium supplemented with 50 ng/ml of ATc for 36 hours. Sporulation efficiency was enumerated as described in the method section. Overexpression of *sinR* in R20291 strain increased its sporulation efficiency 2.5-fold (45%) when compared to the control strain, where the average sporulation efficiency was 18%. Overproduction of SinR' in R20291, however, reduced the sporulation efficiency to 5% (Figure 3.7A).

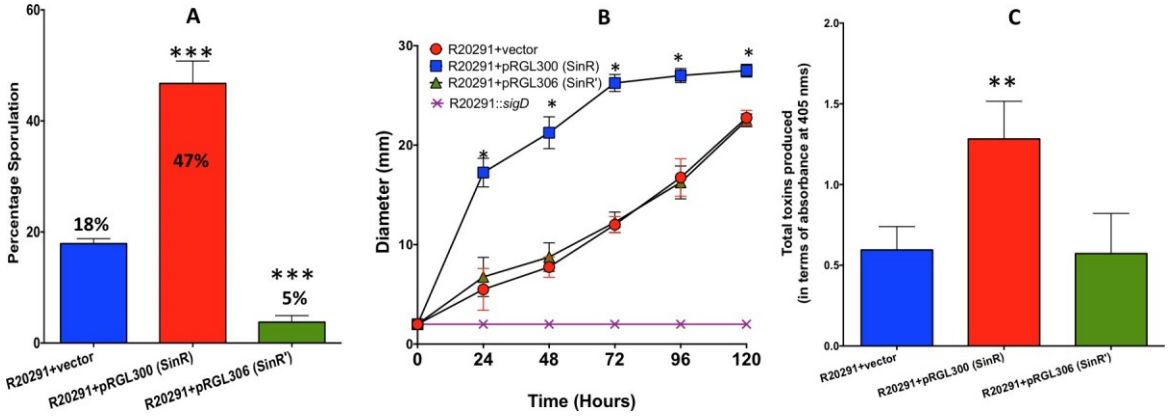


Figure 3.7 Effect of sinR or sinR' overexpression in the R20291 strain.

The *sinR* or the *sinR'* gene was cloned under tetracycline-inducible promoter and the resulting plasmid constructs were introduced into wildtype (WT) R20291 strain for overexpression. (A) Toxin ELISA, (B) Motility assay (C) Sporulation frequency. The data shown are means ± standard errors of three biological replicates. Statistical analysis was performed using one way-ANOVA with Dunnett's multiple comparisons test comparing values to the average of the parent with vector control (***<0.0005, *< 0.05 p-value).

Overproduction of SinR in R20291 resulted in increased motility as well (Figure 3.7B). In *C. difficile*, toxin production is minimal during exponential phase (~4 to 8h) of the bacterial culture and reaches its maximum during the stationary phase (12h -16h) [9]. To detect any positive influence of both SinR and SinR' on toxin production in the parent strain, we chose to use the 8h time point. The bacterial cultures were grown for 6h in TY medium and were induced with 50 ng/ml of ATc for two hours before harvesting their cytosolic protein for Toxin ELISA. Results from these experi- ments showed that overexpression of *sinR* resulted in a nearly 2.5-fold increase in the toxin production in the R20291 strain when compared to the R20291 with vector alone control (Figure 3.7C). No significant effect on toxin production was observed when *sinR'* was overexpressed in R20291 (Figure 3.7C). This could be because *sin* locus is expressed only during the early stationary phase (10-12h) in *C. difficile* (Figure S 3.2). We performed toxin ELISA at 8h time-point when SinR is predicted to be lower in the bacterial cells. If SinR' acts on toxin production primarily by repressing SinR, then overexpression of SinR' at this time-point will not have any effect on toxin production. Nevertheless, results from this overexpression studies demonstrated that increased SinR content in *C. difficile* could result in increased toxin production, motility, and sporulation.

3.6.11 SinR' interacts with SinR

In *B. subtilis*, SinR monomers bind with each other to form a homotetramer, which would then bind to upstream sequences of the target genes to repress their expression [34, 52]. SinI in

B. subtilis binds with SinR and prevents the SinR homotetramer formation and thus blocks its activity [52]. To test the protein-protein interaction of *C. difficile* SinR with SinR', we performed GST pull-down experiments using SinR-6His and SinR'-GST. Purified SinR-6His pro-

tein was mixed with crude lysates from *E. coli* expressing SinR'-GST. When we passed this mixture through the Ni^{++} affinity chromatography column, we pulled out SinR-6His along with SinR'-GST, suggesting the tight association of SinR with SinR' (Figure 3.8A, lanes 5, 7). In control, the GST alone did not interact with the SinR-6His (Figure 3.8A, lanes 6, 8), confirming protein specific interaction between SinR with SinR'. These results provided compelling evidence that SinR' affects toxin production and sporulation indirectly by binding with SinR to inhibit its activity on its target genes.

3.6.12 SinR binds to codY promoter region

Transcriptome analysis of the R20291::*sinRR'* showed up-regulation of *codY*, an important global regulator by ~3 to 30 fold compared to parent strains (Table S 3.5, Table S 3.8). CodY is highly conserved in many Gram-positive bacteria [53–55]. In *B. subtilis* it regulates several metabolic genes and controls competence, sporulation, and motility [56–58]. In *C. difficile*, the *codY* mutant produced more toxins and spores than the parent strains and thus it is a repressor of these pathways [14, 17, 18]. We hypothesized that many phenotypes and transcriptional changes we observe in the *sinRR'* mutant could be related to the up-regulation of *codY* in these mutant strains. To investigate whether SinR and SinR' or both controls *codY* expression by binding to the promoter region of *codY*, we carried out electrophoretic mobility shift assays (EMSAs). We used radiolabeled DNA probe that contained the putative promoter region of the *codY* gene and performed binding reactions using purified SinR-6His or SinR'-6His proteins. First, we tested SinR alone at increasing concentrations and found that it can shift the probe when used above 100 nM concentration (Figure 3.7B).

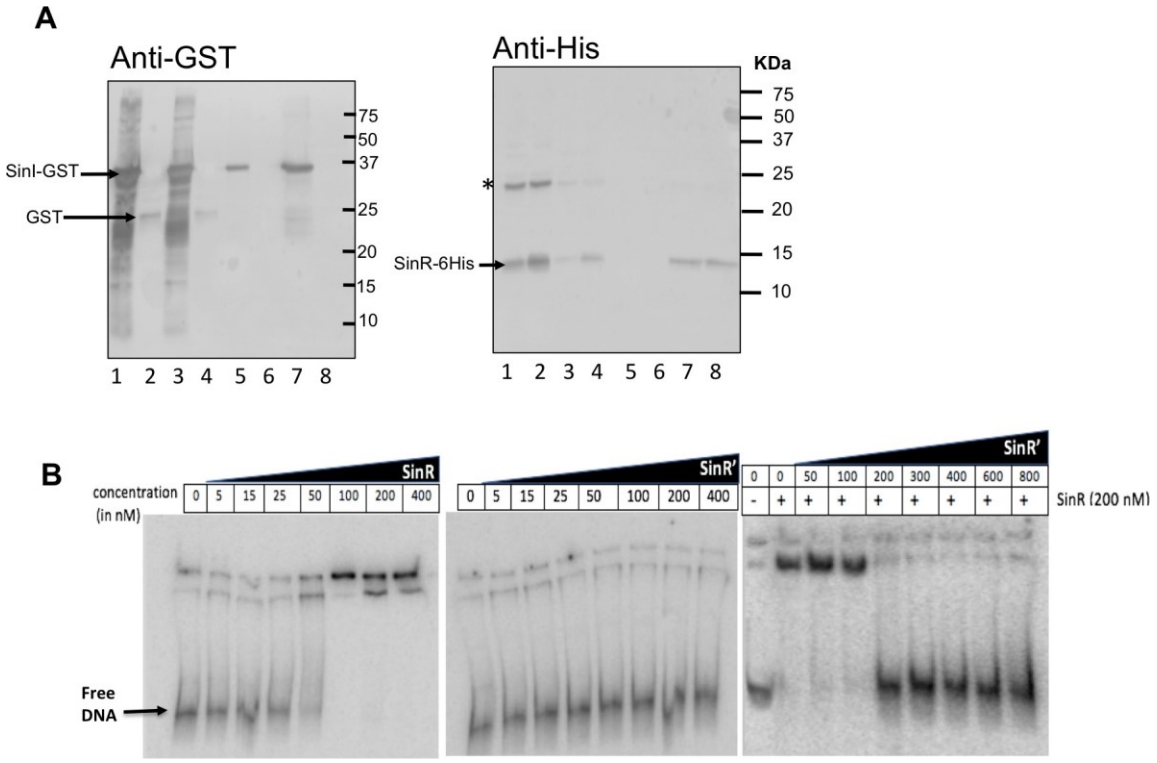


Figure 3.8 SinR' interacts with SinR.

(A) In vitro, protein-protein interactions indicate that SinR' binds tightly to SinR. GST-tagged SinR' protein was incubated with SinR-6His proteins and purified using Ni++ agarose affinity columns. The elutes were probed with anti- GST and with anti-His antibodies. Lanes details are as follows: Input 1: Mixture of SinR'-GST expressing E.coli lysate with purified SinR-6His. 1.Input 2: Mixture of GST expressing E. coli lysate with purified SinR-6His. 2. Unbound from input 1 after passing through Ni++ column. 3. Unbound from input 2 after passing through Ni++ column. 4. Elute with 50 mm imidazole (SinR'-GST + SinR-6His). 5. Elute with 50 mM imidazole (GST+SinR-6His) 6. Elute with 200 mM imidazole (SinR'-GST + SinR-6His). 7. Elute with 200 mM imidazole (GST+SinR-6His). * indicates SinR-His dimer. (B) Interactions of SinR with codY promoter region. EMSA analysis of SinR- 6His, SinR'-6His, a mixture of SinR'-6His and SinR-His binding to codY probe.

When SinR' was used similarly, it was unable to cause the mobility shift of the probe, even at the highest concentration (Figure 3.7B). We then tested whether SinR' would prevent SinR from binding to the *codY* promoter region. To do this, we used increasing amounts of SinR', in the presence of a fixed amount of SinR (Figure 3.7B). The results show that the presence of SinR' in the reaction mix could prevent SinR from binding to the DNA. As a negative control, we used a DNA probe that contained the promoter region of *gluD*, which codes for glutamate dehydrogenase

(GDH). Neither SinR nor SinR' was able to shift the control DNA even at the highest concentrations tested (Figure S 3.10). Based on these results, we conclude that SinR binds specifically to *codY* promoter region to control its transcription. This result also provided evidence that the SinR' interaction with SinR prevents its regulatory activity on its target gene.

3.6.13 CodY regulates sin locus expression

In a recent study, CodY was found to negatively regulate *sinRR'* expression in the *C. difficile* 630Δ*erm* strain [14]. A CodY putative binding site was identified in the *sin* locus upstream sequence, and reporter fusions with the *sin* locus promoter revealed the CodY could negatively regulate *sin* locus expression in this strain. However, in the UK1 strain (belongs to the ribotype 027 as R20291), the promoter fusion revealed a positive regulation of *sin* locus by CodY. Because of these contradictory observations, one could not conclude whether CodY regulates *sin* locus. To examine the role of CodY on *sin* locus expression, we performed EMSA with purified CodY-6His and the putative CodY binding region upstream of *sin* locus. An oligonucleotide with putative CodY binding sequence upstream of *sinR* was synthesized (ORG 721) (Table S 3.2) and was radioactively labeled with [γ- ^{32} P] dATP. A double-stranded DNA probe was generated after annealing with the complementary oligonucleotide (ORG722). It is worth noting no sequence difference was found within this putative *sin* promoter regions of the UK1, R20291, JIR8094 and 630Δ*erm* genomes. We also generated probes with a known CodY binding sequence upstream of the *tcdR* gene (using ORG719 and ORG720) and with non-specific sequence (ORG702 and ORG723) as positive and negative controls respectively. EMSA was performed by incubating the radioactively labeled probes with varying concentrations of purified CodY-6His. We found that CodY could bind to the sequence upstream of *sin* locus at the concentration of 400 nM (Figure

3.9A). As expected the shift was observed with the positive control probe, while no shift could be observed with the non-specific DNA probe even with high protein concentrations (Figure 3.9A). Binding of CodY to its targets most of the time results in repres- sion of their transcription [17, 18, 58]. However, there are few targets where CodY was found to promote transcription [58]. To check whether CodY has any positive influence on *sin* locus expression in UK1 strain as reported [14], we performed western blot analysis and looked for SinR and SinR' in UK1 strain and its *codY* mutant (UK1::*codY*). Results showed that SinR and SinR' protein content in the UK1::*codY* mutant was higher than in the UK1 parent strain (Figure 3.9B). Our data demonstrate that CodY has a negative impact on SinR and SinR' production in this strain. Since our repeated attempts to create *codY* mutants in R20291 and JIR8094 strains failed, we could not include them in this analysis. Nonetheless, our results from the EMSA and the western blot analyses corroborate the negative regulation of the *sin* locus by CodY.

3.6.14 R20291::sinRR' is less virulent in hamster

Since the *C. difficile sin* locus was found to be important for the regulation of many important pathways under *in vitro* growth conditions, we wanted to determine its significance in *C. diffi- cile* pathogenesis. We used the hamster model in which *C. difficile* infection is known to cause severe disease signs [59, 60]. Syrian hamsters were gavaged with 2,000 vegetative cells of *C. dif- ficile* strain R20291 or with R20291::*sinRR'* and monitored for *C. difficile* infection. Fecal pellets were collected daily until animals developed diarrheal symptoms. All ten animals infected with parental strain R20291 succumbed to the disease within five days after bacterial challenge.

Two of the ten animals infected with R20291::*sinRR'* exhibited disease symptoms within two days after challenge (Figure 3.10A). Diseased hamsters were sacrificed (see M&M), and their cecal

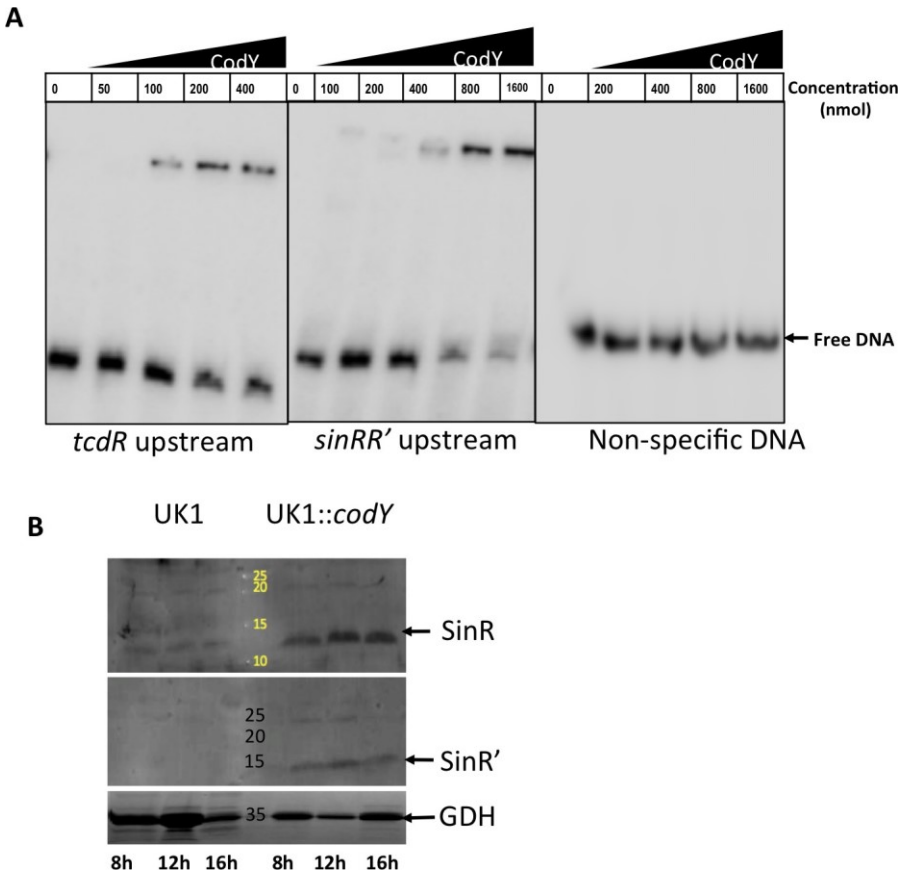


Figure 3.9 CodY controls the sin locus expression.

(A) CodY-6His binding to sin locus promoter region. The *tcdR* upstream and a non-specific DNA probe was as positive and negative controls respectively. (B) Western blot analysis of UK1 and UK1::*codY* mutants to detect SinR and SinR' proteins.

contents were collected for toxin ELISA and CFU count. All surviving *sinRR'* mutant infected hamsters (8 in total) and uninfected control hamsters were also sacrificed fifteen days post-infection, and their cecal contents were also tested for toxins and *C. difficile* cells. Toxins could be detected (Figure S 3.11) in the cecal contents of all the diseased hamsters (10 from R20291 group

and two from R20291::*sinRR'* group), which confirmed the occurrence of CDI in them. However, toxins could not be detected in the eight hamsters that survived the R20291::*sinRR'* challenge. The cecal contents of R20291 infected hamsters contained nearly 10^7 colony-form- ing units per gram. No *C. difficile* could be recovered from the cecal contents of any of the R20291::*sinRR'* challenged animals, including of the two hamsters that came down with CDI in this group (Figure 3.10B). If the *sinRR'* mutant lyses *in vivo*, as we observed in *in vitro* growth conditions, it could explain why we could not recover any *C. difficile* cells but could detect tox- ins in the cecal contents of that two hamsters that came down with the disease after *sinRR'* mutant challenge. Since nearly 80% of the animals survived the R20291::*sinRR'* challenge, we conclude that members of the SinRR' regulon are needed for *C. difficile* successful pathogenesis.

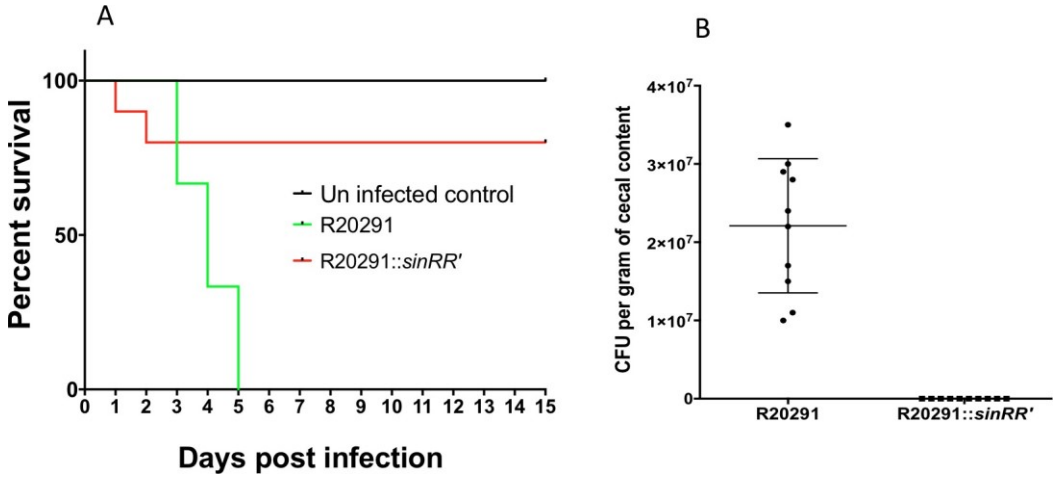


Figure 3.10 Disrupting *sinRR'* decreases morbidity in hamster models of *C. difficile* infection.

(A) Kaplan-Meier survival curve of clindamycin-treated Syrian golden hamsters inoculated with 2,000 vegetative cells of C. difficile R20291 (n = 10) or sinRR' mutant (n = 10). Six animals were used as an uninfected control. Animals were monitored every four hours for the symptoms of lethargy, poor fur coat, wet tail or hunched posture. Moribund animals were euthanized and log-rank statistical analysis was performed; $p<0.001$. (B) Total number of C. difficile colony forming units (CFU) /gm of cecal contents recovered postmortem.

3.7 Discussion

This study aims to decipher the role of the SinRR' regulators in *C. difficile* physiology. In *C. difficile*, there has been no data explaining their function, except for a few expression analyses, where mutations in *sigH*, *tcdR*, *codY*, *spo0A*, *opp*, *app* were found to affect the expression of the *sin* locus [13, 14, 21, 22, 60]. Initial clues about the role of SinR and SinR' in sporulation came from the work performed by Saujet *et al*. where they showed increased expression of *sinR* in the asporogenous *sigH* mutant, suggesting it to be a negative regulator of sporulation as in the case of *B. subtilis* [22]. However, *sinR* was found to be up-regulated in the hyper-sporulating oligopeptide transporter *opp-app* mutant and was down-regulated in the hypo-sporulation *tcdR* mutant [13, 60]. These later studies suggested the positive influence of SinR on sporulation. In this work, we mutated the *sin* locus in two different *C. difficile* strains and conclusively showed that unlike *B. subtilis* SinR, which inhibits sporulation, *C. difficile* SinR has a positive effect on sporulation.

Transcriptome analysis of *sinRR'* mutants revealed that in addition to sporulation, genes involved in motility, transport, stress response, cell wall biogenesis, and various metabolic pathways were also affected. It is worth noting that *cynT*, the gene adjacent to *sin* locus (Figure 3.1B), is one among the many metabolic genes that were found to be down-regulated in the *sinRR'* mutants (Table S 3.4 and Table S 3.6). The analysis also revealed that the *sin* locus mutations could affect the transcription of many important regulators, including *codY*, *sigD*, *spo0A*, and *tcdR*. This observation compelled us to hypothesize that SinRR' might be indirectly influencing transcription of many of these genes by controlling their regulators. For example, changing in the transcription of *codY*, a global regulator can affect the gene regulatory circuits of various pathways.

CodY is known to be a sensor of the metabolic state of the cell. During the exponential growth phase, when the nutrients are abundant, CodY binds to branched-chain amino acids

(BCAAs), and GTP and acts primarily as a repressor of various alternative metabolic pathways [17, 18]. When nutrients become limited in the cell, CodY is no longer bound by the cofactors and the transcriptional repression by CodY is alleviated on its targets. In *C. difficile*, CodY controls toxin production and sporulation in addition to metabolic pathways [14, 17, 18]. The transcription of *codY* was found to be up-regulated in the R20291::*sinRR'* (Table S 3.5), (Figure 3.11). This observation of increased *codY* transcription in the asporogenic *sinRR' C. difficile* mutant is consistent with the recent findings that a *C. difficile codY* mutant hyper- sporulates [14].

To test whether increased CodY activity in the mutant is the reason for its lower toxin produc- tion and sporulation, we tried to isolate a *sinRR'-codY* double mutant and were unsuccessful even after several attempts. However, our EMSA experiments with purified SinR and *codY* upstream DNA showed that the SinR could specifically bind to this region, possibly to repress its transcription. We have also shown that purified CodY, in turn, can bind with the upstream region of *sin* locus upstream region to control its expression. Since the *sin* locus codes for both *sinR* and its antagonist *sinR'*, SinR repression on *codY* would be moderate when compared to CodY's repression on the *sin* locus. Also, when the cells enter the stationary phase, CodY repression on the *sin* locus may be alleviated in the absence of its co-substrates and will result in the *sin locus* expression, which we found to be essential for sporulation initiation. We performed dot blot analysis with cytosolic proteins of R20291 and R20291::*sinRR'* and determined that CodY in R20291::*sinRR'* was only moderately higher than R20291 (Figure S 3.12). This could be due to the cell to cell variation in gene expression within the test population. For example, only 18% of the R20291 population enters sporulation in the growth conditions we tested. In *C. difficile*, only cells with low or inactive CodY enter sporulation. If we consider sporulation as an indirect measure for inactive CodY in a bacterial cell, we can say that the CodY production or activity was affected

only in a fraction of cells in the parent population. To overcome this issue, we compared the CodY content in R20291::*sinR'* cells (which produce more SinR) with R20291:: *sinRR'*. It is worth to note that nearly 50% of R20291::*sinR'* culture enters sporulation. Nearly two- fold more CodY could be detected in R20291::*sinRR'* cells when compared to R20291::*sinR'* cells. Other than modulating CodY content in *C. difficile*, SinR could also affect the CodY activity indirectly by affecting the concentrations of CodY substrates (BCAA and GTP). The transcriptome analysis indeed showed numerous metabolic genes to be affected in the *sinRR'* mutant. In the JIR8094::*sinRR'* mutant, *codY* was not among the differentially regulated genes. How- ever, in this strain *ccpA* was up-regulated nearly 13.5-fold (Table S 3.7). Similar to CodY, CcpA also represses toxin gene expression in *C. difficile* [15, 16]. Thus lower toxin production in JIR8094:: *sinRR'* could be due to the higher CcpA activity in this mutant (Figure 3.11). We are currently testing whether *ccpA* is directly regulated by SinRR'. We are also setting up experiments to check whether increased CcpA has any role in controlling *codY* expression in the JIR8094::*sinRR'* strain. SigD is one other regulator whose expression was found to be affected in the *sinRR'* mutants. In *C. difficile*, the *sigD* expression is repressed by elevated levels of c-di-GMP [50]. The enzyme, diguanylate cyclase coded by *dccA* synthesize c-di-GMP from GTP. In this study, we have shown the expression of *dccA* is up-regulated in *sinRR'* mutant (Table S 3.5 and Table 3.2) and the observation of three-fold higher intracellular concentration of c-di-GMP in the *sinRR'* mutant, corroborated the transcriptome data. These results suggest that SinR and SinR' regu- lates motility and toxin production indirectly by regulating the c-di-GMP production. Another scenario that can result in higher intracellular c-di-GMP concentration is when c-di-GMP degrading phosphodiesterases are reduced within the cell.

In *C. difficile*, *pdcA* codes for a c-di- GMP phosphodiesterases, and it was recently identified to be repressed by CodY [61].

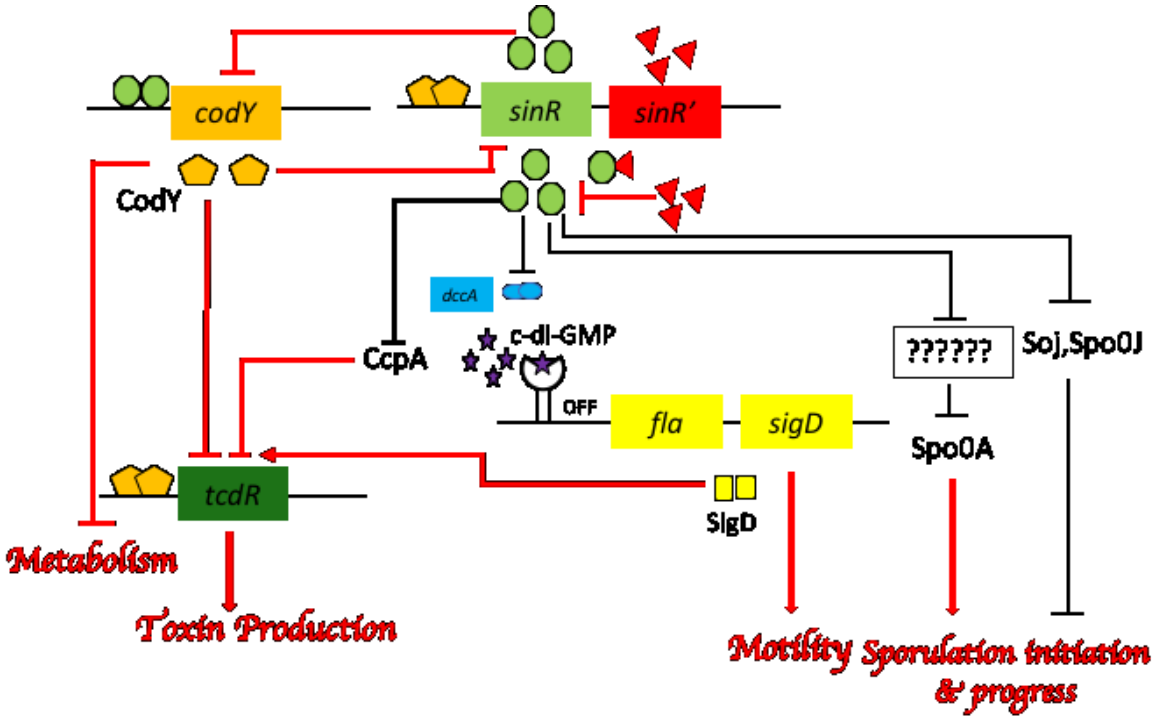


Figure 3.11 Proposed regulatory circuit of sin operon

Schematic diagram showing sin locus regulation of genes involved in toxin production, sporulation, and motility in *C. difficile*. Known genetic interactions are marked in red and the predicted interactions are marked in black.

RNA-- Seq analysis did not identify *pdcA* as one among the differentially regulated genes in R20291:: *sinRR'* strain. However, it was under-expressed nearly 4-fold in *JIR8094*::*sinRR'* (Table S 3.6) mutant. Increased CodY activity in the *sinRR'* mutant could indirectly result in increased c-di- GMP concentration, which in turn can suppress toxin production and motility.

In *B. subtilis*, the SinR's repressor's activity on its target genes is inhibited by SinI, which is coded in the same operon (Figure 3.1A). In *B. subtilis* the polycistronic *sinRI* transcripts are pro-

duced from two upstream promoters. The monocistronic *sinR* transcripts are driven from a promoter located within the coding region of *sinI*. Regulating the transcription rate of *sinRI* and *sinR* helps *B. subtilis* to control its SinR and SinI content. Our RT-PCR and QRT-PCR analysis detected *sinRR'* transcripts in *C. difficile*. We have also shown that disrupting *sinR* by insertion mutagenesis affects both *sinR* and *sinR'* transcription. These results suggest that *sinRR'* is transcribe as a bicistronic message. However, there is a possibility that *sinR'* may have an independent promoter within *sinR* coding sequence as in *B. subtilis*. Our QRT-PCR analysis repeatedly detected lower levels of *sinR'*, *sinRR'* transcripts than the *sinR* transcripts. Western blot analysis also revealed lower levels of SinR' than the SinR in growth conditions tested (Figure 3.2B-lane 2 and Figure S 3.9C). There is a possibility that mRNA degradation from the 3' end can result in lower levels of *sinR'* transcripts, which in turn can result in lower levels of SinR' than SinR. We did not detect any secondary structures upstream of *sinR'* that can influence its translation rate or translation initiation. We, however, noted that the RBS of *sinR'* are just two nucleotides away from the *sinR* stop codon. Ribosome complex occupying the *sinR* stop codon can prevent the assembly of new ribosome complex at the *sinR'* RBS to initiate translation.

Since SinR' has a DNA binding domain, it is also possible that SinR' may work as direct regulator independently from SinR and may have its own targets for regulation. In such case, SinR' may not always be available to inhibit SinR function. A transcriptome analysis of *sinR'* mutant and its comparison with *sinRR'* transcriptome may help us to identify direct targets of SinR'.

In *B. subtilis*, other than SinI, SinR also interacts with SlrR and SlrA to regulate genes involved in matrix formation (the *eps* and *tap-sipW-tas* operon), autolysis (*lytABC*) and motility (hag, encoding flagellin) [29, 31]. In *B. subtilis*, the SlrR is a DNA binding protein, and it is homologous to SinR. Conversely, SlrA is a small protein devoid of any DNA binding domains and is

homologous to SinI [38, 39]. While SlrR can form heterodimers with SinR to repress *lytABC* and *hag* expression, it can also inhibit SinR's repression activity on *eps* and *tap- sipW-tas* operons [38, 39, 62] which are needed for biofilm formation. The *C. difficile sin* locus codes for two DNA binding proteins SinR and SinR' and their interactions resembles the inter- action between *B. subtilis* SinR-SlrR. Similar to *B. subtilis* SlrR, *C. difficile* SinR' carries a DNA binding domain and it will be interesting to analyze whether SinR-SinR' complexes together are needed for the repression of any genes. It is important to note that the autolysis phenotype of *sinRR'* mutant was complemented only when both SinR and SinR' were expressed (Figure 3.2C). This suggests that like *B. subtilis* SinR-SlrR complex, SinR-SinR' complex together repress autolysis in *C. difficile*. No *lytABC* homologs could be identified in *C. difficile* genome, and the precise reason for autolysis in *sinRR'* mutant is not clear yet. However, the RNA-Seq data revealed that nearly 6% of the differentially expressed genes in *sinRR'* mutant plays a role in cell wall synthesis or assembly. This highlights that SinR and SinR' play an important regulatory role in this pathway.

Among the phenotypes tested, asporogenesis of the *sinRR'* mutant was the only one we could not complement. Even the expression of *spo0A* failed to initiate sporulation in this mutant. Transcripts of Spo0A~P activated *sigE* and *sigF* did not show any increase when *spo0A* was expressed in the *sinRR'* mutant, suggesting the Spo0A remain unphosphorylated and inactive. We are currently performing additional experiments to test this hypothesis.

Another regulatory checkpoint for sporulation initiation is chromosomal DNA replication and segregation. This is achieved through the action of Soj and Spo0J in *B. subtilis*, where they repress sporulation until chromosomal segregation has occurred. They block the *spo0A* depen- dent transcription in *B. subtilis* [63]. The *spo0J* and *soj* homologs in *C. difficile* are CD630_3671 and CD630_3672, respectively in an operon, which also carries CD630_3673, an additional Spo0J-

like orthologue. In both JIR8094::*sinRR'* and R20291::*sinRR'*, all three genes were up-reg- ulated ~3 fold. Hence, the inactivation of Spo0A could result partly because of the up-regulation of the *soj* operon in the *sinRR'* mutants (Figure 3.11). But the function of *soj* and *spo0J* in *C. difficile* should be determined before we can speculate their roles in asporogenesis of *sinRR'* mutants.

BLAST search revealed that SinRR' to be unique to *C. difficile* and its close relative *Clostridium sordellii*. The *sin* locus is absent in other *Clostridia*. Even though sporulation-specific sigma factors appear to be conserved among *Clostridia*, recent studies have suggested that sporulation initiation and regulation of *C. difficile* to be distinct [64, 65]. Since the *sin* locus appears to play a significant role in sporulation initiation and regulation, it is reasonable to speculate its presence could be one of the reasons why the regulation of sporulation initiation is distinct in *C. difficile*.

In summary, our study supports earlier reports that in *C. difficile*, virulence, sporulation, metabolism and motility pathways are inter-connected [13–24]. While many regulators in this network are yet to be identified, here we present the evidence that SinRR' play a central role in this regulatory network. SinR regulates multiple pathways by controlling other global regulators. Finding genes that are directly under SinR regulation may lead to the identification of new regulatory genes and gene products that are important for *C. difficile* pathogenesis.

3.8 Acknowledgments

We thank following investigators for sharing their lab resources: David Ho, Rockefeller University, for FliC antibody; Wiep Klass, Leids University and Aimee Shen, Tufts University for Spo0A antibody; Linc Sonenshein, Tufts University for CodY antibody and UK::*codY* mutant; Nigel Minton, University of Nottingham, for the plasmid pMTL007C-E5; Robert Fagan

for the vector pRPF185. We thank Jose E. Lopez for technical assistance throughout the study and Yusuf Ceftci, Varun Govind for editing the manuscript.

3.9 References

1. Curry SR. *Clostridium difficile*. Clin Lab Med. 2017;37(2):341-69. doi: 10.1016/j.cll.2017.01.007. PMID: 28457354.

2. Napolitano LM, Edmiston CE, Jr. *Clostridium difficile disease*: Diagnosis, pathogenesis, and treatment update. Surgery. 2017. doi: 10.1016/j.surg.2017.01.018. PMID: 28267992.

3. CDC. Antibiotic resistance threats in the United States, 2013. 2013; Available from:. http://wwwcdcgov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508pdf.

4. Carter GP, Rood JI, Lyras D. The role of toxin A and toxin B in Clostridium difficile-associated disease: Past and present perspectives. Gut Microbes. 2010;1(1):58-64. doi: 10.4161/gmic.1.1.10768. PMID: 20664812.

5. Kuehne SA, Cartman ST, Heap JT, Kelly ML, Cockayne A, Minton NP. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature. 2010;467(7316):711-3. doi: 10.1038/nature09397. PMID: 20844489.

6. Kuehne SA, Collery MM, Kelly ML, Cartman ST, Cockayne A, Minton NP. Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. J Infect Dis. 2014;209(1):83-6. doi: 10.1093/infdis/jit426. PMID: 23935202.

7. Lyras D, O'Connor JR, Howarth PM, Sambol SP, Carter GP, Phumoonna T, et al. Toxin B is essential for virulence of *Clostridium difficile*. Nature. 2009;458(7242):1176-9. doi: 10.1038/nature07822. PMID: 19252482.

8. Deakin LJ, Clare S, Fagan RP, Dawson LF, Pickard DJ, West MR, et al. The *Clostridium difficile* spo0A gene is a persistence and transmission factor. Infect Immun. 2012;80(8):2704-11. doi: 10.1128/IAI.00147-12. PMID: 22615253.

9. Dupuy B, Sonenshein AL. Regulated transcription of *Clostridium difficile* toxin genes. Mol Microbiol. 1998;27(1):107-20. PMID: 9466260.

10. Mani N, Dupuy B. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci U S A. 2001;98(10):5844-9. doi: 10.1073/pnas.101126598. PMID: 11320220.

11. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, et al. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J Bacteriol. 2002;184(21):5971-8. PMID: 12374831.

12. Karlsson S, Dupuy B, Mukherjee K, Norin E, Burman LG, Akerlund T. Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. Infect Immun. 2003;71(4):1784-93. PMID: 12654792.

13. Edwards AN, Nawrocki KL, McBride SM. Conserved oligopeptide permeases modulate sporulation initiation in *Clostridium difficile*. Infect Immun. 2014;82(10):4276-91. doi: 10.1128/IAI.02323-14. PMID: 25069979.

14. Nawrocki KL, Edwards AN, Daou N, Bouillaut L, McBride SM. CodY-Dependent Regulation of Sporulation in *Clostridium difficile*. J Bacteriol. 2016;198(15):2113-30. doi: 10.1128/JB.00220-16. PMID: 27246573.

15. Antunes A, Camiade E, Monot M, Courtois E, Barbut F, Sernova NV, et al. Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. Nucleic Acids Res. 2012;40(21):10701-18. doi: 10.1093/nar/gks864. PMID: 22989714.

16. Antunes A, Martin-Verstraete I, Dupuy B. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Mol Microbiol. 2011;79(4):882-99. doi: 10.1111/j.1365-2958.2010.07495.x. PMID: 21299645.

17. Dineen SS, McBride SM, Sonenshein AL. Integration of metabolism and virulence by *Clostridium difficile* CodY. J Bacteriol. 2010;192(20):5350-62. doi: 10.1128/JB.00341-10. PMID: 20709897.

18. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL. Repression of *Clostridium difficile* toxin gene expression by CodY. Mol Microbiol. 2007;66(1):206-19. doi: 10.1111/j.1365-2958.2007.05906.x. PMID: 17725558.

19. Edwards AN, Tamayo R, McBride SM. A novel regulator controls *Clostridium difficile* sporulation, motility and toxin production. Mol Microbiol. 2016;100(6):954-71. doi: 10.1111/mmi.13361. PMID: 26915493.

20. Mackin KE, Carter GP, Howarth P, Rood JI, Lyras D. Spo0A differentially regulates toxin production in evolutionarily diverse strains of *Clostridium difficile*. PLoS One. 2013;8(11):e79666. doi: 10.1371/journal.pone.0079666. PMID: 24236153.

21. Pettit LJ, Browne HP, Yu L, Smits WK, Fagan RP, Barquist L, et al. Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. BMC Genomics. 2014;15:160. doi: 10.1186/1471-2164-15-160. PMID: 24568651.

22. Saujet L, Monot M, Dupuy B, Soutourina O, Martin-Verstraete I. The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. J Bacteriol. 2011;193(13):3186-96. doi: 10.1128/JB.00272-11. PMID: 21572003.

23. Anjuwon-Foster BR, Tamayo R. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet. 2017;13(3):e1006701. doi: 10.1371/journal.pgen.1006701. PMID: 28346491.

24. Aubry A, Hussack G, Chen W, KuoLee R, Twine SM, Fulton KM, et al. Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. Infect Immun. 2012;80(10):3521-32. doi: 10.1128/IAI.00224-12. PMID: 22851750.

25. El Meouche I, Peltier J, Monot M, Soutourina O, Pestel-Caron M, Dupuy B, et al. Characterization of the SigD regulon of *C. difficile* and its positive control of toxin production through the regulation of tcdR. PLoS One. 2013;8(12):e83748. doi: 10.1371/journal.pone.0083748. PMID: 24358307.

26. McKee RW, Mangalea MR, Purcell EB, Borchardt EK, Tamayo R. The second messenger cyclic Di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD. J Bacteriol. 2013;195(22):5174-85. doi: 10.1128/JB.00501-13. PMID: 24039264.

27. Barilla D, Caramori T, Galizzi A. Coupling of flagellin gene transcription to flagellar assembly in *Bacillus subtilis*. J Bacteriol. 1994;176(15):4558-64. PMID: 8045886.

28. Kuroda A, Sekiguchi J. High-level transcription of the major *Bacillus subtilis* autolysin operon depends on expression of the sigma D gene and is affected by a sin (flaD) mutation. J Bacteriol. 1993;175(3):795-801. PMID: 8093697.

29. Chu F, Kearns DB, McLoon A, Chai Y, Kolter R, Losick R. A novel regulatory protein governing biofilm formation in *Bacillus subtilis*. Mol Microbiol. 2008;68(5):1117-27. doi: 10.1111/j.1365-2958.2008.06201.x. PMID: 18430133.

30. Kodgire P, Dixit M, Rao KK. ScoC and SinR negatively regulate epr by corepression in *Bacillus subtilis*. J Bacteriol. 2006;188(17):6425-8. doi: 10.1128/JB.00427-06. PMID: 16923912.

31. Chu F, Kearns DB, Branda SS, Kolter R, Losick R. Targets of the master regulator of biofilm formation in *Bacillus subtilis*. Mol Microbiol. 2006;59(4):1216-28. doi: 10.1111/j.1365-2958.2005.05019.x. PMID: 16430695.

32. Gaur NK, Cabane K, Smith I. Structure and expression of the *Bacillus subtilis sin* operon. J Bacteriol. 1988;170(3):1046-53. PMID: 3125149.

33. Voigt CA, Wolf DM, Arkin AP. The *Bacillus subtilis sin* operon: an evolvable network motif. Genetics. 2005;169(3):1187-202. doi: 10.1534/genetics.104.031955. PMID: 15466432.

34. Draughn GL, Bobay BG, Stowe SD, Olson AL, Feldmann EA, Thompson RJ, et al. Solution structures of biofilm-controlling proteins SinI and SinR from *Bacillus subtilis* reveal details of DNA-binding and regulatory mechanism. The FASEB Journal. 2017;31(1 Supplement):907.2-.2.

35. Pflughoeft KJ, Sumby P, Koehler TM. *Bacillus anthracis* sin locus and regulation of secreted proteases. J Bacteriol. 2011;193(3):631-9. doi: 10.1128/JB.01083-10. PMID: 21131488.

36. O'Connor JR, Lyras D, Farrow KA, Adams V, Powell DR, Hinds J, et al. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol. 2006;61(5):1335-51. doi: 10.1111/j.1365-2958.2006.05315.x. PMID: 16925561.

37. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102. doi: 10.1186/gb-2009-10-9-r102. PMID: 19781061.

38. Chai Y, Norman T, Kolter R, Losick R. An epigenetic switch governing daughter cell separation in *Bacillus subtilis*. Genes Dev. 2010;24(8):754-65. doi: 10.1101/gad.1915010. PMID: 20351052.

39. Chai Y, Norman T, Kolter R, Losick R. Evidence that metabolism and chromosome copy number control mutually exclusive cell fates in *Bacillus subtilis*. EMBO J. 2011;30(7):1402-13. doi: 10.1038/emboj.2011.36. PMID: 21326214.

40. Fimlaid KA, Bond JP, Schutz KC, Putnam EE, Leung JM, Lawley TD, et al. Global analysis of the sporulation pathway of *Clostridium difficile*. PLoS Genet. 2013;9(8):e1003660. doi: 10.1371/journal.pgen.1003660. PMID: 23950727.

41. Rosenbusch KE, Bakker D, Kuijper EJ, Smits WK. *C. difficile* 630Deltaerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. PLoS One. 2012;7(10):e48608. doi: 10.1371/journal.pone.0048608. PMID: 23119071.

42. Olmedo G, Ninfa EG, Stock J, Youngman P. Novel mutations that alter the regulation of sporulation in *Bacillus subtilis*. Evidence that phosphorylation of regulatory protein SpoOA controls the initiation of sporulation. J Mol Biol. 1990;215(3):359-72. PMID: 2121995.

43. Burbulys D, Trach KA, Hoch JA. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. Cell. 1991;64(3):545-52. PMID: 1846779.

44. Jiang M, Grau R, Perego M. Differential processing of propeptide inhibitors of Rap phosphatases in *Bacillus subtilis*. J Bacteriol. 2000;182(2):303-10. PMID: 10629174.

45. LeDeaux JR, Grossman AD. Isolation and characterization of *kinC*, a gene that encodes a sensor kinase homologous to the sporulation sensor kinases KinA and KinB in *Bacillus subtilis*. J Bacteriol. 1995;177(1):166-75. PMID: 8002614.

46. Perego M, Cole SP, Burbulys D, Trach K, Hoch JA. Characterization of the gene for a protein kinase which phosphorylates the sporulation-regulatory proteins Spo0A and Spo0F of *Bacillus subtilis*. J Bacteriol. 1989;171(11):6187-96. PMID: 2509430.

47. Trach K, Burbulys D, Strauch M, Wu JJ, Dhillon N, Jonas R, et al. Control of the initiation of sporulation in *Bacillus subtilis* by a phosphorelay. Res Microbiol. 1991;142(7-8):815-23. PMID: 1664534.

48. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, et al. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J Bacteriol. 2009;191(23):7296-305. doi: 10.1128/JB.00882-09. PMID: 19783633.

49. Collery MM, Kuehne SA, McBride SM, Kelly ML, Monot M, Cockayne A, et al. What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and derivatives. Virulence. 2016:1-15. doi: 10.1080/21505594.2016.1237333. PMID: 27652799.

50. Purcell EB, McKee RW, McBride SM, Waters CM, Tamayo R. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. J Bacteriol. 2012;194(13):3307-16. doi: 10.1128/JB.00100-12. PMID: 22522894.

51. Bordeleau E, Fortier LC, Malouin F, Burrus V. c-di-GMP turn-over in *Clostridium difficile* is controlled by a plethora of diguanylate cyclases and phosphodiesterases. PLoS Genet. 2011;7(3):e1002039. doi: 10.1371/journal.pgen.1002039. PMID: 21483756.

52. Newman JA, Rodrigues C, Lewis RJ. Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in *Bacillus subtilis*. J Biol Chem. 2013;288(15):10766-78. doi: 10.1074/jbc.M113.455592. PMID: 23430750.

53. Bennett HJ, Pearce DM, Glenn S, Taylor CM, Kuhn M, Sonenshein AL, et al. Characterization of *relA* and *codY* mutants of *Listeria monocytogenes*: identification of the CodY regulon and its role in virulence. Mol Microbiol. 2007;63(5):1453-67. doi: 10.1111/j.1365-2958.2007.05597.x. PMID: 17302820.

54. Chateau A, van Schaik W, Joseph P, Handke LD, McBride SM, Smeets FM, et al. Identification of CodY targets in *Bacillus anthracis* by genome-wide *in vitro* binding analysis. J Bacteriol. 2013;195(6):1204-13. doi: 10.1128/JB.02041-12. PMID: 23292769.

55. Majerczyk CD, Dunman PM, Luong TT, Lee CY, Sadykov MR, Somerville GA, et al. Direct targets of CodY in *Staphylococcus aureus*. J Bacteriol. 2010;192(11):2861-77. doi: 10.1128/JB.00220-10. PMID: 20363936.

56. Handke LD, Shivers RP, Sonenshein AL. Interaction of *Bacillus subtilis* CodY with GTP. J Bacteriol. 2008;190(3):798-806. doi: 10.1128/JB.01115-07. PMID: 17993518.

57. Ratnayake-Lecamwasam M, Serror P, Wong KW, Sonenshein AL. *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. Genes Dev. 2001;15(9):1093-103. doi: 10.1101/gad.874201. PMID: 11331605.

58. Sonenshein AL. CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr Opin Microbiol. 2005;8(2):203-7. doi: 10.1016/j.mib.2005.01.001. PMID: 15802253.

59. Girinathan BP, Braun S, Sirigireddy AR, Lopez JE, Govind R. Importance of Glutamate Dehydrogenase (GDH) in *Clostridium difficile* Colonization In Vivo. PLoS One. 2016;11(7):e0160107. doi: 10.1371/journal.pone.0160107. PMID: 27467167.

60. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, et al. Effect of *tcdR* Mutation on Sporulation in the Epidemic *Clostridium difficile* Strain R20291. mSphere. 2017;2(1). doi: 10.1128/mSphere.00383-16. PMID: 28217744.

61. Purcell EB, McKee RW, Courson DS, Garrett EM, McBride SM, Cheney RE, et al. A Nutrient-Regulated Cyclic Diguanylate Phosphodiesterase Controls *Clostridium difficile* Biofilm and Toxin Production during Stationary Phase. Infect Immun. 2017;85(9). doi: 10.1128/IAI.00347-17. PMID: 28652311.

62. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R. Sticking together: building a biofilm the *Bacillus subtilis* way. Nat Rev Microbiol. 2013;11(3):157-68. doi: 10.1038/nrmicro2960. PMID: 23353768.

63. Ireton K, Gunther NWt, Grossman AD. spo0J is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. J Bacteriol. 1994;176(17):5320-9. PMID: 8071208.

64. Pereira FC, Saujet L, Tome AR, Serrano M, Monot M, Couture-Tosi E, et al. The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. PLoS Genet. 2013;9(10):e1003782. doi: 10.1371/journal.pgen.1003782. PMID: 24098139.

65. Saujet L, Pereira FC, Serrano M, Soutourina O, Monot M, Shelyakin PV, et al. Genome-wide analysis of cell type-specific gene transcription during spore formation in *Clostridium difficile*. PLoS Genet. 2013;9(10):e1003756. doi: 10.1371/journal.pgen.1003756. PMID: 24098137.

66. Girinathan BP, Braun SE, Govind R. *Clostridium difficile* glutamate dehydrogenase is a secreted enzyme that confers resistance to H2O2. Microbiology. 2014;160(Pt 1):47-55. doi: 10.1099/mic.0.071365-0. PMID: 24145018.

67. Putnam EE, Nock AM, Lawley TD, Shen A. SpoIVA and SipL are *Clostridium difficile* spore morphogenetic proteins. J Bacteriol. 2013;195(6):1214-25. doi: 10.1128/JB.02181-12. PMID: 23292781.

68. Teng F, Murray BE, Weinstock GM. Conjugal transfer of plasmid DNA from *Escherichia coli* to enterococci: a method to make insertion mutations. Plasmid. 1998;39(3):182-6. doi: 10.1006/plas.1998.1336. PMID: 9571134.

69. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010;80(1):49-55. doi: 10.1016/j.mimet.2009.10.018. PMID: 19891996.

70. Govind R, Dupuy B. Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. PLoS Pathog. 2012;8(6):e1002727. doi: 10.1371/journal.ppat.1002727. PMID: 22685398.

71. Criscuolo A, Brisse S. AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. Genomics. 2013;102(5-6):500-6. doi: 10.1016/j.ygeno.2013.07.011. PMID: 23912058.

72. Soutourina OA, Monot M, Boudry P, Saujet L, Pichon C, Sismeiro O, et al. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet. 2013;9(5):e1003493. doi: 10.1371/journal.pgen.1003493. PMID: 23675309.

73. Govind R, Fitzwater L, Nichols R. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. J Bacteriol. 2015;197(15):2600-9. doi: 10.1128/JB.00224-15. PMID: 26013487.

74. Sheffield P, Garrard S, Derewenda Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr Purif. 1999;15(1):34-9. doi: 10.1006/prep.1998.1003. PMID: 10024467.

75. Sambol SP, Tang JK, Merrigan MM, Johnson S, Gerding DN. Infection of hamsters with epidemiologically important strains of *Clostridium difficile*. J Infect Dis. 2001;183(12):1760-6. doi: 10.1086/320736. PMID: 11372028.

76. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol. 2010;192(19):4983-90. doi: 10.1128/JB.00610-10. PMID: 20675492.

77. Mooyottu S, Kollanoor-Johny A, Flock G, Bouillaut L, Upadhyay A, Sonenshein AL, et al. Carvacrol and trans-cinnamaldehyde reduce *Clostridium difficile* toxin production and cytotoxicity *in vitro*. Int J Mol Sci. 2014;15(3):4415-30. doi: 10.3390/ijms15034415. PMID: 24625665.

78. Fagan RP, Fairweather NF. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem. 2011;286(31):27483-93. doi: 10.1074/jbc.M111.263889. PMID: 21659510.

79. Sheffield P, Garrard S, Derewenda Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr Purif. 1999;15(1):34-9. doi: 10.1006/prep.1998.1003. PMID: 10024467.

3.10 Supplemental Data included in this file:

S1 Text. Plasmids construction.
S2 Text. Supplemental methods: I. Quantitative reverse transcription PCR (qRT-PCR) II. Quantification of c-di-GMP in *C. difficile* by High Performance Liquid Chromatography (HPLC)

S1 Table. Bacterial strains and plasmids used in this study.
S2 Table. Oligonucleotides used for PCR reactions**.**
S3 Table. Oligonucleotides used for QRT-PCR reactions**.**
S4 Table. Under-expressed genes in R20291::*sinRR'* compared to R20291.
S5 Table. Over-expressed genes in R20291::*sinRR'* compared to R20291.
S6 Table. Under-expressed genes in JIR8094::*sinRR'* compared to JIR8094.
S7 Table. Over-expressed genes in JIR8094::*sinRR'* compared to JIR8094.
S8 Table. QRT-PCR analysis of selected genes in *sinRR'* mutants.

S1 Figure. Construction and confirmation of the *sinR* mutant in *C. difficile* JIR8094 and R20291.
S2 Figure. Evidence for the read-through transcription of *sin* locus in *C. difficile*.
S3 Figure. Characterizing *sinRR'* mutants.
S4 Figure. Autolysis assay.
S5 Figure. Analysis of sporulation in JIR8094::*sinRR'* mutant.
S6 Figure. Motility analysis of *sinRR'* mutant.
S7 Figure. Toxin production in JIR8094::*sinRR'* mutant.
S8 Figure. Quantification of intracellular c-di-GMP by HPLC.
S9 Figure. Construction and characterization of the *sinR'* mutant in *C. difficile* R20291.
S10 Figure. Gel mobility shift assay reveals neither SinR nor SinR' binds to *gluD* upstream (non-specific control DNA).
S11 Figure. Toxin ELISA to detect *C. difficile* toxins in cecal contents of infected hamsters.
S12 Figure. Dot blot analysis of R20291, R20291::*sinRR'* and R20291::*sinR'* cytosolic proteins using CodY specific antibody.

3.10.1 S1 Supplemental text. Plasmids construction

pRG291: A 670bp fragment containing *sigD* was amplified using primers org536 and org531 and cloned as *BamH1/Sac1* into pRPF185.

pRG300: A 339bp fragment containing *sinR* was amplified using primers org549 and org550 and cloned as *BamH1/Sac1* into pRPF185.

pRG306: A 315bp fragment containing *sinR'* was amplified using primers org553 and org554 and cloned as *BamH1/Sac1* into pRPF185.

pRG307: A 1025bp fragment containing region upstream of *sinR* with *sinR* was amplified using primers org555 and org556 and cloned into the MCS of pGEM-T

pRG308: A 1364bp fragment containing region upstream of *sinR* with *sinRR'* was amplified using primers org555 and org557 and cloned into the MCS of pGEM-T

pRG309: A 1465bp fragment containing region upstream of *spo0A* with *spo0A* was amplified using primers org559 and org560 and cloned into the MCS of pGEM-T

pRG310: The 1025bp upstream *sinR* with *sinR* sequence from pRG307 was cloned as *EcoR1/Kpn1* into pMTL84151.

pRG311: The 1364bp upstream *sinR* with *sinRR'* sequence from pRG308 was cloned as *EcoR1/Kpn1* into pMTL84151.

pRG312: The 1465bp upstream *spo0A* with *spo0A* sequence from pRG309 was cloned as *EcoR1/Kpn1* into pMTL84151.

pRG324: A 363bp fragment containing *sinR* with C term His tag was amplified using primers org582 and org583 and cloned into the MCS of pGEM-T

pRG325: A 339bp fragment containing *sinR'* with C term His tag was amplified using primers org584 and org585 and cloned into the MCS of pGEM-T.

pRG327: The 339bp *sinR'* sequence with 6X His tag from pRG325 was cloned as *Xho1/BamH1* into pET16B for purification.

pRG329: The 363bp *sinR* sequence with 6X His tag from pRG324 was cloned as *Xho1/BamH1* into pET16B for purification.

pRG330: A 315bp fragment containing *sinR'* was amplified using primers org619 and org620 and cloned into the MCS of pGEM-T.

pRG331: A 315bp *sinR'* sequence from pRG330 was cloned as *Nco1/Sal1* into GST parallel II yielding *sinR'* in-frame with GST tag for GST pull down experiment.

pRG334: A 654bp fragment containing *sinRR'* was amplified using primers org549 and org554 and cloned as *BamH1/Sac1* into pRPF185.

pRG359: The 786bp *codY* sequence with 6X His tag was amplified using primers org 675 and org676 and cloned as *Xho1/BamH1* into pET16B for purification.

pMTL007-CE5:Cdi-*sin*R-141a, pMTL007-CE5:Cdi-*sin*R'-129s, pMTL007-CE5:Cdi-*sigD*: Introns specific for *sinR*, *sinR'* and *sigD* were designed following the Perutka algorithm at ClosTron.com. The introns were synthesized and cloned inbetween *BsrGI* and *HindIII* sites in pMTL007-CE5 vector.

3.10.2 S2 Supplemental methods

3.10.2.1 Quantitative reverse transcription PCR (qRT-PCR)

C. difficile cultures were grown in TY medium for 12h and were used for total RNA extraction following a previously described protocol [22, 60, 72]. After DNase (Turbo; Ambion) treatment, a 30 µL reaction was set up with 5 µg of template RNA, 4 µL of deoxynucleoside triphosphates (dNTP; 10mM each), 1 µg of hexamer oligonucleotide primer (5 µg/µL pdN_6; Roche), and 6 µL of reverse transcription (RT) buffer and was heated at 80 °C for 5 min and cDNA was synthesized at 42 °C for 2 hours using avian myeloblastosis virus (AMV) reverse transcriptase (Promega). Final 20 µL reaction volume containing 10 ng of cDNA, 400 nM gene-specific primers, and 12.75 µL of SYBR PCR master mix (BioRad) was used to perform Real-time quantitative PCR using iQPCR real-time PCR instrument (BioRad). Quantity of cDNA of a gene in each sample was normalized to the quantity of *C. difficile* 16S rRNA gene and the ratio of normalized target concentrations (threshold cycle [$2^{-\Delta\Delta Ct}$] method) [22, 60, 72] gives the relative change in gene expression. A minimum of three biological replicates were used per sample.

3.10.2.2 Quantification of c-di-GMP in C. difficile by High Performance Liquid Chromatography (HPLC).

Nucleotides were extracted and quantified from *C. difficile* as previously described with some modifications [61]. Briefly, cells from 50ml of early stationary phase cultures of R20291and R20291::*sinRR'* strains were harvested by centrifugation at 3000g for 20 min. The optical densities of the cultures at the time of harvest was recorded and dilutions were plated to determine the number of CFU extracted. The bacterial cell pellet was

washed in 1ml of TE buffer (10mM Tris (pH7.5), 1mM EDTA, pH8) and vortexed in 500 µl of nucleotide extraction buffer (40% acetonitrile 40% methanol in 0.1N formic acid)) and incubated at −20°C for 30 min. Samples were centrifuged for 5 min at 17,000 × *g* in 4°C, and the supernatant was immediately neutralized by adding 4 µl of 15% (wt/vol) NH_4HCO_3 per 100 ul of sample. One in ten dilution of the samples were made with deionized water and 50 µl of the sample aliquots was injected into high performance liquid chromatography (HPLC, Schimadzu Corp. Nakagyo-ku, Kyoto, Japan) and the analytes were separated on C18 reverse phase HPLC column (4.6 by 250 mm; 5 µm) with the following gradient of solvent A (10 mM ammonium acetate in water) to solvent B (0.1% TFA v/v acetonitrile). The following elution method was used: 0 to 1 min, isocratic hold at 0% (of eluent B); 1 to 2 min, 0 to 20% linear gradient; 2 to 4 min, isocratic hold at 20%; 4 to 5 min, 20 to 40% linear gradient; 5 to 7 min, isocratic hold at 40%; 7 to 8, 40 to 70% linear gradient; 8 to 10 min, isocratic hold at 70%; 10 to 11, 70 to 90% linear gradient; 11 to 13 min, isocratic hold at 90%; 13 to 14 min, 90 to 0% linear gradient; 14 to 16 min, 0% isocratic hold to re-equilibrate the column for the next sample injection (16 min total). The c-di-GMP was determined by fitting the peak area to a linear standard curve obtained by analyzing serial dilutions (250 nM, 125 nM, 62.5 nM, 31.25 nM, 15.625 nM, and 7.8125 nM) of standard c-di-GMP (Sigma-Aldrich, Inc. St. Louis, MO, USA) using the same HPLC method. The intracellular levels of c-di-GMP was calculated by dividing the total amount of c-di-GMP extracted by the total intracellular volume of bacteria extracted. The later was determined by multiplying the number of cells extracted, based on CFU counts by the volume of one bacterial cell. To measure the volume of a single bacterial cell, average cell length and width were measured from the electron microscopic images of the R20291

and R20291::*sinRR'* mutant cells (at least fifty independent cells were measured) and were estimated to be 3.46×10^{-16} L using ImageJ software.

3.10.3 Supplementary tables

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

Bacterial strain or plasmid	Relevant features or genotype	Source or reference
Clostridium difficile JIR8094	Erm^s derivative of strain 630	O'Connor *et al.* (2006)
Clostridium difficile R20291	Clinical isolate - NAP1/027 ribotype, isolated in 2006 following an outbreak in Stoke Mandeville Hospital, UK	Stabler *et al.* (2009)
Escherichia coli DH5α	*endA1 recA1 deoR hsdR17* (r_K^- m_K^+)	NEB
Escherichia coli S17-1	Strain with integrated RP4 conjugation transfer function; favors conjugation between *E. coli* and *C. difficile*	Teng *et al.* (1998)
Clostridium difficile JIR8094::*sinR*R"	JIR8094 with intron insertion within *sinR*	This study
Clostridium difficile R20291::*sinR*R'	R20291with intron insertion within *sinR*	This study
Clostridium difficile R20291::*sinR'*:	R20291with intron insertion within *sinR'*	This study
Clostridium difficile UK1	Clinical isolate	Sorg *et al*., (2010)
Clostridium difficile UK1::codY		Mooyottu *et al*. (2014)
R20291::sigD	R20291with intron insertion within *sigD*	This study
pMTL007-CE5	ClosTron plasmid	Heap *et al.* (2010)
pMTL007-CE5:Cdi-*sin*R-141a	pMTL007-CE5 with group II intron targeted to *sinR*	This study
pMTL007-CE5:Cdi-*sin*R'-a	pMTL007-CE5 with group II intron targeted to *sinR'*	This study
pMTL007-CE5:Cdi-*sigD*	pMTL007-CE5 with group II intron targeted to *sigD*	This study
pRPF185	*E. coli*/*C. difficile* shuttle plasmid	Fagan *et al*. (2011)
pRG300	pRPF185 containing *sinR* under inducible *tet* promoter	This study
pRG334	pRPF185 containing *sinRR'* under inducible tet promoter	This study
pRG306	pRPF185 containing *sinR'* under inducible *tet* promoter	This study

pRG291	pRPF185 containing *sigD* under inducible *tet* promoter	This study
pRG310	pMTL84151 containing 759bp upstream *sinR* with *sinR* gene (*sinR* under own promoter)	This study
pRG311	pMTL84151 containing 759bp upstream *sinR* with *sinRR'* gene (*sinRR'* under own promoter)	This study
pRG312	pMTL84151 containing 300bp upstream *spo0A* with *spo0A* gene	This study
pRG329	pET16B containing *sinR* gene with His tag	This study
pRG327	pET16B containing *sinR'* gene with His tag	This study
pRG359	pET16B containing *codY* gene with His tag	This study
pGST parallel II	GST parallel II vector for GST fusions	Sheffield *et al.* (1999)
pRG331	GST parallel II containing *sinR'* with GST tag	This study

References:

1. O'Connor JR, Lyras D, Farrow KA, Adams V, Powell DR, Hinds J, et al. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol. 2006;61(5):1335-51. doi: 10.1111/j.1365-2958.2006.05315.x. PMID: 16925561.

2. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, et al. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol. 2009;10(9):R102. doi: 10.1186/gb-2009-10-9-r102. PMID: 19781061.

3. Teng F, Murray BE, Weinstock GM. Conjugal transfer of plasmid DNA from *Escherichia coli* to enterococci: a method to make insertion mutations. Plasmid. 1998;39(3):182-6. doi: 10.1006/plas.1998.1336. PMID: 9571134.

4. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol. 2010;192(19):4983-90. doi: 10.1128/JB.00610-10. PMID: 20675492.

5. Mooyottu S, Kollanoor-Johny A, Flock G, Bouillaut L, Upadhyay A, Sonenshein AL, et al. Carvacrol and trans-cinnamaldehyde reduce *Clostridium difficile* toxin production and cytotoxicity *in vitro*. Int J Mol Sci. 2014;15(3):4415-30. doi: 10.3390/ijms15034415. PMID: 24625665.

6. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, et al. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010;80(1):49-55. doi: 10.1016/j.mimet.2009.10.018. PMID: 19891996.

7. Fagan RP, Fairweather NF. *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem. 2011;286(31):27483-93. doi: 10.1074/jbc.M111.263889. PMID: 21659510.

8. Sheffield P, Garrard S, Derewenda Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr Purif. 1999;15(1):34-9. doi: 10.1006/prep.1998.1003. PMID: 10024467.

Table S 3.2 Oligonucleotides used for PCR reactions.

Name	Sequence (5' → 3')	Description
EBS-U	CGA AAT TAG AAA CTT GCG TTC AGT AAAC	Group II intron specific primer
ORG522	CTAAAGTGGAGGGATAATAATTGGCAAATA TAGG	*sinR*-Forward
ORG523	TTATATTTCTTTGTATTTAATGATGCAC	*sinR*-Reverse
ORG549	GTTAACAGATCTGAGCTCCTAAAGTGGAGG GATAATAATTG	*sinR*-Forward with SacI-pRPF185 (pRG300) cloning
ORG550	AAGTTTTATTAAAACTTATAGGATCCTTATAT TTCTTTGTATTTAATGATGCAC	*sinR*-Reverse with BamH1-pRPF185 (pRG300) cloning
ORG551	GTTAACAGATCTGAGCTCCTGTAATAAGAA GATGTTTTTTAATGG	*spo0A*-Forward with SacI-pRPF185 (pRG301) cloning
ORG552	AAGTTTTATTAAAACTTATAGGATCCTTATTT AACCATACTATGTTCTAGTC	*spo0A*-Reverse with BamH1-pRPF185 (pRG301) cloning
ORG536	AGCGTTAACAGATCTGAGCTCGGAGGCGTA GTTAATGAATAG	*sigD*-Forward with SacI-pRPF185 (pRG291) cloning
ORG531	AAGTTTTATTAAAACTTATAGGATCCCTATAT AGAATATTTAAGTTCTTTTATCTTGTTTC	*sigD*-Reverse with BamH1-pRPF185 (pRG291) cloning
ORG553	GTTAACAGATCTGAGCTCGTGAGGGAAATA GTAACAATAATGAATTATATAG	*sigR'* Forward with SacI-pRPF185 (pRG306) cloning
ORG554	AAGTTTTATTAAAACTTATAGGATCCTTATAT TTTATTCTTTTTTATGATGTCTATAATC	*sigR'* Reverse with BamH1 -pRPF185 (pRG306) cloning
ORG555	CTTCTTATTTTTATGGTACCATGTAATATCAC CCTCTTTAAAAATTTTTTTATTATTATATC	*sinR*-Forward with *Kpn1* –pMTL84151 (pRG310 cloning)
ORG556	GGGCATCGAAATAAAAAACTAGTTTATATTT CTTTGTATTTAATGATGCAC	*sinR*-Reverse with *EcoR1*-pMTL84151 (pRG310 cloning)
ORG557	GGGCATCGAAATAAAAAACTAGTTTATATTT TATTCTTTTTTATGATGTCTATAATCTG	*sinR'*-Reverse with *EcoR1*-pMTL84151 (pRG311cloning)
ORG559	CTTCTTATTTTTATGGTACCGGTGCAATAAC TCATGTTTTTAG	*spo0A* upstream + *spo0A*-Forward with *Kpn1*– pMTL84151 (pRG312) cloning
ORG560	GGGCATCGAAATAAAAAACTAGTGACTCTC ATATTTAAACCTCCAC	*spo0A* reverse with *EcoR1*–pMTL84151 (pRG312 cloning)
ORG582	CTCGAGTTGGCAAATATAGGAAAAATAATA GG	*sinR* forward with *XhoI* to clone in pET16B
ORG583	GGATCCTTAGTGGTGATGGTGATGATGTAT TTCTTTGTATTTAATGATGCAC	*sinR* reverse with *BamH1* to clone in pET16B
ORG584	CTCGAGATGAATTATATAGGTAAAAGAC	*sinR'* forward with *XhoI* to clone in pET16B
ORG585	GGATCCTTAGTGGTGATGGTGATGATGTAT TTTATTCTTTTTTATGATGTC	*sinR'* reverse with *BamH1* to clone in pET16B
ORG-19	CGATACGACCGAAAACCTGTATTTTCAGGG CGCCATGGGGATGAATTATATAGG	*sinR'* forward GST with *NcoI* to clone in GST-parallel-II vector
ORG620	GCGGCCGCACTAGTTGAGCTCGTCGACTTA TATTTTATTCTTTTTTATGATGT	*sinR'* reverse GST with *SalI* to clone in GST-parallel-II vector
ORG629	TGGAATAAGCCAACAGGAGAGTTTTGTGAG	*codY* upstream-forward; for SinR binding experiment
ORG630	GTAATCTTCCATGTTCTTCAGGAAAGATTT	*codY* upstream- revers; for SinR binding experiment
ORG72	TAAAAAATAAACTGAGAAAATGATATACTAA TTT	*gluD* upstream-Forward; for SinR binding (control)
ORG73	TATAAATACGTTATAATTATGTATACTCCATT	*gluD* upstream-Reverse; for SinR binding (control)
ORG721	TAAAATAGAAAATTT TTTTAATTTTCAAAATATATTCTACATATCTA ATATGTAATTAC	*sin* locus upstream- with potential CodY binding sequence
ORG722	GTAATTACATATTAGATATGTAGAATATATTT TGAAAATTAAAAAAATTTTCTATTTTA	*sin* locus upstream- with potential CodY binding sequence (complementary to ORG721)

Name	Sequence (5' → 3')	Description
ORG719	TAGTTATAACTTCAAAAAAGACTGAAAATTA AGAAAAAAGAAATATAAAT	*tcdR* upstream- with CodY binding sequence
ORG720	ATTTATATTTCTTTTTTCTTAATTTTCAGTCTT TTTTGAAGTTATAACTA	*tcdR* upstream-with CodY binding sequence-complementary to ORG719
ORG702	GTAATATATCCGATTTTAGCATATGCTAAAA TATCATACATCAAATTTTTAACTACTTAC	Non- specific DNA sequence used for CodY binding (non-specific control)
ORG723	GTAAGTAGTTAAAAATTTGATGTATGATATT TTAGCATATGCTAAAATCGGATATATTAC	Complementary to ORG702 used to generate dsDNA probe (non-specific control)

Table S 3.3 Oligonucleotides used for QRT-PCR reactions.

Primer	Sequence (5' → 3')	Gene target
RG-RT23 (F)	GAGGAGAGTGGAATTCCTAGTGTAG	*16srRNA*
RG-RT24 (R)	GGACTACCAGGGTATCTAATCCTGT	*16srRNA*
RG-RT25 (F)	AGGCAGGTTTACATCCAACATA	*sinR*
RG-RT26 (R)	AGTGGTATGTCTAAAGCAGTAGC	*sinR*
RG-RT27 (F)	AAAGACTTAAAGAAGAACGGAAAA	*sinR'*
RG-RT28 (R)	TTGGATTCTTTTTACCACTTTCG	*sinR'*
RG-RT29 (F)	AAGTTAATATTTTTGAAGATGAGGATG	CD630_2216
RG-RT30 (R)	TGGGATACTTATTTTCACTTAACCAA	CD630_2216
RG-RT33 (F)	CATGAAATAGGAGTACCAGCTCA	*spo0A*
RG-RT34 (R)	CTCCATGCAACCTCTATTGC	*spo0A*
RG-RT35 (F)	TGCTCATGTATTATGGCTGGTATTT	*murG*
RG-RT36 (R)	AGGTTTGGAGCAATCAAGAAAGA	*murG*
RG-RT37 (F)	TGACTTTACACTTTCATCTGTTTCTAGC	*sigE*
RG-RT38 (R)	GGGCAAATATACTTCCTCCTCCAT	*sigE*
RG-RT39 (F)	TGTTTATAGATACTGGGCTCTCTGG	*spoIID*
RG-RT40 (R)	GACTTGCTCTGTGTTAGTTCCATC	*spoIID*
RG-RT41 (F)	CGCTCCTAACTAGACCTAAATTGC	*sigF*
RG-RT42 (R)	GGAAGTAACTGTTGCCAGAGAAGA	*sigF*
RG-RT43 (F)	CCTCAGTGAAATCATCATCATAGTCTTTA	*gpr*
RG-RT44 (R)	CCTGGTAATTGGTCTTGGAAATAGA	*gpr*
RG-RT45 (F)	CAAACTGTTGTCTGGCTTCTTC	*sigG*
RG-RT46 (R)	GTGGTGTTAATACATCAGAACTTCC	*sigG*
RG-RT49 (F)	CATATGTTGCTAATCGAGTTCCTTTAT	*sigK*
RG-RT50 (R)	TCAACGGAAGATCAGGATGATTTA	*sigK*
sleB-RT (F)	GATATTGTAGAGAACCCCTAATCC	*sleB*
sleB-RT (R)	GCAAATCCTAAAGCTAAAAATAC	sleB
Gpr-RT (F)	GGTGTTACTATTAAGTTCTTGTCAT	*gpr*
Gpr-RT (R)	CTGGTGGAGGTGTTGGCAATACTAG	*gpr*
sspA-RT (F)	CTATCTGTTGCTTTTTCCAGCC	*sspA*
sspA-RT (R)	GTATGAGTAATTATCAACAAGTTG	*sspA*
spoVAC RT(F)	GTAGACCAAATAAGCCCAAAACC	*spoVAC*
spoVAC-(R)	CAGAACTAGCACCTAGTTTATC	*spoVAC*

Primer	Sequence (5' → 3')	Gene target
spoVAD-(F)	GTGGCGATTTAATAAATCAAATAG	*spoVAD*
spoVAD- (R)	CACTTCCTGCTCCTGTAACTGTCC	*spoVAD*
cdeC-RT (F)	GATGAAATAAATTCAGAAGACATGA	*cdeC*
cdeC-RT (R)	GGCACTGCATTTGATACAGAGAAG	*cdeC*
sleC-RT (F)	CTGTTCCATAGATACCATCTTC	*sleC*
sleC-RT (R)	GGGCAGTAAAGACTTAGGTGACC	*sleC*
cotCB-RT (F)	GGTACAGAGGAAATGGCTCATGTTG	*cotCB*
cotCB-RT (R)	CTTGTAGTAAAGTTTACTCCATTAG	*cotCB*
cotE-RT (F)	GAATATTGATAAAGCATCATCATATG	*cotE*
cotE-RT (R)	GCCATAAGAGATGTTATAGGGGATG	*cotE*
cotB-RT (F)	GATTTTATCTTACACTGTTCTATTCC	*cotB*
cotB-RT (R)	GGACCATATTATGATGGAACATGCTC	*cotB*
cotA-RT (F)	CTTACCTAGAACTTCAACACCAGTTA	*cotA*
cotA-RT (R)	CATTGTGTAATCTTAAAGCTGTTGC	*cotA*
pdaA-RT (F)	CATCTAATATCTCAGTATTTGTTG	*pdaA*
pdaA-RT (R)	GCGAACAATCTTTAAAATATACACAA	*pdaA*
RG-RT3 (F)	TAATAAAAATACTGCCCTCGACAAA	*tcdA*
RG-RT4 (R)	ATAAATTGCATGTTGCTTCATAACT	*tcdA*
RG-RT7 (F)	CTGGACAATGGAAGGTGGTT	*tcdB*
RG-RT8 (R)	TTGATGGTGCTGAAAAGAAGTG	*tcdB*
RG-RT5 (F)	AACATCTTGGAATATCTGAATTTTTCTCTA	*tcdE*
RG-RT6 (R)	TCTGTCATTGCATCTAGTAAAATTGCT	*tcdE*
RG-RT1 (F)	CAAGAAATAACTCAGTAGATGATTTGCAA	*tcdR*
RG-RT2 (R)	TCTCCCTCTTCATAATGTAAAACTCTACTA	*tcdR*
codY-RT(F)	CACTCTGTCCTACTCCGAGC	*codY*
codY-RT(R)	TCTATCACACCTGCACTCTCA	*codY*
dccA-RT(F)	TCAATTACTGACCCATTAAC	*dccA*
dccA-RT(R)	CTCCTACATTATGACCTTCA	*dccA*
sigD-RT(F)	TGATAGAGAAGAGGAAGCTCCA	*sigD*
sigD-RT(R)	TCTGAAACACCTAGCACTTTTCC	*sigD*
ccpA-RT(F)	AATACCTCATGCGGCTTTTG	*ccpA*
ccpA-RT(R)	TTGCTACTGCTCCCATATCGT	*ccpA*

Table S 3.4 Under-expressed genes in R20291::*sinRR'* compared to R20291.

Transcripts significantly down regulated in C. difficile R20291::*sinRR'* is shown in Appendix C

Table S 3.5 Over-expressed genes in R20291::*sinRR'* compared to R20291.

Transcripts significantly up regulated in C. difficile R20291::*sinRR'* is shown in Appendix D

Table S 3.6 Under-expressed genes in JIR8094::*sinRR'* compared to JIR8094.

Transcripts significantly down regulated in C. difficile JIR8094::*sinRR'* is shown in Appendix E

Table S 3.7 Over-expressed genes in JIR8094::*sinRR'* compared to JIR8094.

Transcripts significantly up regulated in C. difficile JIR8094::*sinRR'* is shown in Appendix F

Table S 3.8 QRT-PCR analysis of selected genes in *sinRR'* mutants.

Gene	R20291/ R20291::*sinRR'*			
	Fold change (12h)	**adj. *p* value**	**Fold change (16h)**	**adj. *p* value**
sinR	4.578	0.0493E-06	6.975	1.486E-09
sinR'	6.893	1.41335E-10	10.456	0.54565E-03
spo0A	5.567	2.34045E-08	8.345	6.70941E-08
murG	4.342	0.00277972	5.234	1.07420E-01
sigE	5.234	0.008494963	25.35	3.44231E-05
spoIID	95.343	1.3567E-12	134.67	0.0045627
sigF	3.456	0.000688614	12.45	8.94523E-10
gpr	2.309	0.002356	5.54	1.28109E-02
sigG	10.35	0.003505323	24.56	2.3947E-08
sigK	4.56	4.28126E-10	15.09	9.42895E-07
sleB	2.39	3.8754E-3	4.67	0.0038671
sspA	54.24	2.9456E-11	78.24	5.49240E-03
spoVAC	245.986	0.005417615	689.78	0.98289E-09
spoVAD	134.734	0.000195106	276.28	0.00679831
cdeC	6.395	3.65845E-05	8.82	1.39475E-03
sleC	102.96	0.002345	240.45	7.39640E-04
cotCB	34.09	1.1946E-09	56.68	0.92587E-12
cotE	14.212	7.12107E-05	64.20	3.56729E-04
cotB	121.39	4.77323E-05	300.57	2.09561E-12
cotA	23.556	0.0056734	39.51	2.78457E-07
pdaA	2.389	0.003080399	19.83	6.27361E-03
CDR20291_2213	1.1253	3.52950E-03	0.946	4.98451E-05

3.10.4 Supplementary Figures

Figure S 3.1 Construction and confirmation of the sinR mutant in C. difficile JIR8094 and R20291.

(A) Schematic representation of ClostTron (group II intron)- mediated disruption of the sinR gene in C. difficile. (B) PCR verification of the intron insertion, conducted with intron-specific primer EBS universal [EBS(U)] with sinR—specific primers ORG-549 and ORG-550.

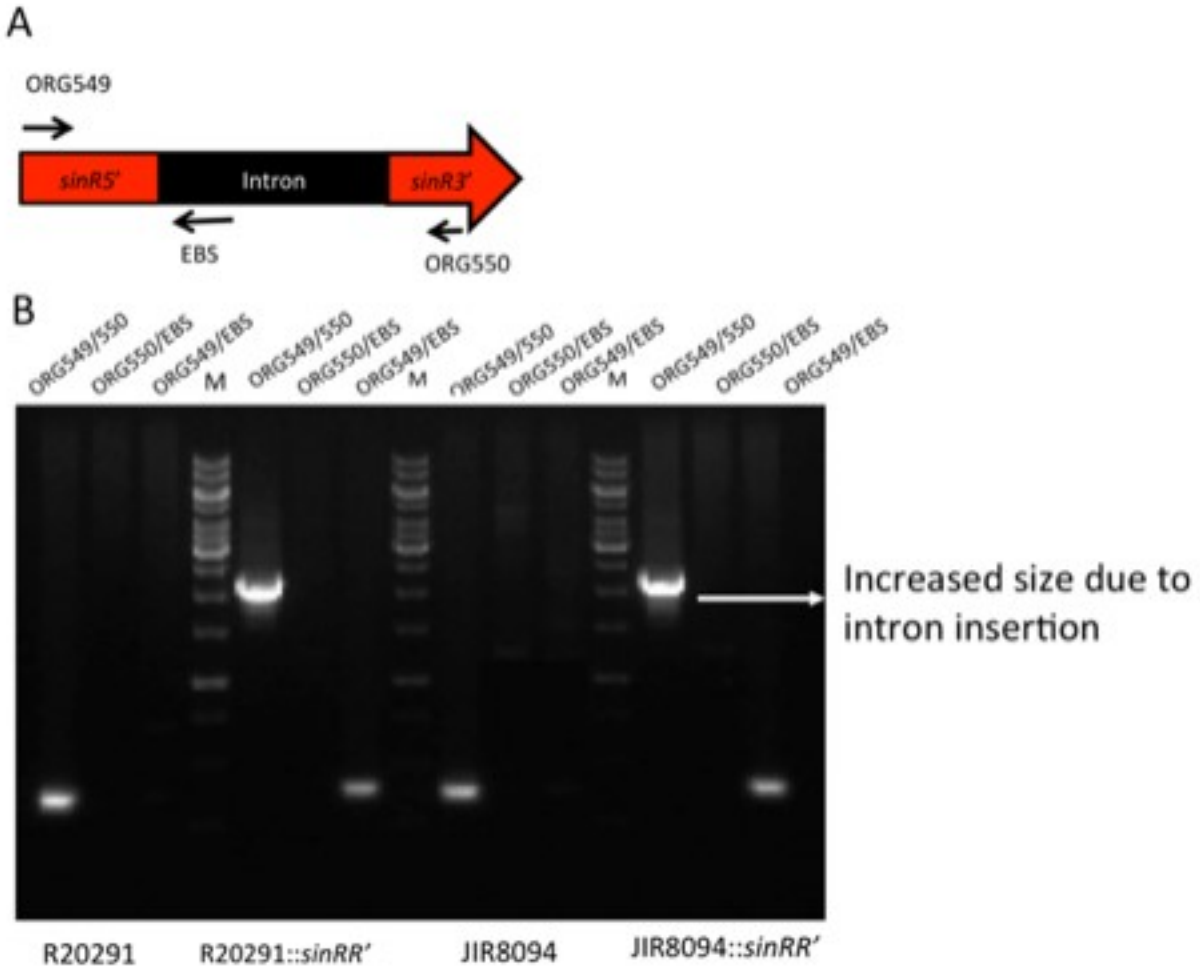
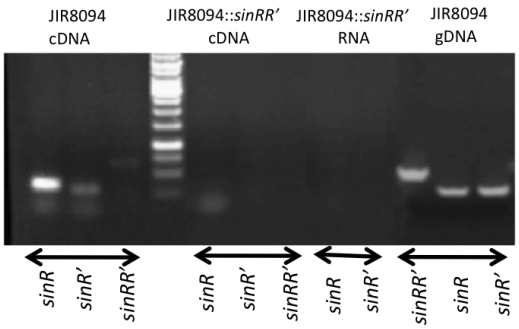


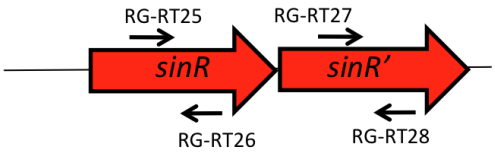
Figure S 3.2 Evidence for the read-through transcription of sin locus in C. difficile.

(A) RT-PCR results of *sinRR'*, *sinR* and *sinR'* using cDNA, RNA and genomic DNA prepared from *C. diffi- cile* JIR8094 and JIR8094::*sinRR'*. (B) Schematic representation of gene structure in *sin* locus and the location of primer design site for each gene products respectively. (C) RT-PCR results of *sin* locus transcripts in JIR8094 and JIR8094::*sinRR'* strains collected at different time points. The representative results from three independent experiments are shown.

A



B



C

Transcripts measured	Age of the bacterial culture from the time of inoculation*.	JIR8094 ΔCt ±SE	JIR8094::sinRR' ΔCt ±SE
sinR	8h	16.45 ± 0.48	33.21 ±1.45
	10h	12.56 ± 0.32	34.51±0.86
	12h	13.67 ±0.76	ND
sinR'	8h	17.23 ± 0.23	34.56±0.91
	10h	13.39 ± 0.45	ND
	12h	14.56 ±0.89	ND
sinRR'	8h	17.78 ±0.67	ND
	10h	14.41 ±0.56	ND
	12h	18.46±0.71	ND

Figure S 3.3 Characterizing sinRR' mutants.

(A) Western blot analysis of parent and mutants using SinR and SinR' specific antibodies.
(B) Growth curve of parent and the mutant strains.

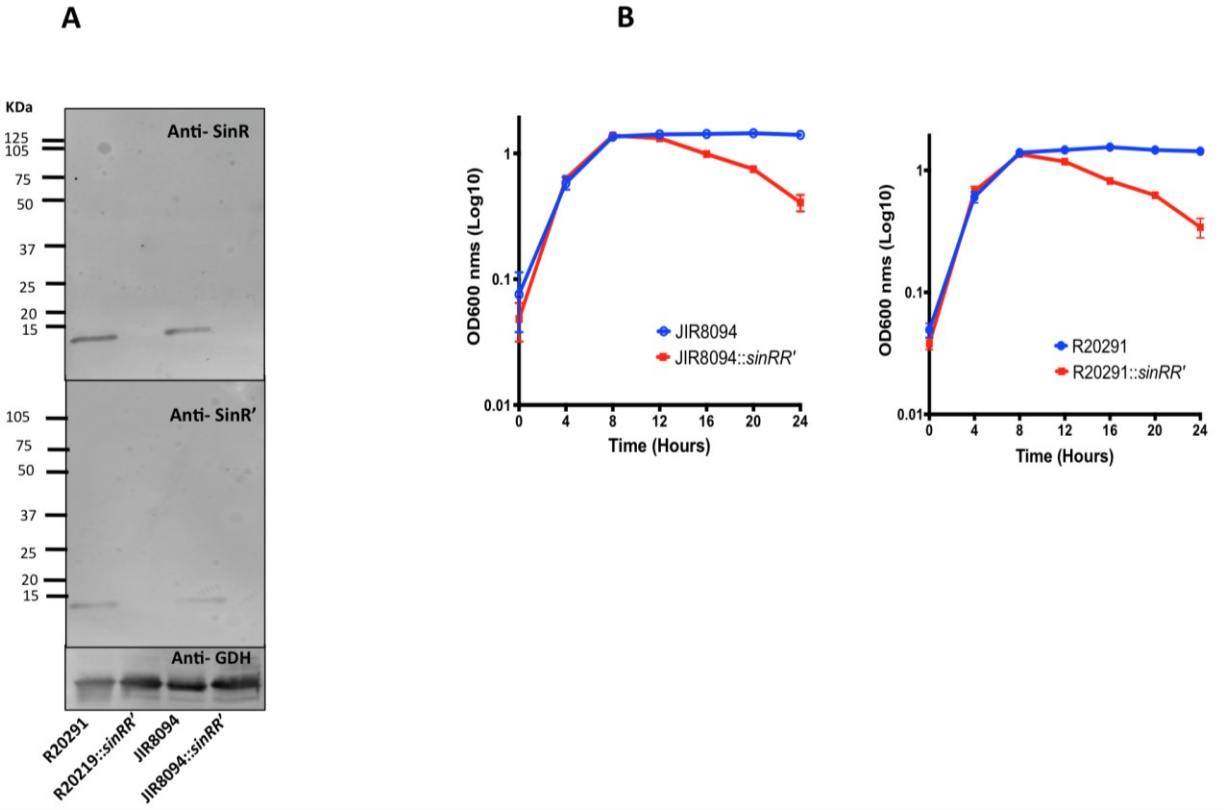


Figure S 3.4 Autolysis assay.

Triton X-100 induced autolysis of R20291::*sinRR'* at the stationary phase showing rapid lysis compared to the parent strain. Expression of *sinRR'* prevented autol- ysis in *sinRR'* mutant. The autolysis is expressed as percent initial absorbance at an optical den- sity of 600nm. Error bars indicate ± standard deviation. The experiments were repeated at least three times independently (*, *p*:::0.05 by a two-tailed Student's *t*-test).

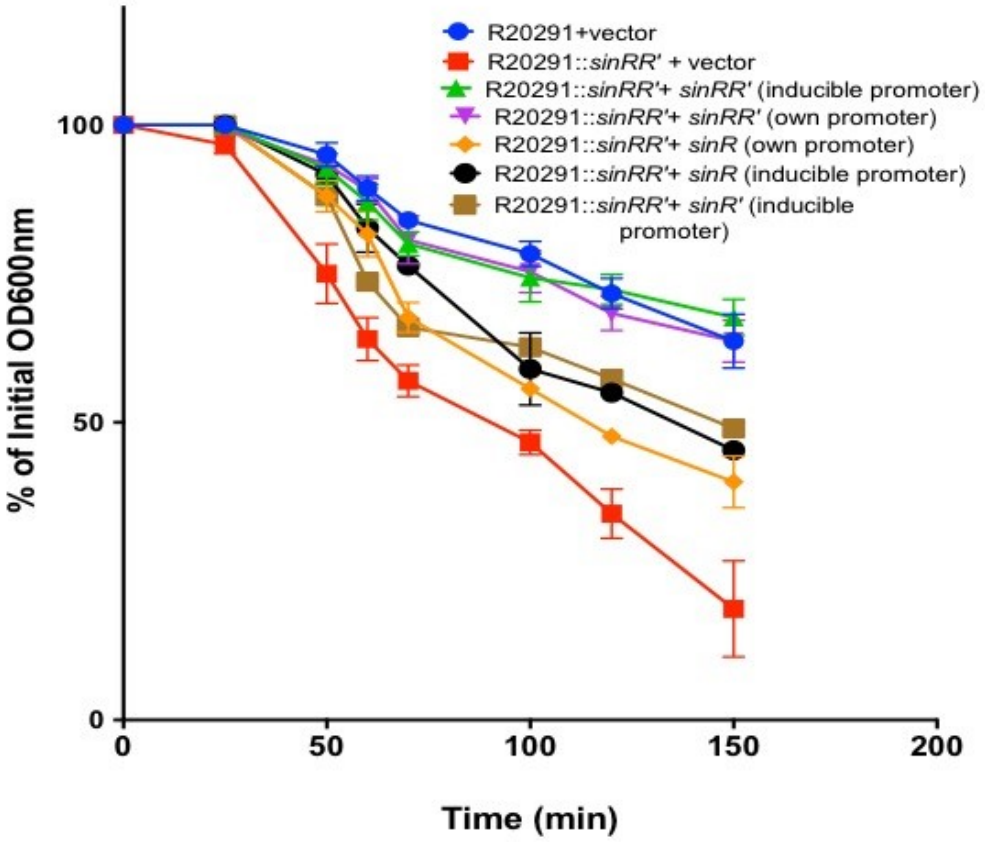


Figure S 3.5 Analysis of sporulation in JIR8094::sinRR' mutant.

(A) Phase contrast microscopy of JIR8094 and JIR8094::*sinRR'* cells. (B) JIR8904::*sinRR'* mutant was asporogenic as shown in the representative TEM images in comparison with the parent strain. Black arrows indicate mature spores in the parent strain. (C) Sporulation frequency of JIR8094 and JIR8094::*sinRR'* strains. The data shown are mean ± standard errors of three replicates. *** $p < 0.0005$ (by two-tailed student's *t*-test). (D) Western blot analysis demonstrating lower Spo0A expression in the *sinRR'* mutant.

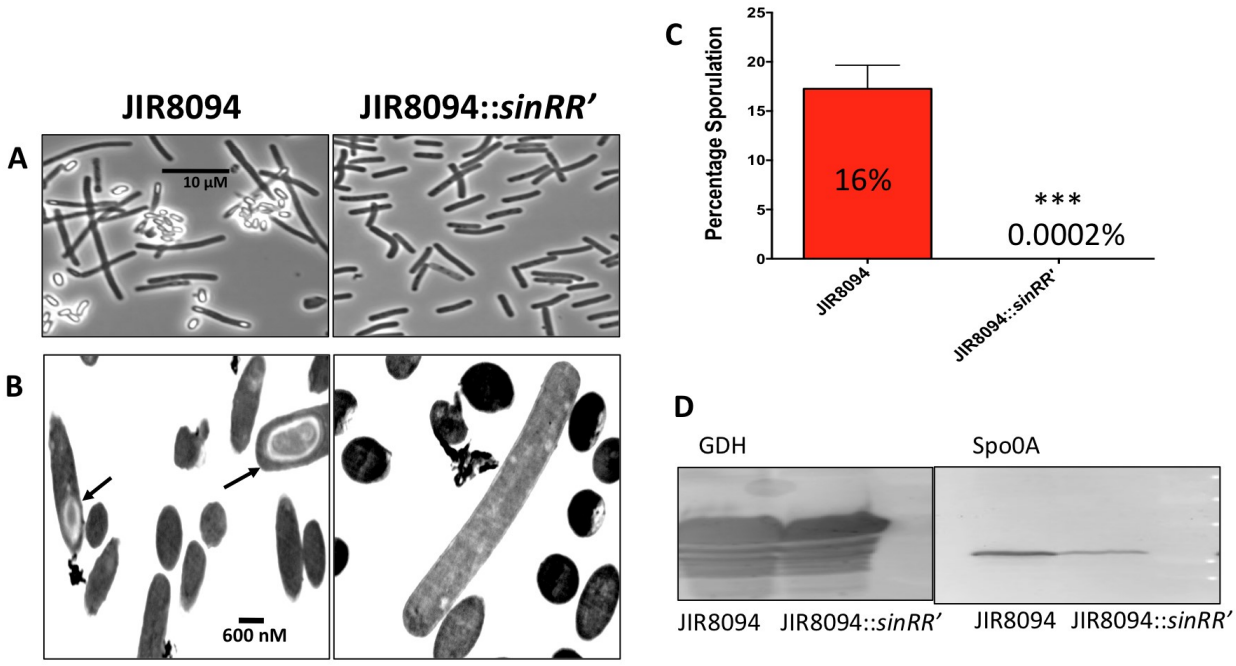


Figure S 3.6 Motility analysis of sinRR' mutant

(A) Dot blot analysis of R20291, R20291::*sinRR'* proteins using FliC and GDH (internal control) specific antibody. (B) Swimming motility of the R20291 and R20291::*sinRR'* strain showing the non-motile phenotype of *sinRR'* mutant in BHIS with 0.3% agar.

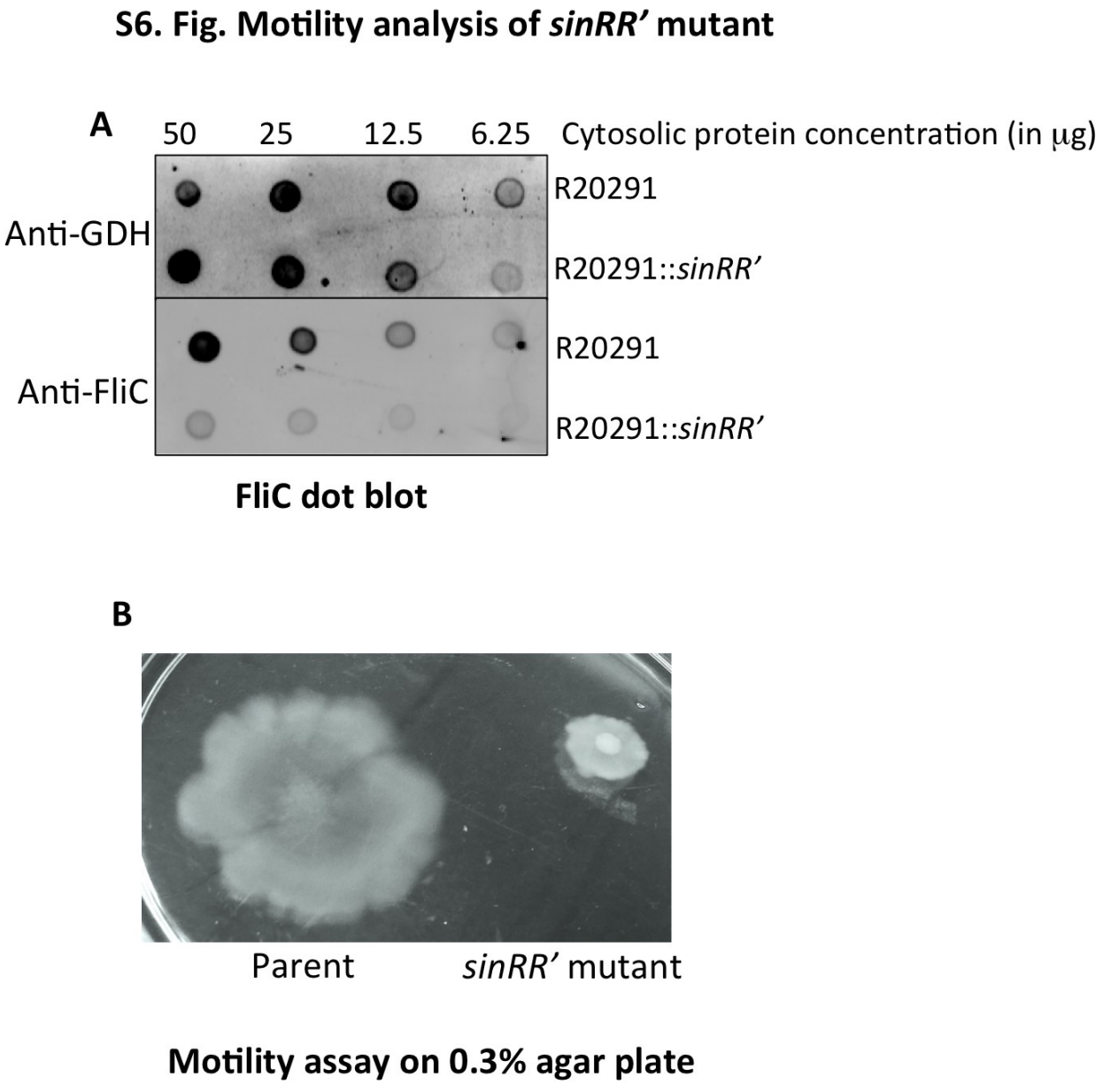


Figure S 3.7 Toxin production in JIR8094::sinRR' mutant.

Toxin ELISA performed with cytosolic proteins harvested from JIR8094 and JIR8094::*sinRR'* mutant. The data shown aremean ± standard errors of three replicates. ** *p*< 0.005 (by two-tailed student's t-test).

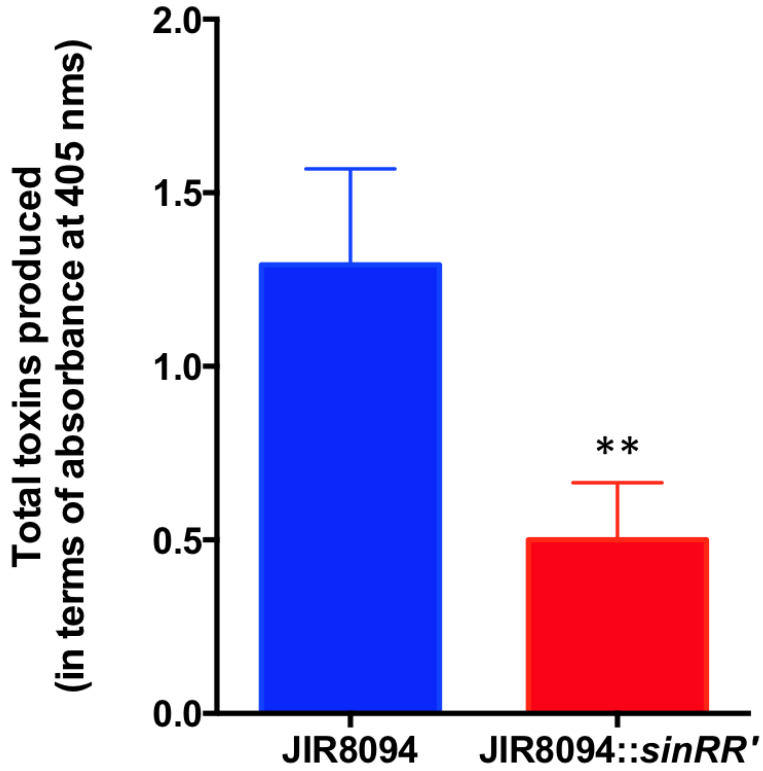


Figure S 3.8 Quantification of intracellular c-di-GMP by HPLC.

(A) The c-di-GMP peak in HPLC. (B) The standard curve was constructed by analyzing samples containing a predeter- mined amount of c-di-GMP and their respective peak area. (C) Analysis of intracellular nucle- otide pools prepared from R20291 and R20291::*sinRR'* cells. Arrows indicate the peak corresponding to c-di-GMP.

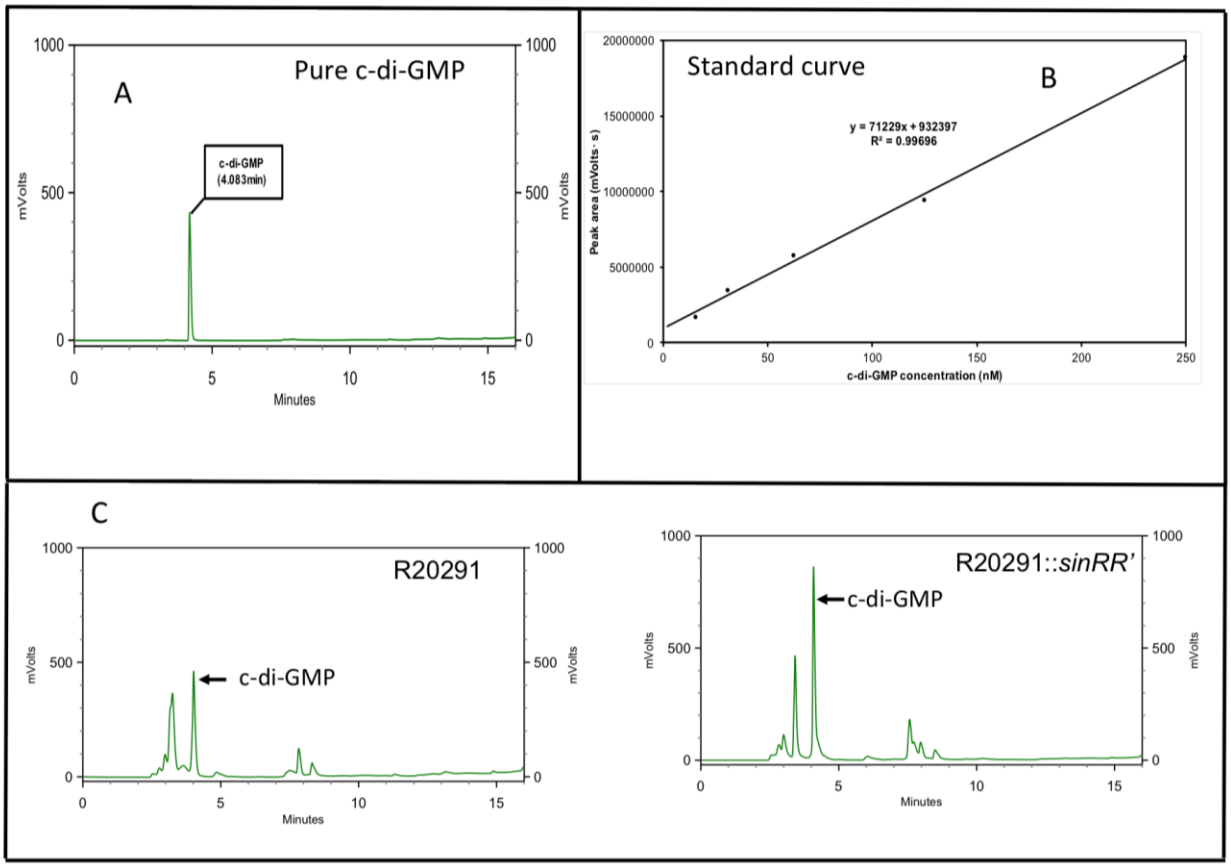


Figure S 3.9 Construction and characterization of the sinR' mutant in C. difficile R20291.

(A) PCR verification of the intron insertion verified with intron-specific primer EBS universal [EBS(U)] with gene-specific primers ORG-553 and ORG-554. (B) Schematic representation of ClostTron (group II intron)- mediated disruption of the *sinR'* gene in *C. difficile* R20291. (C) Western blot analysis of R20291 and R20291::*sinR'* proteins using SinR and SinR' specific anti- bodies. (D) Growth curve of parent R20291 and *sinR'* mutant in TY medium showing no autolysis of *sinR'* mutant.

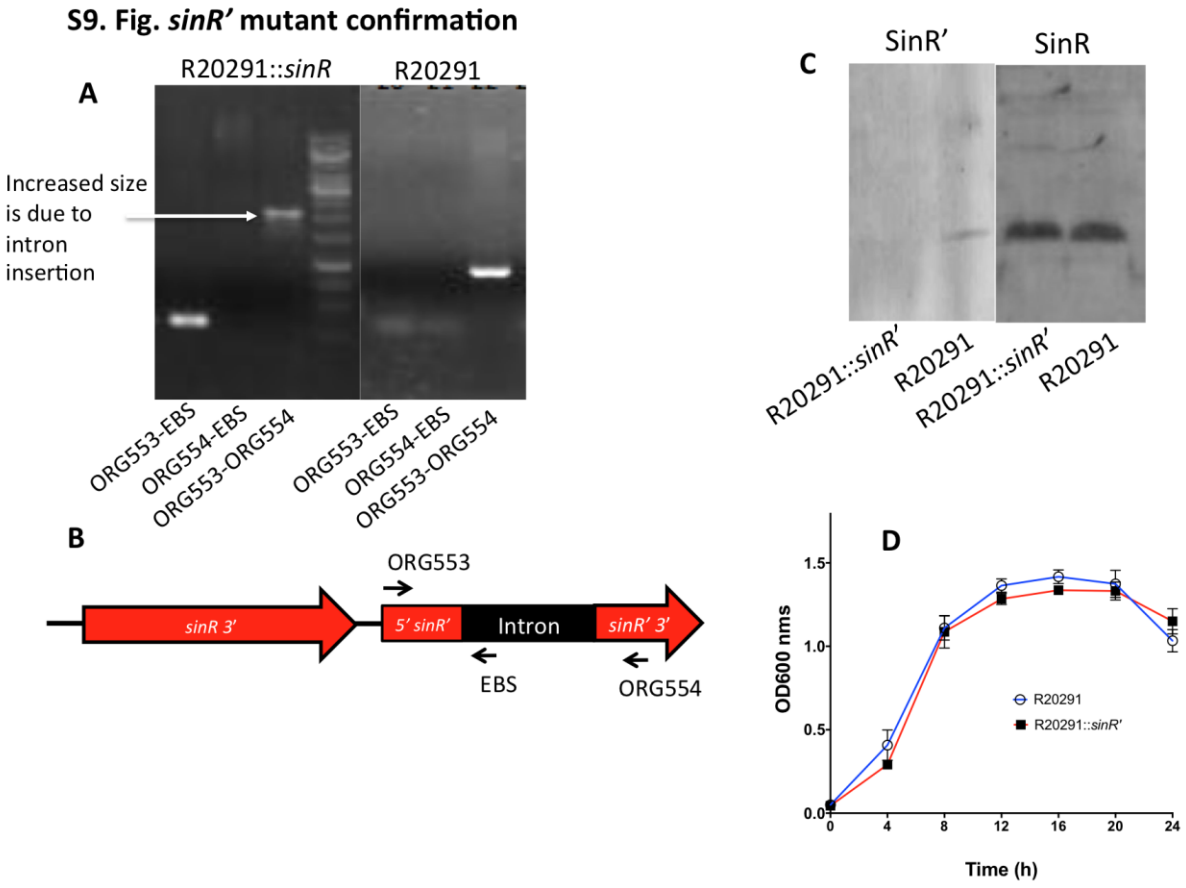


Figure S 3.10 Gel mobility shift assay reveals neither SinR nor SinR' binds to gluD upstream (non-specific control DNA).

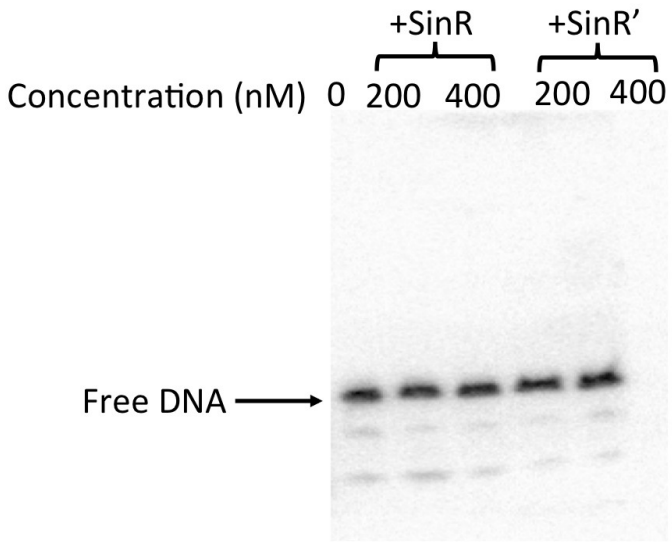


Figure S 3.11 Toxin ELISA to detect *C. difficile* toxins in cecal contents of infected hamsters.

Cecal contents harvested upon post-mortem were analyzed using *C. difficile* premier Toxin A &B ELISA kit from Meridian Diagnostics Inc. (Cincinnati, OH), following manufacturer's instruction. Negative control from the ELISA kit used along with the test samples. Each bar represents one animal.

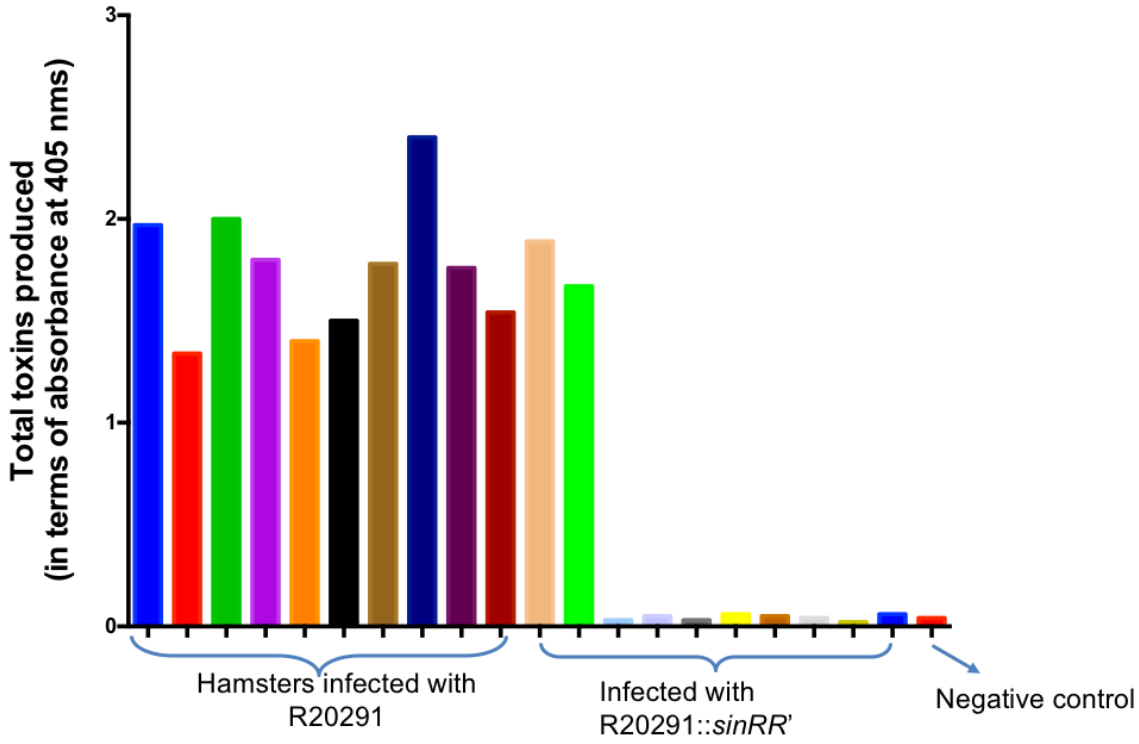
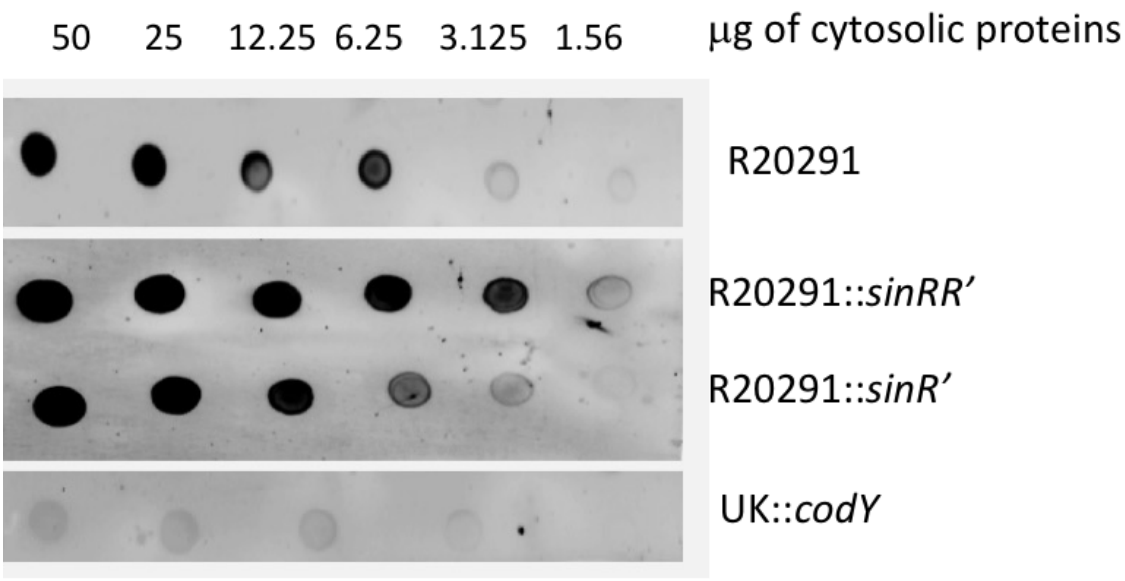


Figure S 3.12 Dot blot analysis of R20291, R20291::sinRR' and R20291::sinR' cytosolic proteins using CodY specific antibody. UK::*codY* mutant was used as a control**.**



Chapter 4 - Y146N mutation blocks roadblock repression of CodY

4.1 Abstract

The incidence of *C. difficile* infection (CDI) has increased dramatically due to the emergence of hypervirulent strains which are highly infectious, has higher symptomatic rate and outcompete endemic strains in the host's gut resulting in severe illness, frequent relapse, and increased mortality [1]. The physiological factors that led to this rapid emergence remain unclear. Here we show that an epidemic hypervirulent strain R20291 belonging to ribotype 027 carries a missense mutation where the 146th tyrosine residue is replaced with asparagine (CodY^{Y146N}) in the dimerization domain of CodY and this residue is found to be highly conserved in CodY homologs. Cody is a unique global transcriptional regulator, which recognizes GTP and ILVs (Branched Chain Amino Acids – Isoleucine, Leucine and Valine) as substrates and represses the transcription of several genes required for metabolism and virulence in *C. difficile*. Here we report that recombinant CodY^{Y146N} formed more tetramers (inactive form) than dimers (active). Through various genetic and biochemical experiments, we have shown that CodY^{Y146N} only partially regulated the genes required for toxin production and sporulation *in vivo* and it remained irresponsive to GTP and ILVs *in vitro*. Our finding suggests that tyrosine residue in the 146th position is critical for dimerization and to fully activate CodY.

4.2 Introduction

Bacterial growth depends on the nutrient supply and this leads to adjustment in its gene expression. In response to nutrient limitation, bacteria adapt to a series of adaptive changes which

includes (i) activation of transport system to uptake amino acids, peptides and other available nutrients; (ii) inducing the intracellular catabolic system and secreting extracellular protease to degrade polymeric nutrients; (iii) inducting antibiotics production to eliminate the competitors for the limited resources; (iv) activating chemotaxis and motility to move to nutrient-rich environments; and, (v) activating competent system to take up foreign DNA to develop genetic advantage or for nutrition. Failure to improve the nutritional status by the responses mentioned above will lead the bacterial cells to enter into sporulation. Hence, it is crucial for the bacteria to link the senses of extracellular stimuli with that of the intracellular stimuli and decide on a specific pattern of expression or repression of genes, operon or regulons.

In Gram-positive bacteria, the adaptation to nutritional status is mainly triggered by CodY and it has been emerged that the bacteria uses CodY to prioritize its gene expression. CodY, a 29kDa, 259 amino-acid dimeric protein is a DNA binding global transcriptional regulator which is highly conserved across low G+C Gram-positive bacteria [1]. In *C. difficile*, CodY is a major nutrient-sensing repressor protein that governs the expression of 146 genes organized in 82 transcriptional units, most of which are involved in the adaptive responses listed above and 52 out of the 82 transcriptional units are controlled directly by CodY [2]. It is also very likely that not all CodY targets can be regulated to the same magnitude at a given level of CodY Based on the availability of the substrates and the accumulation of the end-products, CodY regulates the operons of different metabolic pathways that are required for the utilization of nutrients and biosynthesis of particular cellular constituent [2, 3]. In *C. difficile*, CodY is also found to regulate the expression of virulence gene expression like toxin production and sporulation [2, 4].

CodY is a unique regulator with an allosteric metabolite interaction site [5, 6]. During the growth of cells in rich medium, CodY senses ligands like branched chain amino acids like ILV;

(Isoleucine, Leucine and Valine) and GTP molecules, becomes highly active and binds as a dimer directly to the CodY binding consensus sequence (AATT/WTTCW_G/RAAA/TA/WTT) in the upstream regions to regulate target genes transcription [5, 7, 8]. As the bacteria enters into stationary phase, the intracellular concentration of ILV and GTP decreases [9, 10]. In the absence of its substrates, CodY loses its DNA binding activity, which results in the derepression of genes repressed by CodY [7, 11]. EMSA experiments have shown that both effectors, GTP and ILV influence the function of CodY in pathogenic Gram-positive bacteria like *Bacillus subtilis Bs*CodY*, C. difficile Cd*CodY, and *Staphylococcus aureus Sa*CodY [2, 4, 11-13], however, in *Streptococcus pneumoniae Sp*CodY and *Lactococcus lactis Ll*CodY, GTP has no influence on CodY activity and it responds only to ILV [14, 15].

CodY is a two-domain protein with a N-terminal metabolite binding domain (MBD) and C-terminal DNA binding domain (DBD) and a long helical linker (LHL) in between as revealed by the crystal structure of *Bacillus cereus* CodY [6, 16, 17]. The N terminal MDB includes a highly conserved GAF domain, which was first identified in cyclic GMP [cGMP]-stimulated phosphodiesterase, Adenylate cyclase, and FhlA bacterial transcriptional regulator. It forms a hydrophobic pocket within which the isoleucine and valine binds [16-19]. Altering the residues in the ILV binding pocket were found to have profound effect on the activity of CodY [2]. The GTP binding site in CodY is predicted to be more deeply buried within the GAF domain than the ILV [11, 20]. The C terminal DBD has HTH motif that contains specific amino acids essential for efficient DNA binding and the DNA binding activity of CodY changes due to the structural changes in MBD caused by ILV interaction. The ligand interacting in MBD brings about a slight bend in the helical linker between MBD and DBD which moves the position of DBD in the dimeric CodY, facilitating efficient DNA binding [6, 21].

Over the past decade, *Bs*CodY has been extensively characterized and mutational studies in the effector binding sites of the protein have shown varying extent of gene expression. The residues important for ILV and GTP binding and dimerization in *Bacillus* species is shown in (Figure 4.1). In *B. subtilis*, the residues essential for ligand binding include R61, F71, F98 and they are critical for ILV binding and F40 and F98 are predicted to be involved in GTP interaction [2, 3]. In *C. difficile*, both essential amino acid residues like F71, F98, F40Y are highly conserved, suggesting a similar CodY regulation between *B. subtilis* and *C. difficile*. However, no mutational studies have been done to identify CodY regulation in *C. difficile* yet.

In this study, we have identified that the *codY* gene in *C. difficile* R20291 strain (PCR ribotype 027) carries a missense-mutation (in the 146th residue) that lies in the dimer interface. The wildtype CodY is designated as CodY^{WT} and mutated CodY is designated as CodY^{Y146N} in this paper. To understand the effect of Y146N mutation on CodY activity, we sought to see the difference between ligand binding potential of CodY^{WT} and CodY^{Y146N}. Studies performed by Han et al have showed *B. cereus* CodY exists in two states, an inactive tetramer form which blocks DNA binding under insufficient nutrients and an active dimer form which enhances the DNA binding activity when nutrients are surplus [4]. In this study, we have shown that the purified recombinant CodY^{WT} mostly exists in its active form as dimers however, the purified CodY^{Y146N}, predominantly formed tetramers and exhibited a slightly different ligand binding potential relative to CodY^{WT}, suggesting an obstruction in the DNA binding motif and the ligand binding sites. Consistent with our *in vitro* experiments, we also found differential regulation of CodY^{Y146N} *in-vivo*. Our findings suggest an altered pleiotropic role of CodY on *C. difficile* pathogenesis is due to Y146N mutation.

4.3 Aim of the work described in this chapter

In this chapter, I have used genetic, biochemical experiments to decipher the effect of the Y146N mutation on CodY dimerization potential and its effect on toxin production and sporulation in *C. difficile*.

4.4 Materials and methods

4.4.1 Bacterial strains and growth conditions

The plasmids and bacterial strains used in this study are listed in Table 4.1. *Clostridium difficile* strains were grown anaerobically (10% H_2, 10% $C0_2$ and 80% N_2) in Ty (tryptose yeast extract) agar or broth as described previously [5] Erythromycin (Erm; 2.5 µg/ml), Lincomycin (Linc 20ug/ml), Cefoxitin (Cef; 25 µg/ml), thiamphenicol (Thio; 15 µg/ml) were added to culture medium whenever necessary. Sporulation was induced in respective *C. difficile* strains by growing them in 70:30 sporulation medium (63 g Bacto Peptone, 3.5 g Protease Peptone, 11.1 g BHI, 1.5g yeast extract, 1.06 g Tris base, 0.7 g NH_4SO_4, 15 g agar per liter) with thiamphenicol and induced with ATc (Anhydrous Tetracycline) wherever required [6] *Escherichia coli* strain S17-1 [7] carrying specific plasmids used for complementation was cultured aerobically in Luria-Bertani (LB) broth and supplemented with chloramphenicol (25ug ml^{-1}) as indicated.

4.4.2 Construction of *C. difficile* complemented strains

The *R20291* and *UK1* chromosomal DNA was used to PCR amplify the $codY^{Y146N}$ *and* $codY^{WT}$ ORF region (768bp) along with its ribosomal binding site using gene-specific primers listed in Table 4.2, which, carried restriction sites *BamH1* and *SacI* respectively. The expression vector pRPF185 carrying tetracycline-inducible promoter and resulting PCR product was digested with

BamH1 and *SacI* and ligated to yield *codY*^{WT} (pSB1), *codY*^{Y146N} (pSB2), under tetracycline-inducible promoter respectively (Table1). The empty expression vector pRPF185-gusA (pRGL154), pSB1, pSB2 were introduced into *UK1*::*codY* mutant strains by conjugation. Transconjugants carrying pRGL154, pSB1, pSB2, were grown in the 70:30 medium supplemented with thiamphenicol (15 ug/ml) and ATc and the sporulation efficiency assay was performed as described below.

Table 4.1 Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Relevant features or genotype	Source or reference
Escherichia coli DH5α	*endA1 recA1 deoR hsdR17 (r_K^- m_K^+)*	NEB
Escherichia coli S17-1	Strain with integrated RP4 conjugation transfer function; favors conjugation between *E. coli* and *C. difficile*	[7]
Clostridium difficile UK1	Clinical isolate – NAP1/027 ribotype, isolated in	[8]
Clostridium difficile R20291	Clinical isolate - NAP1/027 ribotype, isolated in 2006 following an outbreak in Stoke Mandeville Hospital, UK	[9]
Clostridium difficile UK1::codY	*UK1*with intron insertion within *codY*	[10]
pRPF185	*E. coli*/*C. difficile* shuttle plasmid	[6]
pRGSB1	pRPF185 containing *codY*^{WT} under inducible *tet* promoter	This study
pRGSB2	pRPF185 containing *codY*^{Y146N} under inducible *tet* promoter	This study
pRG359	pET16B containing *codY*^{WT} gene with His tag	This study
pRG361	pET16B containing *codY*^{Y146N} gene with His tag	This study

4.4.3 General DNA techniques

DNeasy Blood and Tissue Kit (Qiagen) was used to extract chromosomal DNA from the *C. difficile* cultures. PCRs were carried out as described previously [5] using PCR Master Mix

(Promega) and the primers used throughout the study (Integrated DNA Technologies) is listed in Table 2. Gene-clean Kit (mpbio) was used to gel extract the PCR products, and QIAprep Spin Miniprep Kit (Qiagen) was used to obtain plasmid DNA. Standard procedures were used to perform routine cloning.

Table 4.2 Oligonucleotides used for this study

Primer	Sequence (5' → 3')	Purpose (operon)
ORG 675	CTCGAGATGGCAAGTGAAGTGTTAC AAAAAACAAGG	*codY* forward with xho1 for purification
ORG 677	GGATCCTTAATGGTGGTGGTGATGA TGTTGATTGTTTTTTAATTTTTTTAATT CATCTG	*codY* reverse with BamH1 for purification
ORG 678	AGTATGAGCTCCATGAGGAGATGAT TAAATGGCAAGTGAAGTG	*codY* forward with sac1 for pRPF185
ORG 679	GTACGGATCCTTACTTGTCGTCATCG TCTTTGTAGTCTTGATTGTTTTTTAAT TTTTTT	*codY* reverse with BamH1 and flag for pRPF185
ORG 719	TAGTTATAACTTCAAAAAAGACTGAA AATTAAGAAAAAAGAAATATAAAT	*tcdR* forward for CodY binding (T4K labelling)
ORG 720	ATTTATATTTCTTTTTTCTTAATTTTCA GTCTTTTTTGAAGTTATAACTA	*tcdR* reverse for CodY binding (T4K labelling)
ORG 721	TAAAATAGAAAATTTTTTTAATTTTCA AAATATATTCTACATATCTAATATGTA ATTAC	*sinR* upstream forward for CodY binding (T4K labelling)
ORG 722	GTAATTACATATTAGATATGTAGAAT ATATTTTGAAAATTAAAAAAATTTTCT ATTTTA	*sinR* upstream reverse for CodY binding (T4K labelling)
ORG 723	CTCATGATATTCCTCCTTTCTTTTTGC TATTATATTTTAATT	*dccA* upstream forward for CodY binding (T4K labelling)
ORG 724	TTATTAACATTGTTATTTTAAAATAAA TTAAAATTATCTTTC	*dccA* upstream reverse for CodY binding (T4K labelling)
ORG 737-F	GATGAAATTATAAAAATGAATTAACA GAATCTTTATACAATTTATCCTTTAC	*glgC* upstream forward for CodY binding (T4K labelling)
ORG 737-R	GTAAAGGATAAATTGTATAAAGATTC TGTTAATTCATTTTTATAATTTCATC	*glgC* upstream reverse for CodY binding (T4K labelling)

4.4.4 Sporulation efficiency assay

The 70:30 sporulation broth medium was used to perform sporulation assays [28] as described previously [29]. In brief, the UK strain complemented with *codY*^{Y146N} and *codY*^{WT} were grown in 70:30 medium at 37^{0}C. Samples were taken after 30 hours of growth, serially diluted and plated on Ty with 0.1% taurocholate to determine the total number of CFU. The total amount of spore's present was determined by mixing 500ul of the samples from each culture with 95% ethanol (1:1) and incubated for 1 hour to kill all the vegetative cells. The spo0A mutant was used as a negative control to ensure that all vegetative cells died during ethanol treatment. The ethanol-treated samples were then serially diluted, plated on Ty with 0.1% taurocholate and incubated for a minimum of 24 - 48 hours at 37^{0}C to enumerate the number of spores. Dividing the number of spores by the total number of CFU and multiplying the value by 100 determined the percentage of ethanol-resistant spores. The sporulation efficiency was then calculated based on the ratio of ethanol resistant cells for the mutant strains compared to the wild-type.

4.4.5 Western Blot analysis

C. difficile cells for western blot analysis were harvested and washed in 1X PBS solution before suspending in sample buffer (Tris 80mM, SDS 2%, and Glycerol 10%) for sonication. Whole cell extracts were then heated at 100^{0}C for 7 min and centrifuged at 17,000 g for 1 min, and the proteins were separated by SDS-PAGE and electroblotted onto PVDF membrane (20V, 400mAmps, 1h). Immobilized proteins on the membranes were then probed with rabbit antisera to CodY at the dilution of 1:10,000 or mouse antisera to Spo0A at the dilution of 1:10,000. Primary antibodies were then detected using HRP-conjugated goat-anti-rabbit or goat-anti-mouse secondary antibody (Jackson Immuno Research) at a dilution of 1:10000. Immunodetection of

proteins was performed with Pierce ECL 2 Western blotting Substrate Kit (Thermo Scientific) according to the manufacturer's recommendations, and the blots were developed in Typhoon 9410 scanner using 457 nm laser and 520BP 40 CY2 Blue FAM Emission filter.

4.4.6 Cloning, expression and purification of recombinant CodY^{WT}-6His and CodY^{Y146N}-6His proteins in *E. coli*

The 259-residue full-length *C. difficile codY* gene encoding *codY*^{Y146N} and *codY*^{WT}, were overexpressed in Rosetta *E. coli* cells using pET16B expression system as C-terminal hexa-His-tagged proteins. The *codY*^{Y146N} and *codY*^{WT} PCR fragments were obtained by using the chromosomal DNA of *C. difficile* R20291 and UK1 as the template. The PCR fragments were then cloned into Xho1 and BamH1 of pET16B. The resulting plasmid was then transformed into *E. coli* Rosette competent cells to obtain recombinant strain.

To overexpress the *codY*^{WT}, the *E. coli* recombinant strains were grown at 37^0C in LB medium containing chloramphenicol (25ug ml-1) and ampicillin (100ug ml-1). Protein expression was achieved by inducing with 1mM IPTG at a 17^0C overnight with mild agitation. Cells were harvested by centrifugation, and the His-tagged proteins were purified by nickel-affinity chromatography (Sigma) in buffer containing 50mM sodium phosphate (pH 8.0), 0.3M NaCl, and 10mM imidazole. The proteins were eluted with linear gradient of imidazole (50-500 mM). The collected, purified proteins were concentrated using GE Healthcare Vivaspin 2 concentrator (3KD MW/CO) and stored at -70^0C before use. The mutant codY was also overexpressed and purified as above.

4.4.7 Toxin ELISA

Relative quantification of total toxins in the *C. difficile* cultures grown in the Ty medium at different time points were performed as described previously with slight modifications [5] In brief, *C. difficile* cultures were grown for 5hrs and induced with 100ng/ml of ATc for 3 hrs and 2 ml of cells were harvested and suspended in 200ul of sterile PBS, sonicated and centrifuged to get the cytosolic proteins. 100ul of total cytosolic proteins was used to measure the relative toxin levels using *C. difficile* premier Toxin A&B ELISA kit from Meridian Diagnostics Inc. (Cincinnati, OH).

4.4.8 Electrophoretic mobility shift assay (EMSAs)

For gel mobility shift assays, the upstream regions of *tcdR, glgC* as shown in Table 4.2 was synthesized in IDT and was labeled with [□-32 P]dATP-6000 Ci/mmol (PerkinElmer Life Sciences) using T4 polynucleotide kinase. The labeled probe was then annealed with its complementary oligo to generate a double-stranded DNA probe by boiling it at 95^0C for 5 min. The DNA-protein binding reaction was carried out by mixing labeled DNA with increasing amounts of WT and Y146N in a 10ul volume containing 1x binding buffer [10mM Tris pH 8.0, 50mM KCl, 50ug BSA, 0.05% NP40, 10% Glycerol and 250ng of calf thymus DNA]. 10mM each of isoleucine, leucine, and valine (ILV), or 2mM GTP or both was added wherever indicated, and the binding reaction was performed at room temperature for 30 min. The reaction was stopped by adding 5ul of gel loading buffer and samples were separated on a pre-run 12% non-denaturing polyacrylamide gel prepared in Tris-glycine buffer, and electrophoresis was carried out in 35mM HEPES-43mM imidazole buffer (pH 7.4) as described previously [30]. If 10mM ILV was present in the binding reaction, the same concentration of ILV was also added to the electrophoresis buffer.

Gels were then dried under vacuum for 1hr, and the autoradiography was performed with Molecular Dynamics Phosphor-Imager technology.

4.5 Results

4.5.1 CodY^{Y146N} partially complement toxin production and sporulation

CodY is a known repressor of toxin production and sporulation in *C. difficile*. When nutrients are surplus and intracellular levels of ILV and GTP are high, CodY indirectly repress toxin production by repressing *tcdR*, the positive regulator of toxin genes *tcdA* and *tcdB*. However, in nutrient limiting conditions, as the levels of ILV and GTP alleviates, CodY loses its activity and toxin gene expression is turned on [11]. Recent evidences have shown that CodY also repress sporulation, but the exact mechanism is still unclear [12]. Consistent with this observation, hyper sporulating and robust toxin producing phenotypes were observed when *codY* was mutated in 630Δerm (PCR ribotype 012) and UK1 (PCR ribotype 027) strains [2, 4]. Based on these studies, it is evident that mutation or inactivation of CodY can result in increased levels of toxin production and sporulation. The epidemic R20291 strain is known for its hypertoxin production and hypersporulation. In this study, we found that the R20291 carried a missense mutation in the dimerization domain of CodY and we labeled it as CodY^{Y146N}. Multiple sequence alignment using Clustal-Omega showed that the tyrosine residue in position 146 is highly conserved across the pathogenic Gram-positive bacteria (Figure 4.1) suggesting its importance. Our repeated attempts to create a *codY* mutant in R20291 strain was unsuccessful. Since R20291 strain and UK1 strain belong to the same PCR 027 ribotype, we used the UK1Δ*codY* strain to study the effect of the Y146N mutation on the activity of CodY. We expressed the *codY*^{WT}(pSB1) and *codY*^{Y146N}(pSB2) under a tetracycline-inducible promoter in pRPF185 (as described in Materials and methods) and complemented the UK1Δ*codY* strain with pSB1 and pSB2. The expression and production of

CodY^{WT} and CodY^{Y146N} was verified through PCR, and western blot analysis using CodY and Spo0A specific antibody (Figure 4.2). Next, we performed *in vitro* experiments for toxin production and sporulation to see if there was any difference in CodY activity between the WT and mutant complemented strains.

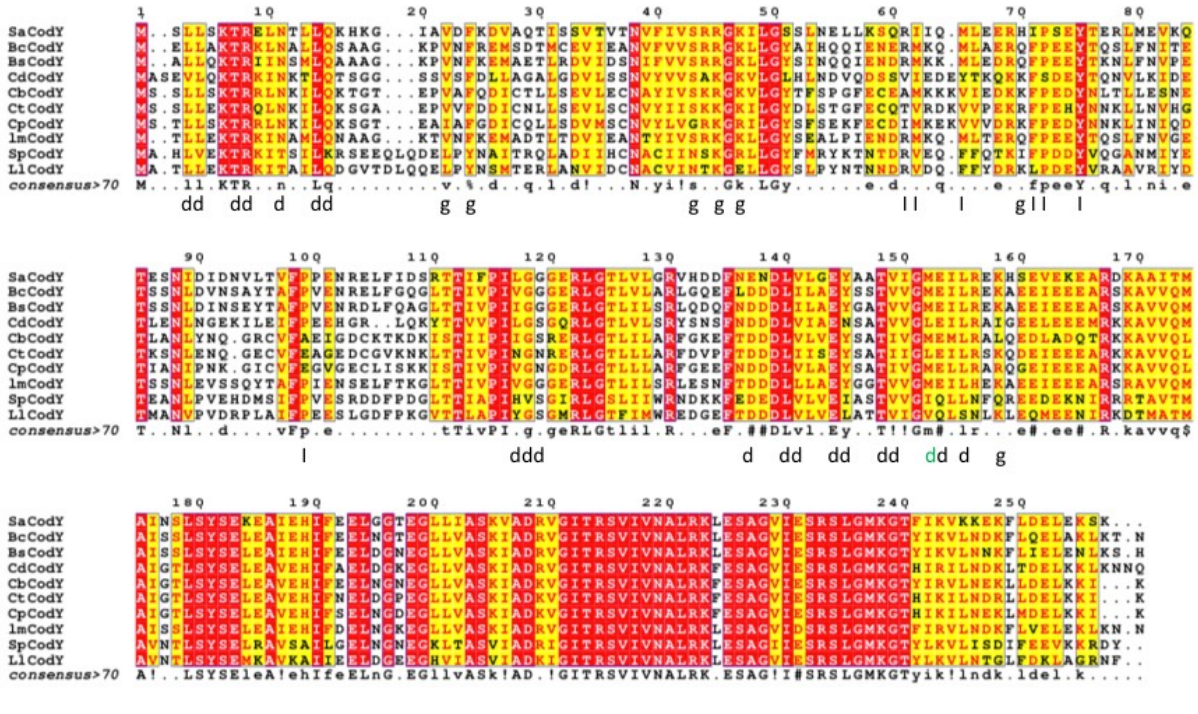


Figure 4.1 Multiple sequence alignment of CodY orthologues.

The secondary structure of CodY from specified Gram-positive bacteria is aligned using Clustal W. Conserved residues are boxed in red and invariant residues are shown in yellow. The Ile binding site, GTP binding sites, and residues whose side chain forming possible dimers is shown below as I, g and d respectively. The species used are: Sa, *Staphylococcus aureus*; Bc, *Bacillus cereus*; Bs, *B. subtilis*; Cd, *Clostridium difficile*; Cb, *C. botulinum*; Ct, *C. tetani*; *Cp. C. perfringens*; Lm, *Listeria monocytogenes*; Sp, *Streptococcus. pneumoniae*; Ll, *Lactococcus lactis*.

The expression and production of CodY^{WT} and CodY^{Y146N} was verified through western blot analysis using CodY specific antibody, and we performed *in vitro* experiments including toxin production and sporulation to see if there was any difference in CodY activity between the parent and the mutant.

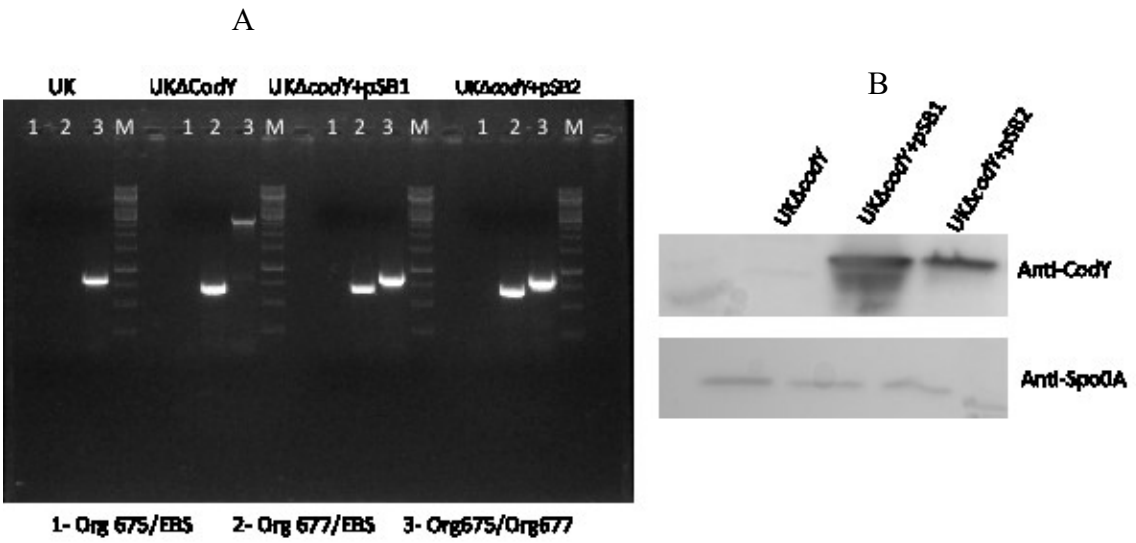


Figure 4.2 Confirmation of the complementation in UKΔ*codY*

(A) PCR verification of the complementation in UKΔ*codY* with CodY^{WT} and CodY^{Y146N} using intron-specific primer EBS universal [EBS(U)] with CodY— specific primers ORG-675 and ORG-677.

To check the effect of CodY^{Y146N} on toxin production, we induced a 6h bacterial cultures expressing UKΔ*codY*+pSB1, UKΔ*codY*+pSB2, and their respective vector alone control UKΔ*codY*+pRGL154 with 100ng/ml of ATc for 5 hrs. Cytosolic proteins were harvested from the induced cultures, and toxin ELISA was performed on it. We observed a two-fold reduction in toxin production in the UKΔcodY+pSB1, almost comparable to the parent strain, however, in UKΔcodY+pSB2 strain, partial complementation of toxin production relative to the UKΔcodY+pRGL154 was observed (Figure 4.3), suggesting an impact of the Y146N mutation on CodY function.

To determine the role of CodY^{Y146N} on sporulation, we grew the test strains in 70:30 sporulation medium with 100ng/ml of ATc for 30h. Similar to the toxin production, we observed partial complementation of sporulation in UKΔcodY+pSB2 and a complete complementation in UKΔcodY+pSB1. However, we did not notice any significant levels in the expression of Spo0A

in the complemented strains (Figure 4.2B). Hence the expression of UKΔcodY+pSB2 in the mutant failed to regulate the toxin production and sporulation to a certain extent suggesting a deferential regulation in CodY^{Y146N} activity. This diverse function in CodY activity could be attributed to binding of effectors like ILV and GTP since activity of CodY is directly proportional to the intracellular pools of ILV and GTP.

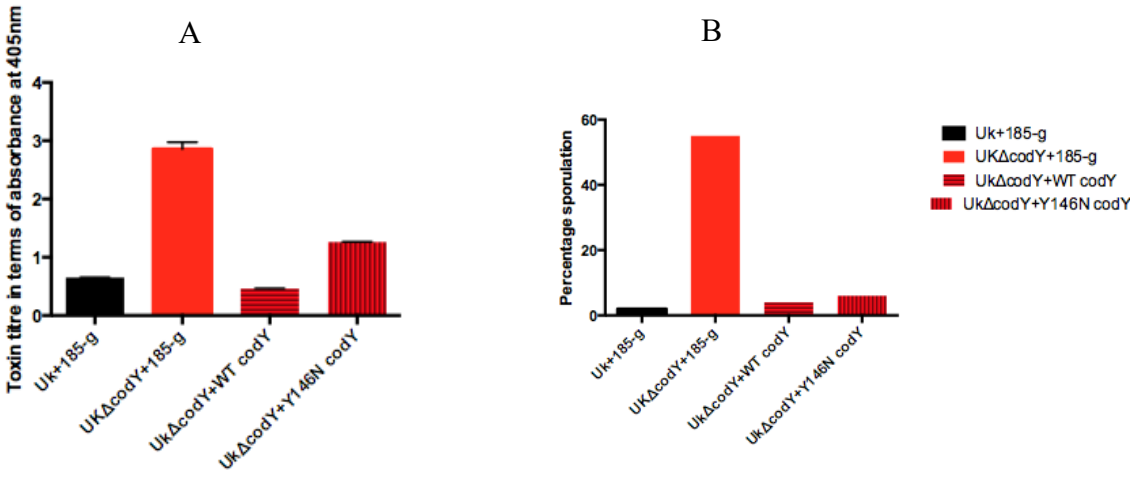


Figure 4.3 Toxin ELISA of UkΔcodY complemented strain.

Toxin production was measured by ELISA and statistical analysis was performed using one way-ANOVA with Dunnett's multiple comparisons test comparing values to the average of the parent with vector control (***<0.0005, **< 0.005 p-value) (B) *C. difficile* cultures were grown for 30 hrs in 70:30 sporulation medium under anaerobic conditions and frequency of sporulation (CFU/ml ethanol resistant spores) of complemented strains relative to mutant was determined.

4.5.2 CodY^{Y146N} predominately existed as tetramers

Since we observed a profound differential CodY^{Y146N} activity with toxin production, when complemented in the UKΔ*codY*, our central question was to see if the mutated CodY is inactive. Biochemical studies performed in *B. cereus* CodY have shown that CodY exist in two forms – active and inactive, based on the nutrient availability. In the presence of sufficient nutrients, CodY exist in active dimer conformation, and transcriptional regulation occurs, and when cells lack in

nutrients, the MBDs faces the opposite side, and the four DBDs assemble at the center of the tetramer, thereby no effectors can bind to CodY nor can CodY bind to its target DNA and becomes inactive (Figure 4.4) [4]. To test the oligomeric state of CodY variants, the $CodY^{WT}$ and $CodY^{Y146N}$ His-tagged proteins were expressed in *E. coli* and purified by Nickel affinity chromatography. When these purified recombinant proteins were analyzed on a native gel, we noticed that $CodY^{Y146N}$ existed mostly as tetramers in comparison with the wild-type CodY which mainly existed as dimers (Figure 4.5). These results provided evidence that $CodY^{Y146N}$ was inactive in its tetramer form which might explain its failure to repress the toxin production and sporulation in the UKΔ*codY*.

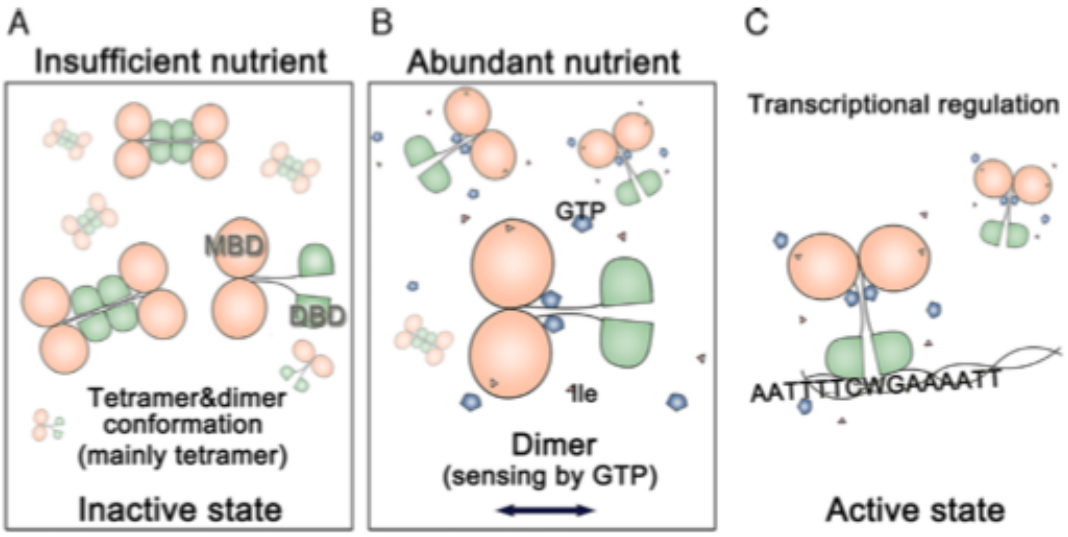


Figure 4.4 Nutrient based oligomeric state of CodY.

(A). When nutrients are insufficient, CodY exists mostly as tetramer than dimers (B). In the presence of nutrients, ligand binding facilitate rearrangement of CodY residues to its dimer conformation and activates it. (C) Ligand activated CodY binds to DNA and regulates the transcription of its target genes. The labels MBD refers to Metabolite binding domain and DBD refers to DNA binding domain LHL refers to the Long helical linker that connects MBD and DBD. The small triangles represent isoleucine molecules and blue pentagon represent GTP molecules. (Adapted from [4])

4.5.3 Differential Interaction of CodY^{WT} and CodY^{Y146N} with target promoter regions

To address the question on how well the CodY^{Y146N} responded to GTP and ILV and interacted with DNA, we performed *in vitro* binding experiment of the purified wild type and mutated CodY protein with known CodY target DNA.

The DNA binding activity of CodY is highly enhanced in the presence of effectors like ILV and GTP [11, 13] and in a recent study it was showed that the tetramer forms of CodY are deficient in binding to these effectors [4]. CodY regulates toxin production by binding directly to the promoter region of *tcdR*, and repressing the expression of toxin genes [14]. Since CodY^{Y146N} exist mostly as a tetramer, we wanted to examine the role of mutated CodY for its ability to bind to the upstream region of *tcdR*. We performed EMSA with purified CodY^{WT} and CodY^{Y146N} using [γ-^{32}P]dATP- labeled 60bp oligonucleotide containing a single putative binding sequence upstream of the *tcdR* gene (using ORG 719 and ORG 720).

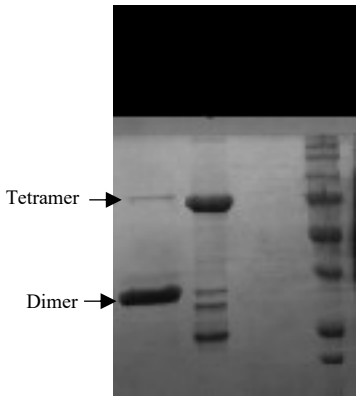


Figure 4.5 Variants of purified wildtype and mutated CodY.
Purified form of wildtype and mutated CodY is shown and their different oligomeric state (dimer and tetramer) is marked.

Probes generated using known CodY binding sequence upstream of the *glgC* gene (using ORG 737-F and ORG 737-R) was also included in the experiment. EMSA results showed that the

CodY^{WT} bound to the upstream regions of *tcdR* and *glgC* and its DNA binding affinity was significantly increased in the presence of ILV and GTP (Figure 4.6). CodY^{WT} bound to the tcdR upstream at a concentration as low as 125nM in the absence of any effectors and it decreased to as low as 31nM and 16nM in the presence of ILV and GTP. However, the CodY^{Y146N} showed diminished binding to the DNA probes with a concentration as high as 250nM in the absence of effectors and no changes were observed in the presence of effectors (Figure 4.6). Similar observation was noticed with the upstream region of glgC.

This data suggests that altering the amino acid in the dimerization domain might have favored the tetrameric state of CodY. This inactive CodY would in turn alter the ILV and GTP binding pockets in CodY, which could subsequently impair the DNA binding affinity of CodY. Alternatively, the stability of CodY protein dimer might also be affected by the CodY^{Y146N} thus, resulting in decreased DNA-binding affinity. A previous study has shown that CodY stops acting as a transcriptional repressor and the CodY's affinity for DNA decreases at low GTP concentration [15-17]. However, in *S. pneumoniae* and *L. lactis*, GTP does not influence CodY activity, and it responds only to ILV and the residue in the 146th position is altered in both the bacteria similar to *C. difficile* R20291 strain (Figure 4.1)[18, 19]. However, additional site-directed mutagenesis experiments are required to elucidate the individual effector's role in determining the role of CodY^{Y146N}. Our EMSA data suggest that the oligomeric state has altered the effector binding site and henceforth the activation of CodY.

4.6 Discussion

CodY responds to effectors like ILV and GTP which stimulates the binding of CodY to the DNA sequence and it regulates more than 100 genes primarily involved in bacterial adaptive

response to nutritional availability [11, 20, 21]. In pathogenic bacteria, CodY mainly regulates functions associated with their virulence [11, 21-24]. However, the exact mechanism by which the effectors stimulate CodY binding is still unclear. CodY acts as a sensor and the effector molecules are regarded as the nutrient and energy status of the cell. CodY in *B. subtilis* is an extensively studied protein and the GAF and HTH domains in CodY are responsible for two distinct physical interactions, the recognition of substrates and DNA binding ability, respectively [25-27]. As shown by partial proteolysis experiment and structural analysis of N terminal domain of CodY, the binding of isoleucine to the GAF domain in the N-terminus, induces a significant conformational change by major repositioning of numerous residues resulting in the formation of hydrophobic pocket around the bound ILV [25].

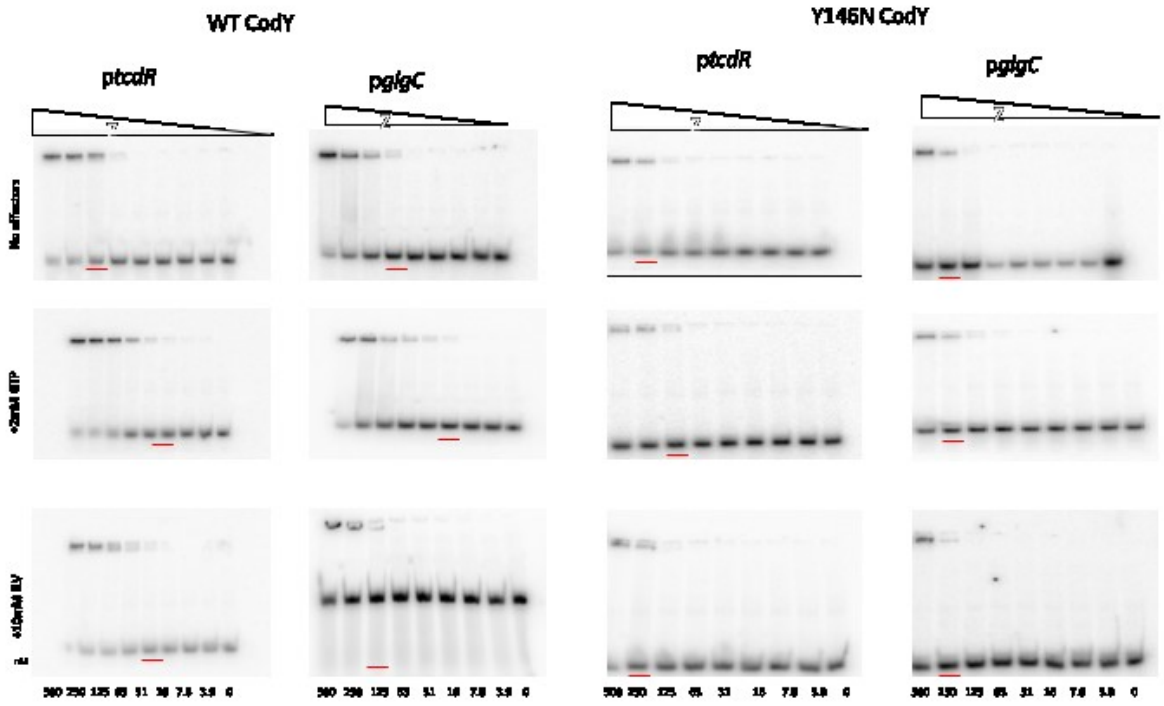


Figure 4.6 EMSA with purified CodY^{WT} and CodY^{Y146N} and known DNA targets

EMSA was performed with purified CodY^{WT} or CodY^{Y146N} and DNA fragments containing upstream regions of *tcdR* and *glgC*. ILV and GTP was added as indicated and Minimal concentration of CodY required to see is visible shift is underlined in red.

Next, the binding of isoleucine or valine to the binding pocket causes a relative orientation of two HTH domains in the CodY dimer and the HTH motif binds to the consensus DNA sequence and repress the transcription of the genes downstream of the binding site [2, 28].

Structural analysis of full length CodY in *B. cereus* have showed the existence of two conformational state of CodY, an inactive tetramer form and an intermediate dimer form with ILV which is ready to be activated and only the dimeric form of CodY is found to be active as a repressor in all the bacteria studied so far [4]. Based on the mutational studies done in *Bs*CodY, ILV binding to the GAF domain needs residues F40, F71, F98 and the dimerization of CodY is mediated by Arg8, Glu144, Thr148 residues [2, 29]. Although no mutational studies have been performed in *C. difficile* CodY, high sequence homology occurs between *C. difficile* CodY and other Gram-positive bacteria like *Bacillus subtilis, Bacillus anthracis, Clostridium perfringens, Enterococcus fecalis, Listeria monocytogenes, Streptococcus agalactiae, Streptococcus thermophiles, Streptococcus pneumoniae, Lactobacillus* and most of the residues listed above are highly conserved in *C. difficile*. However, no direct mutagenesis study was performed on Y146 in any of the CodYs to see its effect on dimer stability. Since most of the ILV, GTP and dimerization residues are highly conserved in *C. difficile*, we postulated that the Y146N mutation in CodY might have an impact on CodY activity. First, we performed *in vitro* experiment by complementing $CodY^{WT}$ and $CodY^{Y146N}$ in UKΔ*codY* and found that $CodY^{Y146N}$ was not complementing the toxin production and sporulation as efficiently as $CodY^{WT}$ (Figure 4.3). This may suggest the inactivity of CodY, which might be due to lack of effector binding or the existence of CodY in an inactive form. So, we did necessary biochemical experiments like overexpressing the proteins to see their oligomeric state of $CodY^{Y146N}$ and found that it existed mostly in tetramer state (Figure 4.5). Studies done in *Bacillus cereus* show that tetramer form of CodY is inactive [4]. To confirm this,

we performed binding experiment with both $CodY^{WT}$ and $CodY^{Y146N}$ on the upstream region of *tcdR, glgC* and found that the mutated CodY was immune to effectors like ILV and GTP (Figure 4.6). It is important to note that $CodY^{Y146N}$ protein is 80% tetramer and 20% dimer which might indeed explain the lack of roadblock repression of transcription in epidemic R20291 strain. The presence of GTP has been shown to separate CodY tetramer into dimers or monomers, however our EMSA show an irresponsive nature of $CodY^{Y146N}$ to both ILV and GTP. Since CodY is the global nutrient and energy sensing regulator, altering the CodY in a way that it remains constitutively derepressed might explain the hypervirulence nature of the epidemic R20291 strain.

CodY is an exclusive pleiotropic regulatory repressor, and we have shown evidence that $CodY^{Y146N}$ is functionally inactive *in vivo* and *in vitro*. However further studies are required to determine the exact mechanism of action. We are not sure if this mutated CodY has differential targets, which requires extensive genetic and biochemical experiments. It is unclear as to how the mutated CodY regulates sporulation which also requires further investigation.

4.7 Acknowledgments

We would like to extend our sincere gratitude to the following investigators for sharing their lab resources: Nigel Minton, University of Nottingham, for the plasmid pMTL007C-E5; Linc Sonenshein, Tufts University for CodY antibody and UK::*codY* mutant; Robert Fagan for the vector pRPF185.

4.8 References

1. Yakob, L., et al., Mechanisms of hypervirulent Clostridium difficile ribotype 027 displacement of endemic strains: an epidemiological model. Sci Rep, 2015. **5**: p. 12666.

2. Villapakkam, A.C., et al., Genetic and biochemical analysis of the interaction of Bacillus subtilis CodY with branched-chain amino acids. J Bacteriol, 2009. **191**(22): p. 6865-76.

3. Levdikov, V.M., et al., The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in gram-positive bacteria. J Biol Chem, 2006. **281**(16): p. 11366-73.

4. Han, A.R., et al., The structure of the pleiotropic transcription regulator CodY provides insight into its GTP-sensing mechanism. Nucleic Acids Res, 2016. **44**(19): p. 9483-9493.

5. Girinathan, B.P., et al., Effect of tcdR Mutation on Sporulation in the Epidemic Clostridium difficile Strain R20291. mSphere, 2017. **2**(1).

6. Fagan, R.P. and N.F. Fairweather, Clostridium difficile has two parallel and essential Sec secretion systems. J Biol Chem, 2011. **286**(31): p. 27483-93.

7. Teng, F., B.E. Murray, and G.M. Weinstock, Conjugal transfer of plasmid DNA from Escherichia coli to enterococci: a method to make insertion mutations. Plasmid, 1998. **39**(3): p. 182-6.

8. Sorg, J.A. and A.L. Sonenshein, Inhibiting the initiation of Clostridium difficile spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol, 2010. **192**(19): p. 4983-90.

9. Stabler, R.A., et al., Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol, 2009. **10**(9): p. R102.

10. Mooyottu, S., et al., Carvacrol and trans-cinnamaldehyde reduce Clostridium difficile toxin production and cytotoxicity in vitro. Int J Mol Sci, 2014. **15**(3): p. 4415-30.

11. Dineen, S.S., S.M. McBride, and A.L. Sonenshein, Integration of metabolism and virulence by Clostridium difficile CodY. J Bacteriol, 2010. **192**(20): p. 5350-62.

12. Nawrocki, K.L., et al., CodY-Dependent Regulation of Sporulation in Clostridium difficile. J Bacteriol, 2016. **198**(15): p. 2113-30.

13. Brinsmade, S.R., et al., Regulation of CodY activity through modulation of intracellular branched-chain amino acid pools. J Bacteriol, 2010. **192**(24): p. 6357-68.

14. Dineen, S.S., et al., Repression of Clostridium difficile toxin gene expression by CodY. Mol Microbiol, 2007. **66**(1): p. 206-19.

15. Geiger, T. and C. Wolz, Intersection of the stringent response and the CodY regulon in low GC Gram-positive bacteria. Int J Med Microbiol, 2014. **304**(2): p. 150-5.

16. Handke, L.D., R.P. Shivers, and A.L. Sonenshein, Interaction of Bacillus subtilis CodY with GTP. J Bacteriol, 2008. **190**(3): p. 798-806.

17. Ratnayake-Lecamwasam, M., et al., Bacillus subtilis CodY represses early-stationary-phase genes by sensing GTP levels. Genes Dev, 2001. **15**(9): p. 1093-103.

18. Hendriksen, W.T., et al., CodY of Streptococcus pneumoniae: link between nutritional gene regulation and colonization. J Bacteriol, 2008. **190**(2): p. 590-601.

19. Petranovic, D., et al., Intracellular effectors regulating the activity of the Lactococcus lactis CodY pleiotropic transcription regulator. Mol Microbiol, 2004. **53**(2): p. 613-21.

20. Belitsky, B.R. and A.L. Sonenshein, Genome-wide identification of Bacillus subtilis CodY-binding sites at single-nucleotide resolution. Proc Natl Acad Sci U S A, 2013. **110**(17): p. 7026-31.

21. Somerville, G.A. and R.A. Proctor, At the crossroads of bacterial metabolism and virulence factor synthesis in Staphylococci. Microbiol Mol Biol Rev, 2009. **73**(2): p. 233-48.

22. Brinsmade, S.R., CodY, a master integrator of metabolism and virulence in Gram-positive bacteria. Curr Genet, 2017. **63**(3): p. 417-425.

23. Lobel, L., et al., The metabolic regulator CodY links Listeria monocytogenes metabolism to virulence by directly activating the virulence regulatory gene prfA. Mol Microbiol, 2015. **95**(4): p. 624-44.

24. Sonenshein, A.L., CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. Curr Opin Microbiol, 2005. **8**(2): p. 203-7.

25. Levdikov, V.M., et al., Structural rearrangement accompanying ligand binding in the GAF domain of CodY from Bacillus subtilis. J Mol Biol, 2009. **390**(5): p. 1007-18.

26. Levdikov, V.M., et al., Structure of the Branched-chain Amino Acid and GTP-sensing Global Regulator, CodY, from Bacillus subtilis. J Biol Chem, 2017. **292**(7): p. 2714-2728.

27. Ling, B., et al., Molecular dynamics simulations of isoleucine-release pathway in GAF domain of N-CodY from Bacillus Subtilis. J Mol Graph Model, 2013. **44**: p. 232-40.

28. Belitsky, B.R. and A.L. Sonenshein, Genetic and biochemical analysis of CodY-binding sites in Bacillus subtilis. J Bacteriol, 2008. **190**(4): p. 1224-36.

29. Stenz, L., et al., The CodY pleiotropic repressor controls virulence in gram-positive pathogens. FEMS Immunol Med Microbiol, 2011. **62**(2): p. 123-39.

Chapter 5 - Final conclusion and future direction

The main objective of the work presented in this dissertation is to identify the interplay between STM (Sporulation, Toxin production, and Motility) pathways. Firstly, we used microscopic and transcriptomic approach to decipher the novel differential role of *tcdR* on sporulation. Secondly, we characterized the pleiotropic role of *sin* locus which was differentially regulated in *tcdR* mutant using genetic, biochemical, microscopic and NGS approaches, which suggested a significant amount of crosstalk between critical regulators of STM pathways.

Taken together, data from this study defines how TcdR and SinR coordinate other regulators in regulating major virulence factors like sporulation, toxin production, motility, metabolism.

5.1 Role of *tcdR* in sporulation

The work presented in Chapter 2, describe for the first time the role of TcdR in sporulation. We show that in the absence of *tcdR*, both the quantity and quality of the spores are affected mainly due to downregulation of sporulation specific sigma factors (*sigE, sigK, sigF, sigG*) and their associated genes such as *spoIIR, spoIIGA, spoIIID, pdaA, sspA, spoVAC, spoVAD, cotJBD, cotA, cotB, cotE, bclA3, bclA2, sleC, cdeC, spoVFB,* and *dpaAB* which are involved in spore structure and germination. SigE and SigK are mother cell specific and SigF and SigG are forespore specific sigma factors. Downregulation of these major sigma factors will affect the overall sporulation efficiency as identified in this chapter. The early sigma factor SigE is activated by proteolytic cleavage of pro-SigE, with the help of enzyme SpoIIGA which is co-expressed with *sigE*. The signal for SpoIIGA activation is the SpoIIR signal protein which is synthesized in the newly formed forespore and is regulated by SigF Hence a reduced *spoIIR* due to lower levels of *sigF* will

decrease the levels of activated SpoIIGA and hence accumulation of more inactive pro-SigE. *spoIIID* is a transcriptional regulator involved in the transcription of *sigK* and in the absence of *spoIIID* little to no transcription of *sigK* would occur. In the absence of SigK, genes such as *cotJBD, cotA, cotB, cotE, bclA3, bclA2, spoVAC, dpaAB, sleC, cdeC, cspB* that synthesis spore coat and exosporium are affected. An unglycosylated BclA3 is shown to be highly susceptible to heat treatment and the gene CD3350 responsible for glycosylation of BclA3 is significantly downregulated in the *tcdR* mutant suggesting the role of BclA3 in heat sensitive spores. Studies have also shown that *spoVAC*, *dpaAB* mutants are highly heat sensitive and Dipicolinate synthase enzyme subunits (SpoVFB and DpaA) are required for production of dipicolinic acid (DPA) which protect spores from heat treatment. CdeC, a major protein needed for exosporium assembly is also down regulated in *tcdR* mutant. In combination, decreased levels of *bclA3, cdeC, spoVAC* and DPA coding operon can contribute to the production of spores with weaker exosporium and make them highly susceptible to heat. Not only were the spores lacking *tcdR* heat sensitive, they also required increased levels of sodium taurocholate to germinate. Germination is induced when bile salt sensing germination receptor (GR) CspC, sense germinants like sodium taurocholate and triggers the release of DPA from the spore core. This leads to the activation of spore cortex lytic enzyme, SleC. that degrade the peptidoglycan (PG) cortex layer, which then allows core hydration. Activation of SleC depends on CspC through CspB mediated cleavage of prodomain to generate active SleC. The *sleC* levels in the *tcdR* mutant was found to be downregulated and *sleC* mutant is defective in germination suggesting lower *sleC* role in inefficient germination of *tcdR* mutant spore. In summary, in the absence of *sigE, sigG, sigK* and their associated genes, less spores are formed, and they are heat sensitive, defective in germination and exosporium formation. All these data suggest that a fully assembled spore requires a functional *tcdR* gene and we also have

illustrated that the credibility of spores to harsh environmental conditions is challenged in the absence of *tcdR*. We also identified an overexpression of *srlR* gene encoding glucitol/sorbitol specific PTS system in the *tcdR* mutant based on our transcriptomic data and sorbitol has been shown to inhibit sporulation in *C. difficile*. Additionally, we provide evidence that mutation in *tcdR* resulted in decreased transcript levels of *sin* regulon, suggesting a positive regulation of TcdR on *sin* regulon which led to its characterization in Chapter 3.

5.2 Role of SinR, SinR' in STM pathway

The work described in Chapter 3 exploits a combined genetic, biochemical and transcriptomic approach to define the pleiotropic role of sin regulon at a whole genome level. We have created a mutation in both the *sinR* and *sinR'* genes of sin regulon and characterized them individually to give a holistic approach to their functionality in *C. difficile* pathogenesis.

We have shown that *sinRR'* mutants are avirulent both *in vitro* and *in vivo*. The *sinRR'* mutants are autolytic, asporogenic, non-motile and produces significantly less toxin compared to the parent R20291 strain. The autolytic nature is due to both SinR and SinR', however effects on motility and toxin production are mainly due to presence of SinR protein as complementing the *sinRR'* with *sinR'* had no effect on these phenotypes. Our overexpression study using *sinR* and *sinR'* in the R20291 wildtype strain and phenotypic characterization of *sinR'*mutant further confirmed our analysis. We also identified that both motility and toxin production were interlinked as SinR was found to regulate the levels of both CodY, the negative regulator of *tcdR*, and c-di-GMP, that inversely regulates motility sigma factor *sigD*. Thereby SinR was found to act as a repressor of several negative regulators (CodY, C-di-GMP, Spo0A, SigD) of STM and metabolic pathways thereby indirectly exerting a positive effect on the virulence factors. We have also shown

that SinR and not SinR' inhibits CodY by binding directly to the promoter region of CodY and CodY also exerts a similar regulation on SinR based on our EMSA experiment and western blot analysis of respective proteins in *codY* and *sinR'* mutants. Using GST pulldown experiments, we have also characterized and shown the inhibitory effects of SinR' on SinR. Importantly, we also have validated the vital role of SinR in the pathogenesis using *in vivo* hamster models.

5.3 Differential role of CodY^{Y146N}

The preliminary work presented in Chapter 4 shows evidence on how bacteria evolve as epidemic strains by introducing SNPs in its global regulators like CodY and alters its function. CodY is a major metabolic protein, which senses the nutritional state of the cell and accordingly regulates several virulence factors. It is a known repressor of toxin gene regulation and sporulation. Here we have shown that a mutation in the dimerization interface is known to affect the oligomeric state of CodY and henceforth alter its regulation on the virulence factors. Using EMSA we have provided evidence that the altered CodY protein is unable to bind efficiently to its target DNA sequence as efficiently as the wild type protein.

5.4 Future directions

C. difficile hypervirulent strains are on the rise, which found to produce increased levels of toxins and spores which makes it very crucial to understand the regulation of toxin production and sporulation in detail and know how they interact with each other.

Our present study demonstrates that TcdR and SinR regulate multiple pathways associated with infection, persistence, and transmission of the disease. However, many questions are still left

unanswered. How *tcdR* regulate *sin* locus? How the levels of SinR and SinR' are regulated by the sin operon? How sinR controls *dccA* and *ccpA*? what leads to the inactivity of spo0A in *sinRR'* mutant and is soj-spo0J under the direct control of SinR? What are the targets of SinR' in addition to opposing SinR, what consensus sequences do SinR recognize? what are the target genes of $CodY^{Y146N}$? In short, this study has opened up a new arena of several significant questions which when answered will help in managing and treating CDI more efficiently.

Appendix A - Effect of *tcdR* mutation on Sporulation

AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere®

RESEARCH ARTICLE
Molecular Biology and Physiology

Effect of *tcdR* Mutation on Sporulation in the Epidemic *Clostridium difficile* Strain R20291

Brintha P. Girinathan,[a] Marc Monot,[b] Daniel Boyle,[a] Kathleen N. McAllister,[c] Joseph A. Sorg,[c] Bruno Dupuy,[b] Revathi Govind[a]

Division of Biology, Kansas State University, Manhattan, Kansas, USA[a]; Department of Microbiology, Laboratoire Pathogenèse des Bactéries Anaérobies, Institut Pasteur, Paris, France[b]; Department of Biology, Texas A&M University, College Station, Texas, USA[c]

ABSTRACT *Clostridium difficile* is an important nosocomial pathogen and the leading cause of hospital-acquired diarrhea. Antibiotic use is the primary risk factor for the development of *C. difficile*-associated disease because it disrupts normally protective gut flora and enables *C. difficile* to colonize the colon. *C. difficile* damages host tissue by secreting toxins and disseminates by forming spores. The toxin-encoding genes, *tcdA* and *tcdB*, are part of a pathogenicity locus, which also includes the *tcdR* gene that codes for TcdR, an alternate sigma factor that initiates transcription of *tcdA* and *tcdB* genes. We created a *tcdR* mutant in epidemic-type *C. difficile* strain R20291 in an attempt to identify the global role of *tcdR*. A site-directed mutation in *tcdR* affected both toxin production and sporulation in *C. difficile* R20291. Spores of the *tcdR* mutant were more heat sensitive than the wild type (WT). Nearly 3-fold more taurocholate was needed to germinate spores from the *tcdR* mutant than to germinate the spores prepared from the WT strain. Transmission electron microscopic analysis of the spores also revealed a weakly assembled exosporium on the *tcdR* mutant spores. Accordingly, comparative transcriptome analysis showed many differentially expressed sporulation genes in the *tcdR* mutant compared to the WT strain. These data suggest that regulatory networks of toxin production and sporulation in *C. difficile* strain R20291 are linked with each other.

IMPORTANCE *C. difficile* infects thousands of hospitalized patients every year, causing significant morbidity and mortality. *C. difficile* spores play a pivotal role in the transmission of the pathogen in the hospital environment. During infection, the spores germinate, and the vegetative bacterial cells produce toxins that damage host tissue. Thus, sporulation and toxin production are two important traits of *C. difficile*. In this study, we showed that a mutation in *tcdR*, the toxin gene regulator, affects both toxin production and sporulation in epidemic-type *C. difficile* strain R20291.

KEYWORDS *Clostridium difficile*, sporulation, toxin gene regulation

Received 22 December 2016 **Accepted** 29 January 2017 **Published** 15 February 2017

Citation 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. https://doi.org/10.1128/mSphere.00383-16.

Editor Craig D. Ellermeier, University of Iowa

Address correspondence to Revathi Govind, rgovind@ksu.edu.

Clostridium difficile is a Gram-positive, spore-forming, anaerobic bacillus and is the leading cause of hospital-acquired diarrheal diseases (1, 2). Nearly 50% of all patients carry *C. difficile* asymptomatically after hospitalization (2, 3). Nearly 10% of all *C. difficile*-infected patients develop pseudomembranous colitis, and 3% develop severe, life-threatening complications such as fulminant colitis and toxic megacolon (4). *C. difficile* infection (CDI) is commonly acquired from *C. difficile* spores present in the hospital environment, and individuals become infected when the normal colonic microbiota is suppressed by antibiotic therapy (5). In the gut, *C. difficile* spores germinate to the toxin-producing vegetative form in response to certain bile acids, e.g., taurocholic acid (TA), and amino acids. *C. difficile* toxins A (TcdA) and B (TcdB) are then

Appendix B - Pleiotropic role of sin regulator

PLOS | PATHOGENS

RESEARCH ARTICLE

Pleiotropic roles of *Clostridium difficile sin* locus

Brintha Parasumanna Girinathan [1], Junjun Ou [2], Bruno Dupuy [3,4], Revathi Govind [1]*

1 Division of Biology, Kansas State University, Manhattan, KS, United Sates of America, **2** Department of Agronomy, Kansas State University, Manhattan, KS, United Sates of America, **3** Laboratoire Pathogène Anaérobies, Institut Pasteur, Paris, France, **4** Université Paris Diderot, Sorbonne Paris Cité, Paris, France

* rgovind@ksu.edu

Citation: GirinathanBP, OuJ, DupuyB, GovindR (2018) Pleiotropicroles of *Clostridiumdifficilesin* locus. PLoS Pathog14(3): e1006940. https://doi.org/10.1371/journal.ppat.1006940

Editor: TheresaM. Koehler, University of Texas Medical School at Houston, UNITED STATES

Received: August16, 2017

Accepted: February13, 2018

Published: March12, 2018

Abstract

Clostridium difficile is the primary cause of nosocomial diarrhea and pseudomembranous colitis. It produces dormant spores, which serve as an infectious vehicle responsible for transmission of the disease and persistence of the organism in the environment. In *Bacillus subtilis*, the *sin* locus coding SinR (113 aa) and SinI (57 aa) is responsible for sporulation inhibition. In *B. subtilis*, SinR mainly acts as a repressor of its target genes to control sporulation, biofilm formation, and autolysis. SinI is an inhibitor of SinR, so their interaction determines whether SinR can inhibit its target gene expression. The *C. difficile* genome carries two *sinR* homologs in the operon that we named *sinR* and *sinR'*, coding for SinR (112 aa) and SinR' (105 aa), respectively. In this study, we constructed and characterized *sin* locus mutants in two different *C. difficile* strains R20291 and JIR8094, to decipher the locus's role in *C. difficile* physiology. Transcriptome analysis of the *sinRR'* mutants revealed their pleiotropic roles in controlling several pathways including sporulation, toxin production, and motility in *C. difficile*. Through various genetic and biochemical experiments, we have shown that SinR can regulate transcription of key regulators in these pathways, which includes *sigD*, *spo0A*, and *codY*. We have found that SinR' acts as an antagonist to SinR by blocking its repressor activity. Using a hamster model, we have also demonstrated that the *sin* locus is needed for successful *C. difficile* infection. This study reveals the *sin* locus as a central link that connects the gene regulatory networks of sporulation, toxin production, and motility; three key pathways that are important for *C. difficile* pathogenesis.

Appendix C - Down regulated genes in R∆*sinRR'* relative to R20291

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_0065	259458.0	17.9	2594.6	0.0	elongation factor	0.012532
CDR20291_1462	49945.0	15.6	499.4	0.0	putative uncharacterized protein next to holin	0.00001967
CDR20291_1078	42340.0	15.3	423.4	0.0	hypothetical protein next to 1070 operon	0.00230895
CDR20291_2278	42264.0	15.3	422.6	0.0	putative peptidoglycan-binding/hydrolysing protein	0.0005621
CDR20291_2092	35318.8	15.1	353.2	0.0	putative lipoprotein	3.42221E-08
CDR20291_0383	34088.0	15.0	340.9	0.0	hypothetical protein	0.000066277
CDR20291_1416	30812.2	14.9	308.1	0.0	putative uncharacterized protein next to phage protein	1.3014E-13
spoIIIAA	28469.8	14.8	284.7	0.0	stage III sporulation protein AA	0.00005831
CDR20291_1443	24808.9	14.6	248.1	0.0	hypothetical protein before phage protein	2.8613E-07
CDR20291_1430	24235.2	14.6	242.3	0.0	hypothetical protein	0.00047475
CDR20291_3285	23820.3	14.5	238.2	0.0	hypothetical protein	0.0078469
CDR20291_3478	21614.5	14.4	216.1	0.0	transposase like protein b	0.001479
CDR20291_3468	21168.5	14.4	211.7	0.0	hypothetical protein	1.28132E-09
CDR20291_1851	19653.3	14.3	196.5	0.0	hypothetical protein	0.000286121
CDR20291_2909	19521.0	14.3	195.2	0.0	type I restriction enzyme r subunit	2.4104E-08
CDR20291_1771	18923.9	14.2	189.2	0.0	hypothetical protein	0.000062748
CDR20291_1278	18854.4	14.2	188.5	0.0	hypothetical protein	4.524E-11
CDR20291_1325	14938.4	13.9	149.4	0.0	hydrolase	0.00009398
CDR20291_0246	14920.6	13.9	149.2	0.0	ornithine cyclodeaminase	4.90116E-13
CDR20291_1759	14083.1	13.8	140.8	0.0	addiction module toxin	0.0069524
CDR20291_1774	13588.5	13.7	135.9	0.0	hypothetical protein	0.00209598
CDR20291_2761	13072.1	13.7	130.7	0.0	hypothetical protein	1.2348E-09
CDR20291_2912	12202.3	13.6	122.0	0.0	type I restriction enzyme m subunit	0.00004851
CDR20291_3289	12147.1	13.6	121.5	0.0	leucine rich repeat protein	3.8046E-07
CDR20291_0863	11684.0	13.5	116.8	0.0	hypothetical protein	0.00007392
CDR20291_0921	11557.1	13.5	115.6	0.0	hypothetical protein adjacent to ccpA	0.000133551

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
tlpB	11166.3	13.4	111.7	0.0	transposase like protein b	0.000361228
CDR20291_0438	11058.8	13.4	110.6	0.0	DNA binding response regulator	0.000077184
CDR20291_0242	10870.0	13.4	108.7	0.0	glycosyl transferase group 2 family protein	0.007
CDR20291_0226	10807.0	13.4	108.1	0.0	dtdp-glucose 4,6 dehydratase	0.00000017
CDR20291_3125	10281.2	13.3	102.8	0.0	hypothetical protein	1.61003E-06
CDR20291_0243	10062.6	13.3	100.6	0.0	glycosyl transferase group 2 family protein	0.0042336
CDR20291_3025	8665.1	13.1	86.6	0.0	gcn5-related n-acetyltransferase	1.44321E-10
CDR20291_2961	8406.3	13.0	84.1	0.0	hypothetical protein	0.000323604
CDR20291_2760	7594.9	12.9	75.9	0.0	hypothetical protein	0.0044891
CDR20291_3124	7532.0	12.9	75.3	0.0	sensor protein	7.0004E-22
CDR20291_1741	7231.1	12.8	72.3	0.0	hypothetical protein	0.0002405
CDR20291_1448	7181.0	12.8	71.8	0.0	hypothetical protein next to phage protein	2.376E-09
CDR20291_0502	7008.7	12.8	70.1	0.0	hypothetical protein	0.00118455
CDR20291_1449	6287.6	12.6	62.9	0.0	phage tail tape measure protein	3.014E-17
CDR20291_1071	6231.5	12.6	62.3	0.0	hypothetical protein in 1070 operon	0.00000924
CDR20291_1748	6016.7	12.6	60.2	0.0	two component response regulator	9.8304E-14
CDR20291_0551	5960.0	12.5	59.6	0.0	ABC transporter ATP binding protein/permease	1.16724E-08
CDR20291_1418	5935.8	12.5	59.3	0.0	hypothetical protein next to phage protein	2.45688E-07
CDR20291_0710	5800.6	12.5	58.0	0.0	hypothetical protein	0.000291368
CDR20291_1419	5520.9	12.4	55.2	0.0	hypothetical protein next to phage protein	0.01624
CDR20291_1423	5356.4	12.4	53.6	0.0	hypothetical protein next to phage protein	0.000069133
yobD	5015.6	12.3	50.1	0.0	transcriptional regulator yobd protein	0.00156
CDR20291_0245	4578.3	12.2	45.8	0.0	carbamoyl-phosphate synthetase	0.0000995
tndX	4359.9	12.1	43.6	0.0	conjugative transposon site specific recombinase	0.0001683
CDR20291_1863	4275.4	12.1	42.7	0.0	hypothetical protein	0.008903
CDR20291_2299	4026.4	12.0	40.3	0.0	transposase	0.00416

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_1847	3996.4	12.0	40.0	0.0	transcriptional regulator	0.0012
CDR20291_0963	3851.0	11.9	38.5	0.0	hypothetical protein	0.00500017
spoIIIAD	3571.7	11.8	35.7	0.0	stage III sporulation protein AD	1.32997E-08
CDR20291_1425	3562.2	11.8	35.6	0.0	virulence associated protein e	9.68949E-13
CDR20291_1420	3489.5	11.8	34.9	0.0	hypothetical protein next to phage protein	1.94429E-09
CDR20291_2709	3471.3	11.8	34.7	0.0	transposase	4.22366E-18
CDR20291_1747	3469.9	11.8	34.7	0.0	hypothetical protein	2.04814E-09
CDR20291_2516	3313.6	11.7	33.1	0.0	cobalt dependent x-pro dipeptidase	1.15334E-11
CDR20291_0239	3251.0	11.7	32.5	0.0	hypothetical protein	0.0052405
CDR20291_0531	3174.3	11.6	31.7	0.0	membrane associated metalloprotease	6.84263E-08
CDR20291_0213	3041.9	11.6	30.4	0.0	hypothetical protein	6.05295E-14
CDR20291_1421	3036.7	11.6	30.4	0.0	hypothetical protein next to phage protein	1.45584E-08
CDR20291_0234	3008.4	11.6	30.1	0.0	hypothetical protein	1.23477E-12
CDR20291_3455	2961.7	11.5	29.6	0.0	hypothetical protein	3.3158E-09
CDR20291_2515	2801.5	11.5	28.0	0.0	amino acid permease family protein	2.39944E-15
CDR20291_0603	2735.6	11.4	27.3	0.0	amidohydrolases 3	6.29116E-09
CDR20291_3050	2689.6	11.4	26.9	0.0	salivaricin lantibiotic ABC transporter permease	3.28484E-17
CDR20291_0962	2570.7	11.3	25.7	0.0	hypothetical protein	0.001
CDR20291_3276	2471.1	11.3	24.7	0.0	transposase, mutator type	1.90443E-09
CDR20291_3282	2459.7	11.3	24.6	0.0	hypothetical protein	0.003039935
CDR20291_3189	2387.0	11.2	23.9	0.0	DNA-binding response regulator, lytr family	0.001985767
CDR20291_2757	2366.4	11.2	23.7	0.0	hypothetical protein	0.001
csrA	2335.3	11.2	23.3	0.0	carbon storage regulator	0.00765
CDR20291_0405	2322.0	11.2	23.2	0.0	transposase like protein b	1.80966E-14
CDR20291_1424	2303.5	11.2	23.0	0.0	DNA directed DNA polymerase	5.55112E-15
CDR20291_3049	2267.4	11.1	22.7	0.0	ABC transporter ATP binding protein	0.004321
CDR20291_3280	2256.6	11.1	980.3	0.4	hypothetical protein	3.59089E-10
CDR20291_3051	2153.9	11.1	21.5	0.0	sensor histidine kinase	3.77587E-06

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_0861	2122.5	11.1	21.2	0.0	sensor protein	0.001
CDR20291_2960	2117.0	11.0	21.2	0.0	hypothetical protein	0.0186
CDR20291_1461	2033.4	11.0	20.3	0.0	holin	0.000301694
CDR20291_0862	1994.1	11.0	19.9	0.0	ABC transporter ATPase	8.0669E-10
CDR20291_1107	1989.5	11.0	19.9	0.0	ABC transporter permease	0.005
CDR20291_2911	1900.0	10.9	19.0	0.0	restriction modification system dna specificity domain	2.51426E-07
CDR20291_3290	1890.3	10.9	18.9	0.0	transposase mutator type	1.57732E-10
CDR20291_3188	1783.9	10.8	17.8	0.0	sensor histidine kinase virs	0.003871663
CDR20291_2980	1687.9	10.7	16.9	0.0	ABC transpporter permease	1.36664E-08
CDR20291_0724	1637.7	10.7	16.4	0.0	hypothetical protein	0.000206207
CDR20291_2990	1609.3	10.7	1507.3	0.9	hypothetical protein	0.001362581
CDR20291_2335	1593.4	10.6	15.9	0.0	hypothetical protein	0.000976112
CDR20291_2758	1527.1	10.6	15.3	0.0	lantibiotic ABC transporter ATP-binding protein	1.44329E-14
CDR20291_1744	1501.6	10.6	15.0	0.0	site specific recombinase resolvase family	0.003006322
CDR20291_0045	1406.2	10.5	566.8	0.4	putative uncharacterized protein	3.62481E-06
CDR20291_1788	1397.2	10.4	1200.0	0.8	hypothetical protein	1.47658E-05
CDR20291_3010	1299.9	10.3	1627.4	1.2	hypothetical protein	2.27751E-12
CDR20291_2275	1203.6	10.2	2055.4	1.7	putative uncharacterized protein (fragment)	0.05
CDR20291_0043	1128.3	10.1	578.0	0.5	thymidylate synthase	0.05
CDR20291_0922	1065.9	10.1	463.0	0.4	hypothetical protein adjacent to ccpA	0.000812098
ispD	933.4	9.9	405.5	0.4	major intracellular serine protease	4.44089E-16
CDR20291_0382	857.7	9.7	190.6	0.2	ABC transporter permease	3.63265E-06
CDR20291_2996	828.5	9.7	711.5	0.8	hypothetical protein	0.045
spoIIIAG	794.0	9.6	344.9	0.4	stage III sporulation protein ag	2.08494E-09
CDR20291_2994	742.3	9.5	1874.3	2.5	CRISPR-associated helicase cas3	1.19438E-05
CDR20291_3278	737.6	9.5	3109.4	4.2	hypothetical protein	1.45749E-08
CDR20291_3281	719.1	9.5	1205.4	1.7	transposon Tn21 resolvase	0.003718108
CDR20291_2988	690.0	9.4	939.2	1.4	hypothetical protein	9.80101E-07
CDR20291_0924	680.8	9.4	295.7	0.4	hypothetical protein adjacent to ccpA	0.005000024

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
sleC	642.4	9.3	258.9	0.4	SleC- spore peptidoglycan hydrolase/ germinant receptor complex	0.005062748
CDR20291_2997	620.3	9.3	317.7	0.5	CRISPR-associated protein cas6	0.005
CDR20291_1465	612.1	9.3	265.9	0.4	manganese containing catalase	0.00509398
CDR20291_1760	609.4	9.3	5923.5	9.7	addiction module toxin	0.005
CDR20291_2991	599.8	9.2	1278.9	2.1	frg domain containing protein	0.0119524
CDR20291_2989	583.5	9.2	730.4	1.2	hypothetical protein	0.0345
CDR20291_2986	566.2	9.1	3457.4	6.1	hypothetical protein	3.73627E-08
CDR20291_0247	565.2	9.1	227.8	0.4	hypothetical protein	1.76775E-11
CDR20291_1415	559.7	9.1	463.2	0.8	site specific recombinase, phage integrase family	2.53081E-08
CDR20291_1464	547.1	9.1	237.6	0.4	penicillin binding protein	3.06532E-05
CDR20291_1780	529.5	9.0	970.0	1.8	hypothetical protein	1.13721E-11
CDR20291_1772	511.9	9.0	10872.7	21.2	hypothetical protein	8.3309E-07
CDR20291_1778	511.1	9.0	438.9	0.8	hypothetical protein	0.00106688
CDR20291_0711	485.3	8.9	210.8	0.4	cytidine/deoxycytidylate deaminase family protein	0.002079352
CDR20291_2998	483.7	8.9	415.4	0.8	CRISPR-associated protein cas5 family	5.27088E-05
CDR20291_2197	483.5	8.9	194.9	0.4	full- transposase	0.009394451
CDR20291_2987	480.4	8.9	790.3	1.6	hypothetical protein	0.001572206
spoIIIAH	480.2	8.9	571.1	1.2	stage III sporulation protein AH	1.13377E-06
CDR20291_3293	479.9	8.9	208.5	0.4	hypothetical protein	0.000453
CDR20291_2983	477.0	8.9	207.2	0.4	Fe3+ ABC transporter periplasmic component like protein	0.00420458
CDR20291_2524	455.0	8.8	183.4	0.4	putative membrane protein	0.0456
CDR20291_2995	453.4	8.8	987.8	2.2	CRISPR -associated auto-regulator	4.5883E-05
CDR20291_1105	452.5	8.8	3190.0	7.0	regulatory protein	0.000523642
CDR20291_2277	451.3	8.8	948.0	2.1	putative beta-lactamase repressor	3.72506E-05
CDR20291_1749	450.8	8.8	422.3	0.9	sensor protein	1.42728E-10
licB	438.3	8.8	18167.9	41.4	PTS system lichenan-specific transporter subunit IIB	4.40327E-10
CDR20291_1417	421.7	8.7	348.9	0.8	HipA like protein	0.007653

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_3461	416.9	8.7	358.1	0.8	chloramphenicol 0-acetyltransferase	3.6294E-08
CDR20291_0044	411.5	8.7	178.8	0.4	dihydrofolate reductase region	0.00223179
CDR20291_0046	392.5	8.6	645.7	1.6	thiamine biosynthesis protein	0.003978415
CDR20291_0224	368.9	8.5	148.7	0.4	glucose-1-phosphate thymidylyltransferase	0.000343975
celB	365.1	8.5	42573.8	116.6	PTS system lichenan-specific transporter subunit IIB	4.35425E-05
CDR20291_1447	360.6	8.5	1091.6	3.0	hypothetical protein next to phage protein	0.00345
CDR20291_1429	350.1	8.5	1643.4	4.7	hypothetical protein	0.004259837
CDR20291_3457	348.7	8.4	151.5	0.4	hypothetical protein	1.0204E-09
CDR20291_0501	348.2	8.4	151.2	0.4	hypothetical protein	3.79854E-07
CDR20291_1779	348.2	8.4	1065.0	3.0	hypothetical protein	0.000992012
CDR20291_3469	333.9	8.4	145.1	0.4	transcriptional regulator	4.53873E-09
CDR20291_1360	332.0	8.4	1443.2	4.3	hypothetical protein	1.03995E-07
CDR20291_3286	309.5	8.3	435.6	1.4	hypothetical protein	1.17839E-06
spoVT	309.1	8.3	3883.9	12.6	Stage V sporulation protein T	9.59726E-10
CDR20291_1777	308.8	8.3	158.2	0.5	hypothetical protein	8.04335E-05
spoIVA	307.2	8.3	9469.8	30.8	stage IV sporulation protein A	2.35729E-06
sspB	305.5	8.3	795.2	2.6	small acid-soluble spore protein B	5.95351E-08
CDR20291_2276	294.9	8.2	1823.4	6.2	putative beta-lactamase inducer	0.04521
CDR20291_2145	290.4	8.2	148.7	0.5	hypothetical protein	0.002948414
CDR20291_2908	286.3	8.2	124.4	0.4	hypothetical protein	0.001278356
CDR20291_1072	268.8	8.1	59.7	0.2	integrase, catalytic region	0.004532
flgM	259.3	8.0	132.8	0.5	Negative regulator of flagellin biosynthesis	0.00018669
CDR20291_2196	257.1	8.0	103.6	0.4	integrase, catalytic region	8.07941E-07
CDR20291_0439	255.1	8.0	102.8	0.4	sensor histidine kinase	0.000000067
CDR20291_0526	239.4	7.9	623.1	2.6	hypothetical protein	0.000616459
CDR20291_3456	230.1	7.8	117.9	0.5	hypothetical protein	8.74396E-10
CDR20291_2183	203.1	7.7	183.9	0.9	beta-lactamase inducer	4.02335E-08
sspA	196.8	7.6	8752.3	44.5	small acid-soluble protein A	5.77836E-07
CDR20291_1826	188.1	7.6	75.8	0.4	site specific recombinase	3.53909E-07

Gene Name	**Real Fold Change**	**Log2 fold change**	**Parent-Mean expression**	**R20291 *sinRR'* Mean expression**	**Function (known or predicted)**	**adj. *p*-Value**
CDR20291_1279	183.5	7.5	321.8	1.7	exported polysaccharide deacetylase	6.93525E-08
dacF	181.0	7.5	1039.9	5.7	D-alanyl-D-alanine carboxypeptidaase	7.12346E-07
CDR20291_3140	180.3	7.5	92.3	0.5	PTS system transporter subunit IIB	2.92249E-07
CDR20291_2019	175.8	7.5	3880.9	22.1	hypothetical protein	0.000176891
CDR20291_1606	170.2	7.4	146.1	0.8	C4 dicarboxylate anaerobic carrier	0.000470897
CDR20291_3511	156.7	7.3	201.0	1.3	peptidase	0.008533382
CDR20291_1324	155.1	7.3	133.2	0.8	Amino acid aminotransferase	1.8639E-05
flgL	151.9	7.2	567.0	3.7	flagellar hook-associated protein	3.45645E-09
CDR20291_3387	146.5	7.2	63.6	0.4	hypothetical protein	1.33227E-15
celC	143.7	7.2	6472.3	45.0	PTS system lichenan-specific transporter subunit IIA	0.007180793
CDR20291_1196	139.7	7.1	255.9	1.8	hypothetical protein	2.90007E-09
cotJC2	139.6	7.1	130.8	0.9	cotJC2 - spore coat protein	0.056
CDR20291_0130	138.0	7.1	172.7	1.2	hypothetical protein	0.002934188
CDR20291_2985	133.7	7.1	348.0	2.6	transcriptional regulator	0.000576333
celF	129.3	7.0	49526.4	383.0	6-phospho beta gluosidase - chitobiose degradation	0.000978006
spoIIIAF	119.5	6.9	51.9	0.4	stage III sporulation protein AF	0.008257844
CDR20291_2984	116.9	6.9	50.8	0.4	hypothetical protein	0.003456
CDR20291_1773	116.7	6.9	50.7	0.4	hypothetical protein	1.27233E-06
CDR20291_2777	115.5	6.9	17004.1	147.2	hypothetical protein	0.00786
CDR20291_2964	114.4	6.8	46.1	0.4	transcriptional regulator	7.21645E-15
CDR20291_1924	113.8	6.8	49.4	0.4	hypothetical protein	1.78968E-12
CDR20291_2289	102.9	6.7	128.8	1.2	hypothetical protein	0.000740096
CDR20291_0926	101.6	6.7	8447.7	83.2	hypothetical protein- gene adjacent to *ccpA*	1.05863E-06
CDR20291_1073	98.9	6.6	265.2	2.7	hypothetical protein in 1070 operon	0.00821132
CDR20291_1335	98.9	6.6	557.3	5.6	lipoprotein	0.008961977
CDR20291_1338	98.2	6.6	32.6	0.3	transposase like protein b (pseudogene)	0.001404359
CDR20291_2523	95.3	6.6	408.5	4.3	putative membrane protein	1.36557E-14

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_1110	94.7	6.6	38.2	0.4	two component sensor histidine kinase	0.007905788
CDR20291_2698	94.1	6.6	481.7	5.1	conserved hypothetical protein	0.003929811
CDR20291_3139	94.0	6.6	85.1	0.9	PTS system transporter subunit IIA	0.004483046
CDR20291_0919	93.6	6.5	206.7	2.2	hypothetical protein adjacent to ccpA	0.000476639
fliS1	92.5	6.5	238.0	2.6	flagellarprotein FliS	6.28391E-05
spoIIIAB	90.8	6.5	72.3	0.8	stage III sporulation protein AB	8.33334E-06
CDR20291_0569	89.2	6.5	232.3	2.6	fic family protein	2.91592E-05
fliD	88.8	6.5	760.8	8.6	flagellar cap protein	0.004196392
CDR20291_1142	88.7	6.5	38.5	0.4	hypothetical protein	0.00031591
tcdB	88.4	6.5	2942.4	33.3	ToxinB	5.80669E-12
CDR20291_1214	87.0	6.4	37.8	0.4	phage protein	0.002643783
spoVAC	86.6	6.4	34.9	0.4	stage V sporulation protein AC	0.000105562
spoVAD	84.5	6.4	172.2	2.0	stage V sporulation protein AD	0.000161277
CDR20291_0316	81.8	6.4	949.9	11.6	spore coat assembly asparagine rich protein	0.000577483
CDR20291_0225	81.0	6.3	35.2	0.4	dtdp- r dehydrorhamnose 3,5 epimerase	4.5E-09
CDR20291_3359	80.3	6.3	100.5	1.2	hypothetical protein	2.22045E-16
CDR20291_3187A	76.6	6.3	33.2	0.4	autoinducer prepeptide	0.009879222
CDR20291_1558	76.3	6.3	329.4	4.3	hypothetical protein	4.33826E-08
slpA	70.4	6.1	293272.2	4165.7	S layer precursor protein	0.007561607
CDR20291_1623	70.1	6.1	122.9	1.7	hypothetical protein next to AraC family transcriptional regulator	0.000387775
CDR20291_1277	69.3	6.1	35.5	0.5	hypothetical protein	0.002035128
CDR20291_0241	68.6	6.1	89.1	1.3	glycosyltransferase	0.000203367
spoIIID	68.0	6.1	262.6	3.9	stage III sporulation protein D	6.24846E-05
cwlD	67.7	6.1	200.7	3.0	germination specific N acetylmuramoyl L ananine amidase	1.01834E-09
CDR20291_1740	66.1	6.0	84.8	1.3	putative RNA methyltransferase	0.000730506
flgK	65.2	6.0	504.5	7.7	flagellar hook -associated protein	3.78766E-05
CDR20291_1308	64.5	6.0	10033.1	155.5	5 nitroimadazole reductase	0.005749942

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_0156	63.4	6.0	281.4	4.4	hypothetical protein	0.002928365
CDR20291_0713	63.0	6.0	27.3	0.4	hypothetical protein	0.006668258
CDR20291_1852	62.9	6.0	157.8	2.5	hypothetical protein	6.74842E-11
CDR20291_0231	59.1	5.9	23.8	0.4	flagellar biosynthesis protein	6.97553E-13
CDR20291_2522	57.7	5.9	47.7	0.8	conserved hypothetical protein	1.11717E-05
gltC	57.5	5.8	49.4	0.8	sodium/glutamate symporter	0.008778122
CDR20291_0124	56.9	5.8	6022.8	105.8	cell wall endopeptidase	0.05
CDR20291_1106	55.5	5.8	69.4	1.2	ABC transporter permease	4.42552E-05
CDR20291_3159	54.8	5.8	113.4	2.1	ATP/GTP binding protein	3.56382E-14
CDR20291_2697	53.6	5.7	1341.9	25.0	putative membrane protein	2.75093E-05
CDR20291_1138	51.1	5.7	262.2	5.1	hypothetical protein	0.000258498
CDR20291_3136	50.2	5.6	1427.6	28.4	phosphosugar isomerase	2.05209E-11
CDR20291_0128	50.2	5.6	88.0	1.7	hypothetical protein	0.00533707
CDR20291_0958	49.5	5.6	1759.7	35.5	AraC family transcriptional regulator	2.28781E-06
rbr	48.2	5.6	1159.3	24.0	ruberythrin	2.08671E-10
CDR20291_0925	47.2	5.6	247.4	5.2	hypothetical protein adjacent to ccpA	0.05
CDR20291_3477	47.2	5.6	15621.1	331.3	potassium porter antitransporter	5.97413E-10
CDR20291_0610	47.1	5.6	445.3	9.4	ATP dependent peptidase	0.0000789
CDR20291_2050	45.6	5.5	36.3	0.8	transcriptional regulator	1.79956E-12
CDR20291_1210	45.6	5.5	575.3	12.6	hypothetical protein	1.35473E-07
CDR20291_1137	45.2	5.5	84.2	1.9	hypothetical protein	4.98532E-05
tcdE	44.2	5.5	40.0	0.9	TcdE-Holin like protein	6.57037E-05
sigG	44.0	5.5	1012.3	23.0	sporulation sigma factor G	2.10125E-11
CDR20291_3137	43.8	5.5	266.4	6.1	PTS system transporter subunit IID	0.05
CDR20291_2517	42.2	5.4	18.3	0.4	putative transcriptional regulator	1.24533E-07
CDR20291_1392	41.7	5.4	1031.6	24.7	acyl carrier protein phosphodiesterase	3.33067E-16
CDR20291_2755	41.4	5.4	285.0	6.9	esterase/halogenase	1.84716E-06
spmB	41.4	5.4	56.3	1.4	spore maturation protein B	0.00081103
CDR20291_1216	40.4	5.3	163.4	4.0	phage protein	7.67533E-09
CDR20291_2300	38.2	5.3	217.3	5.7	hypothetical protein	0.00054917

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2732	36.8	5.2	322.1	8.8	amidohydrolase	0.001699346
CDR20291_3138	36.2	5.2	376.2	10.4	PTS system transporter subunit IIC	1.8516E-10
spoIIE	36.1	5.2	576.8	16.0	stage II sporulation protein e	7.04757E-05
CDR20291_2872	36.0	5.2	5150.3	143.0	putative regulatory protein	0.000000965
CDR20291_3116	35.3	5.1	188.0	5.3	iron only dehydrogenase	1.59904E-05
CDR20291_2048	33.6	5.1	69.6	2.1	D-alanine D-alanine carboxypeptidase	1.67009E-11
CDR20291_1789	33.1	5.0	120.4	3.6	conjugative transposon membrane protein	0.05
tcdR	32.9	5.0	57.7	1.7	Toxin genes specific sigma factor	1.11E-16
CDR20291_1217	32.1	5.0	81.5	2.5	phage tail fiber protein	0.009847647
CDR20291_1215	31.9	5.0	123.0	3.9	phage protein	1.60813E-10
fliS2	31.7	5.0	32.2	1.0	flagellar protein	0.006963256
CDR20291_0468	29.6	4.9	90.9	3.1	hypothetical protein	0.008205754
CDR20291_0343	29.4	4.9	23.4	0.8	two component sensor histidine kinase	0.05
CDR20291_2233	29.2	4.9	5444.6	186.3	hypothetical protein	2.3222E-05
gpr	28.8	4.8	336.2	11.7	germination protease	1.83321E-07
sodA	28.4	4.8	64.1	2.2	**superoxide dismutase**	0.000179593
fliC	28.2	4.8	7484.8	265.0	Flagellin subunit	0.00382371
CDR20291_1654	27.4	4.8	34.3	1.2	hypothetical protein next to radical SAM	5.67298E-07
CDR20291_0688	26.5	4.7	149.6	5.6	**Calcium/sodium antiporter**	9.49138E-09
CDR20291_2574	24.8	4.6	20.6	0.8	putative membrane protein	0.0034
CDR20291_0342	24.6	4.6	32.7	1.3	ABC transporter permease	0.000456567
CDR20291_1684	24.5	4.6	22.9	0.9	hypothetical protein	3.16225E-12
CDR20291_1917	24.4	4.6	29.8	1.2	hypothetical protein	0.00054321
CDR20291_2204	24.3	4.6	22.7	0.9	hypothetical protein	5.21345E-08
CDR20291_3095	24.1	4.6	354.7	14.7	single-strand DNA binding potein	0.00654
cat1	23.4	4.5	10542.1	450.3	succinyl coA transferase - TCA cycle VII acetate producers	0.05
CDR20291_2028	23.3	4.5	70.9	3.0	hypothetical protein	1.84826E-08
CDR20291_1810	22.7	4.5	2189.6	96.6	hypothetical protein	4.4526E-07
CDR20291_0319	22.5	4.5	149.4	6.6	hypothetical protein	1.25608E-05

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2121	22.4	4.5	9.7	0.4	SinR	0.006
CDR20291_1907	22.3	4.5	102.9	4.6	hypothetical protein	2.34989E-07
CDR20291_1795	22.2	4.5	394.3	17.8	helicase	5.24123E-05
CDR20291_1446	22.1	4.5	46.8	2.1	prophage antirepressor related protein	5.04186E-09
CDR20291_2867	21.8	4.4	50.3	2.3	aminotransferase	0.00854
CDR20291_1444	21.6	4.4	9.4	0.4	phage protein	2.11888E-08
CDR20291_1207	21.5	4.4	270.5	12.6	hypothetical protein	1.05732E-11
spmA	21.5	4.4	46.1	2.1	spore maturation protein A	2.80032E-12
CDR20291_1649	21.4	4.4	384.5	18.0	ABC transporter permease	4.38978E-07
CDR20291_2229	21.0	4.4	3605.3	171.8	hypothetical protein	0.007854
CDR20291_1205	20.4	4.3	155.4	7.6	hypothetical protein	3.3036E-09
cspBA	19.5	4.3	303.3	15.6	germinant receptor complex cold shock protein	7.88015E-07
CDR20291_2128	19.5	4.3	50.4	2.6	hypothetical protein	5.69054E-07
CDR20291_0884	19.2	4.3	42.0	2.2	signaling protein	9.88098E-15
CDR20291_2261	18.7	4.2	184.6	9.8	hypothetical protein	0.000684158
CDR20291_1282	18.7	4.2	178.5	9.5	bifunctional protein peroxiredoxin/chitinase	0.000291694
CDR20291_3404	18.3	4.2	2014.6	109.9	cell wall hydrolase	8.71277E-05
CDR20291_1218	18.1	4.2	703.9	38.9	hypothetical protein in phage protein operon	0.001919924
CDR20291_1622	18.0	4.2	15.5	0.8	AraC family transcriptional regulator	0.000354597
CDR20291_2075	18.0	4.2	3868.4	215.2	iron-sulfur binding protein	0.00954
CDR20291_1077	17.9	4.2	36.0	2.0	hypothetical protein next to 1070 operon	0.00018875
sucD	17.8	4.2	14367.4	805.5	succinate semialdehyde dehydrogenase (NADP) - 4 aminobutyrate degradation II	1.08354E-08
CDR20291_1434	17.8	4.2	128.5	7.2	phage protein	7.05055E-07
CDR20291_1108	17.7	4.1	14.6	0.8	ABC transporter ATP binding protein	0.002315528
CDR20291_2049	17.4	4.1	83.8	4.8	hypothetical protein next to D alanine D alanine carboxypeptidase	2.65088E-06

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2781	17.4	4.1	473.6	27.2	GntR family transcriptional regulator	1.42924E-05
CDR20291_2756	17.2	4.1	17.4	1.0	bacterioferritin	8.54594E-05
CDR20291_1648	17.2	4.1	237.3	13.8	ABC transporter permease	1.95091E-05
CDR20291_1211	17.0	4.1	37.9	2.2	hypothetical protein	4.83821E-10
CDR20291_2184	16.9	4.1	488.2	28.9	GntR family transcriptional regulator	0.000554059
CDR20291_2236	16.8	4.1	2928.5	174.2	Xaa pro dipeptidase	0.000413823
CDR20291_1954	16.8	4.1	2596.0	154.8	hypothetical protein	0.00654
abfD	16.5	4.0	19664.9	1194.5	hypothetical protein	3.33567E-12
cspC	16.3	4.0	145.8	8.9	CspC germinant receptor complex	0.009214675
CDR20291_1212	16.2	4.0	429.8	26.5	phage cell wall hydrolase	4.05723E-08
CDR20291_3154	15.3	3.9	72.5	4.7	hypothetical protein	0.008754
CDR20291_3094	15.2	3.9	359.9	23.6	hypothetical protein	0.000306378
CDR20291_3497	15.2	3.9	105.3	6.9	hypothetical protein	9.73708E-10
spoIIP	15.1	3.9	285.3	18.9	stage II sporulation protein P	2.67203E-11
CDR20291_0381	15.0	3.9	1565.8	104.4	cell surface protein	0.000223243
CDR20291_1206	14.9	3.9	245.2	16.5	hypothetical protein	3.10521E-06
CDR20291_2500	14.8	3.9	39.2	2.6	putative membrane protein	0.000160056
flgB	14.6	3.9	101.3	6.9	flagellar basal body rod protein	0.0009876
CDR20291_2859	14.6	3.9	14.8	1.0	hypothetical protein	0.00065
CDR20291_1075	14.5	3.9	22.0	1.5	hypothetical protein next to 1070 operon	4.24214E-05
fliA	14.4	3.8	31.4	2.2	FliA/SigD/RNA polymerase sigma factor for flagellar operon	4.13806E-08
CDR20291_1141	14.3	3.8	13.4	0.9	hypothetical protein	0.002355205
CDR20291_3400	14.0	3.8	167.7	11.9	spore cortex-lytic enzyme	5.7983E-09
spoIIIAE	13.9	3.8	29.8	2.1	spoIIIAE- stage III sporulation protein AE	1.48072E-09
CDR20291_2262	13.9	3.8	230.6	16.6	hypothetical protein	0.000882324
motA	13.8	3.8	280.7	20.3	chemotaxis protein	0.000310215
abfT	13.8	3.8	11390.8	827.0	4 hydroxybutyrate CoA transferase	0.002
CDR20291_1905	13.5	3.8	107.7	8.0	transcriptional regulator	7.33545E-05
tcdA	13.1	3.7	18604.0	1415.3	Toxin A	0.000304519

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_1824	13.1	3.7	37.5	2.8	ABC transporter permease	0.005
comE	13.1	3.7	47.7	3.6	competence protein	0.007291591
bclA3	13.1	3.7	818.0	62.5	bclA3 putative exosporium glycoprotein	0.00543
CDR20291_1918	12.7	3.7	244.0	19.2	decarboxylase (valine degradation II) 2-oxoglutarate decarboxylation to succinyl-coA	0.000125987
CDR20291_2495	12.4	3.6	5340.1	429.1	putative amino acid permease	0.000138258
CDR20291_1208	12.4	3.6	155.9	12.6	hypothetical protein	0.009814884
def1	12.4	3.6	681.6	55.0	peptide deformylase I	1.12773E-08
CDR20291_2784	12.2	3.6	54.5	4.5	ABC transporter permease	0.0005643
CDR20291_0600	12.2	3.6	70.8	5.8	RNA polymerase sigma factor sigT	2.39808E-14
CDR20291_0499	12.1	3.6	2559.6	210.7	Hypothetical	8.61187E-11
grpE	12.1	3.6	23360.5	1928.1	heat shock protein	0.002950212
pflE	12.1	3.6	1886.6	156.5	pyruvate formate-lyase 3 activating enzyme	0.002097911
CDR20291_1160	12.0	3.6	401.2	33.4	polysaccharide deacetylase	0.000448404
CDR20291_1646	11.8	3.6	1386.3	117.2	**TetR family transcriptional regulator**	3.37369E-05
CDR20291_0618	11.5	3.5	492.8	42.9	hypothetical protein	4.77396E-15
CDR20291_1667	11.4	3.5	134.5	11.8	hypothetical protein	1.29878E-05
CDR20291_2165	11.3	3.5	242.4	21.4	oxidoreductase	0.007543
abfH	10.9	3.5	7838.0	716.1	NAD-dependent 4-hydroxybutyrate dehydrogenase	0.000434735
CDR20291_1650	10.9	3.5	533.8	48.8	transcriptional regulator	8.41549E-14
CDR20291_3418	10.8	3.4	82.9	7.7	hypothetical protein	6.211E-10
fliJ	10.8	3.4	60.8	5.6	flagellar protein	5.41685E-05
uppS	10.7	3.4	118.2	11.0	UDP pyrophosphate synthetase	2.53579E-10
pflD	10.4	3.4	18604.6	1781.6	mixed acid fermentation respiration	1.60974E-09
flgE	10.4	3.4	377.8	36.5	flagellar hook protein	0.006156236
CDR20291_1647	10.3	3.4	887.0	86.1	ABC transporter ATP binding protein	2.89688E-06
CDR20291_0400	10.1	3.3	32.5	3.2	ABC transporter ATP binding protein	0.005643
CDR20291_1354	10.1	3.3	679.8	67.3	drug / sodium antiporter	6.06958E-08

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2575	10.0	3.3	26.5	2.6	hypothetical protein	2.37797E-16
CDR20291_3115	9.9	3.3	54.9	5.6	polysaccharide deacetylase	9.97943E-22
CDR20291_3152	9.8	3.3	38.5	3.9	hypothetical protein	4.96686E-16
fleN	9.8	3.3	80.9	8.2	flageller number regulator	1.61961E-15
CDR20291_2161	9.7	3.3	356.9	36.8	hypothetical protein	0.001752083
fliL	9.7	3.3	62.0	6.4	fliL flageller basel body - associated protein	0.0006754
CDR20291_3155	9.6	3.3	37.4	3.9	type IV pillin	0.00034
CDR20291_1312	9.5	3.3	999.2	104.7	hypothetical protein	0.000045
anmK	9.5	3.3	72.4	7.6	anhydro-N-acetylmuramic acid kinase	1.08435E-12
CDR20291_1683	9.5	3.2	8.9	0.9	hypothetical protein	0.000166102
CDR20291_3326	9.4	3.2	299.0	31.7	oligopeptidase	2.92818E-05
CDR20291_2491	9.4	3.2	656.7	69.7	cdta (adp-ribosyltransferase enzymatic component)	0.007096445
CDR20291_2260	9.0	3.2	756.0	83.5	carbon starvation protein	0.000056
CDR20291_3080	9.0	3.2	63.8	7.1	small acid-soluable spore protein	6.98416E-06
CDR20291_1342	9.0	3.2	1111.8	123.9	two-component sensor histidine kinase	0.00056
CDR20291_2151	8.9	3.2	4884.4	548.1	hypothetical protein	4.05212E-05
CDR20291_1920	8.9	3.2	29.9	3.3	ArsR family transcriptional regulator	7.93315E-06
CDR20291_0228	8.4	3.1	29.9	3.5	hypothetical protein	0.000454331
bclA2	8.4	3.1	49.5	5.9	exosprium glycoprotein	0.00288464
CDR20291_2337	8.4	3.1	59.4	7.1	hypothetical protein	4.55191E-15
cwp66	8.3	3.1	3253.9	391.4	celll surface protein	0.000092274
CDR20291_3153	8.2	3.0	95.2	11.5	hypothetical protein	0.000102489
CDR20291_1389	8.2	3.0	13.8	1.7	ABC transporter permease	0.000131599
CDR20291_1076	8.2	3.0	163.1	19.9	Hypothetical	0.00346
CDR20291_0969	8.1	3.0	194.9	24.0	hypothetical protein	0.00017306
CDR20291_2323	8.1	3.0	54.1	6.7	nitrite sulfite reductase (alkynitronates degradation)	2.1678E-06
fliK	8.0	3.0	181.0	22.5	flagellar hook-length control protein	0.004
CDR20291_0718	8.0	3.0	42.5	5.3	metal-binding protein	1.72429E-12

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2533	8.0	3.0	1336.7	166.4	putative extracellular solute-binding protein	6.53582E-06
CDR20291_0532	8.0	3.0	66.9	8.4	hypothetical protein	0.005581262
CDR20291_2857	8.0	3.0	28.2	3.5	hypothetical protein	0.000500209
CDR20291_0433	7.9	3.0	49.5	6.3	sugar-phosphate dehydrogenase	6.07594E-06
dpaL2	7.8	3.0	1283.9	164.2	diaminopropinate ammmonia-lysase	0.0003421
CDR20291_0440	7.6	2.9	1427.5	188.3	hemagglutinin /adhesin	5.20906E-11
CDR20291_0561	7.5	2.9	280.6	37.3	hypothetical protein	0.000959879
CDR20291_2120	7.4	2.9	58.5	7.9	carbonic anhdrase (gluconeogenesis II before sinR)	0.000727506
vanR	7.3	2.9	1801.6	248.4	two-component response regulator	0.005
CDR20291_0495	7.2	2.8	237.4	33.1	hypothetical protein	0.0432
hisG	7.1	2.8	24.2	3.4	ATP phosphoribosyltransferase	7.62501E-13
CDR20291_1343	7.1	2.8	340.7	47.7	3-methyladenine DNA glycosylase	0.000550671
flgC	7.1	2.8	109.1	15.3	flageller basal body rod protein	0.04521
CDR20291_0938	7.1	2.8	820.8	115.2	LysR family transcriptional regulator	0.000127298
CDR20291_2874	7.1	2.8	2726.4	383.6	hypothetical protein	0.000453152
proC2	7.0	2.8	540.8	76.8	pyrroline-5-carboxylate reductase	4.38701E-06
hrcA	7.0	2.8	6533.9	933.0	heat-inducible transcription repressor	0.002605907
flgD	7.0	2.8	187.2	26.9	basal body rod modification protein	1.44644E-07
CDR20291_1224	6.9	2.8	50.3	7.3	hypothetical protein	2.53384E-09
CDR20291_1653	6.8	2.8	581.8	85.1	radical SAM protein	0.00442828
CDR20291_2419	6.8	2.8	1236.3	182.4	aminotransferase	0.002344517
CDR20291_3158	6.7	2.8	406.1	60.4	hypothetical protein	1.02128E-05
CDR20291_1353	6.7	2.7	12.8	1.9	ABC transporter permease	0.001568147
fliI	6.7	2.7	616.5	92.3	flagellum-specific ATP synthase	0.000191876
motB	6.6	2.7	234.2	35.3	chemotaxis protein	0.001745817
CDR20291_0559	6.6	2.7	22.7	3.4	acetyltransferase	1.12012E-07
tdcF	6.5	2.7	738.7	113.7	regulatory endoribonuclease	3.29323E-07

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_1652	6.5	2.7	83.0	12.8	hypothetical protein next to def1	0.0321
CDR20291_3350	6.5	2.7	874.3	135.5	pillin	7.80931E-13
CDR20291_2166	6.4	2.7	8.5	1.3	membrane-associated caaX amino terminal protease	0.000313031
CDR20291_1104	6.4	2.7	4011.9	627.0	hypothetical protein before regulatory protein	1.77636E-15
CDR20291_2089	6.4	2.7	124.1	19.4	n-acetylmuramoyl-alanine amidase	0.00214281
CDR20291_0263	6.4	2.7	35.6	5.6	flagellar protein	1.29506E-05
CDR20291_0341	6.3	2.7	31.8	5.0	ABC transporter ATP binding protein	6.53333E-12
CDR20291_2097	6.3	2.7	115.4	18.2	phosphoesterase	2.34845E-09
vanS	6.3	2.7	1629.7	258.2	two-component sensor histidine kinase	0.002781938
spoIVB	6.2	2.6	34.0	5.5	spoIVB Stage IV sporulation protein B	6.61403E-06
flhA	6.2	2.6	265.5	43.0	flhA flagellar biosynthesis protein FlhA	0.000120411
CDR20291_2608	6.1	2.6	1155.9	189.9	polysaccharide deacetylase	3.24702E-09
CDR20291_2675	6.0	2.6	1959.5	325.2	cell surface protein	6.18838E-13
CDR20291_2724	6.0	2.6	161.0	26.8	transport related ATPase	0.000945575
fliG	6.0	2.6	147.3	24.6	fliG flagellar motor switch protein	1.74834E-05
CDR20291_0599	6.0	2.6	106.3	17.8	hypothetical protein	0.008077949
fliH	6.0	2.6	308.2	51.7	fliH flagellar assembly protein	3.04518E-08
CDR20291_2269	5.9	2.6	37332.9	6341.5	putative aspartate aminotransferase (aspartate biosynthesis aspartate degradation I)	1.14571E-07
CDR20291_3156	5.9	2.6	151.3	25.8	type IV pilus-assembly protein	2.92397E-07
CDR20291_2058	5.8	2.5	273.3	47.1	glutamyl-aminopeptidase	2.14619E-05
CDR20291_1244	5.8	2.5	55.7	9.7	peptidase	5.20509E-11
CDR20291_0001	5.7	2.5	67.8	11.8	small-molecule-binding protein	4.2845E-05
CDR20291_1550	5.7	2.5	1550.0	272.6	telluirum resistance protein	0.0003722
CDR20291_1067	5.7	2.5	100.3	17.7	penicillin-binding protein	0.003438827
CDR20291_1005	5.7	2.5	33.5	5.9	hypothetical protein	5.9337E-06
CDR20291_2111	5.6	2.5	426.4	75.6	esterase	9.95705E-07

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_1431	5.6	2.5	38.1	6.8	phage DNA-binding protein	0.000403526
CDR20291_0663	5.6	2.5	8447.7	1514.1	hypothetical protein	6.05397E-07
CDR20291_0435	5.6	2.5	17.6	3.1	PTS system transporter subunit IIB	0.000172933
CDR20291_0363	5.6	2.5	1305.8	234.2	radical SAM protein (heme biosynthesis from uroporphyrinogen-III II anaerobic)	0.001381718
fliE	5.5	2.5	72.7	13.1	fliE flagellar hook-basal body protein	8.68438E-07
CDR20291_1264	5.5	2.5	4904.6	883.9	ATP-binding protein	0.0040436
CDR20291_0885	5.5	2.5	66.4	12.0	cell wall anchored protein	0.000876
CDR20291_2188	5.5	2.5	489.1	88.8	two-component response regulator	6.2611E-12
CDR20291_1890	5.4	2.4	382.4	70.2	hypothetical protein	2.40715E-08
CDR20291_1750	5.4	2.4	15.5	2.9	lantibiotic ABC transporter ATP-binding protein	0.000129663
fliF	5.4	2.4	490.0	90.8	fliF flagellar MS-ring protein	0.003635569
CDR20291_0200	5.4	2.4	26.9	5.0	hypothetical protein	1.25675E-10
flhF	5.3	2.4	84.9	15.9	flagellar biosynthesis regulator	0.000896
CDR20291_2417	5.3	2.4	266.1	50.3	hypothetical protein	0.000860408
CDR20291_0774	5.3	2.4	5267.0	996.6	cell surface protein	0.003498
CDR20291_1932	5.2	2.4	294.4	56.2	transporter	0.001152589
CDR20291_3448	5.2	2.4	1886.0	360.3	hypothetical protein	9.87986E-10
CDR20291_3039)	5.2	2.4	816.8	156.6	peptidase	4.59228E-09
flhB	5.2	2.4	168.6	32.4	bifunctional flagellar biosynthesis protein FliR/FlhB	0.0008654
pbuX	5.2	2.4	365.0	70.4	Xanthine permase x 2	0.000161486
CDR20291_1800	5.2	2.4	53.7	10.4	conjugative transposon mobilization protein	5.36631E-06
ldh	5.2	2.4	103.5	20.1	L-lactate dehydrogenase (pyruvate fermentation to lactate heterlactic fermentation L-lactaldehyde degradation (Aerobic)	0.00051385
modA	5.1	2.4	910.2	177.2	molybdenum ABC transporter substrate-binding protein	0.001156081
CDR20291_0401	5.1	2.4	111.7	21.8	ABC transporter permease	1.42344E-10

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_3300	5.1	2.3	1096.7	215.8	lipoportein - GerS - germination regulator	0.009228505
CDR20291_0886	5.1	2.3	32.2	6.3	glycosyl transferase	4.73107E-05
CDR20291_1480	5.0	2.3	1687.9	335.9	hypothetical protein	2.39762E-09
CDR20291_1672	5.0	2.3	726.8	145.2	arsenate reductase	1.91975E-05
gatA	5.0	2.3	1060.7	212.1	PTS system galactitol-specific transporter subunit-IIA	0.002663847
CDR20291_1971	5.0	2.3	272.7	54.6	hypothetical protein	1.52737E-07
CDR20291_1243	5.0	2.3	210.3	42.2	hypothetical protein before peptidase	2.18426E-06
CDR20291_1931	5.0	2.3	45.8	9.2	AraC family transcriptional regulator	1.9528E-08
CDR20291_2796	5.0	2.3	109.1	21.9	hypothetical protein	3.78905E-08
CDR20291_1625	4.9	2.3	71.5	14.6	hypothetical protein next to AraC family transcriptional regulator	0.00007654
CDR20291_1911	4.9	2.3	32.7	6.6	cell surface protein	1.88196E-07
CDR20291_0661	4.9	2.3	2929.8	597.2	hypothetical protein	7.654E-07
CDR20291_0511	4.9	2.3	531.1	108.4	hypothetical protein	0.000133749
cspD	4.9	2.3	616.1	126.0	cold shock protein/germinant receptor complex	5.18201E-06
fliP	4.9	2.3	104.6	21.4	fflagellar biosynthesis protein	4.81837E-14
cbiM	4.8	2.3	8876.1	1831.1	cobalt transport protein	3.75631E-09
veg	4.8	2.3	48789.8	10139.5	hypothetical protein	1.10797E-05
CDR20291_1126	4.8	2.3	1850.8	387.8	hypothetical protein	0.000765
CDR20291_0172	4.7	2.2	784.1	165.4	hypothetical protein	0.000161211
CDR20291_0859	4.7	2.2	98.7	20.8	hypothetical protein	0.001732335
CDR20291_3141	4.7	2.2	147.5	31.1	transcriptional antiterminator	0.000208197
CDR20291_1881	4.7	2.2	2116.1	448.6	redyctive dehalogenase	3.71947E-05
CDR20291_0901	4.7	2.2	50.3	10.7	Hypothetical	9.10006E-10
ctsR	4.7	2.2	2246.6	479.7	transcriptional regulator	7.6543E-06
ddl	4.7	2.2	36.3	7.8	D-alanine D alanine ligase (UDP N acetylmuramoyl pentapeptide biosynthesis I (meso-DAP-containing) and lysie containing	1.22349E-06
CDR20291_2674	4.7	2.2	47.5	10.2	putative membrane protein	6.50806E-09

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_0770	4.7	2.2	2841.6	609.1	hydrolase	5.64422E-05
CDR20291_3496	4.6	2.2	51.3	11.0	hypothetical protein	0.007947725
CDR20291_0719	4.6	2.2	1537.3	333.3	ATP-GTP-bining protein	3.77476E-15
CDR20291_0434	4.6	2.2	39.4	8.6	PTS system transporter subunit-IIA	0.000054
CDR20291_0825	4.6	2.2	1377.9	300.6	metallo-beta-lactamase	2.88158E-05
CDR20291_1725	4.6	2.2	115.4	25.2	cation transport-related membrane protein	2.38744E-07
CDR20291_0481	4.6	2.2	247.0	54.1	endonuclease	2.00835E-05
CDR20291_2783	4.5	2.2	207.5	45.7	signaling protein	0.000254698
CDR20291_0676	4.5	2.2	2750.2	605.3	signaling protein	3.61286E-05
CDR20291_3037	4.5	2.2	406.4	89.8	chlorohydrolase/aminohydrolase	0.007620035
CDR20291_0199	4.5	2.2	484.3	108.1	membrane-associated nucleotidase (adenosine nucleotide degradation II guanosine nucleotides degradation III urate biosynthesis /inosine 5'phosphate degradation)	1.70974E-14
ogt2	4.5	2.2	35.8	8.0	methylated -DNA-portein cysteine methyltransferase 2	4.32987E-15
CDR20291_0189	4.5	2.2	2870.0	640.7	ketopantoate reductase	0.000718099
CDR20291_0916	4.5	2.2	227.1	50.7	pseudoridylate synthase	0.000595582
CDR20291_3008	4.5	2.2	100.0	22.5	hypothetical protein	3.83831E-10
CDR20291_2194	4.4	2.2	400.0	90.0	hypothetical protein	1.67338E-07
CDR20291_0567	4.4	2.2	27.0	6.1	lantibiotic ABC transporter ATP-binding protein	0.000180802
sat	4.4	2.2	601.4	135.4	streptogramin A acetyltransferase	8.48546E-05
CDR20291_0283	4.4	2.1	1568.5	354.6	hypothetical protein	0.000367792
CDR20291_3157	4.4	2.1	842.1	190.5	type IV pilus-assembly protein	0.00703431
gatB	4.4	2.1	6526.2	1482.7	PTS system galactitol-specific transporter subunit-IIA	1.61016E-05
CDR20291_0437	4.4	2.1	86.3	19.6	PTS system transporter subunit II D	0.007280641
CDR20291_0047	4.4	2.1	55.0	12.6	putative thymidylate synthase	0.002870937
CDR20291_1070	4.3	2.1	1379.6	319.0	cell surface protein	0.003437877

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_0960	4.3	2.1	87.1	20.2	hypothetical protein	6.14286E-13
CDR20291_2673	4.3	2.1	173.8	40.3	putative membrane protein	6.14338E-05
CDR20291_3298	4.3	2.1	945.2	220.5	CopG family transcriptional regulator	0.003909778
CDR20291_3079	4.3	2.1	161.3	37.9	methyltransferase	0.001827166
CDR20291_0662	4.3	2.1	9206.2	2165.7	hypothetical protein	0.000285918
CDR20291_1963	4.2	2.1	2117.7	499.0	hypothetical protein	0.002900441
mviN	4.2	2.1	247.5	58.3	putative transmembrane virulence factor MviN family protein	2.52716E-08
alr	4.2	2.1	7219.4	1701.4	alanine racemase	0.001239873
CDR20291_1812	4.2	2.1	150.9	35.7	hypothetical protein	0.005903069
CDR20291_1330	4.2	2.1	747.0	177.7	hypothetical protein	0.008810854
CDR20291_1814	4.2	2.1	163.5	39.0	plasmid-realted protein	3.80455E-05
CDR20291_0942	4.2	2.1	694.1	165.8	PTS system transporter subunit-IIB	5.10703E-14
CDR20291_3038	4.2	2.1	1284.2	306.8	D-aminoacylase	0.004652333
CDR20291_2843	4.2	2.1	226.5	54.2	hypothetical protein	1.43981E-05
CDR20291_2022	4.2	2.1	3418.4	821.8	copper-transporting P-type ATPase	1.96115E-07
CDR20291_0892	4.2	2.1	547.1	131.6	N-acetylmuramoyl-L-alanine amidase	2.96889E-10
CDR20291_0955	4.1	2.1	406.6	98.0	hypothetical protein	6.09218E-06
CDR20291_1878	4.1	2.1	228.1	55.0	ABC transporter ATP binding protein	2.42737E-10
CDR20291_2412	4.1	2.0	987.2	238.9	metal dependent phosphohydrolase	0.000870344
CDR20291_1805	4.1	2.0	101.6	24.7	ABC transporter permease	0.001684025
CDR20291_0121	4.1	2.0	26.0	6.3	hypothetical protein	0.0006543
CDR20291_1909	4.1	2.0	92.4	22.5	hypothetical protein	3.18745E-13
CDR20291_3109	4.1	2.0	189.2	46.2	hypothetical protein	0.002764003
spoIV	4.1	2.0	102.6	25.1	stage IV sporulation protein	9.98145E-06
CDR20291_0155	4.1	2.0	158.4	39.0	membrane associated CAAX amino terminal protease	1.50969E-05
CDR20291_1466	4.1	2.0	2186.7	539.1	hypothetical protein next to manganese containing catalase	2.44681E-05
CDR20291_2076	4.0	2.0	210.3	52.0	cNMP-binding regulatory protein	0.000173441

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2187	4.0	2.0	1384.8	343.4	two-component sensor histidine kinase	0.000286203
CDR20291_1393	4.0	2.0	353.4	87.7	hypothetical protein	0.006112635
CDR20291_1884	4.0	2.0	506.0	125.7	ABC transporter ATP binding protein	3.6941E-06
CDR20291_3494	4.0	2.0	42.8	10.7	hypothetical protein	1.93057E-06
CDR20291_3160	4.0	2.0	1276.1	320.2	probable transporter	2.71927E-06
CDR20291_3493	4.0	2.0	39.8	10.0	hypothetical protein	0.000960107
spoVS	4.0	2.0	435.8	110.1	stage V sporulation protein S	1.24548E-06
argG	4.0	2.0	165.2	41.8	argininosuccinate synthase	8.25638E-05
CDR20291_2563	3.9	2.0	262.2	66.4	LysR-family transcriptional regulator	4.5012E-10
CDR20291_1532	3.9	2.0	5051.6	1281.0	putative glycerophosphoryl diester phosphodiesterase	0.001577126
malX	3.9	2.0	140.6	35.7	PTS system, maltose and glucose-specific IIbc component	0.000710202
CDR20291_0475	3.9	2.0	53.8	13.7	putative membrane protein	0.004637793
asrC	3.9	2.0	22.7	5.8	anaerobic sulfite reductase subunit C	0.00341444
CDR20291_0227	3.9	2.0	80.9	20.6	putative transglycosylase	0.007313714
CDR20291_2274	3.9	2.0	1731.7	443.8	probable dehydrogenase	3.22506E-06
CDR20291_1983	3.9	2.0	5341.1	1373.0	conserved hypothetical protein	0.0009355
CDR20291_1904	3.9	2.0	142.8	36.7	conserved hypothetical protein	5.92503E-10
CDR20291_0539	3.9	2.0	100.3	25.8	MerR family transcriptional regultor	4.51528E-13
CDR20291_0633	3.9	2.0	6844.4	1769.1	putative signaling protein	0.000505606
CDR20291_1922	3.9	1.9	48.0	12.4	transcriptional regulator, AraC family	0.000168613
CDR20291_3529	3.9	1.9	2376.7	615.9	putative exported protein	0.009215404
CDR20291_0945	3.8	1.9	77.8	20.3	putative peptidase	1.05794E-08
tal1	3.8	1.9	2261.8	596.8	putative transaldolase	0.054543
CDR20291_1687	3.8	1.9	705.5	187.0	putative L-threonine dehydrogenase	0.0017865

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2853	3.8	1.9	19.7	5.2	putative sugar uptake protein	0.000375541
CDR20291_1202	3.7	1.9	123.5	33.0	putative phage repressor	2.74998E-05
CDR20291_2641	3.7	1.9	2538.3	684.6	conserved hypothetical protein	3.80494E-05
CDR20291_2932	3.7	1.9	32.5	8.8	putative glutamine amidotransferase	0.00000453
CDR20291_1591	3.7	1.9	124980.4	33770.3	putative dinitrogenase iron-molybdenum cofactor	0.000665025
CDR20291_2266	3.7	1.9	112349.3	30519.4	butyrate kinase	1.31213E-05
CDR20291_3362	3.7	1.9	38.5	10.5	putative iron ABC transporter, solute-binding protein	7.25107E-08
CDR20291_2499	3.7	1.9	825.5	225.7	hypothetical protein	7.21041E-06
CDR20291_1882	3.6	1.9	247.9	68.2	two-component system response regulator	0.001291137
ssuA	3.6	1.9	6503.7	1793.1	putative aliphatic sulfonates ABC transporter, substrate-binding lipoprotein	2.73052E-07
CDR20291_1144	3.6	1.9	180.3	49.8	putative NAD-dependent deacetylase (Sir2-family regulatory)	0.000957367
CDR20291_2492	3.6	1.8	2696.2	748.3	adp-ribosyltransferase binding component	1.77636E-15
CDR20291_1291	3.6	1.8	154.2	42.9	hypothetical protein	0.005106084
fur	3.6	1.8	5263.2	1480.7	ferric uptake regulation protein	2.19074E-06
CDR20291_1930	3.5	1.8	125.1	35.3	putative phage regulatory protein	3.38924E-05
ogt2	3.5	1.8	311.5	88.1	methylated-DNA--protein-cysteine methyltransferase 2	0.05
CDR20291_0827	3.5	1.8	49.6	14.0	conserved hypothetical protein	2.04241E-05
modB	3.5	1.8	27.3	7.7	ABC transporter, permease protein	0.005433038
CDR20291_3199	3.5	1.8	751.3	213.8	putative nitroreductase	2.93936E-06
CDR20291_1968	3.5	1.8	2148.3	611.6	conserved hypothetical protein	2.9976E-15
CDR20291_2782	3.5	1.8	49.1	14.0	putative drug/sodium antiporter	8.78554E-06

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_2135	3.5	1.8	146.3	41.8	putative glucokinase	9.2219E-05
CDR20291_0513	3.5	1.8	760.3	217.4	hypothetical protein	1.05341E-05
CDR20291_0375	3.5	1.8	98.6	28.4	conserved hypothetical protein	0.000100458
CDR20291_3071	3.5	1.8	358.5	103.3	hypothetical protein	0.000847521
nadD	3.5	1.8	607.3	175.8	nicotinate-nucleotide adenyltransferase	8.14488E-06
CDR20291_0182	3.5	1.8	230.3	66.7	putative membrane-associated metalloprotease	9.87E-08
CDR20291_2200	3.5	1.8	49.8	14.4	ABC transporter, substrate-binding protein0	0.002821301
CDR20291_0448	3.4	1.8	476.8	138.2	putative signaling protein	0.005681507
CDR20291_1323	3.4	1.8	113161.9	32843.8	putative ruberythrin	1.76224E-08
CDR20291_0714	3.4	1.8	467.1	136.2	putative stage IV sporulation protein	3.14534E-07
CDR20291_1347	3.4	1.8	325.0	94.8	hypothetical protein	0.001
CDR20291_0474	3.4	1.8	384.0	112.5	putative exported protein	0.006619176
CDR20291_0675	3.4	1.8	1669.4	491.8	putative lipoprotein	2.20951E-05
CDR20291_0436	3.4	1.8	70.6	20.8	PTS system, IIc component	0.000321887
CDR20291_1180	3.4	1.8	43433.7	12819.5	aspartate aminotransferase	0.000008965
CDR20291_2670	3.4	1.8	711.2	210.2	cell surface protein	4.97614E-06
CDR20291_1590	3.4	1.8	17522.9	5202.2	ArsR-family transcriptional regulator	0.05
CDR20291_1877	3.3	1.7	594.9	177.7	putative ABC transporter, permease protein	0.005759827
CDR20291_0017	3.3	1.7	3198.3	955.5	putative DNA-binding protein	2.54207E-06
CDR20291_2112	3.3	1.7	2021.7	605.6	putative membrane protein	4.44089E-16
CDR20291_0402	3.3	1.7	64.3	19.3	two-component sensor histidine kinase	0.008624649
CDR20291_0954	3.3	1.7	262.8	78.9	putative phosphoesterase	0.005575514
oppA	3.3	1.7	21233.9	6411.0	oligopeptide ABC transporter, substrate-binding lipoprotein	6.77459E-06
CDR20291_1921	3.3	1.7	45.0	13.7	putative membrane protein	0.006505287
CDR20291_0957	3.3	1.7	73296.3	22346.8	nitroreductase-family protein	0.006804837
CDR20291_3330	3.3	1.7	125.0	38.1	putative membrane protein	0.001849102

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_3297	3.3	1.7	188.4	57.8	putative regulator of cell growth	0.001476217
CDR20291_1908	3.3	1.7	56.0	17.2	NUDIX-family protein	0.00627968
ssuB	3.2	1.7	262.1	80.7	putative aliphatic sulfonates ABC transporter, ATP-binding protein	0.001181623
flgG	3.2	1.7	60.9	18.8	flagellar basal-body rod protein FlgG	6.75E-09
CDR20291_1521	3.2	1.7	182251.7	56506.9	putative nitric oxide reductase flavoprotein	3.63176E-06
CDR20291_1832	3.2	1.7	475.3	147.4	two-component sensor histidine kinase	2.45376E-05
CDR20291_2195	3.2	1.7	212.5	66.4	putative transcriptional regulator	0.009374281
aroD	3.2	1.7	150.7	47.1	3-dehydroquinate dehydratase	0.00982452
radA	3.2	1.7	5274.0	1664.0	DNA repair protein	2.03582E-12
CDR20291_0833	3.2	1.7	343.4	108.4	aldo/keto reductase	1.65E-12
rpiB1	3.2	1.7	523.9	165.8	ribose-5-phosphate isomerase 1	4.12521E-10
CDR20291_2897	3.1	1.7	61.3	19.5	putative exported protein	0.000734899
CDR20291_3088	3.1	1.7	4400.4	1398.4	putative membrane protein	6.01878E-07
hbd	3.1	1.6	9373.5	3006.4	3-hydroxybutyryl-CoA dehydrogenase	0.000193473
thlA1	3.1	1.6	51153.9	16445.3	acetyl-CoA acetyltransferase	8.87601E-06
CDR20291_1717	3.1	1.6	669.3	215.9	conserved hypothetical protein	0.005408216
CDR20291_2206	3.1	1.6	60.4	19.5	two-component sensor histidine kinase	0.007280575
CDR20291_3329	3.1	1.6	113.9	37.0	putative septum formation protein	0.000170466
CDR20291_3331	3.1	1.6	50.9	16.5	conserved hypothetical protein	6.754E-09
CDR20291_1794	3.1	1.6	261.2	85.2	putative uncharacterized protein	1.0433E-07
oppF	3.1	1.6	14650.4	4785.5	oligopeptide ABC transporter, ATP-binding protein	0.0012

Gene Name	Real Fold Change	Log2 fold change	Parent-Mean expression	R20291 *sinRR'* Mean expression	Function (known or predicted)	adj. *p*-Value
CDR20291_3106	3.0	1.6	412.6	135.9	probable polysaccharide deacetylase	0.000675
CDR20291_0496	3.0	1.6	269.6	88.8	putative membrane protein	0.001215274
CDR20291_2689	3.0	1.6	150.2	49.5	putative membrane protein	0.002642629
CDR20291_2671	3.0	1.6	5700.9	1885.2	putative glycosyltransferase	0.000123997
cbiN	3.0	1.6	2533.2	838.4	cobalt transport protein	0.00100167
CDR20291_3211	3.0	1.6	34.9	11.5	hypothetical protein	3.54277E-05
pheA	3.0	1.6	40.5	13.4	P-protein	8.87959E-05
CDR20291_2818	3.0	1.6	76.3	25.4	conserved hypothetical protein	0.000132095
CDR20291_2938	3.0	1.6	40.4	13.5	putative C4-dicarboxylate anaerobic carrier	4.40862E-06
CDR20291_0677	3.0	1.6	505.4	168.6	putative ATP-dependent DNA helicase	2.29794E-09
CDR20291_2217	3.0	1.6	362.6	122.4	sigma 54 interacting transcription antiterminator	6.21383E-06
CDR20291_1880	2.9	1.6	173.3	58.8	two-component response regulator	0.002893449
CDR20291_0956	2.9	1.6	354.2	120.2	hypothetical protein	0.000788674
CDR20291_3424	2.9	1.6	97.5	33.2	two-component sensor histidine kinase	0.002340679
dnaK	2.9	1.6	108445.6	37020.6	DnaK	4.5993E-05
mleN	2.9	1.5	48.7	16.6	putative malate-2H(+)/Na(+)-lactate antiporter	0.05
CDR20291_2496	2.9	1.5	1086.9	372.4	probable peptidase	9.38346E-07
CDR20291_2967	2.9	1.5	73.0	25.1	conserved hypothetical protein	0.008062885
CDR20291_2649	2.9	1.5	494.8	170.5	putative N-acetylmuramoyl-L-alanine amidase	1.35226E-06
CDR20291_1359	2.9	1.5	1281.5	441.6	conserved hypothetical protein	3.27758E-08
glnA	2.9	1.5	1166.0	403.3	glutamine synthetase	9.00368E-08
CDR20291_1674	2.9	1.5	614.1	212.6	conserved hypothetical protein	4.93012E-13
CDR20291_0404	2.9	1.5	33.3	11.5	TetR-family transcriptional regulator	3.34137E-06
spo0A	2.9	1.5	5171.9	1790.2	stage 0 sporulation protein A	2.99663E-06

Gene Name	**Real Fold Change**	**Log2 fold change**	**Parent-Mean expression**	**R20291 *sinRR'* Mean expression**	**Function (known or predicted)**	**adj. *p*-Value**
CDR20291_0540	2.9	1.5	15.6	5.4	putative membrane protein	9.64198E-06
CDR20291_0660	2.9	1.5	310.8	108.3	putative biotin/lipoate-protein ligase	1.14579E-05
CDR20291_1902	2.9	1.5	664.8	232.5	ABC transporter, substrate-binding protein	1.27639E-05
CDR20291_0190	2.9	1.5	103.6	36.2	LysR-family transcriptional regulator	3.86235E-09
CDR20291_0715	2.9	1.5	1015.6	355.4	putative N-acetylmuramoyl-L-alanine amidase	1.64701E-06
CDR20291_1883	2.9	1.5	263.5	92.3	two-component sensor histidine kinase	1.00424E-11
hisC	2.9	1.5	29.8	10.4	putative histidinol-phosphate aminotransferase	5.71554E-15
CDR20291_1187	2.9	1.5	63.4	22.2	padr-family transcriptional regulator	5.26672E-06
CDR20291_2418	2.9	1.5	250.5	87.9	putative membrane protein	1.91823E-07
CDR20291_3202	2.8	1.5	143.0	50.3	AsnC-family transcriptional regulator fragment	1.11432E-05
rpe	2.8	1.5	161.0	56.9	putative ribulose-phosphate 3-epimerase	1.34427E-11

Appendix D - Up regulated genes in R∆*sinRR'* relative to R20291

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
int-Tn	0.0	-7.5	1.6	304.9	integrase	6.65319E-08
CDR20291_0707	0.0	-5.7	15.9	847.0	putative membrane protein	2.68E-03
atpC	0.0	-5.7	72.3	3773.2	ATP synthase epsilon chain	1.40784E-07
atpG	0.0	-5.4	245.6	10501.7	ATP synthase subunit gamma	1.1118E-10
act	0.0	-5.4	51.3	2094.8	putative beta-alanine CoA-transferase	1.39E-06
CDR20291_0708	0.0	-5.3	117.2	4732.4	putative aminohydrolase	2.0271E-12
atpH	0.0	-5.2	25.5	970.1	ATP synthase gamma chain	6.79403E-13
eutA	0.0	-5.2	3.7	140.8	putative ethanolamine/propanediol utilisation protein	6.13264E-11
CDR20291_0706	0.0	-5.2	59.8	2238.4	putative membrane protein	5.87E-03
rbsC	0.0	-5.2	27.9	1039.4	ribose ABC transporter, permease protein	1.96011E-10
rbsA	0.0	-5.2	75.3	2750.3	ribose ABC transporter, ATP-binding protein	6.623E-08
CDR20291_0709	0.0	-5.0	100.3	3225.3	conserved hypothetical protein	8.83E-03
atpA	0.0	-4.9	376.8	11578.6	ATP synthase alpha chain	3.6844E-05
argE	0.0	-4.9	39.6	1199.3	putative acetylornithine deacetylase	7.06331E-13
eutC	0.0	-4.9	4.1	124.6	putative ethanolamine/propanediol ammonia-lyase light chain	5.08836E-11
eutL	0.0	-4.8	1.2	32.8	putative ethanolamine/propanediol utilisation protein	1.46682E-10
CDR20291_0729	0.0	-4.8	15.6	421.2	putative membrane protein	2.26503E-12
atpF	0.0	-4.7	197.0	5166.9	ATP synthase B chain	1.84082E-07
rbsB	0.0	-4.7	80.5	2087.5	D-ribose ABC transporter, substrate-binding protein	8.55805E-09
eutB	0.0	-4.6	11.1	275.4	putative ethanolamine/propanediol ammonia-lyase heavy chain	5.04107E-12
CDR20291_0307	0.0	-4.6	28.6	704.2	conserved hypothetical protein	4.36449E-07
gatY	0.0	-4.6	5.8	140.3	tagatose bisphosphate aldolase	3.37E-04

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
valS	0.0	-4.6	55.9	1317.7	valyl tRNA synthase	9.75255E-11
CDR20291_1838	0.0	-4.6	7.7	179.9	putative ethanolamine/propanediol utilisation aldehyde alcohol dehydrogenase	5.92E-10
CDR20291_2686	0.0	-4.5	580.5	13194.2	putative exported protein	2.50969E-08
CDR20291_2586	0.0	-4.5	7.2	160.7	putative permease	5.98E-02
CDR20291_0191	0.0	-4.5	45.5	1011.9	putative membrane protein	1.15E-17
atpD	0.0	-4.5	2208.4	48799.4	ATP synthase beta chain	2.66394E-14
CDR20291_2585	0.0	-4.5	45.7	1006.9	putative aminohydrolase	8.53034E-08
CDR20291_2685	0.0	-4.5	156.1	3412.9	cell surface protein	9.42663E-11
eutM	0.0	-4.4	4.0	88.3	putative ethanolamine/propanediol utilisation protein	2.91556E-23
CDR20291_2454	0.0	-4.3	8.4	169.0	PTS system, IIC component	3.48678E-06
CDR20291_2587	0.1	-4.2	14.8	271.2	putative membrane protein	1.41188E-09
CDR20291_0529	0.1	-4.1	2041.3	36210.3	chloride ion channel protein (pseudogene)	8.90659E-08
CDR20291_1697	0.1	-4.1	40.4	692.6	putative hydrolase	7.50058E-38
CDR20291_1837	0.1	-4.1	2.4	41.5	putative ethanolamine/propanediol utilisation protein	2.75677E-14
hymC	0.1	-4.1	952.2	15851.2	putative iron-only hydrogenase, catalytic subunit	1.10551E-09
mtlD	0.1	-4.0	887.1	14465.9	mannitol-1-phosphate 5-dehydrogenase	6.98098E-08
CDR20291_1702	0.1	-4.0	20.1	324.9	putative NADPH-dependent FMN reductase	4.07392E-15
rbsR	0.1	-4.0	80.7	1296.1	putative ribose operon repressor	1.57813E-21
pduQ	0.1	-4.0	2.8	44.2	putative ethanolamine/propanediol utilization propanol dehydrogenase	7.89616E-15
atpB	0.1	-4.0	114.6	1792.8	ATP synthase A chain	7.46636E-17
CDR20291_2251	0.1	-4.0	1.5	22.9	putative membrane protein	2.4679E-16
CDR20291_0748	0.1	-3.9	58.6	879.8	putative 6-phospho-beta-glucosidase	2.06035E-20

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
abgB1	0.1	-3.9	8.3	123.7	putative aminobenzoyl-glutamate utilization protein	5.46963E-40
hom2	0.1	-3.9	5.0	74.0	homoserine dehydrogenase	1.41779E-12
mtlF	0.1	-3.9	256.3	3813.2	PTS system, mannitol-specific IIa component	1.31013E-11
eutN	0.1	-3.8	1.7	24.7	putative ethanolamine/propanediol utilisation protein	5.03148E-14
mtlR	0.1	-3.8	1284.6	17812.0	putative transcription antiterminator	1.21117E-39
atpZ	0.1	-3.8	9.7	133.2	putative ATP synthase protein	5.15997E-30
agaS	0.1	-3.8	3.2	44.4	putative tagatose-6-phosphate ketose/aldose isomerase	1.32018E-24
CDR20291_2253	0.1	-3.7	109.3	1433.6	putative exported protein	8.15757E-28
CDR20291_2248	0.1	-3.7	227.5	2919.3	putative aliphatic sulfonate ABC transporter, ATP-binding protein	8.02139E-12
CDR20291_0059	0.1	-3.7	334.0	4229.6	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase	3.86E-02
CDR20291_1698	0.1	-3.7	105.2	1320.8	cell surface protein	1.73151E-10
CDR20291_2247	0.1	-3.6	67.1	837.9	hypothetical protein	3.50086E-08
CDR20291_2487	0.1	-3.6	4046.2	50207.0	putative carbon starvation	4.51355E-34
eutT	0.1	-3.6	2.0	24.8	putative ethanolamine/propanediol utilisation cobalmine adenosyl transferase	1.54E-03
atpE	0.1	-3.6	183.1	2231.6	ATP synthase C chain	2.50844E-42
CDR20291_0564	0.1	-3.6	1.5	17.8	putative transcriptional regulator	6.20312E-43
atpI	0.1	-3.5	334.0	3892.0	ATP synthase protein I	1.16849E-10
CDR20291_2349	0.1	-3.5	11.2	127.8	ABC transporter, ATP-binding protein	3.76933E-08
CDR20291_3313	0.1	-3.5	91.2	1031.2	putative cytidine and deoxycytidylate deaminase	1.75329E-27
bioY	0.1	-3.5	2.9	32.1	putative biotin synthase	3.54291E-28
CDR20291_2272	0.1	-3.5	39.0	429.0	putative signaling protein	1.73588E-14
CDR20291_2249	0.1	-3.4	144.7	1577.6	abc transporter permease protein (pseudogene)	3.43E-04
serS1	0.1	-3.4	885.5	9597.5	seryl-tRNA synthetase	2.00212E-27
asnA	0.1	-3.4	26.8	284.3	aspartate-ammonia ligase	1.30E-02

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_0747	0.1	-3.4	33.1	351.6	PTS system, IIabc component	3.34E-05
CDR20291_1841	0.1	-3.4	2.2	23.1	putative ethanolamine/propanediol utilisation protein	9.78069E-26
CDR20291_2373	0.1	-3.4	299.3	3109.1	putative exported protein	3.31E-05
rpmG	0.1	-3.4	5.8	60.0	50S ribosomal protein L33	3.40497E-26
rbsK	0.1	-3.3	208.3	2072.9	putative ribokinase	2.41E-04
dccA	0.1	-3.3	32.1	318.4	putative signaling protein	2.05E-24
CDR20291_2849	0.1	-3.3	16.1	158.4	PTS system, IIc component	1.06302E-07
CDR20291_3190	0.1	-3.3	1.8	17.7	conserved hypothetical protein	1.96E-05
CDR20291_2557	0.1	-3.3	4.3	41.9	putative Na(+)/H(+) antiporter	3.01119E-08
mtlA	0.1	-3.3	791.1	7649.6	PTS system, mannitol-specific IIbc component	8.49426E-08
rnpA	0.1	-3.3	28.6	274.8	ribonuclease P protein component	4.29269E-07
CDR20291_0750	0.1	-3.2	42.2	392.5	two component response regulator	3.31E-02
CDR20291_1581	0.1	-3.2	33.3	306.2	putative membrane protein	2.06583E-07
CDR20291_2688	0.1	-3.2	11.4	103.2	cell surface protein	3.08025E-05
ruvB	0.1	-3.2	60.8	547.7	holliday junction DNA helicase	1.19187E-07
hymB	0.1	-3.2	603.4	5421.8	putative iron-only hydrogenase, electron-transferring subunit	6.35128E-08
CDR20291_3481	0.1	-3.2	12.2	109.7	putative acetyltransferase	2.0667E-06
CDR20291_1615	0.1	-3.2	0.7	6.7	probable permease	6.06941E-07
CDR20291_2699	0.1	-3.2	9.2	83.0	putative membrane protein	5.49E-02
CDR20291_1233	0.1	-3.2	352.9	3144.6	putative allophanate hydrolase subunit 2	9.77574E-08
uxaA	0.1	-3.1	47.6	419.6	putative altronate hydrolase	1.33872E-12
CDR20291_3539	0.1	-3.1	2.9	25.4	conserved hypothetical protein	3.15791E-07
CDR20291_2256	0.1	-3.1	191.3	1654.5	conserved hypothetical protein	1.82394E-13
CDR20291_0164	0.1	-3.1	2.8	24.4	putative membrane protein	7.00652E-09
CDR20291_2692	0.1	-3.1	39.9	344.0	putative membrane protein	1.00571E-15
gidB	0.1	-3.1	91.1	782.9	methyltransferase (putative glucose inhibited division protein B)	1.50925E-07
CDR20291_1232	0.1	-3.1	51.1	437.3	putative allophanate hydrolase subunit 1	1.01E-02

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
acoC	0.1	-3.1	33.8	289.1	E2 component of acetoin dehydrogenase enzyme system	1.14314E-07
ulaA	0.1	-3.1	7.8	66.3	PTS system abscorbate-specific transporter subunit IIC	7.06078E-08
CDR20291_2245	0.1	-3.1	50.0	424.7	putative oligoendopeptidase	3.18E-03
CDR20291_0358	0.1	-3.1	12.4	105.0	putative membrane protein	1.00943E-07
acoB	0.1	-3.1	44.9	376.8	acetoin:2,6-dichlorophenolindophenol oxidoreductase beta subunit	1.13E-03
acoA	0.1	-3.1	21.9	182.3	acetoin:2,6-dichlorophenolindophenol oxidoreductase alpha subunit	1.14935E-06
eutH	0.1	-3.1	6.0	50.3	putative ethanolamine/propanediol transporter	7.45996E-08
CDR20291_0296	0.1	-3.0	75.0	609.4	ABC transporter, ATP-binding protein	1.29435E-16
gidA	0.1	-3.0	334.3	2707.1	glucose inhibited division protein A	1.89595E-08
CDR20291_2255	0.1	-3.0	46.9	378.0	hypothetical protein	2.58275E-14
CDR20291_1406	0.1	-3.0	69.4	554.1	putative peptidyl-prolyl isomerase	3.37307E-16
CDR20291_2842	0.1	-3.0	18.9	150.6	probable alcohol dehydrogenase	5.95702E-07
hymA	0.1	-3.0	168.6	1342.3	putative iron-only hydrogenase, electron-transferring subunit	1.85E-04
CDR20291_1578	0.1	-3.0	2.2	17.4	putative FMN-binding exported protein	5.52745E-08
CDR20291_0039	0.1	-3.0	2275.1	17906.1	putative dual-specificity prolyl/cysteinyl-tRNA synthetase	3.14E-06
trmE	0.1	-3.0	177.2	1379.9	putative tRNA modification GTPase	4.60541E-08
CDR20291_1289	0.1	-3.0	15.0	116.8	putative membrane protein	1.66638E-09
CDR20291_2871	0.1	-3.0	33.5	260.2	proton-dependent oligopeptide transporter	8.63048E-14
CDR20291_3192	0.1	-2.9	75.9	581.5	conserved hypothetical protein	6.98139E-09
queA	0.1	-2.9	148.7	1136.9	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	2.75915E-08
CDR20291_1268	0.1	-2.9	97.8	735.8	putative signaling protein	4.87356E-07
CDR20291_3223	0.1	-2.9	55.3	411.5	putative phosphoesterase	1.08E-03
CDR20291_0295	0.1	-2.9	65.8	488.2	DNA binding protein	1.36086E-08

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_2273	0.1	-2.9	4.3	31.7	conserved hypothetical protein	5.36754E-09
ruvA	0.1	-2.9	31.7	232.4	holliday junction DNA helicase	2.59739E-07
ruvC	0.1	-2.8	36.7	263.1	crossover junction endodeoxyribonuclease RuvC	6.22E-04
CDR20291_2254	0.1	-2.8	220.7	1569.6	putative permease	6.73968E-10
CDR20291_1230	0.1	-2.8	48.8	345.8	conserved hypothetical protein	6.06465E-09
cls	0.1	-2.8	863.4	6033.7	putative cardiolipin synthetase	4.12482E-10
iunH	0.1	-2.8	229.4	1588.8	Inosine uridine preferring nucleoside hydrolase	2.8636E-24
adhE	0.1	-2.8	29.1	201.1	bifunctional acetaldehyde-CoA/alcohol dehydrogenase	4.5333E-11
CDR20291_0728	0.1	-2.8	14.5	99.7	hydroxymethylglutaryl-coA lyase	3.12374E-19
CDR20291_2390	0.1	-2.8	91.6	624.1	D-alanyl-D-alanine carboxypeptidase	2.53796E-06
CDR20291_2402	0.1	-2.7	6.7	44.9	PTS system transporter subunit IIBC	1.39924E-11
CDR20291_2701	0.1	-2.7	4.7	31.2	putative gluconate permease	3.83672E-09
CDR20291_2287	0.2	-2.7	31.2	203.2	exonuclease	5.55E-04
CDR20291_2826	0.2	-2.7	13.3	86.7	ABC transporter permease	1.62119E-07
CDR20291_2721	0.2	-2.7	1184.2	7693.1	hypothetical protein	3.31623E-08
malX	0.2	-2.7	8.6	55.5	PTS system, maltose and glucose-specific IIbc component	2.86537E-07
rplK	0.2	-2.7	28.3	180.1	50S ribosomal protein L11	1.32E-05
acoL	0.2	-2.7	64.2	408.9	E3 component of acetoin dehydrogenase enzyme system	7.21418E-08
CDR20291_1486	0.2	-2.7	30.5	194.2	ribose ABC transporter, permease protein	1.13E-05
CDR20291_1579	0.2	-2.7	5.1	32.3	TetR family transcirptional regualtor	2.03987E-14
CDR20291_0563	0.2	-2.7	11.5	72.4	hypothetical protein	1.24E-04
CDR20291_0735	0.2	-2.7	9.9	62.3	electron transfer flavoprotein subunit beta	9.99E-02
CDR20291_0287	0.2	-2.7	7.2	45.4	PTS transporter subunit IIA	1.50782E-07
tgt	0.2	-2.7	221.9	1394.4	queuine tRNA riobsyltransferase	1.44826E-09
CDR20291_2378	0.2	-2.6	8.2	51.0	Xaa-Pro dipeptidase	4.40791E-17
CDR20291_2850	0.2	-2.6	5.6	34.4	PTS system, IIb component	8.39488E-08
CDR20291_2582	0.2	-2.6	4.2	25.8	putative membrane protein	8.61203E-15

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_2687	0.2	-2.6	158.1	970.0	cell surface protein	3.53225E-07
CDR20291_0749	0.2	-2.6	99.3	607.3	putative transcription antiterminator	9.5409E-08
srlE	0.2	-2.6	16.6	100.7	PTS system, glucitol/sorbitol-specific IIbc component	1.45069E-07
CDR20291_0635	0.2	-2.6	227.8	1375.3	DNA mismatch repair protein	4.79749E-14
CDR20291_3221	0.2	-2.6	210.2	1258.1	putative exported protein	1.95773E-07
CDR20291_2288	0.2	-2.6	23.9	142.8	hypothetical protein	1.91069E-13
fbpA	0.2	-2.6	41.1	244.7	fbaA fibronectin binding protein	3.32374E-12
lplA	0.2	-2.6	91.5	541.2	lipoate protein ligase	5.22255E-35
adhE	0.2	-2.6	10674.2	63100.2	aldehyde-alcohol dehydrogenase	4.68467E-11
CDR20291_0596	0.2	-2.6	13.1	77.6	hypothetical protein	3.0142E-09
CDR20291_2626	0.2	-2.6	435.0	2567.4	carbon-nitrogen hydrolase	2.48324E-09
CDR20291_1405	0.2	-2.6	33.7	198.9	polysaccharide deacetlyase	1.57744E-08
CDR20291_1100	0.2	-2.5	2579.9	15068.8	branched chain amnioacid transport system carrier protein	1.43403E-08
CDR20291_1582	0.2	-2.5	105.8	617.4	radical SAM protien	1.95778E-07
CDR20291_3322	0.2	-2.5	50.6	294.5	radical SAM-superfamily protein	2.50767E-14
CDR20291_2690	0.2	-2.5	429.7	2494.5	protein translocase subunit	2.77458E-08
rplL	0.2	-2.5	79.2	459.2	50S ribosomal protein L7/L12	6.47E-02
CDR20291_1378	0.2	-2.5	223.7	1281.1	ABC transporter permease	1.24281E-07
CDR20291_0170	0.2	-2.5	120.1	685.8	ABC transporter ATP-binding protein	2.61468E-08
pheT	0.2	-2.5	144.2	820.9	phenylalanine-tRNA-synthetase subunit beta	3.72958E-11
murB	0.2	-2.5	145.1	822.9	UDP-N-acetylenolpyruvoylglucosamine reductase	1.78703E-09
nusG	0.2	-2.5	82.5	461.4	transcription antitermination protein	3.50E-03
fusA	0.2	-2.5	850.9	4722.8	translation elongation factor G	1.41754E-14
upp	0.2	-2.5	122.8	680.2	uracil phosphoribosyltransferase	1.64289E-09
rho	0.2	-2.5	2074.8	11433.5	transcription termination factor Rho	7.68258E-15
CDR20291_2807	0.2	-2.5	12.6	69.2	hypothetical protein	5.99381E-08
CDR20291_1518	0.2	-2.4	24.2	132.3	putative permease	2.53043E-11

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
srlB	0.2	-2.4	5.1	27.7	PTS system, glucitol/sorbitol-specific IIa component	2.24E-06
CDR20291_0357	0.2	-2.4	8.3	45.5	conserved hypothetical protein	4.1606E-14
CDR20291_3222	0.2	-2.4	137.8	747.6	putative ATP-binding protein	1.29819E-13
CDR20291_1376	0.2	-2.4	293.1	1576.1	ABC transporter permease	7.66568E-14
CDR20291_3128	0.2	-2.4	7.9	42.6	two component response regulator	3.38569E-10
CDR20291_0636	0.2	-2.4	45.0	241.2	hypothetical protein	4.48428E-11
CDR20291_1408	0.2	-2.4	49.7	265.9	hypothetical protein	3.38149E-35
CDR20291_1710	0.2	-2.4	42.5	226.6	conserved hypothetical protein	7.49948E-10
CDR20291_1487	0.2	-2.4	106.0	564.1	ribose ABC transporter substrate-binding protein	4.00481E-09
feoB3	0.2	-2.4	416.0	2181.9	putative ferrous iron transport protein B	5.45054E-08
CDR20291_1099	0.2	-2.4	16.6	87.0	branched chain amnioacid transport system carrier protein	3.47E-03
CDR20291_1411	0.2	-2.4	53.1	277.4	hypothetical protein	2.17333E-07
CDR20291_2408	0.2	-2.4	104.9	546.8	putative L-asparaginase	1.56E-04
CDR20291_2512	0.2	-2.4	21.0	108.1	putative exported protein	1.34191E-07
CDR20291_1377	0.2	-2.4	151.4	776.8	ABC transporter ATP-binding protein	1.14E-02
CDR20291_2438	0.2	-2.3	10.3	52.3	hypothetical protein	5.20641E-43
CDR20291_2804	0.2	-2.3	29.3	148.7	putative AMP binding protein	8.3096E-10
CDR20291_2271	0.2	-2.3	120.6	609.7	putative signaling protein	8.56653E-08
uraA	0.2	-2.3	14.5	73.0	uracil permease	5.14207E-11
scrK	0.2	-2.3	20.4	101.7	putative fructokinase	7.52915E-12
CDR20291_3525	0.2	-2.3	224.9	1117.4	conserved hypothetical protein adjacent to soj operon	6.05932E-10
CDR20291_2741	0.2	-2.3	10.8	53.3	DeoR-family transcriptional regulator	1.34669E-07
sigV	0.2	-2.3	23.9	117.7	RNA polymerase sigma factor	3.28829E-08
CDR20291_1571	0.2	-2.3	15.0	73.7	putative NADPH-dependent FMN reductase	2.18E-05
oxaA1	0.2	-2.3	191.4	940.6	putative sporulation membrane protein	1.91752E-09
CDR20291_2252	0.2	-2.3	2070.2	10168.1	putative sulfonate ABC transporter, solute-binding lipoprotein	8.20E-02

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
rplA	0.2	-2.3	121.6	590.8	50S ribosomal protein L1	8.18942E-08
grdD	0.2	-2.3	13698.2	66498.4	glycine/sarcosine/betaine reductase complex component C alpha subunit	2.74747E-07
CDR20291_2040	0.2	-2.3	86.6	418.8	putative signaling protein	6.26411E-06
CDR20291_2656	0.2	-2.3	288.0	1391.7	putative cell-wall hydrolase	1.33427E-13
CDR20291_1409	0.2	-2.3	51.7	249.5	hypothetical protein	2.54573E-37
plfB	0.2	-2.3	2223.2	10614.2	formate acetylate transferase	1.56731E-08
rpiB2	0.2	-2.3	125.8	599.9	ribose 5 phosphate isomerase 2	6.1132E-07
jag	0.2	-2.2	209.4	993.4	SpoIIIJ-associated protein	3.52E-03
pykF	0.2	-2.2	764.2	3624.5	pyruvate kinase - gluconeogenesis II	1.81592E-07
CDR20291_2191	0.2	-2.2	36.3	171.5	pilin protein	1.42598E-08
CDR20291_2825	0.2	-2.2	14.3	67.2	ABC transporter ATP-binding protein	8.06E-23
rplU	0.2	-2.2	450.1	2113.8	50S ribosomal protein L21	2.14164E-09
CDR20291_1288	0.2	-2.2	18.7	88.0	putative membrane protein	1.9922E-07
proC1	0.2	-2.2	1434.8	6728.2	pyroline 5-carboxylate reductase	6.13764E-08
CDR20291_1116	0.2	-2.2	7.3	34.3	recombinase factor protein RarA	9.43E-05
hemK	0.2	-2.2	59.9	279.6	protein methyltransferase	1.55423E-08
CDR20291_3113	0.2	-2.2	90.4	421.1	two component response regulator	4.60E-02
CDR20291_2096	0.2	-2.2	38.6	179.6	cyclomaltodextrinase	9.79242E-15
CDR20291_0179	0.2	-2.2	26.1	121.3	ATP/GTP binding proteins	1.6616E-12
CDR20291_1657	0.2	-2.2	90.3	419.2	hypothetical protein	4.47E-02
secE	0.2	-2.2	15.5	71.8	preprotein translocase SecE subunit	6.01227E-08
pheS	0.2	-2.2	28.1	129.4	phenylalanine-tRNA-synthetase subunit alpha	2.45E-04
CDR20291_0009	0.2	-2.2	2.8	12.9	hypothetical protein	2.47968E-11
CDR20291_1273	0.2	-2.2	186.6	856.8	hypothetical protein	3.67E-03
CDR20291_2294	0.2	-2.2	64.6	294.8	DNA topoisomerase	3.73045E-11
cysK	0.2	-2.2	170.0	771.9	cysteine synthase A	2.65864E-08
ptsG	0.2	-2.2	116.4	528.0	PTS system, glucose-specific IIbc component (fragment)	1.63E-03
argS	0.2	-2.2	774.1	3504.7	arginyl tRNA synthetase	6.93635E-08

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_1239	0.2	-2.2	27.8	125.3	hypothetical protein	8.06E-02
CDR20291_2372	0.2	-2.2	5.6	25.4	conserved hypothetical protein	4.95E-28
rpmH	0.2	-2.2	101.2	454.9	50S ribosomal protein L34	2.79E-31
CDR20291_1010	0.2	-2.2	3819.7	17158.0	FAD/FMN containing dehydrogenase	1.13E-16
rplJ	0.2	-2.2	227.8	1020.1	50S ribosomal protein L10	1.00153E-07
CDR20291_1284	0.2	-2.2	182.9	815.7	hypothetical protein	3.43464E-08
feoA3	0.2	-2.2	67.7	301.5	ferrous iron transport protein A	7.18105E-08
CDR20291_2584	0.2	-2.2	66.1	294.3	putative aminotransferase	6.15858E-06
infC	0.2	-2.2	591.3	2626.7	translation initiation factor IF-3	7.77584E-07
CDR20291_1410	0.2	-2.1	151.8	666.2	dehydrogenase accessory protein	3.23453E-08
CDR20291_1555	0.2	-2.1	1164.0	5023.2	bifunctional glycine dehydrogenase/aminomethyl transferase protein	1.50E-02
CDR20291_2270	0.2	-2.1	233.9	1002.2	putative sigma-54 interacting regulatory protein	9.34752E-08
CDR20291_1231	0.2	-2.1	87.4	373.1	probable transporter	1.14731E-07
CDR20291_0051	0.2	-2.1	4042.9	17221.0	elongation factor TU	1.4959E-09
CDR20291_0290	0.2	-2.1	8.0	34.3	PTS system transporter subunit IIB	4.64E-04
CDR20291_0764	0.2	-2.1	113.2	481.6	isocitrate/3-isopylmalate dehydrogenase	6.52397E-08
CDR20291_2246	0.2	-2.1	41.4	175.7	putative hydrolase	1.44261E-06
CDR20291_0211	0.2	-2.1	30.6	129.3	hydrolase	1.38E-11
CDR20291_3145	0.2	-2.1	179.7	760.2	protease	1.86923E-08
CDR20291_2453	0.2	-2.1	11.2	47.0	PTS system, IIb component	1.18982E-07
CDR20291_1519	0.2	-2.1	70.8	296.1	mechanosensitive iron channel protein	2.04E-23
addA	0.2	-2.1	483.8	2019.7	ATP dependent nuclease subunit A	6.32449E-08
CDR20291_2703	0.2	-2.1	23.0	95.7	GntR family transcriptional regulator	3.00E-19
CDR20291_1379	0.2	-2.1	27.9	115.9	two component sensor histidine kinase	1.3913E-07
gloA	0.2	-2.1	18.4	76.1	lactoylglutathione lyase	6.52E-03
CDR20291_0395	0.2	-2.0	9.2	38.3	AraC family transcriptional regulator	1.60E-06

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_1485	0.2	-2.0	80.0	330.2	ribose ABC transporter ATP binding protein	1.30423E-08
CDR20291_1028	0.2	-2.0	18.6	76.7	acetyltransferase	4.52724E-10
CDR20291_3220	0.2	-2.0	92.6	380.8	putative DNA repair protein (nucleotide pyrophosphatase)	3.94E-06
rpsG	0.2	-2.0	372.7	1528.5	30S ribosomal protein S7	6.15891E-08
CDR20291_3317	0.2	-2.0	42.8	175.3	low molecular weight protein tyrosine phosphatase	1.62026E-07
cafA	0.2	-2.0	354.1	1442.5	ribonuclease G	6.59502E-11
CDR20291_3181	0.2	-2.0	14.7	59.8	two component response regulator	4.45912E-08
CDR20291_0285	0.2	-2.0	176.9	717.1	hydrolase	4.68625E-10
CDR20291_2379	0.2	-2.0	68.0	275.0	PTS system transporter subunit IIC	3.25075E-09
CDR20291_1510	0.2	-2.0	58.3	235.3	amidohydrolase	1.53876E-08
abgB2	0.2	-2.0	44.3	178.8	aminobenzoyl-glutamate utilization protein	1.86E-02
CDR20291_1583	0.2	-2.0	38.4	154.3	conserved hypothetical protein	3.23147E-10
CDR20291_2490	0.2	-2.0	42.6	170.4	putative response regulator	3.07015E-07
rpmE	0.3	-2.0	765.3	3060.5	50S ribosomal protein L31	1.27745E-08
CDR20291_0443	0.3	-2.0	38.0	150.7	hypothetical protein	8.62E-03
srlE	0.3	-2.0	12.5	49.3	PTS system, glucitol/sorbitol-specific IIbc component x3	2.12065E-07
ribC	0.3	-2.0	3.8	14.8	riboflavin biosynthesis protein	6.23E-03
CDR20291_1476	0.3	-2.0	252.9	990.4	two component sensor histidine kinase (CD1579 homologue in JIR)- sporulation specific kinase	7.41723E-08
CDR20291_2827	0.3	-2.0	23.2	90.9	hypothetical protein	1.93468E-09
CDR20291_2627	0.3	-2.0	286.0	1118.8	cytosine permease	2.61848E-09
CDR20291_0291	0.3	-2.0	29.8	116.1	PTS system transporter subunit IIC	1.73E-02
CDR20291_0210	0.3	-2.0	357.6	1394.2	sugar phosphate kinase	5.47942E-09
CDR20291_2735	0.3	-2.0	2.5	9.7	MarR family transcriptional regulator	4.21529E-13
CDR20291_1692	0.3	-2.0	36476.9	142054.4	pyridine nucleotide-disulfide oxidoreductase	9.32039E-10
CDR20291_0321	0.3	-2.0	46.5	180.6	ABC transporter permease	9.54E-03

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
crr	0.3	-2.0	34.3	133.1	PTS system, glucose-specific IIa component	6.72457E-13
CDR20291_3219	0.3	-2.0	339.8	1315.0	conserved hypothetical protein	2.31E-03
rplC	0.3	-2.0	108.7	420.1	50S ribosomal protein L3	1.57146E-10
gcvPB	0.3	-1.9	2141.4	8200.3	glycine dehydrogenase subunit 2	1.41E-02
CDR20291_1287	0.3	-1.9	747.5	2856.2	putative multiprotein complex assembly protein	7.50E-03
clpX	0.3	-1.9	3877.6	14809.8	ATP dependent protease - ATP binding subunt ClpX	1.22781E-09
CDR20291_2848	0.3	-1.9	61.6	235.1	putative glycosyl hydrolase	6.32E-06
srlE	0.3	-1.9	18.6	70.4	PTS system, glucitol/sorbitol-specific IIbc component x3	1.51485E-08
etfA3	0.3	-1.9	2716.7	10291.4	electron transfer flavoprotein subunit alpha	2.06E-21
CDR20291_2700	0.3	-1.9	15.2	57.1	conserved hypothetical protein	2.38301E-08
ribC	0.3	-1.9	129.8	488.3	riboflavin biosynthesis protein	1.60E-03
CDR20291_0763	0.3	-1.9	383.3	1440.1	aconitate hydratase	5.62E-03
scrR	0.3	-1.9	357.4	1340.3	putative sucrose operon repressor (LacI-family transcriptional regulator)	1.02459E-07
CDR20291_2292	0.3	-1.9	43.9	164.2	cell wall hydrolase	1.82568E-08
CDR20291_2805	0.3	-1.9	59.0	220.3	putative thiolase	1.59439E-07
CDR20291_1584	0.3	-1.9	65.6	244.5	putative DNA binding protein	8.58E-03
srlB	0.3	-1.9	2.6	9.9	PTS system, glucitol/sorbitol-specific IIA component x2	7.33718E-08
rpsL	0.3	-1.9	223.4	829.0	30S ribosomal protein S12	2.77E-03
CDR20291_2093	0.3	-1.9	23.9	88.2	phage related cell wall hydrolase	6.07379E-08
sacA	0.3	-1.9	14.7	54.2	putative sucrose-6-phosphate hydrolase	2.57E-02
CDR20291_3314	0.3	-1.9	26.5	97.2	hypothetical protein	1.44534E-07
CDR20291_0639	0.3	-1.9	149.2	547.1	adenosylcobamide radical SAM protein	5.28203E-08
CDR20291_2082	0.3	-1.9	149.2	546.2	aminoacid ABC transporter permease	7.73E-11
CDR20291_3526	0.3	-1.9	142.8	519.9	hypothetical protein adjacent to soJ operon	2.23594E-07
gltX	0.3	-1.9	304.1	1101.0	glutamyl-tRNA synthetase	5.29357E-08

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
nrdE	0.3	-1.8	98.3	353.5	ribonucleoside-diphosphate reductase subunit alpha	4.54E-28
clpP1	0.3	-1.8	1787.4	6423.9	ATP dependent Clp protease proteolytic subunit	1.38914E-07
CDR20291_1468	0.3	-1.8	992.7	3507.8	putative lysophospholipase pseudogene	8.24867E-06
CDR20291_0996	0.3	-1.8	289.7	1022.6	hypothetical protein	1.05881E-07
pyc	0.3	-1.8	1134.9	3996.9	pyruvate carboxylase	9.20027E-08
CDR20291_0712	0.3	-1.8	626.7	2200.6	penicillin binding protein	1.11E-16
CDR20291_2863	0.3	-1.8	4.2	14.9	phosphosugar isomerase	2.85764E-07
CDR20291_2080	0.3	-1.8	271.3	947.9	aminoacid ABC transporter substrate binding protein	1.89782E-18
ileS	0.3	-1.8	190.0	663.8	isoleucyl tRNA synthetase	1.47243E-23
dhaB2	0.3	-1.8	36.4	126.4	glycerol dehydratase activator	1.42128E-18
CDR20291_0624	0.3	-1.8	11.1	38.5	RNA methylase	1.6398E-12
grdA	0.3	-1.8	10937.2	37909.5	glycine/sarcosine/betaine reductase complex protein A	3.46435E-21
CDR20291_1551	0.3	-1.8	21.8	75.6	lipoprotein	1.22E-04
prmA	0.3	-1.8	97.2	335.0	ribosomal protein L11 methyltransferase	7.68581E-17
CDR20291_3527	0.3	-1.8	395.3	1359.3	hypothetical protein adjacent to soj	1.17E-02
CDR20291_2078	0.3	-1.8	68.0	233.5	aminoacid ABC transporter ATP binding protein	1.15643E-14
CDR20291_0781	0.3	-1.8	11.9	40.6	hypothetical protein	3.56451E-12
CDR20291_1318	0.3	-1.8	430.9	1468.8	penicillin binding protein	1.76154E-19
CDR20291_0300	0.3	-1.8	32.9	111.9	Biotin synthase	1.11E-03
CDR20291_1712	0.3	-1.8	35.5	120.4	acyltransferase	4.44E-05
CDR20291_2380	0.3	-1.8	24.4	82.9	PTS system transporter subunit IIB	2.35753E-12
CDR20291_2099	0.3	-1.8	193.6	654.6	cell surface protein	3.19988E-13
bglA2	0.3	-1.8	39.2	132.2	6 phospho-beta-glucosidase	5.56E-05
CDR20291_2581	0.3	-1.8	11119.4	37432.1	putative sodium:dicarboxylate symporter	7.71091E-13
CDR20291_0734	0.3	-1.8	24.6	82.7	acyl-CoA dehyrogenase	1.29E-65
CDR20291_2342	0.3	-1.7	64.7	217.3	hypothetical protein	8.12083E-14
CDR20291_1617	0.3	-1.7	38.2	128.0	putative hydantoinase	2.22E-03

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
spaG	0.3	-1.7	4.8	16.1	lantibiotic ABC transporter permease	1.54391E-15
appA	0.3	-1.7	131.7	439.5	appA oligopeptide ABC transporter, substrate-binding protein	9.30E-08
rpsS	0.3	-1.7	20.3	67.5	30S ribosomal protein S19	3.88718E-12
CDR20291_1506	0.3	-1.7	4.9	16.4	ABC transporter permease	1.08E-03
CDR20291_1344	0.3	-1.7	1780.8	5897.7	transcriptional regulator	6.24263E-12
CDR20291_3131	0.3	-1.7	47.3	155.5	cation transporter	1.94E-03
CDR20291_0169	0.3	-1.7	77.7	254.8	drug/sodium antiporter	2.24355E-12
topA	0.3	-1.7	356.6	1165.6	DNA topoisomerase I	2.01837E-10
CDR20291_3318	0.3	-1.7	81.9	267.7	hypothetical protein	2.82909E-14
CDR20291_1942	0.3	-1.7	150.3	491.2	lipoprotein	3.15105E-20
CDR20291_0678	0.3	-1.7	86.5	282.3	aminoacid ABC transporter substrate binding protein	1.62693E-14
CDR20291_0995	0.3	-1.7	857.4	2791.3	radical SAM protein	1.50573E-12
CDR20291_0293	0.3	-1.7	44.4	143.9	hypothetical protein	3.97159E-13
CDR20291_0178	0.3	-1.7	2483.0	8013.7	cyclopropane-fatty-acyl-phospholipid synthase	4.21E-02
rpsQ	0.3	-1.7	31.0	100.0	30S ribosomal protein S17	1.3319E-13
CDR20291_0736	0.3	-1.7	22.5	72.4	electron transfer flavoprotein subunit alpha	9.38E-09
CDR20291_1709	0.3	-1.7	73.6	236.6	phosphoglycerate mutase	1.27838E-12
soj	0.3	-1.7	167.2	536.6	soj sporulation initiation inhibitor	7.54E-03
CDR20291_2079	0.3	-1.7	124.1	396.4	peptidase	1.72875E-15
grdB	0.3	-1.7	32524.5	103602.4	glycine reductase complex component B gamma subunit	5.55574E-12
gutD	0.3	-1.7	58.0	184.4	sorbitol-6-phosphate 2-dehydrogenase	8.06E-02
pyrH	0.3	-1.7	73.5	233.8	uridylate kinase	1.16447E-12
CDR20291_3212	0.3	-1.7	32.4	103.1	putative GTPase	8.67957E-13
panB	0.3	-1.7	92.6	294.0	3-methyl-2-oxobutanoate hydroxymethyltransferase	1.90E-22
pgk	0.3	-1.7	566.0	1792.7	phosphoglycerate kinase	1.15629E-25
aksA	0.3	-1.7	154.7	488.8	putative homocitrate/2-isopropylmalate synthase	1.46E-43
rplW	0.3	-1.7	91.6	289.3	50S ribosomal protein L23	1.33566E-13

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_3127	0.3	-1.7	27.4	86.6	two-component sensor histidine kinase	4.14E-35
prfA	0.3	-1.7	369.9	1166.0	peptide chain release factor 1	1.30696E-14
CDR20291_2711	0.3	-1.7	17.2	54.2	putative membrane protein	1.27E-71
grdE	0.3	-1.7	16511.7	51851.8	glycine reductase complex component B alpha and beta subunits	1.09048E-12
proS	0.3	-1.6	741.0	2325.5	prolyl-tRNA synthetase	3.12E-37
serS2	0.3	-1.6	286.4	898.0	seryl-tRNA synthetase	8.54505E-13
tsf	0.3	-1.6	253.2	790.3	elongation factor Ts	4.72E-04
CDR20291_0396	0.3	-1.6	26.1	81.3	conserved hypothetical protein	5.71882E-12
CDR20291_2341	0.3	-1.6	214.7	666.7	putative radical SAM superfamily protein	4.41777E-13
CDR20291_2394	0.3	-1.6	146.4	454.0	putative histidinol-phosphate aminotransferase	1.55E-06
rpsC	0.3	-1.6	395.6	1222.7	30S ribosomal protein S3	3.2616E-12
cmk	0.3	-1.6	151.4	467.9	cytidylate kinase	1.75415E-13
ntpD	0.3	-1.6	8.7	26.9	V-type sodium ATP synthase subunit D	3.12E-02
CDR20291_1616	0.3	-1.6	6.1	18.7	conserved hypothetical protein	1.3272E-12
CDR20291_2052	0.3	-1.6	67.5	206.8	DNA mismatch repair protein	5.45416E-16
grdC	0.3	-1.6	21695.9	66233.8	glycine/sarcosine/betaine reductase complex component C beta subunit	4.55279E-13
CDR20291_0999	0.3	-1.6	791.3	2414.3	conserved hypothetical protein	1.40342E-12
sbp	0.3	-1.6	14.9	45.4	putative membrane protein	1.93E-25
CDR20291_2455	0.3	-1.6	34.2	104.1	putative glycosyl hydrolase	3.48531E-13
rplN	0.3	-1.6	119.5	361.7	50S ribosomal protein L14	1.62E-06
ispH	0.3	-1.6	84.5	255.0	putative 4-hydroxy-3-methylbut-2-enyl diphosphate reductase	4.99144E-12
CDR20291_2014	0.3	-1.6	191.2	575.6	Xanthine/uracil/thiamine/ascorbate permease family protein	9.56547E-15
mtnN	0.3	-1.6	72.5	217.4	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	5.22124E-13
rplR	0.3	-1.6	25.9	77.5	50S ribosomal protein L18	1.38E-02
CDR20291_2081	0.3	-1.6	278.9	831.4	aminoacid ABC transporter permease	3.7592E-12

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_0638	0.3	-1.6	87.2	259.2	amino acid ABC transporter, substrate-binding protein	1.17E-09
kamA	0.3	-1.6	322.1	955.5	L-lysine 2,3-aminomutase	1.48426E-12
hydN	0.3	-1.6	44.8	132.9	electron transport protein	5.76E-31
aspC	0.3	-1.6	529.7	1558.4	aspartate aminotransferase	2.21105E-12
comR	0.3	-1.6	3321.6	9763.4	polyribonucleotide nucleotidyltransferase	1.82E-19
CDR20291_2978	0.3	-1.6	35.4	103.9	PTS system, IIabc component	3.96118E-13
rpsE	0.3	-1.6	88.1	258.8	30S ribosomal protein S5	4.88E-21
CDR20291_0595	0.3	-1.6	16.5	48.4	putative uncharacterized protein (pseudogene)	4.36002E-13
CDR20291_0206	0.3	-1.6	370.2	1084.9	putative transcription antiterminator	1.30E-42
addB	0.3	-1.5	469.5	1368.4	ATP-dependent nuclease subunit B	2.02E-03
CDR20291_0117	0.3	-1.5	19875.2	57667.7	putative subunit of oxidoreductase	1.30745E-12
CDR20291_2577	0.3	-1.5	26.6	77.3	putative lipoprotein	1.07E-03
CDR20291_1691	0.3	-1.5	25499.4	73687.2	putative nitrite and sulfite reductase subunit	2.26424E-30
plfA	0.3	-1.5	97.2	280.6	pyruvate formate-lyase activating enzyme	1.57138E-12
CDR20291_1505	0.3	-1.5	30.8	88.7	ABC transporter, ATP-binding protein	1.23E-21
hydN1	0.3	-1.5	310.7	894.0	electron transport protein	3.45E-07
nrdF	0.3	-1.5	45.2	129.7	ribonucleoside-diphosphate reductase subunit beta	3.25083E-17
hydA	0.3	-1.5	606.8	1740.0	hydrogenase	3.62913E-13
appF	0.3	-1.5	20.4	58.5	appF oligopeptide ABC transporter, ATP-binding protein	3.08051E-23
CDR20291_2483	0.3	-1.5	36.3	103.7	pseudoridylate synthase	4.30E-03
codY	0.4	-1.5	156.0	443.5	CodY	7.94154E-18
CDR20291_2434	0.4	-1.5	73.6	209.0	conserved hypothetical protein	8.77848E-14
CDR20291_3177	0.4	-1.5	16.5	46.9	hypothetical protein	6.45E-41
CDR20291_3021	0.4	-1.5	1353.3	3842.9	phophatidylethanolamine-binding regulatory protein	9.17688E-18
CDR20291_0356	0.4	-1.5	22.4	63.7	putative transcriptional regulator	9.84364E-20
ksgA	0.4	-1.5	50.9	143.6	dimethyladenosine transferase	4.24137E-15

Gene NAME	RealFC	Log fold 2	Mean Expression (Parent)	Mean expression (*sinRR*' mutant)	Function (actual or predicted)	adj. *p* value
CDR20291_3117	0.4	-1.5	29.0	81.8	putative membrane protein	1.51E-10
dapB1	0.4	-1.5	213.9	602.4	dihydrodipicolinate reductase	1.80277E-20
CDR20291_1241	0.4	-1.5	7.2	20.3	putative membrane protein	2.5902E-14
dltA	0.4	-1.5	556.6	1562.8	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)	3.17407E-13
garR	0.4	-1.5	7.4	20.8	2-hydroxy-3-oxopropionate reductase	7.00E-07
CDR20291_1528	0.4	-1.5	9.7	27.2	putative membrane-associated metal-dependent hydrolase	1.3642E-11
CDR20291_3441	0.4	-1.5	73.0	204.6	putative exported phosphoesterase	6.59231E-13
CDR20291_0444	0.4	-1.5	21.1	59.0	putative membrane protein	1.53E-04
CDR20291_1585	0.4	-1.5	256.2	715.8	putative lipoprotein	6.73759E-13
CDR20291_0509	0.4	-1.5	43.8	122.3	putative PEP-utilising kinase	3.72013E-14
rplE	0.4	-1.5	80.2	223.9	50S ribosomal protein L5	9.78E-02
CDR20291_2085	0.4	-1.5	272.0	757.6	putative formate dehydrogenase	3.18928E-31
ispF	0.4	-1.5	32.5	90.4	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	1.18259E-20
CDR20291_2740	0.4	-1.5	131.7	366.3	putative bifunctional protein	6.72504E-13
CDR20291_2397	0.4	-1.5	102.6	285.0	TetR-family transcriptional regulator	1.59E-07
CDR20291_2398	0.4	-1.5	241.7	671.3	putative transporter	5.59606E-14
CDR20291_3393	0.4	-1.5	421.7	1170.3	putative membrane protein	3.54661E-13
CDR20291_3361	0.4	-1.5	45.1	125.0	putative DNA-binding protein	6.00E-08
CDR20291_1500	0.4	-1.5	23.5	65.0	PTS system, IIb component	1.59713E-12
stp	0.4	-1.5	57.6	159.1	serine/threonine phosphatase	5.69009E-13
CDR20291_2828	0.4	-1.5	32.1	88.3	hypothetical protein	3.02E-52
CDR20291_0684	0.4	-1.5	61.5	169.1	putative reductase	3.41234E-11
CDR20291_0640	0.4	-1.5	193.7	530.6	conserved hypothetical protein	2.19E-31
rplB	0.4	-1.5	381.5	1044.8	50S ribosomal protein L2	9.97E-02
CDR20291_2007	0.4	-1.5	6.7	18.3	putative [2Fe-2S]-binding subunit of oxidoreductase	3.27649E-14
rplP	0.4	-1.5	228.2	624.0	50S ribosomal protein L16	3.59E-02
rplV	0.4	-1.4	29.1	79.4	50S ribosomal protein L22	1.72E-13
CDR20291_1484	0.4	-1.4	54.1	147.4	hypothetical protein	0.0000432

Gene NAME	**RealFC**	**Log fold 2**	**Mean Expression (Parent)**	**Mean expression (*sinRR*' mutant)**	**Function (actual or predicted)**	**adj. *p* value**
rplX	0.4	-1.4	85.0	231.5	50S ribosomal protein L24	5.81586E-15
rplF	0.4	-1.4	84.4	229.6	50S ribosomal protein L6	6.20E-03

Appendix E - Down regulated genes in J∆sinRR' relative to JIR8094

Gene ID (gene name)	Real Fold change	Log_2-Fold change	Mean Expression-Parent	Mean Expression-mutant	Function (known or predicted)	adj. *p* value
tlpB	30621.1	14.9	306.2	0.0	transposase-like protein B	1.8796E-14
CD630_30880	793.5	9.6	41762.1	52.6	cellobiose-phosphate degrading protein	6.4717E-19
CD630_30890	742.2	9.5	151931.9	204.7	PTS system glucose-like transporter subunit IIBC	5.7467E-09
treA	491.1	8.9	45178.6	92.0	trehalose-6-phosphate hydrolase	2.8813E-33
CD630_30870	79.2	6.3	358.8	4.5	RpiR family transcriptional regulator	7.9812E-09
treR	37.3	5.2	718.0	19.2	GntR family transcriptional regulator	1.0867E-09
CD630_10650	36.6	5.2	36.6	1.0	hypothetical protein	1.2538E-16
CD630_12382	22.3	4.5	53.5	2.4	hypothetical protein	9.7889E-10
ddl	19.9	4.3	403.2	20.2	D-alanine--D-alanine ligase	5.5399E-19
CD630_23440	17.7	4.1	6751.5	381.9	membrane protein	2.1207E-17
CD630_15110	16.0	4.0	53.4	3.3	hypothetical protein (spore coat protein)	0.00015606
CD630_23400	15.3	3.9	4488.3	293.6	hypothetical protein	7.8872E-11
mngA	15.0	3.9	70.7	4.7	PTS system 2-O-a-mannosyl-D-glycerate specific transporter subunit IIABC	4.8962E-17
CD630_01570	14.8	3.9	371.9	25.1	hypothetical protein	6.491E-16
CD630_17930	14.7	3.9	789.9	53.9	hypothetical protein	4.0788E-30
CD630_18311	14.6	3.9	16.1	1.1	hypothetical protein	2.5442E-13
hbd	14.3	3.8	17319.3	1215.0	4-hydroxybutyrate dehydrogenase	6.5372E-15
cat	13.9	3.8	19013.7	1371.0	4-hydroxybutyrate CoA-transferase	6.1844E-07
CD630_32090	13.4	3.7	13.1	1.0	PadR family transcriptional regulator	5.0472E-10
atpF	13.2	3.7	354.7	26.9	V-type ATP synthase subunit F	1.234E-18
abfD	12.3	3.6	22070.3	1795.8	gamma-aminobutyrate metabolism dehydratase/isomerase	3.9481E-17
atpA	12.3	3.6	2859.8	233.0	V-type ATP synthase subunit A	7.1701E-11
atpC	11.5	3.5	968.0	84.2	V-type ATP synthase subunit C	1.3796E-10

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
CD630_32060	11.4	3.5	11.4	1.0	membrane protein	4.0604E-16
CD630_06280	11.4	3.5	52.4	4.6	membrane protein	3.5317E-15
CD630_15970	11.3	3.5	673.3	59.7	hypothetical protein	1.0329E-09
sinR	11.0	3.5	33.5	3.0	HTH-type transcriptional regulator	2.1076E-15
CD630_23450	10.9	3.4	2722.7	250.4	LysR family transcriptional regulator	6.9052E-14
cat1	10.1	3.3	8412.0	832.7	succinyl-CoA:coenzyme A transferase	9.1197E-16
CD630_32100	10.1	3.3	28.9	2.9	hypothetical protein	6.5506E-14
atpB	10.1	3.3	1154.2	114.8	V-type ATP synthase subunit B	1.5403E-25
atpE	9.7	3.3	1234.2	127.2	V-type ATP synthase subunit E	1.7434E-08
atpK	9.3	3.2	1309.8	141.4	V-type ATP synthase subunit K	3.5905E-10
CD630_21110	9.2	3.2	469.0	50.9	DeoR family transcriptional regulator	6.3453E-74
CD630_06271	9.2	3.2	70.4	7.7	ferredoxin	7.3621E-26
atpD	8.9	3.2	430.5	48.5	V-type ATP synthase subunit D	2.7779E-14
CD630_29610	8.9	3.1	1205.9	136.1	hypothetical protein	1.4288E-09
glyA	8.7	3.1	2492.1	285.1	serine hydroxymethyltransferase	8.3319E-12
hbd	8.6	3.1	25875.6	2995.4	3-hydroxybutyryl-CoA dehydrogenase	1.223E-06
atpI	8.6	3.1	5228.8	607.8	V-type sodium ATP synthase subunit I	3.3266E-10
CD630_23460	8.6	3.1	148.5	17.3	membrane protein	1.7427E-15
CD630_19401	8.4	3.1	79.7	9.5	membrane protein	5.8263E-05
CD630_14400	8.4	3.1	951.9	113.7	LrgB family transporter	2.8093E-05
sucD	7.8	3.0	13420.5	1727.6	succinate-semialdehyde dehydrogenase	1.653E-07
CD630_14920	7.5	2.9	1567.5	208.5	sporulation-associated two-component sensor histidine kinase spo0A	4.0277E-25
thlA1	7.4	2.9	148975.4	20159.3	acetoacetyl-CoA thiolase 1	7.5391E-10
tRNA-Cys	7.3	2.9	173.0	23.6		8.7707E-16
tRNA-Cys	7.3	2.9	173.0	23.6		6.9693E-09
argG	7.1	2.8	266.0	37.4	argininosuccinate synthase	2.7348E-42
panB	6.8	2.8	600.1	88.1	hypothetical protein	8.6794E-17

Gene ID (gene name)	Real Fold change	Log_2-Fold change	Mean Expression-Parent	Mean Expression-mutant	Function (known or predicted)	adj. *p* value
CD630_29640	6.8	2.8	821.0	120.7	tRNA-binding protein	4.032E-05
CD630_17570	6.5	2.7	56.6	8.6	hypothetical protein	6.4672E-68
CD630_08470	6.5	2.7	136.6	21.0	hypothetical protein	2.641E-10
CD630_14930	6.4	2.7	506.6	78.6	3-methyladenine DNA glycosylase	1.0273E-26
CD630_17750	6.2	2.6	27.9	4.5	amino acid family ABC transporter permease	9.2675E-07
CD630_30370	6.2	2.6	1272.2	204.5	CarD family transcriptional regulator	2.1508E-05
CD630_08310	6.1	2.6	77.3	12.6	ATP-dependent RNA helicase	7.9875E-06
CD630_29650	6.0	2.6	1930.2	321.5	signaling protein	8.1347E-17
CD630_19500	6.0	2.6	663.4	110.7	membrane protein	1.1055E-14
CD630_08500	5.8	2.5	105.4	18.0	NifU-like protein	3.4682E-10
CD630_16330	5.8	2.5	683.0	117.4	serine protease	6.1084E-07
CD630_10530	5.8	2.5	5900.8	1024.8	oligonucleotide binding protein	3.73E-10
spoIVA	5.6	2.5	132.3	23.5	stage IV sporulation protein A	8.2144E-12
leuS	5.6	2.5	1147.6	204.7	leucyl-tRNA ligase	3.4905E-10
CD630_26870	5.5	2.5	18.9	3.4	hypothetical protein	7.7712E-11
CD630_08480	5.5	2.5	358.9	65.4	hypothetical protein	3.0795E-10
cspC	5.4	2.4	356.8	65.6	major cold shock protein CspC	1.6382E-16
CD630_32070	5.4	2.4	2179.4	404.2	multi antimicrobial extrusion protein	1.2398E-13
CD630_19660	5.4	2.4	29.1	5.4	acyl-CoA thioesterase	6.0558E-42
vanTG	5.3	2.4	220.7	41.4	alanine racemase 1	1.6308E-14
CD630_29620	5.3	2.4	515.3	96.9	hypothetical protein	5.6872E-05
etfA	5.2	2.4	31649.3	6038.2	electron transfer flavoprotein subunit alpha	7.6498E-07
CD630_19870	5.2	2.4	54.8	10.5	cell wall binding protein	1.2914E-16
CD630_26500	5.2	2.4	22.4	4.3	cell division protein FtsQ	6.5212E-11
CD630_01330	5.1	2.4	17.1	3.3	phosphoribosyl transferase	9.6664E-28
CD630_19410	5.1	2.3	410.1	81.0	hypothetical protein	8.6437E-18
hisG	5.1	2.3	44.3	8.7	ATP phosphoribosyltransferase	1.7661E-09
panB	5.0	2.3	1344.2	268.5	3-methyl-2-oxobutanoate hydroxymethyltransferase	1.558E-39
CD630_22930	5.0	2.3	15.9	3.2	hypothetical protein	3.1173E-05

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
crt2	5.0	2.3	21236.1	4257.0	3-hydroxybutyryl-CoA dehydratase	2.4368E-09
CD630_24050	5.0	2.3	118.2	23.8	MarR family transcriptional regulator	3.6686E-10
CD630_17250	4.9	2.3	22.7	4.6	AraC family transcriptional regulator	4.2833E-10
CD630_17610	4.9	2.3	37.4	7.6	phenazine biosynthesis PhzC/PhzF protein	5.8293E-12
CD630_16400	4.9	2.3	837.4	172.1	hypothetical protein	3.7332E-19
CD630_17300	4.9	2.3	1608.7	331.5	ATP-binding protein	1.6928E-12
CD630_26120	4.8	2.3	1419.3	293.6	amino acid permease	2.9994E-10
CD630_16410	4.8	2.3	438.8	90.9	hypothetical protein	0.00016526
CD630_14970	4.8	2.3	729.7	152.3	hypothetical protein	7.0877E-17
bcd2	4.8	2.3	53322.9	11157.0	butyryl-CoA dehydrogenase	1.0302E-22
panC	4.7	2.2	1559.8	329.2	pantoate--beta-alanine ligase	9.8003E-08
CD630_20280	4.7	2.2	614.8	130.5	aspartate racemase RacX	8.7664E-30
etfB	4.7	2.2	24231.8	5200.2	electron transfer flavoproteins subunit beta	2.4387E-23
CD630_08400	4.6	2.2	4082.5	892.4	isomerase/hydrolase	5.3902E-06
CD630_14280	4.6	2.2	21.2	4.6	HxlR family transcriptional regulator	4.7132E-07
CD630_08490	4.5	2.2	1294.2	286.5	glutamate carboxypeptidase	5.9491E-10
CD630_26130	4.5	2.2	1425.1	317.2	M24 family peptidase	1.7173E-10
CD630_17760	4.5	2.2	88.6	19.8	amino acid family ABC transporter ATP-binding protein	1.7364E-05
CD630_32080	4.5	2.2	712.5	159.5	MarR family transcriptional regulator	1.0281E-23
CD630_10310	4.5	2.2	68.4	15.3	cell wall anchored protein	2.6468E-10
CD630_19670	4.4	2.2	546.6	123.0	hypothetical protein	5.559E-13
CD630_35200	4.4	2.1	64.5	14.6	cation efflux protein	1.7322E-23
CD630_12730	4.4	2.1	97.0	22.1	DNA processing Smf single strand binding protein	1.5056E-10
CD630_26940	4.4	2.1	95.7	21.9	hypothetical protein	1.3293E-10
CD630_14390	4.4	2.1	566.0	129.4	LrgA family transporter	3.1622E-10
mngB	4.4	2.1	84.5	19.3	alpha-mannosidase	1.2729E-39
CD630_19520	4.4	2.1	421.1	96.6	LysR family transcriptional regulator	1.725E-09
CD630_21810	4.4	2.1	948.2	217.9	alpha/beta hydrolase	8.6164E-16

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
CD630_22590	4.3	2.1	356.2	83.3	hypothetical protein	4.2652E-10
sspA	4.2	2.1	184.0	43.5	small, acid-soluble spore protein alpha	1.3092E-11
CD630_26960	4.2	2.1	1426.6	337.2	pyridoxal phosphate-dependent transferase	3.4907E-11
CD630_22910	4.1	2.0	31.3	7.7	HTH-type transcriptional regulator	1.1325E-14
proC	4.0	2.0	440.1	109.0	pyrroline-5-carboxylate reductase	4.691E-11
CD630_06960	4.0	2.0	257.6	63.9	cation transport protein	6.6936E-15
CD630_22600	4.0	2.0	71.8	17.9	major facilitator superfamily transporter	1.0136E-14
CD630_28140	4.0	2.0	425.5	106.5	GntR family transcriptional regulator	5.5275E-12
CD630_12390	3.9	2.0	43.5	11.0	beta-lactams repressor	2.7253E-10
CD630_27270	3.9	2.0	237.8	61.4	membrane protein	6.1522E-24
CD630_28270	3.9	2.0	705.3	182.5	MarR family transcriptional regulator	2.798E-16
CD630_13440	3.9	1.9	127.2	33.0	cell wall biosynthesis protein	1.8601E-16
CD630_16460	3.9	1.9	459.7	119.4	peptidase	4.0107E-14
CD630_16290	3.8	1.9	668.8	175.4	hypothetical protein	1.8146E-15
pdcA	3.7	1.9	1605.2	431.8	diguanylate kinase signaling protein	6.139E-12
CD630_10320	3.7	1.9	41.2	11.1	hypothetical protein	2.0624E-12
potA	3.7	1.9	43.5	11.7	spermidine/putrescine ATP-binding protein	8.308E-15
dapA2	3.7	1.9	324.8	87.9	dihydrodipicolinate synthase 2	0.00022985
rpmB	3.7	1.9	1922.7	521.0	50S ribosomal protein L28	3.0323E-11
CD630_06720	3.7	1.9	306.2	83.1	hypothetical protein	1.7025E-28
CD630_08410	3.7	1.9	21097.8	5766.6	5-aminoimidazole-4-carboxamide ribonucleotide transformylase	2.204E-19
CD630_14140	3.6	1.8	351.8	97.9	oxidoreductase Fe-S subunit	7.5258E-13
CD630_14070	3.6	1.8	230.6	64.3	GntR family transcriptional regulator	2.0722E-37
CD630_17600	3.5	1.8	28.6	8.1	iron-sulfur protein	2.9898E-05
CD630_01800	3.5	1.8	64.8	18.3	hypothetical protein	9.8005E-13
CD630_25020	3.5	1.8	2143.5	610.6	pyridoxal phosphate-dependent transferase	1.2542E-44

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
CD630_17380	3.5	1.8	476.9	135.8	two-component sensor histidine kinase	5.6551E-06
CD630_05460	3.5	1.8	802.6	229.1	hypothetical protein	2.845E-15
CD630_30730	3.5	1.8	1035.9	296.3	membrane protein	9.5876E-12
CD630_35610	3.5	1.8	70.0	20.1	hypothetical protein	1.7577E-05
spo0A	3.5	1.8	11191.6	3230.4	stage 0 sporulation protein A	1.3737E-09
CD630_10790	3.4	1.8	613.0	178.3	LysR family transcriptional regulator	4.985E-05
CD630_15200	3.4	1.8	127.8	37.8	membrane protein	0.00017952
CD630_22560	3.4	1.8	35.5	10.5	PTS system transporter subunit IIC	2.9079E-17
CD630_01810	3.4	1.7	271.8	81.0	membrane-associated metalloprotease	1.332E-16
CD630_16450	3.3	1.7	481.8	144.1	membrane protein	7.2534E-10
CD630_18930	3.3	1.7	1580.7	472.9	oligonucleotide binding regulator	3.0663E-13
CD630_21520	3.3	1.7	150.6	45.2	glutamyl-aminopeptidase	4.2913E-05
CD630_21530	3.3	1.7	501.3	150.8	hypothetical protein	3.8429E-19
def1	3.3	1.7	247.8	74.7	peptide deformylase 1	9.0138E-23
CD630_15080	3.3	1.7	353.0	106.4	iron-sulfur binding protein	6.2141E-10
hisZ	3.3	1.7	37.4	11.3	ATP phosphoribosyltransferase regulatory subunit	1.3073E-10
CD630_06520	3.3	1.7	247.8	74.8	GntR family transcriptional regulator	4.3098E-05
mraW	3.3	1.7	256.1	77.5	16S rRNA m(4)C1402 methyltransferase	0.00022359
CD630_12210	3.3	1.7	55.6	16.8	membrane protein	2.1463E-09
spoIIAA	3.3	1.7	122.5	37.1	anti-sigma F factor antagonist	1.118E-09
CD630_20270	3.3	1.7	1477.7	448.5	N-carbamoyl-L-amino acid hydrolase	3.3625E-15
CD630_02791	3.3	1.7	46.8	14.2	hypothetical protein	3.3115E-10
CD630_17390	3.3	1.7	644.5	197.9	two-component response regulator	1.6581E-10
CD630_14710	3.2	1.7	944.3	291.4	membrane protein	3.0061E-10
CD630_05840	3.2	1.7	253.1	78.4	diguanylate kinase signaling protein	3.1213E-14
hydD	3.2	1.7	150.2	46.5	hydrolase	1.6441E-10
CD630_03110	3.2	1.7	89.9	27.9	hypothetical protein	7.5237E-05

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
vanZ	3.2	1.7	216.2	67.7	teicoplanin resistance protein	4.3892E-16
CD630_25490	3.2	1.7	123.3	38.6	sugar family ABC transporter permease	1.6254E-19
CD630_19490	3.2	1.7	93.8	29.4	two-component sensor histidine kinase	7.1064E-21
CD630_09930	3.2	1.7	951.8	298.7	thioesterase	5.0077E-10
hisC	3.2	1.7	38.3	12.1	histidinol-phosphate aminotransferase	1.6619E-14
CD630_28120	3.2	1.7	68.0	21.5	gluconate permease	8.6575E-13
hisK	3.2	1.7	112.2	35.6	histidinol-phosphatase	2.7741E-06
secG	3.1	1.6	522.6	167.4	protein-export membrane protein SecG	2.7792E-07
CD630_15831	3.1	1.6	509.2	164.0	hypothetical protein	1.8648E-13
CD630_19630	3.1	1.6	274.8	88.7	methyltransferase	5.8425E-13
CD630_23960	3.1	1.6	727.3	235.3	hypothetical protein	7.8366E-11
CD630_21540	3.1	1.6	1498.2	485.1	ATP/GTP-binding protein	1.394E-33
CD630_28260	3.1	1.6	2463.8	798.7	aldolase	3.7503E-40
npdA	3.1	1.6	220.9	71.9	NAD-dependent deacetylase	5.6747E-08
oppF	3.1	1.6	23246.0	7574.7		1.4318E-26
ftnA	3.1	1.6	2630.5	859.2	ferritin	4.6602E-12
CD630_19640	3.1	1.6	547.8	179.2	transcriptional regulator	3.0347E-09
CD630_16510	3.0	1.6	290.2	95.4	signaling protein	7.4795E-12
CD630_07960	3.0	1.6	448.9	148.6	LacI family transcriptional regulator	1.2987E-09
CD630_34890	3.0	1.6	64.9	21.6	oligoendopeptidase F	1.2652E-13
CD630_10600	3.0	1.6	173.5	57.8	pseudouridylate synthase	1.5071E-11
CD630_21650	3.0	1.6	119.5	39.9	HTH-type transcriptional regulator	6.6783E-16
CD630_17580	3.0	1.6	366.3	123.1	radical SAM superfamily protein	9.961E-10
CD630_13020	3.0	1.6	810.3	272.3	MazG-like pyrophosphohydrolase	6.9457E-10
CD630_14750	3.0	1.6	119.0	40.1	hydrolase	2.2296E-10
spoVD	3.0	1.6	242.9	81.9	stage V sporulation protein D	1.4327E-09
CD630_17640	3.0	1.6	282.2	95.4	phenylalanyl-tRNA synthetase subunit beta	3.262E-10
CD630_26150	3.0	1.6	262.8	88.8	TetR family transcriptional regulator	7.7685E-09

Gene ID (gene name)	**Real Fold change**	**Log_2-Fold change**	**Mean Expression-Parent**	**Mean Expression-mutant**	**Function (known or predicted)**	adj. *p* value
CD630_05830	2.9	1.5	254.8	87.0	diguanylate kinase signaling protein	7.3157E-46
fapR	2.9	1.5	2293.5	788.8	fatty acid biosynthesis transcriptional regulator	6.7335E-10
CD630_16520	2.9	1.5	1885.9	649.1	tellurium resistance protein	1.0603E-09
glyQ	2.9	1.5	3752.4	1293.9	glycyl-tRNA ligase subunit alpha	2.58E-32
CD630_26890	2.9	1.5	218.3	75.3	lipoprotein	3.0566E-09
CD630_10020	2.9	1.5	30.7	10.7	nucleotide/oligonucleotide binding protein	2.571E-13
CD630_17650	2.9	1.5	385.7	135.1	NUDIX family hydrolase	6.636E-15
hydR	2.8	1.5	48.7	17.1	TetR family transcriptional regulator	2.6032E-45
CD630_06530	2.8	1.5	205.8	72.4	multidrug family ABC transporter ATP-binding protein	1.4978E-21
sigF	2.8	1.5	492.0	173.8	RNA polymerase sigma-F factor	1.7166E-06
CD630_14760	2.8	1.5	1271.9	449.8	signaling protein	1.9707E-08
CD630_35700	2.8	1.5	286.8	101.7	peptidase	3.5033E-11
mviN	2.8	1.5	353.2	125.5	transmembrane virulence factor	3.1016E-45
CD630_08220	2.8	1.5	46.5	16.6	multidrug family ABC transporter ATP-binding protein	3.7086E-10
spoIIAB	2.8	1.5	355.0	126.6	anti-sigma F factor	0.00019304
oppD	2.8	1.5	23462.4	8401.1	ABC transporter ATP-binding protein	4.8763E-13
spoVS	2.8	1.5	1078.1	386.1	stage V sporulation protein S	7.4213E-15
CD630_16300	2.8	1.5	57.0	20.6	membrane protein	7.8906E-24
CD630_10390	2.8	1.5	309.4	112.2	hypothetical protein	3.0297E-12
CD630_24920	2.8	1.5	5895.5	2142.2	sporulation-associated two-component sensor histidine kinase spo0A	5.7815E-28

Appendix F - Up regulated genes in J∆sinRR' relative to JIR8094

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_11870	0.0	-10.6	0.0	15.9	hypothetical protein	3.7463E-11
eutM	0.0	-9.2	0.2	117.3	ethanolamine carboxysome structural protein	9.5475E-10
ermB	0.0	-8.2	4.5	1293.1	ribosomal RNA adenine N-6-methyltransferase	4.61E-14
CD630_21690	0.0	-7.4	9.7	1639.1	iron-sulfur binding protein	1.9242E-05
hcp	0.0	-7.4	91.6	15394.7	hybrid cluster protein	3.7994E-12
eutC	0.0	-7.4	2.2	359.7	ethanolamine ammonia-lyase small subunit	8.8192E-15
eutT	0.0	-7.1	0.4	58.9	ethanolamine corrinoid cobalamin adenosyltransferase	3.4563E-26
CD630_23240	0.0	-7.0	63.6	8229.8	sugar-phosphate dehydrogenase	9.5357E-11
CD630_23230	0.0	-6.8	96.4	11082.7	sugar-phosphate dehydrogenase	4.6144E-12
eutB	0.0	-6.6	8.3	815.1	ethanolamine ammonia lyase large subunit	1.2577E-13
eutA	0.0	-6.4	4.2	353.1	reactivating factor for ethanolamine ammonia lyase	3.0457E-05
eutE	0.0	-6.1	6.1	411.8	ethanolamine acetaldehyde oxidoreductase	1.4279E-09
eutK	0.0	-6.0	4.9	311.7	ethanolamine carboxysome structural protein	9.2379E-12
CD630_32770	0.0	-6.0	1.2	72.9	PTS system mannose/fructose/sorbose transporter subunit IIC	1.5781E-29
CD630_11880	0.0	-5.9	0.6	35.6	gamma-glutamyl-gamma-aminobutyrate hydrolase	4.273E-10
CD630_23250	0.0	-5.9	152.7	9040.1	PTS system fructose/mannitol family transporter subunit IIC	1.8338E-45
CD630_15460	0.0	-5.7	154.1	7877.0	channel-forming hemolysin	5.9882E-09
CD630_19210	0.0	-5.4	1.0	42.5	ethanolamine utilization protein	5.2398E-05
CD630_23260	0.0	-5.4	103.1	4409.0	PTS system fructose/mannitol family transporter subunit IIB	7.9678E-13
CD630_32760	0.0	-5.2	1.4	54.2	PTS system mannose/fructose/sorbose transporter subunit IID	7.4394E-19
eutL	0.0	-5.2	1.4	50.3	ethanolamine utilization protein EutL	1.5664E-09
CD630_27970	0.0	-5.1	535.0	18613.1	calcium-binding adhesion protein	5.6219E-29
CD630_19230	0.0	-5.1	1.6	56.3	ethanolamine utilization protein	3.0935E-09
CD630_11890	0.0	-5.1	3.3	111.6	amino acid/polyamine transporter I	2.8217E-15
CD630_01900	0.0	-5.0	43.2	1370.9	membrane protein	6.7302E-08
CD630_08600	0.0	-4.9	2.5	75.3	hypothetical protein	2.2014E-05

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_02930	0.0	-4.9	54.2	1597.1	bacitracin/multidrug family ABC transporter ATP-binding protein	1.6992E-14
mtlR	0.0	-4.9	596.4	17242.5	PTS operon transcription antiterminator	5.8487E-06
eutG	0.0	-4.8	3.5	100.2	ethanolamine iron-dependent alcohol dehydrogenase	2.4228E-13
CD630_02920	0.0	-4.8	40.2	1137.6	HTH-type transcriptional regulator	6.3552E-08
etfB	0.0	-4.8	12.6	350.8	electron transfer flavoprotein subunit beta	3.3689E-06
CD630_02940	0.0	-4.8	13.7	378.5	bacitracin/multidrug family ABC transporter permease	3.9669E-14
eutQ	0.0	-4.7	2.8	75.0	ethanolamine utilization protein	8.6904E-74
CD630_23270	0.0	-4.7	24.6	642.5	PTS system fructose/mannitol family transporter subunit IIA	9.5872E-32
mtlD	0.0	-4.7	652.7	16837.3	mannitol-1-phosphate 5-dehydrogenase	2.5419E-15
rpiB1	0.0	-4.7	12.5	321.4	ribose-5-phosphate isomerase B	1.8093E-09
mtlF	0.0	-4.7	128.3	3266.1	PTS system mannitol-specific transporter subunit IIA	5.0288E-10
CD630_30360	0.0	-4.7	13.6	345.1	major facilitator superfamily transporter	1.0528E-14
CD630_08030	0.0	-4.6	18.7	468.8	acyl-CoA dehydrogenase	4.87E-12
CD630_02921	0.0	-4.6	10.6	265.4	hypothetical protein	2.9496E-15
CD630_26990	0.0	-4.6	26.8	647.4	membrane protein	5.158E-15
mtlA	0.0	-4.6	333.8	7987.7	PTS system mannitol-specific transporter subunit IICB	1.916E-07
xdhA	0.0	-4.6	1.4	33.5	xanthine dehydrogenase FAD-binding subunit	1.6609E-16
CD630_19200	0.0	-4.6	1.9	44.8	propanediol utilization phosphotransacylase	7.4049E-15
CD630_20820	0.0	-4.5	3.5	80.8	amidohydrolase	2.484E-08
CD630_17410	0.0	-4.5	48.8	1112.0	hypothetical protein	6.3721E-13
CD630_10921	0.0	-4.5	0.2	4.5	conjugative transposon protein	2.6694E-05
etfA	0.0	-4.5	24.8	561.4	electron transfer flavoprotein subunit alpha	3.4592E-10
CD630_08620	0.0	-4.5	7.6	171.2	PTS system lactose/cellobiose-family transporter subunit IIC	8.9318E-12
CD630_01660	0.0	-4.4	11.9	253.3	peptidase	1.9149E-09
CD630_26970	0.0	-4.4	139.2	2899.3	peptidase	0.0004192
CD630_22830	0.0	-4.3	5.3	107.0	PTS operon transcription antiterminator	7.6715E-30
CD630_26980	0.0	-4.3	20.9	420.7	hypothetical protein	3.1413E-10
CD630_20890	0.1	-4.3	17.9	349.4	metal-dependent hydrolase	1.4245E-09
CD630_22800	0.1	-4.3	2.0	39.5	PTS system fructose/mannitol family transporter subunit IIC	2.4296E-12
CD630_27960	0.1	-4.3	239.3	4628.9	cell surface protein	7.7944E-28

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_00410	0.1	-4.3	3.2	61.1	PTS system galactitol-specific transporter subunit IIA	8.8871E-19
CD630_32780	0.1	-4.2	0.6	12.1	PTS system mannose/fructose/sorbose transporter subunit IIA	1.9242E-29
CD630_08020	0.1	-4.2	11.8	218.4	CoA-transferase	3.5043E-10
dpaL	0.1	-4.2	7.4	137.1	diaminopropionate ammonia-lyase	2.6933E-16
CD630_32430	0.1	-4.1	287.9	4818.2	hypothetical protein	5.5559E-10
CD630_22810	0.1	-4.0	2.4	40.3	PTS system fructose/mannitol family transporter subunit IIB	7.0971E-16
CD630_14630	0.1	-4.0	31.8	519.4	hypothetical protein	7.3799E-14
eutH	0.1	-4.0	7.0	111.9	ethanolamine utilization protein EutH	1.4181E-12
CD630_00650	0.1	-4.0	488.0	7672.4	NADP-dependent dehydrogenase	2.5073E-11
ord	0.1	-4.0	9.9	154.4	2,4-diaminopentanoate dehydrogenase	1.35E-08
CD630_27370	0.1	-4.0	306.9	4772.8	nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase	7.7139E-09
CD630_22820	0.1	-3.9	1.3	20.1	PTS system fructose/mannitol family transporter subunit IIA	5.603E-14
CD630_08610	0.1	-3.9	0.8	12.6	PTS system lactose/cellobiose-family transporter subunit IIB	8.7864E-13
CD630_32750	0.1	-3.9	15.8	234.5	phosphosugar isomerase	1.1159E-12
CD630_20910	0.1	-3.9	15.4	222.8	xanthine/uracil permease	3.1385E-14
prdB	0.1	-3.8	1058.4	15031.9	proline reductase	6.4612E-30
CD630_03400	0.1	-3.8	362.8	5070.4	hypothetical protein	4.2176E-10
crt1	0.1	-3.8	4.8	66.0	3-hydroxybutyryl-CoA dehydratase	1.7551E-15
ccpA	0.1	-3.8	242.2	3295.6	LacI family transcriptional regulator	8.2261E-11
CD630_10800	0.1	-3.7	435.9	5570.1	lipoprotein	4.2031E-16
CD630_08640	0.1	-3.7	16.4	208.5	maltose-6'-phosphate glucosidase	3.4903E-23
CD630_20840	0.1	-3.7	11.2	141.0	peptidase	8.6875E-15
CD630_21151	0.1	-3.7	336.7	4233.1	hypothetical protein	1.7245E-12
CD630_33680	0.1	-3.6	12.7	155.1	ribosome biogenesis GTPase RsgA	6.3486E-43
CD630_08010	0.1	-3.6	16.6	202.0	divalent anion symporter family permease	8.7294E-13
CD630_20900	0.1	-3.6	11.6	139.6	xanthine/uracile permease	1.6735E-15
CD630_01650	0.1	-3.5	10.2	117.9	amino acid transporter	8.932E-08
CD630_20830	0.1	-3.5	5.1	58.6	D-hydantoinase	2.2027E-20
ptsG-BC	0.1	-3.5	70.9	804.2	PTS system glucose-specific transporter subunit IIBC	6.4371E-15
atpC	0.1	-3.5	626.6	7031.8	ATP hydrolase subunit epsilon	8.4928E-12
CD630_03410	0.1	-3.5	53.6	600.3	hypothetical protein	1.4855E-10

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_07990	0.1	-3.5	315.1	3481.0	beta-alanine CoA-transferase	2.1961E-16
CD630_17400	0.1	-3.5	28.1	309.7	glycine/sarcosine/betaine reductase complex component B subunits alpha and beta	9.9615E-38
CD630_04900	0.1	-3.5	5.9	64.8	sugar-phosphate dehydrogenase	1.0148E-16
CD630_24160	0.1	-3.5	8.6	94.2	PTS operon transcription antiterminator	8.2822E-14
cooS	0.1	-3.4	2561.6	27454.8	oxidoreductase Fe-S subunit	1.7168E-10
CD630_27380	0.1	-3.4	187.0	1989.1	cytosine permease	6.625E-06
CD630_00400	0.1	-3.4	20.2	214.8	PTS operon transcription antiterminator	1.2649E-09
xdhA	0.1	-3.4	19.0	199.3	xanthine dehydrogenase molybdenum binding subunit	9.3316E-13
nth	0.1	-3.4	87.1	907.0	endonuclease III	5.0355E-10
uxaA	0.1	-3.4	46.7	478.5	D-galactate dehydratase/altronate hydrolase	1.3961E-21
CD630_02860	0.1	-3.3	5.0	51.3	PTS system mannose/fructose/sorbose transporter subunit IIA	2.475E-09
CD630_29360	0.1	-3.3	25.6	259.4	hypothetical protein	1.313E-27
CD630_32840	0.1	-3.3	172.9	1727.6	serine protease, HrtA family	2.3953E-09
atpD	0.1	-3.3	6448.7	64295.2	F0F1 ATP synthase subunit beta	2.0126E-14
prdD	0.1	-3.3	374.1	3723.4	proline reductase PrdD	7.7458E-11
CD630_01760	0.1	-3.3	9450.8	91956.3	oxidoreductase NAD/FAD binding subunit	9.3735E-11
prdE	0.1	-3.3	264.1	2565.6	proline reductase PrdE	1.4756E-07
CD630_36140	0.1	-3.3	382.6	3708.8	hypothetical protein	5.5034E-17
fliG	0.1	-3.3	5.1	48.5	flagellar motor switch protein FliG	1.9242E-22
CD630_02840	0.1	-3.3	7.5	71.5	PTS system mannose/fructose/sorbose transporter subunit IIA	5.6945E-06
eutS	0.1	-3.2	3.6	34.1	ethanolamine carboxysome structural protein	1.5305E-14
tkt	0.1	-3.2	931.8	8669.8	transketolase	4.8048E-15
CD630_21340	0.1	-3.2	64.1	593.7	signaling protein	3.8217E-12
CD630_29370	0.1	-3.2	9.0	83.2	hypothetical protein	1.2836E-44
prdA	0.1	-3.2	2437.0	22371.5	D-proline reductase PrdA	2.6722E-22
ptsG-A	0.1	-3.2	27.2	248.0	PTS system glucose-specific transporter subunit IIA	2.315E-16
cooS	0.1	-3.2	13136.2	119529.3	carbon monoxide dehydrogenase	2.4375E-05
gpmI	0.1	-3.1	724.9	6364.9	phosphoglyceromutase	8.6384E-12
CD630_24170	0.1	-3.1	22.8	198.3	PTS system glucitol/sorbitol-specific transporter subunit IIB	1.2218E-11
CD630_17940	0.1	-3.1	5.1	43.3	hypothetical protein	8.673E-08

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_25400	0.1	-3.1	298.4	2535.3	coenzyme A disulfide reductase	1.4687E-09
fliH	0.1	-3.1	11.6	98.2	flagellar assembly protein FliH	1.827E-17
fliF	0.1	-3.1	16.2	137.4	flagellar MS-ring protein	3.277E-05
CD630_29440	0.1	-3.1	78.7	660.8	recombination function protein	1.8399E-26
tkt	0.1	-3.1	319.4	2666.1	transketolase	6.151E-17
CD630_30380	0.1	-3.0	169.8	1389.8	hypothetical protein	5.4665E-37
CD630_29450	0.1	-3.0	60.6	491.7	bacteriophage resistance protein	7.8816E-11
CD630_24870	0.1	-3.0	22.1	179.0	PTS system fructose-like transporter subunit IIB	8.4664E-42
CD630_24860	0.1	-3.0	66.8	533.4	PTS system fructose-like transporter subunit IIC	3.5038E-16
CD630_15430	0.1	-3.0	8.3	66.0	FMN-dependent NADH-azoreductase	9.6058E-10
CD630_02880	0.1	-3.0	37.1	295.9	PTS system mannose/fructose/sorbose transporter subunit IIC	6.4753E-10
CD630_20860	0.1	-3.0	18.5	145.1	NAD(P)H-dependent FMN reductase	8.1131E-28
CD630_04970	0.1	-3.0	47.1	368.6	conjugative transposon protein	3.0242E-14
hadI	0.1	-2.9	40807.1	313661.4	activator of 2-hydroxyisocaproyl-CoA dehydratase	5.6095E-33
srlM	0.1	-2.9	21.4	163.8	sorbitol operon activator protein	2.4434E-27
CD630_07500	0.1	-2.9	93.4	712.8	amino acid family ABC transportersubstrate-binding protein	4.3961E-12
clpP1	0.1	-2.9	2016.1	15327.1	ATP-dependent Clp protease proteolytic subunit 1	6.1912E-10
rph	0.1	-2.9	1144.0	8674.9	bifunctional tRNA nucleotidyltransferase/nucleoside-triphosphatase	5.2473E-10
ortB	0.1	-2.9	15.2	115.0	2-amino-4-ketopentanoate thiolase subunit beta	1.0136E-10
CD630_32380	0.1	-2.9	585.0	4380.1	proline reductase PrdE-like protein	1.0903E-09
CD630_10780	0.1	-2.9	7.7	57.1	PTS system mannose/fructose/sorbose transporter subunit IID	5.4649E-21
adhE	0.1	-2.9	18536.2	137608.6	bifunctional acetaldehyde-CoA/alcohol dehydrogenase	1.5227E-39
CD630_29380	0.1	-2.9	6.5	48.5	hypothetical protein	1.4034E-15
CD630_20870	0.1	-2.9	25.7	190.4	xanthine dehydrogenase	8.0893E-42
CD630_33070	0.1	-2.9	195.2	1437.9	phosphoesterase	2.4516E-13
bglA	0.1	-2.9	37.8	278.2	6-phospho-beta-glucosidase	2.5394E-14
hadB	0.1	-2.9	60320.7	442583.3	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase B	5.0633E-05
CD630_10940	0.1	-2.9	13.2	95.8	RNA polymerase sigma factor	2.8608E-09

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_07980	0.1	-2.9	69.7	504.0	membrane protein	3.0008E-07
CD630_24130	0.1	-2.8	108.8	782.9	NAD(P)-binding protein	6.5421E-12
rplX	0.1	-2.8	38.4	271.7	50S ribosomal protein L24	3.9528E-09
CD630_02890	0.1	-2.8	75.4	533.7	PTS system mannose/fructose/sorbose transporter subunit IID	3.3695E-10
CD630_24180	0.1	-2.8	5.9	41.0	PTS system glucitol/sorbitol-specific transporter subunit IIC	6.9726E-08
23s rRNA	0.1	-2.8	786.6	5464.3		1.5132E-25
CD630_29460	0.1	-2.8	44.8	311.3	hypothetical protein	3.637E-16
clpX	0.1	-2.8	4675.1	31708.7	ATP-dependent protease ATP-binding subunit ClpX	1.7991E-14
rpe	0.1	-2.7	14.4	96.4	ribulose-phosphate 3-epimerase	1.1352E-08
tpi	0.2	-2.7	175.5	1168.2	triosephosphate isomerase	6.8357E-20
CD630_04960	0.2	-2.7	23.8	158.2	conjugative transposon protein	1.5193E-07
rbsA	0.2	-2.7	178.2	1179.1	ribose-specific ABC transporter ATP-binding protein	2.7552E-05
acdB	0.2	-2.7	65496.3	432334.4	acyl-CoA dehydrogenase	4.6903E-40
CD630_24880	0.2	-2.7	3.4	22.1	PTS system fructose-like transporter subunit IIA	1.6141E-09
srlR	0.2	-2.7	138.1	901.4	PTS operon transcription antiterminator	0.00300867
rpsH	0.2	-2.7	33.1	214.7	30S ribosomal protein S8	7.2074E-14
srlEa	0.2	-2.7	46.6	299.1		7.1283E-16
atpG	0.2	-2.7	2381.9	15094.4	F1 ATP synthase subunit gamma	1.2596E-15
hadA	0.2	-2.7	77696.1	491271.2	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	1.0906E-09
CD630_32320	0.2	-2.6	6806.5	42488.3	hypothetical protein	1.9215E-15
pgk	0.2	-2.6	445.6	2770.8	phosphoglycerate kinase	5.7481E-15
CD630_20780	0.2	-2.6	12.4	76.2	S-adenosylhomocysteine deaminase	4.0946E-45
CD630_30940	0.2	-2.6	16.7	102.3	sigma-54 dependent transcriptional regulator	6.0563E-12
dnaJ	0.2	-2.6	2836.5	17363.8	chaperone protein DnaJ	2.2696E-16
CD630_12800	0.2	-2.6	385.8	2356.8	NifU-like protein	1.6089E-16
uxaA	0.2	-2.6	5.8	35.3	carbohydrate hydrolase, SAF domain-containing protein	7.017E-10
uvrC	0.2	-2.6	251.3	1524.5	Exconuclease ABC subunit C	4.3075E-44
CD630_24890	0.2	-2.6	15.0	91.0	PTS operon transcription antiterminator	2.3495E-07
CD630_00511	0.2	-2.6	4.9	29.5	hypothetical protein	6.2433E-09
CD630_29470	0.2	-2.6	8.2	49.8	hypothetical protein	5.0435E-15
rplE	0.2	-2.6	44.6	268.8	50S ribosomal protein L5	2.1367E-09

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_04130	0.2	-2.6	1.3	7.7	single-strand DNA-binding protein	5.4267E-15
rpsQ	0.2	-2.6	17.1	101.9	30S ribosomal protein S17	2.6095E-19
CD630_07390	0.2	-2.6	47.2	280.3	hypothetical protein	1.9654E-18
rplR	0.2	-2.6	14.9	88.4	50S ribosomal protein L18	1.6739E-11
serS1	0.2	-2.6	2958.8	17436.0	seryl-tRNA ligase	3.5851E-05
eutV	0.2	-2.6	45.8	269.9	ethanolamine specific two-component response regulator	2.1868E-30
CD630_00430	0.2	-2.6	17.2	100.7	PTS system galactitol-specific transporter subunit IIC	8.8825E-09
CD630_23680	0.2	-2.5	96.8	559.9	hypothetical protein	0.00014597
CD630_25100	0.2	-2.5	10.5	60.8	PTS system glucose-like transporter subunit IIBC	4.9892E-05
CD630_25680	0.2	-2.5	29.2	168.4	PTS system mannose-specfic transporter subunit IIC	1.2287E-18
CD630_07840	0.2	-2.5	240.8	1377.2	N-acetylmuramoyl-L-alanine amidase	1.0846E-18
fhuC	0.2	-2.5	42.5	239.6	ferrichrome-specific ABC transporter ATP-binding protein	3.6141E-21
flgB	0.2	-2.5	2.1	11.6	flagellar basal-body rod protein FlgB	5.145E-10
CD630_15680	0.2	-2.5	490.4	2748.0	hypothetical protein	3.811E-32
CD630_30400	0.2	-2.5	114.8	641.6	hypothetical protein	1.5029E-05
rbsB	0.2	-2.5	166.3	925.9	ribose-specific ABC transporter substrate-binding protein	5.095E-09
hadC	0.2	-2.5	31528.3	174691.3	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase C	3.8355E-11
23s rRNA	0.2	-2.5	5.8	31.9		2.3989E-10
rbsK	0.2	-2.5	238.4	1308.1	PfkB family ribokinase	1.2783E-08
5s rRNA	0.2	-2.4	77.1	419.6		1.3707E-09
rbsC	0.2	-2.4	73.0	396.0	ribose-specific ABC transporter permease	1.6935E-14
CD630_30390	0.2	-2.4	282.5	1526.8	ATPase	2.5369E-17
CD630_29430	0.2	-2.4	161.6	871.8	replication protein	0.00011886
tRNA-Val	0.2	-2.4	92.5	495.0		1.7209E-15
CD630_04980	0.2	-2.4	275.2	1465.9	cell-division FtsK/SpoIIIE-family protein	0.00017029
rplF	0.2	-2.4	46.7	248.3	50S ribosomal protein L6	4.5238E-10
CD630_14740	0.2	-2.4	25072.7	132774.9	rubrerythrin	8.6059E-15
gapA	0.2	-2.4	743.9	3935.9	glyceraldehyde-3-phosphate dehydrogenase	1.539E-14
CD630_19300	0.2	-2.4	37.4	196.9	hypothetical protein	6.3997E-09
CD630_02870	0.2	-2.4	8.3	43.7	PTS system mannose/fructose/sorbose transporter subunit IIB	6.9995E-10

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
srlD	0.2	-2.4	64.0	336.9	sorbitol-6-phosphate dehydrogenase	1.7686E-14
CD630_09480	0.2	-2.4	2.7	14.2	major capsid protein	2.578E-09
CD630_29160	0.2	-2.4	2.7	14.2	major capsid protein	0.00017578
CD630_01420	0.2	-2.4	10067.4	52788.0	RNA-binding protein	2.0168E-15
prdF	0.2	-2.4	3702.3	19399.9	proline racemase	2.0127E-14
rplP	0.2	-2.4	120.2	627.2	50S ribosomal protein L16	4.8342E-25
rplN	0.2	-2.4	66.3	343.3	50S ribosomal protein L14	0.00021788
flgD	0.2	-2.4	6.1	31.6	basal-body rod modification protein FlgD	1.6899E-05
23s rRNA	0.2	-2.4	1165.2	6015.1		8.6905E-12
nfo	0.2	-2.4	285.4	1472.0	endonuclease IV	4.8839E-16
CD630_04910	0.2	-2.4	8.6	44.4	PTS system mannose/fructose/sorbose transporter subunit IIA	1.2036E-11
CD630_04930	0.2	-2.3	21.2	106.1	PTS system mannose/fructose/sorbose transporter subunit IIC	0.0004937
eutP	0.2	-2.3	15.4	76.7	ethanolamine utilization protein	3.8539E-07
CD630_23050	0.2	-2.3	87.7	436.1	pilin protein	2.1768E-11
scrR	0.2	-2.3	655.7	3260.9	LacI family transcriptional regulator	1.248E-12
CD630_20460	0.2	-2.3	133.3	661.4	hypothetical protein	0.00021228
CD630_28740	0.2	-2.3	49.1	241.8	MATE family drug/sodium antiporter	3.4805E-10
CD630_13640	0.2	-2.3	3.8	18.8	XkdM-like protein	3.3306E-09
CD630_31010	0.2	-2.3	22.5	110.1	PTS system glucose-like transporter subunit IIBC	2.1439E-14
argE	0.2	-2.3	82.1	399.1	acetylornithine deacetylase ArgE	1.5751E-09
CD630_02850	0.2	-2.3	9.5	45.5	PTS system mannose/fructose/sorbose transporter subunit IIB	2.1139E-07
CD630_28310	0.2	-2.2	374.8	1778.0	adhesin	5.005E-05
iscS2	0.2	-2.2	1104.6	5235.8	cysteine desulfurase	3.9785E-05
tig	0.2	-2.2	2976.1	14055.3	trigger factor	1.0225E-26
nrdF	0.2	-2.2	40.4	189.3	ribonucleoside-diphosphate reductase subunit beta	7.2146E-12
CD630_05100	0.2	-2.2	215.9	1006.8	RNA polymerase sigma factor	5.9199E-17
rpsC	0.2	-2.2	252.9	1173.7	30S ribosomal protein S3	2.5133E-13
CD630_32460	0.2	-2.2	8864.2	40936.4	surface protein	1.4016E-16
hprK	0.2	-2.2	262.3	1210.3	HPr kinase/phosphorylase	1.2174E-09
CD630_23670	0.2	-2.2	466.3	2151.3	multidrug family ABC transporter permease	1.0393E-08

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_05000	0.2	-2.2	12.7	58.7	antirestriction protein	1.1416E-10
srlA	0.2	-2.2	17.4	80.1	PTS system glucitol/sorbitol-specific transporter subunit IIC	0.00076181
etfB1	0.2	-2.2	79632.0	365785.9	electron transfer flavoprotein subunit beta	1.6591E-05
rpsE	0.2	-2.2	55.1	251.9	30S ribosomal protein S5	5.4196E-12
xynA	0.2	-2.2	589.6	2688.2	PTS system xyloside-specific transporter subunit IIA	2.1508E-09
CD630_24850	0.2	-2.2	12.8	58.4	Xaa-Pro aminopeptidase	4.4785E-17
CD630_02900	0.2	-2.2	74.7	335.1	hypothetical protein	8.2851E-05
etfA1	0.2	-2.1	107535.3	477098.0	electron transfer flavoprotein subunit alpha	4.2245E-05
flgC	0.2	-2.1	4.9	21.5	flagellar basal-body rod protein FlgC	8.2241E-10
acoB	0.2	-2.1	52.4	228.9	acetoin dehydrogenase E1 component subunit beta	2.1527E-09
uvrB	0.2	-2.1	416.6	1807.3	excinuclease ABC subunit B	0.00083748
rpsK	0.2	-2.1	138.5	599.5	30S ribosomal protein S11	2.3009E-06
kdgT1	0.2	-2.1	11.8	50.9	2-keto-3-deoxygluconate permease	7.1703E-07
CD630_05020	0.2	-2.1	56.6	242.1	ATPase	3.0302E-10
eutW	0.2	-2.1	108.6	463.1	ethanolamine specific two-component sensor histidine kinase	1.6572E-09
mcsB	0.2	-2.1	2704.4	11525.6	ATP:guanido phosphotransferase	1.2778E-08
CD630_26700	0.2	-2.1	30.7	130.7	ABC transporter ATP-binding protein	4.5078E-11
groL	0.2	-2.1	3430.9	14609.3	60 kDa chaperonin	2.5588E-14
23s rRNA	0.2	-2.1	1436.4	6112.1		1.904E-14
glgC	0.2	-2.1	1861.4	7915.0	glucose-1-phosphate adenyl transferase	0.00032831
grdE	0.2	-2.1	18872.5	80060.8	glycine reductase complex component B subunits alpha and beta	1.449E-08
pfkA	0.2	-2.1	723.2	3060.7	6-phosphofructokinase	1.2244E-08
clpC	0.2	-2.1	17419.2	73581.7	class III stress response-related ATPase	7.9532E-23
CD630_29480	0.2	-2.1	85.3	359.8	hypothetical protein	2.1608E-06
CD630_11250	0.2	-2.1	17807.1	75094.6	nitroreductase-family protein	1.5428E-15
glgD	0.2	-2.1	986.7	4153.7	glycogen biosynthesis protein	8.0187E-09
CD630_15420	0.2	-2.1	884.3	3718.8	hypothetical protein	0.0002745
CD630_25171	0.2	-2.1	26.3	110.3	hypothetical protein	2.0109E-11
CD630_23630	0.2	-2.1	299.8	1247.1	hypothetical protein	1.0423E-05
CD630_23180	0.2	-2.0	327.9	1355.2	phosphohexomutase	8.0205E-11
CD630_26691	0.2	-2.0	160.1	658.4	Na(+)/H(+) antiporter	2.3271E-09

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_16230	0.2	-2.0	41771.8	171591.8	oxidoreductase	4.9938E-10
CD630_20980	0.2	-2.0	26.4	108.1	hypothetical protein	5.1656E-15
CD630_28300	0.2	-2.0	3621.7	14803.1	hypothetical protein	4.0436E-13
atpA	0.2	-2.0	3120.2	12655.6	F1 ATPase subunit alpha	4.3511E-25
rplO	0.2	-2.0	153.4	621.5	50S ribosomal protein L15	3.613E-09
CD630_34140	0.3	-2.0	11.2	44.9	sugar family ABC transporter substrate-binding protein	2.5291E-10
CD630_26040	0.3	-2.0	197.1	786.9	hypothetical protein	3.9388E-11
CD630_10770	0.3	-2.0	11.6	46.1	PTS system mannose/fructose/sorbose transporter subunit IIC	6.6639E-10
23s rRNA	0.3	-2.0	1881.5	7451.8		9.7785E-05
rbsR	0.3	-2.0	193.2	764.8	LacI family transcriptional regulator	5.4623E-09
pyk	0.3	-2.0	1303.5	5148.0	pyruvate kinase	3.6542E-16
CD630_06052	0.3	-2.0	124.7	490.8	hypothetical protein	1.2234E-15
CD630_04940	0.3	-2.0	33.1	130.1	PTS system mannose/fructose/sorbose transporter subunit IID	8.2219E-08
CD630_03040	0.3	-2.0	82.1	322.3	hypothetical protein	6.7033E-16
CD630_07170	0.3	-2.0	879.8	3450.3	bifunctional carbon monoxide dehydrogenase	3.3432E-10
dnaK	0.3	-2.0	26994.0	105580.2	molecular chaperone DnaK	2.7727E-22
eno	0.3	-2.0	3955.8	15466.7	enolase	1.0858E-15
CD630_04490	0.3	-2.0	6.6	25.7	NhaA family Na+/H+ antiporter	4.4099E-10
CD630_17970	0.3	-2.0	57261.2	222362.4	coenzyme A disulfide reductase	1.1014E-06
CD630_26050	0.3	-2.0	668.9	2593.4		4.527E-10
infB	0.3	-2.0	1475.1	5712.1	translation initiation factor IF-2	6.9982E-13
gloA	0.3	-2.0	44.9	173.6	lactoylglutathione lyase	3.9707E-05
thrS	0.3	-1.9	4838.3	18693.0	threonyl-tRNA ligase	9.069E-09
5s rRNA	0.3	-1.9	34.8	134.3		5.0286E-15
flgE	0.3	-1.9	13.0	50.0	flagellar hook protein FlgE	0.0004984
fliK	0.3	-1.9	6.6	25.4	flagellar hook-length control protein FliK	1.0302E-10
CD630_32150	0.3	-1.9	29.8	114.0	glycine betaine/carnitine/choline ABC transporter ATP-binding protein	8.3059E-10
fliD	0.3	-1.9	73.1	276.8	flagellar hook-associated protein 2 FliD	5.0887E-09
CD630_31830	0.3	-1.9	109.0	412.2	peptidase	3.6444E-05
CD630_31890	0.3	-1.9	370.6	1400.6	NADPH-dependent FMN reductase	1.1415E-08
CD630_23660	0.3	-1.9	273.4	1029.0	hypothetical protein	2.4862E-10

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
uraA	0.3	-1.9	35.8	134.0	uracil-specific ABC transporter permease	1.3869E-12
CD630_32160	0.3	-1.9	59.4	222.2	glycine betaine/carnitine/choline ABC transporter permease	9.5068E-06
bglF	0.3	-1.9	8.5	31.9	PTS system beta-glucoside-specific transporter subunit IIABC	2.3471E-12
CD630_15700	0.3	-1.9	2126.0	7909.4	hypothetical protein	6.973E-10
CD630_17960	0.3	-1.9	36504.6	135707.8	nitrite and sulfite reductase subunit	8.8534E-13
pyrD	0.3	-1.9	89.7	333.0	dihydroorotate dehydrogenase 1B	4.0083E-10
cls	0.3	-1.9	48.0	176.4	cardiolipin synthetase 1	1.3592E-09
tRNA-Thr	0.3	-1.9	14.1	51.6		1.1625E-05
tRNA-Thr	0.3	-1.9	14.1	51.6		1.197E-09
tRNA-Thr	0.3	-1.9	14.1	51.6		4.7291E-14
tRNA-Thr	0.3	-1.9	14.1	51.6		4.7897E-22
tRNA-Thr	0.3	-1.9	14.1	51.6		8.6472E-33
atpF	0.3	-1.9	1538.7	5629.0	F0 ATP synthase subunit B	1.9665E-09
rplW	0.3	-1.9	82.3	301.2	50S ribosomal protein L23	3.8995E-45
CD630_35120	0.3	-1.9	371.8	1354.0	type IV pilus transporter system	1.1349E-12
carA	0.3	-1.9	144.9	527.5	carbamoyl phosphate synthase small subunit	3.706E-14
tRNA-Val	0.3	-1.9	81.9	297.7		0.00051342
tRNA-Val	0.3	-1.9	81.9	297.7		0.00020911
tRNA-Val	0.3	-1.9	81.9	297.7		1.6724E-23
oraE	0.3	-1.9	16.1	58.5	D-ornithine aminomutase E component	7.3726E-15
CD630_34080	0.3	-1.9	522.7	1892.7	DNA mismatch repair ATPase MutS	0.0001213
CD630_17292	0.3	-1.8	53.7	193.2	hypothetical protein	2.4033E-05
CD630_32360	0.3	-1.8	1308.7	4695.1	membrane protein	3.3255E-13
CD630_23650	0.3	-1.8	2565.0	9110.3	nitrate/sulfonate/taurine ABC transporter substrate-binding protein	3.9487E-05
mcsA	0.3	-1.8	1041.8	3697.5	activator of protein kinase McsB	8.8031E-16
carB2	0.3	-1.8	779.2	2765.6	carbamoyl phosphate synthase large subunit	2.067E-33
23s rRNA	0.3	-1.8	6054.8	21434.7		5.8155E-14
flgK	0.3	-1.8	25.4	89.0	flagellar hook-associated protein FlgK	6.3171E-06
rplD	0.3	-1.8	175.8	610.7	50S ribosomal protein L4	8.6931E-08
pyrD	0.3	-1.8	24.0	83.0	dihydroorotate dehydrogenase catalytic subunit	7.1149E-21
CD630_17680	0.3	-1.8	290.7	1002.6	membrane protein	1.2356E-10
CD630_23590	0.3	-1.8	64.2	221.1	HAD superfamily hydrolase	1.5234E-10

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
glgA	0.3	-1.8	1850.7	6369.4	glycogen synthase	6.0679E-07
CD630_21150	0.3	-1.8	553.8	1892.1	copper-transporting P-type ATPase	3.4639E-11
16s rRNA	0.3	-1.8	754.2	2568.6		1.5315E-09
16s rRNA	0.3	-1.8	754.2	2568.6		5.1019E-20
pro)	0.3	-1.8	10265.1	34681.4	prolyl-tRNA ligase	2.2683E-10
CD630_24270	0.3	-1.7	8.1	27.3	flavodoxin/ferredoxin oxidoreductase subunit gamma	2.8772E-15
16s rRNA	0.3	-1.7	496.0	1654.4		0.00026473
16s rRNA	0.3	-1.7	496.0	1654.4		0.00011073
16s rRNA	0.3	-1.7	496.0	1654.4		4.968E-08
16s rRNA	0.3	-1.7	496.0	1654.4		4.5463E-16
CD630_36100	0.3	-1.7	236.1	784.9	hypothetical protein	8.6162E-05
trxB	0.3	-1.7	9138.5	30374.9	thioredoxin-disulfide reductase	1.0332E-09
grdB	0.3	-1.7	47554.6	157919.3	glycine reductase complex component B subunit gamma	0.00014829
CD630_27520	0.3	-1.7	1454.2	4828.7	hypothetical protein	2.884E-21
uvrA	0.3	-1.7	581.6	1929.8	excinuclease ABC subunit A	2.624E-10
CD630_05070	0.3	-1.7	25.4	84.1	conjugative transposon protein	8.8801E-06
rplV	0.3	-1.7	24.8	81.5	50S ribosomal protein L22	3.3603E-16
CD630_31900	0.3	-1.7	118.9	389.4	MerR family transcriptional regulator	1.1464E-24
CD630_05940	0.3	-1.7	210.3	688.0	hypothetical protein	1.309E-12
CD630_15410	0.3	-1.7	491.8	1601.8	MATE family drug/sodium antiporter	4.2073E-19
rplB	0.3	-1.7	328.6	1067.2	50S ribosomal protein L2	8.07E-14
xynB	0.3	-1.7	201.5	654.0	PTS system xyloside-specific transporter subunit IIB	1.1224E-07
23s rRNA	0.3	-1.7	582.7	1888.0		3.976E-54
nrdE	0.3	-1.7	157.6	510.6	ribonucleoside-diphosphate reductase subunit alpha	1.6402E-09
rimI	0.3	-1.7	78.0	251.5	alanine acetyltransferase	0.00011398
CD630_13480	0.3	-1.7	136.6	440.0	lipoprotein	1.0312E-13
CD630_31820	0.3	-1.7	214.0	688.6	D-aminoacylase	8.3242E-12
CD630_23620	0.3	-1.7	494.3	1589.3	hypothetical protein	4.7069E-10
CD630_31810	0.3	-1.7	63.1	202.7	chlorohydrolase/aminohydrolase	6.2383E-13
CD630_01530	0.3	-1.7	585.3	1880.4	4-hydroxyphenylacetate decarboxylase catalytic subunit HpdB	9.2132E-15
CD630_01640	0.3	-1.7	12.5	40.3	membrane protein	1.0195E-09
xynD	0.3	-1.7	308.5	989.0	PTS system xyloside-specific transporter subunit IID	7.5034E-10

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
CD630_13740	0.3	-1.7	33.5	107.3	beta-lactamase-inhibitor protein II	2.1818E-06
CD630_31290	0.3	-1.7	52.0	165.7	glycosyl hydrolase/phosphorylase	1.5322E-18
CD630_20190	0.3	-1.7	75.8	239.2	transporter drug resistance	2.1557E-14
xpt	0.3	-1.7	34.2	107.8	xanthine phosphoribosyltransferase	2.1926E-09
CD630_33100	0.3	-1.7	662.5	2085.4	D-isomer specific 2-hydroxyacid dehydrogenase	8.8533E-05
proC	0.3	-1.7	4566.9	14333.1	pyrroline-5-carboxylate reductase	2.5171E-05
CD630_31770	0.3	-1.6	22.5	70.7	xanthine dehydrogenase	9.1744E-15
CD630_05950	0.3	-1.6	56.0	175.4	ParB-like nuclease	4.4938E-10
grdA	0.3	-1.6	15377.0	48012.2	glycine/sarcosine/betaine reductase complex protein A	2.5215E-36
CD630_05030	0.3	-1.6	74.1	230.2	membrane protein	0.00025044
16s rRNA	0.3	-1.6	502.7	1552.7		7.3497E-12
16s rRNA	0.3	-1.6	502.7	1552.7		1.0663E-14
16s rRNA	0.3	-1.6	502.7	1552.7		1.2316E-15
trxB3	0.3	-1.6	13122.9	40522.1	thioredoxin reductase 3	7.6114E-10
fliC	0.3	-1.6	577.2	1779.5	flagellin C	2.4679E-11
trxA2	0.3	-1.6	3659.4	11253.4	thioredoxin 2	1.8144E-08
dnaF	0.3	-1.6	671.5	2062.0	PolC-type DNA polymerase III	3.6773E-08
gene2938 CD630_27110	0.3	-1.6	89.7	275.0	LysR family transcriptional regulator	4.0869E-10
CD630_29300	0.3	-1.6	100.7	307.2	anti-repressor protein	1.6571E-16
motA	0.3	-1.6	24.4	73.8	flagellar motor rotation protein MotA	1.6403E-10
rpsD	0.3	-1.6	220.6	665.6	30S ribosomal protein S4	4.9479E-13
CD630_05490	0.3	-1.6	50.4	151.4	hypothetical protein	1.6547E-08
gapN	0.3	-1.6	65.4	196.2	glyceraldehyde-3-phosphate dehydrogenase	6.2616E-07
feoB	0.3	-1.6	1566.2	4691.9	ferrous iron transport protein B	4.4649E-05
CD630_32920	0.3	-1.6	7.5	22.6	hypothetical protein	6.2158E-11
CD630_06060	0.3	-1.6	12991.1	38835.7	hypothetical protein	0.00056157
prdC	0.3	-1.6	707.7	2114.1	electron transfer protein	2.4962E-37
CD630_23610	0.3	-1.6	806.2	2398.4	nitrate/sulfonate/taurine ATP-binding protein	5.7562E-23
pyrK	0.3	-1.6	58.4	173.8	dihydroorotate dehydrogenase electron transfer subunit	3.8528E-14
CD630_03650	0.3	-1.6	27.5	81.9	multidrug family ABC transporter permease	6.3748E-10
rpmE	0.3	-1.6	3162.5	9370.6	50S ribosomal protein L31	0.05
CD630_08810	0.3	-1.6	3146.8	9317.6	protein modulating protease activity	1.6336E-15

Gene Name	Real Fold Change	Log2 Fold Change	Expression - Parent	Expression- Mutant	Function (Known or Predicted)	adj. *p* value
rplL	0.3	-1.6	197.1	583.6	50S ribosomal protein L7/L12	2.7338E-18
CD630_27860	0.3	-1.6	469.6	1388.5	cell surface protein	8.4842E-06
fusA	0.3	-1.6	1874.5	5529.4	elongation factor G	4.9683E-15
CD630_05101	0.3	-1.6	53.9	158.7	conjugative transposon protein	2.0361E-09
CD630_18010	0.3	-1.6	772.5	2275.0	membrane protein	3.3372E-06
bcp	0.3	-1.6	142.1	418.2	thiol peroxidase	1.7838E-33
CD630_31670	0.3	-1.6	3785.2	11113.6	phosphatidylethanolamine-binding regulatory protein	9.5511E-14
CD630_03160	0.3	-1.5	76.5	223.0	ABC transporter permease	4.1151E-10
CD630_29410	0.3	-1.5	36.7	107.1	single-strand DNA-binding protein	6.1411E-18
XdhC	0.3	-1.5	12.3	35.9	xanthine dehydrogenase iron-sulfur binding subunit XdhC3	0.00019634
rpsS	0.3	-1.5	24.2	70.1	30S ribosomal protein S19	1.116E-09
23s rRNA	0.3	-1.5	4014.9	11632.5		3.4299E-15
rbo	0.3	-1.5	15156.7	43911.0	rubredoxin oxidoreductase	1.2482E-09
prlA	0.3	-1.5	167.5	484.6	preprotein translocase subunit SecY	1.2545E-15
CD630_31780	0.3	-1.5	24.5	70.2	D-hydantoinase	1.0843E-06
grdD	0.4	-1.5	32609.9	92911.3	glycine reductase complex component C subunit alpha	6.8379E-13
flgL	0.4	-1.5	33.2	93.7	flagellar hook-associated protein FlgL	6.811E-05
CD630_18410	0.4	-1.5	411.5	1159.7	two-component sensor histidine kinase	0.00027043
CD630_01500	0.4	-1.5	72.1	203.2	peptidase	4.3372E-15
lysA	0.4	-1.5	317.5	889.9	diaminopimelate decarboxylase	2.0242E-09
CD630_01550	0.4	-1.5	333.8	935.1	4-hydroxyphenylacetate decarboxylase activating subunit HpdA	3.8693E-25
pyrD	0.4	-1.5	42.9	119.9	dihydroorotate dehydrogenase catalytic subunit	2.6836E-08
grpE	0.4	-1.5	4114.8	11423.6	protein GrpE	1.2507E-22
CD630_01160	0.4	-1.5	96216.0	267062.8	2-ketoisovalerate ferredoxin reductase	0.05
CD630_14940	0.4	-1.5	4793.7	13289.0	HTH-type transcriptional regulator	2.2614E-09
feoA	0.4	-1.5	276.7	766.5	ferrous iron transport protein	6.9544E-10
CD630_31910	0.4	-1.5	258.9	717.0	oxidoreductase	2.5164E-24
rpsM	0.4	-1.5	44.0	121.5	30S ribosomal protein S13	2.8165E-11
CD630_01180	0.4	-1.5	25141.0	69328.4	ferredoxin/flavodoxin oxidoreductase subunit gamma	6.4183E-06
CD630_05610	0.4	-1.5	599.2	1650.7	aldo/keto reductase ferredoxin	1.9645E-18
CD630_31950	0.4	-1.5	54.9	150.8	UvrABC system protein A 2	4.5076E-14