The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Investigation of the effect of insertional mutation in *agr* and *sigB* loci in *Clostridium difficile* R20291

By

Jorge Mauricio Montfort-Gardeazabal

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

April 2017

Abstract

Clostridium difficile has become the major cause of healthcare acquired diarrhoea in the world. The disease caused by this pathogen is largely mediated by the production of toxins. The *C. difficile* genome contains an incomplete accessory gene regulator (*agr*) locus, designated *agr1*. In *Staphylococcus aureus,* it has been shown that the *agr* locus regulates virulence factors through Quorum Sensing (QS). In addition to the incomplete *agr1* locus found in all of the *C. difficile* strains sequenced to date, the strains belonging to the BI/NAP1/027 group possess an additional *agr* operon, designated *agr2*. These strains have been found to produce higher toxin levels and a more severe disease in patients. The studies described in this thesis have shown that the two *agr* systems present in the *Clostridium difficile* BI/NAP1/027 strain R20291 are involved in the regulation of spore formation and toxin production. The *agr1* locus (*agrB1, agrD1)* is responsible for the production of the autoinducing peptide (AIP) AgrD1, while the second *agr2* locus (*agrC, agrA, agrB2, agrD2*) mediates production of the AgrD2. Initial mutational analysis using a University of Nottingham isolate (R20291 NM) suggested that AgrD1 both positively regulates sporulation and negatively regulates toxin production, whereas AgrD2 positively regulates toxin production but negatively regulates spore formation.

It was subsequently discovered that strain R20291 NM exhibited significantly different phenotypes to R20291 BW. The former possessed a single polar flagellum, whereas R20291 BW was peritrichously flagellated. The NM strain

exhibited impaired motility and increased biofilm formation, demonstrated different growth rates, produced greater quantities of toxins and exhibited a relative delay in the onset of sporulation compared to R20291 BW. The equivalent *agr* mutants made in R20291 BW indicated that the regulatory control exerted by AgrD1 in sporulation was broadly the same, while AgrD2 does not seem to play any role in the regulation of spores, contrasting with observations made in the *agrB2* mutant created in R20291 NM. Surprisingly, the effects of the *agr* mutants made in R20291 BW on toxin production were opposite to those observed in the NM strain.

The subtle differences in the behaviour of the two R20291 isolates was most likely due to the presence of mutations in R20291 NM revealed by whole genome sequencing. Of the four genes affected, a mutation in the anti-sigma factor RsbW, was considered the most likely culprit. An investigation of its possible role in toxin production and sporulation was therefore undertaken through complete ClosTron-mediated inactivation of *rbsW* in both R20291 isolates and the creation of a *sigB* ClosTron mutant in R20291 BW. The sporulation and toxin production phenotypes of the *rbsW* mutants of the two strains mirrored that of their respective *agr2* mutants. Thus, inactivation of *rbsW* brought forward the onset of sporulation in R20291 NM, but had no effect on the initiation of sporulation in R20291 BW, while toxin production in R20291 NM was reduced, but increased in R20291 BW. Motility and biofilm formation in the *rbsW* mutants of both strains was unaffected. These data suggest that interference with the RbsW/SigB interaction preferentially affects AgrD2-mediated regulatory processes. In contrast, however, inactivation of SigB in R20291 BW caused a delay in the initiation of sporulation, but did not affect toxin production. Mutation of *sigB*, however, did not affect motility or biofilm formation, although, in common with other bacteria, the resistance of the R20291 BW *sigB* mutant to oxidative stress was reduced.

The picture that emerges is of a complex regulatory interrelationship between the two *agr* QS systems and SigB in this important nosocomial pathogen, a relationship that has been subtly subverted in strain R20291 NM. These findings emphasise the importance of knowing the genome sequence of strains under investigation if valid conclusions are to be drawn.

Declaration

The work presented in this thesis is my own, except where acknowledged otherwise. No part of it has been presented to any university in support of an application for any degree.

Jorge Mauricio Montfort-Gardeazabal

April 2017

Dedication

To my mother, who is the inspiration of my life, you have showed me the meaning of strength, dedication and hard work. Without you none of this would have been possible.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisors, Nigel Minton and Sarah Kuehne, for their guidance, constant support and encouragement during my PhD. I am also very grateful to Steve Cartman for his advice and support during my first year. Many thanks to all the members of the Clostridia Research Group who assisted me when I needed it and for their good advice. I also want to thank the members of the Advance Microscopy Unit for their assistance in this project.

I am deeply grateful to my family for loving and supporting me during this time, as well as inspiring me to be better every day. I want to express my gratitude and love to Jeena Velzen. You have helped me to grow as a person, challenging me to be a better person every day. Your love and support have been invaluable to me and you have changed my life for the better. I would like to thank Nunzia Lopiccolo and Esteban Paredes for their constant support and friendship. You are amazing people. You made this an unforgettable experience.

I am grateful to CONACYT (Consejo Nacional de Ciencia y Tecnologia, Mexico) for believing in me and funding this project; giving me the opportunity to achieve another goal in my life.

Table of Contents

List of Figures

List of Tables

List of abbreviations

aa	Amino-acid
ACE	allele-coupled exchange
agr	Accessory gene regulator
AHL	N-acylhomoserine lactone
AIP	Autoinducing peptide
BHIs	Brain heart infusion media supplemented
bp	Base pair
°C	Degrees centigrade
CDAD	*Clostridium difficile* associated disease
CDI	*Clostridium difficile* infection
CDMM	*Clostridium difficile* minimal media
CFU	Colony forming unit
cm	centimetres
CRISPR	Clustered regularly interspaced short palindromic repeats
DEPC	Diethyl pyrocarbonate
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
Erm	Erythromycin
g	Gram
g	Gravitational force
gDNA	Genomic DNA

h	Hours
kb	Kilobase
kV	Kilovolt
l	Litre
LB	Luria Bertani
M	Molar
µg	Microgram
mg	Milligram
µl	Microliter
ml	millilitre
µm	Micrometer
µM	Micromolar
mM	Millimolar
mRNA	Messenger RNA
NAP1	North American pulsed-field type 1
ng	Nanogram
OD_{600}	Optical density at a 600nm wavelength
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pH	Potential of hydrogen
qPCR	Real-time PCR
QS	Quorum sensing
RNA	Ribonucleic acid
RPM	Revolutions per minute

RT-PCR	Reverse transcriptase PCR
s	Seconds
SNV	Single nucleotide variant
SNP	Single nucleotide polymorphism
TAE	Tris-acetate-EDTA
UV	Ultraviolet
V	Volt
v/v	Percent volume per volume
w/v	Percent weight per volume

Chapter 1

Introduction

1.1 The genus Clostridium

The genus *Clostridium* is one of the largest prokaryotic genera and has the following general characteristics: the bacteria are Gram-positive rods which are usually arranged in short chains; they are capable of endospore production; are strictly anaerobic and are not able to reduce sulphate to sulphite (Cato *et al.,* 1986). Whilst the majority of the genus share these characteristics, there are some members that do not precisely conform to this criteria, for example, *Clostridium ramosum* is Gram-negative, while *Clostridium histolyticum* can grow on agar plates exposed to air. Clostridia are ubiquitous in nature, being found from soil to lakes, and even in the guts of some animals. Although there are around 100 clostridial species, only a few are associated with human diseases (Hatheway, 1990). The most important clostridial diseases are tetanus (*Clostridium tetani*), botulism (*Clostridium botulinum*), gas gangrene (*Clostridium perfringens*) and antibiotic associated diarrhoea (*Clostridium difficile*) (Murray *et al*., 2013; Ryan & Ray, 2004).

1.2 Clostridium difficile

C. difficile was first described by Hall and O'Toole (1935), after they successfully isolated an anaerobic Gram-positive bacterium capable of toxin production from the stool of neonates. Due to the difficulties encountered in its isolation and culture they named this bacterium *Bacillus difficilis*, although it was reclassified sometime after as belonging to the genus *Clostridium*. In 1978, *C. difficile* was identified as the primary cause of antibiotic-related pseudo-membranous colitis (Bartlett *et al*., 1978), after which point, several reports associated the consumption of antibiotic and colonisation with *C. difficile* to the pseudo-membranous colitis in patients (Aronsson *et al*., 1981; Bartlett *et al*., 1980; Don & Davis, 1981; George *et al*., 1982). Since then, *C. difficile* has become the leading cause of antibiotic associated diarrhoea in hospitals around the world and has become a growing concern in healthcare, as recent studies have revealed increased incidence of community and hospital-acquired *Clostridium difficile* infection (CDI) (Baker *et al.*, 2010; Hensgens *et al.*, 2012; Honda & Dubberke, 2014).

1.2.1 *C. difficile* associated disease (CDAD)

Clostridium difficile is an opportunistic pathogen that proliferates when the gut microbiota is disrupted. The risk group for *C. difficile* associated disease (CDAD) includes the elderly, long-term hospitalized patients, patients undergoing surgery and immunocompromised patients (van Werkhoven *et al.*, 2015). Although healthy individuals are not likely to develop CDAD, the chronic consumption of broad spectrum antibiotics affects the indigenous intestinal microflora, leading to the disruption of the barrier effect, which increases the risk of infection by *C. difficile* (Owens *et al*., 2008).

The symptoms usually appear 5 to 10 days after antimicrobial therapy is applied. Although some patients are asymptomatic, the infections with *C. difficile* usually vary from mild diarrhoea to severe diarrhoea with abdominal pain, although sometimes the disease can evolve into life-threatening conditions such as pseudo-membranous colitis, toxic mega colon and bowel perforation (Sayedy *et al.,* 2010; Voth *et al.,* 2005; Kelly & LaMont, 2008). It has been shown that a single dose of intravenous antibiotics, the administration of antineoplastics or antiviral agents can cause CDAD (Ainstein *et al*., 1981; Colarian, 1988; Ambrose *et al*., 1985).

Several techniques have been used to diagnose the infection by *C. difficile* in symptomatic patients, such as the culture of the etiological agent, cellular cytotoxicity assays using cell cultures, latex agglutination tests, enzyme-linked assay systems, counter immunoelectrophoresis fluorescent-antibody assays and polymerase chain reactions (Staneck *et al.,* 1996; LaSala *et al.,* 2012; Collins *et al.,* 2015). The treatment of CDI depends on the severity of the case. Metronidazole and Vancomycin are used as first line therapy (Bagdasarian *et al.,* 2015). The 2010 Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA) guidelines (Table 1.1.) establish that the treatment for CDAD should consist, depending on the severity of the infection, of the administration of metronidazole or vancomycin or a mixture of both.

Table 1.1. Treatment of CDI Based on 2010 SHEA/IDSA Guidelines. These guidelines are based on the best available evidences and practices, as determined by the Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA) (Cohen *et al.*, 2010). The criteria proposed for defining severe or complicated CDI are based on expert opinion.

Category	Definition	Recommendation
Initial, Mild-moderate	Leukocytosis with a white blood cell count of 15,000 cells/µl or lower and a serum creatinine level less than 1.5 times the premorbid level.	Metronidazole 500 mg 3 times per day by mouth for 10–14 days.
Initial, Severe	Leukocytosis with a white blood cell count of 15,000 cells/µl or higher or a serum creatinine level greater than or equal to 1.5 times the premorbid level.	Vancomycin 125 mg 4 times per day by mouth for 10–14 days
Initial, Severe, complicated	Shock, hypotension, ileus	Vancomycin 500 mg 4 times per day by mouth or by nasogastric tube plus intravenous metronidazole 500 mg every 8 hours. If complete ileus, vancomycin can be given rectally.
First recurrence		Like initial episode
Second recurrence		Vancomycin in a tapered and/or pulsed regimen.

1.2.2 *C. difficile* virulence factors

There are several factors that are involved in the complex set of interactions needed to lead to CDAD. It has been demonstrated that these virulence factors are not equally distributed in all the *C. difficile* serotypes (Delmée & Avesani, 1990; Hunt & Ballard, 2013), leading to some strains being able to cause a more severe disease than others. Although the toxins produced by *C. difficile* are responsible for the symptoms associated with CDI, the bacteria uses other virulence factors to establish the colonisation of the gut.

1.2.2.1 Cell wall proteins and S-layer

It has been observed that *C. difficile* is able to adhere to the epithelium mucosa of animal models such as mice and hamsters (Spigaglia *et al.,* 2013). Many bacterial species use surface proteins to recognise and interact with adhesive matrix molecules on the host cells. A series of cell wall proteins (CWP) associated with adhesion have been identified in *C. difficile* (Wright *et al.,* 2005). The CWP family consists of 29 paralogous members that contain the Cell Wall Binding 2 (CWB2) motif (Fagan & Fairweather, 2014). A cluster of 17 genes, 11 of which encode CWP proteins, has been identified in *C. difficile* (Savariau-Lacomme *et al.,* 2003); however, 17 CWP loci have been found to be distributed throughout the *C. difficile* genome (Biazzo *et al.,* 2013). The *slpA* gene, which encodes for the surface layer protein A (SlpA), is located within the *cwp* cluster as well as outside of it. The SlpA protein is a polypeptide precursor that needs to undergo proteolytic cleavage by the Cwp84 cysteine protease to produce the two surface layer protein (SLP) subunits that form the paracrystalline S-layer (Kirby *et al.,* 2009).

The paracrystalline S-layer consists of two subunits, known as the high molecular weight surface layer protein (HMW-SLP) and the low molecular weight surface layer protein (LMW-SLP). These form a sTable non-covalent complex that is the basic structure of the S-layer, which has been shown to play an important role in the host-cell adherence of *C. difficile (*Merrigan *et al.,* 2013). Additionally, SLPs have been found to be able to bind the surface epithelium lining of the digestive cavities as well as the lamina propria mucosae beneath the epithelium in human gastrointestinal tissue as well as extracellular matrix proteins (ECM) such as type I collagen, thrombospondin and fibronectin (Calabi *et al.*, 2002). The two subunits that form the S-layer are translocated across the cytoplasmic membrane to the cell wall through a Sec translocase system. Sec systems usually consists of an ATPase, SecA, and a heteromeric membrane channel, SecYEG (Driessen *et al.,* 2008), although an accessory Sec locus has been reported in several bacteria, consisting of a second copy of SecA and SecY, named SecA2 and SecY2, or the presence of the SecA2 protein alone (Rigel & Braunstein, 2008). *C. difficile* contains a complete Sec system as well as an accessory SecA2 protein which have been found to play non-redundant roles in the cell; SecA recognises and translocate housekeeping substrates while SecA2 is involved in the translocation of SlpA and several CWPs (Fagan & Fairweather, 2011).

In Gram-positive bacteria, sortase enzymes covalently attach proteins to the cell wall surface or polymerise proteins in order to construct pili (Hendrickz *et al.,* 2011). All characterised sortases to date are membrane-bound cysteine transpeptidases that join proteins that contain a cell wall sorting signal (CWSS) to an amino group on the cell surface (Clancy *et al.,* 2010).There are 6 different sortase classes, A to F, that catalyse similar transpeptidation reactions but are classified depending on the substrate motifs that they are able to recognise and their function in the cell (Spirig *et al.*, 2011). Class A sortases have been found in almost all Gram-positive bacteria and are consider housekeeping genes as this class anchors a large number of distinct proteins to the cell wall (Marraffini *et al.*, 2006). In *S. aureus*, SaSrtA, a class A sortase that recognises an LPXTG motif (where X can be any amino acid), has been shown to anchor several proteins involved in the immune response evasion and adherence (Mazmanian *et al.*, 1999). Furthermore, it was observed that SaSrtA is essential for virulence in animal models (Mazmanian *et al.*, 2000). In *C. difficile*, the gene *CD2718* encodes the only conserved sortase protein present in all sequenced strains and presents a structure that is highly similar to the class B sortase SrtB found in *S. aureus* and *B. anthracis* Accordingly, *CD2718* is also annotated as *srtB* (Sebaihia *et al.*, 2006). In *S. aureus, srtB* is part of an iron-regulated locus (*isd*) and SrtB anchors the iron transporter protein IsdC by recognising an NPQTN motif (Maresso *et al.*, 2006). In *C. difficile*, SrtB cleaves an (S/P) PXTG motif between threonine and glycine residues and covalently anchors the threonine residue to a nucleophile such as glycine (Donahue *et al.*, 2014). Several substrates for SrtB have been identified, including the collagen-binding protein CD2831 and the putative adhesin CD0386 (Peltier *et al.*, 2015; Chambers *et al.*, 2015). The putative adhesin CD0386 possesses a SPKTG motif whereas the collagen-binding protein CD2831 has a PPKTG motif; both of these proteins are processed and anchored on the surface of the cell wall by SrtB. However, the role of SrtB in virulence is unclear, as insertional inactivation of *srtB* has been shown to have no impact on the virulence of the mutant strain generated in a hamster model (Chamber *et al.*, 2015). More studies are needed to have a better understanding of the role of SrtB in the virulence of *C. difficile*.

Additionally, the function of several CWPs has been characterised in *C. difficile*. Cwp84 is a cysteine protease that is responsible for the cleavage of the SlpA precursor making it essential for the correct assembly of the S-layer (de la Riva *et al.,* 2011). Cwp13 is responsible of the activation of Cwp84 as this protein is translated as a proenzyme that needs to be activated by proteolytic cleavage. It has been observed that Cwp84 is capable of degrading the extracellular matrix (ECM) proteins *in vitro* (Janoir *et al.,* 2007), however, the role of this protein in the colonisation of the gut is unclear as the *cwp84* mutant possesses the same virulence as the wild type in a hamster model (Kirby *et al.,* 2009). Cwp66 is a protein produced at the surface of the cell wall after heat-shock and it is known to be involved in adhesion to Vero cells (Waligora *et al.,* 2001). Lastly, *cwpV* encodes a protein precursor that undergoes posttranslational cleavage to produce two subunits that form a sTable complex on the surface of the bacteria that shows autoaggregative properties and has been associated with phage resistance (Reynolds *et al.,* 2011; Sekulovic *et al.*, 2015).

A number of adhesins have been described in *C. difficile*, including the collagen-binding protein CbpA, lipoprotein CD630_08730, fibronectin-binding protein Fbp68 and heat-shock protein GroEL. The collagen-binding protein CbpA has been shown to be able to bind collagen type I and IV *in vitro,* although a *cbpA* insertional mutant created in *C. difficile* adheres to collagen at the same rate as the wild type (Tulli *et al*., 2013); furthermore, the mutant shows the same levels of colonisation in a competitive dixenic mouse model as the parental strain (Janoir *et al.*, 2013). This results suggest that CpbA doesn't play a crucial role in adhesion. Studies using a fibronectin protein Fbp68 insertional mutant showed that the adherence to various cell lines *in vitro* is unaffected (Lin *et al.*, 2011), although *in vivo* experiments in a dixenic mouse model showed that there was a lower rate of adhesion to caecal mucosa, suggesting that it is a factor that contributes to the colonisation process (Barketi-Klai *et al.*, 2011). Lipoprotein CD630_08730 is thought to be involved in adherence, as insertional mutants show a lower adherence to Caco-2 cells (Kovacs-Simon *et al*., 2014), although its role in host colonisation is yet to be described. Lastly, immunisation against heat-shock protein GroEL has been shown to reduce the gut colonisation in a mouse model (Pechine *et al.*, 2013), supporting the idea that this protein plays a role in the colonisation of the host.

1.2.2.2 Flagella

Flagellum mediated motility is recognised as an essential feature for a variety of gastrointestinal bacteria that plays an important role in host invasion and colonisation (Ormonde *et al.*, 2000; O'Neil & Marquis, 2006; Erdem *et al.*, 2007). The production of flagella in *C. difficile* is regulated by three operons, designated F1, F2 and F3 (Sebaihia *et al.*, 2006). It has been observed that genes associated with motility are expressed at a higher rate during exponential growth and that the sigma factor H plays an important role in the regulation of the expression of these genes (Saujet *et al.*, 2011). The F3 operon regulates the expression of early stage genes and contains the FliA sigma factor, which controls the expression of late stage components encoded by the F1 locus, including the FliC flagellin and the FliD cap protein (Tasteyre *et al.,* 2000; Tasteyre *et al.,* 2001). The F2 locus encodes factors involved in the post-translational modification of flagella required for correct assembly and function (Faulds-Pain *et al.*, 2014). Tasteyre *et al.* (2001) performed the first phenotypic study of the FliC and FliD proteins and demonstrated that the purified proteins are able to bind to murine mucus, suggesting that these proteins could be involved in the colonisation process. However, *C. difficile* 630 *fliC* and *fliD* mutant strains showed an increased adherence to Caco2 cells, while *C. difficile* R20291 *fliC* and *fliD* mutants had a decreased adherence to Caco2 cells compared to the wild type (Baban *et al.,* 2013), suggesting that the role of flagella in colonisation may vary among strains. Moreover, allelic exchange was used to create a paralysed flagella in R20291 (strain R20291 *motB*: D23A) which retained the same number of flagella as the wild type but was non-motile. This strain showed less adherence to Caco2 cells compared to the wild type, although the adherence of R20291 *motB*:D23A was higher than that of the *fliC* and *fliD* mutants. This results suggest that the flagellar structure have adhesive properties that contribute to the adhesion of the bacteria to certain surfaces, independent of the motility of the flagella.

Although flagella seem to play a role in adherence and the colonisation of the gut, the occurrence of non-flagellated strains indicates that motility and the presence of flagella is not required for survival and colonisation. Most noTable are those strains belonging to PCR ribotype 078, which have been recognised as particularly virulent (Wilcox *et al.,* 2012). In PCR ribotype 078 ribotype strains, whilst the F1 region is present the F3 region is completely absent (STable *et al.,* 2006). The emergence of this strain as a health concern suggests that these strains are capable of producing the same disease as flagellated strains despite the lack of flagella and motility.

It has been observed that the disruption of flagellar genes in *C. difficile* 630 located in the F3 locus (*fliF, fliG, fliM, fliE*) decreases the amount of toxin produced in supernatants (Baban *et al.,* 2013; Aubrey *et al.,* 2012). In contrast, flagellar mutants of genes located in the F1 locus (*fliC, fliD*) produced increased amounts of toxins, equating to 44.4- and 7.4-fold higher, respectively, compared to the wild type. Baban *et al*. (2013) also observed that the R20291 strain *fliC* mutant caused higher rates of mortality (70%) compared to the wild type (0%) in a mice model. However, another study found that the cytotoxicity of the *fliC* mutant and the parental strain *in vitro* was similar and the use of *in vivo* transcriptomics data revealed that the toxin gene expression in the mutant and the wild type was broadly equivalent at 14 hours post-infection (Barketi-Klai *et al.,* 2014). Furthermore, the transcriptomic data revealed 310 genes that were differentially regulated in the *fliC* mutant, including genes related to flagella, sporulation, membrane transport and metabolism. The differences seen in the *fliC* mutant in 630 and R20291 suggest that the role of the flagellar genes in the internal regulation of the cell seems to vary depending on the strain and that they are not only involved in the assembly of the flagella.

1.2.2.3 *Clostridium difficile* toxins

The main virulence factors of *C. difficile* are the two large toxins, toxin A (TcdA) and toxin B (TcdB). TcdA is a 308kDa enterotoxin, while toxin B (TcdB) is a 270kDa cytotoxin. The two toxins share 49% identity and 63% similarity (Carter *et al.,* 2012). They are encoded by genes that are located in a 19.5Kb long region known as the pathogenicity locus, PaLoc (Braun *et al*., 1996), which encompasses of 5 genes: *tdcA*, *tdcB*, *tdcC*, *tcdR* (also known as *tcdD*) and *tcdE* (Fig.1.1).

The *tcdA* and *tcdB* genes are responsible for the production of toxin A and toxin B, respectively. TcdR is an alternative sigma factor that is located upstream of all the other genes and has a key role as a positive regulator of the expression of *tcdA* and *tcdB* (Mani *et al.*, 2001). It has been observed that *tcdR* is poorly expressed in early and exponential growth phase but highly expressed in stationary phase (Hundsberger *et al*., 1997). The *tcdR* gene autoregulates its own expression according to the growth phase and the environment (Mani *et al*., 2002).

The role of TcdC is controversial. It has been proposed to be an anti-sigma factor that blocks the interaction of TcdR and the RNA polymerase and prevents the association of the sigma factor with the core enzyme, acting as a negative regulator of the toxin production in the cell (Matamouros *et al.,* 2007).

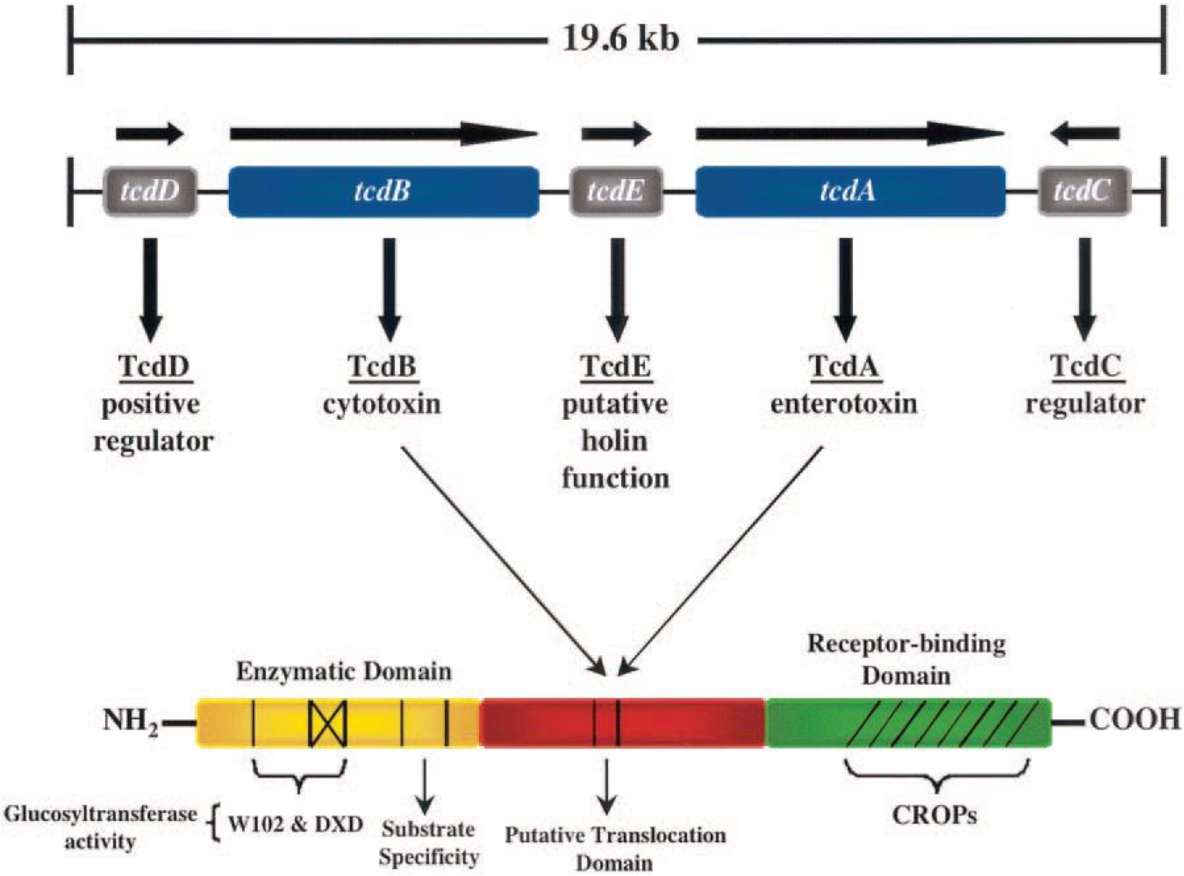


Figure 1.1. Genetic arrangement of the pathogenicity locus in *C. difficile.* The pathogenicity locus is formed by the *tcdA* and *tcdB* genes, which encode the TcdA and TcdB toxins, respectively. *tcdD* (also referred to as *tcdR)* encodes a proposed positive regulator while *tcdC* codes for a putative negative regulator of the toxin expression. *tcdE* is a predicted to encode a holin protein. The structure of the toxins is formed by the glucosyltransferase domain at the N-terminus followed by a predicted translocation domain in the middle; the binding domain is located at the C-terminus (Voth & Ballard 2005).

The emergence of hypervirulent strains capable of producing high levels of toxins that carry a modification in TcdC (Warny *et al.*, 2005) seemed to support this hypothesis. However, a number of studies present conflicting data that shows that the modification or deletion of TcdC does not affect the toxin production in strain 630 or R20291 (Baker *et al.*, 2012; Cartman *et al.*, 2012). The role of TcdC remains unclear and further study of the protein is necessary to understand its relevance in the cell. *tcdE* is located between *tcdA* and *tcdB*, and is transcribed in the same direction. The encoded TcdE protein shares similarity with holins proteins (pore forming proteins) and therefore is believed to help in the liberation of the toxin from within the cell to the external environment (Govin & Dupuy, 2012). It has been observed that TcdE mutant's of strains 630 and CDC646 show no difference in toxin secretion compared to the wildtype strains (Olling *et al.*, 2012). In contrast, a mutant created in strain R20291 was defective in the release of toxin into the medium (Govin *et al.*, 2015). This suggest that the role of TcdE varies dependent on the strain. Further studies are necessary to understand the role of TcdE in the liberation of the toxins from the cell.

TcdA and TcdB are mono glycosyltransferases belonging to the large clostridia toxins (LCT) family that inactivate Rho family GTPases, including Rho, Ras, Ral, Rac and Cdc42, by glycosylation (Pruitt & Lacy, 2012). GTPases are a large family of hydrolase enzymes that can bind and hydrolyse guanosine triphosphate (GTP). They play key roles in important cellular processes, such as the maintenance of the actin cytoskeleton, adhesion, immune cell signalling, and are essential for epithelial barrier function, among others (Etienne-Manneville & Hall, 2002). The inactivation of the GTPases results in the dysregulation of F-actin depolymerisation, which compromises the integrity of the actin cytoskeleton causing the characteristic rounding of the cell, followed by apoptosis mediated by caspase-3 and caspase-9 (Gerhard *et al.*, 2008). The death of colonocytes results in increased permeability and fluid accumulation, leading to the characteristic symptom of CDI, diarrhoea.

It has been proposed that TcdA and TcdB intoxication of the target cell occurs via the ABCD (Activity, Binding, Cutting, and Delivery) domain model (Jank & Aktories, 2008). The Activity domain consists of a glycotransferase domain located at the N-terminal region, while the Binding domain resides at the C-terminal end of the protein. The Cutting domain is constituted by a cysteine protease domain and the Delivery domain is an internal hydrophobic region. The intoxication process initiates with the receptor being recognise by the Binding domain, which causes the endocytosis of the toxin by the cell via an uncharacterised clathrin-dependent pathway, resulting in the formation of endosomes containing the toxin within the host cell (Papatheodorou *et al.*, 2010). The acidification of the endosome leads to a conformational change in the toxins that results in the insertion of the toxin into the endosomal membrane with the subsequent formation of a pore (Pruitt *et al.*, 2010). The toxin is then free to interact with inositol hexakiphosphate, which leads to an autocatalytic cleavage of the toxin and the posterior release of the N-terminal glycotransferase domain into the cytosol of the cell (Egerer *et al.*, 2009; Pruitt *et al.*, 2009). The glycotransferase domain then transfers a glucose moiety to the target Rho GTPases, locking them in an inactive conformation and interrupting their functions in the cell (Jank & Aktories 2008).

Early experimental data showed that the intragastric administration of purified toxin A in a hamster model was capable of causing all of the symptoms seen in CDI, including diarrhoea, intestinal inflammation and death. In contrast, the administration of purified toxin B did not cause any of the symptoms unless the intestinal wall had suffered prior damage or toxin B was combined with low doses of toxin A (Lyerly *et al.,* 1985). A similar study performed in a rabbit model supported the idea that toxin A was essential for CDI to develop (Mitchell *et al.,* 1986). Moreover, a strong immune response to toxin A has been reported to reduce the severity and relapse of CDI in humans (Kyne *et al.,* 2001), which led to believe that toxin A played a more important role in the development of the disease. However, the emergence of virulent strains producing toxin B but not toxin A (Drudy *et al.,* 2007) capable of producing the full spectrum of CDI symptoms suggested that the toxin B had a more prominent role than earlier work suggested. This finding was supported by other studies where TcdB was able to cause damage to the host in the absence of TcdA (Savidge *et al.,* 2003; Hamm *et al.,* 2006; Steele *et al.,* 2013). Recent studies of TcdA and TcdB mutants in R20291 and 630 background have shown that both toxins are capable of producing the disease in a hamster model (Kuehne *et al.,* 2010; Kuehne *et al.,* 2013), although it is interesting to notice that hamsters infected with strains able to produce toxin B (TcdA mutants and wild type) succumbed to the infection faster than the hamster infected with strains able to produce only toxin A. The exact role of toxin A and toxin B in CDI is still unclear, although this results suggest that the role of TcdB in the development of the disease is more important than previously reported.

Additional to the two large clostridial cytotoxin, some strains of *C. difficile* are able to produce a third toxin, the *C. difficile* binary toxin (CDT) (Stubbs *et al.,* 2000). This toxin it is part of the clostridial ADP-ribosylating family of toxins, belonging specifically to the Iota-like subfamily of binary toxins as it exhibits activity and structure similar to the *C. perfringens* iota toxin (Popoff *et al.*, 1988). As with all the other binary toxins, CDT comprises two subunits, CDTa (Catalytic subunit) and CDTb (Binding subunit); the catalytic subunit shows an actin-specific ADP ribosyltransferase activity while the binding subunit displays affinity for Lipolysis-stimulated lipoprotein receptors (LSR), a type 1 transmembrane protein (Papatheodorou *et al.,* 2011). The binding subunit interacts with an Ig-like domain mediated by the amino acids 757-866, which induces clustering of the LRS into lipid rafts (Hemmasi *et al.,* 2015; Papatheodorou *et al.,* 2013). The uptake of the toxin and the release into the cytosol are proposed to mimic the Iota-toxin model found in *C. perfringens* (Blocker *et al.*, 2001). Once inside the cytosol, the catalytic domain interacts with G-actin and ribosylates arginine 177, resulting in inhibition of actin polymerisation (Aktories *et al.,* 2012). The role of CDT in the development of the disease is still unclear. *In vitro*, low doses of the toxin induces the formation of protrusions in epithelial cells, which increases the adherence of *C. difficile* to Caco-2 cells. Hamster models infected with CDT-only strains were not affected by the disease, suggesting that CDT is not able to produce the CDI by itself (Geric *et al.,* 2006). However, CDT seems to work synergistically with TcdA, as a TcdA^{+} TcdB^{-} CDT^{+} strain appeared to be more virulent than an isogenic TcdA^{-} TcdB^{-} CDT^{+} strain in the hamster model (Kuehne *et al.*, 2014). In this study the TcdA^{-} TcdB^{-} CDT^{+} strain was able to kill 33% of the population, which indicates that the strain retains some virulence even in the absence of the two large clostridial toxins. Indeed, four cases of CDI produced by TcdA^{-} TcdB^{-} CDT^{+}

strains have been described in France (Eckert *et al.*, 2015), highlighting the importance of the further investigation of CDT and its role in CDI.

1.2.2.4 *Clostridium difficile* endospore

C. difficile possesses the ability to produce an endospore under stress conditions, which can resist radiation, heat and many common disinfectants used to clean in hospital facilities, making them difficult to eradicate from these environments (Lawley *et al.*, 2010; Maillard 2011). The spore is formed by a complete copy of the genome inside a protective spore core surrounded by a thick layer of peptidoglycan, the spore coat, and an outer layer referred to as the exosporium (Henriquez *et al.*, 2007; Lawley *et al.*, 2009). In *Bacillus* species the exosporium is formed by a paracrystalline basal layer of the protein BclA (Sylvestre *et al.*, 2002) that interacts with the ECM and regulates germination. Three BclA-like (BclA1, BclA2 and BclA3) proteins have been described in *C. difficile* (Pizarro-Guajardo *et al.*, 2014) although only BclA1 has been demonstrated to be localised to the exosporium layer. The exosporium in *C. difficile* has been linked to the ability of the spore to germinate, as it has been observed that spores with an intact exosporium germinate less efficiently than the spores without this outer layer (Brahmbhatt *et al.*, 2007; Escobar-Cortez *et al.*, 2013). The spore coat is responsible for the resistance of the spore to hostile environments and it is composed of several proteins (Lawley *et al.*, 2009), of which only a few have been characterised, including CotA, CotD, CotG, SodA and CspC. The CspC protein has been identified to be a key component of a system that detects bile acids and controls germination (Francis *et al.,* 2013). The germination of the spores depends on the presence of bile salts, like taurocholate, and the amino acid glycine, although germination receptors are yet to be described in *C. difficile* (Sorg & Sonenshein, 2008).

The sporulation process is initiated as a response to hostile or unfavourable conditions and it is controlled by the master regulator Spo0A (Heap *et al.,* 2007; Deakin *et al.,* 2012). The insertional inactivation of *spo0A* leads to a sporulation deficient phenotype. The sporulation pathway of *C. difficile* has been compared to that of *B. subtilis.* A number of studies have shown the importance of sigma factors in the regulation of sporulation, with several hundred genes demonstrated to be differentially regulated when this sigma factor was mutated (Saujet *et al.,* 2013; Fimlaid *et al.,* 2013). Some of the pathway is conserved in both species, for instance, the dependence of sigma factor K on sigma factor E, whereas others are differently regulated in *C. difficile* (Paredes-Sabja *et al.,* 2014). The production of spores by the bacterium allows it to survive in hostile conditions for an extended period of time, making the spore the major agent responsible for disease transmission.

1.2.2.5 Other virulence factors

The presence of putative Type IV pili has been detected in the majority of clostridia species that have been sequenced (Melville & Craig, 2013). These surface appendages have been shown to be involved in adhesion and cell aggregation in Gram-positive bacteria (Piepenbrink *et al.,* 2016). A cluster of 9 pilin and putative pilin-like (T4P) genes has been reported to be present in *C. difficile* 630 and R20291 (Maldarelli *et al.,* 2014). The expression of the type IV pili has been shown to be present *in vivo* through immunogold-labelling of the PilA component (Goulding *et al.,* 2009). The inactivation of the *pilA1* and *pilB1* genes produced a pili-deficient phenotype that corresponded to a reduced cell aggregation under high c-di-GMP conditions (Bordeleau *et al.,* 2015). Furthermore, a type II c-di-GMP riboswitch has been identified upstream of the *pilA1* gene and it has been shown that a high intracellular concentration of c-di-GMP increases the transcription of the T4P cluster.

The formation of biofilms *in vitro* by *C. difficile* has been reported to be regulated by central regulators such as Spo0A and by surface components like flagella or Cwp84 (Dawson *et al.,* 2012; Dapa *et al.,* 2013). The biofilm is a matrix of proteins, such as Cwp84 and toxins, DNA and polysaccharides in which the bacteria are embedded (Semenyuk *et al.,* 2014). Biofilms have been shown to increase the resistance of *C. difficile* to oxygen stress and antibiotics as well as the persistence of the spores *in vitro* on abiotic surfaces (Dawson *et al.,* 2012). There is, however, little evidence that *C. difficile* produces biofilms *in vivo.* It has been observed that *C. difficile* resides in multicellular communities in association with the loose outer layer of the mucus in the gut (Semenyuk *et al.,* 2015). Nevertheless, additional studies are needed to define the role of these communities and if they are able to form biofilms *in vivo*.

A protease designated Zmp1 has been identified in *C. difficile* (Cafardi *et al.*, 2013). This Zinc-metalloprotease (Zmp1) has been shown to cleave several host proteins such as fibronectin, fibrinogen and IgA2 antibodies, among others. A collagen binding protein (CD630_28310) located next to the *zmp1* gene has been shown to be a Zmp1 target (Cafardi *et al.*, 2013). Two riboswitches were identified upstream of *zmp1* and *CD630_28310*; a type 1 riboswitch is located upstream *zmp*1, whereas a type 2 riboswitch was found upstream of *CD630_28310* (Soutourina *et al.*, 2013). These type of riboswitches are activated by variation in the concentration of c-di-GMP in the cell, a messenger used in signal transduction in bacteria (Bordeleau *et al.*, 2011). It is possible that at high c-di-GMP concentrations the production of Zmp1 is upregulated resulting in the cleavage of CD630_28310, thereby reducing the adhesion of the bacteria to the host. A reduction in the concentration of c-di-GMP would lead to the upregulation of CD630_28310 and downregulation of Zmp1, leading to the promotion of adhesion to the host (Hensbergen *et al.*, 2014).

1.3 Quorum sensing in bacteria

In several bacteria, toxin production, sporulation and other processes are regulated by a cell–to-cell communication system known as quorum sensing (QS) (William and Camara, 2013; Novick and Geisinger, 2008; Davies *et al.*, 1998). Communication between cells is achieved by the release of a low mass molecule referred to as autoinducer (Fuqua & Greenberg, 1994). The autoinducer is released into the environment as the population grows, enabling individual cells to sense and respond to alterations in population density through the differential expression of certain sets of genes. When the population is low, the concentration of the molecule in the environment is not enough to be detected and thus, genes regulated by the quorum sensing systems are not expressed; however, once the population has increased and the molecule reaches a critical threshold concentration in the environment, its detection by receptors in the bacterial membrane triggers a series of events inside the cell, leading to a change in gene expression (Winzer *et al*., 2003). The genes involved in QS are often related to toxin production, spore production, biofilm formation and virulence factor production (Lerat & Moran, 2004; Kleerebezem & Quadri, 2001; Novick & Geisinger, 2008).

Gram-positive and Gram-negative bacteria use different autoinducers and QS systems (Rutherford & Bassler, 2012). Gram-positive bacteria use autoinducing peptides (AIP) as signalling molecules (Miller & Bassler, 2001). The AIPs are processed and secreted from the cell via an ATP binding cassette (ABC) transporter to the outside of the cell. There are two models of QS detection described for Gram-positive bacteria (Rutherford & Bassler, 2012). In one of them, a transmembrane two-component histidine kinase receptor is activated by the AIP, causing it to phosphorylate itself and pass the phosphate group to a cognate response regulator, which activates transcription in the QS regulon. In the second model of QS, the AIP is transported back into the cell, where it interacts with transcription factors in order to modulate gene expression.

The Gram-negative bacteria signalling molecule is an acyl homoserine lactone (AHL) or other molecules derived from S-adenosylmethionine (Wei *et al.*, 2011). Once the AHL is synthesised, it can freely diffuse across the membranes of the cell (Kaplan & Greenberg, 1985). As the population grows, the concentration of the AHL in the environment increases and, when a critical concentration of the autoinducer is reached, the AHL diffuses back to the cell and interacts with transcription factors within the cell which leads to changes in gene expression (Callahan & Dunlap, 2000). Some Gram-negative bacteria use a two-component system histidine kinase receptor that is an analogue of the two-component system described in some Gram-positive bacteria. An overview of the QS models in bacteria are shown in Fig. 1.2

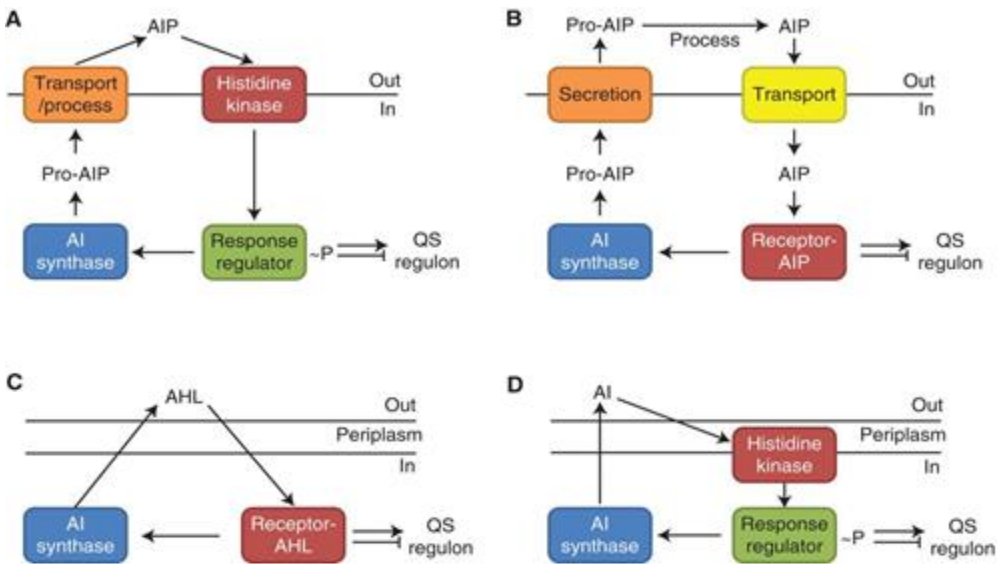


Figure 1.2. General models for quorum sensing in bacteria. Quorum sensing in Gram-positive bacteria can be achieve by A) the detection of the AIP and phosphorylation of the response regulator by a membrane-bound histidine kinase or B) the transport of the AIP into the cell and its interaction with transcriptional factors. Gram-negative QS is mediated by a C) LuxI/LuxR system or D) two component system (Rutherford and Bassler, 2012)

1.3.1 Quorum sensing in Gram-negative

Quorum sensing was first described in two marine bacterial species, *Vibrio fischeri* and *Vibrio harveyi,* where it controls light emission in both species (Bassler *et al.,* 1997). Light emission occurs in response to cell density and the accumulation of the autoinducer in the environment (Nealson *et al*., 1970; Nealson & Hastings, 1979). In *V. fischeri*, the autoinducer identified was an acylated homoserine lactone, N-(3-oxohexanoyl)-homoserine lactone (Eberhard *et al*., 1981). Two regulatory proteins are involved in the control of the luminescence, LuxI and LuxR (Fuqua *et al.,* 1994). LuxI, an acyl-homoserine-lactone synthase, is involved in the production of the signalling molecule, while LuxR detects the autoinducer and promotes the transcription of the LuxICEDABE operon responsible for the transcription of the luciferase enzymes required for light emission (Engebrecht *et al*., 1983; Schaefer *et al*., 1993). LuxI/LuxR homologues have been found in over 100 Gram-negative bacterial species (Case *et al.,* 2008). *V. harveyi* uses two QS systems working in parallel to regulate light emission; two autoinducer molecules, AI-1 and AI-2, and their cognate sensors, LuxN and LuxPQ, respectively (Bassler *et al.*, 1993; Bassler *et al.,* 1994). Synthesis of AI-1 is dependent on two genes, *luxS* and *luxM,* while the gene require for the enzymatic synthesis of the autoinducer AI-2 has been designated *luxS* (Schauder *et al*., 2001).

LuxS/AI-2 homologous genes have been reported to be present in both, Gram-positive and Gram-negative bacteria (Surette *et al*., 1999; Bassler, 1999). Some reports show that LuxS is involved in the expression of virulence factors in *E. coli* O157, *S. pyogenes* and *Vibrio vulfinicus* (Sperandio *et al*., 1999; Kim *et al*., 2000; Lyon *et al.,* 2001).

Pseudomonas aeruginosa is a ubiquitous Gram-negative bacteria that causes acute and chronic infections in humans (Mena & Gerba, 2009). The infection is typically found in patients with a compromised immune system or in the lungs of patients with cystic fibrosis (Zemanick *et al.*, 2011). Three QS systems have been identified in this pathogen; two of them are LuxI/LuxR homologues and one is a non-LuxI/LuxR system called *Pseudomonas* quinolone signalling system (Jimenez *et al.,* 2012). The QS systems in *P. aeruginosa* control the production of virulence factors (i.e. exotoxin A, elastase, proteases, swarming motility, rhamnolipids), highlighting the importance of QS in the expression of virulence in bacteria (Dusane *et al.,* 2010; Schuster *et al.*, 2006).

1.3.2 Quorum sensing in Gram-positive bacteria

Two component systems are one of the most prevalent strategies by which bacteria sense and respond to changes in their environment. These pathways respond to a wide variety of stimuli, including changes in osmolarity, quorum signals and antibiotics, among others (West *et al.,* 2001). Two component systems are pathways that comprise a sensor histidine kinase, capable of detecting a specific signal, and its cognate response regulator (Bijlsma *et al.,* 2003). Archetypically, the stimulation of the sensor histidine kinase by a signalling molecule catalyses an autophosphorulation reaction on a conserved histidine residue with subsequent transfer of the phosphoryl group to a conserved aspartate residue located in the response regulator (Stock *et al.,* 2000). The phosphorylation of the response regulator activates the output domain. The most common output domains are DNA-binding domains, which leads to transcriptional changes in the cell (Gao *et al.,* 2007). It has been observed that, in most cases, the sensor histidine kinases act as phosphatases for their cognate response regulators when the signalling molecule is not present, modulating the output response of the system (Laub & Goulian 2007).

The histidine kinases contain two conserved domains, a dimerization and histidine phosphotransfer domain (DHp), where the conserve histidine residue that is subjected to the autophosphorylation and phosphotransfer reaction is located, and a catalytic and ATP binding (CA) domain (Casino *et al.,* 2010). Several additional domains can be found at the N-terminal of the DHp domain, including 1 to 13 transmembrane domains with signal recognition. Signal recognition usually occurs in the extracellular portion of the protein, although it can also occur in the periplasmic portion of the cell.

The response regulators share two conserved domains, a receiver domain and an output domain (Laub & Goulian, 2007). The receiver domain harbours the aspartate residue that is subjected to phosphorylation as well as highly conserved amino acids that catalyse phosphotransfer from its cognate histidine kinase (Capra & Laub, 2012). The phosphorylation of the response regulator leads to a conformational change that allows the protein to produce its output response. Although there are several output domains, i.e. diguanylate cyclases and methyltransferase, the most common output response is binding of DNA, which leads to transcriptional changes in the cell (Gao *et al.*, 2007).

The first Gram-positive QS model that was identified is based on two-component adaptive response proteins. The signalling peptide is usually synthesised as an inactive precursor which is later cleaved to produce the active autoinducer peptide (AIP) (Okada *et al.*, 2005). The AIP is then secreted out of the cell via an ABC transporter (Kleerebezem *et al.*, 1994). Once the critical threshold concentration of AIP is reached, a histidine sensor kinase protein of a two component signalling system detects the autoinducer and promotes changes in the transcription of the cell (Inouye & Dutta, 2002; West & Stock, 2001). This QS model has been most thoroughly studied in *Staphylococcus aureus*, where it controls virulence factors as described in the 1.3.3 section.

Recent work has revealed the existence of an alternative QS model in some Gram-positive bacteria. This alternative QS system does not rely on a sensor histidine kinase, instead, the pro-AIP is secreted out of the cell and processed by extracellular proteases. After the activation of the AIP, the signalling molecule is transported back to the cell, where the active AIP directly interacts with transcriptional factors, leading to transcriptional changes in the cell (Slamti & Lereclus, 2002). This type of QS system has been identified in *Bacillus cereus*. In *B. cereus,* this QS system is reliant on the interaction between PlcR, a pleiotropic regulator of extracellular virulence, and the activated form of PapR. Following the secretion of the pro-AIP to the environment, the pro-AIP is cleaved and modified by a neutral protease B, NprB, to produce the active AIP (Pomerantsev *et al.,* 2009). NrpB cleaves the pro-AIP into peptides of 5, 7, 8 and 11 amino acids. The peptides consisting of 5 and 7 amino acids are the only peptides that activate PlcR, with the heptapeptide causing maximum activation; furthermore, it has been observed that the heptapeptide is more prevalent *in vivo* (Boillaut *et al.,* 2008). Once the AIP is processed and activated, it is imported back to the cell via an oligopeptide permease system (OPP) (Gominet *et al.,* 2009). The imported AIP interacts with PlcR and promotes conformational changes in the DNA-binding domain of PlcR, facilitating the interaction with DNA and regulating transcription of the target genes (Derclerck *et al.,* 2007). PlcR controls several virulence factors of *B. cereus* such as haemolysins, enterotoxins, proteases and phospholipases, as well as the expression of *papR* (Gohar *et al.,* 2008).

1.3.3 Quorum sensing in *Staphylococcus aureus*

In *S. aureus,* the accessory gene regulator (*agr*) quorum sensing is a global regulator known to be involved in the regulation of several genes, including virulence factors (i.e. haemolysins, serine protease, DNase, fibrinolysin and enterotoxins A-D) (Otto *et al*., 1998; George & Muir, 2007). S. *aureus* is a member of the human microbiota as a commensal; however, the activation of the *agr* system leads to its transformation from an inoffensive commensal to an aggressive pathogen (Roux *et al*., 2009). The *agr* system also plays an important role in biofilms. Biofilm formation is a multi-step process which involves two independent mechanisms (Parsek & Tolker-Nielsen, 2008). The first step is the attachment to a surface and it is mediated by polysaccharide intercellular adhesion (PIA). After this, the colonies grow into macro-colonies, which later fuse together to form the biofilm. The last step involves the detachment of the biofilm, thought to be important for dissemination during the infection process. The *agr* system is thought to be involved in the detachment of the biofilm (Lauderdale *et al*., 2009). Thus, detachment of cells coincides with the expression of the *agr* system (Kong *et al*., 2006; Yarwood *et al*., 2004). Furthermore, a study by Novick (Novick & Geisinger, 2008) shows that *agr* mutants stay in the adhesion stage with enhanced adherence and are not capable of detachment.

The *agr* locus is approximately 3.5Kb long and comprises 2 operons, RNAII and RNAIII, and two promoters, P2 and P3, respectively (Novick *et al*., 1995). The RNAII operon consists of 4 genes (*agrA*, *agrB*, *agrC* and *agrD*). The RNAIII operon encodes for RNA III, a 514 nucleotide regulatory RNA that increases the transcription of several secreted virulence factors (Roux *et al*., 2009).

The *agrD* gene encodes the autosignalling molecule which is a 46 amino acids propeptide that needs to be processed before being released to the environment in order to become the mature autoinducing peptide (AIP) (Lyon & Novick, 2004).

Once the propeptide is produced inside the cell, it anchors to the cell membrane where AgrB, a transmembrane protein essential for the processing and secretion of the AIP, cleaves the N- and the C- ends of the peptide. A cyclisation reaction between a central conserved cysteine and the C-terminal carboxyl produces an 8 amino acid thiolactone ring, which is the mature AIP (Zhang *et al*., 2004).

AgrC is a membrane-bound sensor kinase capable to bind the AIP (Lina *et al*., 1998). Once the AIP reaches a critical concentration, the sensor histidine kinase detects the signalling molecule and is activated via autophosporylation (Lyon & Novick *et al*., 2004). AgrA is a member of the LytR family of response regulators which undergo conformational changes upon phosphorylation, allowing them to interact with their cognate response regulator (Koenig *et al*., 2004). AgrA is involved in the activation of the P2 and P3 promoters by interacting with the intergenic region where the promoter for this operons is located. This region is 120 bp long and it is located between the P2 and P3 promoters. It contains a strong 17bp inverted repeat, which is believed to be a bidirectional regulatory protein binding site (Bayer *et al*., 1996).

RNAIII is produce by the P3 transcript (Novick *et al*., 1993). It codes for the δ haemolysin toxin but also plays a role in the regulation of virulence factors in *S. aureus*. The genetic changes within the cell lead to a decrease in transcription of genes encoding cell surface proteins (e.g. adhesins) and an increase in the transcription of genes encoding virulence factors such as toxins and proteases (Novick & Geisinger, 2008). RNAIII mediates the translational repression of *rot* (repression of toxins) mRNA through a limited number of base pairings involving two loop–loop interactions (Boisset *et al.*, 2007). RoT is a global regulator that negatively regulates the transcription of several virulence factors in the cell (Said-Salim *et al.*, 2003). The general model for the *S. aureus* quorum sensing is represented in Fig. 1.3.

There are at least 4 different *agr* groups, categorized as I-IV, based on variation in the B-D-C region (Dufour *et al*., 2002). The AIP from any of these groups can bind the receptor of any of the other groups, but it will only activate the receptor from those in the same group, while inhibit those belonging to the other groups (Mayville *et al*., 1999). There is only cross-talk between the groups I and IV, meaning that they are cross-active. A single change in the amino acid sequence can lead to a change in the group specificity (Dufour *et al*., 2002).

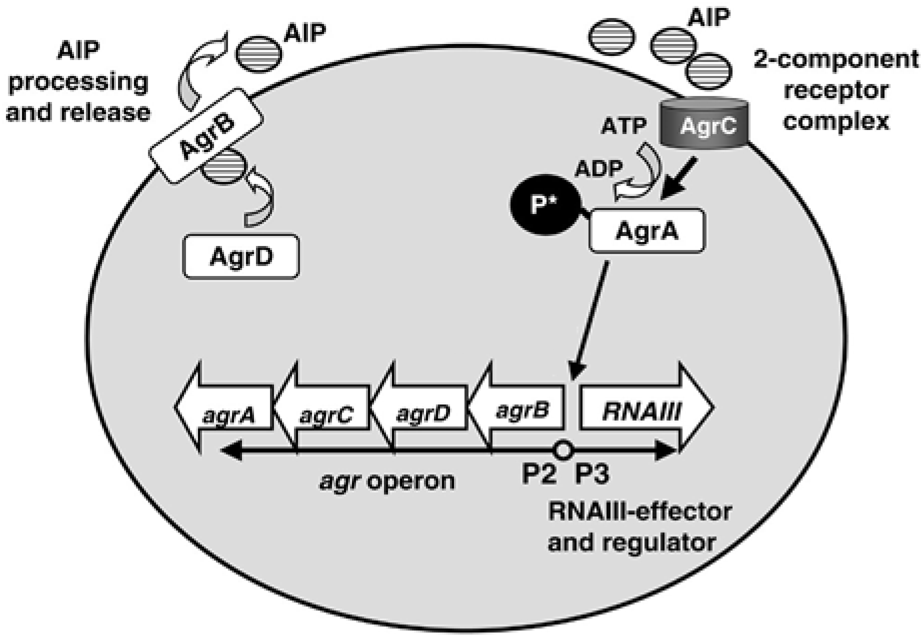


Figure 1.3. General model for the *Staphylococcus aureus* quorum sensing. The P2 and P3 operons regulate virulence factors in *S. aureus* through quorum sensing. AgrD is the pro-AIP that is processes by AgrB and excreted to the outside of the cell; when a critical concentration is reached, AgrC interacts with the AIP and autophosphorylates itself, passing the phosphate group to AgrA, which then interacts with the intergenic region that regulates the expression of the P2 and P3 operons (Asad & Opal, 2008).

1.3.4 Quorum sensing in Clostridia

Putative *agr* regions homologous to that in *S. aureus* have been found in several clostridia species, including *C. botulinum*, *C. perfringens*, *C. acetobutylicum* and *C. difficile*. In *C. perfringens* the *agr* operon plays an important role in the regulation of toxins (Chen *et al*., 2012). In this study, the *C. perfringens* type D *agrB* mutants were unable to produce epsilon-toxin (ETX), while the type A had a drastic decrease in the production of alpha-toxin (CPA) and perfringolysin O (PFO). When the mutants were complemented, the toxin production was restored to nearly wild-type levels. It was proposed that the toxin production could be regulated by the *agr* operon and a VirS/VirR two-component regulatory system. The VirS/VirR mutants were unable to produce PFO and could only produce low levels of CPA. However, the levels of ETX remained the same, suggesting that the regulation of ETX is VirS/VirR independent.

C. botulinum contains two *agr* systems, *agr1* and *agr2*, that encode proteins homologous to the AgrB and AgrD proteins found in *S. aureus*. Insertional mutants of *agrD1* and *agrD2* have lower levels of sporulation and toxin production compared to the wild type; however, the reduction in spore formation is more prominent in the *agrD1* mutants, while the decrease in toxin production is greater in the *agrD2* mutants (Cooksley *et al*., 2010). These findings suggest that the *agrBD1* operon is involved in the regulation of spore formation, while *agrBD2* have a more important role in toxin production.

It has been shown that the *agr* system also plays an important role in *C. acetobutylicum* (Steiner *et al*., 2012). *agrA*, *agrB* and *agrC* insertional mutants were shown to reduced levels of sporulation and granulose formation compared to the wild type. These defects could be reversed by complementing the mutants with a vector carrying the native *agr* locus. These findings revealed that the *agr* locus plays an important role in the regulation of virulence factors in clostridia.

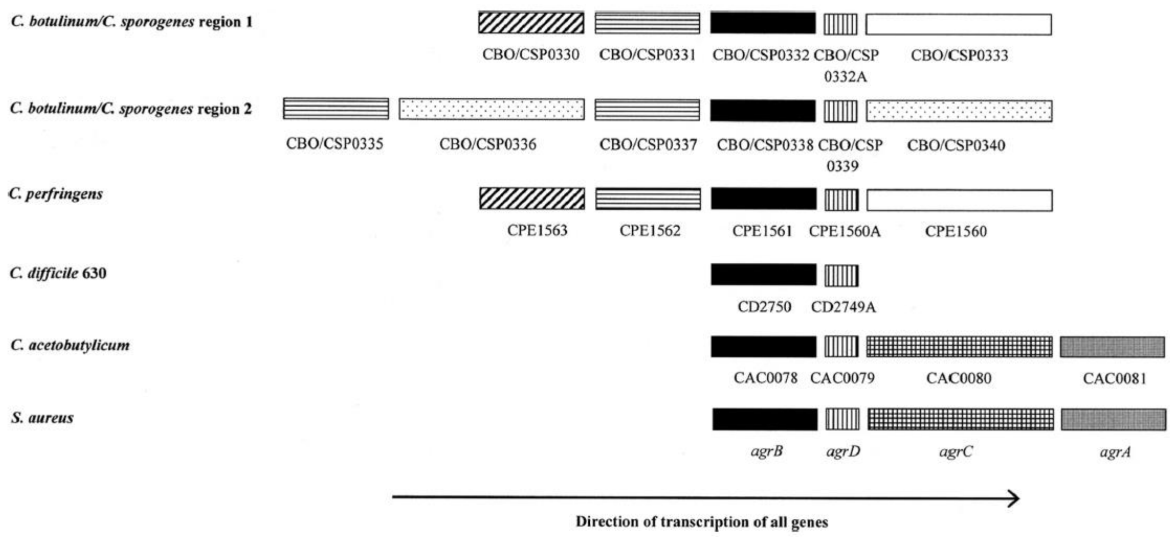


Figure 1.2. Putative *agr* homologous regions in clostridia. Homologue regions to the *agr* system found in *S. aureus*. Some elements of the *agr* operon are missing in some species (Cooksley *et al*., 2010).

1.3.5 Quorum sensing in *C. difficile*

It has been proven that the production of toxin A and B in *C. difficile* is a cell population density dependant process (Dupuy & Sonenshein, 1998). A *luxS* homologue has been identified in the genome of the *C. difficile* 630; moreover, this strain is also able to produce the AI-2 molecule (Carter *et al*., 2005). The addition of exogenous AI-2 showed no effect on toxin regulation and antisense directed against the AI-2 mRNA didn't show downregulation of this process. The production of an isogenic mutant is yet to be achieved and thus the role of the AI-2 in the cell remains unclear.

C. difficile strain 630 also possesses a putative *agr* systems comprising genes encoding *agrB* and *agrD* but lacking any genes in their vicinity analogous to the *agrC* and *agrA* genes of *S. aureus*. This *agrBD* locus has been designated *agr*1 as other *C. difficile* strains carry a second *agr* locus, termed *agr*2. This second *agr* locus contains homologues to the four *agr* genes found in *S. aureus* (*agrBDAC*) (STabler *et al*., 2009). PCR ribotype 027 strains, such as R20291, are generally considered to be more virulent (often referred to as hypervirulent) as they are responsible for higher rates of mortality compared to other strains (Goorhuis *et al*., 2007; MacCannell *et al*., 2006; Hernandez-Rocha *et al.*, 2012). A comparative genome analysis between 630 and two 027 strains, the epidemic R20291 strain and the historical CD196 strain, revealed that 630 possesses 505 genes not found in the 027 PCR ribotype strains, while the 027 strains contain 234 unique genes (STabler *et al.*, 2009). Furthermore, the R20291 strain contains 47 unique extra genes not found in CD196, including antibiotic resistance and two-component systems among others, suggesting that the presence of these genes increase the virulence of this strain compared to the historical non-epidemic CD196 strain.

1.4 Other factors involved in the regulation of virulence in bacteria

The modulation of virulence factors in bacteria is often controlled by a complex set of interactions between different regulation systems. As mentioned in section 1.3, the *agr* locus is involved in the upregulation of virulence factors through QS in several bacteria*;* however, the regulation of virulence factors is also affected by other pathways in the cell. For instance, CodY, a protein that is involved in the transition between the exponential growth phase and early stationary phase found in several members of the low G+C (DNA containing less than 50% G-C base pairs) family of Gram-positive bacteria (Sonenshein 2005), has been shown to negatively regulate some virulence factors during early exponential growth in *S. aureus,* including the haemolytic alpha-toxin, the *agr* locus and the polysaccharide intracellular adhesin (Pohl *et al.*, 2009; Majerczyk *et al.*, 2008). CodY is a direct repressor of the transcription of the haemolytic alpha-toxin gene (*hysA*) in *S. aureus* through the binding of a consensus motif known as CodY binding box (AATTTTCWGAAAATT) located upstream of *hysA,* preventing the transcription of this gene (Ibberson *et al.,* 2014). A CodY binding box containing 3 mismatches is located upstream of the RNA III gene and several CodY binding boxes with no fewer than five mismatches are found in the intergenic region containing the P2 and P3 promoters of the *agr* locus (Pohl *et al.,* 2009); however, a study conducted by Roux *et al.* (2014) showed that CodY can only weakly interact with the intergenic region *in vitro*, suggesting that CodY regulates the *agr* locus indirectly, although the mechanism of this regulation has not yet been determined. In *C. difficile*, CodY has been shown to be a potent repressor of toxin genes during the exponential growth by binding to the *tcdR* promoter region, effectively downregulating the transcription of the PaLoc in the cell (Dineen *et al.,* 2007). Additionally, CodY has also been reported to repress sporulation in *C. difficile* by interacting with two loci associated with the initiation of sporulation in the cell (*opp* and *sinR*) although the regulatory mechanism by which this is achieved is not fully elucidated, as many factors involved in the initiation of sporulation are still unknown or not fully understood (Nawrocki *et al.*, 2016).

Another operon that has been observed to affect the regulation of virulence factors in several Gram-positive bacteria is the locus containing the sigma factor B (SigB). Sigma factors are required for the initiation of transcription as they confer specificity to the RNA polymerase holoenzyme by directing the catalytic core to the appropriate transcription start site as well as initiating the strand separation of the double-helical DNA (Feklistov *et al.*, 2014). In Gram-positive bacteria, sigma factor A (SigA) is involved in essential housekeeping functions such as cellular growth and reproduction (Paget & Helmann 2003) while alternative sigma factors mediate specialized functions such as biofilm formation and cell differentiation, among others (Feklistov *et al.*, 2014). In *B. subtilis*, SigB is the master regulator of the general stress response that is triggered when the cell is subject to adverse conditions, including high and low temperatures; ethanol, salt and oxidative stress; cell wall stress caused by antibiotics; starvation for glucose, oxygen and phosphate as well as inhibitors that reduce the ATP pool in the cell (Hecker *et al.*, 2007). The activity of SigB is regulated by a partner-switching mechanism between the RsbW protein, an anti-sigma factor, and RsbV, the antagonist protein of RsbW, through reversible serine and threonine phosphorylation (Sonenshein *et al.*, 2002). RsbW phosphorylates RsbV in growing cells and forms a sTable complex with SigB that prevents the interaction of SigB with the RNA polymerase core (Benson & William 1993). SigB can only be released from the RsbW/SigB complex when RsbV is dephosphorylated by either of two RsvV~P phosphatases, RsbU and RsbP (Vijay *et al.*, 2000); thus, the phosphorylation status of RsbV determines the activation state of SigB and the general stress response. In *B. subtilis, sigB* mutants do not seem to be impaired in sporulation at

37°C; however, at lower temperatures the efficiency of sporulation is reduced, suggesting that SigB plays a supporting role in sporulation (Mendez *et al.,* 2004).

In *L. monocytogenes,* SigB is not only involved in the general stress response in the cell but is also intimately involved in the production of PrfA, the main global regulator for the expression of virulence factors (Nadon *et al.,* 2002). P2, the strongest promoter for *prfA,* is dependent on SigB as well as some PrfA-dependent genes (Schwab *et al.,* 2005). In *S. aureus*, many studies (Ziebandt *et al.,* 2004; Nicholas *et al.,* 1999; Bischoff *et al.,* 2004) have shown that SigB influences, both directly and indirectly, the expression of several virulence factors. SigB indirectly represses the transcription of the virulence factor toxic shock syndrome toxin 1 (TSSST-1) that is encode by *tst* via a mechanism involving the *agr* locus (Andrey *et al.,* 2015). Several cell wall-associated adhesion molecules as well as the global regulator SarA are positively regulated by SigB (P2 SigB dependent) (Cerca *et al.,* 2008). It has also been observed that SigB acts in the opposite way to RNAIII, inhibiting the transcription of extracellular virulence factors (lipases, toxins, proteases) (Ziebandt *et al.,* 2004). A *sigB* locus has also been identified in *C. difficile* (Sebaihia *et al.,* 2006) although more studies need to be conducted to determine the influence of SigB in the regulation of virulence factors.

As evidenced by the influence of CodY and SigB in the regulation of virulence factors in several Gram-positive bacteria, the expression of virulence factors is affected directly and indirectly by multiple levels of regulation that allow the cell to adapt to different environmental conditions as needed, highlighting the importance to study the mechanisms by which the different virulence factors can be regulated in the cell through these pathways.

Another factor that can dramatically affect the expression of virulence factors in a cell is the appearance of single nucleotide polymorphism (SNP) in regulatory genes. It has been observed that depending on the site where the SNP mutation occurs, the virulence of the bacteria can be increased or decreased. For instance, in *P. aeruginosa,* a SNP mutation in the *mexT* gene, a regulator of the MexE-MexF-OprN multidrug efflux system (Köhler *et al.*, 1999), leads to a premature stop codon resulting in a non-functional truncated protein (Olivas *et al.*, 2012). This mutation leads to a more virulent phenotype compared to the wild type, as the mutated strain shows enhanced collagenase activity, high swarming motility, altered pyocyanin production and a destructive phenotype against intestinal epithelial cells. In *V. cholera*, a hypervirulent clinical isolate has been shown to contain SNPs mutation in four regulatory genes (*hapR, hns, luxO and vieA*) (Carignan *et al.*, 2016). It was observed that the introduction of these SNPs in the reference strain increases the production of toxin-coregulated pilus (TCP) and ToxT, the master virulence regulator (Krukonis *et al.*, 2000) effectively increasing the virulence of the cell. The presence of SNPs in two independently isolated erythromycin-derivatives of the *C. difficile* 630 strain (630E and 630ΔE) has been shown to greatly affect the expression of virulence factors such as toxin production, sporulation and motility (Collery *et al.*, 2016). These examples highlight the importance to evaluate the effect of SNPs in the cell as they could be affecting different pathways that are involved in the regulation of virulence factors, greatly affecting the overall virulence of the cell.

1.5 Project aims

Open reading frames homologues to the *agr* locus in *S. aureus* have been found in several clostridia species. In *C. perfringens* and *C. sporogenes* this putative *agr* locus play an important role in the regulation of sporulation and other virulence factors. An incomplete *agr* locus consisting of only the BD genes is present in 630, lacking the putative sensor kinase and the response regulator. In the *C. difficile* PCR ribotype 027 strain R20291, an additional complete *agr* locus consisting of the *BDCA* genes is present. Several 027 strains have a higher toxin production and cause a more severe disease than other strains lacking this second *agr* locus. However, it is important to note that the 027 strain CD196 is a non-epidemic strain in spite of harbouring the *agr*2 operon. A comparative genome analysis showed that R20291 contains 47 genes not found in CD196, including drug resistance genes, transposon genes, phage islands and two-component response regulators. These additional genes not present in CD196, along with the *agr*2 operon, seem to play an important role in the increased virulence of the R20291 strain. In many bacteria the *agr* locus plays an important role in the regulation of virulence factors through an *agr*-QS regulatory system. Thus we theorise that the presence of this *agr*2 locus in R20291 plays an important role in the regulation of its pathogenesis, as seen in other clostridia species. The aim of this thesis was to investigate the role of the *agr2* operon in the 027 strain R20291 and its effect on the regulation of toxins and spores. The characterisation of the QS regulatory system in *C. difficile* could revealed new therapeutic targets and alternatives to the use of antibiotic by targeting key elements in the regulation cascade in order to reduce the expression of virulence factors.

During the course of this work, it was discovered that the isolate of the PCR ribotype strain R20291 being used exhibited dramatic differences in phenotypic properties compared to isolates from other laboratories. Whole genome sequencing of the Nottingham R20291 isolate revealed a number of Single Nucleotide Polymorphsms (SNPs) and Insertions and Deletions (InDels) that were not present in an equivalent isolate sourced from the Brendan Wren laboratory. This necessitated a revaluation of data and strategies and prompted an investigation of the likely causes of the observed phenotypic differences, in particular, a SNP in a gene encoding a putative anti-sigma factor, RbsW and the role of the *sigB* operon in virulence.

Chapter 2

Materials and methods

2.1 Bacteriological methods

2.1.1 Bacterial strains. Bacterial strains are listed in Table 2.1.

Table 2.1. Bacterial strains used in this study.

OLIGONUCLEOTIDE	SEQUENCE 5’ – 3’
agrA_HT_NdeI_F1	gctcccagtatacgcgcccatatgcaccatcaccatcaccataatagaaattttatatatttattaaaaatatataagaaagg
agrA_NdeI_F1	gctcccagtatacgcgcccatatgaatagaaattttatatatttattaaaaatatataagaaagg
agrA_XbaI_R1	cccggtagttttcgcgcctctagattagcatactatatcccctaatatattagttatagc
agrA_HT_XbaI_R1	cccgctagtatacgcgcctctagattaatggtgatggtgatggtggcatactatatcccctaatatattagttatagc
agrA_F1	tgaattttcatcaggagaagagttac
agrA_R1	cctcataaccttcttgcataaactc
agrB2_F1	catagttaatatttcggtagc
agrB2_R1	ttagctgaaaaaatagttatc
agrD1_F1	tgaagtttgctagctcattggc
agrD1_R1	cttctttagattgctgatttctttggg
agrD2_F1	cgttaaacttacttaaaaacatctcagc
agrD2_R1	gtaaagcttgtggctcttttgc
EBS universal	cgaaattagaaacttgcgttcagtaaac
lacZ_R1	gcgccattcgccattcaggctgcg
M13_F1	tgtaaaacgacggccagt
M13_R1	caggaaacagctatgacc
pyrE_F1	gagaaggaataaaaagtttagacg
pyrF_R1	accttcatttataaggctaactgc

rpoA_F1	ggagatatttgtagataaagg
rpoA_R1	ctcttgaggatttatactacc
rsbW_F1	tatatagggataaactttggcttgtg
rsbW_prom_F1	ttttttgcggccgcgtaaagattaatcttaaagatttagattatataga
rsbW_prom_R1	aaaaaacctaggttaaatatcaactcctaaatatttagtcatttttattttt
rsbW_screening_F1	aagctttgatgatgagggag
rsbW_screening_R1	cttcatactctatacttccaaagttacc
rsbW_homarm_F1	ttttttgtttaaacgtggcaaattggagcagtgt
rsbW_homarm_R1	ttttttgtttaaacaagaagcctccatagcctct
rsbW_HA_crispr_F1	atatatatcgcgcgcggtggcaaattggagcagtgt
rsbW_HA_crispr_R1	atatatatggcgcgccggtaactttggaagtatagagtatgaag
sigB_promoter_F1_ notI	ttttttgcggccgccataaataaggttgttgaaaaggc
sigB_F1_notI_true	ttttttgcggccgcggagttgatatttaatgaaaaatgtagc
sigB_R1_xbaI	tttttttctagacatctttaaaacagtaggtcaatac
sigB_SOE_F1	tttttttttaggaggtttttacatgtctatgaaaaatgtagctaatgc
sigB_SOE_R1	ttttttgcattagctacatttttcatagacatgtaaaaacctccta
sgRNA_flanking_2.0_F1	atatattctagatttatatttagtcccttgcctt
sgRNA_flanking_3.1_R1	atatatggcgcgccatatatatgcgatcgcataaaaataagaagcctgcaaatg
SOE_F1_intron1	tcacagtaatgtgaacaagg gttttagagctagaaatagcaagtt
SOE_R1_intron1	ccttgttcacattactgtga gtcgacgaaaactcctcct
SC7_F1	gacggatttcacatttgccgttttgtaaacgaattgcagg
SC7_R1	agatcctttgatcttttctacggggtctgacgctcagtgg

2.1.2 Bacterial growth conditions

2.1.2.1 Aerobic bacterial growth conditions

Luria Bertani (LB) broth or agar was used for the routinely culture of *Escherichia coli* strains. Cultures were incubated aerobically at 37°C in a shaker at 200 rpm. *E. coli* TOP10 was used as a vector host, while *E. coli* CA434 and *E. coli* NEB Sexpress were used as donor strains to transfer plasmid DNA into *C. difficile* R20291 NM and *C. difficile* R20291 BW, respectively. Chloramphenicol was added at a concentration of 12.5 µg/ml (in broth) or 25 µg/ml (in agar) where appropriate.

2.1.2.2 Anaerobic bacterial growth conditions

C. difficile strains were routinely grown in BHIS (brain heart infusion) supplemented with L- Cysteine [0.1%; Sigma, United Kingdom] and yeast extract [5 mg/ml; Oxoid]) broth or agar. Cultures were incubated at 37°C in an anaerobic cabinet (Don Whitley, United Kingdom) with an atmosphere consisting of 80% (v/v) nitrogen, 10% (v/v) hydrogen and 10% (v/v) carbon dioxide. The media was supplemented with thiamphenicol (15 µg/ml), when necessary. Prior to inoculation, solid media was pre-reduced for 4 hours in the cabinet whereas liquid media was pre-reduced for 8 hours under the same conditions. To assess the production of toxins, Tryptose-yeast (TY) broth medium (3%[w/v] Bacto tryptose, 2%[w/v] yeast extract and 0.1% [w/v] Thioglycolate) was used as growth media.

2.1.3 Measurement of *C. difficile* growth in liquid media

C. difficile cultures were measured spectrophotometrically with a Novaspec II visible spectrophotometer at a wavelength of 600nm. A fresh overnight culture was used to inoculate the growth media with a 1/100 inoculum. The culture was incubated anaerobically at 37°C. 1 ml samples were taken from the main culture and measured at different time points.

2.1.4 Preparation of electro-competent *E. coli* cells

A fresh overnight culture of *E. coli* TOP10 or CA434 was used to inoculate 200ml of pre-warmed LB broth with a 1/100 inoculum. The culture was incubated at 37°C with 200 rpm shaking until the bacterial culture reached exponential growth phase (0.5-0.8 OD_{600}). The culture was transferred into 50ml falcon tubes and chilled on ice for 30 minutes. The cells were harvested by centrifugation at 4000xg for 15 minutes at 4°C. The cell pellets were washed with ice-cold sterile distilled water and centrifuged as before. The washing step was repeated two more times. The cell pellets were resuspended in 4ml of sterile 10% glycerol and centrifuged as described above. The cell pellets were resuspended in 600µl of sterile 10% glycerol and stored in 50µl aliquots at -80°C.

2.1.5 Transformation of electro-competent *E. coli* strains by electroporation

Electroporation was carried out in a cold 0.2 cm electrode gap Gene pulser cuvette (Biorad) containing a 50µl aliquot of electro-competent *E. coli* thawed on ice and the desired DNA plasmid. An electric pulse of 2.5kV voltage, 25µF capacitance and 200Ω resistance was applied across the cuvette. A 0.3ml volume of SOC medium (Invitrogen, UK) was added to the cuvette. The cells were transferred to a sterile Eppendorf tube and were incubated under aerobic conditions at 37°C with 200 rpm shaking for 1 hour. The cells were then plated in LB agar plates with the appropriate antibiotics for the selection of transformants.

2.1.6 Transfer of plasmid DNA by conjugation

Conjugation between a donor *E.coli* CA434 and a recipient *C. difficile* strain was carried out using the “spotting” method described by Purdy (Purdy *et al.*, 2002). In short, 1 ml of overnight culture of donor *E. coli* was centrifuge at 4000xg for 1.5 minutes at room temperature. The pellet was washed with sterile PBS (phosphate-buffered saline) and centrifuged as before. The PBS was discarded and the cell pellet was transferred to the anaerobic work station. The pellet was re-suspended with 200µl of an overnight culture of the recipient *C. difficile* strain. The suspension was plated in a BHIS agar plate and it was incubated under anaerobic conditions at 37°C for 24 hours. The cells were harvested with a sterile plastic loop and re-suspended in 500µl of sterile PBS. 100µl of the suspension was plated onto BHIS agar containing cycloserine (250µg/ml), cefoxitin (8 µg/ml) and thiamphenicol (15µg/ml) to counter select against the growth of the donor E. coli strain and to select for *C. difficile* transformants. Plates were incubated at 37°C under anaerobic conditions for 24-72 hours.to allow growth of transconjugants colonies. Colonies were picked and restreaked on selective plates for further analysis.

2.1.7 Sporulation assay

Clostridium difficile strains were incubated in BHIS broth for five days at 37°C in an anaerobic cabinet. A starter culture was prepared using a 1/100 inoculum of a *C. difficile* overnight culture which was incubated until an OD_{600} between 0.2 and 0.5 was reached. The sporulation medium (BHIS broth) was then inoculated with a 1/100 inoculum of this culture. A 500µl sample was taken every 24 hours. This sample was removed from the cabinet and heated at 65°C for 30 minutes in order to kill all the vegetative cells, but not the spores. A non-heated sample was also removed from the cabinet to control for the effects of oxygen exposure. After this, both samples were returned to the cabinet and serially diluted in PBS; the dilutions were then plated onto BHIS agar supplemented with 0.1% taurocholate

(Sigma, UK.). Total CFU/ml were enumerated as described above before heat treatment. After 24 hour incubation period, colonies were counted and multiplied by the dilution factor. A *C. difficile spo0A* mutant was used as a sporulation negative control.

2.1.8 Supernatant assay

A starter culture was prepared using a 1/100 inoculum of a *C. difficile* overnight culture and it was incubated until an OD_{600} between 0.2 and 0.5 was reached. The sporulation medium (BHIS broth) was then inoculated with a 1/100 inoculum of this culture. After 48 hours the medium was filter sterilized with a 0.2µm filter and it was placed in the anaerobic cabinet for 8 hours. The supernatants were used to perform sporulation assays as described above.

2.2 Molecular methods

2.2.1 Extraction of genomic DNA

2.2.1.1 Extraction and purification of DNA with Spin column-based method

Genomic DNA was extracted and purified from *C. difficile* using the GenElute Bacterial Genomics DNA Kit (Sigma-Aldrich) according to the manufacturer's instructions.

2.2.1.2 Extraction and purification of genomic DNA with phenol/ chloroform/isoamyl alcohol

5 – 10 ml of *C. difficile* overnight culture were centrifuged at max speed for 5 minutes. The supernatant was discarded and the pellet re-suspended in 180µl of lysis buffer (10mg/ml lysozyme of egg chicken white [Sigma-Aldrich] in PBS) and it was incubated at 37°C for 30 minutes. 4µl of RNase A were added to the lysate and it was incubated for 30 minutes at 37°C. After the incubation, 25µl of proteinase K solution (20mg/ml [Qiagen]), 85µl dH_2O and 110µl 10% SDS solution were added to the lysate and it was incubated at 65°C for 30 minutes. An equal volume (400µl) of Phenol:Chloroform:Isoamyl alcohol mix (Sigma-Aldrich) was added and it was mixed thoroughly by inversion. The lysate was transferred to a phase-lock tube and it was centrifuged at max speed for 3 minutes. The aqueous phase (top phase) was transferred to a new phase-lock tube containing 400µl of Phenol:Chloroform:Isoamyl alcohol mix and it was mixed by inversion. The tube was centrifuged at max speed for 3 minutes. The top phase was transferred one more time in the same conditions and centrifuged as mentioned above. The top phase was transferred to a Eppendorf tube containing 40µl of 3M NaAc and 800µl of 100% ethanol and it was mixed by inversion. The tube was incubated at -80°C for 30 minutes. After the incubation time, the tube

was centrifuged at max speed for 15 minutes. The supernatant was discarded and the pellet was washed by adding 1ml of 70% ethanol; the pellet was then centrifuged for 3 minutes at max speed and the supernatant was discarded. The pellet was air dried for 45 minutes at room temperature. The pellet was then re-suspended in 50µl dH_2O.

2.2.2 Extraction and purification of plasmid DNA

Genomic DNA was extracted from *C. difficile* using the QIAprep spin mini prep (Qiagen) according to the manufacturer’s instructions.

2.2.3 DNA analysis in Agarose gel electrophoresis

A 1% (w/v) agarose gel in TAE (Tris-acetate-EDTA) buffer was used to separate DNA by applying a voltage of 85 – 110 V for 50 – 70 minutes. For the visualisation of the DNA, the liquid gel was stained 1:10 with SYBR Safe (Life technologies).

2.2.4 Purification of DNA from agarose gel

The DNA was extracted from the agarose gel using the QIAGEN QIAquick gel extraction kit according to the manufacturer's instructions. In short, the band was cut from the agarose and weight. The band was placed inside an Eppendorf tube and 3 volumes of Buffer QG were added. The tube was incubated at 59°C for 10 minutes or until the agarose piece was dissolved. 1 volume of isopropanol was added to the sample and it was mixed by inverting the tube. The mixture was transferred to a QIAquick column and centrifuged at 13000RPM. The flow-through was discarded and 50µl were added to the column. The column was incubated at room temperature for 1 minute and centrifuged as described above.

2.2.5 Quantification of nucleic acids using spectrophotometry

The ND-1000 spectrophotometer (NanoDrop; Thermo-Scientific, UK) was used to quantify chromosomal and plasmid DNA according to the manufacturer's instructions.

2.2.6 Digestion of DNA with restriction endonucleases

The restriction endonucleases were used according the manufacturer's protocol. In short, the reaction mixture were set up containing 1µg of DNA, 1µl of the restriction endonuclease and 1x buffer and dH_2O in a final volume of 20µl. The reaction was incubated 60 minutes at 37°C.

2.2.7 De-phosphorylation of DNA free ends

Shrimp Alkaline Phosphatase (NEB) was used following the manufacturer's instructions to remove 5' phosphate groups to prevent self-ligation of linearized DNA. Briefly, 1pmol of DNA ends was mixed with 2µl of CutSmart buffer and 16µl of dH_2O. 1µl of phosphatase was added to the mixture and incubated at 37°C for 30 minutes. The reaction was stopped by heating the mixture at 65°C for 5 minutes.

2.2.8 Ligation of DNA fragments

To ligate DNA fragments previously digested with restriction endonucleases, the T4 DNA ligase (NEB) was used according to the manufacturer's protocol. In short, the vector and the fragments were mixed in a 1:3 ratio, a 1 µl buffer was added and dnase free water was used to complete 9µl; a 1µl t4 ligase enzyme was added and the mixture was incubated for 10 minutes at room temperature.

2.2.9 Dialysis of DNA ligation reactions

After the incubation time for the ligation of DNA fragments, the entire mix was placed onto a 0.025µm nitrocellulose membrane over dH_2O for 30 minutes.

2.2.10 Purification of DNA from PCR reactions.

DNA from PCR reactions was purified using the QIAquick PCR purification kit (QIAGEN) according to the manufacturer's instructions. In short, 5 volumes pf PB buffer were added to 1 volume of PCR reaction. The mixture was transferred to a QIAquick spin column and centrifuged for 60 seconds at 13000 RPM. The flow-through was discarded and 0.75ml of PE buffer were added to the tube and the tube was centrifuged as before. The QIAquick column was transferred to a new tube and 50µl of dH_2O were added to the column. The column was incubated at room temperature for 1 minute and then centrifuged as mentioned above.

2.2.11 Amplification of DNA by standard Polymerase Chain Reaction

Routine PCR amplification used to screen for fragments was performed with the GoTaq 2X green master mix (Promega) according to the manufacturer's protocol. Whenever the PCR fragment needed to have high-fidelity for cloning, the Pfu DNA polymerase (Promega) was used instead.

Table 2.2. Oligonucleotides used in this study

OLIGONUCLEOTIDE	SEQUENCE 5' – 3'
rsbW_prom_F1	gcggccgctagactcaacaggcttaggagctatgatagg
rsbW_prom_R1	tctagagcattagctacatttttcattaaatatcaactcc
q5sdm_rsbW watermark_F1	tccaatagatgatatcgaagatataaagg
q5sdm_rsbW_watermark_R1	aaacccatcttattggcaattc
agrB2_F1	catagttaatatttcggtagc
agrB2_R1	ttagctgaaaaaatagttatc
agrD1_F1	tgaagtttgctagctcattggc
agrD1_R1	cttctttagattgctgatttctttggg
agrD2_F1	cgttaaacttacttaaaaacatctcagc
agrD2_R1	gtaaagcttgtggctcttttgc
pyrE_F1	gagaaggaataaaaagtttagacg
pyrF_R1	accttcatttataaggctaactgc
rpoA_F1	ggagatatttgtagataaagg
rpoA_R1	ctcttgaggatttatactacc
sigB_promoter_F1_ notI	ttttttgcggccgccataaataaggttgttgaaaaggc
sigB_F1_notI_true	ttttttgcggccgcggagttgatatttaatgaaaaatgtagc
sigB_R1_xbaI	tttttttctagacatctttaaaacagtaggtcaatac
sigB_SOE_F1	ttttttttaggaggtttttacatgtctatgaaaaatgtagctaatgc
sigB_SOE_R1	ttttttgcattagctacatttttcatagacatgtaaaaacctccta
rsbW_HA_crispr_F1	atatatatcgcgcgcggtggcaaattggagcagtgt
rsbW_HA_crispr_R1	atatatatggcgcgccggtaactttggaagtatagagtatgaag
sgRNA_flanking_2.0_F1	atatattctagatttatatttagtcccttgcctt
sgRNA_flanking_3.1_R1	atatatggcgcgccatatatatgcgatcgcataaaaataagaagcctgcaaatg

SOE_F1_intron1	tcacagtaatgtgaacaagg gttttagagctagaaatagcaagtt
SOE_R1_intron1	ccttgttcacattactgtga gtcgacgaaaactcctcct
rsbW_F1	ttttttgcggccgcgtaaagattaatcttaaagatttagattatataga
rsbW_R1	aaaaaacctaggttaaatatcaactcctaaatatttagtcatttttattttt
rsbW_screening_F1	aagctttgatgatgagggag
rsbW_screening_R1	cttcatactctatacttccaaagttacc
rsbW_homarm_F1	ttttttgtttaaacgtggcaaattggagcagtgt
rsbW_homarm_R1	ttttttgtttaaacaagaagcctccatagcctct
agrA_ht_ndei_F1	gctcccagtatacgcgcccatatgcaccatcaccatcaccataatagaaattttata tatttattaaaaatatataagaaagg
agrA_ndei_F1	gctcccagtatacgcgcccatatgaatagaaattttatatatttattaaaaatatataa gaaagg
agrA_xbai_R1	cccggtagttttcgcgcctctagattagcatactatatcccctaatatattagttatag c
agrA_ht_xbai_R1	cccgctagtatacgcgcctctagattaatggtgatggtgatggtggcatactatat cccctaatatattagttatagc
agrA_F1	tgaattttcatcaggagaagagttac
agrA_R1	cctcataaccttcttgcataaactc
EBS universal	cgaaattagaaacttgcgttcagtaaac
RAM-R	acgcgtgcgactcatagaattatttcctcccg
lacZ_R1	gcgccattcgccattcaggctgcg
M13_F1	tgtaaaacgacggccagt
M13_R1	caggaaacagctatgacc
SC7_F1	gacggatttcacatttgccgttttgtaaacgaattgcagg
SC7_R1	agatcctttgatcttttctacggggtctgacgctcagtgg

2.2.12 TOPO TA cloning

The TOPO TA cloning kit (Life technologies) was used to insert a PCR fragment into the pCR 2.1-TOPO vector for sub-cloning according to the manufacturer's instructions.

2.2.13 Insertional mutants in *C. difficile*

In order to inactivate genes in *C. difficile*, the ClosTron system was used as previously described by Heap, et al, 2007. A computer algorithm available at http:/www.clostron.com was used to identify target sites in the gene of interest. The target sites were synthesised and integrated into the pMTL007C-E2 plasmid by an external company (DNA 2.0, USA; Heap,et al, 2010).

2.2.14 RNA extraction

To protect the RNA from degradation, 10ml of RNAprotect (QIAGEN) were added to 5ml of *C. difficile* liquid culture and mixed thoroughly in anaerobic conditions. The mixture was incubated for 5 minutes at room temperature and then it was centrifuged at 5000 rpm for 10 minutes. The supernatant was discarded and the pellet was stored at -80°C. The pellet was re-suspended in 1ml of RNApro solution (MP Biomedicals) and transferred to a matrix tube. The matrix tube was processed in a homogeniser at 6400rpm for 40 seconds and then it was centrifuged at max speed for 15 minutes at 4°C. The supernatant was transferred to an Eppendorf tube and it was incubated at room temperature for 5 minutes to increase RNA yield. After the incubation time, 300µl of chloroform were added to the tube and it was mixed by vortexing for 10 seconds. The mixture was incubated for 5 minutes at room temperature to permit nucleoprotein dissociation and increase RNA purity. The mixture was centrifuged at 13000 rpm for 15 minutes at 4°C. The aqueous phase (upper phase) was transferred to a fresh Eppendorf tube containing 500µl of 100% ethanol and it was mixed by inversion. The mixture was then incubated at -20°C for 60 minutes. After the incubation period, the mixtures was centrifuged at 13000 rpm for 15 minutes at 4 °C. The supernatant was discarded and the pellet was air dried for 10 minutes at room temperature. The pellet was then re-suspended in 50µl of DEPC-treated water and stored at -80°C.

2.2.15 DNAse treatment of RNA

Contaminating DNA was removed from RNA samples using the Turbo DNA free kit (Ambion). Each sample (50µl) was mixed with 5.5µl of 10X buffer and 2µl of Turbo DNase and then incubated for 30 minutes at 37°C. Following the incubation time, 5µl of DNase inactivation reagent was added and mixed by vortexing the tube. The mixture was incubated for 5 minutes at room temperature and it was then centrifuged at 10000rpm for 1.5 minutes and the supernatant containing DNA-free RNA was transferred to a new Eppendorf tube.

2.2.16 Purification of DNase-treated RNA

The RNA samples treated with DNases were purified with the RNeasy mini-kit (QIAGEN) following the manufacturer's protocol.

2.2.17 RT-PCR

The RT-PCR was performed using the Quantitec reverse transcription kit (QIAGEN) following the manufacturer's protocol. Briefly, 2 µl of gDNA-Wipeout Buffer was added to an Eppendorf tube along with 1 µg of RNA and distilled water to a volume of 14 µl. The mixture was incubated for 2 minutes at 42°C. The tube was placed on ice after the incubation time. 4µl of Quantiscript RT buffer, 1µl of RT primer mix and 1µl of Quantitec reverse transcriptase were added to the tube. The content of tube was mixed and was incubated at 42°C for 30 minutes. After the incubation time the tube was heated at 95°C to inactivate the reverse transcriptase.

2.2.18 PCR optimisation

The primers used in the qPCR were optimised by performing a test with different concentrations. The primers were used in the following concentrations: 10mM, 5mM, 2.5mM and 1.25mM. The dissociation curve showed that the *rpoA* primers performed better when the forward primer had a concentration of 1.25mM, while optimal concentration for the reverse primer was 10mM. For the *agr*A primers, the forward primer was used at 10mM and the reverse primer at 5mM.

2.2.19 qPCR

The real-time PCR was performed using the QuantiTect SYBR green pcr kit (Qiagen) according to the protocol suggested by the manufacturer. In short, the reaction was set up mixing in a tube the following: 12.5µl Quantitec SYBR green, 1µl of each primer, 1µl of template and 4.5µl dH_2O. The reaction was mixed thoroughly and placed in the real-time cycler ABI7500 (Applied biosciences). The following amplification conditions were used: A holding stage at 95°C for 10 minutes, the cycling stage consisting of 40 cycles of 95°C for 15 seconds, 60°C for 1 minute and the melt curve stage consisting of 95°C for 15 seconds, 60°C for a minute, 90°C for 30 seconds and 60°C for 15 seconds. The default manual Ct threshold value of 0.2 was changed to 0.02.

2.2.20 Generation of R20291 BW Δ*pyrE* through Allele coupled exchange.

The plasmid pMTL-YN18 was introduced in *E. coli* NEB Sexpress through electroporation. The resulting strain was conjugated with *C. difficile* R20291 BW for 24 h in BHIs plates under anaerobic conditions. The growth was harvested and plated in BHIs plates supplemented with thiamphenicol (15µg/ml) cycloserine (250µg/ml) and cefoxitin (8µg/ml). Fast growing colonies were restreaked onto the same media and colonies were grown in BHIs broth supplemented with thiamphenicol (15µg/ml) overnight. PCR screenings with the pyrE_F1 primer (sequence located in the genome) and M13 primer (sequence located in the plasmid) revealed the colonies in which a single cross-over event had occurred. The colonies were restreaked in BHIs plates without antibiotic to promote the occurrence of the second cross-over event and the loss of the plasmid. After 5 days the growth was harvested and plated in minimal media containing FOA (2mg/ml) and Uracil (5µg/ml). Only colonies that have loss the function of the *pyrE* gene are able to survive, as FOA becomes toxic if an active *pyrE* gene is present. Six FOA-resistant colonies that had lost their resistance to thiamphenicol were tested for the *lacZ* fragment in the genome with the forward primer pyrE_F1 and the lacZ_R1 primers. Only true Δ*pyrE* mutants are able to produce the expected fragment (ca. 550bp), while the wild type is not able to amplify any region.

2.2.21 Restoration of *pyrE* into R20291 BW Δ*pyrE strains*

The *C. difficile* R20291 BW Δ*pyrE* recipient strains were grown overnight in BHIs broth supplemented with uracil (20µg/ml) and conjugated with the *E. coli* NEB Sexpress donor carrying the appropriate repair plasmid (YN2 or YN2C) overnight in BHIs agar plates supplemented with uracil (20µg/ml). The growth was harvested and plated in BHIs agar plates supplemented with uracil (20µg/ml), thiamphenicol (15µg/ml), cycloserine (250µg/ml) and cefoxitine (8µg/ml). After 3 days, the 6 largest colonies were re-streaked in the same medium. The largest colony from each re-streak was re-streak in the same medium after 24 h. Fast growing cells were harvested and plated onto *C. difficile* minimal medium (CDMM) without uracil. The 6 largest and 6 smallest colonies found in the CDMM plates were re-streaked in the same medium and incubated for 24 hours. Single colonies were taken from the CDMM plates and inoculated in BHIs broth containing cycloserine (250µg/ml) and cefoxitin (8µg/ml). PCR was used to confirm correction of the *pyrE* gene. The culture that restored the *pyrE* gene was plated in BHIs plates to isolate single colonies. Replica-plating was used to identify colonies that lost the plasmid-encoded Tm-resistance. The Tm-sensitive colonies were re-screened with PCR and amplicons were sent for sequencing to confirm that the *pyrE* gene was reverted back to wild type.

2.3 *In vitro* methods

2.3.1 Vero cell culture.

Vero cells were taken from storage and re-suspended in Dulbecco's Modified Eagle Medium (DMEM) (Sigma), supplemented with 10% (v/v) heat inactivated foetal calf-serum (Sigma), 2mM L-glutamine (Sigma) and penicillin-streptomycin (Sigma) pre-warmed at 37°C to a density of 2 x 10^5 cells/ml. 20ml of this solution was then transferred to a 75 cm^2 tissue culture flask. The flask was incubated at 37°C *in 5%* CO_2 until monolayers reached confluence, normally after 2 days. After this the media was poured off the flask and the monolayer was washed with PBS twice. 5ml trypsin-EDTA was added and the flask was incubated at 37°C for 5 minutes. 5ml of fresh media was added to the flask and the content was transferred to a falcon tube. The tube was then centrifuged for 15 minutes at 1000 RPM at room temperature. The supernatant was discarded and the cells were re-suspended in fresh media to a final density of 2 x 10^5 cells/ml. 100 µl of this solution were poured in each well of a 96 well plate. The plates were incubated for 48 hours in 5% CO_2 at 37°C.

2.3.2 Cytotoxicity assay

C. difficile strains were grown in TY medium as described above. A starter culture was prepared using a 1/100 inoculum of a *C. difficile* overnight culture and it was incubated until an OD_{600} between 0.2 and 0.5 was reached. The growth medium (TY broth) was then inoculated with a 1/100 inoculum of this culture. A 1 ml sample was taken every 6, 12, 24, 48 and 72 hours. The samples were centrifuged and filter-sterilized. Supernatants were diluted in fourfold series and 20µl of each dilution was added to a 96-well plate containing monolayers of Vero cells prepared as described above. The plates were incubated in 5% CO_2 at 37°C for 24 hours. Cytotoxicity effects were evaluated after the incubation period by microscopy.

2.3.3 Toxin ELISA

The *C. DIFFICILE* TOX A/B II ELISA (Techlab) kit was used for the detection of *C. difficile* toxins A and B. The ELISA test for toxin A and B was performed according to the manufacturer recommendations. A starter culture was prepared using a 1/100 inoculum of a *C. difficile* overnight culture and it was incubated until an OD_{600} between 0.2 and 0.5 was reached. The growth medium (TY broth) was then inoculated with a 1/100 inoculum of this culture. Samples were collected at the 72 hour mark. The samples were centrifuged and filter-sterilized. The supernatants were diluted in tenfold series with PBS. 50µl of conjugate were applied to each well to be used in the plate, followed by 100µl of diluted sample. A positive and a negative control were run simultaneously with the samples. A standard curve was prepared using the native toxins A and B from *C. difficile* (List Biological Laboratories Incorporated, USA). The plate was incubated for 50 minutes at 37°C. The plate was then washed 4 times with the wash solution provided in the kit. The plate was dried and 100µl of substrate were applied to each well, followed by incubation at room temperature for 10 minutes. 50µl of stop solution were added to each well and the plate was incubated for 2 minutes at room temperature before reading it in a spectrophotometer at 450nm.

2.3.4 Biofilm assay

To measure the biofilm formation, a starter culture of *C. difficile* was prepared by inoculating fresh BHIs broth containing 0.1M glucose with an overnight culture of the desired strain in a 1:100 dilution. Tissue culture 24-well plates (Costar,USA) were pre-reduced in the anaerobic cabinet for 48 hours prior to use. 1 ml of the starter culture was added per well and the plate was incubated for 120 hours. To avoid the risk of liquid evaporation, the plate was wrapped in parafilm. After the incubation time, the wells were washed with PBS and the plate was allowed to dry for 10 minutes. The biofilm was stained with 1ml of filter-sterilised 0.2% (w/v) crystal violet solution per well and it was incubated for 30 minutes at 37°C in anaerobic conditions. The staining solution was removed and the wells were washed twice with PBS. The plate was removed from the anaerobic cabinet and 1ml methanol was added to the wells for the removal of the dye from the biofilm and it was incubated for 30 minutes at room temperature. The methanol-extracted dye was diluted 1:1, 1:10 and 1:100 and the absorbance A570 was measured with Ultrospec 500 pro spectrophotometer.

2.3.5 Motility assay

2.3.5.1 Swarming assay

To assess the swarming motility of the different strains, 0.4% (w/v) BHIs agar was prepared and cooled down at 50°C for 30 minutes. The agar was poured into plates and it was allowed to solidify for 2 hours. The plates were dried at 37°C for 60 minutes before placing them inside the anaerobic cabinet to pre-reduce for 4 hours. 5µl from an overnight culture of *C. difficile* were applied onto the middle of the plate and it was incubated for 48 hours inside the cabinet. After the incubation time, the halo diameter was measured.

2.3.5.2 Swimming assay

For the swimming motility assessment, 0.3% (w/v) BHIs agar plates were prepared and allowed to solidify overnight. The plates were dried at 37°C for 60 minutes and then they were placed inside the anaerobic cabinet to pre-reduce for 4 hours. One single colony was inoculated by stabbing the middle of the plate. After 48 hours the halo diameter was measured.

2.3.6 Peptide complementation assay

A series of synthetic peptides were used to evaluate their potential to restore the phenotype of the *agrB* mutants made in *C. difficile* R20291 NM. The synthetic peptides were dissolved individually in sterile DMSO to a final concentration of 10mM. BHIs broth was supplemented with these AIP's to a final concentration of 20µM. The supplemented BHIs broth was used as culture media for the *agrB1* and *agrB2* mutants and the strains were assay for spore formation over 5 days.

2.3.7 Transmission electron microscopy for the visualization of flagella in *C. difficile.*

To evaluate the number of flagella in *C. difficile*, a 20µl droplet of mid-exponential phase liquid culture was placed on a formvar/carbon 200 mesh cooper grid (Agar Scientifics) and it was incubated at room temperature for 5 minutes. After the incubation period, the excess of bacterial culture was wicked off with a small piece of Whatman filter paper. The cells were fixed to the grid by adding 20µl of glutaraldehyde to the wells and incubating for 1 minute. The grid was washed 3 times with sterile distilled water. The grid was then negatively stained with uranyl acetate (pH 4.5, Agar Scientifics) and the excess was removed as mentioned before. Electron micrographs of the samples were taken using a Tecnai G2 12 Bio Twin microscope with a fully embedded digital imaging system with SIS Megaview III camera (Advance Microscopy Unit, Queen's Medical

Centre, University of Nottingham, United Kingdom). The transmission electron microscope was run at a 100kv.

2.3.8 Resistance of *C. difficile sigB* mutant strains to oxidative stress.

Cultures of *C. difficile* R20291 BW and *C. difficile* R20291 BW sigB mutant strains were grown overnight in BHIs broth under anaerobic conditions. Fresh BHIs media was inoculated with a 1:100 inoculum and the strains were grown until they reached exponential growth (OD_{600} 0.3 to 0.5). Hydrogen peroxide (stock 9.8M, Sigma) was added to the media to a final concentration of 1.4mM or 1.6mM. Negative controls of each culture were taken previous to the addition of the hydrogen peroxide to the media. The cultures were incubated for 1h under anaerobic conditions. The number of viable bacteria was determined by plating aliquots of the appropriate dilutions on BHIs agar and incubating overnight.

2.4 Statistical analysis

All statistical analysis was done using the GraphPad Prism software, using student's t-test with a p value cut off $p<0.01$.

Chapter 3

Characterisation of the *agr* loci in *C. difficile* R20291

3.1 Introduction

3.1.1 *Clostridium difficile* BI/NAP-1/027 epidemic strain

In recent years the worldwide emergence of epidemic strains of *C. difficile* has exacerbated its harmful effects in healthcare settings with increased morbidity and mortality as a consequence of CDI. A prominent grouping of these epidemic strains are so-called BI/NAP1/027 strains on account of them belonging to the restriction endonuclease analysis group BI, pulse-field gel electrophoresis type NAP1 and PCR-Ribotype 027 (RT-027). These strains (from here on referred to as RT-027 strains) have been reported to be the leading cause of CDI outbreaks between 2002 and 2003 in eight healthcare facilities in the United States (Warny *et al.*, 2005). In 2005 it was reported that a RT-027 strain was responsible for an outbreak in Canada (McDonald *et al.*, 2005), where 82% of the stool samples tested positive to the 027 strain. The rising in the incidence of the RT-027 strains in Europe has been reported in 16 countries across the continent (Kuijper *et al.*, 2008). One RT-027 strain, *C. difficile* R20291, was responsible for an outbreak at the Stoke Mandeville hospital in the United Kingdom in 2005 (Buckley *et al.*, 2011; Brazier *et al.*, 2007).

A surveillance study performed between 2007 and 2008 by Brazier *et al.* (2008) analysed 677 isolates obtained from 186 hospitals across nine geographical regions of England and revealed that the *C. difficile* R20291 strain was the most common strain isolated from symptomatic patients. RT-027 strains are characterised by a higher production of toxins A and B, resistance to fluoroquinolone and the production of a binary toxin. Due to the higher virulence seen in these strains, they are sometimes referred to as "hypervirulent". They are largely associated with an incremental increase in the incidence of hospital-acquired CDAD causing a more severe disease as well as higher relapse rates and mortality (Goorhuis *et al*., 2007; Redelings *et al*., 2007). Another study suggests that RT-027 strains are able to produce a higher amount of spores compared to non-outbreak strains (Akerlund *et al.,* 2008).

A comparative genomic and phenotypic study performed by STabler *et al.* (2009) showed the similarities and differences between a historic non-epidemic RT-027 strain (CD196), a previously sequenced RT-012 strain (630) and the epidemic RT-027 strain (R20291). It was observed that the genomes were highly conserved among the strains, but also that the RT-027 strains genomes contain 234 additional unique genes not present in the genome of strain 630. A large portion of the RT-027-specific genes are regulatory genes, including 8 two-component regulators and 15 other transcriptional regulators. An additional complete copy of the *agr* locus was also found in the RT-027 genome and named *agr2.* It consists of the genes *agrA, agrB2, agrC* and *agrD2*, whereas the genome of *C. difficile* 630 contains only a partial *agr* locus named *agr1*, also present in R20291, with single copies of the *agrB1* and the *agrD1* genes. It is important to notice that CD196 strain belonging to the RT-027 is non-epidemic in spite of harbouring the *agr*2 operon. This study also showed that the R20291 strain contains five genetic regions consisting of 47 genes not found in CD196; these additional genes present in R20291 seem to be required in order to achieve the higher virulence seen in this strain compared to CD196.

The *agr* locus plays an important role in the regulation of virulence factors in *S. aureus* (Novick & Geisinger, 2008) as well as in several clostridia species such as *C. perfringens, C.* and *C.acetobutylicum* (Chen *et al.,* 2012; Cooksley *et al*., 2010; Steiner *et al*,. 2012). The possibility exists that the presence of the second *agr* locus in R20291 could play a role in the regulation of virulence factors.

During the course of this work, a study published by Martin *et al.* (2013) concerned with the analysis of an R20291 *agrA* mutant suggested that there were major differences between the R20291 strain in their laboratory (that of Brendan Wren at the London School of Hygiene and Tropical Medicine, LSHTM) and the R20291 that is in use here at the University of Nottingham. The Nottingham strain (R20291 NM) has previously been shown to possess just a single flagellum (Baban *et al*., 2013), in contrast the same strain from LSHTM (R20291 BW) was found to be multi-flagellated (Martin *et al.,* 2013) . Our further characterisation of the two strains, in terms of genome sequencing and phenotypic characterisation, indeed demonstrated that the differences between the two were highly significant (see Chapter 4) and that whereas R20291 BW represents an authentic copy of the strain isolated from Stoke Mandeville, and adheres to the published genome sequence, strain R20291 NM does not. In view of this, we elected to recreate all of the *agr* mutants described in Section 3.3.2.1 in R20291 BW and re-assess their phenotypes. Moreover, as the two strains differ significantly in terms of motility and in the formation of biofilms, these analysis were added to phenotypic characterisation of the *agr* mutants created in the R20291 BW strain.

3.2 Aim of this chapter

The rise of the so called "hypervirulent" strains has been related to their ability to produce a higher amount of toxins and spores than other strains, as well as their resistance to fluoroquinolones. The RT-027 strain R20291, responsible for outbreaks in the UK, contains several regulatory genes not present in strain 630 which may play an important role in the expression of virulence factors. One set of regulatory genes that have been reported to have an important regulatory function in other bacteria is the *agr* locus. In *S. aureus,* the *agr* locus is involved in the regulation of virulence factors such as toxins, biofilm production and haemolysins. It has been shown that the *agr* locus present in other *Clostridium* species is involved in the regulation of spores and toxins. The development of the ClosTron technology (Heap *et al.*, 2010) offers the opportunity to create insertional mutants in the genome. In this study, the role of the *agr* loci in the regulation of virulence factors is investigated through the generation of insertional mutants.

3.3 Results

3.3.1 Phenotypic characterization of *C. difficile* 630Δ*erm* containing the *agr2* locus from R20291 NM

3.3.1.1 Construction of the *C. difficile* 630Δ*erm* mutant strains containing the *agr2* locus

We wanted to evaluate the impact of the *agr2* locus in the production of toxins and spores as the genome of *C. difficile* 630Δ*erm* does not contain a copy of the *agr2* locus found in the *C. difficile* R20291 genome. The *C. difficile* 630 strain used in our laboratory is a strain isolated in the Mullany laboratory (UCL, London, UK) after 30 repeated subcultures in non-selective media designated *C. difficile* 630Δ*erm*; this strain is a variant of *C. difficile* 630 that possesses a specific deletion in the *ermB* gene and has become erythromycin-sensitive, thus allowing the use of *ermB* as a selective marker. In our laboratory, the use of the modular vector pMTL84151 has been extensively used for the introduction of sequences into *C. difficile.* The pMTL84151 vector (Genbank accession number FJ797649) harbours the pCD6 Gram-positive replicon from the *C. difficile* indigenous plasmid pCD6; the Gram-negative replicon from the ColE1 plasmid, the *traJ* transfer function of the RP4 oriT region; the thiamphenicol-resistance gene *catP* from *C. perfringens* and the *lacZ* gene encoding the alpha fragment of the *E. coli* β-galactosidase containing a multiple cloning site (MCS) derived from plasmid pMTL20 (Heap *et al.*, 2009). The pMTL84151::*agr2* plasmid, which contains the 2,808 bp *agr2* locus, as well as 400 bp upstream and 269 bp downstream, was previously created in our laboratory (Patel *et al.,* unpublished) by cloning this sequence into the MCS region of pMTL84151 (Fig. 3.1). The plasmids were sent for sequencing to confirm their sequences using the M13_F and M13_R pair of primers. The pMTL84151::*agr2* plasmid was introduced to the donor strain *E. coli* CA434 via electroporation. Subsequently, the pMTL84151::*agr2* plasmid was transferred to *C. difficile* 630Δ*erm* via conjugation as described in Chapter 2, using *E. coli* CA434 harbouring the pMTL84151::*agr2* plasmid as a donor strain. The resulting strain was designated *C. difficile* 630Δ*erm* pMTL84151::*agr2*. Additionally, the empty vector pMTL84151 was also conjugated in *C. difficile* 630Δ*erm* as a control and the resulting strain was designated *C. difficile* 630Δ*erm* pMTL84151.

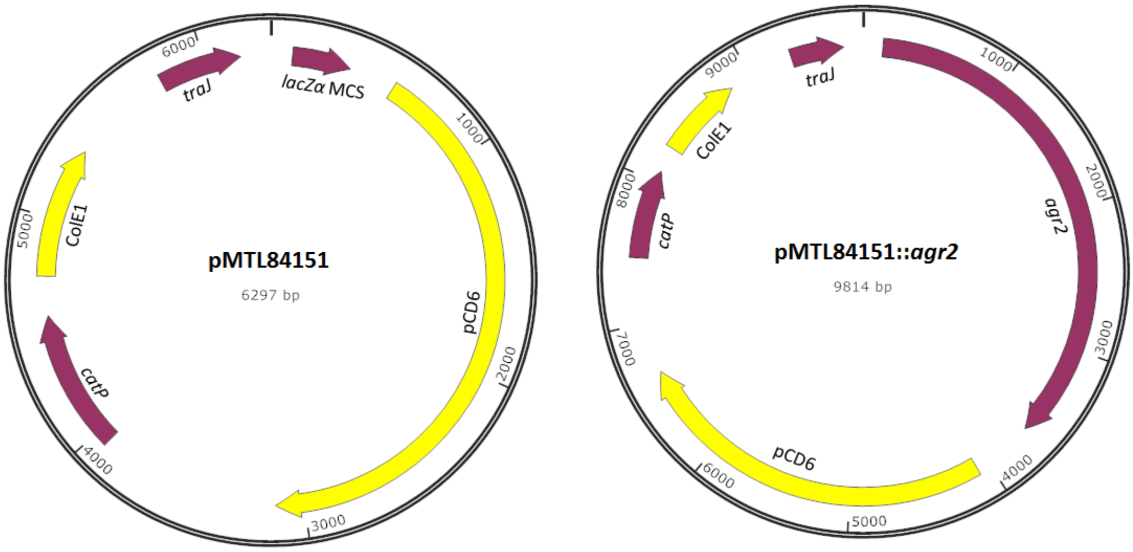


Figure 3.1 Plasmid maps of shuttle vector pMTL84151 and pMTL84151::*agr2*. Common plasmid components are: *catP* gene of *C. perfringens* conferring thiamphenicol resistance; ColE1, the replication region of the *E. coli* plasmid ColE1, and; *traJ*, transfer function of the RP4 oriT region; pCD6, replication region of *C. difficile* indigenous plasmid pCD6. The modular vector pMTL84151 also harbours *lacZα* MCS, the *lacZ* gene encoding the alpha fragment of the *E. coli* β-galactosidase containing a multiple cloning site region, MCS, derived from plasmid pMTL20. In pMTL84151::*agr2*, the *agr2* locus found in strain R20291 NM has been inserted into the MCS region.

3.3.1.2 Sporulation in *C. difficile* 630Δ*erm* mutant strains

To evaluate the effect of the *agr2* operon from R20291 on the phenotype of *C. difficile* 630Δ*erm*, the mutant strains were tested to determine the number of heat-resistant CFU present after a 5-day period. Following heat shock, cells were plated on BHIs agar supplemented with the bile salt taurocholate to promote spore germination. In order to enumerate total CFU/ml (spores and vegetative cells) and heat-resistant CFU/ml (spores), two samples of each strain were taken every 24 h; one sample was serially diluted and plated in BHIs agar plates containing 0.1% of taurocholate (total CFU/ml) and the second sample was heat-treated at 65° C for 30 min before being serially diluted and plated in BHIs agar plates supplemented with 0.1% taurocholate (spores). After the 5-day period, no difference was observed in the ability to form spores between strains since all the mutants produced the same amount of heat-resistant CFU/ml (Fig. 3.2). The strain containing the empty vector was able to produce the same amount of spores as the parental strain, suggesting that the presence of the vector doesn't affect the sporulation process in *C. difficile* 630Δ*erm*. The mutant strain containing the *agr2* locus also showed the same amount of CFU/ml observed in the parental strain. These data indicate that the introduction of the *agr2* locus does not affect the regulation or production of spores in *C. difficile* 630Δ*erm*.

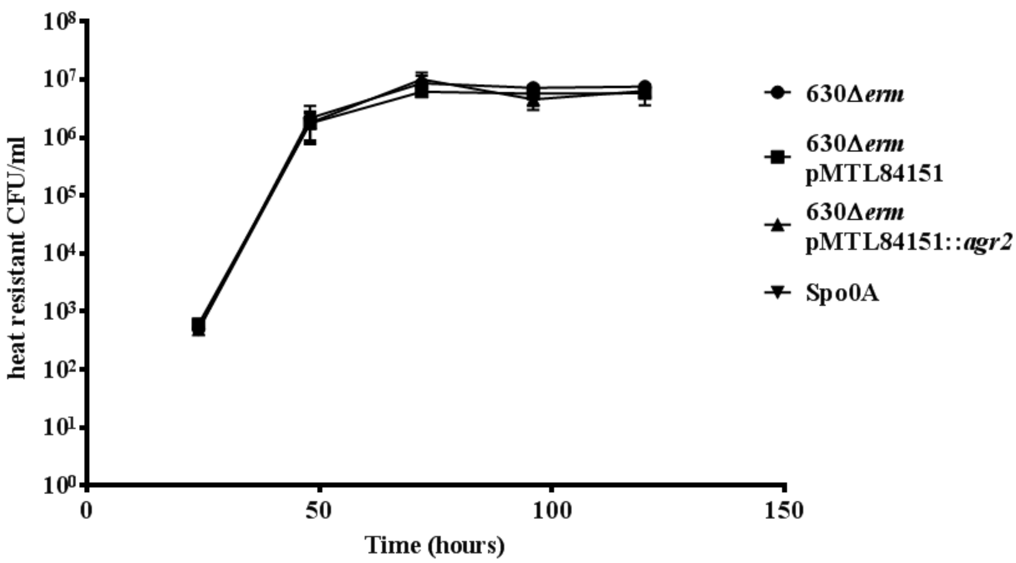


Figure 3.2. Production of heat-resistant CFU in *C. difficile* 630Δ*erm* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of *C. difficile* 630Δ*erm*) could not be measured. The detection limit was 50 CFU/ml.

3.3.1.3 Toxin production in *C. difficile* 630Δ*erm* mutant strains

To observe the effect of the *agr2* locus from R20291 NM on the production of toxins A and B, *in vitro* cell cytotoxicity assays were performed in Vero (African green monkey kidney) cells. This cell line is sensitive to both toxin A and B, rounding up and dying in the presence of the toxins. *C. difficile* 630Δ*erm* strains were cultured in TY medium and samples were collected, centrifuged and filter-sterilised at the times mentioned in Chapter 2 in order to collect the cell-free supernatants. The supernatants were diluted in a 4-fold serial dilution and 20 µl of each dilution was added to the cell line. The end point titre was defined as the first dilution in the series in which the morphology of the Vero cell is the same as the negative control. The negative control was inoculated with sterile TY medium. *C. difficile* R20291 is known to produce a higher concentration of toxins compared to strain 630 and it was included as a reference to observed if the mutant strain containing the *agr2* locus was able to produce similar concentrations.

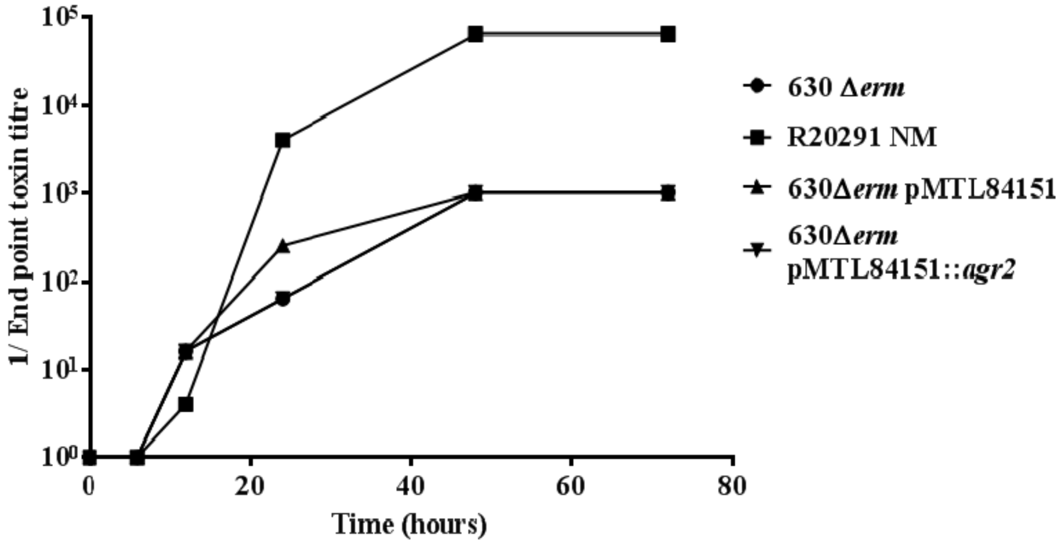


Figure 3.3. Toxin production of *C. difficile* 630Δ*erm* mutant strains and *C. difficile* R20291 over 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time points as described in Materials and Methods. Strains used are indicated in the accompanying key. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means.

It was observed that the introduction of the empty vector in 630Δ*erm* didn't have any impact on the production of toxins since the supernatant of the mutant contained equivalent amounts of toxin to the wild type (Fig. 3.3). The introduction of the *agr2* locus in 630Δ*erm* also resulted in no measureable increase in toxin production compared to the parental strain. In order to establish if the *agr2* operon was being expressed following its introduction into 630Δ*erm*, RT-qPCR was undertaken to demonstrate transcriptional activity. *C. difficile* 630Δ*erm* cultures were grown anaerobically and the RNA was extracted as described in Chapter 2. The extracted RNA was subjected to reverse transcription and primers were designed to amplify the *agrA* (agrA_F1 and agrA_R1) and *agrB2* (agrB2_F1 and agrB2_R1) sequences in the resulting cDNA using the SYBR green PCR master mix (ThermoFischer). As *agrB2* and *agrD2* are transcribed from the same mRNA, as well as *agrA* and *agrC* (Patel et al., unpublished), it was not felt necessary to also target *agrC* or *agrD2*. The amplification of the sequences was quantified using the ABI7500 instrument (Applied Biosystems). RT-qPCR was also undertaken against *rpoA*, a house keeping gene used as an endogenous control, using the rpoA_F1 and rpoA_R1 pair of primers for the amplification of the *rpoA* fragment. The endogenous control gene corrects for variation in RNA content, variation in reverse-transcription and degradation of the RNA sample.

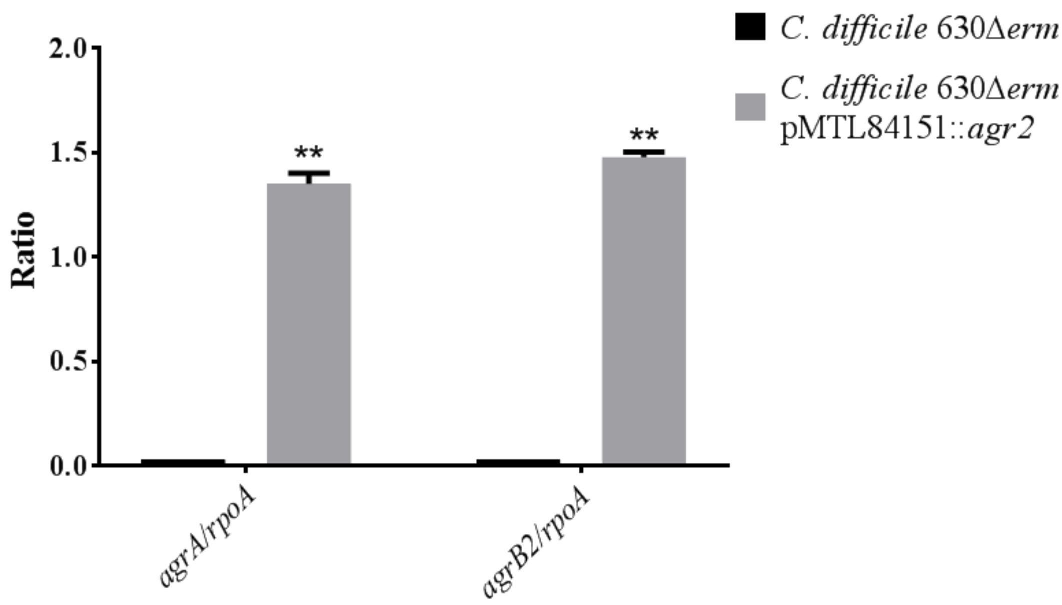


Figure 3.4 Expression of the *agrB2* and *agrA* genes by *C. difficile* 630Δ*erm* harbouring the vector plasmid pMTL84151::*agr2*. RT-PCR was used to evaluate the expression of the *agrA* and *agrB2* genes present in the pMTL84151::*agr2* plasmid. The housekeeping gene *rpoA* was used as an endogenous control. The ratio between the expression of the desired gene and *rpoA* was compared. *C. difficile* 630Δ*erm* does not naturally harbours the *agr2* operon and thus was used as a negative control. No expression of *agrA* or *agrB2* was detected in *C. difficile* 630Δ*erm.* Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the wild type.

The data obtained by RT-qPCR revealed that the genes in the *agr*2 operon are being expressed in the cell (Fig. 3.4). This indicates that the fact that strain 630Δ*erm* does not produce more toxin, equivalent to R20291 NM, is not due to lack of expression of *agr*2 and indicates that the reasons for higher toxin production in R20291 is more complex than initially thought and that other factors are involved, possibly related to the additional genes that are present in R20291 NM but not in 630Δ*erm* (STabler *et al.*, 2009).

3.3.2 Phenotypic characterization of *C. difficile* R20291 NM *agr* mutants

3.3.2.1 Construction of mutant strains in R20291 NM

To assess the role of the *agr* genes in *C. difficile* R20291 NM, the ClosTron system (Heap *et al*., 2007) was used to create insertional mutants in the *C. difficile* R20291 NM strain targeting the *agrA*, *agrB1* and *agrB2* genes. The resulting strains were designated *C. difficile* R20291 NM *agrA*::CTermB (NM *agrA* mutant), *C. difficile* R20291 NM *agrB1*::CTermB (NM *agrB1* mutant) and *C. difficile* R20291 NM *agrB2*::CTermB (NM *agrB2* mutant); the insertional mutants in *C. difficile* R20291 NM were previously created by Manisha Patel in our laboratory (Patel *et al*., unpublished). The integrity of the ClosTron-derived mutants were checked by PCR screening using a forward flanking primer and the EBS universal primer specific to intron encoded sequences (Fig 3.5). The latter sequence is only located in the intron, while the flanking primer is specific to a region just upstream of the intended insertion site. Consequently a DNA fragment will only be amplified by PCR if the intron is present in the genome in the intended gene target. The NM *agrA* and NM *agrB2* mutants were complemented with the plasmid pMTL84151::*agr2*, which contains the 2,808 bp *agr2* locus as well as 400 bp upstream and 269 bp downstream the *agr2* locus. Additionally, the NM *agrA* mutant was complemented with the pMTL84151::*agrA* plasmid, which contains the native *agrA* gene and the 400bp upstream region. The NM *agrB1* mutant was complemented with the pMTL84151::*agrBD1* plasmid, carrying the 726 bp *agrBD1* sequence and 452 bp upstream the *agrBD1* locus. The plasmids used for complementation (Fig. 3.6) in the mutants were previously created by Patel *et al.* (unpublished). The complementation plasmids were extracted and sent for sequencing to confirm their sequences. The complementation plasmids were introduced to their respective *agr* mutants made in *C. difficile* R20291 NM via conjugation as described in Chapter 2, using *E. coli* CA434 as a donor strain. The modular vector pMTL84151 was also introduced in the R20291 NM *agr* mutants as control.

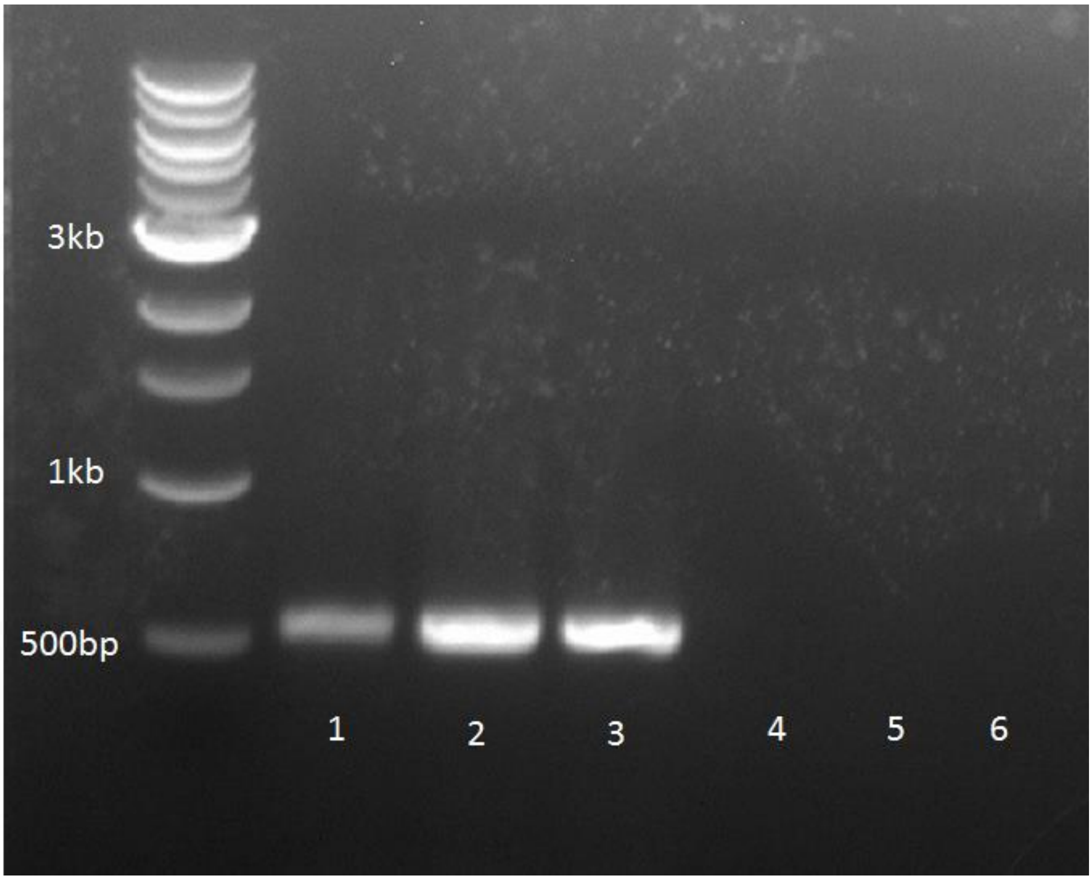


Figure 3.5 Authentication of ClosTron insertions in *C. difficile* R20291 NM *agr* mutants. Chromosomal DNA from the three *agr* mutants and the parental strain were used as the template for PCR with a primer pair composed of the EBS universal primer and a primer specific to the mutant in question; namely agrA_F1 (*agrA*), agrB1_F1 (*agrB1*) and agrB2_F1 (*agrB2*). Respective PCR mutant samples are: R20291 NM *agrA* mutant (lane 1); R20291 NM *agrB1* mutant (lane 2) and R20291 NM *agrB2* mutant (lane 3). Lanes 4-6 are the control reactions using the R20291 parental template DNA and EBS universal primer in conjunction with agrA_F1 (lane 4), agrB1_F1 (lane 5) and agrB2_F1 (lane 6). The lane on the extreme left of the gel contains a 1KB DNA ladder.

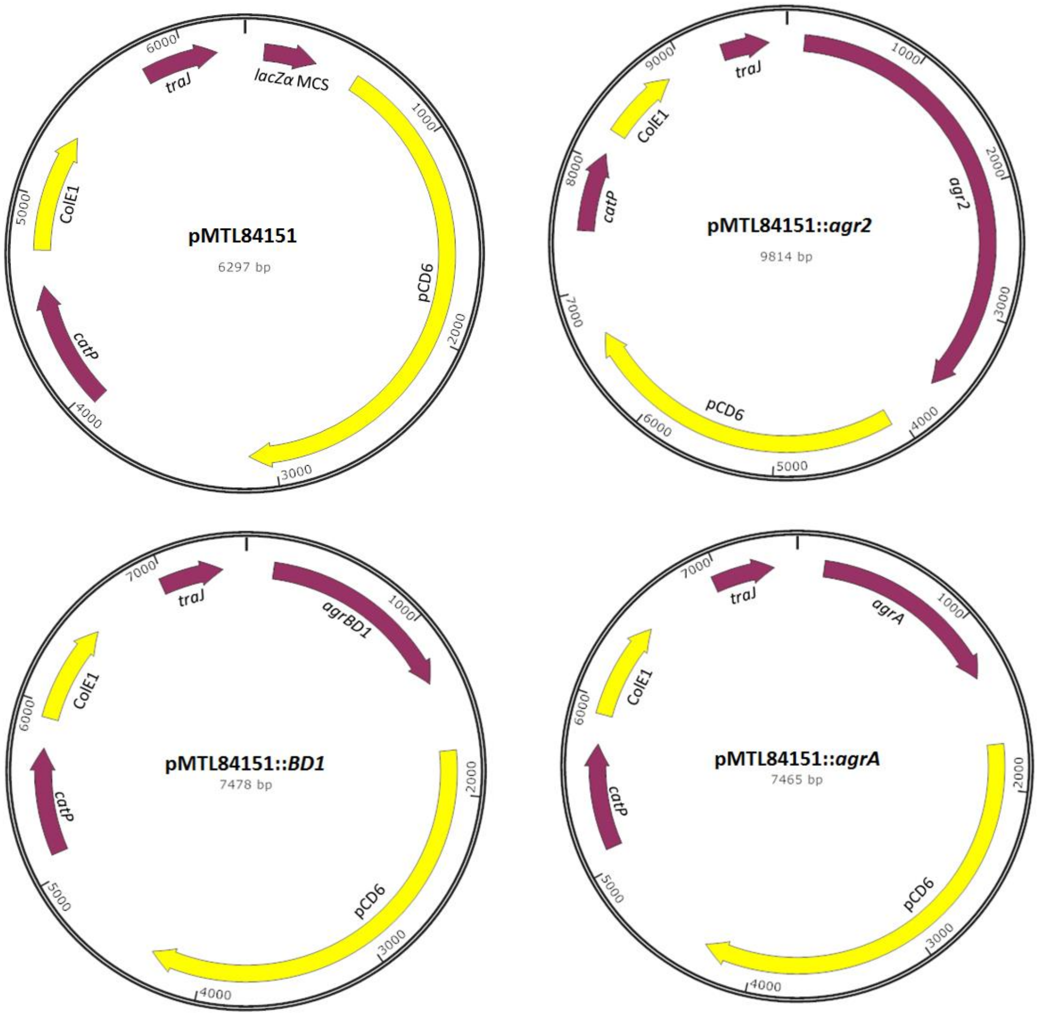


Figure 3.6 Plasmids used for the complementation of the *agr* mutants made in *C. difficile* R20291 NM. The complementation plasmids were created by cloning the desired sequence into the MCS found in the modular vector pMTL84151. Common plasmid components are: *catP* gene of *C. perfringens* conferring thiamphenicol resistance; ColE1, the replication region of the *E. coli* plasmid ColE1, and; *traJ*, transfer function of the RP4 oriT region; pCD6, replication region of *C. difficile* indigenous plasmid pCD6. The modular vector pMTL84151 also harbours *lacZα* MCS, the *lacZ* gene encoding the alpha fragment of the *E. coli* β-galactosidase containing a multiple cloning site, MCS, region derived from plasmid pMTL20.

3.3.2.2 Sporulation in *C. difficile* R20291 NM mutants

The *agr* mutants in R20291 NM were subjected to the sporulation assay described in Chapter 2 to evaluate the effect of the inactivation of the *agr* genes. Briefly, 2 samples of each strain were taken every 24 h over 5 days, one of them being heat-treated at 65°C for 30 min, serially diluted and plated in BHIs agar plates containing 0.1% of the bile salt taurocholate.

Spores, as measured by heat-resistant CFU formation, were detecTable in the R20291 NM wild type cultures after 48 h. In contrast, no spores were evident in cultures of the NM *agrB1* mutant at the 48 h time point. Rather, the first evidence for the presence of heat-resistant CFU occurred at the 72 h time point, some 24 h later than the wild type (Fig 3.7). The NM *agrB1* mutant also had a lower count of heat-resistant colonies in relation to the parental strain. The NM *agrB1* mutant containing the empty vector (pMTL84151) exhibited a similar profile to that of the NM *agrB1* mutant, having a delayed time to sporulation and a lower heat-resistant CFU count compared to the R20291 NM wild type. Interestingly, the presence of the complementation plasmid in the NM *agrB1* mutant strain restored the time at which spores first began to appear to the 48 h time point. Nevertheless, the heat-resistant CFU count was lower than the wild type, similar to the NM *agrB1* mutant. This inability to match the parental levels of heat-resistant CFU could be due to fact that although spores can be detected at the same time point as the wild type, the production of spores occurs at a slower rate.

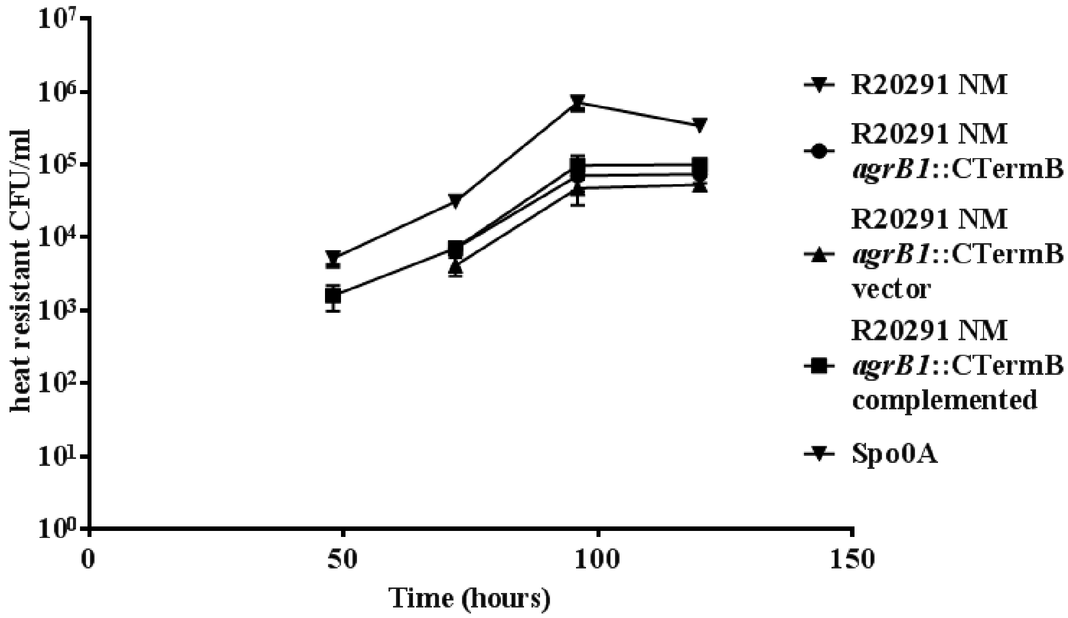


Figure 3.7. Production of heat-resistant CFU in *C. difficile* R20291 NM *agr*B1 mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and the complemented strain harbours pMTL84151::*agrBD1*. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

In contrast to the NM *agrB1* mutant, the spores could be detected in the NM *agrB2* mutant after just 24 h, 1 day before the wildtype strain (Fig 3.8). Complementation of the *agrB2* defect through addition of the entire *agr2* operon on a plasmid (pMTL84151::*agr2*) restored the timing of spore appearance to that of the wild type, with heat-resistant CFU becoming detecTable at 24 h, albeit at slightly higher counts than the wild type. Confirmation that this alteration to the sporulation phenotype was attribuTable to the presence of the *agr2* operon and not the plasmid was obtained by showing that the vector alone (pMTL84151) had no effect on the sporulation phenotype of the NM *agrB2* mutant. It is important to notice that, even though the NM *agrB2* mutant and the NM *agrB2* mutant containing pMTL84151 sporulate at an earlier time point and display a higher spore titre at this earlier phase of growth, the levels of heat-resistant CFU is similar between all the strains at the end of the experiment.

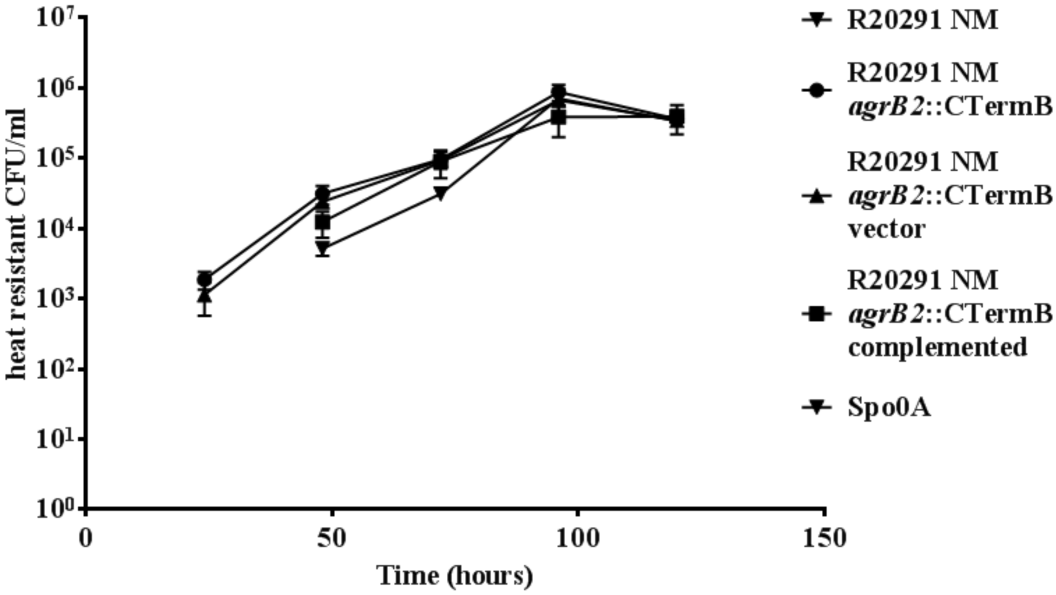


Figure 3.8. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrB2* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and the complemented strain harbours pMTL84151::*agr2*. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

In contrast to the result obtained with the NM *agrB* mutants, inactivation of *agrA* appeared to have no effect on the time at which spores could be detected (Fig. 3.9). As expected, the profile of the NM *agrA* mutant carrying the empty vector and the NM *agrA* mutant containing the vector with *agr*2 locus are similar to the *agrA* mutant, with heat-resistant CFU being detected at the 48 h time point. This suggest that the *agrA* gene doesn't play an important role in the regulation of sporulation in R20291 NM.

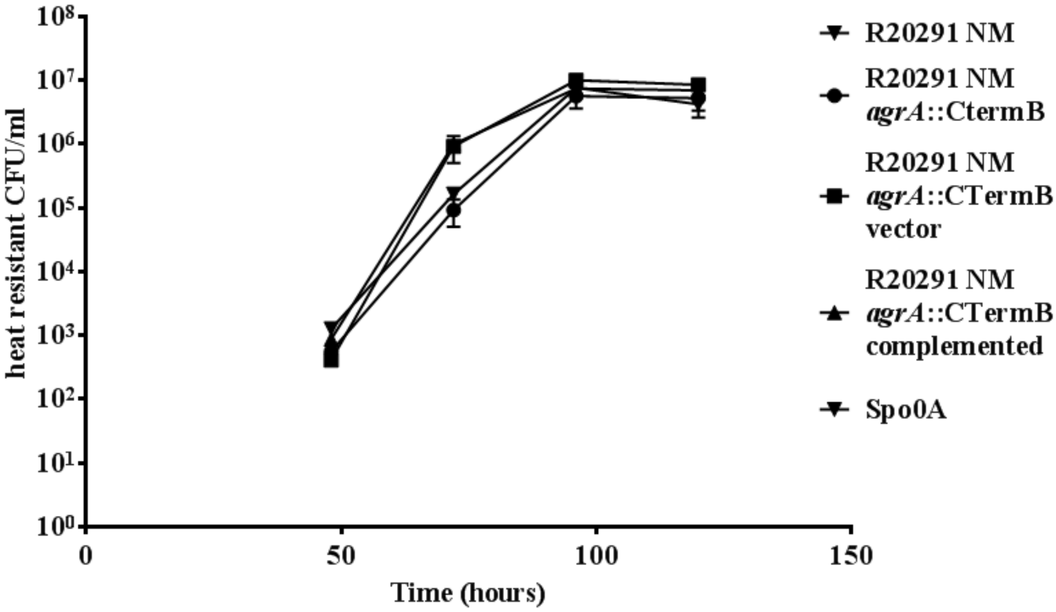


Figure 3.9. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrA* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and the complemented strain harbours pMTL84151::*agr2*. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

Since the samples for the detection of spores were taken every 24 h, there is the possibility that the initiation of sporulation in some of the strains starts just a few hours after the sample is taken but it is not detected until the next time point, making it unclear if the difference is significant or not. In order to establish the time difference in sporulation between the R20291 NM wild type and the NM *agrB* mutant strains, a more in-depth sporulation assay was performed. Samples were taken every 4 h for 24 h prior to the time point where heat-resistant CFU were first detected in the strains (Fig.3.10). Heat-resistant CFU were detected after 20 h in the NM *agrB2* mutant while the appearance of spores in the R20291 NM wild type was observed after 36 h; as expected, sporulation in the NM *agrB1* mutant was delayed compared to the wild type and it was detected after 64 h. The results show that the timing of initiation of sporulation varies greatly between the strains and that the onset of sporulation occurs more than 10 h apart. This confirms that the difference in sporulation time is significant among the NM *agrB* mutants.

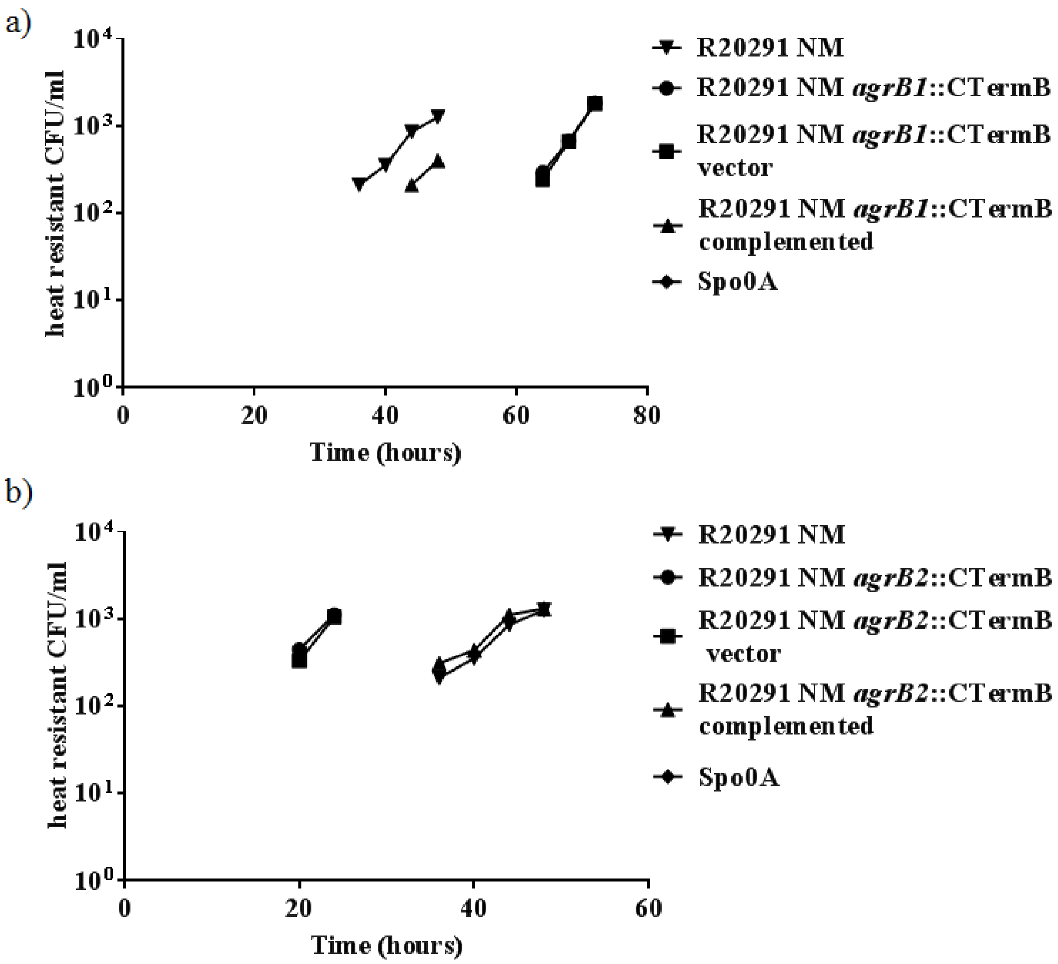


Figure 3.10. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrB* mutant strains. a) R20291 NM *agrB1* mutants. b) R20291 NM *agrB2* mutants**.** The samples were taken every 4 h for 24 h prior to the time spores were first detected in the previous assay. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector; the complemented strain refers to pMTL84151::*BD1* for the NM *agrB1* mutant and pMTL84151::*agr2* for the NM *agrB2* mutant. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

3.3.2.3 Toxin production in *C. difficile* R20291 NM *agr* mutants

Due to the difficulties and challenges of handling cell lines, added to the number of strains that needed to subject to this test, an alternative ELISA (Enzyme-linked immunosorbent assay) method for the detection of toxins was chosen. The commercially available *C. difficile* TOX A/B II ELISA kit was used to detect the amount of toxin produced by the NM *agr* mutants. In short, a sample of *C. difficile* culture was taken after 72 h and the supernatant was filter-sterilised to remove all the vegetative cells and spores. The supernatants were prepared in a 10-fold dilution series and were added to the ELISA plate wells. The experiment then was performed as described in Chapter2.

The R20291 NM *agrA* mutant exhibited an interesting phenotype as its toxin production was reduced significantly compared to the wild type (Fig. 3.11). The R20291 NM *agrA* mutant carrying the empty vector also shared this reduction in toxin production, confirming that the presence of the empty vector has no effect in the toxin production of the strain. The toxic production in the R20291 NM *agrA* mutants complemented with the *agr*2 locus or the *agrA* gene were able to produce similar toxin levels to the parental strain. This phenotype suggests that despite the fact that the *agrA* gene did not play an important role in the regulation of sporulation, it is a key component in the regulation of toxins; although it is not yet clear if it has a direct or indirect role.

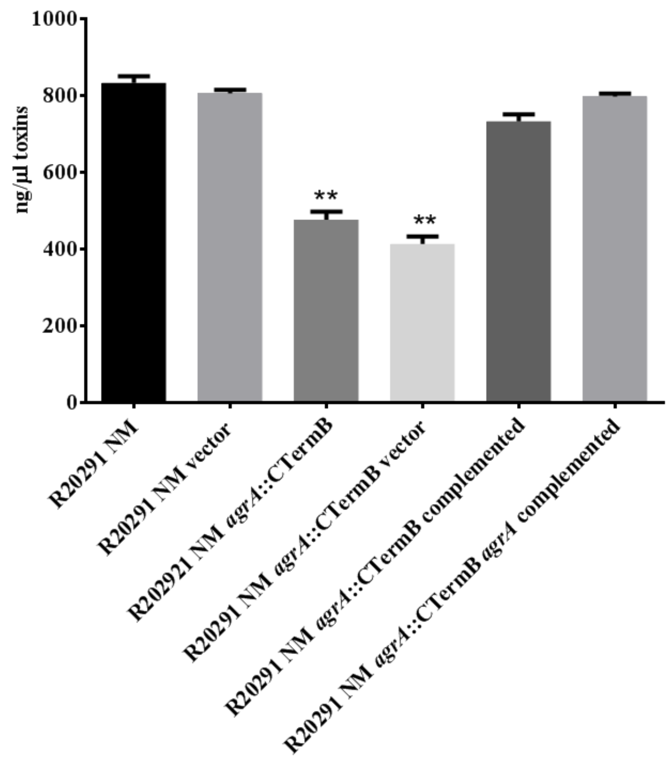


Figure 3.11. Toxin production of *C. difficile* R20291 NM *agrA* mutant strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and complemented refers to pMTL84151::*agr2* and pMTL84151::*agrA* respectively. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the wild type.

The R20291 NM *agrB* mutants also exhibited variations in the level of toxins compared to the parental strain (Fig 3.12). The NM *agrB1* mutant produced a significantly higher amount of toxins than the parental strain; in contrast, the toxin levels in the NM *agrB2* mutant decreased. Interestingly, the mutants carrying the vector, whether it was empty or carrying the complementing gene, had a lower amount of toxins compared to the R20291 NM strain and the NM *agr*B mutants. This behaviour was not seen in the sporulation phenotype, where the mutants carrying the empty vector displayed the same profile as the corresponding *agr* mutant, and suggests that the presence of the plasmid interferes with the toxin production in a manner yet to be determined.

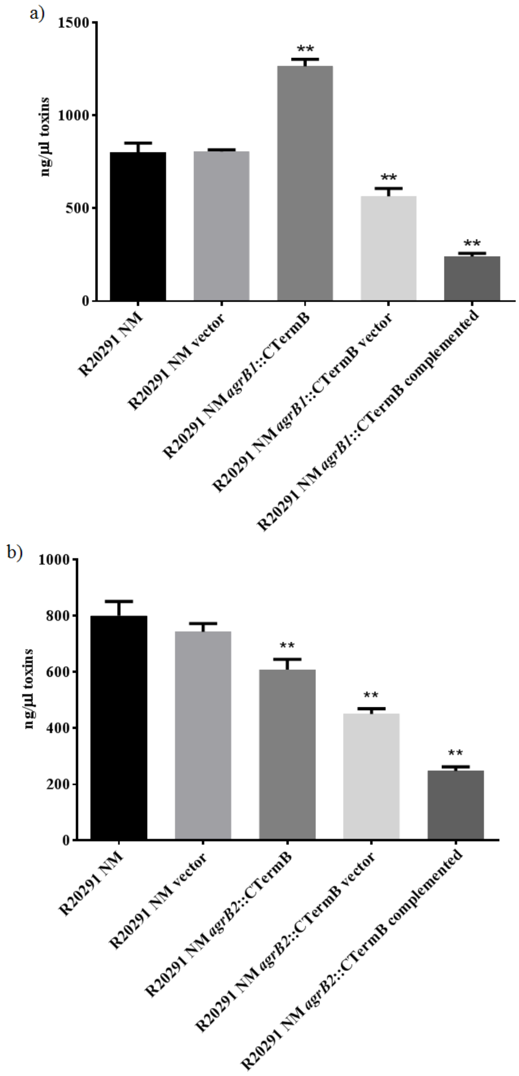


Figure 3.12. Toxin production of *C. difficile* R20291 NM *agrB* mutant strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. pMTL84151 has been abbreviated as vector and the complemented strain refers to pMTL84151::*BD1* and pMTL84151::*agr2* respectively. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's t test) compared to the wild type.

3.3.3 Use of supernatants for the complementation of R20291 NM *agrB* mutants

It has been shown in *C. perfringens* type C that the phenotypes seen in *agrBD* mutants, the downregulation of *pfoA*, the gene responsible for the transcription of the theta-toxin, can be reverted to that of the wild type by growing the strains in stationary-phase supernatants derived from wild type cultures (Ohtani *et al.,* 2009). The supernatants taken from the wild type are capable of restoring the phenotype observed in the *agrBD* mutants as it contains the putative AIP. As seen in the sporulation assay, the NM *agrB1* and NM *agrB2* mutants have a different sporulation onset in relation to the wild type, implicating that the *agrB* genes play a role in the regulation of sporulation. The NM *agrB1* mutant, which possesses an intact *agrB2* gene, shows a delayed sporulation time (72 h), while the NM *agrB2* mutant strain, which has a functional *agrB1* gene, displays an earlier onset of sporulation (24 h). These findings suggest that the *agrB1* gene promotes the onset of sporulation while the *agrB2* negatively regulates the sporulation process in this strain. We wanted to evaluate the ability of the supernatant taken from the R20291 NM wild type to restore the sporulation time of the NM *agrB1* and NM *agrB2* mutants back to the time observed in the wild type (48 h). Since the R20291 NM wild type spores can be detected at the 48 h time point, this time point was chosen for the recollection of the supernatant as the presence of sporulation suggests that the concentration of AIP in the media is enough to trigger the sporulation process. Additionally, supernatants from the NM *agrB1* and *agrB2* mutants were collected to assess their ability to delay or bring forward the onset of sporulation in the other strains.

Briefly, starter cultures were grown in BHIs broth overnight and the culture media was inoculated with a 1/100 inoculum. After 48 h, the supernatants were filter sterilised with a 0.025μM filter and reduced in the anaerobic cabinet for at least 8 h. These supernatant were used as the culture media and sporulation assays were performed as described in Chapter 2. Table 3.1 shows the time point at which spores where first detected during the sporulation assay in the different strains. Contrary to what was expected, the NM *agrB* mutant strains grown in the supernatant from the R20291 wild type did not restore the onset of sporulation to the time point seen in the parental strain (48 h) and displayed a similar behaviour as seen in the fresh media. The only exception to this finding was the NM *agrB2* mutant strain carrying the complementation vector which showed an earlier sporulation time, with spores being detected at the 24 h time point, similar to the R20291 NM *agrB2* mutant strain. The use of the NM *agr*B1 mutant supernatant also showed no difference in the sporulation time of the strains with the exception of the NM *agrB2* mutant strain carrying the complementation vector, which displayed a similar behaviour to the *agrB2* mutant, as seen in the R20291 supernatant.

Table 3.1 First detection of spores in R20291 NM and *agr*B mutants grown in supernatants taken after 48 h. R20291 NM supernatants were used to evaluate their ability to restore the sporulation onset of the *agrB* mutants created in this strain. Additionally, supernatants from the *agrB* mutants were used to evaluate their impact in the initiation of sporulation in the different strains. The Table shows the first time point where spores can be detected.

Strains\Supernatant	Fresh BHIs Supernatant (Reference)	R20291 NM Supernatant	R20291 NM *agrB1*::CtermB Supernatant	R20291 NM *agrB2*::CtermB Supernatant	630Δerm Supernatant
R20291 NM	48h	48h	48h	24h	24h
R20291 NM *agrB1*::CTermB	72h	72h	72h	24h	24h
R20291 NM *agrB1*::CTermB pMTL84151	72h	72h	72h	24h	24h
R20291 NM *agrB1*::CTermB pMTL84151::*agrB D1*	48h	48h	48h	24h	24h
R20291 NM *agrB2*::CTermB	24h	24h	24h	24h	24h
R20291 NM *agrB2*::CTermB pMTL84151	24h	24h	24h	24h	24h
R20291 NM *agrB2*::CTermB pMTL84151::*agr2*	48h	24h	24h	24h	24h

The NM *agrB2* mutant supernatant had a significant impact. The onset of sporulation in all of the strains grown in this supernatant was similar to that of the NM *agrB2* mutant, with spores being detected in all the strains at the 24 h time point. This suggests that the NM *agrB2* supernatant contains the necessary factors to induce the expression of the sporulation pathway. Since the *agrB1* is the only AgrB protein working in the NM *agrB2* mutant, it is possible that the *agrBD1* operon is responsible for the production of the AIP involved in the onset of sporulation. The *agrB1* gene present in strain 630Δ*erm* shares a 100% similarity to the one present in R20291 NM. Since the supernatant from the *agrB2* mutant showed an improvement in sporulation in all the other strains, the *agrB1* gene could be responsible for the signalling molecule that is involved in the onset of sporulation. In order to evaluate if the supernatants of strain 630Δ*erm* had a similar effect on the onset of sporulation in the NM *agrB* mutants, stationary-phase supernatants were taken after 48 h, filter-sterilised and used as media to grow R20291 NM and the *agr*B mutants. The sporulation assays was performed as previously described. The results show that the behaviour of all the strains is similar to the behaviour observed when grown in the presence of supernatants derived from the NM *agrB2* mutant. All the strains included in this study initiated sporulation at an earlier time (24 h) and exhibited an increase in final spore titres. This suggests that the signalling factors needed to initiate sporulation in R20291 NM are also present in the supernatant of strain 630Δ*erm*. Since 630Δ*erm* doesn't have the *agrB2* present in R20291, it is likely that the *agrB1* gene is the one involved in the positive regulation of spores.

3.3.4 Complementation of *agr*B mutants with synthetic peptides

It has been shown in *C. acetobutylicum* that the addition of synthetic peptides resembling the putative AgrD sequence of that strain was able to revert the impaired sporulation and granulose formation phenotype observed in the *agr*B null mutant (Steiner *et al.*, 2012). A protein alignment of the AgrD proteins from several *Clostridium* species (Fig. 3.13) showed the similarities of the sequences compared to the encoded AgrD polypeptides of *C. difficile* R20291. The predicted sequences are based on the AgrD signal molecules found in *S. aureus*, consisting of a 5-aminoacid ring, containing a thioester bond between cysteine and the first amino acid, and a residual tail of up to 4 amino-acids.

```
S. aureus group 1       -MNTLFNLFFDFlTGlLKNlGNlAAYSTCDFIMDEVEVPKELTQLHE--
S. aureus group 2       MNT-LVNMFFDFllKLAKAlGlVGGVNACSSLFDEPKVPAELTNLYDK-
S. aureus group 3       MKK-LLNKVlELLVDFFNSlGYRAAYINCDFLLDEAEVPKELTQLHE--
S. aureus group 4       MNT-LLNlFFDFlTGVLKNlGNVASYSTCYFIMDEVEVPKELTQLHE—

C. difficile agrD1      -MKKFlVRFMKFASSLALSTAlLSANSTCPWllHQPKVPKElSNLKKTN
C. difficile agrD2      -MKKIALNLLKNISALSFGIAVLSANSASSWVAHQAKEPQALQKLKK--
C. botulinum agrD1      -MKKLSKKVLMLVATFTTLLASIVASSACVWCVYQPEEPKCLREE----
C. botulinum agrD2      MKKQLKEKCTKVTAKLLKSVAYSTADSACHLGlYQPKEPKSLRK-----
C. sporogenes agrD1     -MKKLSKKVLMLVATFTTLLASVVASSACVWCVYQPEEPKCLREE----
C. sporogenes agrD2     MKKQLKEKCTKVTAKLLKSVAYSTADSACHLGlYQPKEPKSLRK-----
C. beijerinckii         -MKRLNNKVLMAVAAFATVIASVVSTSACFWAMYQPEEPKALRDE----
C. butyricum            ----MKNKILMGIATFATVIASVVATSACIWGHYQPEEPKCLRDE----
C. tunisiense           -MKKLNKKFLALIATITTLMATTVATSACTWYFYQPEEPKCLSEE----
```

Figure 3.13. Protein alignment of AgrD proteins from *S. aureus* and several clostridia species. The different AgrD proteins found in *S. aureus* are shown at the top while putative sequences found in different clostridial species are shown below. The amino acids coloured red represent a conserved cysteine present in all the predicted proteins except *C. difficile* AgrD2. The grey residues show the residues forming the 5-amino acid ring. The residues coloured blue represent the residues found in *S. aureus* AIPs; in clostridia, the blue residues represent potential tail residues. The yellow residues are conserved sequences in clostridia species while the green residues are semi-conserved regions.

A series of synthetic peptides were used to evaluate their ability to complement the NM *agrB* mutants and revert the phenotype seen in the onset of sporulation in these strains (Fig 3.14). The synthetic peptides were obtained from Dr. Klaus Winzer (Centre for Biomolecular Sciences, Nottingham, UK). The cyclic peptides were synthesised and purified using HPLC by Peptide Protein Research Ltd (UK). The manufacturer ensured the cyclic structure and purity (ranging between 70-95%) of the peptides by conducting RP-HPLC and electrospray mass spectrometry. The peptides were derived from the predicted AgrD1 and AgrD2 amino acid sequences, assuming the formation of a 5 ring structure and a residual tail. The length of the N-terminal tail sequence was based on AIPs found in other bacteria (Nakayama J *et al.*, 2006; Sturme et al., 2005; Novick & Geisinger, 2008).

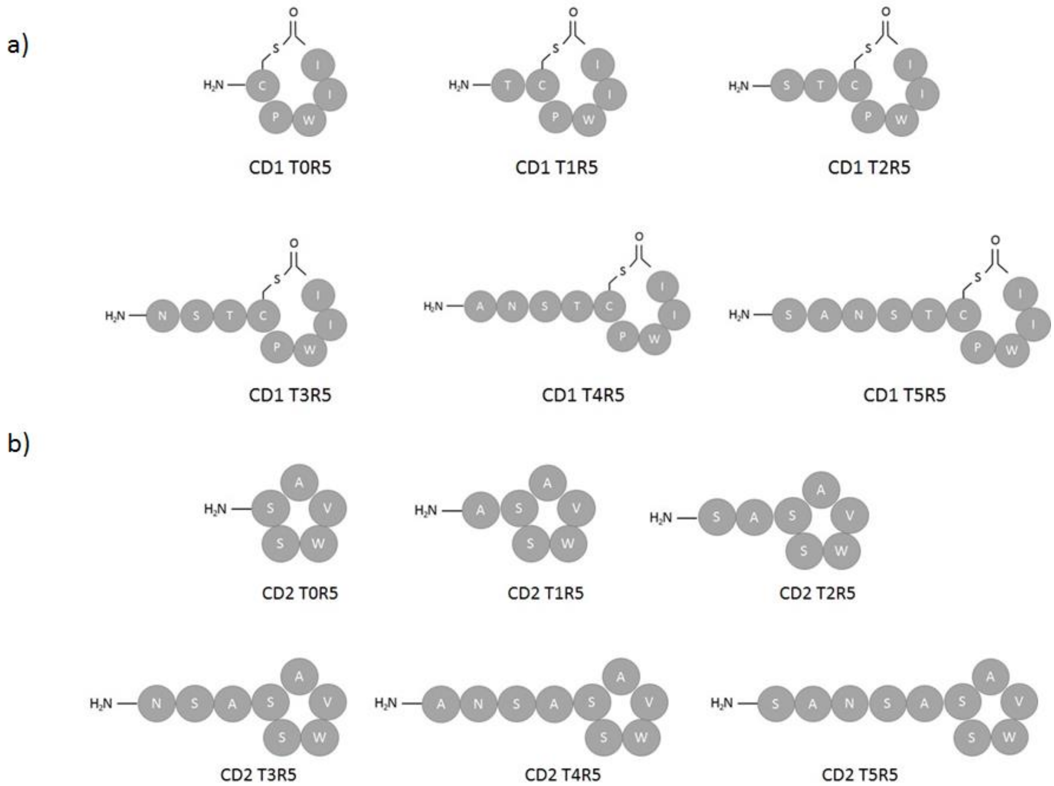


Figure 3.14. Synthetic peptides derived from the putative AgrD sequences found in *C. difficile* R20291. a) Representation of peptides derived from the predicted AgrD1 AIP, consisting of a 5-aminoacid ring with a thiolactone bond between the first aminoacid and the cysteine positioned in the 5^{th} place. The tail vary in length from 1 to up to 5 aminoacids. b) Synthetic peptides dervied from the predicted AgrD2 AIP consisting of a 5-aminoacid ring and a tail of up to 5 aminoacids. The cyclic peptides were synthesised and purified using HPLC by Peptide Protein Research Ltd (UK).

In *C. acetobutylicum*, the *agrB* insertional mutant displays an impaired sporulation phenotype; *C. acetobutylicum agrB*::CTermB cultures supplemented with synthetic peptides to a final concentration of 20µM had a more pronounced effect on the restoration of sporulation compared to the cultures supplemented with a final concentration of 2μM (Steiner *et al.,* 2009). Due to these findings, BHIs broth was supplemented with the synthetic peptides to a final concentration of 20 μM was used as growth media for the NM *agrB* mutants; the ability of the synthetic peptides to restore the onset of sporulation in the *agrB1* and *agrB2* mutants made in strain R20291 NM was evaluated over 5 days. The peptides derived from the AgrD1 sequence were used to complement the R20291 NM *agrB1* mutant and the peptides derived from AgrD2 were tested in the R20291 NM *agrB2* mutant.

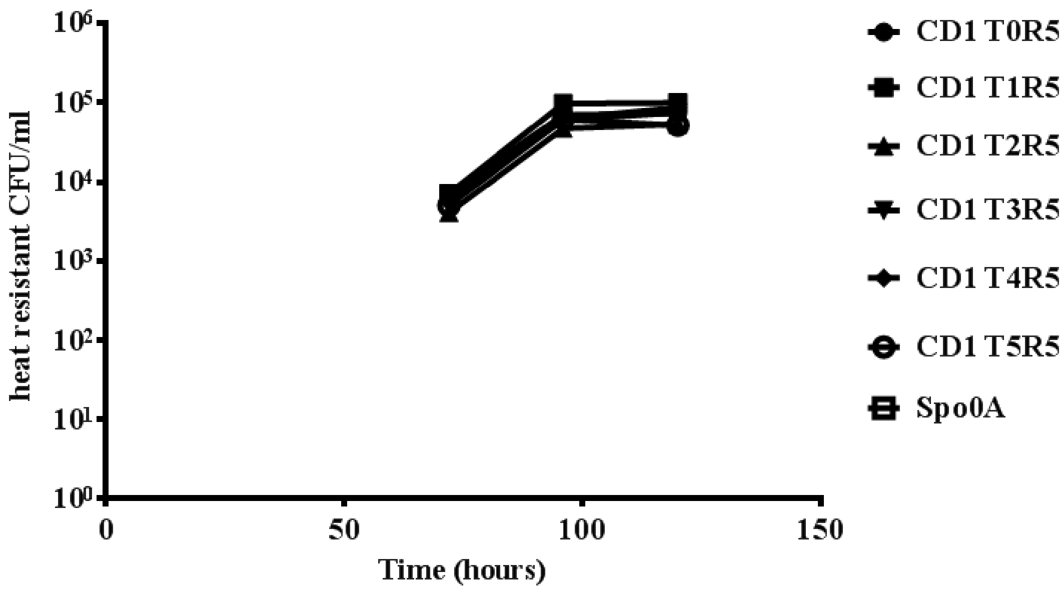


Figure 3.15. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrB1* mutant grown in BHIs containing 20 μM of the respective synthetic peptide. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Synthetic peptides used are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

The addition of exogenous AIPs to the NM *agrB* mutants did not show any effect on the sporulation profile of these strains (Fig 3.15, Fig 3.16). The NM *agrB* mutants performed as previously seen in fresh BHIs broth, as no change in the timing or rate of sporulation was observed. One of the reasons this behaviour is seen could be due to the inadequate quantity of AIP in the culture media. The AIPs with a 3-amino acid N-terminal tail (CD1-T0R3 and CD2-T0R3) were tested in different concentrations to evaluate if this could trigger changes in the sporulation profiles, as these are the predicted AIPs. These AIPs were selected as they have the average number of amino acids comprising the tail. The variation on the concentration of the AIPs in the culture media was tested to evaluate the effects on sporulation in the NM *agrB1* and NM *agrB2* mutants. BHIs broth used as culture media was complemented with the AIPs containing the 3-amino acid N-terminal tail to a final concentration of 200µM, 20µM, 2µM and 0.2µM. The sporulation assays were performed over 5 days as described in Chapter 2.

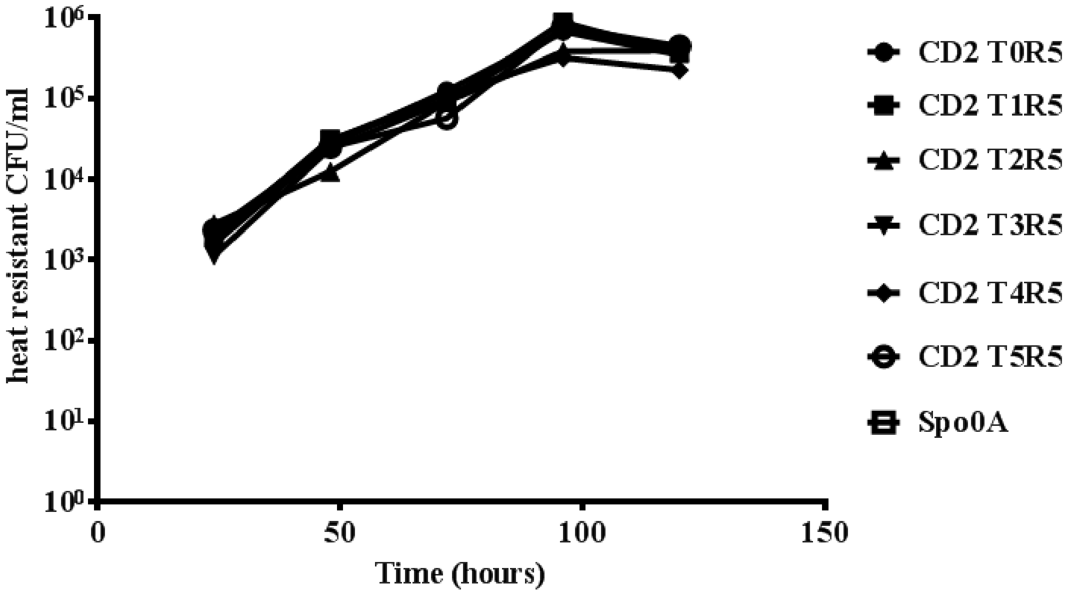
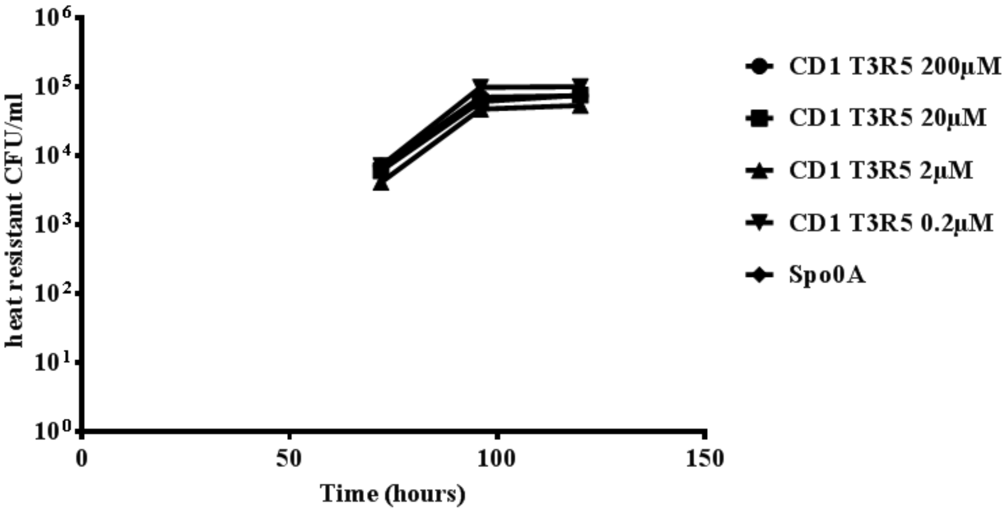


Figure 3.16. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrB2* mutant grown in BHIs containing 20µM of synthetic peptide. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Synthetic peptides used are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

The sporulation assays revealed that none of the concentrations used had any significant effect on the sporulation profile of the *agrB* mutants (Fig 3.17). The AIPs used in this study were not able to restore or modify the phenotype in any of the *agrB* mutants.

a)



b)

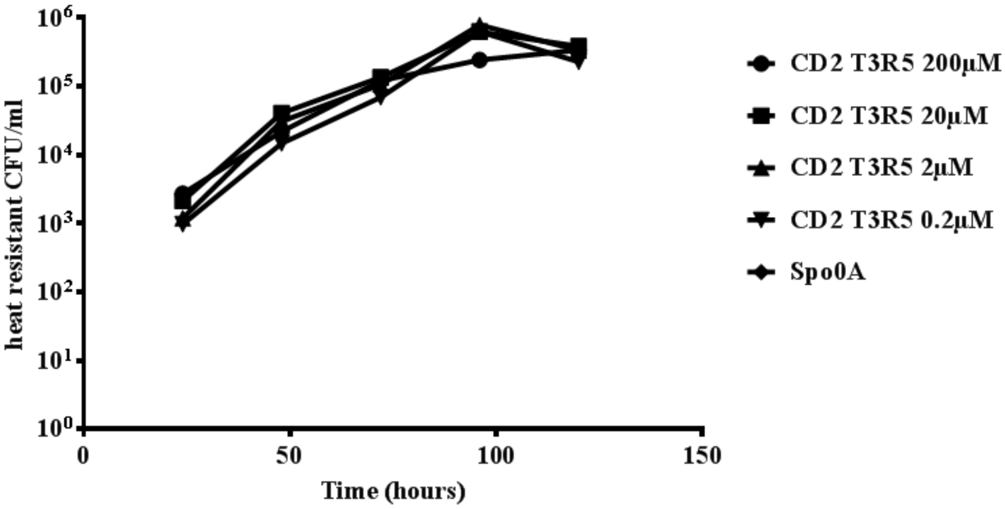


Figure 3.17. Production of heat-resistant CFU in *C. difficile* R20291 NM *agrB* mutants grown in BHIs containing different concentrations of synthetic peptide. The AIP containing a 3-amino acid tail was supplemented in the growth media at different concentrations in a) R20291 NM *agrB1*::CTermB and b) R20291 NM *agrB2*::CTermB. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Synthetic peptides used are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

3.3.5 Use of polyhistidine-tag to isolate DNA bound to the putative AgrA response regulator

The encoded AgrA protein present in *C. difficile* R20291 NM is part of the LytTR family of regulators. These proteins act as response regulators of the microbial two-component signal transduction system and are activated after being phosphorylated by a sensor histidine kinase (Nikolskaya & Galperin, 2002). The LytTR family members share a C-terminal DNA binding domain and an N-terminal receiver domain; these domains are common in all response regulators. In order to identify the DNA domains that interact with AgrA, a polyhistidine-tag consisting of six histidine residues (His-tag) was added to the *agrA* sequence in order to facilitate the isolation of the AgrA protein and the DNA sequence that it binds. The His-tag sequence was added to the C-terminal or the N-terminal of the *agrA* sequence found in *C. difficile* R20291 NM. The tagged sequence was cloned into the vector plasmid pMTL84152 containing the *thl* (thiolase) gene promoter (P_{thl}) for the overexpression of the tagged protein. The resulting plasmids were introduced in *C. difficile* R20291 NM to evaluate the quantity of mRNA being expressed in the strains and the effect of the His-tag on the protein's function.

3.3.5.1 Construction of the mutants

The *agrA* gene found in the R20291 NM strain was amplified using the agrA_NdeI_F1 and agrA_XbaI_R1 pair of primers to introduced an *Nde*I restriction site at the N-terminus and a *Xba*I restriction site at the C-terminus. The *agrA*_HT_NdeI_F and *agrA*_XbaI_R primer pair were used to introduce the His-tag sequence at the N-terminus (HT_*agrA*), whereas the *agrA*_NdeI_F and *agrA*_HT_XbaI_R pair were used to introduce the His-tag at the C-terminus (*agrA*_HT) along with the *Nde*I and *Xba*I restriction sites. The modular plasmid pMTL84152, which contains the *C. acetobutylicum* thiolase gene promoter (P_{thl}) as well as *catP*, was used to clone the fragments generated by PCR. The fragments and the vector were restricted with *Nde*I and *Xba*I to create sticky ends and T4 ligase was used according to the instructions given by the manufacturer to ligate the fragments to the vector, resulting in pMTL84152::*agrA*, pMTL84152::HT_*agrA* and pMTL84152::*agrA*_HT (Fig. 3.18). The plasmids were introduced in *E. coli* CA434 and then transferred into *C. difficile* R20291 NM using conjugation as described in Chapter 2; the resulting strains were designated R20291 NM pMTL84152::*agrA* (native *agrA* found in R20291 NM), *C. difficile* R20291 NM pMTL84152::HT_*agrA* (*agrA* gene displaying the His-tag sequence at the N-terminus) and *C. difficile* R20291 NM pMTL84152::*agrA*_HT strains (*agrA* gene displaying the His-tag sequence at the C-terminus).

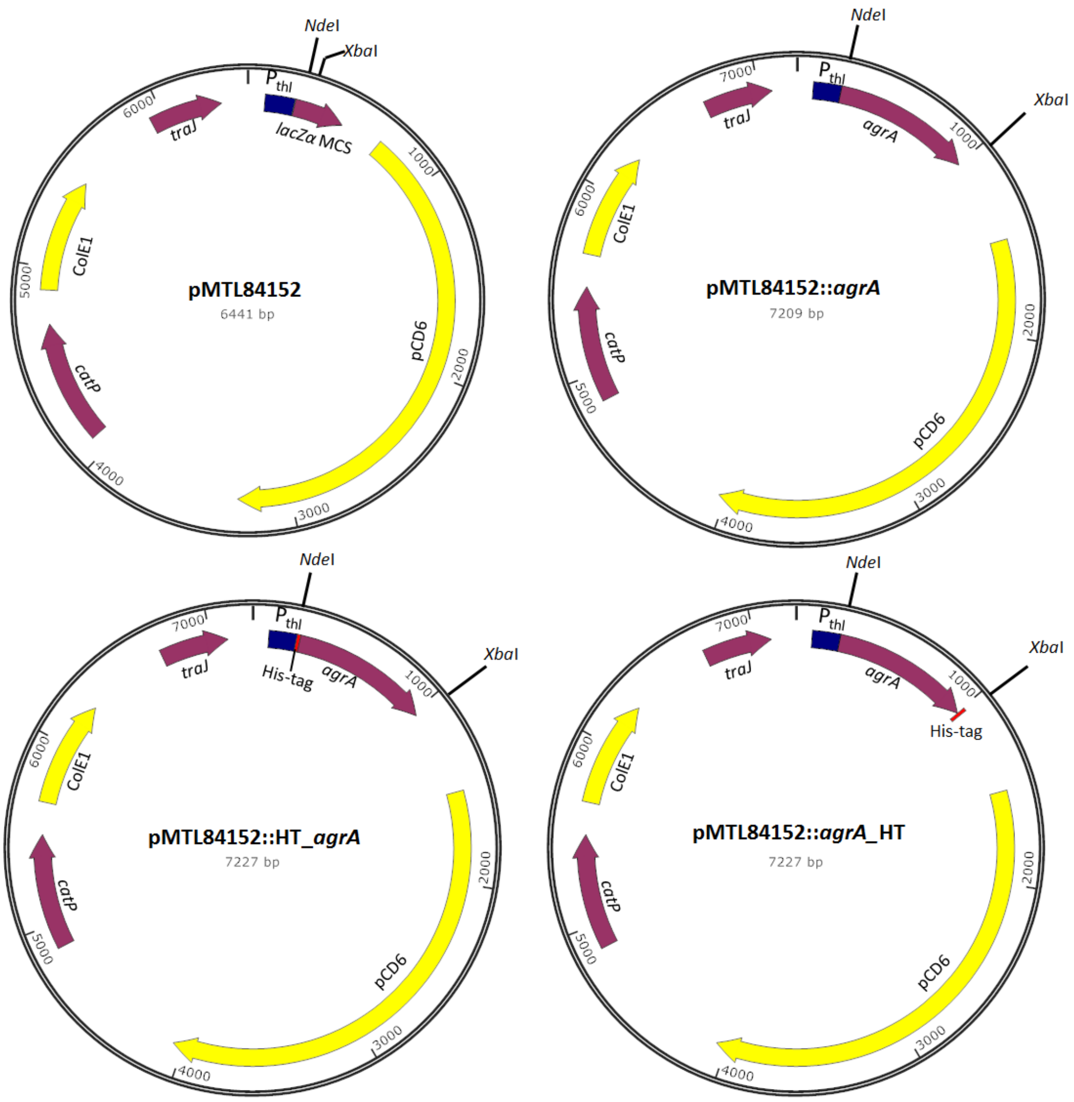


Figure 3.18 Plasmids used for the over-expression of the *agrA* gene containing a six residue polyhistidine-tag (His-tag). Common plasmid components are: P_{thl}, *thl* (thiolase) gene promoter from *C. acetobutilycum; catP* gene of *C. perfringens* conferring thiamphenicol resistance; ColE1, the replication region of the *E. coli* plasmid ColE1, and; *traJ*, transfer function of the RP4 oriT region; pCD6, replication region of *C. difficile* indigenous plasmid pCD6. The modular vector pMTL84152 also harbours *lacZα* MCS, the *lacZ* gene encoding the alpha fragment of the *E. coli* β-galactosidase containing a multiple cloning site, MCS, region derived from plasmid pMTL20. The desired *agrA* fragment (the native *agrA* gene from strain R20291 NM, the *agrA* gene containing the His-tag at the N terminus or the *agrA* gene containing the His-tag at the C terminus) was inserted in the modular vector pMTL84152 using the MCS region.

3.3.5.2 RT-PCR to evaluate the overexpression of the *agrA* sequences from the pMTL84152 plasmid

The overexpression of the *agrA* sequences from the pMTL84152 plasmid was evaluated by measuring the quantity of mRNA present in the cell. Overnight cultures of the mutants were used to inoculate the culture media with a 1/100 inoculum. The cultures were grown anaerobically and, after 8 h, the RNA was extracted as described in Chapter 2. The extracted RNA was subjected to reverse transcription to produce cDNA. The cDNA was used to screen for the *agrA* sequence by real-time PCR using the SYBR green PCR master mix (ThermoFisher) and quantified by the ABI7500 instrument (Applied Biosystems). The housekeeping gene *rpoA* was used as an endogenous control. *C. difficile* R20291 NM was used as control.The *agrA*/*rpoA* ratio revealed that the strains carrying the different plasmids harbouring the *agrA* sequences contain a higher amount of *agrA* mRNA in the cell than that of the housekeeping gene *rpoA*, which suggests that the *agrA* gene is being effectively overexpressed by these strains (Fig. 3.19).

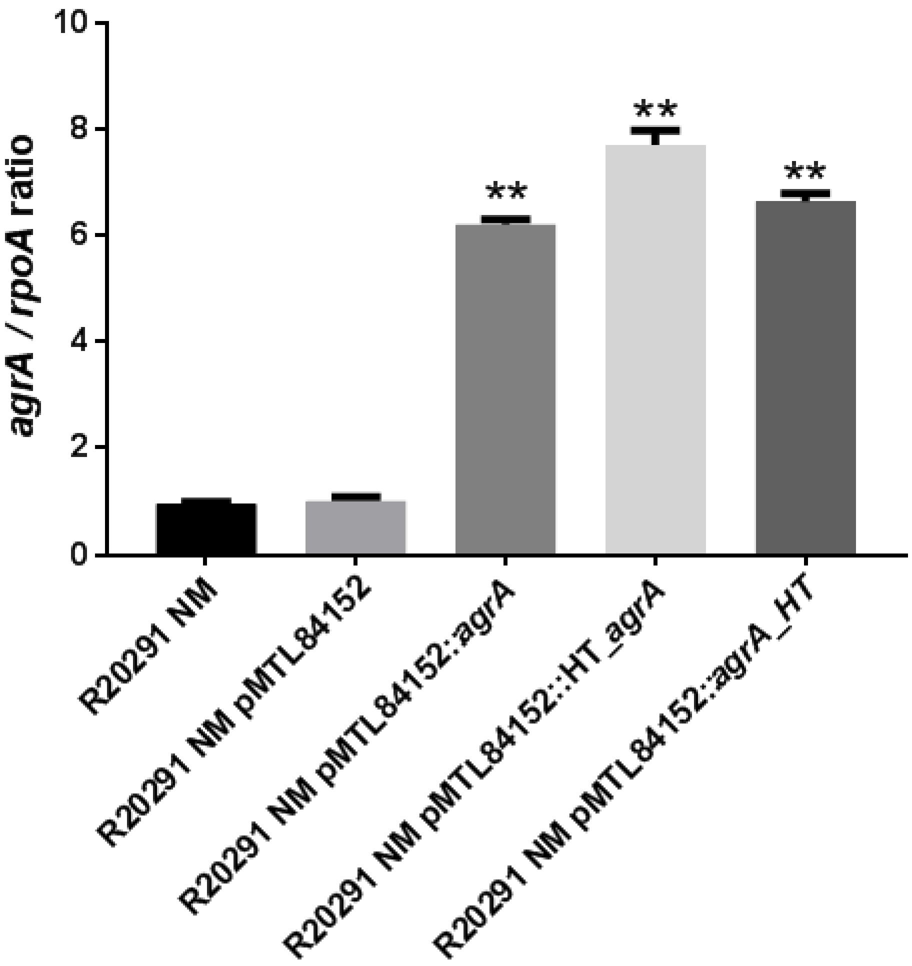


Figure 3.19. RT-PCR of the *agrA* His-tag mutants showing the ratio differences between strains. R20291 NM was included as control. Mutant strains carrying the pMTL84152 plasmid containing the P_{thl} promoter show a higher expression of *agrA* mRNA. All experiments were run in triplicates. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the wild type

It was previously observed that AgrA plays an important role in the toxin production, as the NM *agrA* mutant produces lower toxin levels compared to the parental strain. The overexpression of *agrA* in the cell should lead to an increase in toxin production. In order to see how the overexpression of *agrA* affects the production of toxins in the cell, the strains harbouring the pMTL84152 plasmids containing the *agrA* sequences were test for their ability to produce toxins. The strain carrying the *agrA* gene without the His-tag produced a higher amount of toxins compared to the wild type (Fig. 3.20). Interestingly, the strains carrying the modified *agrA* genes with the His-tag underperformed in this test, being unable to produce the same amount of toxins as the strain carrying the *agrA* without His-tag and also failing to reach the same levels as the wild type. These findings suggest that the presence of the His-tag at either the N- or C-terminus affects the function of the resulting protein, although it is not clear why.

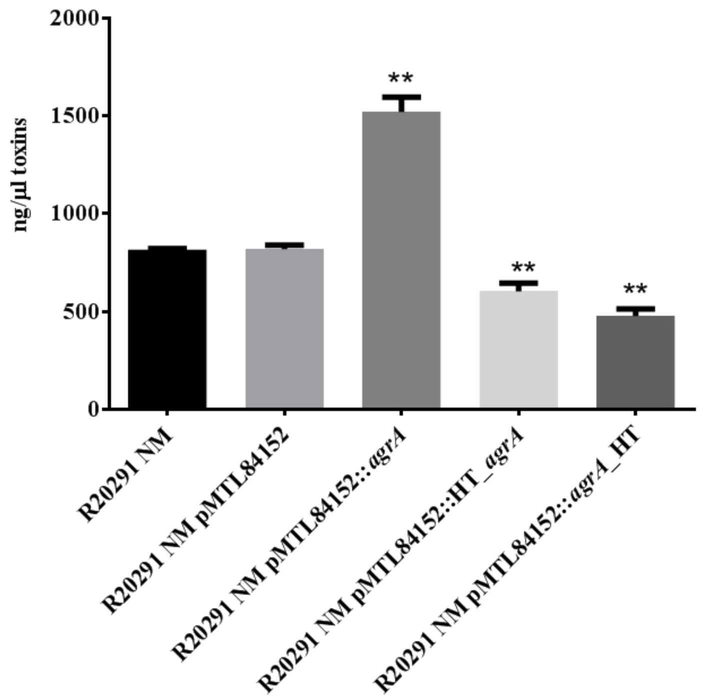


Figure 3.20. Toxin production of *C. difficile* R20291 NM *agrA* His-tag mutant strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's t test) compared to the wild type.

3.3.6 Construction and phenotypical analysis of *agr* mutants in R20291 BW strain.

As previously mentioned, during the course of this study it was revealed that the *C. difficile* isolate used in our laboratory (*C. difficile* R20291 NM) presents several genetic variations unique to this strain (see Chapter 4). As these genetic variations affect greatly the expression of sporulation and toxin production, we concluded that the information drawn from this strain was not reliable. The *C. difficile* strain obtained from Brendan Wren (designated *C. difficile* R20291 BW) at the London School of Hygiene and Tropical Medicine (London, UK) represents an authentic copy of the strain isolated from Stoke Mandeville, as this strain does not present the genetic variations found in *C. difficile* R20291 NM. In light of this, we elected to recreate the *agrB1* and *agrB2* mutants described in *C. difficile* R20291 BW.

3.3.6.1 Construction of the *agr* mutants in R20291 BW

One of the features of R20291 BW is its relatively poor efficiency as a recipient in conjugation experiments with the traditional donor CA434. All too often conjugations have to be repeated several times before transconjugants can be obtained, and then only a handful per plate. Recently, in a parallel project within the group it was shown that conjugation frequencies for *Clostridium* *autoethanogenum* could be improved by over 10-fold by using a modified version of *E. coli* NEB Express in which the conjugative R-factor R702 was introduced from the donor strain usually employed, CA434. This strain is essentially HB101 carrying R702. The new strain was designated NEB Sexpress. Its use with R20291 BW also significantly increased transformation frequencies, allowing ClosTron transconjugants to be routinely obtained. The reason for the improved frequencies are not immediately apparent. The genotypes of the two strains are presented in Table 3.2. Both donors lack type I restriction systems, as well as Mcr and Mrr systems. The latter systems only cleave methylated DNA. R20291 is predicted to carry a single Type I system, a Type II methylase (with no predicted cognate restriction enzyme-encoding gene) and a McrBCP system (http://tools.neb.com/genomes/view.php?view_id=2794). Aside from NEB Express being a K12 strain, and HB101 a type B strain, the major difference between them is that NEB Express is *dcm-*. Presently the reason for the increased frequency is not clear on the basis of RM systems, although such systems may not be the reason for the increased frequency.

Table 3.2 Genotype of *E. coli* strains and the conjugative R-factor used for conjugation with *C. difficile* R20291 BW. Donor strains used in our laboratory for the transmission of vectors to *C. difficile* via conjugation. The strains contain the R-factor R702 (also shown) needed for the transfer of plasmids into *C. difficile.*

Donor	**Genotype**
E. coli HB101	(F - *mcrB mrr hsdS20(rB- mB-)* *recA13 leuB6 ara-14 proA2 lacY1 galK2 xyl-5 mtl-1 rpsL20* (Sm^R) *glnV44* λ-)
E. coli NEB Express	*fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10*--Tet^S*)2 [dcm] R(zgb-210::Tn10*--Tet^S*) endA1 Δ(mcrC-mrr)114::IS10*
R702 (R-factor)	Tc^R , Sm^R, Su^R, Hg^R (Tra^+, Mob^+)

The ClosTron vectors for the insertional inactivation of the *agrB* genes were made using the algorithm at www.clostron.com to retarget the the ClosTron system to the desired gene (Heap *et al*., 2007). The re-targeted plasmids and their insertion site are shown in Table 3.3. The ClosTron plasmids were transformed into the newly derived conjugative donor NEB Sexpress competent *E. coli* (Craig Woods, unpublished) and then conjugated with *C. difficile* R20291 BW for 24 h in anaerobic conditions. The growth was harvested and plated in BHIs supplemented with cefoxitin (8µg/ml), cycloserine (250 µg/ml) and thiamphenicol (15 µg/ml). After 3 days fast growing colonies were restreaked onto the same media for 24 h. The colonies were then transferred to BHIs plates supplemented with lincomycin (20 µg/ml). Lincomycin-resistant colonies were grown in liquid BHIs containing lincomycin. The gDNA from these colonies was extracted and screened looking for the intron. After confirming the presence of the intron, the colonies were plated in BHIs and BHIs supplemented with thiamphenicol to evaluate the loss of the plasmid. Thiamphenicol sensitive colonies were screened for the intron once again.

Table 3.3 ClosTron vectors used for the generation of insertional mutants in *C. difficile* R20291 BW. The pMTL007C-E2 vector harbouring the group II intron was re-targeted to create insertional mutants in the desired genes. The insertion of the group II intron can occur in a sense orientation (s) or an antisense orientation (a).

ClosTron vector	**Target gene**	**Site of insertion in the gene sequence**
pMTL007C-E2::CdiR2029l*agrBl*-362a	*agrB1*	362a
pMTL007C-E2::CdiR2029l*agrBl*-217s	*agrB1*	217s
pMTL007C-E2::CdiR2029l*agrB2*-262s	*agrB2*	262s
pMTL007C-E2::CdiR2029l*agrDl*-92s	*agrD1*	92s
pMTL007C-E2::CdiR2029l*agrD2*-99a	*agrD2*	99a

Initially both *agrB1* and *agrB2* were targeted using the same ClosTrons previously used to make these mutants in R20291 NM. The *C. difficile* R20291 BW *agrB2::*CTermB strain (BW *agrB2* mutant) was readily isolated, with an intron insertion in a sense orientation at position 261/262 bp. Inexplicably, it proved impossible to isolate a mutant of *agrB1*, despite the construction of two different re-targeted ClosTron plasmids (pMTL007C-E2::CdiR2029l*agrBl*-362a and pMTL007C-E2::CdiR2029l*agrBl*-217s). As an alternative, a ClosTron mutant of the *agrD1* (encoding the AIP) gene in which the intron was inserted in a sense orientation at position 261/262 bp was successfully made; this strain was designated *C. difficile* R20291 BW *agrD1::*CTermB (BW *agrD1* mutant). To make data comparable, a ClosTron mutant of *agrD2* in which the intron was inserted in an anti-sense orientation at position 98/99 bp was also isolated. The resulting strain was designated *C. difficile* R20291 BW *agrD2::*CTermB (BW *agrD2* mutant). The insertions were verified by PCR (Fig. 3.21); for the insertional mutants with a sense orientation insertion, BW *agrD1* mutant and BW *agr*B2 mutant, the forward flanking primer of the gene (*agrD1*_F1 and *agrB2*_F1, respectively) and the EBS universal primer were used to generate an intergenic fragment between the gene and the group II intron. For the BW *agrD2* mutant, the reverse flanking primer (*agrD2*_R1) and the EBS universal primer were used to generate the intergenic fragment. This fragment can only be amplified if the intron is present in the gene. The expected sequence of each fragment was confirmed by Sanger sequencing. The frequency of insertion of all mutants is shown in Table 3.4.

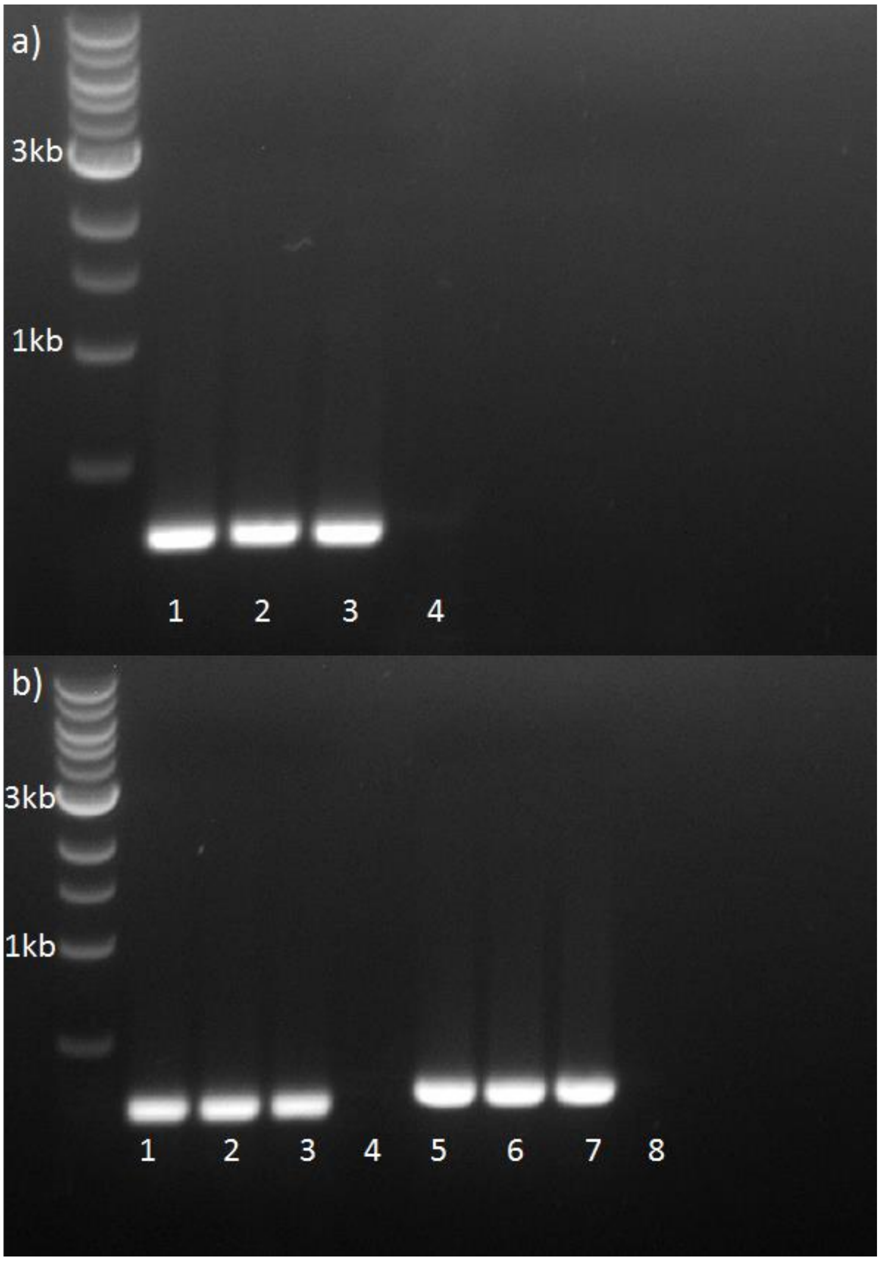


Figure 3.21 Example of PCR screening to confirm the presence of the intron in the genome of the *C. difficile* R20291 BW strain for the *agrB2, agrD1* and *agrD2* mutants. a) Lincomycin-resistant colonies harbouring pMTL007C-E2::CdiR2029l*agrB2*-262s (lane 1-3) and the parental strain R20291 BW (lane 4) were tested with agrB2_F1 + EBS universal primers to detect the insertion of the intron in the *agrB2* gene. b) Lincomycin-resistant colonies harbouring pMTL007C-E2::CdiR2029l*agrDl*-92s (lane 1-3) and the parental strain R20291 BW (lane 4) were tested with agrD1_F1 + EBS universal primers to detect insertion of the intron in the *agrD1* gene. Lincomycin-resistant colonies harbouring pMTL007C-E2::CdiR2029l*agrD2*-99a (lane 5-7) and the parental strain R20291 BW (lane 8) were tested with agrD2_R1 + EBS universal primers to detect insertion of the intron in the *agrD2* gene. At the extreme left, a 1KB DNA ladder is used as a weight marker.

Table 3.4 ClosTron insertion frequency with lincomycin selection. Frequency of the intron in the colonies screened by PCR with their respective flanking primers and the EBS universal primer.

Strain/target site	**Frequency of mutant among clones screened**	
	Percentage (%)	No. screened/No. positive
C. difficile **R20291 BW/** ***agrB1*** **217s**	0	30+/0
C. difficile **R20291 BW/** ***agrB1*** **362a**	0	20/0
C. difficile **R20291 BW/** ***agrB2*** **262s**	90	10/9
C. difficile **R20291 BW/** ***agrD1*** **92s**	100	8/8
C. difficile **R20291 BW/** ***agrD2*** **99a**	100	8/8

The BW *agr*B2 and *agrD2* mutants were complemented with the parental *agr*2 locus using the pMTL84151::*agr2* plasmid while the BW *agrD1* mutant was complemented with the pMTL84151::*agrBD1* used to complement the *agr*B1 mutant constructed in the R20291 NM strain.

3.3.6.1 Sporulation in *C. difficile* R20291 BW *agr* mutants

The *agr* mutants constructed in the R20291 BW strain were tested for their ability to form heat-resistant spores over time in BHIs containing taurocholate, as described in Chapter 2. The sporulation profile of the BW *agr* mutants is presented in Fig. 3.22. The BW *agr*B2 and *agrD*2 mutants show a similar behaviour as the parental strain, both in terms of the onset of sporulation at 24 h and the amount of spores produced over time. In contrast, the BW *agrD1* mutant exhibited a significant delay in the onset of sporulation as spores were only detected after 72 h. Although the appearance of spores in the BW *agrD1* mutant is delayed compared to the other strains, the final number of spores attain the same level as those seen in the other strains at the end of the assay.

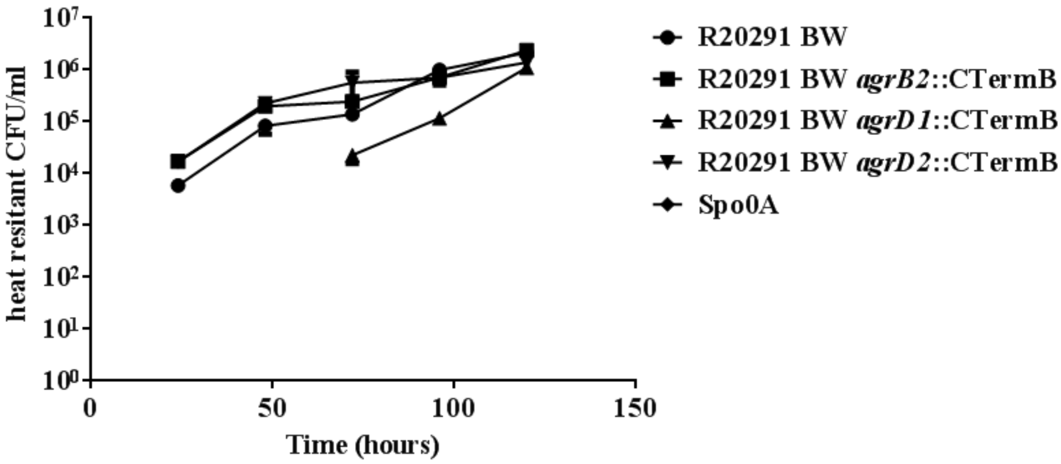


Figure 3.22. Production of heat-resistant CFU in *C. difficile* R20291 BW *agr* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

In order to see if the discrepancy in the onset of sporulation of the BW *agrD1* mutant could be restored to the wild type, the complementation plasmid pMTL84151::*agrBD1* was introduced by conjugation and the sporulation properties of the resultant transconjugant were tested as described in Chapter 2. In parallel, a transconjugant carrying the vector only control (pMTL84151) was also prepared. The sporulation profile of the R20291 BW *agrD1* mutants and control strains are presented below (Fig. 3.23).

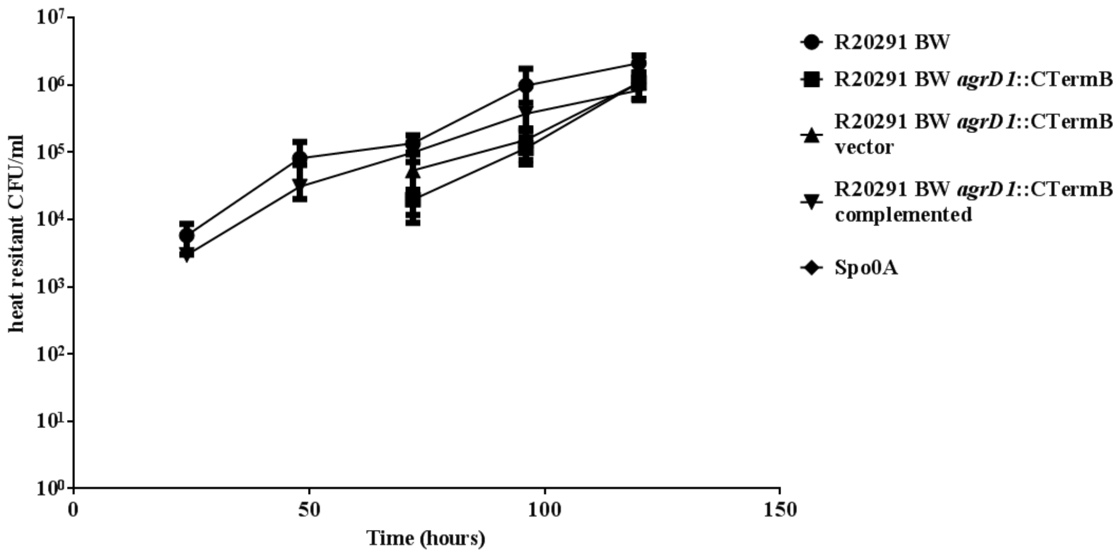


Figure 3.23. Production of heat-resistant CFU in *C. difficile* R20291 BW *agrD1* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and the complemented strain refers to pMTL84151::*BD1*. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

The R20291 BW *agrD1*::CTermB strain containing the empty vector showed a similar behaviour to the wild type with the onset of sporulation at the beginning at the 72 h time point. The strain harbouring the complementation vector was able to start the sporulation process earlier than the R20291 BW *agrD1* mutant, as spores were detected after only 24 h, matching the time of sporulation initiation seen in the R20291 BW wild type strain.

The presence of the complemented plasmid in the R20291 BW *agrB2* and *agrD2* mutants did not have any effect on their sporulation profiles, as these remained unchanged throughout the assay (Fig 3.24).

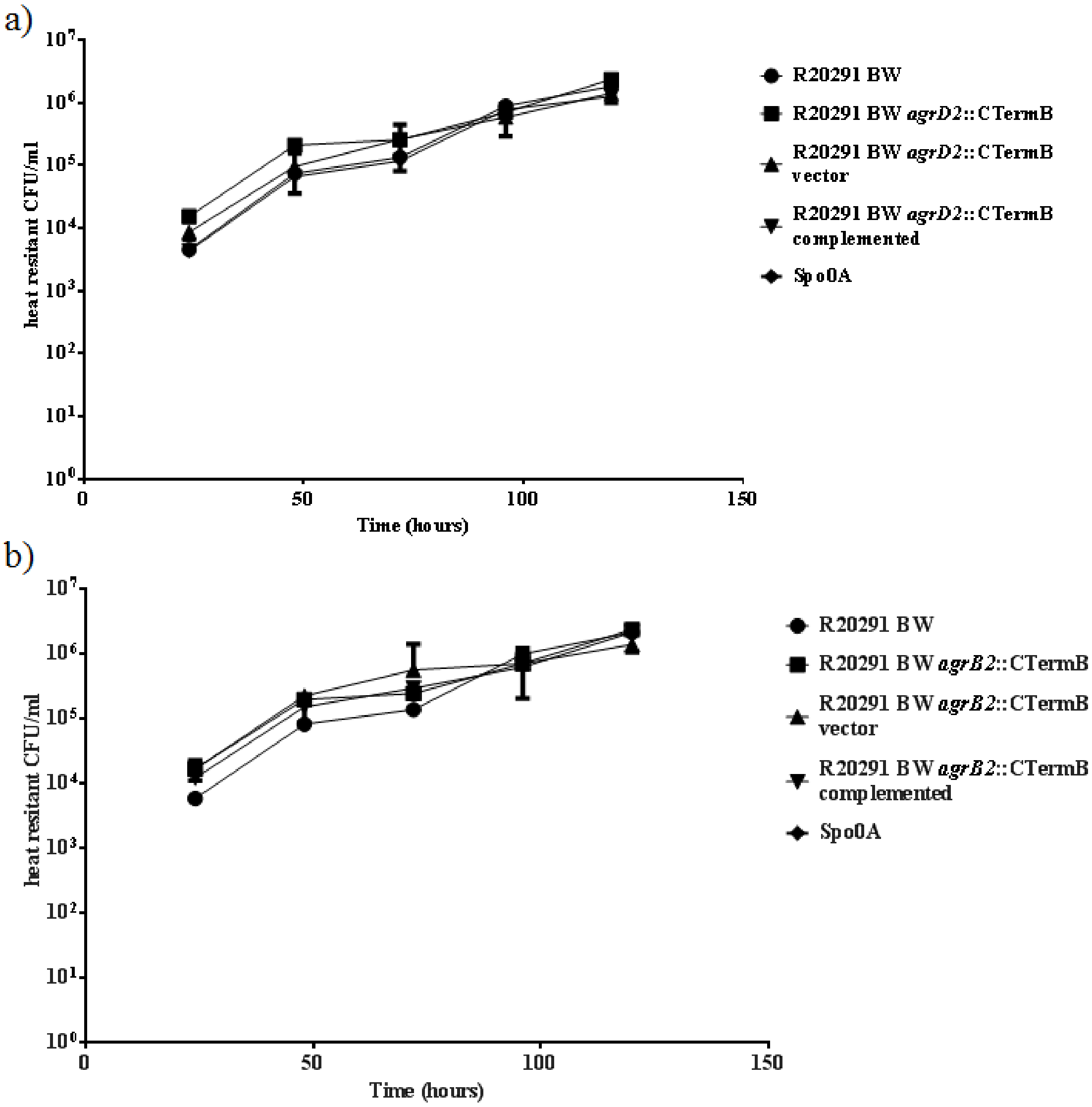


Figure 3.24 Production of heat-resistant CFU in *C. difficile* R20291 BW *agrD2* and *agrB2* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. pMTL84151 has been abbreviated as vector and the complemented strain refers to pMTL84151::*agr2*. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

3.3.6.2 Toxin production in *C. difficile* R20291 BW *agr* mutants

The *agr* mutants constructed in the R20291 BW strain were tested for their ability to form toxins over time (Fig. 3.25). The BW *agrD1* mutant had a reduction in the amount of toxin produce after 72 h, while the BW *agrD1* mutant strain carrying the complementation vector showed a slight increase in toxin levels (Fig. 3.25a). The slight increase in toxin production seen in the complementation strain compared to the parental strain could be a consequence of the presence of the *agrBD1* genes at a high copy, being plasmid-borne, leading to elevated levels of D1 protein in the cell. The data obtained in this experiment suggest that the *agrD1* gene is involve in the positive regulation of toxins. Interestingly, even though the interruption of the *agrD2* and *agrB2* genes had no effect in the sporulation process, the R20291 BW *agrD2* and *agr*B2 mutants produced greater amounts of toxin than the parental strain (Fig 3.25b and Fig 3.25c respectively). The increased levels of toxin in these mutant was restored to wild type amounts following the introduction of the complementation vector. These results suggest that the *agrB2* and *agrD2* genes have a negative effect in the regulation of toxins in R20291.

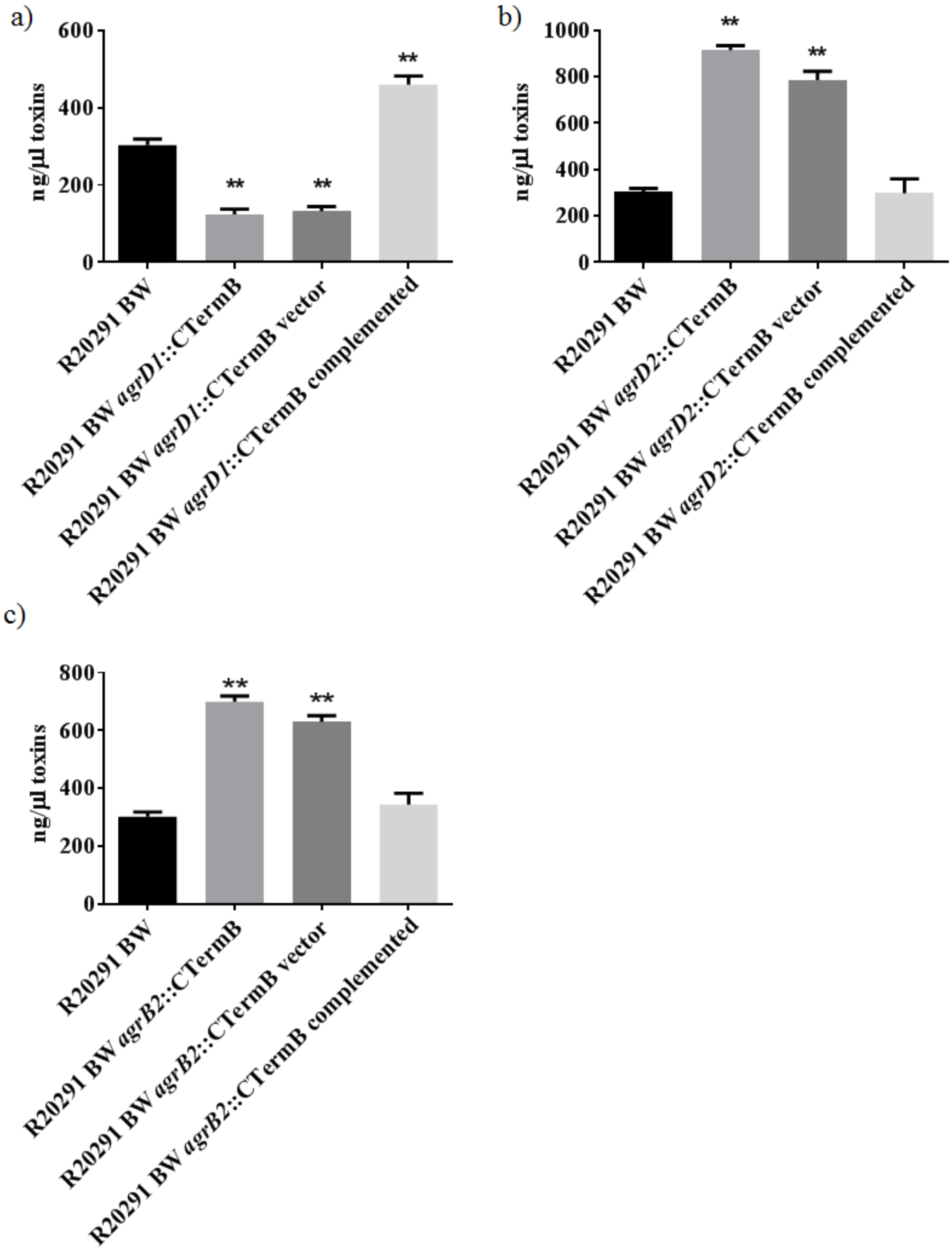
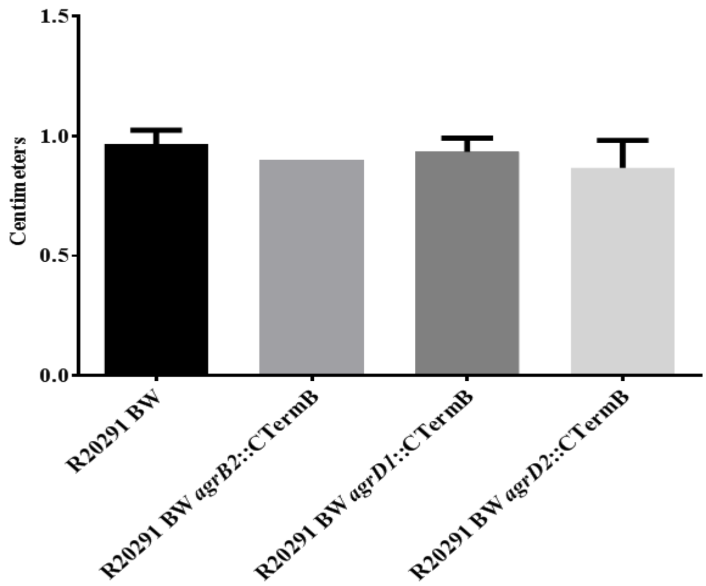


Figure 3.25 Toxin production of the R20291 BW *agr* mutants after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. pMTL84151 has been abbreviated as vector and the complemented strain refers to the strain harbouring pMTL84151::*BD1* for the BW *agrB1* mutant and pMTL84151::*agr2* for the BW *agrB2* and *agrD2* mutants. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the wild type.

3.3.6.3 Motility and biofilm production of the *C. difficile* R20291 BW *agr* mutants

As seen in the study published by Martin *et al.* (2013), the *agrA* mutant produced in the R20291 BW strain had reduced motility compare to the parental strain due to a reduction in the number of flagella around the cell. Here we evaluate the swimming and swarming ability of the R20291 BW *agrD1, agrD2* and *agrB2* mutants in semi-solid media as described in Chapter 2. No significant difference in the swimming or swarming ability was seen in any of the mutant strains compared to the parental R20291 BW strain (Fig 3.26). Transmission electronic Microscopy (TEM) was used to visualise the number of flagella around the cell. In summary, 20µl of mid-exponential culture was added to a copper grid and the cells were fixed for 1 min using 1% glutaraldehyde, washed with PBS and negatively stained with 4% uranyl acetate for 30 s. The grid was air-dried and visualised using a Tecnai G2 12 Bio Twin microscope with a fully embedded digital imaging system with SIS Megaview III camera (Advance Microscopy Unit, Queen's Medical Centre, University of Nottingham, United Kingdom). The TEM images show no significant difference in the number of flagella around the cell in the different strains (Fig. 3.27). The results suggest that, contrary to the observations made in the *agrA* mutant created by Martin *et al.* (2013), the *agrB2, agrD1* and *agrD2* genes play no role in the regulation of flagella in *C. difficile* R20291 BW.

a)



b)

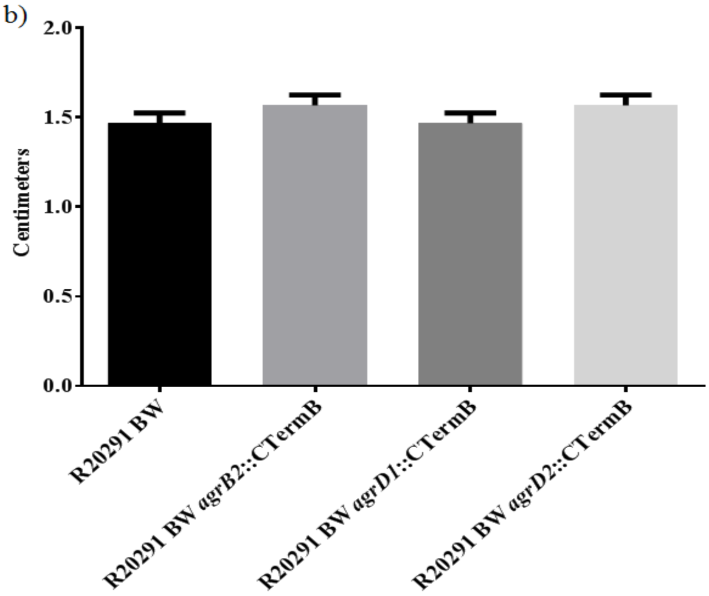


Figure 3.26 Motility of R20291 BW *agr* mutant strains. a) Swarming ability of the R20291 strains in BHIs 0.4% agar plates. b) Swimming assay in BHIs 0.3% agar plate. After 48 h under anaerobic conditions the halo formed by the growth of the strain was measured. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means.

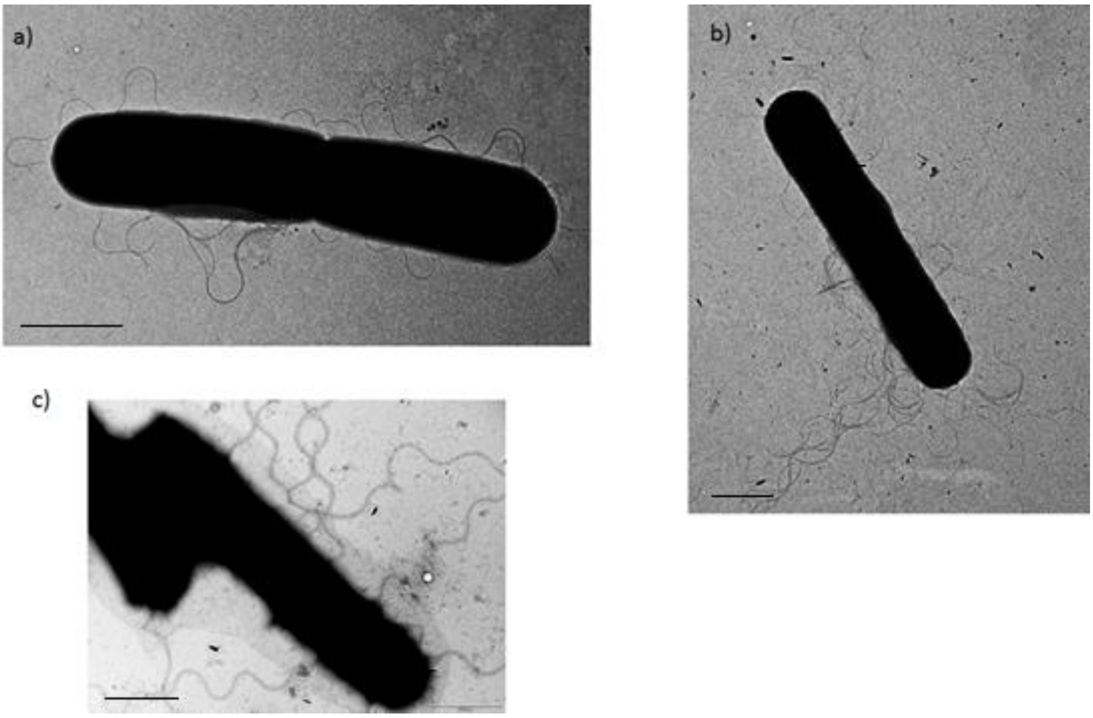


Figure 3.27 Transmission Electronic Microscopy images of the *agr* mutants generated in *C. difficile* R20291 BW. Transmission Electronic Microscopy was used to visualise the flagella present in the *C. difficile* R20291 BW *agr* mutants. a) R20291 BW *agrD1* mutant. b) R20291 BW *agrD2* mutant. c) R20291 BW *agrB2* mutant. Scale bar, 1µm.

The biofilm production by the R20291 BW *agr* mutants after 5 days was tested according to the methodology described in Chapter 2. The production of biofilm showed no significant difference in all the R20291 BW *agr* mutants compared to the parental strain (Fig. 3.28). These results revealed that the *C. difficile* R20291 BW *agr* mutants tested here had no influence in the regulation of biofilm formation after 5 days.

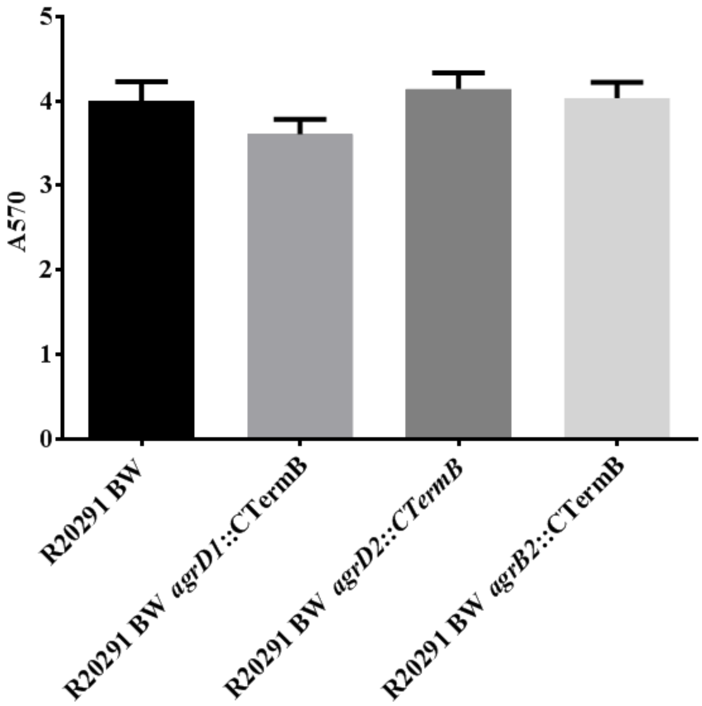


Figure 3.28 Biofilm production of the R20291 BW *agr* mutant strains after 5 days. The strains were grown in BHIs supplemented with 0.1M glucose. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means.

3.4 Discussion

In recent times, the appearance of epidemic strains of *C. difficile* has become a world-wide problem. These epidemic strains cause more severe disease and increase mortality, as well as being linked to the emergence of community-associated CDAD. The BI/NAP1/027 group has become one of the prevalent epidemic strains across Europe, being reported in 16 countries across the continent (McDonald *et al.,* 2005). The PCR-Ribotype 027 (RT-027) strains have a series of genes not present in other strains, several of which are regulatory genes (STabler *et al.*, 2009). It has been described that several RT-027 strains are capable of producing a higher amount of toxins *in vivo* and are responsible for a higher mortality rate (Boone *et al*., 2014). These strains carry a number of genes not present in other strains, including an extra copy of the *agr* locus, designated *agr2*, which play an important role in the regulation of virulence factors in other bacteria. In this study, the role of the extra complete *agr* locus, *agr*2, in the regulation of sporulation and toxin production in two *C. difficile* R20291 strains was evaluated.

The first strain of *C. difficile* sequenced, designated strain 630, lacks 281 genes present in R20291, including the additional complete *agr*2 locus (STabler *et al.*, 2009). The 630 strain used in our laboratory has been designated *C. difficile* 630Δ*erm*. The 630 strain is a non-epidemic strain that lacks the *agr2* operon found in the PCR-Ribotype 027 strains. In order to evaluate if the presence of the *agr*2 locus had any effect in sporulation or toxin production in *C. difficile* 630Δ*erm*, the pMTL84151 vector plasmid, based on the modular system for *Clostridium* shuttle plasmids published by Heap *et al.* (2009), carrying the complete *agr*2 locus was introduced in the *C. difficile* 630Δ*erm* strain. Contrary to what we expected, no difference in the sporulation or toxin production of the strain was observed in the mutant strain carrying the extra *agr*2 locus compared to the wild type. The extraction of RNA from this strain and subsequent RT-qPCR analysis showed that the *agr2* locus was being expressed in the cell, confirming that the lack of difference between the mutant and the parental strain is not due to the lack of expression of the *agr2* operon in the cell. One explanation of these results could be that some of the 281 genes that are present in R20291 but not in 630 are also required to produce the same levels of toxins, as many of these genes are two component regulatory systems and transcriptional regulators (STabler *et al.*, 2009) that could be key elements in the pathway regulated by the *agr2* operon. In a recent article published by Darkoh *et al.* (2016), an *agrBD1* deletion mutant created by allelic exchange in 630 showed a lack of toxins in the supernatant. This indicates that the *agr1* locus play an important role in the regulation of toxins in 630 but it is not clear if both genes are responsible for this change in the toxin production or if only one of these genes is responsible for the phenotype seen in the mutant as a previously *agrB1* mutant generated in our laboratory showed no significant difference in the production of toxins compared to the parental strain (Patel *et al.*, unpublished).

The so called hypervirulent RT-027 strain *C. difficile* R20291 contains the incomplete *agrBD1* locus and an extra complete *agr2* locus within its genome (STabler *et al.*, 2009). We hypothesised that the *agr* loci play a more important role in the regulation in this strain. Insertional mutants created in the R20291 strain used in our laboratory, designated R20291 NM, were phenotypically characterised. The results observed in the R20291 NM *agrB* mutants suggest that these genes play an important role in the regulation of toxins and sporulation, as the interruption of any of these genes greatly affected the production of toxins and spores. In *S. aureus*, the *agrB* gene codes for a transmembrane protein involved in the processing of the pro-peptide encoded by *agrD* and it is group specific (Zhang *et al*., 2004). The interruption of the *agrB* genes should stop the processing of the respective pro-peptide, consequently interrupting the pathway regulated by the AIPs. At first look, the results seen in the sporulation assay suggest that the AgrB1 protein is involved in the production of the AIP that promotes the onset of sporulation, presumably AgrD1, as the mutant lacking this gene has a delayed initiation in the production of spores, while the AgrB2 protein could be involved in the production of a repressor, possibly AgrD2, based on the behaviour seen in the *agrD2* mutant, where the sporulation starts earlier compared to the wild type. The predicted structure of the AIPs are based on homologous regions that resemble the AIPs found in *S. aureus* (Fig. 3.11). An important difference between the *agrD1* and *agrD2* genes is that the AgrD2 sequence does not contain the conserved cysteine residue present in the predicted sequences of the AIPs from other clostridia species, as it is replaced with a serine amino acid. This change seems to be important due to the AIPs found in *S. aureus* requiring the thiol group found in the cysteine residue to cyclize and form the thiolactone peptide, the active form of the AgrD protein (Lyon & Novick, 2004).

The toxin profile of the mutants also show that the insertional interruption of the R20291 NM *agrB* genes has an effect on toxin production. Indeed, the effect seen in the toxin production is the inverse of that seen in the onset of sporulation. It has been reported that Spo0A, the master regulator of sporulation, is able to repress toxin production in the RT-027 strains (Deakin *et al.,* 2012; Mackin *et al.,* 2013). The difference in the onset of sporulation could explain the difference in toxin production, as the R20291 NM *agrB1* mutant with a delayed sporulation time has more time to produce toxins before Spo0A is activated, whereas the *agrB2* mutant shows lower toxin production due to the earlier onset of sporulation. The negative effect of the vector and the complementation plasmids in toxin production of the R20291 NM *agrB1* and *agrB2* mutants prevented us from evaluating the restoration of toxin production in these strains. It is uncertain why this effect was only seen in the *agrB1* and *agrB2* mutants but not in the R20291 NM *agrA* mutant or the wild type carrying the vector. From internal communication within our group, it was revealed that this phenomena has been observed in other insertional mutants harbouring the pMTL84151 plasmid containing the complementation sequence. One method that could be used to evaluate if the restoration of the gene reverts the toxin phenotype could be by allelic exchange, removing the insertion and restoring the native gene in the genome.

In *C. perfringens,* the use of supernatants taken from the wild type was able to restore the impaired sporulation phenotype seen in *agr* mutants (Ohtani *et al.,* 2009). We intended to test the ability to restore, repress or promote the onset of sporulation by using the supernatants from the *C. difficile* R20291 NM parental strain and the NM *agrB* mutants, following our initial hypothesis that the *agrB1* mutant lack of AIP1 but presence of AIP2 in the supernatant would be able to restore the phenotype seen in the *agrB2* mutant and vice versa, as well as evaluate the effect of this supernatant in the phenotype of the parental strain and the own mutant. Contrary to expectations, the only supernatant that was able to cause any effect on the strains was the supernatant from the NM *agrB2* mutant; as *agrB1* is not interrupted in this strain, the AIP produce by AgrB1 should be present in the media. The supernatant promoted the earlier onset of sporulation in all the strains, supporting our suspicion that the AgrB1 produced AIP was related to the positive regulation of sporulation. The supernatant from the 630Δ*erm* strain also show the same phenotype, reinforcing the idea that the AIP1 produce by AgrB1 is the responsible for the phenotype seen in the strains, as this strain does not harbour the *agr2* locus, and thus it plays an important role in the sporulation process. The role of the AIP produce by AgrB2 remains unclear as it is yet to be stablish if this AIP2 has an active role in the regulation of spores and toxins in the cell or if it is solely the lack of AIP1 that is responsible for the phenotype seen in the *agrB1* mutant.

It has been shown in *C. acetobutylicum* that the addition of exogenous synthetic peptides resembling the predicted sequence of the AIP restores the impaired sporulation phenotype observed in an *agrB* mutant (Steiner *et al.,* 2012). Accordingly, synthetic peptides derived from the predicted sequences of the AgrD proteins found in R20291 were tested to evaluate their ability to revert the sporulation phenotypes seen in the NM *agrB* mutants. The predicted sequence of the AgrD AIP has 3 residues in the tail, but different tail lengths were used to evaluate the effect of the tail residues in the strains. Unfortunately none of the peptides used in this study was able to significantly effect on the sporulation of the *agrB* mutants. Different concentrations of the 3-residue tail AIP were used to observe if the concentration was the reason why the other AIPs failed to have any significant effect in the mutants, but the results confirmed that none of the concentrations used was able to modify the phenotype of the *agrB* mutants. A study regarding the regulation of toxins via QS also confirmed that the addition of exogenous predicted AIP with different tail length in the media has no effect in the production of toxins (Darkoh *et al*., 2015). One possible reason why the peptides used in this study had no effect on the strains used is that the *in vivo* AIP might undergo different modifications other than just the cycling of the amino acid chain in order to become the active AIP that is able to be detected by the cell; however, this remain to be described.

AgrA is a protein that belongs to the LytR family response regulator. In *S. aureus*, AgrA interacts with the intergenic region containing the promoters for the *agr* locus and RNA III, a regulatory RNA effector molecule, positively regulating the expression of both regions (Koening *et al.*, 2004). Contrary to what we observed in the R20291 NM *agrB* mutants, where toxins and spores were affected by the insertional mutants, the R20291 NM *agrA* mutant showed reduced toxin production but no differences in the sporulation assay. This results suggest that the AgrA protein is involved in the regulation of toxins independently of the regulation of spores. In the *agr* model described in *S. aureus,* the *agr* operon works as a single unit with all of the components forming part of the same pathway (Asad & Opal, 2008). However, this does not seem to be the case for the AgrA protein in R20291 NM. If the AgrA protein followed the predicted *agr* model, it would be expected that the *agrA* mutant should display a phenotype resembling that seen in the *agrB2* mutant, as these genes form part of the same operon. This opens up the possibility of the *agr2* operon following a different pathway than the one described in *S. aureus*.

The R20291 NM *agrA* gene was tagged with a polyhistidine tail and overexpressed it in the cell in an effort to facilitate the isolation of this protein and to use it in pull down experiments to identify the DNA region that it interacts with, i.e. promoters and genes under the control of *agrA.* The overexpression of the mRNA was confirmed using RT-qPCR and it was observed that the strain carrying the native R20291 NM *agrA* gene and the strains carrying the his tag versions had an increased amount of *agrA* mRNA. The toxin assay also confirmed that the strain carrying the native *agrA* gene had an increased toxin production compared to the wild type, reinforcing the idea of the AgrA protein playing an important role in the regulation pathway for toxins. However, the strains carrying the tagged *agrA* gene showed lower amount of toxins compared to the wild type. The unexpected results seen in the tagged proteins suggest that the presence of the His-tag has a detrimental effect on AgrA function. It is not only unclear why these strains are not only incapable of producing a higher amount of toxins, as seen in the strain carrying the native protein, but why toxin production is affected and unable to match the parental strain. It is possible that the presence of the His-tag affects the interaction of the protein with the target region and thus the regulation of toxins. This outcome meant that using purified His-tag proteins in pull-down experiments to identify promoter regions under AgrA control was unlikely to succeed. Therefore, this avenue of research was abandoned. In the future, if this approach was to be revisited, then other affinity tag methods could be explored to get the required protein, such as ChIP-seq (Parker, 2009).

A comparative RNA sequencing study published by the London School of Hygiene and Tropical Medicine between their R20291 isolate (R20291 BW) and its respective *agrA* mutant showed down regulation of the *tcdA* (Toxin A) gene and a corresponding reduction in the amount of toxin produced (Martin *et al.,* 2013), agreeing with the phenotype seen with the R20291 NM *agrA* mutant made in the present study. Furthermore, they observed down regulation of FliC, the major structural unit of the flagella (Barketi-Klai *et al.,* 2014). Transmission Electronic Microscopy (TEM) images revealed that the parental R20291 strain used in this study possessed several flagella around the cell, while the *agrA* mutant exhibited a total lack of flagella. This observation is in stark contrast with observations made in this laboratory, which clearly shown that the Nottingham R20291 strain (R20291 NM) displays a single polar flagellum (Baban *et al.,* 2013). This discrepancy set us to investigate the differences between the two strains. Both were obtained from the Anaerobe Reference Centre in Cardiff, UK. The genetic variations found are described in further detail in Chapter 4. In essence, the R20291 Nottingham strain (R20291 NM) contains 5 unique single nucleotide polymorphisms (SNP) located in different regions, affecting a series of genes, including the anti-sigma factor RsbW. The regulation of toxins and spores in R20291 NM strain are significantly affected and do not correlate to the phenotype seen in the R20291 BW strain (from LSTMH), confirming that the SNPs found in our strain greatly affect the regulation of toxins and spores. This also means that the conclusions about toxin and spore regulation based in our R20291 NM strain are not reliable, as these virulence factors do not seem to follow the same regulation pathway due to the SNPs. The disparities found in our strain were big enough to make us reconsider working with our strain and prompted us to remake and analyse mutants in strain R20291 BW, obtained from Brendan Wren at the London School of Hygiene and Tropical Medicine (London, UK).

Once we obtained the R20291 BW strain that does not present the genetic variations found in R20291 NM, it was the intention to make the same *agrB1* and *agrB2* mutants that had been previously made in R20291 NM. Inexplicably, the equivalent mutants of *agrB1* could not be made, despite designing two different retargeted ClosTron plasmids; lincomycin resistant colonies were isolated from the cultures but the expected fragment between the flanking primer and the EBS primer could not been amplified, meaning that the presence of the group II intron in the gene could not be confirmed. From communication within our group, it was revealed that this phenomena has occurred with the ClosTron system when targeting other genes, although it is not clear why the insertion in the intended gene cannot be confirmed through PCR. As an alternative, *agrD1* was targeted together with *agrD2* in order to make the comparisons between the two *agr* systems equivalent. In the study published by Darkoh *et al.* (2016), it was reported that they failed to create an *agrD1* mutant through allelic exchange due to the small size of the sequence. However, the ClosTron methodology relies on base-pairing between the target DNA and the intron RNA (Heap *et al.,* 2010) and thus, the size of the gene did not represented an inconvenience, as both genes, *agrD1* and *agrD2,* were successfully interrupted. In many respects, targeting the gene encoding the actual AIP is likely to be more informative than inactivating the gene encoding the AgrB processing enzyme, given that the possibility that cross-talk could be occurring, i.e., AgrB1 could also process AgrD2. We successfully obtained both the *agrD1* and *agrD2* mutants in the *C. difficile* R20291 BW strain, as we as the *agrB2* mutant..

The *agrD1* mutant created in the R20291 BW strain showed a reduction in the toxin levels detected as well as having a delay in the onset of sporulation compared to the parental strain. This contrasted with the toxin phenotype observed in the *agrB1* mutant created in the R20291 NM strain, where the toxin production was increased; however, the onset of sporulation was also delayed in this strain. Following the *S. aureus agr* model, the AgrD1 pro-peptide needs to be cleaved and cycled by AgrB1 in order to create the mature AIP. This would suggest that insertional mutants created in *agrB1* and *agrD1* should present a similar phenotype, as the interruption of any of these genes would prevent the formation of the mature AIP. In a study published by Darkoh *et al.* (2016), the effect of the deletion of the *agrBD1* locus in toxins was evaluated. The toxin profile of the *agrBD1* deletion mutant showed a lack of detecTable toxin production in the cultures. Although we did not see a total lack of toxins in the *agrD1* mutant created in strain R20291 BW, both mutants show that the interruption of the *agrBD1* locus negatively affects the expression of toxins, contrasting with the positive regulation of toxins observed in the *agrB1* mutant created in strain R20291 NM. It has been suggested that the AgrD1 protein is the pre-cursor for a signalling AIP capable of inducing toxin production, as AgrD1 presents the conserved cysteine residue needed for the cyclization of the pro-peptide in order to create the mature AIP, and that the lack of this signalling AIP causes an impaired toxin phenotype (Darkoh *et al.*, 2016). These results imply that the toxin phenotype observed in the BW *agrB1* mutant created in strain R20291 NM is anomalous, suggesting that the production of toxins in *C. difficile* R20291 NM is either regulated by a different regulatory pathway or the normal pathway is in some way impaired.

The *agrB2* and *agrD2* insertional mutants created in strain R20291 BW showed a similar behaviour concerning toxin production, both producing a higher toxin level compared to the parental strain. This is in agreement with the behaviour of the *agrBD2* deletion mutant created by Darkoh *et al.* (2016), where the *agrBD2* mutant was shown to exhibit elevated transcription of the *tcdA* and *tcdB* genes. The sporulation assay did not reveal any difference between the *agrB2* and *agrD2* mutants created in R20291 BW and the wild type. The phenotype of the *agrB2* mutant created in R20291 BW contrast greatly with that of the *agrB2* mutant created in R20291 NM. The R20291 NM *agrB2* mutant has a lower production of toxins and a higher sporulation rate detected after 24 h compared to the parental strain. In contrast, the R20291 BW *agrB2* and *agrD2* mutants exhibit no difference in sporulation and produce higher levels of toxins compared to the parental strain. The differences observed between the *agr* mutants created in R20291 NM and R20291 BW are substantial and confirm that the regulation of virulence factors in *C. difficile* R20291 NM is distinct from that of R20291 BW. In view of this anomaly, further work on R20291 NM was judged as inadvisable.

It has been demonstrated that the *agrA* gene plays an important role in the regulation of flagella in the cell (Martin *et al.,* 2013). However, none of the *agr* mutants produced in the R20291 BW strain showed differences in their motility, the number of flagella or biofilm production. It is an interesting finding that these *agr* members do not affect the production of flagella as the *agr* model found in *S. aureus* would suggest that the *agr2* operon acts as a single unit and thus, the *agrB2* and *agrD2* insertional mutants should present the same phenotype. These findings suggest that AgrA and possibly AgrC might be connected to another regulation pathway that is not directly influenced by *agrBD2* or *agrD1*.

The present study has shown that the *agrD1* gene plays an important role in the regulation of toxins and spores and that AgrD1 is the probable signalling peptide that controls the expression of some virulence factors. The role of the *agr2* operon remains unclear. The evidence suggests that the *agrB2* and *agrD2* are involved in the regulation of toxin production, although they do not appear to be directly affected by the *agrA* gene, a finding that is contrary to the *S. aureus agr* model. The extraction and subsequent comparative analysis of RNA between the mutants and the parental strains could provide more useful information about the genes that are affected and to assess the role of the *agr2* operon in the cell.

Chapter 4

Genetic variations in R20291 strains

4.1 Introduction

The division of a bacterial cell generally produces two daughter cells that are identical to the parent cell, although occasionally a spontaneous genetic change can occur in one of the daughter cells. Since the genetic change is heriTable, a single mutant cell, provided that the mutation does not affect the fitness of the cell, is capable of producing a large population of identical cells (clone) that possesses different characteristics compared to the parent cell. Spontaneous genetic variations can be classified as DNA rearrangements, Insertions and deletions (Indels) or Single Nucleotide Polymorphisms (SNP) (Hedrick, 2011).The DNA rearrangement produces changes in the genetic structure due to insertions, deletions or inversions. If the insertion or deletion of DNA bases into a gene is not divisible by 3, it will alter the reading frame of the sequence from the point of the mutation to the end of the gene, causing a frameshift mutation (Brown, 2007). A frameshift mutation invariably leads to the creation of a premature stop codon, which results in the end of translation and the production of a truncated protein. An in-frame mutation occurs when the insertion or deletion is divisible by 3; in this type of mutations the reading frame of the gene is not altered, although new codons are added to, or deleted from, the sequence of the protein, compromising the structure and folding of the resulting protein.

Single Nucleotide Polymorphisms (SNPs) are random DNA variations in which 1 base pair is changed to another nucleotide different from the reference strain

(Kimchi-Safarty *et al.*, 2007). The SNPs can occur in the coding region of a gene or in an intergenic region. The change of a base pair in the coding region of a protein can lead to a synonymous or a non-synonymous SNPs; a synonymous change codes for the same amino acid and does not alter the resulting protein, whereas a non-synonymous SNP alters the amino acid sequence of the protein, which may compromise its function. Non-synonymous SNPs are divided into missense and nonsense SNPs. The missense SNP change results in a codon that result in the incorporation of a different amino acid into the translated protein, whereas the nonsense SNP introduces a stop codon or a nonsense codon into the mRNA, resulting in a truncated incomplete protein. These changes in the DNA could affect not only the sequence of the protein and its function, but also the transcription, mRNA transport and translation (Kahali *et al.,* 2007).

A study conducted by the London School of Hygiene and Tropical Medicine (Martin *et al.*, 2013) published images taken with transmission electron microscopy (TEM) of their *C. difficile* R20291 wild type and the *C. difficile* R20291 *agrA* mutant created using this strain. The TEM images taken revealed that the R20291 strain used (R20191 BW) possessed multiple flagella. This information contrasted with previous TEM images taken of the R20291 strain used at Nottingham (strain R20291 NM) in which it can be observed that the R20291 cells have a single polar flagellum (Soza *et al.,* 2013). The discrepancy in the number of flagella in the R20291 strains from different laboratories could indicate that the strains possess SNPs or Indels that have resulted in different phenotypes. It has been observed in several bacteria that the presence of SNPs in the genome can affect the virulence of the cell; for instance, the presence of SNPs in 4 regulatory genes in a clinical isolate of *V. cholerae* increases the virulence of the strain (Krukonis *et al.*, 2000). It has also been observed that a SNP mutation in the *mexT* gene in *P. aeruginosa* leads to a more aggressive phenotype compared to the wild type (Olivas *et al.,* 2012). These examples highlight the importance of determining whether a particular strain being used has the expected genome sequence and does not contain additional mutations that could be contributing to an observed phenotype.

4.1.1 Aim of this chapter

Spontaneous genetic changes in a cell can lead to dramatic changes in phenotype. Colonies containing random mutations can be isolated unknowingly when the mutation is not apparent or easily identifiable. The study presented in this chapter sets out to investigate the possible differences found in the R20291 strains from different laboratories. In this chapter, the R20291 strain from our laboratory in the Centre for Bimolecular Sciences, Nottingham (NM), the R20291 strain used in the London school of Hygiene and Tropical medicine laboratory (BW) and the R20291 strain obtained from the Novartis laboratory in Sienna, Italy (Novartis) are compared in order to assess if there are significant phenotypic and genomic differences between them.

4.2 Results

4.2.1 Growth differences in R20291 strains

The bacterial growth curve of the R20291 strains were evaluated over 24 h in BHIs media. Additionally, the strains were also grown in minimal media containing 1% (w/v) glucose, fructose or mannitol in order to evaluate their ability to use these sugars. In short, an overnight culture was used to inoculate the culture media with a 1/100 inoculum. The culture media was incubated under anaerobic conditions over 24 h, reading the O.D.600 every hour.

The behaviour of the R20291 strains in BHIs broth is similar, although the R20291 NM strain reached a lower optical density at 600nm (OD_{600}) compared to the other strains in this culture media (Fig. 4.1a). Surprisingly, the R20291 NM strain had a better performance in minimal media containing the different sugars compared to the BW and Novartis strain (Fig. 4.1b, Fig. 4.1c and Fig. 4.1d). The behaviour of the BW and Novartis strains was similar throughout all the assays, with little OD_{600} variations between them.

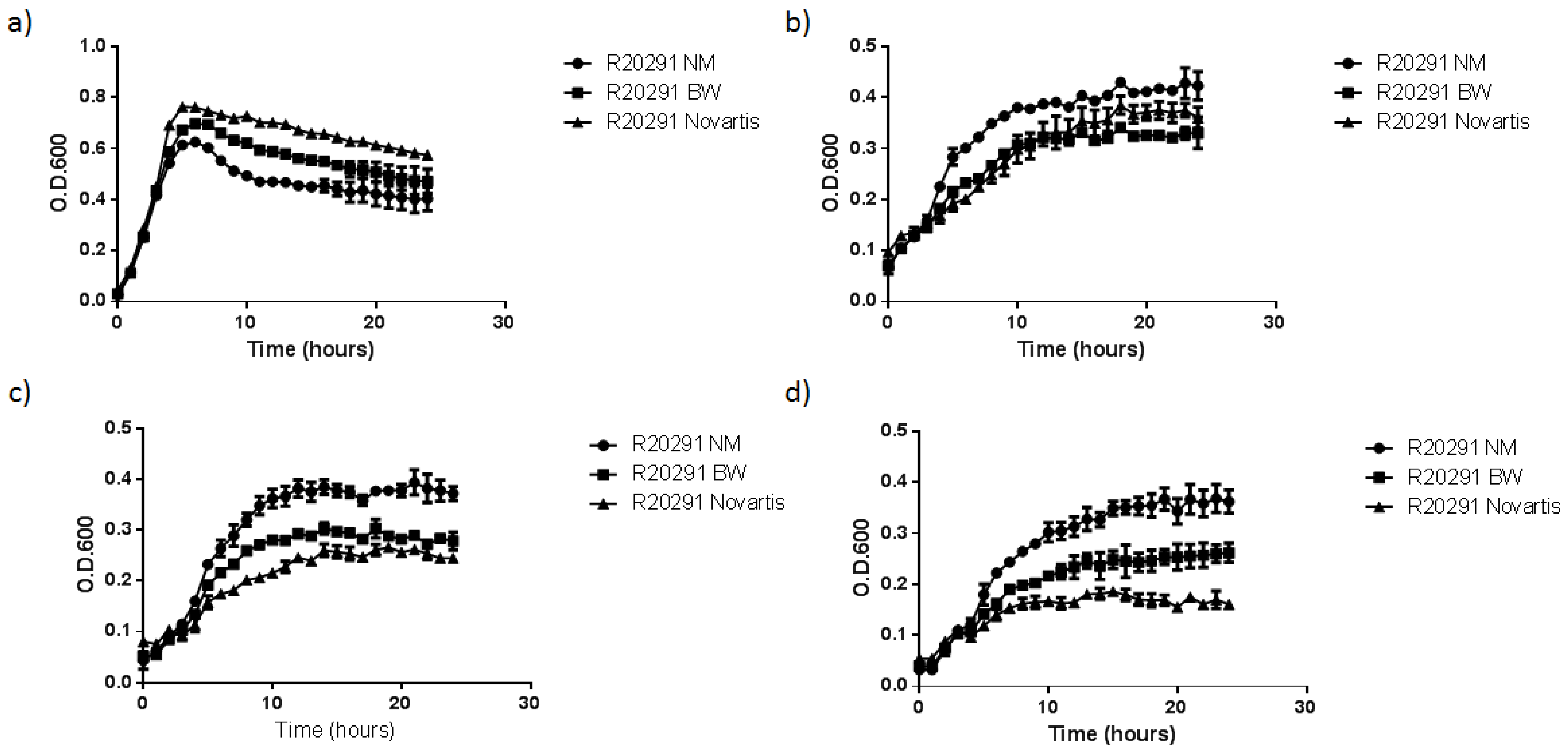


Figure 4.1. Growth of R20291 strains in different culture media over 24 hours. The strains were grown in a) BHIs broth, b) minimal media with 1% (w/v) glucose, c) minimal media with 1% (w/v) fructose and d) minimal media with 1% (w/v) mannitol. Strain were cultivated as described in chapter 2. The OD_{600} was measured every hour. The symbols represent the averages of three independent experiments, and the error bars represent standard errors of the means.

4.2.2 Motility assays and transmission electron microscopy of the R20291 strains.

In order to evaluate the motility of the strains, their swarming and swimming ability were tested. Briefly, the strains were assessed for their swarming ability by adding 5µl of an overnight culture to a 0.4% BHIs agar plate. The plates were incubated for 48 h under anaerobic conditions and, after the incubation period, the halo formed by the strains was measured. The swimming assay was performed by stabbing the middle of a BHIs 0.3% agar plate with a single colony taken with a toothpick. The plates were incubated for 48 h and after the incubation period, the diameter of the halo formed was measured. The results of the swimming assay revealed that the R20291 NM strain had a slightly impaired motility compared to the other two strains (Fig. 4.2). The performance of the R20291 BW and R20291 Novartis strain was similar in both assays (Fig. 4.3), suggesting that these strains are similar and the R20291 NM strain is the only one that has reduced motility.

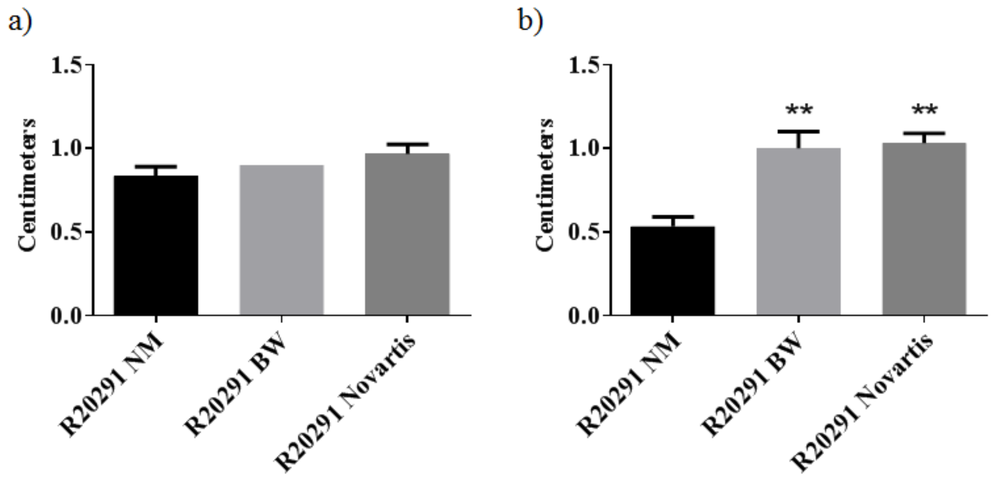


Figure 4.2. Motility of R20291 strains. a) Swarming ability of the R20291 strains in BHIs 0.4% agar plates. b) Swimming assay in BHIs 0.3% agar plate. After 48 h under anaerobic conditions, the diameter of the halo formed by the growth of the strains was measured. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's t test) compared to the R20291 NM strain.

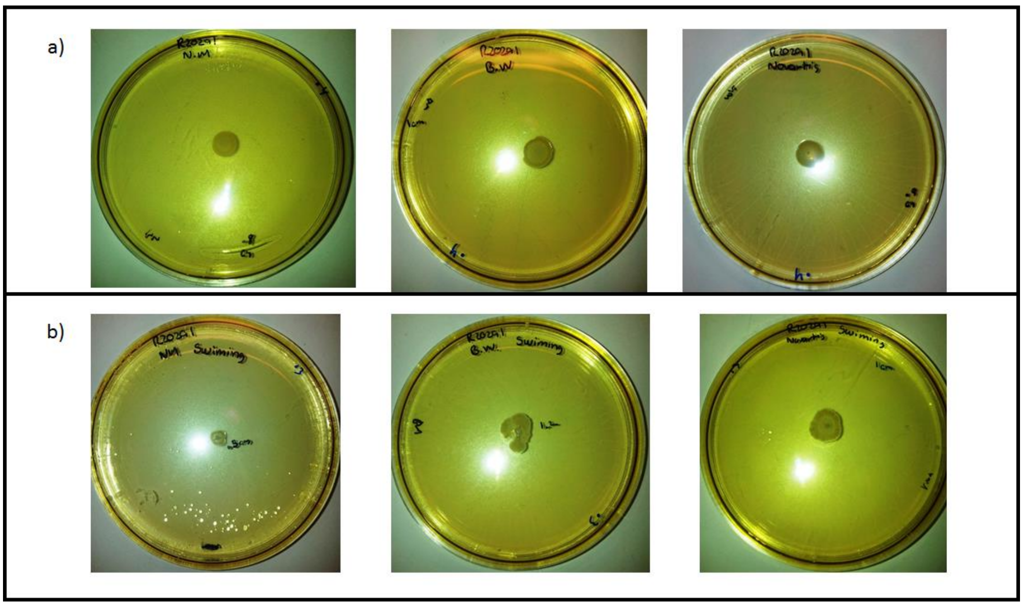


Figure 4.3. Pictures of the R20291 motility assays in BHIs plates. Plates showing the halo formed by the R20291 strains in the a) swarming assay and the b) swimming assay.

Transmission electron microscopy (TEM) was used to visualise the flagella in the different R20291 strains. In summary, 20µl of mid-exponential culture was added to a copper grid and the cells were fixed for 1 min using 1% glutaraldehyde, washed with PBS and negatively stained with 4% uranyl acetate for 30 s. The grid was air-dried and visualised using a Tecnai G2 12 Bio Twin microscope with a fully embedded digital imaging system with SIS Megaview III camera (Advance Microscopy Unit, Queen's Medical Centre, University of Nottingham, United Kingdom). The R20291 BW and the R20291 Novartis strains possessed multiple flagella (Fig. 4.4b and Fig. 4.4c respectively), while the R20291 NM strain showed a single polar flagellum (Fig. 4.4a). The difference of motility seen in the swarming and swimming assay could be explained by the different flagella number in the strains. As the phenotype of the R20291 NM strain is different from the other two strains, it is important to determine if there are any other phenotypical differences between them.

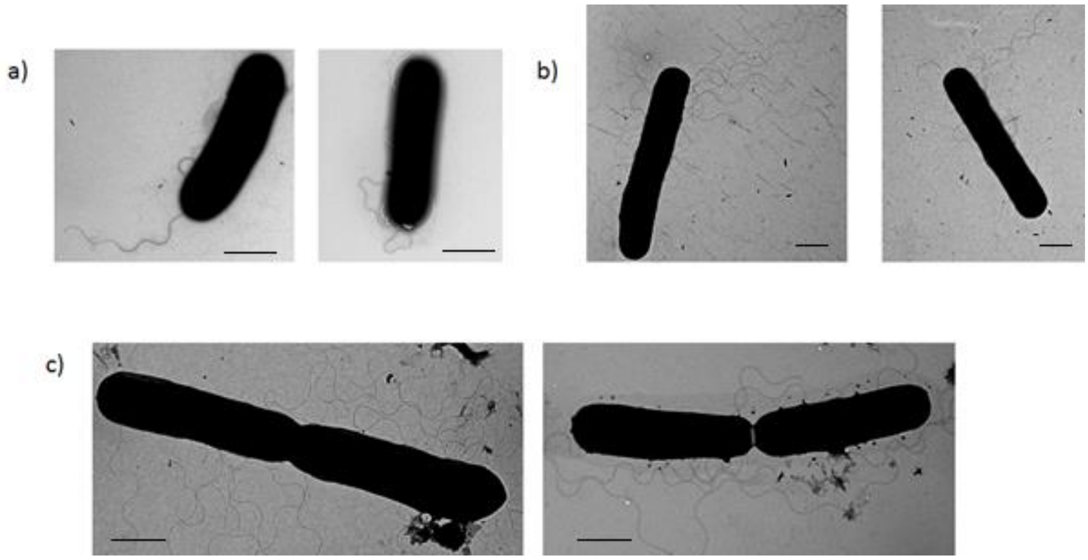


Figure 4.4. Transmission electron microscopy images of the R20291 strains. The presence of flagella around the cell was evaluated with the use of transmission electron microscopy. Images showing the flagella present in a) R20291 NM strain, b) R20291 BW strain and c) R20291 Novartis strain. Scale bar, 1µm.

4.2.3 Toxin production in R20291 strains

It has been reported that the flagellar operon in *C. difficile* R20291 plays an important role in the regulation of toxins. In a study conducted by Baban *et al.* (2013) it was observed that insertional mutants in *fliC* and *fliD* had an increased expression of the *tcdA* gene. The difference in the number of flagella seen in the different R20291 strains could also mean that they may be affected in toxin production. The levels of toxin being produced by the different isolates was evaluated using an ELISA method. Briefly, overnight cultures were prepared in TY medium to allow the production of toxins and they were used to prepare starter cultures as previously described. The starter cultures were used to inoculate TY broth with a 1/100 inoculum. After 72 h under anaerobic conditions, the cultures were centrifuged and the supernatant was filter-sterilised to remove the cells. The supernatant was tested using the *C. difficile* TOX A/B ELISA (TECHLAB) kit.

The R20291 NM strain was able to produce more than twice the quantity of toxins detected in the other strains (Fig. 4.5). Interestingly, the R20291 BW and the R20291 Novartis strains had a similar toxin profile. These results show that the NM strain possesses a different behaviour compared to the other two strains. The difference in toxicity might be related to the difference in the number of flagella seen in the TEM images.

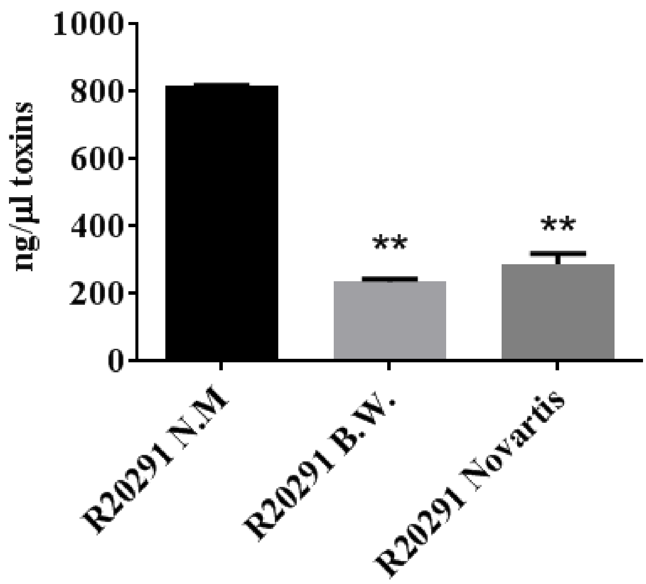


Figure 4.5. Toxin production of the R20291 strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the R20291 NM strain.

4.2.4 Sporulation in R20291 strains

As observed in the *agrB* mutants created in the R20291 NM strain, the regulation pathways of sporulation and toxin production are interconnected, although is not fully understood yet how. After seeing the difference in the toxin profile among the strains, their ability to produce heat-resistant CFU was evaluated. The strains were tested for their ability to form heat-resistant CFU over 5 days, as described in Chapter 2. In short, overnight cultures were used to inoculate starter cultures with a 1/100 inoculum. The starter cultures were grown under anaerobic conditions for 4 h and a 1/100 inoculum was added to the culture media. 1 ml samples were taken every 24 h over 5 days. The samples were heat treated at 65°C for 30 min and spotted onto BHIs agar plates supplemented with taurocholate in order to promote the germination of spores. The BW and Novartis strains were able to produce spores at the 24 h time point (Fig. 4.6). As seen before, in the NM strain spores are only detected after 48 h. The number of spores produced by the NM strain after 48 h were lower than the number of spores produced by the other 2 strains after 24 h. However, the final count of heat-resistant CFU after 5 days was similar between all three strains, despite the delay in the initiation of sporulation in the R20291 NM strain compared to the other 2 strains. The data obtained in the sporulation assay implies that the regulation of sporulation in the R20291 NM strain is affected, although is not yet clear what is the cause of these differences.

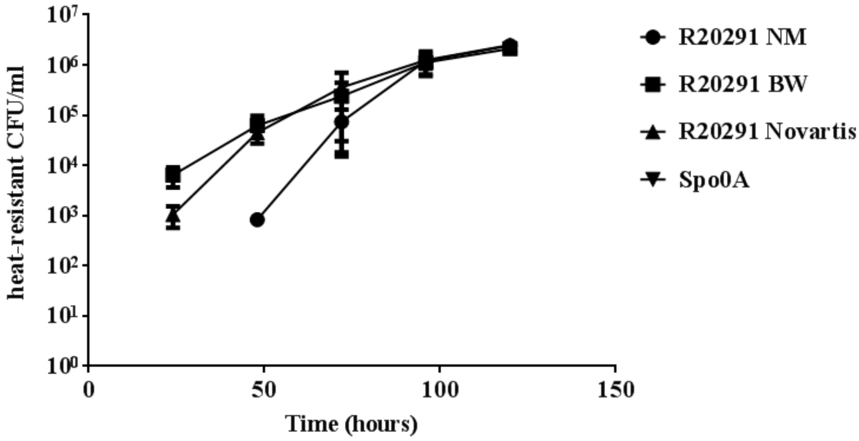


Figure 4.6. Sporulation profile of the R20291 strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

4.2.5 Biofilm formation in R20291 strains

The results from the sporulation and the toxin assays show that the R20291 NM strain behaves in a different manner compared to the BW and Novartis strains. This indicates that there are differences in the regulation of virulence factors among these strains. The production of biofilm, another factor that is believed to be involved in the pathogenesis of *C. difficile,* was assessed in order to determine if any differences existed between strains.

Briefly, overnight cultures were used to inoculate the growth media (BHIs broth containing 0.1M glucose) with a 1/100 inoculum. Tissue culture plates were used to help the biofilm adhere to the bottom of the plate. After 5 days under anaerobic conditions, the wells were washed with PBS and stained with a 0.2% crystal violet solution for 30 min. The wells were washed with PBS and the stained biofilm was washed with methanol in order to extract the dye from the biofilm. Finally, the extracted dye was quantified at A570.

After 5 days, the R20291 NM strain showed a higher quantity of dye extracted from the biofilm compared to the BW and Novartis strains, indicating that the ability to form biofilms in this strain was higher than either of the two other strains (Fig. 4.7). The BW and Novartis strains presented similar levels of extracted dye, being unable to produce the same amount of biofilm compared to R20291 NM strain.

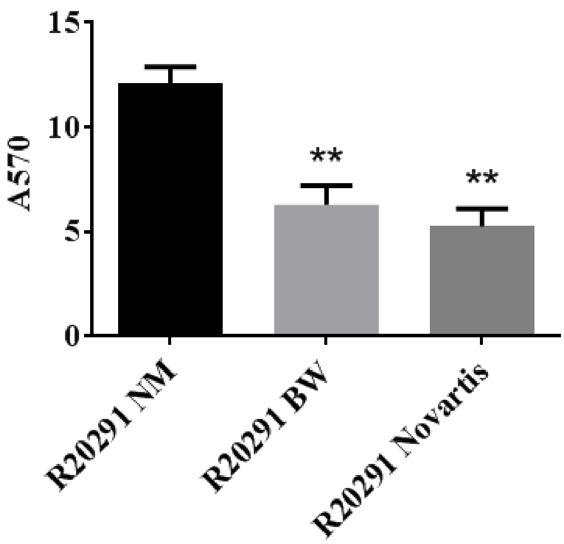


Figure 4.7. Biofilm production of the R20291 strains after 5 days. The strains were grown in BHIs broth supplemented with 0.1M glucose. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, P < 0.01 by Student's t test) compared to the wild type.

4.2.6 Whole genome analysis of the R20291 strains.

The phenotypic characterisation of the R20291 strains revealed that the R20291 NM strain possesses different characteristics in terms of toxin production, the initiation of sporulation, biofilm formation and motility. These differences clearly show that the strains that were thought to be similar have important differences. Whole genome sequencing has been used to evaluate the presence of mutations in other clostridia species (Ehsaan *et al.,* 2016; Collery *et al.,* 2016). In order to observe the genetic differences between strains, the gDNA from the different R20291 strains was sent for whole genome sequencing (Deepseq, University of Nottingham). Preparation of paired-end libraries as well as sequencing was performed as described by the manufacturers. In short, the gDNA from the different strains was obtained using phenol/ chloroform/isoamyl alcohol extraction as described in the section 2.2.1.2. gDNA concentrations were measured with Qubit dsDNA BR assay kit (Thermofisher Scientific). DNA integrity was assessed with conventional gel electrophoresis. Standard fragment sequencing libraries were prepared according to Illumina TrueSeq PCR free library preparation kit (Illumina). 2 µg of gDNA was fragmented to 400-600bp using Covaris S2 sonicator. Fragmented DNA was purified using Agencourt Ampure XP beads (Beckman Coulter) at 1:1.6 DNA:Bead ratio and then end repaired at 30°C for 30 min. End repaired DNA was size-selected using Ampure XP beads at the desired fragment length. Illumina sequencing Y adapters were ligated at 30°C for 10 min followed by incubation at 70°C for 5 min. Unligated adapters were removed with further 2 rounds of purification using Ampure XP beads at 1:1 DNA:Bead ratio. Purified adapter ligated DNA was quantified by qPCR (Kapa biosystems). Equimolar library pool was loaded according to manufacturer's instruction on the MiSeq sequencing platform. Sequencing was done using Illumina MiSeq benchtop sequencer (Illumina). The resulting files were mapped to the reference R20291 genome published by the Sanger institute (GenBank *accession* no. *FN545816*) using CLC Genomics workbench. After the comparison with the reference genome, we defined an alternative allele as a minor call with an observed count or frequency greater than that of a set cut-off value. A frequency of 90% was used as a cut-off for the single nucleotide variation (SNV) calling, in accordance to the method used to identify SNV in other clostridia used in our laboratory (Ehsaan *et al.*, 2016; Collery *et al.*, 2016; Sheng *et al.*, 2016).

The comparative genomic analysis revealed a series of polymorphisms in the strains (Table 4.1). Some of these polymorphisms were shared among the strains in comparison to the reference R20291 genome. These differences are likely to be errors made in the original genome sequencing. Nevertheless, polymorphisms that are strain specific were found. The R20291 Novartis strain shows 3 unique SNVs. Two of these SNVs are present in the coding regions of *codY*, a GTP-sensing transcriptional pleiotropic repressor, and *CD*R20291_*2781*, a putative GntR-family transcriptional regulator, while the third is located in an intergenic region.

The R20291 NM strain presents 5 unique polymorphisms, 3 deletions and 2 SNV, 4 of which are located in coding regions. The deletion mutations in the coding regions are present in the *rbsK* gene, a ribokinase belonging to the *pfkB* family, as well as *CD*R20291_*2456*, a putative two-component sensor histidine kinase. The deletions lead to the translation of incomplete proteins. The SNPs found in *rsbW*, an anti-sigma B factor, and *vncR*, a two-component response regulator Tn1549-like, produce missense mutations, changing the sequence of the resulting protein.

The genomic differences observed in the R20291 strains confirm that there are variations in the genome of the strains. These genetic variations affecting regulatory factors (i.e., anti-sigma factors, transcriptional regulators and kinases) have been shown to have a great impact in the phenotypic characteristics of the strains. The variations found in the genomes are presented in Table 4.1.

Table 4.1. Polymorphisms found in the R20291 strains NM, BW and Novartis.

Reference Position [a]	Type [b]	Reference Allele [c]	Mutant Allele [d]	Coverage [e]	Frequency [f] (%)	Gene [g]	Coding Region Change [h]	Amino Acid Change [i]
9694	SNP	G	T	152	99.34211	*rsbW*	c.245G>T	Gly82Val
132924	Insertion	-	A	20	100			
132939	SNP	G	T	15	100	CDR20291_t40		
132955	SNP	C	A	23	100	CDR20291_t40		
132958	SNP	T	G	18	100	CDR20291_t40		
132959	SNP	T	C	18	100	CDR20291_t40		
143465	Insertion	-	A	82	97.56098	16S RNA		
206403	Insertion	-	A	216	97.22222			
358264	Deletion	A	-	104	98.07692	*rbsK*	169delA	Met57fs
581484	Insertion	-	A	178	94.94382			
581491	Insertion	-	A	177	97.74011			
581498	Insertion	-	A	181	94.47514			
1340128	SNP	T	A	162	100	*codY*	436T>A	Tyr146Asn
1564437	Deletion	A	-	156	96.79487			
1568676	SNP	C	A	141	100	CDR20291_1323	412C>A	Gln138Lys
1578174	Deletion	T	-	136	96.32353			
1578209	Insertion	-	A	140	94.28571			
1592813	SNP	A	T	157	100			
1864421	Insertion	-	T	138	97.82609	CDR20291_1576		
1899604	Deletion	A	-	129	94.57364			
2077305	Deletion	C	-	97	100			
2120669	SNP	A	G	176	99.43182	*vncR*	605A>G	Asp202Gly
2235743	Deletion	T	-	141	99.29078	CDR20291_1913	241delA	Lys81fs
2262066	Insertion	-	A	135	94.81481			
2264191	Deletion	T	-	164	98.78049			
2298116	Insertion	-	T	131	95.41985			
2361948	SNP	C	A	140	100			
2361961	Insertion	-	A	144	97.91667			
2367948	Insertion	-	T	146	97.94521			
2578164	Deletion	T	-	155	99.35484			
2674749	Deletion	T	-	155	99.35484			
2680792	Insertion	-	T	158	96.20253			
2772184	Deletion	T	-	208	97.59615	CDR20291_2368		
2881470	Deletion	T	-	112	98.21429	CDR20291_2456	1302delA	Leu434fs
3077994	Deletion	A	-	163	98.77301			
3162105	Deletion	T	-	182	96.7033			
3292465	SNP	T	C	170	100	CDR20291_2781	174A>G	Ile58Met
3361923	Deletion	A	-	167	97.00599			
3472928	SNP	G	A	171	100			

Illumina, paired-end reads of the three strains were mapped to the R20291 reference genome using CLC Workbench. Polymorphisms shared by all three strains are highlighted yellow. Those unique to the R20291 NM are highlighted green and those specific to R20291 Novartis are highlighted orange.

[a] – the reference sequence is derived from DNA supplied by Brendan Wren's laboratory at the LSHTM, ostensibly strain R20291 BW;

[b] – is the type of mutant allele present, being either a SNP, Insertion of Deletion;

[c] – is the sequence present in the Reference genome (Accession number *FN545816*);

[d] – indicates the sequence present in the genome of the three R20291 isolates sequenced here;

[e] – shows the number of Illumina reads;

[f] – shows the number of reads that contain the identified SNP or Indel;

[g] – indicates the identity of the gene in which the SNP or Indel is located;

[h] – shows the nucleotide change when the mutation resides within a gene, and;

[i] – indicates any resultant change in encoded amino acid caused by the SNP or Indel. The natural amino acid is indicated first, followed by its position in the encoded protein, and then the new encoded amino acid. A frame shift is indicated by fs.

4.3 Discussion

The discrepancies observed in the number of flagella between the R20291 NM strain from our laboratory (Soza *et al.*, 2013) and that from the LSHTM (R20291 BW) used in the publication of Martin *et al.* (2013) led us to investigate the phenotypical differences between strains as both were obtained from the Anaerobe Reference unit in Cardiff and therefore were expected to be isogenic.

The phenotypical tests demonstrated that R20291 NM has different phenotypic properties compared to the R20291 BW and a third strain, R20291 Novartis. The R20291 NM growth rate was different in the culture media used compared to the other two strains; the results suggest that R20291 NM possesses a different ability to use the various sugars present in the media, which suggests a difference in the metabolism of this strain. The motility assays (Section 4.2.2) revealed the impaired motility of the R20291 NM strain as it was not able to match the diameter of the halo formed in the swimming assay by the other two strains. The TEM images (Fig. 4.4) showed that the discrepancy in motility between strains could be explained by the difference in the number of flagella around the cell; the R20291 NM strain possesses only 1 polar flagella whereas the other two strains exhibit a *peritrichous* arrangement of *flagella* around the cell. The contrast in the number of flagella between strains suggest that the flagellar regulation in R20291 NM is different to that of the other two strains.

Biofilms are cell aggregates bound by a matrix of extracellular polymers (Mond & O'Toole, 2009); it has been observed that *Bacillus*, *Pseudomonas*, *Vibrio* and *Escherichia* species undergo a motility-to-biofilm transition that consists of two parts, the inhibition of rotation and the subsequent inhibition of transcription (Guttenplan & Kearns, 2013). As the R20291 NM strains possesses only 1 flagella, it is possible that the increased biofilm formation observed in this strain is caused by the impaired motility of the strain, achieving the inhibition of rotation and transcription faster than the other two strains and triggering the onset of biofilm formation earlier.

The R20291 NM strain is also characterised by a relative delay in the initiation of sporulation as the first spores can only be detected at the 48 h time point in the sporulation assay, contrasting with the 24 h time point observed in the R20291 BW and Novartis strains. Spo0A is the master regulator of sporulation in *C. difficile* and coordinates sporulation, metabolism and some virulence factors (Pettit *et al.,* 2014). The decision to enter sporulation in the cell is regulated by histidine kinases capable to phosphorylate Spo0A, triggering the onset of sporulation (Paredes-Sabja *et al.,* 2014). The delayed onset of sporulation in the cell could be related to a late phosphorylation of the Spo0A protein as it was observed that once the strain starts the sporulation process, it is capable of producing a similar amount of spores after 5 days compared to the other two strains, which would suggest that the sporulation pathway is not impaired but just delayed.

A difference in the toxin production between R20291 NM and the R20291 BW and Novartis strains was also observed. The R20291 Novartis and BW strains produced similar amounts of toxin while the R20291 NM strain produced greater quantities of toxins. It has been reported that the flagellar regulon plays an important role in the modulation of toxin production in *C. difficile* (Aubry *et al.*, 2012; Martin-Verstraete *et al.*, 2016); thus, the increased toxin production seen in the R20291 NM strain could be related to the different flagellar regulation also seen in this strain. Furthermore, it has been observed that the toxin production and the initiation of sporulation are related in *C. difficile* as Spo0A negatively regulates the production of toxin A and toxin B in the cell (Mackin *et al.*, 2013; Pettit *et al.*, 2014). It is possible that the elevated toxin phenotype observed in the R20291 NM strain could be caused not only by the different regulation of the flagellar operon but also to the late activation of Spo0A.

The whole genome analysis revealed that some SNPs were unique to the R20291 NM strain and consisted of 3 deletions and 2 SNPs. One of the deletions is found in an intergenic region, while the other two are located within genes. The deletions present in the *rbsK* and *CD*R20291*_2456* genes result in frameshift mutations. The frameshift mutation of the gene sequence might alter the amino acid sequence of the translated protein, producing an incomplete and possibly non-functional protein. The SNPs found in the *rsbW* and the *vncR* genes lead to a change in the amino acid sequence of the proteins, possibly compromising the folding of the proteins and other interactions. This could lead to the proteins having a reduced or ablated activity.

The *vncR* gene is part of the *vncR/S* two-component system, where it acts as a response regulator. In *Streptococcus pneumoniae*, the *vncR/S* system regulates an autolysin that provides tolerance to penicillin and vancomycin (Mitchell *et al.*, 2002). Furthermore, it has been observed that the mutation of *vncR* gene in *S. pneumoniae* leads to a phenotype that ablates tolerance to penicillin and vancomycin.

The *rbsK* is the structural gene of the ribokinase enzyme (Hope *et al.*, 1986); this enzyme is involved in the metabolism of ribose and catalyses the phosphorylation of ribose to ribose 5-phosphate, being the first step of the ribose pathway. The deletion found in our strain produces a sequence with a premature stop codon that leads to a truncated protein. The alteration of the RbsK protein could explain the difference in growth observed in the media containing different sugars.

*CD*R20291*_2456* is a putative two-component sensor histidine kinase that works along *CD*R20291*_2457*, the putative response regulator. Since the function of the *CD*R20291*_2456/ CD*R20291*_2457* two-component system has not yet been described, it is not clear the effect of the deletion in our strain. This two-component system presents a good opportunity to characterise it in a strain that doesn't contain the SNP and evaluate its role in the cell.

The *rsbW* gene is part of the *sigB* regulon, comprising *sigB*, the sigma factor B, *rsbw*, anti-sigma factor B and *rsbV*, a negative regulator of anti-sigma factor B. In *B. subtilis* the *rsbW* anti-sigma factor B interacts with sigma factor B, a globular regulator that controls the general stress response (Bandow *et al.,* 2002; Young *et al.,* 2013). Anti-sigma factors regulate numerous processes in the cell via negative regulation, including sporulation, flagella production and cellular growth (Brown & Hughes, 1995). The mutation of *rsbW* gene in the R20291 NM strain could be linked to the difference seen in the onset of sporulation between strains as the sporulation pathway and the sigma B regulon have been shown to be interconnected pathways in *B. subtillis* (Reder *et al.,* 2012). The change in the amino acid sequence of *rsbW* caused by the SNP may alter the function of this anti-sigma factor and the regulation of the entire operon, which could lead to significant changes in the internal regulation of the cell.

It is important to note that the behaviour of the R20291 BW strain, which does not present any unique genetic variation and the Novartis strains were similar throughout the different phenotypical assays that the strains were subjected to. This implies that the phenotype seen in the R20291 NM strain is anomalous and that the genetic variations found in this strains make it behave considerably different compared to the other strains. The R20291 strains were initially provided to the different laboratories by the Anaerobe Reference Unit, located in the University hospital of Wales, United Kingdom. Our laboratory uses a traffic light system for the long-term storage of strains, which consists of a green, a yellow and a red tube stored at -80°C. The green tube is used as the current stock, being replaced with a culture grown from the yellow tube once it is finished. The red tube is used only in emergencies and it is kept as the reference stock. We found that cultures from the red tube also contain the SNPs found in our current stock, implying that the SNPs were present in the strain that was sent to our laboratory from the Anaerobe Reference Unit and it was not a result from our stock system.

From communications with Dr. Val Hall at the Anaerobe Reference unit it transpired that popular strains were maintained on agar plates in an anaerobic cabinet and subcultured onto new agar plates every week, discarding the former plate once growth is seen in the new agar plate. This repeated subculture presents an opportunity for SNPs to arise in the strain. If the cultures are not properly curated, the presence of SNPs can be overlooked and it could lead to the formation of a mutant strain that possesses different characteristics compared to the parental strain. In this respect it is interesting to note that R20291 Novartis was originally derived from R20291 BW via the Sanger Institute. Thus, the SNPs that arose most likely occurred during its propagation at the Sanger Institute.

The phenotypic characteristics observed among the R20291 strains revealed that the behaviour of the strains that were thought to be the same are in fact not similar. The differences seen in the growth of the strains, the expression of virulence factors (toxin production and sporulation), their ability to form biofilms and the number of flagella found in the cell established that the presence of genetic variations in their genome have a great impact in the regulatory cascade within the cell. After evaluating the results observed in the R20291 NM strain, we concluded that the behaviour of the strain used in our laboratory strays greatly from the behaviour seen in the other R20291 strains and therefore further research of the *agr* system in this strain was not advisable.

Chapter 5

Role of the *sigB* operon in the regulation of virulence factors in R20291.

5.1 Introduction

Sigma factors are bacterial transcriptional factors that are needed for the initiation of RNA synthesis by RNA polymerase (RNAP). Through their interaction with RNAP they provide promoter selectivity, and guide the binding of the RNAP-sigma factor complex to a specific set of promoter sequences and the subsequent initiation of gene transcription. . Different types of sigma factor recognise different subsets of promoter element that are responsible for the transcription of a variety of genes, i.e. housekeeping genes, virulence genes and general stress response genes (Gruber & Gross, 2003). Anti-sigma factors bind to the sigma factor and inhibit their activity, negatively regulating the expression of genes. For example, in *E. coli* the anti-sigma factor 70Rsd is expressed in the stationary phase and binds to sigma factor 70, which regulates genes essential for the growth of the cell. Formation of this anti-sigma/sigma factor complex, results in the downregulation of housekeeping genes and the expression of stationary genes (Hofmann *et al.,* 2011). In *B. subtilis,* sigma B regulates the general stress response (Hecker *et al*., 2007). During the exponential phase of growth, when stress proteins are not needed, SigB is negatively regulated by RsbW as a consequence of the formation of the RsbW/SigB complex, inhibiting the association of SigB with the RNAP (Benson & Haldenwang, 1993). When stress proteins are required, RsbW is prevented from interacting with SigB by RsbV, an antagonistic protein capable of forming an alternative RsbV/RsbW complex (Dufour & Haldenwang, 1994). RsbV remains inactive during the growth phase as it is rapidly phosphorylated during this stage by RsbW, rendering it inactive.

When the cell is exposed to stress, RsbV is dephosphorylated by RsbU or RsbP and therefore able to interact with RsbW, liberating the SigB protein (Yang *et al.,* 1996; Vijay *et al.,* 2000). It has been shown that sigma B factor functions as a global regulator of virulence genes in *S. aureus* (Kullik *et al.,* 1998; Mitchell *et al.,* 2013) and that there is a connection between it, the *agr* system and biofilm maturation (Moisan *et al.,* 2006; Lauderdale *et al.,* 2009). As reported in other Gram-positive bacteria, the interaction between sigma B and its cognate anti-sigma factor is involved in the regulation of different processes by the cell (Ferreira *et al.,* 2001; van Schaik *et al.,* 2007; Guldimann *et al.,* 2016). The comparative whole genome sequencing analysis revealed that the *rsbW* gene in the R20291 NM strain is affected by a SNP that results in the change of 1 amino acid in the protein sequence; it is not clear if the change in the amino acid sequence could compromise the regulation of the SigB operon and the regulation pathways under its control.

5.2 Aim of this chapter

As described in the previous chapter (Chapter 4), the R20291 NM strain used in our laboratory possesses a unique phenotype not seen in other R20291 strains. It was observed that some virulence factors were differentially expressed in this strain, including toxin production, sporulation, flagella expression and biofilm formation. The observed differences seen in this strain could be explained by the unique genetic variations found in the R20291 NM strain following comparative whole genome sequence analysis. The genetic variations found in the R20291 NM strain affect two two-component systems, a ribokinase and the anti-sigma factor RsbW. RsbW is part of the SigB operon, which regulates the general stress response in the cell (Hecker *et al*., 2007). As observed in other Gram-positive bacteria, the SigB operon could take part in the regulation of virulence factors in the cell (Guldimann *et al.,* 2016). In this chapter, the role of the anti-sigma factor RsbW and the sigma factor SigB in the expression of virulence factors in *C. difficile* R20291 was investigated.

5.3 Results

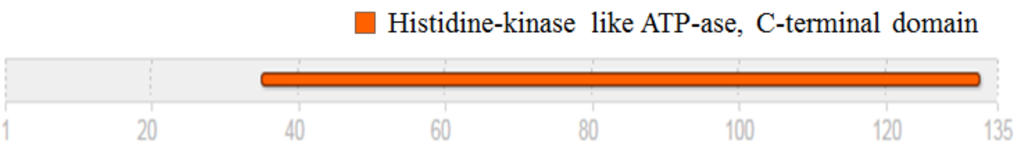
5.3.1 *rsbW* insertional mutant in *C. difficile* R20291 NM

The anti-sigma factor RsbW is a 135-amino acid long protein that is able to bind and form a sTable complex with sigma factor B (SigB), negatively regulating the genes controlled by SigB (Guldimann *et al.,* 2016). The presence of a SNP in the *rsbW* gene found in *C. difficile* R20291 NM causes a change in the amino acid located at the position 82, from glycine (Gly) to valine (Val), altering the sequence of the resulting RsbW protein (Fig. 5.1a). The SNP found in the R20291 NM strain is located in the predicted C-terminal domain, a histidine-kinase like ATP-ase (SMART accession no.SM00387) (Fig. 5.1b). As RsbW needs to phosphorylate RsbV in order to interact with SigB, the effect of the amino acid substitution in the protein domain is uncertain, as valine possesses an isopropyl group not present in the glycine amino acid, which could compromise the folding of the protein or its active site by steric hindrance.

a)

```
       10         20         30         40         50
MACETIKMEI TTNPDYVSII RLTASGIANK MGFPIDDIED IKVAVSEACT
       60         70         80         90        100
NAIKHSKDDV FNIVFNILDN AINIEIQDNG KG/VYDVSSIST PDLENPKESG
      110        120        130
LGLFIIKTLM DDVDIESEHN QGTKIKMTKY LGVDI
```

b)



c)

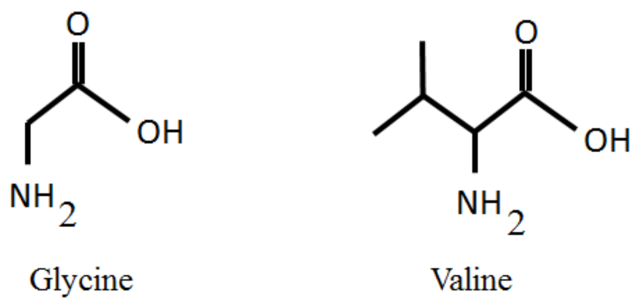


Figure 5.1 Sequence of the RsbW protein in *C. difficile* R20291. a) The sequence of the RsbW protein was taken from the reference genome (accession no. FN545816) published by the Wellcome Trust Sanger. The SNP found in the *C. difficile* R20291 NM strain produces a change in the amino acid located at the position 82; the amino acid found in the reference genome (Glycine, G) is depicted in bold whereas the amino acid found in the *C. difficile* R20291 NM strain (Valine, V) is coloured in red. b) Predicted domain in the RsbW protein sequence (SMART accession no.SM00387). c) Glycine and valine amino acid structures.

5.3.1.1 Construction of the mutants

In order to evaluate if the amino acid change in RsbW protein sequence results in the loss of activity, an insertional *rsbW* mutant was created using the ClosTron technology (Heap *et al.*, 2007) in *C. difficile* R20291 NM. The ClosTron vector was created using the algorithm at www.clostron.com. The resulting plasmid, pMTL007C-E2::CdiR2029l*rsbW*-130bp, was re-targeted to insert the group II intron in a sense orientation at the 130bp position relative to the start of the *rsbW* open reading frame (ORF) (Fig. 5.2) . If the change in the amino acid sequence of RsbW has led to a loss of activity, the insertional interruption of its gene should have no effect on the phenotype.

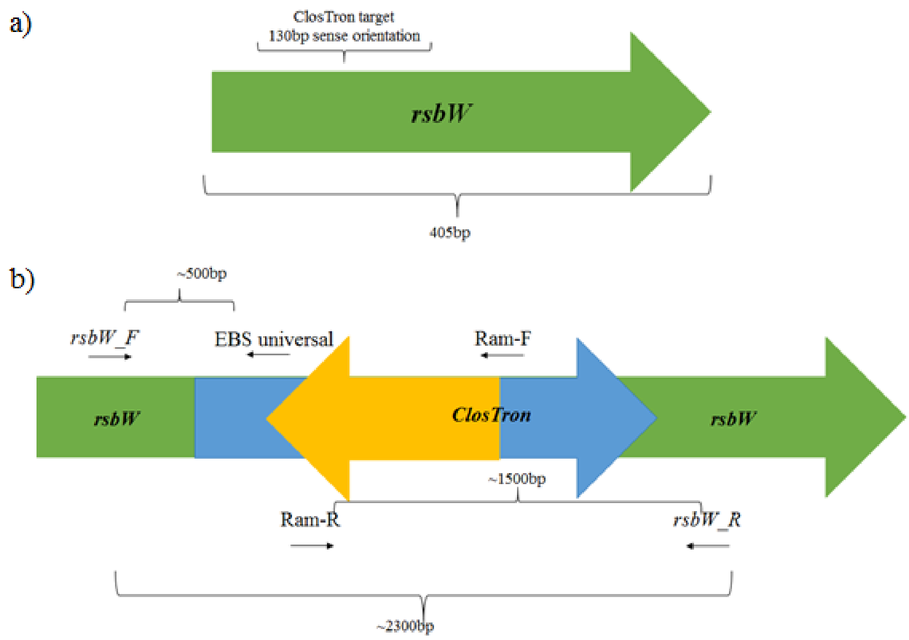


Figure 5.2. Schematic representation of the insertion in the *rsbW* gene. The ClosTron intron was retargeted to be inserted in the *rsbW* gene. a) Representation of the *rsbW* gene and location of the ClosTron insertion in the gene. b) Primers that can be used for PCR screenings; the length of the expected sequences is shown.

The pMTL007C-E2::CdiR2029l*rsbW*-130bp plasmid was transformed into *E. coli* NEB Sexpress using electroporation and the resulting strain was subsequently used as donor strain to transfer the ClosTron vector into *C. difficile* R20291 NM as described in Chapter 2. In order to authenticate the correct insertion of the intron into the *rsbW* gene, four independent colonies were screened using PCR. The rsbW_F1 primer, located at the beginning of the gene and the EBS universal primer, located in the ClosTron insertion, were used to amplify the junction between these two segments, which is only present if the group II intron has been inserted in *rsbW* (Fig. 5.3). The fragments were sent for sequencing to confirm the junction between the gene and the intron. The resulting *rsbW* mutant was designated *C. difficile* R20291 NM *rsbW*::CTermB (NM *rsbW* mutant).

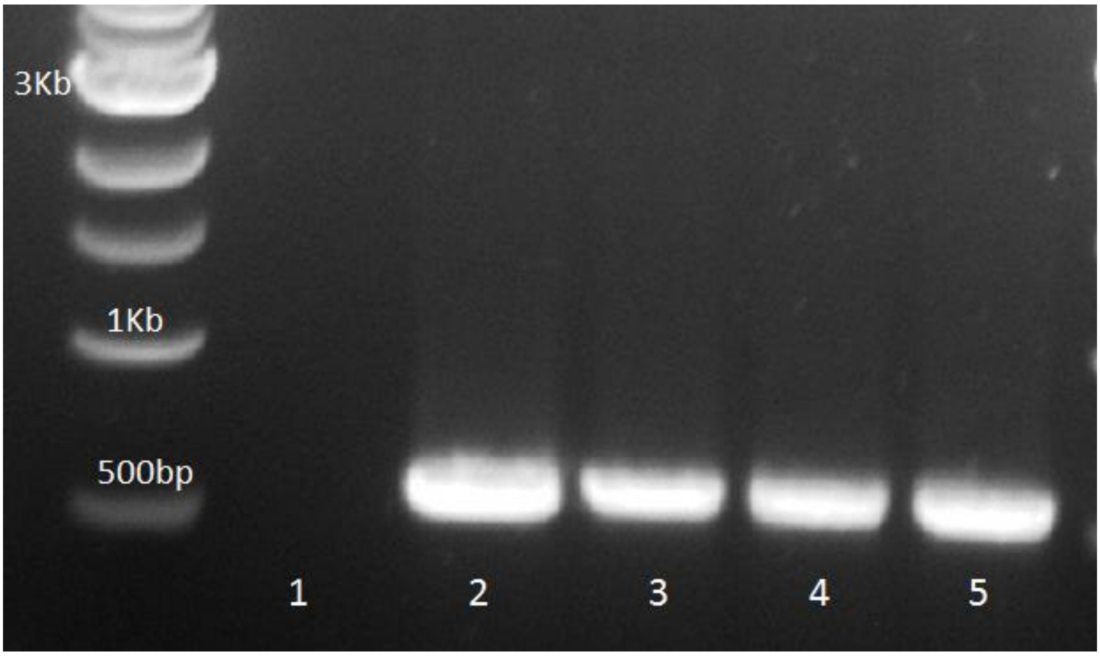


Figure 5.3 PCR screening to confirm the presence of the intron in the *rsbW* gene of the *C. difficile* R20291 NM strain. rsbw_F1 + EBS universal primers were used for the amplification of the junction. The fragment can only be detected if the group II intron has been inserted in the target gene. The parental strain (lane 1) and four independent lincomycin-resistant colonies harbouring the pMTL007C-E2::CdiR2029l*rsbW*-130bp plasmid (lane 2-5) were screened. At the extreme left, a 1KB DNA ladder is used as a weight marker.

The plasmid pMTL84151::rsbW_NM was constructed in order to complement the *C. difficile* R20291 NM *rsbW*::CTermB strain. The plasmid pMTL84151::*rsbW*_NM contains the *rsbW* gene present in the *C. difficile* R20291 NM strain and the 120bp region immediately upstream of the ORF. This region was selected as being likely to encompass the *rsbW* native promoter using the Genome2D webserver for analysis and visualization of bacterial genomes and transcriptome data (University of Groningen). The rsbW_prom_F1 and rsbW_prom_R1 pair of primers were used to amplify the desired sequence containing *Not*I and *Xba*I restriction sites at the 5' and 3' end, respectively. The multiple cloning site (MCS) located in the pMTL84151 shuttle vector harbours one *Not*I restriction site and one *Xba*I restriction site. The purified fragment and the pMTL84151 plasmid were restricted with *Not*I and *Xba*I restriction enzymes, purified and ligated using the T4 DNA ligase kit (New England Biolabs), resulting in the pMTL84151::*rsbW*_NM plasmid (Fig. 5.4). *C. difficile* R20291 *rsbW*::CTermB pMTL84151::*rsbW*_NM was constructed by introducing the complementation plasmid via conjugation into *C. difficile* R20291 *rsbW*::CTermB. The plasmid was recovered from the complemented strains and sent for sequencing to confirm the insertion.

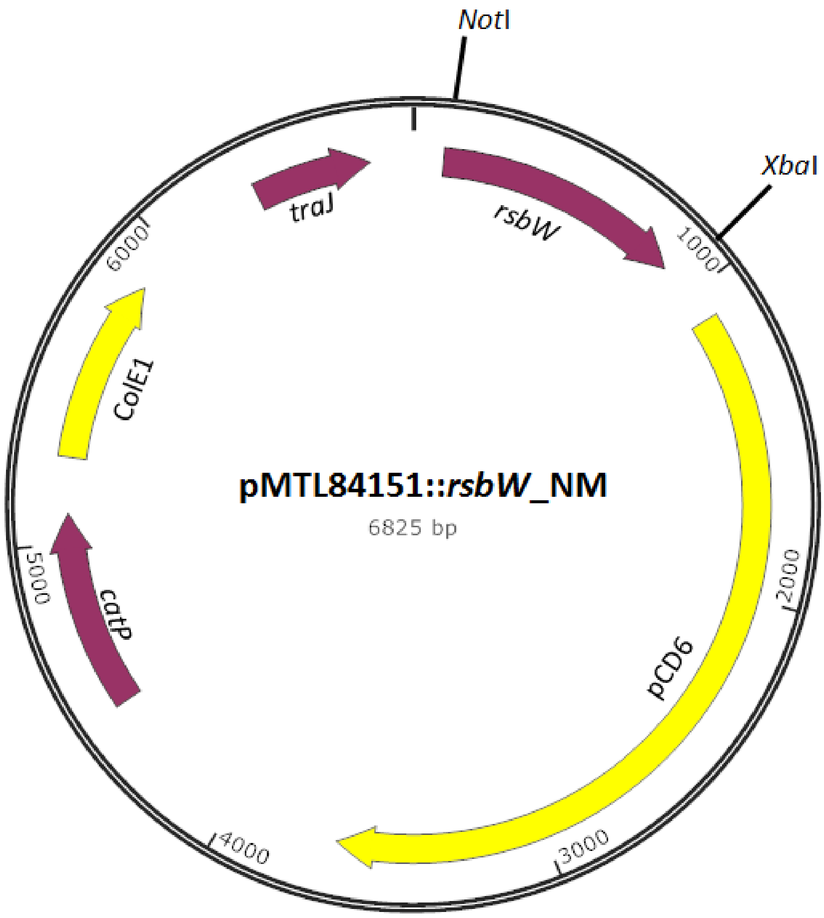


Figure 5.4 Plasmid used for the complementation of *C. difficile* R20291 NM *rsbW*::CTermB. Plasmid components are: *catP* gene of *C. perfringens* conferring thiamphenicol resistance; ColE1, the replication region of the *E. coli* plasmid ColE1, and; *traJ*, transfer function of the RP4 oriT region; pCD6, replication region of the *C. difficile* indigenous plasmid pCD6; *rsbW*, sequence of the *rsbW* gene found in *C. difficile* R20291 NM.

5.3.1.2 Phenotypical analysis of the *C. difficile* R20291 NM *rsbW* mutants

As seen in the previous chapter, spores can be detected in R20291 NM after 48 h, while the R20291 BW and Novartis have initiated sporulation after just 24 h. In order to assess if the interruption of the *rsbW* gene could affect the sporulation in the R20291 NM strain, the production of heat-resistant CFU by the mutants was evaluated over 5 days (Fig. 5.5).

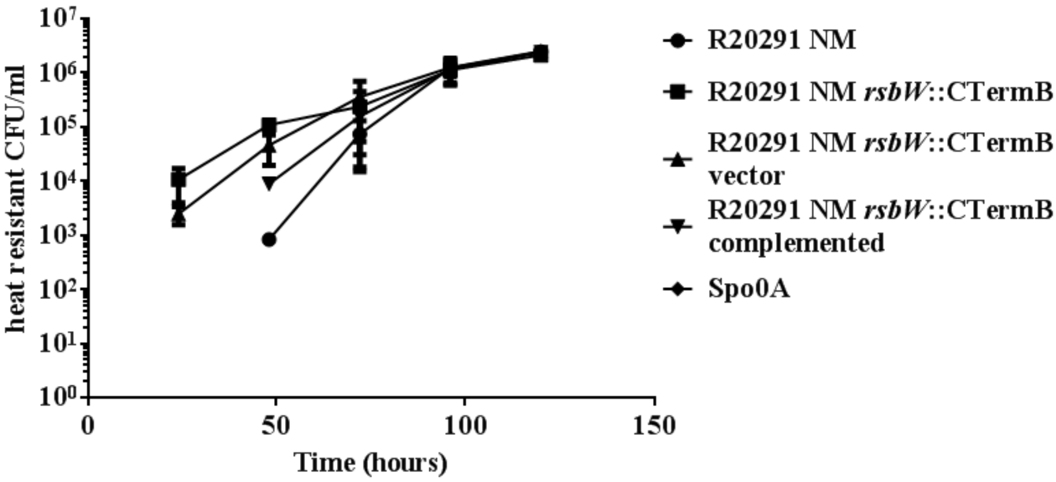


Figure 5.5. Sporulation profile of the R20291 *rsbW* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. The R20291 NM *rsbW*::CTermB mutant harbouring pMTL84151 has been abbreviated as vector and the complemented strain refers to R20291 NM *rsbW*::CTermB carrying pMTL84151::*rsbW*_NM. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

Interestingly, spores could be detected in the NM *rsbW* mutant cultures after only 24 h, in contrast to the 48 h required for detecTable sporulation in the parental strain. This indicates that the insertion in the *rsbW* gene affects the regulation of spores and brings the onset of sporulation forward in time. Sporulation in the NM *rsbW* mutant carrying the complementation vector was restored to that of the parental strain, with spores being detected after 48 h; nevertheless, the complemented strain was able to produce a higher number of spores after 48 h compared to the parental *rsbW* mutant. The R20291 NM strains was previously shown (as described in Chapter 4) to produce higher toxin levels compared to the other R20291 strains (BW and Novartis). As the regulation of toxin production and sporulation seem to be connected in R20291 NM, the ability to produce toxins by the NM *rsbW* mutant was tested after 72 h of growth in TY medium incubated in anaerobic conditions.

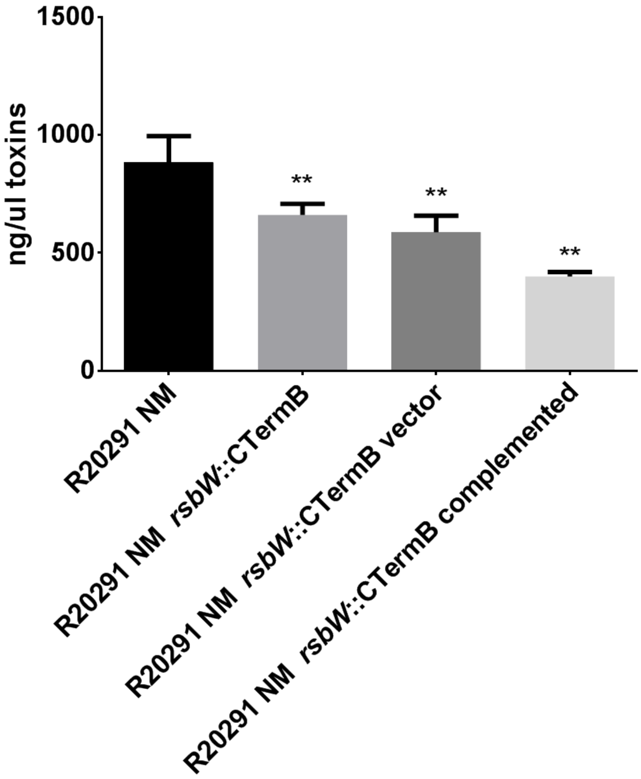


Figure 5.6. Toxin production of the R20291 NM *rsbW* mutant strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. pMTL84151 has been abbreviated as vector and the complemented strain refers to R20291 NM *rsbW*::CTermB carrying pMTL84151::*rsbW*_NM. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's t test) compared to the wild type.

The NM *rsbW* mutant showed a reduced toxin production compared to the parental strain (Fig. 5.6). The complementation vector carrying the *rsbW* sequence brought about a further reduction in toxin levels. This same phenomenon was seen in the R20291 NM *agrB* mutant complemented using the same vector plasmid (as described in Chapter 3). It is unclear why the presence of the complementation plasmids affect the toxin production whereas the presence of the empty vector does not have the same effect. Due to this phenomenon, it is unclear if the toxin phenotype observed can be reverted

In order to determine if the interruption of the *rsbW* gene has any effect on the motility of the strain, swarming and swimming assays (Fig. 5.7a and Fig. 5.7b) were performed as described in Chapter 2. Biofilm formation after five days was also evaluate in the *rsbW* mutants (Fig. 5.7c). No significant difference was found among the strains compared to the parental strain in any of the assays. These results suggest that, even though the interruption of the gene affects the expression of spores and toxin production, it plays no role in the regulation of motility or biofilm formation in the cell.

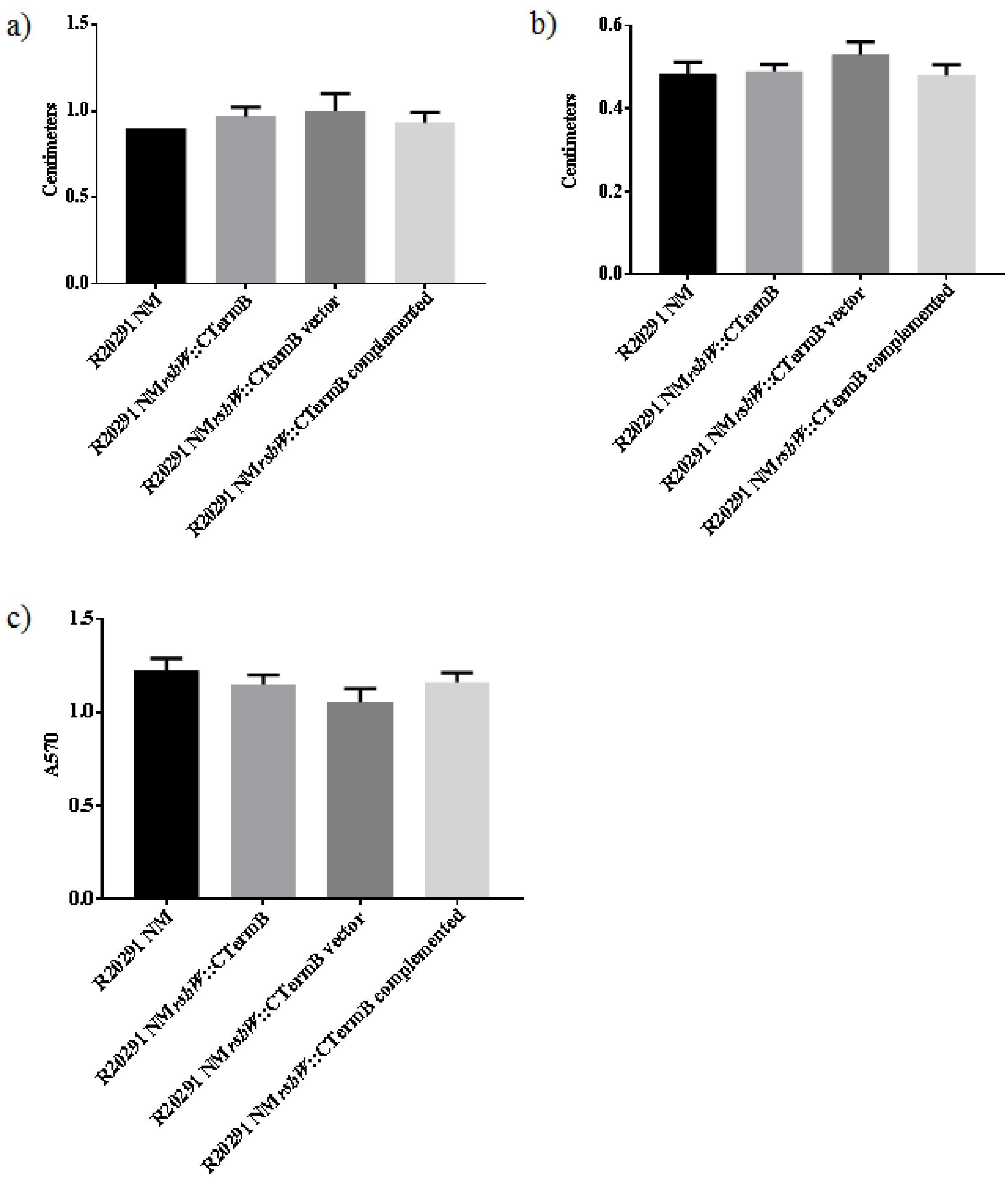


Figure 5.7 Motility and biofilm production of the R20291 NM *rsbW*::CTermB mutants. a) Swarming ability of the R20291NM *rsbW*::CTermB strains in BHIs 0.4% agar plates. b) Swimming assay in BHIs 0.3% agar plate. c) Biofilm production of the R20291 NM *rsbW*::CTermB strains after 5 days, the strains were grown in BHIs supplemented with 0.1M glucose. pMTL84151 has been abbreviated as vector and the complemented strain refers to R20291 NM *rsbW*::CTermB carrying pMTL84151::*rsbW*_NM. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means.

5.3.1.3 Construction of recombinant allele for allelic exchange in R20291 NM

As observed in the NM *rsbW* insertional mutant, the interruption of *rsbW* modifies the expression of sporulation and toxin production in the R20291 NM strain. These results suggest that the *rsbW* SNP that leads to a change in the RsbW amino acid sequence present in R20291 NM does not result in the inactivation of the protein, as the phenotype seen in the mutant should match the phenotype seen in the wild type otherwise. Allelic-coupled exchange (ACE) technology allows the precise manipulation of the genome and provides the tools required to make single nucleotide changes to individual genes, such as *rsbW* (Ng *et al.,* 2013). The method is reliant on two important events: (I) the integration of the vector plasmid into the genome, producing single cross-over integrants through homologous recombination, and (II) the excision of the plasmid from the genome, producing double cross-over excision derivatives. The latter are selected using a negative selection marker (Cartman *et al*., 2012). The use of this technology offers the opportunity to restore the SNP found in the *rsbW* gene of *C. difficile* R20291 NM.

In order to restore the SNP in R20291 NM, an appropriate allelic exchange vector was constructed. The homology arm was created using the *rsbW* gene from the R20291 BW strain, as this strain does not possess the SNP seen in the R20291 NM strain. The homology arm consists of the *rsbW* gene from the R20291 BW strain, the 579bp region upstream and the 468bp region downstream of the gene. *Pme*I restriction sites were added to the primers used to amplify the homology arm in order to flank the sequence with this restriction site. The vector plasmid pMTLSC7215 containing the negative selection marker *codA* and the homology arm were restricted with *Pme*I and later ligated to construct the allelic exchange vector, resulting in the pMTLSC7215::BW*rsbW* ACE vector (Fig. 5.8). The vector was sent for sequencing in order to verify the sequence inserted in the plasmid. The pMTLSC7215::BW*rsbW* vector was introduced into *E. coli* CA434 via electroporation and subsequently conjugated with strain R20291 NM as described in Chapter 2.

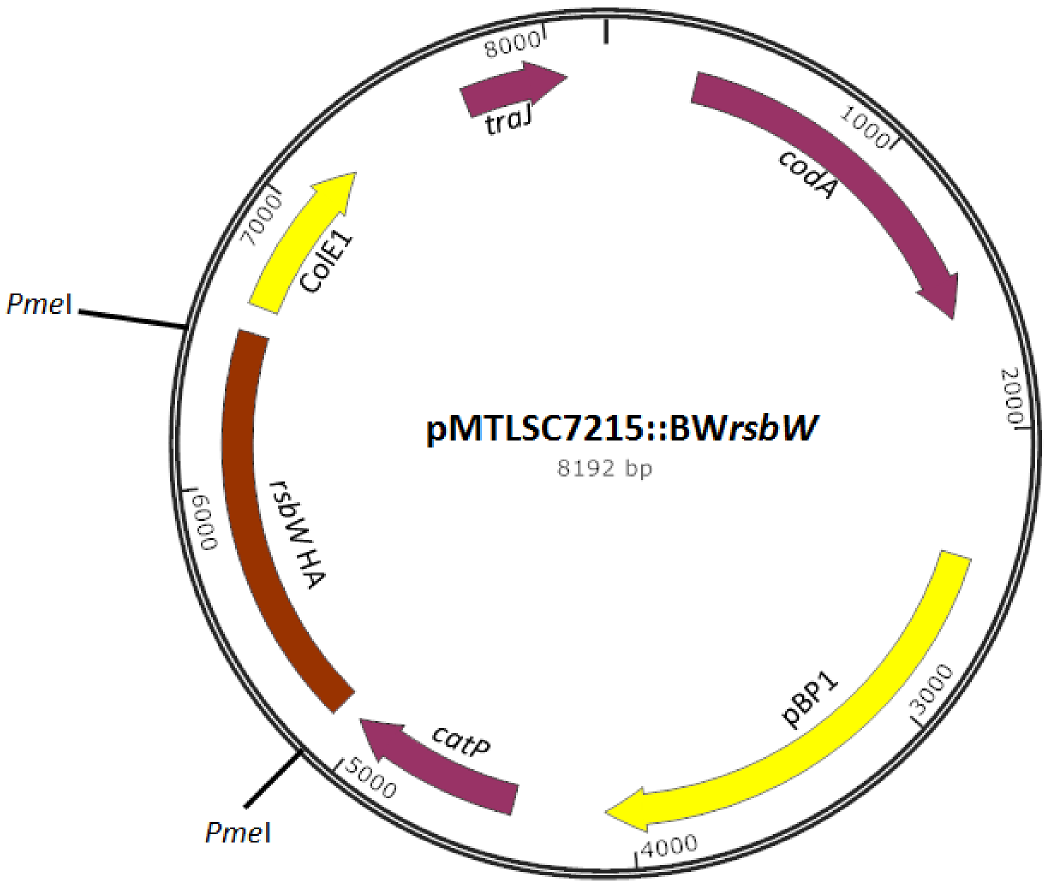


Figure 5.8 Plasmid map of ACE vector pMTLSC7215::BW*rsbW*. Allele-Coupled Exchange (ACE) vector designated to replace the *rsbW* gene harbouring the group II intron in the R20291 NM *rsbW*::CTermB strain and replace it with the *rsbW* gene from the R20291 BW strain. Plasmid components are: *catP* gene of *C. perfringens* conferring thiamphenicol resistance; *rsbW* HA, the homology arm consists of the *rsbW* gene found in the R20291 BW strain, the 579bp region upstream and the 468bp region downstream of the gene; ColE1, the replication region of the *E. coli* plasmid ColE1, and; *traJ*, transfer function of the RP4 oriT region; *codA*, cytosine deaminase gene (*codA*) of *E. coli* used as a negative selection marker; pBP1, replication region of *C. botulinum* plasmid pBP1.

Fast-growing thiamphenicol-resistant colonies were isolated and PCR screened to observe if the colonies were single cross-over mutants. The RsbW_screening_F and the SC7_R pair of primers were used to screen for the single cross-over in the 5' end, while the RsbW_screening_R primer and the SC7_F were used to look for a single cross-over at the 3' end (Fig. 5.9). The RsbW_screening_R and SC7_F pair were unable to yield any band, implying that the single cross-over event at the 3' end did not occurred. The RsbW_screening pair of primers are situated outside the homology arm and they are designed to amplify with the SC7 pair of primers, located inside the vector, only if the first cross-over event has occurred. The PCR screening showed faint bands of the expected size (~900bp) for the PCR where the RsbW_screening_F primer and the SC7_R primer were used. The faint bands suggested that there was a small portion of the population where the single cross-over event had occurred, meaning that there was a mixed population of wild type and single cross-overs. Nevertheless, subsequent attempts to isolate single cross-overs from these cultures were unsuccessful. The isolation of single cross-overs is important in order to continue with the next step in the methodology, as the presence of wild type in a mixed population will show false-positive colonies growing on the negative selection plate.

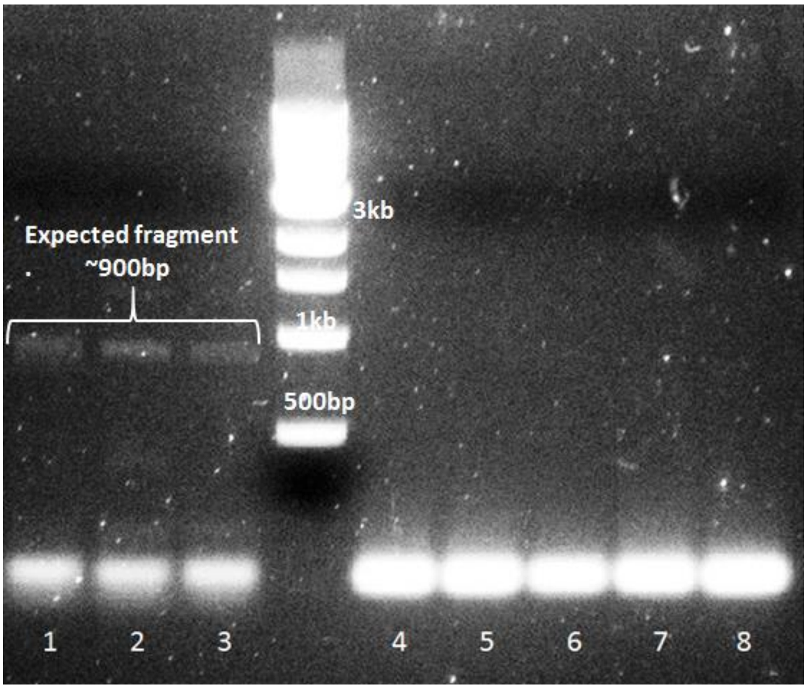


Figure 5.9 PCR screening of *C. difficile* R20291 NM *rsbW* single crossover mutants. Fast growing thiamphenicol-resistant colonies (lane 1-3) were screened with RsbW_screening_F primer and the SC7_R primer for the integration of the ACE plasmid at the 5' end. The positions 4, 5 and 6 represent the same colonies (respectively) were screened with the RsbW_screening_R and SC7_F to look for the integration of the vector at the 3' end. Position 7 and 8 represent amplification of the parental strain R20291 NM with the RsbW_screening_F - SC7-R pair and RsbW_screening_R - SC7-F pair (respectively). A 1KB DNA ladder between lane 3 and 4 is used as a weight marker.

5.3.2 *rsbW* insertional mutant in *C. difficile* R20291 BW

The phenotypical characterisation of the *rsbW* mutant created in the R20291 NM strain revealed that the RsbW protein expressed in this strain is not inactive, as the phenotype of the mutant would mirror the phenotype of the wild type otherwise. The insertional interruption of the *rsbW* gene in strain R20291 NM affects sporulation and toxins production in the cell. However, the evaluation of the *rsbW* gene in a strain that does not contain the SNPs found in R20291 NM is necessary to confirm the role of *rsbW* gene in sporulation and toxin production, as it has been observed that other insertional mutants created in R20291 NM (namely NM *agrB1* and NM *agrB2* mutants) had a different effect in sporulation and toxin production compared to those created in R20291 BW (see Chapter 3).

5.3.2.1 Construction of the mutants

Previously, the conjugation of plasmids into the R20291 BW strain had proven challenging and transconjugants from conjugative donor *E. coli* CA434 could seldom be obtained. Nevertheless, the use of the *E. coli* NEB Sexpress strain as a donor was shown to solve the problem, at least in the case of transposons (as observed in the construction of ClosTron mutants in Chapter 3). Building on this success, we therefore set about making ClosTron mutants of *rsbW* in R20291 BW, following the same procedure used to create the *rsbW* mutant in the R20291 NM (as described in section 5.3.1.1) with the exception of using *E.coli* NEB Sexpress as a conjugative donor instead of *E.coli* CA434. Five lincomycin-resistant colonies were screened by PCR to confirm the junction between the *rsbW* gene and the group II intron (Fig. 5.10). The resulting mutant was designated *C. difficile* R20291 BW *rsbW*::CTermB (BW *rsbW* mutant).

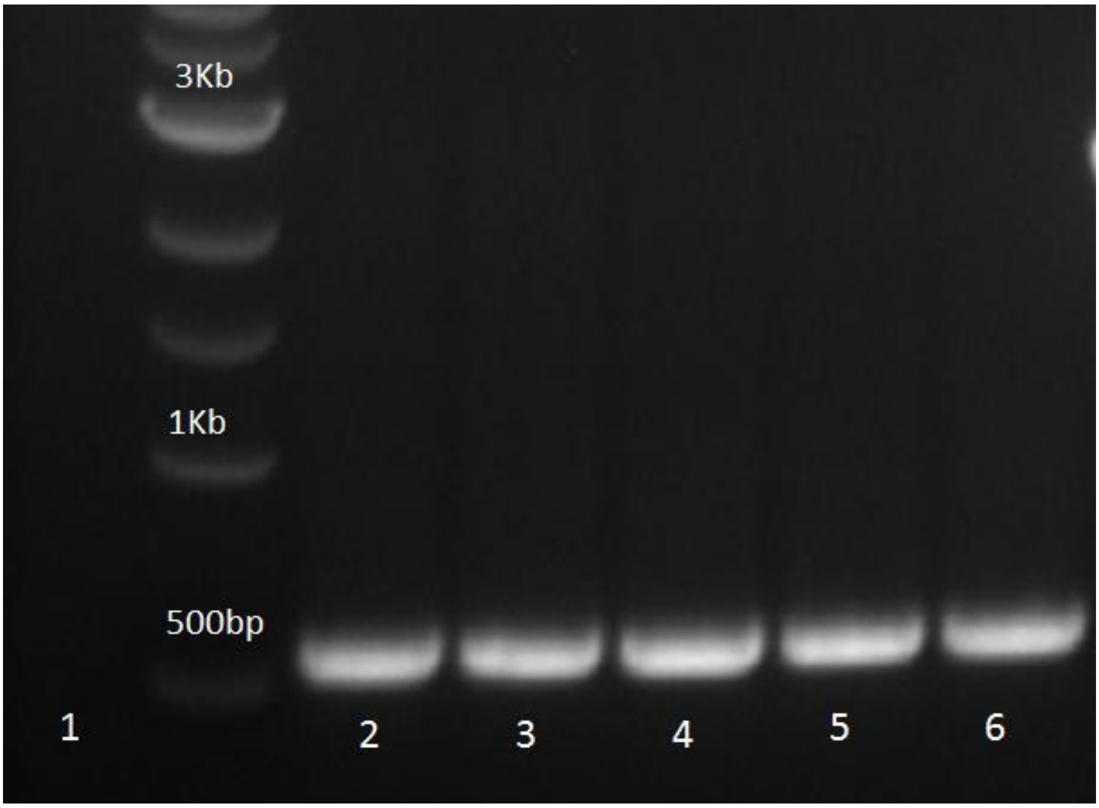


Figure 5.10 PCR screening to confirm the presence of the intron in the *rsbW* gene of the *C. difficile* R20291 BW strain. The junction between the flanking primer of the gene and the EBS universal sequence in located the intron was amplified to confirm the insertion of the group II intron in the genome. The parental strain (lane 1) and five independent lincomycin-resistant colonies (lane 2-6) were screened for the insertion of the group II intron. rsbw_F1 + EBS universal primers were used for the amplification of the junction. A 1KB DNA ladder between lane 1 and 2 is used as a weight marker.

The phenotype of the R20291 BW *rsbW*::CTermB strain was investigated to determine if there was any change in the production of toxins, sporulation, motility or biofilm formation. In terms of sporulation, the BW *rsbW* mutant resembled the phenotype observed in the parental strain (Fig. 5.11a). However, the mutant appeared to produce greater amounts of toxin at 72 h compared to the parent strain (Fig. 5.11b). In all other aspects (motility and biofilm formation after 5 days), the behaviour of the *rsbW* mutant was identical to the parental strain (Fig. 5.11c, Fig. 5.11d and 5.11e). Accordingly, the only tested phenotype affected in the R20291 BW *rsbW* mutant was the amount of toxin produced. The only effect observed in the BW *rsbW* mutant is the amount of toxin produced compared to the wild type. The complementation of the R20291 BW *rsbW* mutant with the pMTL84151 plasmid was not pursued as it has been observed that this methodology is sometimes detrimental for the evaluation of toxins in the mutants (as seen in the complementation of the *agrB2*, *agrB1* and *rsbW* insertional mutants made in the R20291 NM strain). However, new methodologies for complementation into the genome that do not present this disadvantage, such as Allelic exchange and the use of CRISPR/Cas9 system, are being investigated in our laboratory.

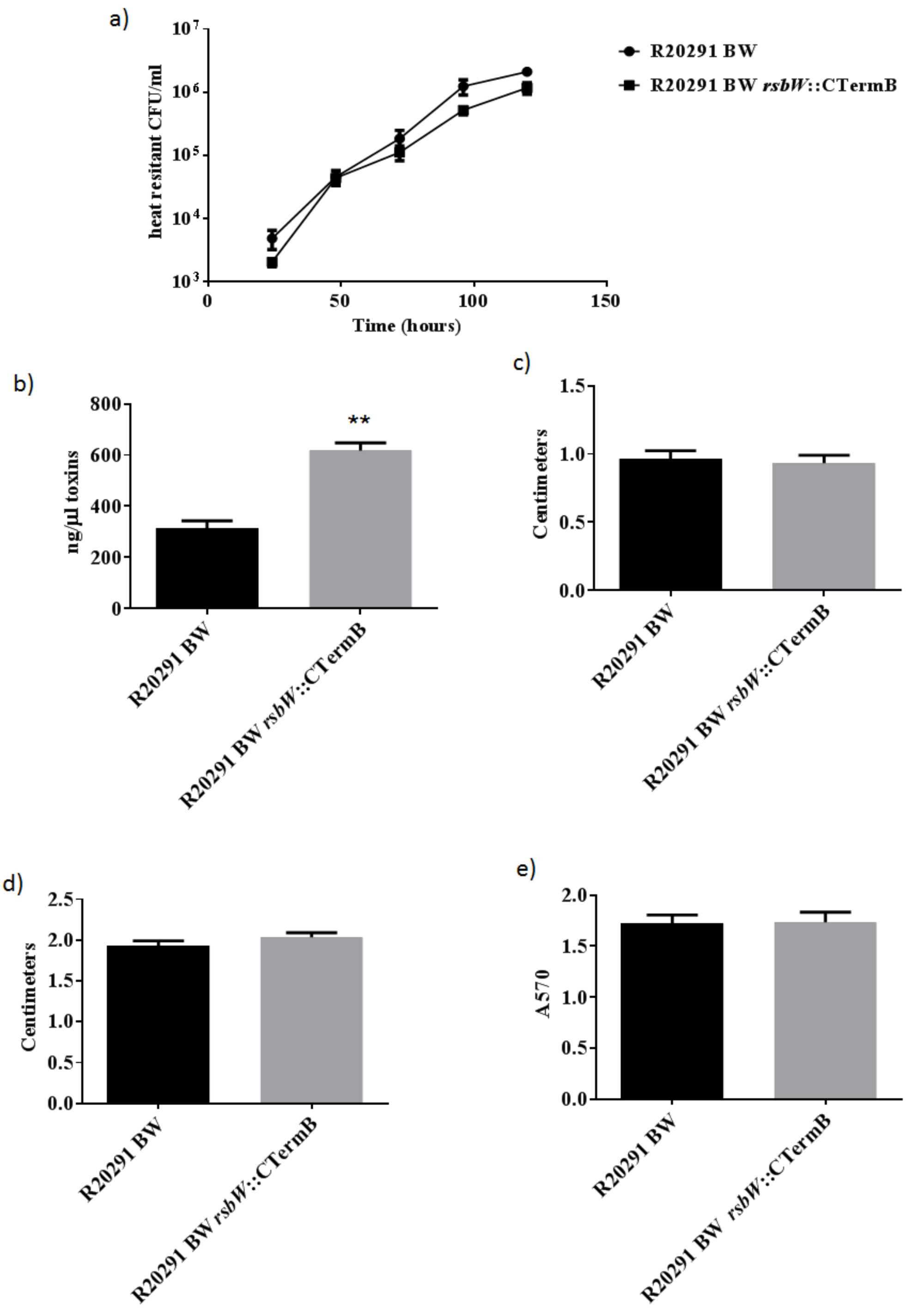


Figure 5.11 Phenotypical analysis of the *rsbW* mutant made in *C. difficile* R20291 BW. a) Sporulation assay in BHIs broth supplemented with 0.1% taurocholate over 5 days. b) Toxin production after 72 h in TY medium. c) Swarming ability of the R20291 BW *rsbW*::CTermB strain in BHIs 0.4% agar plates. d) Swimming assay in BHIs 0.3% agar plate. e) Biofilm production after 5 days; the strains were grown in BHIs supplemented with 0.1M glucose. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's t test) compared to the wild type.

5.3.2.2 Creation of a CRISPR/Cas9 system method for the removal of the group II intron in ClosTron mutants

Attempts to obtain mutants using allelic exchange appeared to fail at the first stage of the procedure, and single crossover integrants could not be obtained. This may imply that such an event is detrimental to the cell. An alternative solution would be to directly select double crossover mutants. The use of systems based on CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) /Cas (CRISPR-associated system) genome editing technology offer an opportunity to achieve this goal.

CRSIPR and the associated Cas is an immune system found in bacteria and archaea that provides protection against foreign DNA by cleaving specific target sequences (Sorek *et al.,* 2013). Short DNA sequences between the CRISPR repeats designate the targets of the system. When the CRISPR-array is transcribed the repeat sequences produce RNA fragments that are referred to as CRISPR-RNA; this guide RNA directs the Cas nuclease to the target site. The Cas nuclease requires the presence of a specific protospacer-adjacent motif (PAM) in the target site in order to cleave the target sequence (Oh & van Pijkeren, 2014). In the type II CRISPR-Cas system, an RNA complex is formed between Cas9 and the sgRNA, which is formed by the guide RNA combined with a transactivating CRISPR-RNA that is required for the Cas9 nuclease to be functional (Chylinski *et al.,* 2013). The guide RNA can be modified in order to retarget the system to specific locations to promote double-stranded breaks in a desire DNA sequence (Jiang *et al.,* 2013). The CRISPR/Cas9 system from *Streptococcus pyogenes* has been successfully used to create deletion mutants in *C. beijerinckii, C. acetobutylicum* and *C. ljungdahlii* (Wang *et al.,* 2015; Li *et al.,* 2016; Huang *et al.,* 2016). In short, the deletions were achieved by using a plasmid containing two homology arms flanking the target gene (the editing template or donor DNA) and the CRISPR/Cas9 system retargeted to the desired gene. The CRISPR/Cas9 system cleaves the genome at the desire location and only cells that repair the genome through homologous recombination with the homology arms contained in the plasmid are able to survive (Fig. 5.12). The possibility of manipulating the genome through the use of the CRISPR/Cas9 system offers the advantage that the system is continuously active in the cell, so that mixed populations will eventually become pure recombinants after a certain number of restreaks. The possibility to deliver plasmids into R20291 BW opens up the opportunity to use this newly developed tool in this strain. In order to evaluate the effect of the SNP found in the *rsbW* gene from the R20291 NM strain in a clean background, we decided to insert the SNP found in the R20291 NM strain into the R20291 BW strain genome. One of the challenges we had to circumvent was the fact that both, the native gene and the homology arm, share the same DNA sequence except for 1 base.

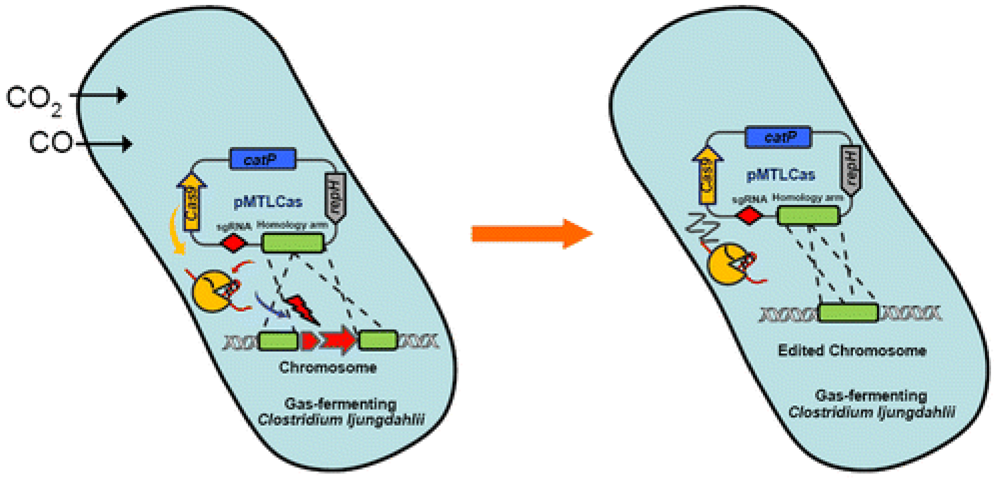


Figure 5.12 Schematic representation of the CRISPR/Cas9 system used in *C. ljungdahlii*. The CRISPR/Cas9 system is targeted to the desire gene in the chromosome where it cleaves the DNA strand, compromising the integrity of the genome. The homology arm located in the plasmid is integrated into the genome through homology recombination and the genome is repair. Cells that do not integrate the homology arm are unable to survive (Huang *et al.*, 2016).

Targeting the gene with CRISPR/Cas9 would result in the cleavage of both DNA sequences, compromising the survivability of the cells. To avoid this scenario, we decided to create an *rsbW* insertional ClosTron mutant in R20291 BW and retarget the CRISPR/Cas9 system to cleave the group II intron. This offers the opportunity to use the native gene found in the R20291 BW strain as the homology arm for the restoration of the DNA in the genome. As the group II intron confers resistance to lincomycin, true recombinants become sensitive to this antibiotic, offering a further screening option for identifying the desired recombinant colonies. Furthermore, the same retargeted plasmid can be used to target any other mutant created with the ClosTron technology as long as the homology arm is replaced with the corresponding homology arm.

The pMTLCas-pta plasmid (Huang *et al.*, 2016) used to create the deletions mutants in *C. ljungdahlii* was modified to target the group II intron found in the insertional mutants created by the use of the ClosTron technology. The replicon in the pMTLCas-pta plasmid was changed to the pBP1 replicon found in the pMTL84251 plasmid (Heap *et al.*, 2009) as this replicon has been successfully used in *C. difficile* in our laboratory. The 20bp target sequence was selected using the computer algorithm for the identification of suiTable regions from Benchling (Benchling.com). The retargeted sequence fragment was created using SOE-PCR in order to merge two fragments created using the sgRNA_2.0_F1 and SOE_intron_R1 and SOE_intronF1 and sgRNA_3.1_R1. The SOE fragment contains an *Xba*I restriction site at the 5' end and a MauBI restriction site at the 3' end. Additionally, an AsiSI restriction site was inserted six bases prior to the AscI restriction site. The retargeted fragment and the pMTLCas-pta plasmid were digested using *Xba*I and *Mau*BI restriction enzymes, purified and ligated, resulting in the pMTLCas-intron plasmid. The homology arm for the *rsbW* gene consists of the *rsbW* gene found in the R20291 BW strain, the 579bp region upstream and the 588bp region downstream of the gene and it was created using the rsbW_HA_crispr_F1 and rsbW_HA_crispr_R1. The resulting fragment contains a MauBI restriction site at the 5' end and an *Asc*I restriction site at the 3' end. The pMTLCas-Intron plasmid and the homology arm fragments were digested with *Mau*BI and *Asc*I restriction enzymes, purified and ligated to create the pMTLCas-Intron:: rsbw plasmid (Fig. 5.13a). The resulting plasmid was transferred to *E. coli* Top10 by electroporation. The final plasmid was digested with *BsmB*I and *Asc*I to check the integrity of the plasmid; the expected fragment (~2200 bp) consists of the retargeted sgRNA and the *rsbW* homology arm (Fig.

5.13b). The use of *Xba*I in the final plasmid is not recommended as the homology arm presents an additional *XbaI* restriction site whiting the coding region, resulting in two fragments of similar size instead of one. The plasmid was sent for sequencing to ensure the integrity of the inserted sequences.

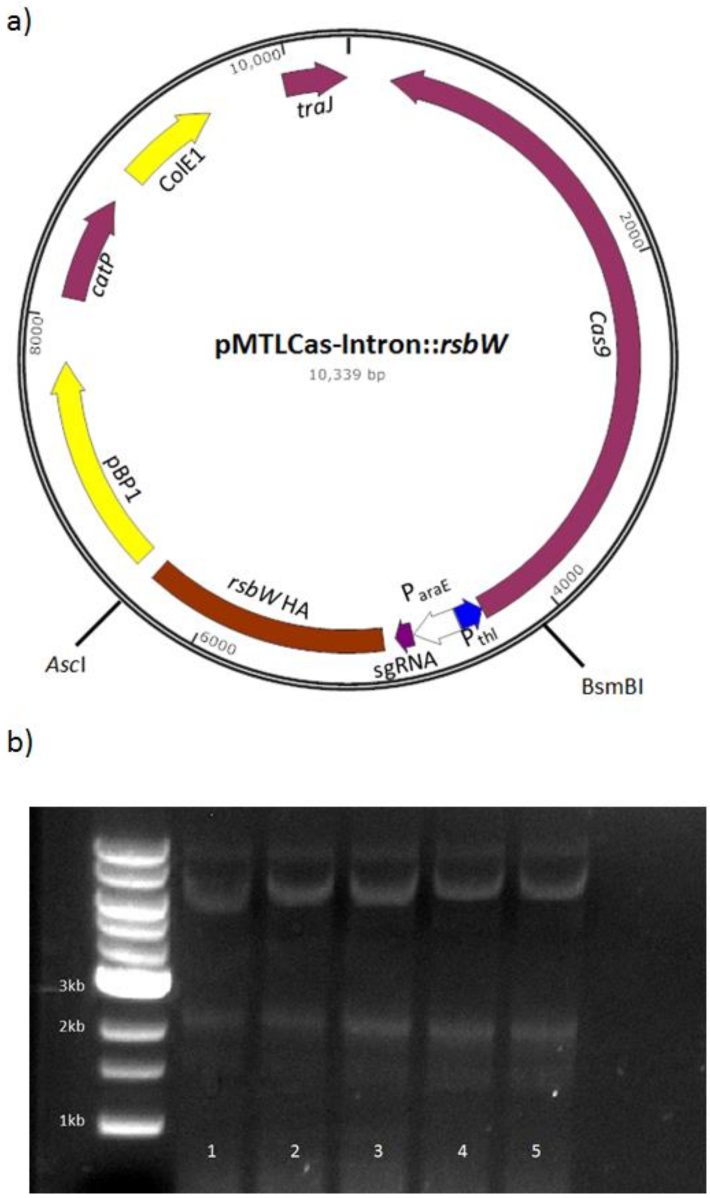


Figure 5.13 Plasmid pMTLCas-Intron::SNP*rsbW*. a) The pMTLCas-Intron::SNP*rsbW* plasmid was design to target the Cas9 nuclease to the group II intron found in ClosTron mutants. The plasmid contains the *catP* gene that confers thiamphenicol resistance; the replication region of the *E. coli* plasmid ColE1; the transfer function of the RP4 *oriT* region TraJ; the Cas9 nuclease from *S. pyogenes*; the P_{thl} promoter and the P_{araE} promoter from *C. acetobutilycum;* the retargeted sgRNA; the homology arm consisting of the *rsbW* gene found in *C. difficile* R20291 NM plus the 579bp region upstream and the 468bp region downstream of the gene; pBP1, replication region of the *C. botulinum* pBP1 plasmid, RepA and Orf2. b) Digestion of pMTLCas-Intron::*rsbW* with *Bsm*BI and *Asc*I; lane 1-5 represent plasmids isolated from five independent thiamphenicol-resistant colonies. The expected fragment (~2200bp) containing the retargeted sgRNA and the homology arm was detected in the agarose gel. At the extreme left, a 1KB DNA ladder is used as a weight marker.

The pMTLCas-Intron::SNPrsbw was introduced into *E.coli* NEB Sexpress through electroporation; subsequently, this strain was conjugated with the R20291 BW *rsbW::*CTermB strain for 24 h under anaerobic conditions. The growth was harvested and plated in BHIs plates supplemented with thiamphenicol (15µg/ml), cefoxitin (8µg/ml) and cycloserine (250µg/ml). The colonies that grew on the plates were restreaked in the same media and screened by PCR for the loss of the intron. Unfortunately, none of the colonies screened after the first restreak showed the loss of the plasmid. The colonies were restreaked two more times in the same media and screened once more for the recombinants. This second screening also showed no recombinant colonies. This results suggest that the group II intron sequence is not being cleaved by the CRISPR/Cas9 system, as the colonies screened show the group II insertion in the gene.

5.3.3 SigB role in the regulation of virulence factors

In many Gram-positive bacteria, SigB plays an important role in the regulation of oxidative stress and virulence (Guldimann *et al.*, 2016). Having observed the effect on toxin production in the *rsbW* insertional mutant made in *C. difficile* R20291 BW, we wanted to evaluate the role of the *sigB* gene in this strain, as R20291 BW does not possesses the SNPs found in the R20291 NM strain. As observed in the R20291 NM *rsbW*::CTermB mutant, the complementation of sigma factors mutants with vector plasmids is not the ideal form of complementation, as gene dosage from the plasmid might differ when compared to the expression of the gene from the genome. Currently, the use of the CRISPR/Cas9 system for the removal of the group II intron found in the R20291 BW *rsbW*::CTermB mutant is being investigated in our laboratory. However, this method has not been successfully developed yet, necessitating the development of an alternative methodology to complement the insertional mutants created in R20291 BW. In *C. difficile* R20291 NM, a procedure for making precise alterations in the genome using *pyrE*-based allelic exchange has been developed (Ng *et al.*, 2013). This systems allows for complementation studies to be undertaken at an appropriate gene dosage, as the complementation sequence is introduced directly into the genome. The generation of a Δ*pyrE* mutant in R20291 BW is the first step required for the use of this methodology. Once the Δ*pyrE* mutant has been created, the generation of an insertional mutant has to be generated in this Δ*pyrE* strain. The resulting strain is then restored with two different sequences, the native *pyrE* gene to evaluate the mutation in a clean background and the insertion of the *pyrE* gene along with the complementation sequence of the insertional mutant. As we have already generated an *rsbW* insertional mutant in *C. difficile* R20291 BW and we are investigating the CRISPR/Cas9 methodology to restore the native gene in the genome of that strain, we decided to evaluate this new approach in the generation of a *sigB* mutant in *C. difficile* R20291 BW.

5.3.3.1 Allelic exchange for the construction of a Δ*pyrE* mutant in R20291 BW.

The Allele-Coupled Exchange (ACE) method has been developed for the rapid insertion of heterologous DNA directly into the genome (Heap *et al.*, 2012). The ACE method is based in the occurrence of two cross-over events through homologous recombination. The *pyrE* gene is part of the pyrimidine biosynthesis and it has been used as a positive/negative selection marker as this gene is essential in the absence of exogenous sources; additionally, it renders 5-fluoro-oratate toxic to the cells (Ehsaan *et al.,* 2016). One of the advantages of creating mutants in a *pyrE*$^-$ background strain is that mutants created by any means can be rapidly complemented into the genome concomitant to the restoration of the *pyrE* allele through ACE. The complementation of the gene in the genome offers considerable advantages over the complementation with plasmids as the latter option can result in anomalies due to gene dosage and the effects of added antibiotics.

A *pyrE*$^-$ mutant was created in the R20291 BW in order to produce a strain that can be easily complemented in the genome through the use of Allele-Coupled Exchange (ACE) as describe in Ng *et al.* (2013). The plasmid pMTL-YN18 (Fig. 5.14) carries two asymmetric homology arms, a 300bp left homology arm composed of a *pyrE* allele that lacks 50 nucleotides from the 5' end and 235 bp from the 3' end and a 1200bp right homology arm that comprises the downstream region of the *pyrE* gene. These homology arms flank a *lacZ* site that contains a multiple cloning site region. The generation of the Δ*pyrE* mutant in *C. difficile* R20291 BW was carried out as described in Chapter 2.

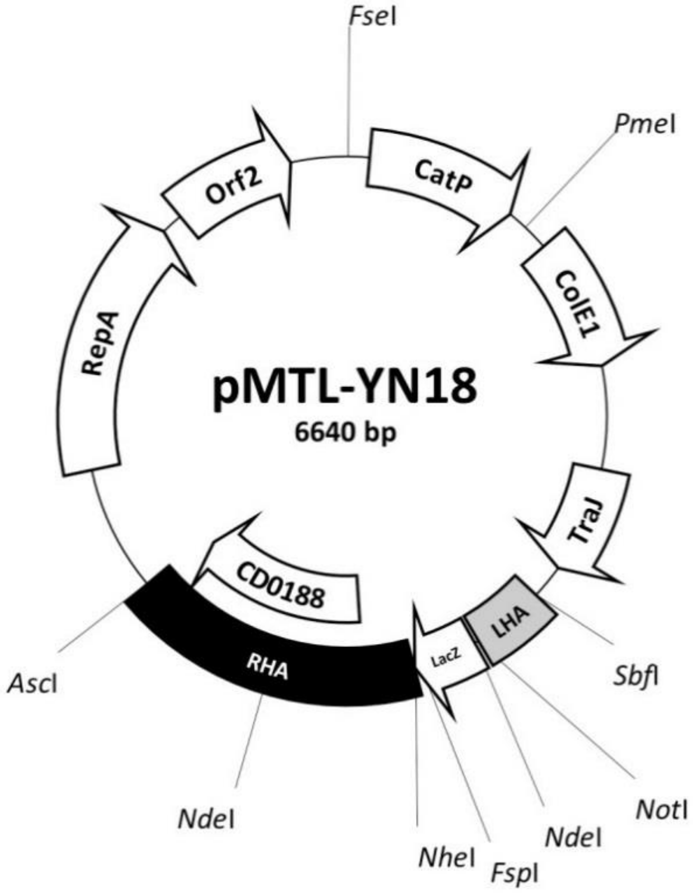


Figure 5.14 Representation of the pMTL-YN18 vector (Ng *et al.,* 2013). The plasmid contains the *catP* gene that confers thiamphenicol resistance; the replication region of the *E. coli* plasmid ColE1; the transfer function of the RP4 *oriT* region TraJ; a left homology arm consisting of 300bp internal fragment of the *pyrE* found in R20291 that lacks 50bp from the 5' end and 235bp from the 3'; the *lacZ* alpha fragment of the beta-galactosidase from *E. coli* containing a multicloning site; a right homology arm consisting of the 1200bp downstream region from *pyrE* and the replication region of the *C. botulinum* pBP1 plasmid, RepA and Orf2.

In short, the plasmid pMTL-YN18 was transformed into *E. coli* NEB Sexpress by electroporation. The resulting strain was conjugated with R20291 BW for 24 h in BHIs plates under anaerobic conditions. The growth was harvested and plated onto BHIs agar supplemented with thiamphenicol (15µg/ml) cycloserine (250µg/ml) and cefoxitin (8µg/ml). Fast growing colonies were restreaked onto the same media and colonies were grown in BHIs broth supplemented with thiamphenicol (15µg/ml) overnight. PCR screenings with the pyrE_F1 primer (sequence located in the genome) and M13 primer (sequence located in the plasmid) revealed the colonies in which a single cross-over event had occurred (Fig. 5.15a). The colonies were restreaked in BHIs plates without antibiotic to promote the occurrence of the second cross-over event and the loss of the plasmid. After 5 days the growth was harvested and plated in minimal media containing FOA (2mg/ml) and Uracil (5µg/ml). Only colonies that carry an inactivated pyrE gene are able to survive, as FOA becomes toxic if an active *pyrE* gene is present. Five FOA-resistant colonies that had lost their resistance to thiamphenicol were tested for the *lacZ* fragment in the genome with the forward primer pyrE_F1 and the lacZ_R1 primers. Only true Δ*pyrE* mutants are able to produce the expected fragment (ca. 550bp), while the wild type is not able to amplify any region (Fig. 5.15b). The R20291 BW Δ*pyrE* mutant was sent for a whole genome sequencing to verify that no other genetic variation is present in the genome compared to the parental strain. The comparison of the data received showed the presence of one SNP, a substitution of guanine for adenine, present in the position 3472928; however, this SNP is located in an intergenic region.

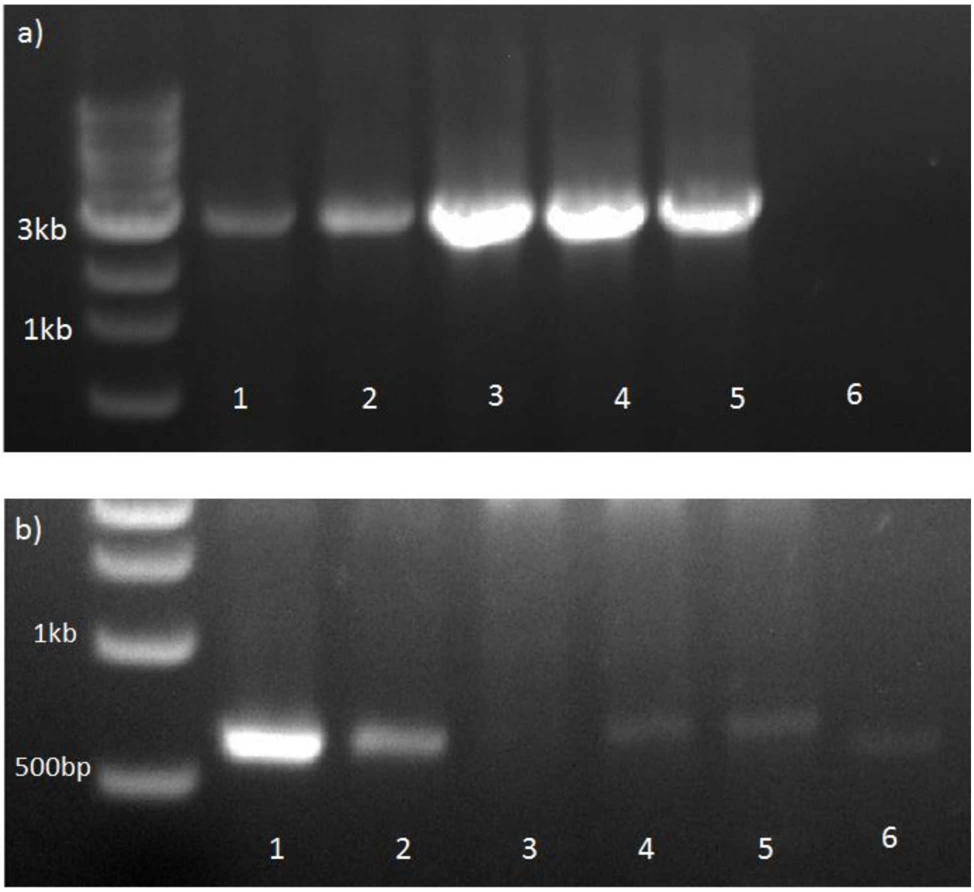


Figure 5.15. R20291 BW Δ*pyrE* mutants. a) Colonies where the first cross-over event occurred. Five fast-growing thiamphenicol-resistant colonies (lane 1-5) and the parental strain (lane 6) were screened with the M13_F and pyrE_F primer. The insertion of pMTLYN-18 plasmid in the genome yields a fragment of ca. 3kb. b) Six FOA-resistant thiamphenicol-sensitive colonies (lane 1-6) were screened to confirm the second cross-over event. The use of pyrE_F1 and lacZ_R1 produce a fragment only if the *lacZ* site has been inserted in the genome. The presence of the expected band confirms the change of the functional *pyrE* gene with the version present in the plasmid YN18 that is interrupted. The lane on the extreme left of the gel contains a 1KB DNA ladder.

5.3.3.2 Construction of *sigB* insertional mutants

As seen in the *rsbW* mutants, the *sigB* operon plays a role in the regulation of toxins and spore production in *C. difficile* R20291. An insertional *sigB* mutant was created using the ClosTron technology (Heap *et al* 2007) in the R20291 BW Δ*pyrE* strain. The retargeted ClosTron plasmid was designed to insert the group II intron into the *sigB* gene in a sense orientation at the position 192bp from the start of the open reading frame (ORF), resulting in the *C. difficile* R20291 BW Δ*pyrE sigB*::CTermB strain. To verify that the intron had inserted at the intended location, PCR screening were used to confirm the site of insertion using the sigB_F1 and EBS universal primers (Fig. 5.16). The expected fragment (ca. 400bp) can only be amplified if the insertion is present at the intended location.

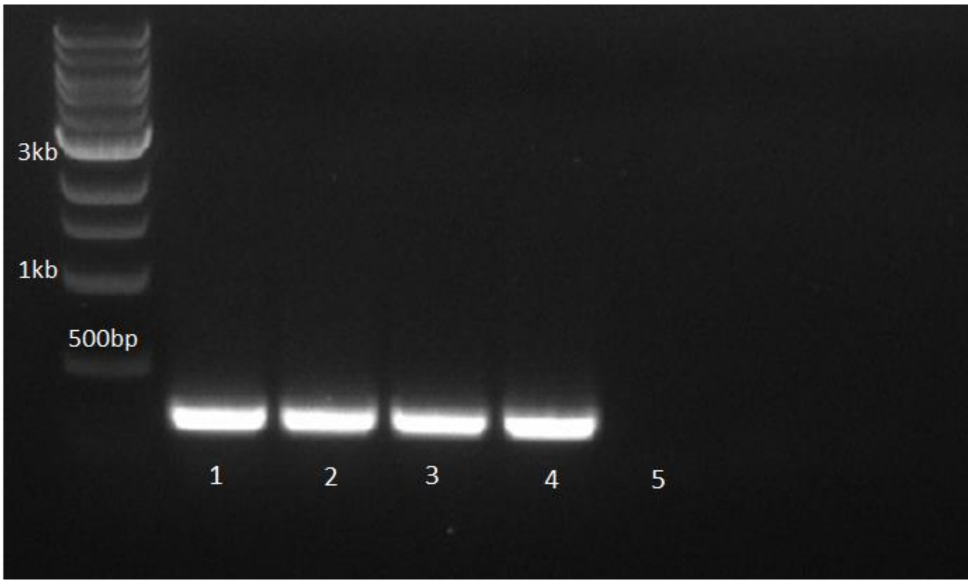


Figure 5.16 *sigB* insertional mutants in R20291 BW Δ*pyrE*. PCR screenings confirming the insertion of the group II intron into the *sigB* gene. Four independent lincomycin-resistant colonies (lane 1-4) and the parental strain (lane 5) were screened using the sigB_F1 and EBS universal primers to amplify the junction between the gene and the insertion. The lane on the extreme left of the gel contains a 1KB DNA ladder.

5.3.3.3 Complementation of R20291 BW Δ*pyrE* *sigB*::CTermB using the *pyrE* allele.

In order to evaluate the function of *sigB* in a clean background, the *pyrE* gene was restored using the plasmid pMTL-YN2. This plasmid is similar to the pMTL-YN18 plasmid except that the left homology arm consists of a 300bp *pyrE* allele that lacks the first 50 nucleotides from the ORF and the plasmid lacks the *lacZ* region between the homology arms. The plasmid pMTL-YN2C is a variation of the pMTL-YN2 plasmid. This version is similar to the parental plasmid except that contains an additional segment of DNA in the middle of the homology arms containing the transcriptional terminator (T1) of the *fdx* gene from *C. pasteurianum*, the *lacZ* gene and a multicloning site region that can be used to complement a desire gene into the genome (Ng *et al.*, 2013). In order to complement the *sigB* gene in the *sigB* insertional mutants, a complementing region containing the 300bp region upstream the *rsbV* gene and the *sigB* gene was created. The primers sigB_promoter_F1 and sigB_SOE_R1 were used to create the fragment containing the putative promoter region that presents a *Not*I restriction site at the 5' end; primers sigB_SOE_F1 and sigB_R1_XbaI were used to amplify the native *sigB* gene with an *Xba*I restriction site at the 3' end. SOE-PCR was used to merge the fragments and create the complementing sequence containing the putative promoter region and the *sigB* gene. The fragment was sent for sequencing to verify its sequence. The plasmid YN2C and the SOE fragment were digested with the restriction enzymes NotI and XbaI, purified and ligated to create the complementation plasmid pMTL-YN2C::promsigB.

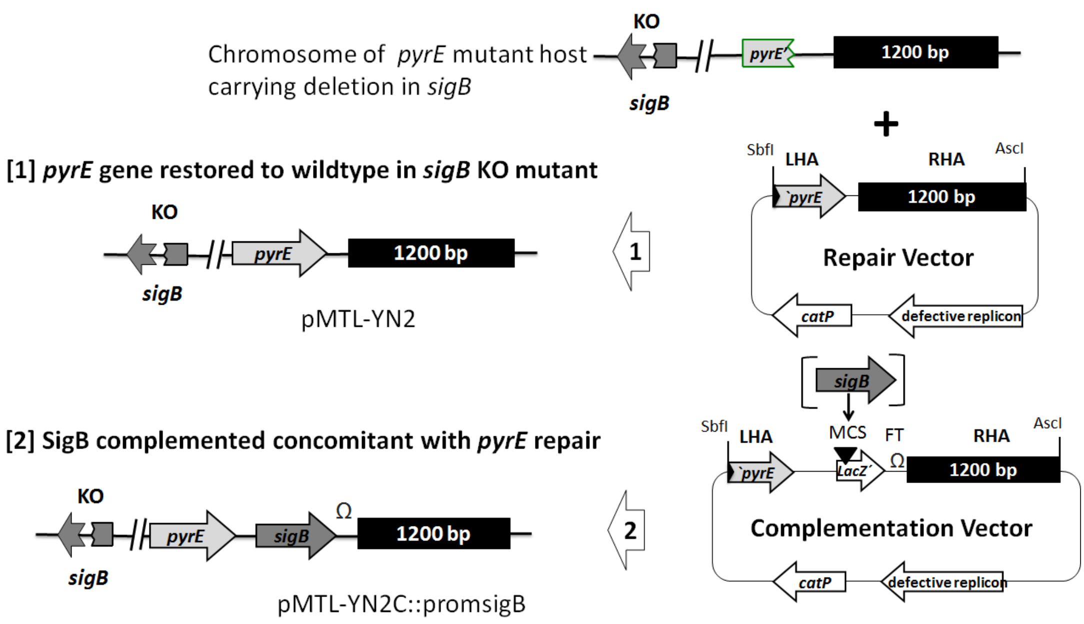


Figure 5.17 PyrE ACE correction vectors for *C. difficile* R20291 BW Δ*pyrE*. The pMTL-YN2 plasmid [1] is use to restore the *pyrE* gene into the genome. The pMTL-YN2C::promsigB plasmid [2] contains the *sigB* gene inserted in the multiple cloning site and the *sigB* gene is restored along with the *pyrE* gene allele.

The plasmids pMTL-YN2 and pMTL-YN2C::promsigB (Fig. 5.17) were transformed in the *E. coli* NEB express conjugative donor and the plasmid were conjugated in the R20291 Δ*pyrE sigB*::CTermB strain as described in Chapter 2. The colonies that have restored the *pyrE* gene are able to grow on minimal media without the addition of uracil in the media. Six random colonies that were able to grow on minimal media without uracil and had lost their resistance to thiamphenicol were selected and screened by PCR. The colonies restored with the pMTL-YN2 plasmid were tested using the pyrE_F1 and lacZ_R1 primers; the fragment cannot be amplified in the cells that have restored the native *pyrE* gene as a result of the removal of the *lacZ* region from the genome, resulting in clone R20291 BW *sigB*::CTermB (BW *sigB* mutant) (Fig. 5.18a). The colonies restored with the pMTL-YN2C::promsigB plasmid were screened using the sigB_promoter_F1 and pyrF_R1 primers. The expected fragment can only be amplified if the complementation region has been inserted into the genome (Fig. 5.18b). The resulting strain was designated R20291 *sigB*::CTermB complemented (BW sigB mutant complemented). The PCR fragments were isolated and sent for sequencing to verify they had the correct structure.

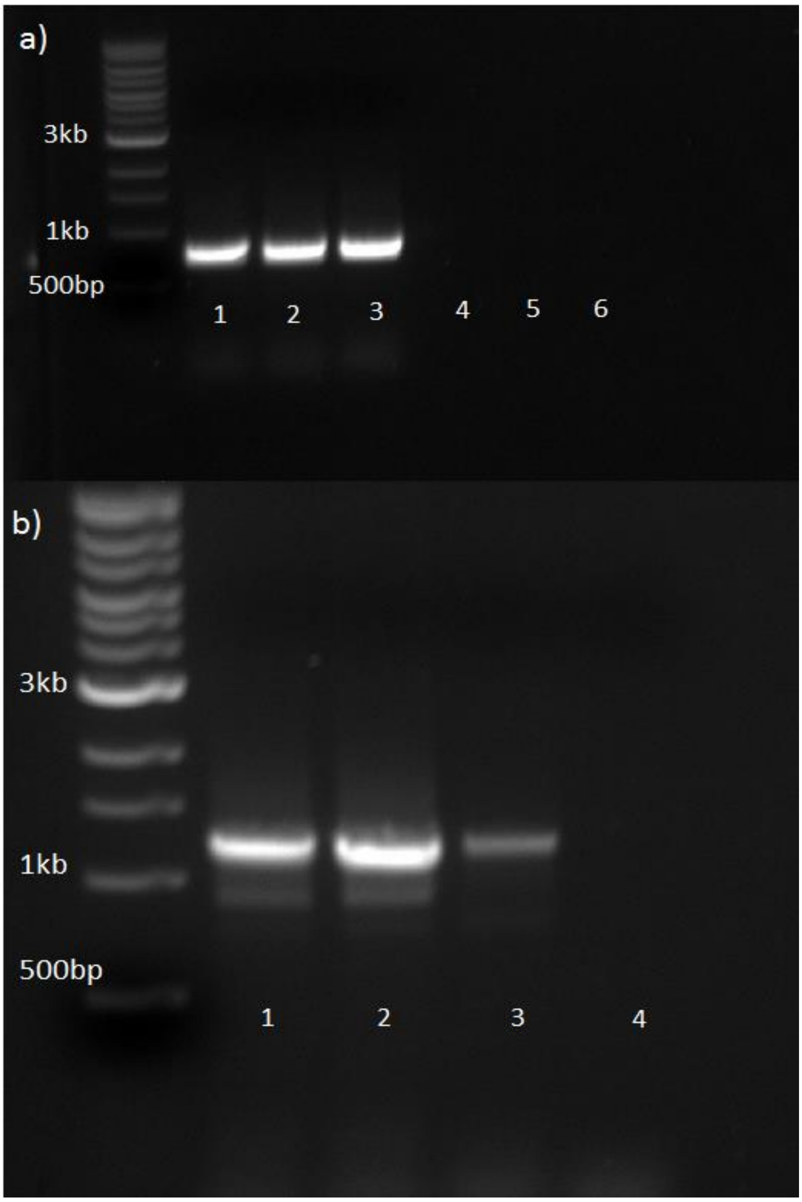


Figure 5.18 *sigB* insertional mutants complemented with the *pyrE* allele in R20291 BW. a) Restoration of the native *pyrE* gene that does not contain the *lac*Z site using the plasmid pMTL-YN2. Lane 1-3 represent three independent colonies screened with pyrE_F1 and pyrF_R1 to amplify the native *pyrE* gene and lane 4-6 represent the same colonies (respectively) screened with pyrE_F1 and lacZ_R1 to ensure the removal of the *lacZ* fragment found in the defective *pyrE* allele . b) Insertion of the *sigB* gene along the native *pyrE* allele. PCR fragment produced by the junction between the sigB gene and the downstream region of the *pyrE* gene; the fragment is only amplified if the *pyrE* gene has been restored along with the *sigB* gene using the pMTL-YN2C plasmid. Lane 1-3 represent three independent colonies screened with the sigB_promoter_F1 and pyrF_R1 primers and lane 4 represents the parental strain. The lane on the extreme left of the gel contains a 1KB DNA ladder.

5.3.3.4 Phenotypical analysis of R20291 BW *sigB*::CTermB insertional mutants.

It was been reported that *sigB* plays an important role in the regulation of spores in *B. subtilis* (Reder *et al.,* 2010). In order to test the effect of the insertion in the *sigB* gene, the *sigB* mutants were subjected to a 5 days sporulation assay to determine if the inactivation of this gene affects the regulation of spores in the cell (Fig. 5.19). The assays was performed as described in Chapter 2.

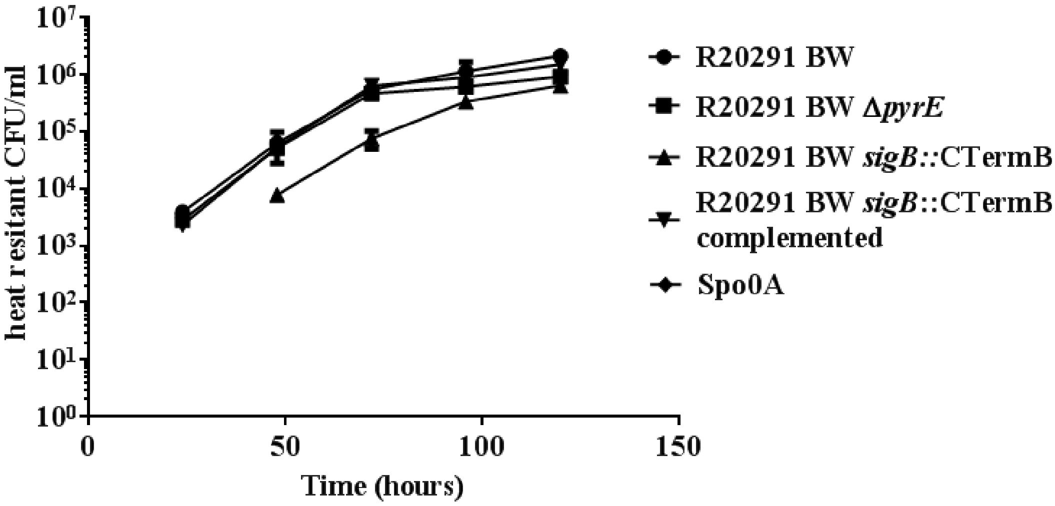


Figure 5.19 Production of heat-resistant CFU in *C. difficile* R20291 BW *sigB* mutant strains over 5 days. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Strains included are indicated in the accompanying key. Sporulation in the negative control (Spo0A, a *spo0A* mutant of R20291 NM) could not be measured. The detection limit was 50 CFU/ml.

The onset of sporulation in the BW *sigB* mutant was delayed compared to the wild type; the first spores in the BW *sigB* mutant can be detected at the 48 h time point, which contrasts with the parental strain, where spores could first be detected at 24

h. This delay in the onset of sporulation reveals that the *sigB* gene has a role in the regulation of sporulation.

The *sigB* mutants were subjected to a test to evaluate their ability to produce toxins over a 72 h time period, in order to evaluate if the sporulation pathway is not the only pathway affected by the insertion of the intron in *sigB*. The level of toxins produce by the different strains is shown in Fig. 5.20.

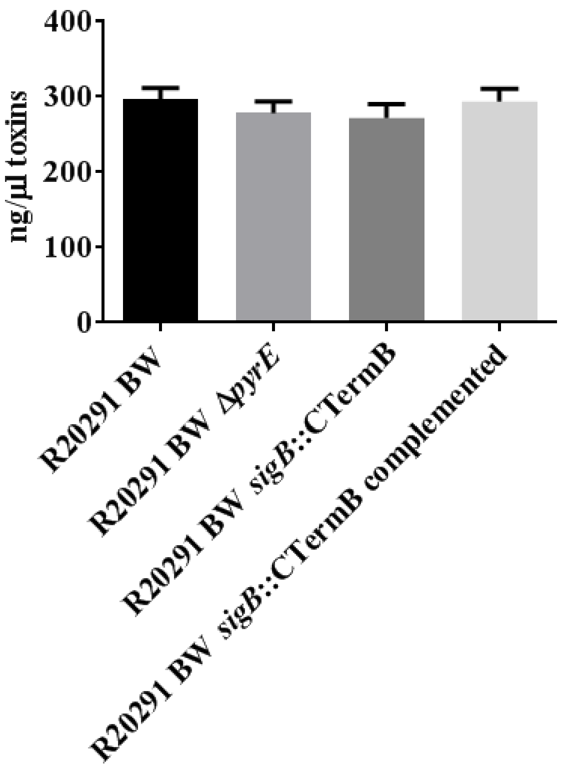


Figure 5.20 Toxin production of the R20291 BW *sigB* mutant strains after 72 h. The indicated strains were cultured in TY medium and the toxicity of supernatants tested at the indicated time point as described in Materials and Methods. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means. Asterisks indicate statistically significant differences (**, $P < 0.01$ by Student's *t* test) compared to the wild type.

Interestingly, the toxin levels of the different strains are similar to the parental strain even though the timing of sporulation is different in the *sigB* mutants. It seems that the regulation of toxins is not under the control of *sigB*. The motility of the *sigB* mutants was tested to evaluate if the interruption of *sigB* could affect the regulation of flagella as well as its effect on the production of biofilm. The experiments were performed as described in Chapter 2. The motility of the different strains did not show any significant difference compared to the wild type as all the strains had a similar performance in both assays, swarming and swimming (Fig. 5.21a and Fig. 5.21b). The interruption of the *sigB* gene also had no discernible effect in the production of biofilm by the different strain, as all the strain presented a similar behaviour after 5 days (Fig. 5.21c).

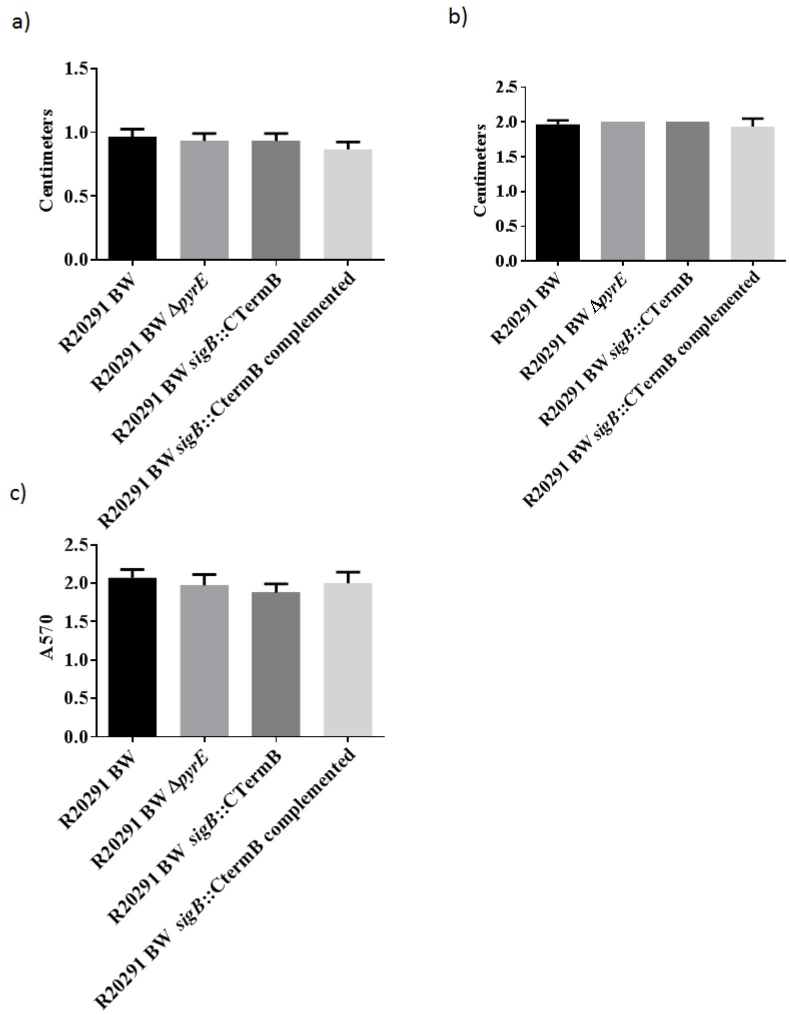


Figure 5.21 Motility and biofilm formation of the *sigB* mutants. Motility and biofilm production of the R20291 BW *sigB*::CTermB mutants. a) Swarming ability of the R20291BW *sigB*::CTermB strains in BHIs 0.4% agar plates. b) Swimming assay in BHIs 0.3% agar plate. c) Biofilm production of the R20291 BW *sigB*::CTermB strains after 5 days; the strains were grown in BHIs supplemented with 0.1M glucose. The symbols represent the averages of three independent experiments and error bars indicate the standard errors of the means.

In *S. aureus* and *B. subtilis* sigma factor B, also called general stress regulon, plays an important role in the environmental stress response (Horsburgh *et al.*, 2002; Nannapaneni *et al.*, 2012). Sigma B has been reported to confer oxidative resistance in *B. subtilis* and it has been shown that *sigB* mutants have an impaired oxidative stress response (Voelker *et al.*, 1995; Engelmann *et al.*, 1996). The ability of the *sigB* mutants to resist oxidative stress was tested using hydrogen peroxide as described in Chapter 2. In short, the strains were grown to mid log phase in BHIs broth under anaerobic conditions and were treated with 1.2 mM and 1.6 mM H_2O_2 for one hour. Cells were plated after the incubation time in BHIs plates. Untreated samples were also plated as a control group.

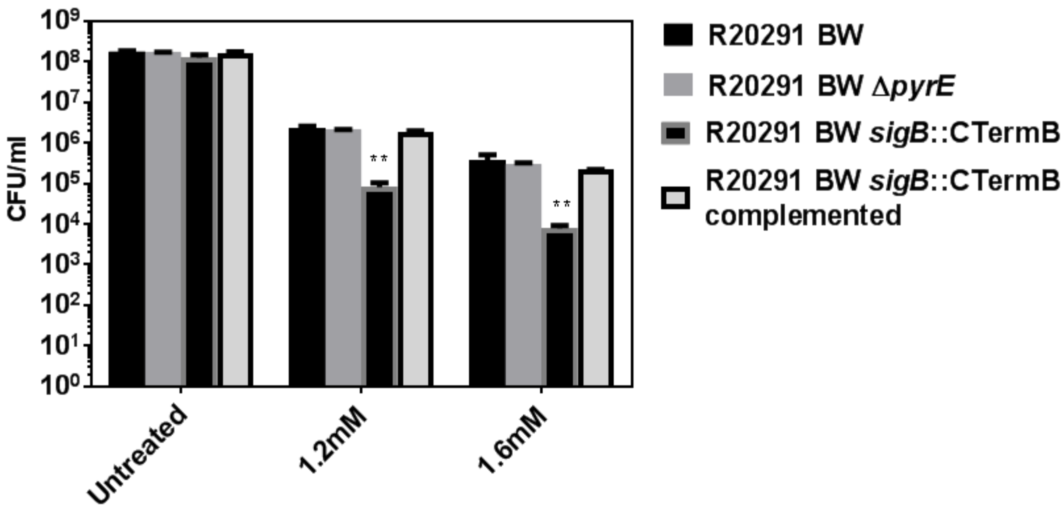


Figure 5.22 Viability of *C. difficile* R20291 BW *sigB* mutants during oxidative stress. Numbers of CFU/ml in strains treated with 1.4 mM or 1.6mM H_2O_2 for 1 h. The viable cell number was determined by plating before and after treatment with H_2O_2. 'Untreated' means numbers of CFU/ml before addition of H_2O_2. The bars indicate the averages of three independent experiments and the error bars indicate standard error of the means. Asterisks (**) indicate statistical significance ($p<0.01$). The detection limit for the assay was 50 CFU/ml.

As expected, the BW *sigB* mutant presented a lower amount of viable cells after the treatment with the oxidative agent compared to the wild type (Fig. 5.22). The complemented strain showed a similar performance to the wild type. The role of the sigma B factor in *C. difficile* R20291 BW seems to be linked to the regulation of resistance to oxidative stress, similar to the findings reported in *B. subtilis* (Engelmann *et al.,* 1996).

5.3.4 Discussion

The role of the SigB operon has been linked to the regulation of the *agr* operon in *S. aureus* (Bischoff *et al.,* 2001; Lauderdale *et al.,*2009) and the regulation of virulence in *L. monocytogenes* (Zhang *et al.,* 2013). In this work, we found that the R20291 NM strain has a different phenotype compared to the reference R20291 BW strain in terms of toxin production and sporulation. The R20291 NM strain possesses various unique SNPs, of which one is located in the *rsbW* gene. The SNP in the *rsbW* gene leads to an amino acid change in the position 82, replacing a glycine with a valine; this change is located in the predicted histidine-like ATPase domain. The structural difference between the two amino acids is the presence of an isopropyl group found in the valine structure (Fig 5.1c). This isopropyl group could affect the folding of the protein or the interactions of the predicted domain, compromising its function; however, it is unclear if the resultant variant protein has lost activity due to this change. To determine the role of the *rsbW* gene in the R20291 NM strain, we created an insertional mutant in its *rsbW* gene (R20291 NM *rsbW::*CTermB) which was shown to affect toxin production and sporulation in the strain containing the intron. The observed phenotype suggests that the change in the amino acid sequence does not result in the inactivation of the protein, as the phenotype seen in the mutant would otherwise be similar to the parental strain. These data indicate that the SigB operon is involved in the regulation of toxin and spore production in R20291 NM. Complementation of the *C. difficile* R20291 NM *rsbW*::CTermB mutant strain with the *rsbW* gene found in the R20291 NM strain restored the time of initiation of sporulation to that of the parental strain. Nevertheless, the number of spores produced by the complemented strain was higher than that seen in the wild type. Sigma factors are tightly regulated as they are crucial for the regulation of different physiological processes in the cell (Kazmierczak *et al.,* 2005); therefore, it is possible that the expression of the *rsbW* gene from the complementation plasmid is at a higher rate compared to the expression of the gene when localised in the genome, a consequence of gene dosage. Furthermore, the use of the pMTL84151 as a vector for complementation has presented some disadvantages in some of our strains affecting the expression of toxins. Due to these findings, new methodologies for complementation, such as the use of the *pyrE* allele to repair the gene into the genome and the CRISPR/Cas9 system are currently under investigation in our laboratory.

The use of allelic exchange for the restoration of the *rsbW* SNP in the R20291 NM strain was investigated. Normally, ACE relies on two recombination events where a single crossover event leads to integration of the plasmid into the genome through the recombination at one homology arm followed by a second recombination event that promotes the excision of the plasmid from the genome (Cartman *et al.,* 2012). The PCR screenings showed a faint band using a flanking primer for the *rsbW* gene and a primer that is located inside the vector, meaning that the expected band can only appear if there is integration of the plasmid in *rsbW*. However, the isolation of single-crossover colonies from these cultures was not possible, as single colonies screened from these cultures did not show the expected band. These results reveal that there is a mixed population in the cultures, where only a small fraction of the cells have integrated the plasmid into the genome. It is possible that the integration of the plasmid in the *rsbW* gene creates an unsTable intermediate that is detrimental to cell in the media containing antibiotics, which could explain the low single-crossover population and the failure of the isolation of pure single crossovers.

The various SNPs present in the R20291 NM strain have been shown to greatly affect the regulation of virulence factors in this strain, which makes it difficult to evaluate regulatory pathways concerned with sporulation and toxin production, for instance, in a R20291 NM background. Previously, the delivery of plasmids into the R20291 BW strain was proven difficult, as no transconjugants could be achieve using *E. coli* CA434 strain as a conjugative donor. The use of *E. coli* NEB Sexpress strain as a donor for conjugations using R20291 BW as the recipient significantly improved the frequency of DNA transfer. This provided the opportunity to create insertional mutants in R20291 BW and evaluate the effect of the interruption in the *rsbW* gene in a clean background.

The *rsbW* insertional mutant created in the R20291 BW strain affected the regulation of toxin. RsbW acts as an anti-sigma B factor in the cell, forming a complex that inhibits the interaction of the SigB factor with the core RNA polymerase (Benson *et al.*, 2013). The RsbW anti-sigma factor is regulated by the anti-anti-sigma factor RsbV (Dufour *et al.*, 1994), which interacts with RsbW, forming a complex that prevents RsbW from inhibiting the activity of SigB. In early stages of growth in *B. subtilis*, the RsbV protein is phosphorylated by the kinase activity of RsbW, rendering RsbV inactive; after exposure to stress, the RsbV protein is dephosphorylated by RsbU and RsbP, two specific PP2C-type phosphatases (Yang *et al.,* 1996; Vijay *et al.,* 2000). Once dephosphorylated, RsbV is able to interact with RsbW to form the RsbV/RsbW complex, liberating the SigB protein from the RsbW/SigB complex. The insertional inactivation of RsbW in R20291 BW leads to elevated levels of toxins. Following the model seen in *B. subtilis*, the inactivation of RsbW would result in the lack of phosphorylation of RsbV, which would lead to an early activation of the SigB protein. In *Listeria monocytogenes*, the role of RsbV has been linked to the positive regulation of several virulence genes (Zhang *et al.,* 2013). The early activation of the RsbV gene could be the reason for the increase in toxin production in the R20291 BW *rsbW* mutant, as this would provide more time for the cell to express virulence factors regulated by this protein. The creation and evaluation of an RsbV insertional mutant in R20291 BW presents a good opportunity to further investigate the role of this gene.

The CRISPR/Cas9 system has been successfully used to create mutants in other clostridia species (Wang *et al.,* 2015; Li *et al.,* 2016; Huang *et al.,* 2016). The pMTLCas-Intron::SNPrsbw created to target the group II intron was successfully introduced into *C. difficile* R20291 BW *rsbW*::CTermB. However, the results indicate that the group II intron is not being cleaved, as the colonies screened demonstrated that the intron was still present in the *rsbW* gene. This also revealed that the plasmid is being correctly replicated and expressed in the cell, as the colonies containing this plasmid are resistant to thiamphenicol. The promoter used for the expression of the Cas9 nuclease, that of the thiolase gene from *C.*

acetobutylicum, has been successfully used in our laboratory for the overexpression of genes in *C. difficile,* most notably the *ermB* gene present in the ClosTron RAM. The P_{araE} promoter used to express the sgRNA in the CRISPR/Cas9 plasmid constructed in this study is the same as that used in *C. ljungdahlii*, where it enabled effective targeting of Cas9 to the intended target genes (Huang *et al*., 2016). However, it is not clear yet if this promoter is as effective in *C. difficile* as in *C. ljungdahlii*. The inability to obtain a mutant would suggest not. If the sgRNA is not being correctly expressed in the cell, this could explain the lack of Cas9 nuclease-mediated cleavage of the targeted sequence in the intron, as the sgRNA guides Cas9 to its target sequence. The substitution of the P_{araE} promoter with alternatives known to express efficiently in *C. difficile* should overcome this deficiency. Once such promoter has been identified, it should prove possible to replace the native SNP-containing *rsbW* gene of R20291 NM with that of R20291 BW. CRISPR/Cas9 systems cannot currently be used to effect single base change alterations to target genes, as the target sequence will also be present in the new allele and so will also be cleaved; this approach would represent a general strategy for accomplishing such SNP changes in any gene. Thus, following the creation of a ClosTron mutant in the gene target, the modified CRISPR/Cas9 vector described here could be used to replace the intron-inactivated mutant allele with a functional copy of the gene containing any desired amino acid modification. One caveat of this approach is that the gene under investigation must not be an essential gene. That is to say, cells in which it is inactivated must remain viable.

An alternative strategy which could be used to implement CRISPR genome editing capability in *C. difficile* would be to make use of the endogenous CRISPR system. This approach has been successfully reported in *Clostridium pasteurianum* (Pyne *et al.,* 2016). In this approach it is only necessary to introduce a plasmid carrying the sgRNA together with the editing template (knock-out cassette), as the approach is reliant of the activity of a native Cas nuclease. This means that the plasmids to be constructed are more easily assembled and are considerably smaller in size. To implement such an approach it is necessary to know the target sequence of the native Cas nuclease. Fortunately, a recent analysis of *C. difficile* R20291 has identified the likely sequences involved (Boudry *et al*., 2015).

The altered sporulation phenotype observed in the R20291 BW *sigB* mutant implicates sigma factor B in the regulation of sporulation in *C. difficile*. In *B. subtilis*, SigB positively regulates the expression of Spo0E, which dephosphorylate Spo0A, deactivating the sporulation master regulator and preventing the initiation of sporulation (Reder *et al.*, 2012). These findings show that *sigB* negatively regulates sporulation in *B. subtilis*. In *C. difficile* R20291 BW, inactivation of *sigB* leads to a delay in the onset of sporulation, suggesting that, contrary to what has been observed in *B. subtilis*, the SigB protein in *C. difficile* R20291 BW positively regulates the onset of sporulation, although SigB is not essential for the sporulation process. The regulatory pathway of *C. difficile* seems to incorporate different intracellular signals that regulate the activation of Spo0A as *C. difficile* does not possess several key regulatory components of the sporulation regulon present in *B. subtilis* (Edwards & McBride *et al.,* 2014).

Furthermore, the *sigB* operon found in *B. subtilis* consists of eight genes, *rsbR, rsbS, rsbT, rsbU, rsbV, rsbW, sigB* and *rsbX* (Wise & Price, 1995), while the locus found in *C. difficile* consists of only five gene, *CD0007, CD0008* and homologues to *rsbV*, *rsbW* and *sigB* (Fig. 5.21). These differences in the regulatory components may be responsible for the different function of the SigB protein in the regulation of sporulation in *C. difficile*.

The *sigB* operon in *S. aureus* consists of 6 genes; *mazE, mazF, rsbU, rsbV, rsbW, sigB* (Fig. 5.21); *mazE* and *mazF* are two small genes that encode for a toxin-antitoxin system (Fu *et al.*, 2007); *rsbU* codifies a protein phosphatase 2C (PP2C) capable of dephosphorylate RsbV, while *rsbV*, *rsbW* and *sigB* genes encode the "partner-switching" module between RsbV, RsbW and SigB that plays an important role in the regulation of the general stress response in the cell (Pane-Farre *et al.*, 2009). The *sigB* operon in *C. difficile* consists of 5 genes; *CD0007, CD0008* and homologues to *rsbV, rsbW* and *sigB (CD0009, CD0010* and *CD0011* respectively). CD0007 and CD0008 are hypothetical proteins that have not been described yet. The only conserved homologues to the *sigB* operon present in *S. aureus* are RsbV, RsbW and SigB. The oxidative stress response seems to be conserved in both species as inactivation of *sigB* impaired the oxidative stress response in the *C. difficile* R20291 *sigB*::CTermB mutant, confirming the hypothesis that the SigB protein is part of the stress response regulation, as seen in *B. subtilis* (Engelmann & Hecker, 1996, Hecker *et al.,* 2007), *L. monocytogenes* (Ferreirra *et al.*, 2001) and *S. aureus* (Bischoff *et al.,* 2001); however, *C. difficile* lacks an RsbU homologue in its genome, raising the question of how RsbV is activated in the cell, as RsbU regulates the phosphorylation state of RsbV in the *S. aureus* model, which is the decisive factor that determines the state of SigB (Vijay *et al.,* 2000); when RsbV is phosphorylated by RsbW this protein is not able to form a complex with RsbW and thus RsbW forms a complex with SigB. When *S. aureus* is subjected to environmental stress, RsbU dephosphorylates RsbV and consequently allows SigB to be released from the RsbW/SigB complex, turning SigB into its activated state. The genome of *C. difficile* R20291 harbours two PP2C phosphatases, SpoIIE and CD2685. SpoIIE is thought to be part of the sporulation pathway, as *spoIIE* is among several genes under the control of Spo0A (Abecasis *et al*., 2013). Nevertheless, the target of CD2685 has yet to be described; the characterisation of a *CD2685* insertional mutant could offer the opportunity to evaluate the role of this phosphatase and determine if it is involved in the regulation of the *sigB* operon. SigB also plays an important role in the biofilm formation in *S. aureus,* where it was found that the inactivation of *sigB* leads to a reduce biofilm formation in the mutants when compared to the parental strain (Atwood *et al.*, 2015). The *sigB* operon regulates the biofilm formation in the cell indirectly by a mechanism involving the *agr* operon, as SigB downregulates the expression of *agr,* which regulates the expression of nucleases, proteases and haemolysins that destabilise the biofilm (Guldimann *et al.,* 2016). However, inactivation of *sigB* in R20291 BW did not affected the biofilm formation of the mutants; indicating that the SigB regulon is not involved in the formation of biofilm. It has also been observed that SigB strongly represses the expression of toxins in *S. aureus* through two distinct regulatory pathways dependent upon *sarA* and *agr* (Andrey *et al.,* 2015).

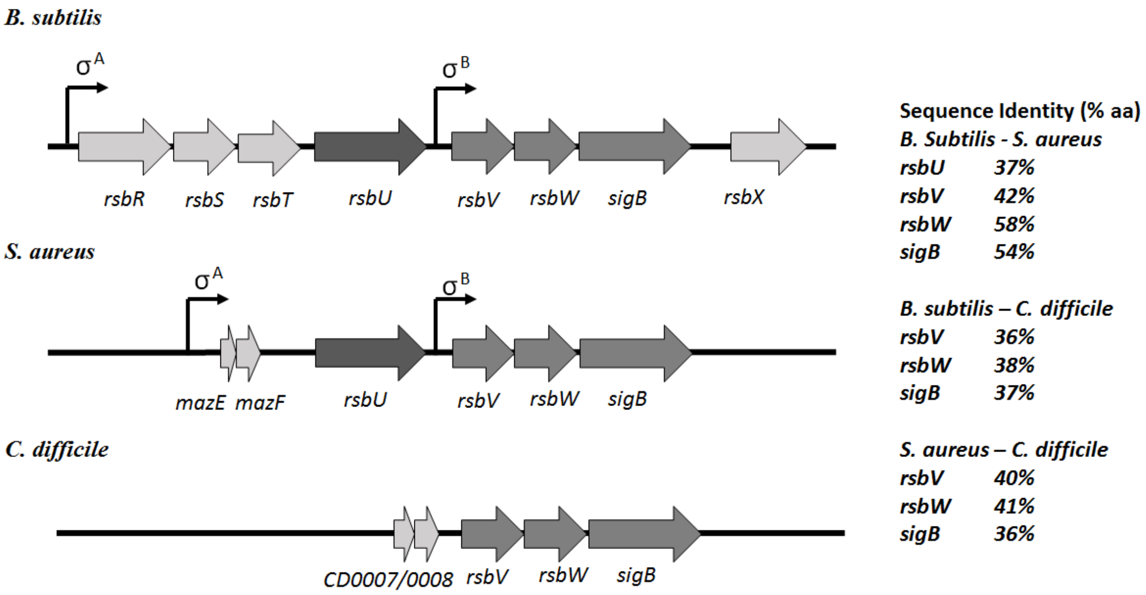


Figure 5.23 *sigB* operons found in *B. subtilis, S. aureus* and *C. difficile.* Comparison of the genes found in the *sigB* operon between *B. subtilis, S. aureus* and *C. difficile*. The identity of the sequences are shown on the right. The black lines shown in the *B. subtilis* and *S. aureus* operons represent experimentally demonstrated transcriptional start sites (Pane-Farré *et al.*, 2009). No transcriptional start sites have been experimentally demonstrated in *C. difficile* to date.

Nevertheless, the toxin phenotype of the R20291 BW *sigB* mutant was similar to the one observed in the parental strain. The results observed in Chapter 3 and the study presented by Darkoh *et al.* (2016) show that the *agr1* operon and the *agrA* gene located in the *agr2* operon are involved in the production of toxins in *C. difficile*; in *S. aureus*, several virulence factors (toxins, biofilm, proteases, nuclease) are indirectly controlled by SigB through the regulation of the *agr* operon. The result of the toxin assay in the R20291 BW *sigB* mutant suggests that the SigB protein present in *C. difficile* R20291 BW does not interact with the *agr* loci found in this strain. These findings also imply that the toxin synthesis is not associated with sporulation, as the R20291 *sigB* mutant presents a delayed sporulation but no difference in the toxin levels produced by the mutant, contrary to what has been observed in *C. perfringens*, where the production of the enterotoxin CPE is controlled by sporulation specific sigma factors SigE and SigF (Harry *et al.*, 2009).

During the course of this work, the advantage of undertaking complementation studies based on integration of the complementing gene (*sigB*) into the genome at the *pyrE* locus was emphasised. The work presented in this chapter provides evidence that the SigB operon plays an important role in the regulation of toxins and sporulation in *C. difficile* R20291 *in vitro*.

Chapter 6

Conclusion

6.1 Key findings of this study

C. difficile has become the leading cause of antibiotic-associated diarrhoea in hospitals around the world (Baker *et al.* 2010; Hensgen *et al.* 2012; Honda & Dubberke 2014). Its two principle virulence factors are the production of the toxins (TcdA and TcdB) that mediate the disease and the formation of endospores, the main agent of disease transmission. The means by which these processes are regulated is, therefore, of considerable importance. A relatively common system for regulating multiple processes in Gram-positive bacteria is through Quorum Sensing and the *agr* locus. The role of the *agr* operon in the regulation of virulence factors, such as spores and toxins, has been reported in *S. aureus* and several clostridia species, among others (George & Muir, 2007; Chen *et al.,* 2012; Ohtani *et al.,* 2009; Steiner *et al.,* 2012). The *agr* operon found in *S. aureus* consists of four genes, *agrA, agrB, agrC* and *agrD* (Novick & Geiser, 2008). The epidemic *C. difficile* R20291 strain possesses 281 genes , including an additional *agr* operon, that are not present in the first sequenced *C. difficile* strain, designated *C. difficile* 630, which only carries a partial *agr* operon consisting of the *agrB1* and *agrD1* genes (STabler *et al.,* 2009). We initiated our investigation by focusing on the role of the *agr2* operon in the regulation of toxin production and sporulation.

6.1.1 Role of the *agr* loci in the regulation of virulence factors.

The introduction of the *agr2* operon into strain 630Δ*erm* had no impact on sporulation or the production of toxins in this strain. This indicates that the increased virulence observed in R20291 and several other RT-027 strains is not simply a consequence of the presence of the *agr2* operon in the genome. Indeed, the non-epidemic RT-027 CD196 strain harbours the *agr2* locus but does not exhibit the increased virulence observed in the R20291 strain (STabler *et al.*, 2009). Even though strain CD196 also possesses the *agr2* locus in its genome, it lacks 47 coding sequences present in the R20291 strain. One or more of the additional genes present in R20291 may be responsible for the differences seen between the historic PCR ribotype strain CD196 and the epidemic strain R20291 in terms of virulence. Alternatively, subtle changes in genes common to both strains, and in particular those involved in regulatory processes, may have altered the phenotype of R20291 compared to CD196 in terms of virulence. Through whatever mechanism, it seems likely that the greater virulence of strain R20291 is a consequence of the complex interaction between different pathways and regulatory pathways rather a sole consequence of the presence of the *agr2* operon in its genome.

The studies described in this thesis have shown that both of the *agr* loci in strain R20291 NM influence the regulation of toxin and spore production. These findings are consistent with observations made in other clostridia, including *C. perfringens* and *C. botulinum* (Ohtani *et al.*, 2009; Cooksley *et al.*, 2010), where the *agr* system regulates different virulence factors. Although *C. difficile* harbours homologues to the predicted AIP, the work done with synthetic AIPs demonstrated that the cycling of the putative AgrD1 and AgrD2 peptides based on the *S. aureus* model is not enough to produce any effect in the cell, suggesting that other modifications may take place in order to create the active AIP. However, what became obvious during the course of this work was the fact that the R20291 strain being studied at Nottingham was significantly different from strains being worked on elsewhere (London School of Hygiene and Tropical Medicine and Novartis in Siena, Italy). These differences manifested themselves as significantly dissimilar behaviour in terms of sporulation, toxin production, motility and biofilm formation. Whilst it is clear that these differences must be associated with the distinct mutational differences (SNVs and Indels) discovered between the two strains through NGS, it did not prove possible to conclusively associate any one particular mutation to the phenotypic changes.

Having established that *C. difficile* R20291 BW was a more appropriate model for studying the two *agr* operons, their involvement in virulence factor regulation was re-visited. The *agrD1* gene was found to positively regulate toxin and spore production, as an insertional mutant showed a reduction in both compared to the parental strain. In contrast, inactivation of either *agrD2* or *agrB2* genes led to increased toxin production, indicating that they were most likely involved in the negative regulation of toxin production as opposed to the positive regulatory role of the *agr1* system. These results showed that the effect of the insertional interruption of the *agr* genes was drastically different depending on which strain was used for the mutant construction. Even though the *agrD1* mutant created in

R20291 BW is not equivalent to the *agrB1* mutant created in R20291 NM, both genes are consider to form part of the same regulatory pathway. Thus, AgrB1 processes the AgrD1 propeptide to derive the active AIP. It follows that inactivation of either gene would be expected to prevent production of AIP and, as a consequence, an *agrB1* and *agrD1* mutant would be expected to share a similar phenotype. . The interruption of the *agrB1* gene in R20291 NM delayed the onset of sporulation but promoted the production of higher toxin levels. However, whilst inactivation of *agrD1* in R20291 BW also caused a delayed in the onset of sporulation, the levels of toxin produced was reduced. The latter observation is in agreement with the phenotype of an *agrBD1* mutant created using allelic exchange by Darkoh *et al.* (2016).

The insertional interruption of the *agrB2* and *agrD2* genes in R20291 BW did not affect the onset of sporulation but resulted in an increase in toxin production. These data are in agreement with the observed phenotype of the *agrBD2* mutant created by Darkoh *et al.* (2016). The increase in toxin production in the *agrB2* and *agrD2* mutants of R20291 BW was, however, in contrast to the results observed with the *agrB2* mutant of R20291 NM, in which the onset of sporulation was brought forward and toxin expression levels were reduced. . These differences fully justified the decision to recreate the *agr* mutants in R20291 BW and suggest that the different phenotype of the R20291 NM mutant is a consequence of the SNPs and Indels in its genome. .

In the study of Martin *et al.* (2013), inactivation of *agrA* (encoding the AgrA response regulator) in R20291 BW resulted in a reduction in toxin production.

Inactivation of *agrA* should lead to a reduction in the production of the AgrD2 peptide, and as consequence the same phenotype as seen here – increased toxin production. The reason for this anomaly is not clear. However, as the *agr1* operon lacks any obvious equivalents to the *agrCA* two-component system within its immediate vicinity it may be the case that *agrB1*/*agrD1* genes also under the control of the *agrCA* two-component system of the *agr2* locus. Inactivation of *agrA* may, therefore, also be affecting the production of AgrD1, which is a positive regulator of toxin production. Mutation of *agrD1* would negatively impact on toxin production.

6.1.2 Genetic variations among *C. difficile* R20291 strains.

The genetic variations found in the *C. difficile* R20291 strain used in Nottingham (R20291 NM) were shown to greatly affect the phenotype of the strain compared to other two R20291 strains used in different laboratories, namely the Novartis laboratory (Sienna, Italy) and the London School of Hygiene and Tropical Medicine (London, UK). The SNPs within the genome of R20291 NM are located in *rbsK*, a ribokinase belonging to the *pfkB* family, *CD*R20291*_2456*, a putative two-component sensor histidine kinase, *rsbW*, an anti-sigma-B factor, and *vncR*, a two-component response regulator Tn*1549*-like. Given the role of the sigma B in other bacteria and given the phenotypes seen here in terms of the effects on sporulation and toxin production of the *rsbW* ClosTron mutant created in *C. difficile* R20291 NM (Chapter 5), it is tempting to lay the blame for the phenotypic changes seen in strain R20291 NM on the SNP in *rsbW*. However, such assumptions are fraught with danger. In the recent analysis of the two difference variants of the erythromycin sensitive derivatives of strain 630 (630E and 630Δ*erm*), it was generally assumed that the frameshift in the topoisomerase gene (essential in some bacteria) in strain 630E most likely explained its somewhat attenuated phenotype, and that the inversion of the promoter region in front of its flagella operon was the cause of its lack of motility (Collery *et al.*, 2016). However, the use of allelic exchange to both reverse the orientation of the flagella operon promoter and restore the topoisomerase gene to wildtype did not correct either phenotype. Notwithstanding these issues, both, the Collery *et al.* (2016) publication and the discoveries made here, emphasise two points. Firstly, any study should always bear in mind the possibility that SNPs may arise at any point during processes which involve taking strains through procedures that involve single cell selection, such as picking single colonies. Repeatedly taking single cells (restreaking colonies) risks the possibility of selecting a clonal population carrying a SNP. Such steps, while unavoidable when undertaking molecular genetics, should be kept to a minimum. Secondly, laboratory strains should always be genome sequenced before investing any significant time in their analysis. Had this been done at an earlier stage with R20291 NM, then the extra work involved in re-making mutants could have been avoided.

6.1.3 Role of the *sigB* operon in the regulation of virulence factors.

The SigB operon has been reported to play a role in the regulation of virulence factors in bacteria such as *S. aureus* and *L. monocytogenes* (Mitchell *et al.*, 2013; Zhang *et al.*, 2013). One of the SNPs found in R20291 NM is located in the *rsbW* gene. Our hypothesis was that the *rsbW* SNP was involved in some of the changes, relative to R20291 BW, in virulence factors seen in R20291 NM. The research conducted in Chapter 5 focused on the SigB operon, specifically the *rsbW* and *sigB* genes. The results obtained confirmed that the SigB operon is not only involved in the cells' stress response, it also plays a role in the regulation of toxin and spore production. The *rsbW* mutant created in the R20291 BW strain showed an increased production of toxins compared to the parental strain. This result suggests that the *rsbW* gene participates in the regulation of toxins in the cell. However, this does not mean that the *rsbW* gene is directly involved in the regulation of toxins; rather, the toxin phenotype observed in the *rsbW* mutant might be explained by the interaction of RsbW with SigB (Guldimann *et al.,* 2016). During exponential growth, SigB remains inactive as a consequence of its interaction with RsbW. RsbW is an anti-sigma factor capable of forming a sTable RsbW/SigB complex that prevents SigB from interacting with its targets. The activation of SigB depends on the phosphorylated state of the RsbV protein, an anti-anti-sigma factor capable of binding RsbW and releasing SigB from the RsbW/SigB complex when the cells faces environmental stress (de Been *et al.,* 2011). In *L. monocytogenes,* it has been observed that SigB contributes to the transcription of PrfA, a key transcriptional regulator of virulence, from the *prfA* P2 promoter (Nadon *et al.,* 2002). Furthermore, it has been reported that SarA, a global regulator of virulence, is positively regulated by SigB in *S. aureus* (Bischoff *et al.,* 2004; Atwood *et al.,* 2015). The lack of the RsbW protein in the *C. difficile* BW *rsbw*::CTermB strain would mean that the SigB protein is not inactivate during exponential growth, being able to interact with SigB-dependant promoters at earlier times compared to the parental strain. It is possible that the early activation of SigB in the *rsbW* mutant could directly or indirectly promote the expression of toxins in the cell, explaining the higher toxin level observed in this strain.

The initiation of sporulation in *C. difficile* is regulated by Spo0A, the global regulator of sporulation in the cell (Deakin *et al.*, 2012). The R20291 BW *sigB*::CTermB mutant exhibits a delayed onset of sporulation compared to the wild type; however, it was observed that once the sporulation has initiated, the production of spores follows the same behaviour as seen in the parental strain. This suggests that the mutation affects the signalling involved in the initiation of sporulation as it has been observed that once Spo0A is activated, the sporulation process is auto-regulated by Spo0A (Rosenbuch *et al.*, 2012). The activation of Spo0A in *B. subtilis* is regulated through a phosphorylation-mediated signal transduction pathway comprised of five sensor kinases, KinA through KinE, (Jiang *et al.*, 2000). In this pathway, the initiation of sporulation is negatively regulated by promoting a series of phosphatases that remove the phosphoryl group from the active Spo0A (Spo0A~P), preventing the initiation of sporulation in the cell (Perego, 2001). One of this phosphatases, Spo0E, has been demonstrated to be positively regulated by SigB, preventing sporulation of the cells when it is overexpressed in the cell (Reder *et al.*, 2012). Nevertheless, it has also been reported that SigB is necessary for the efficient sporulation of *B. subtilis* cultures at low temperatures (Mendez *et al.*, 2004). This suggests that the decision-making process of sporulation in *B. subtilis* integrates SigB in a different way depending on the environmental conditions. Nevertheless, *C. difficile* lacks several key regulatory components found in *B. subtilis*, including phosphorelay transferases and phosphatases such as Spo0E, among others (Edwards & McBride, 2014). In

C. difficile, Spo0A seems to be regulated through the action of three orphan histidine kinases (CD1492, CD1579 and CD2492) that share similarity to the KinA-E phosphatases found in *B. subtilis* and possibly other regulators that have not been identified yet (Underwood *et al.,* 2009; Paredes *et al.,* 2005). The delay in the onset of sporulation observed in the *sigB* mutant created in the R20291 BW strain implies that SigB positively promotes the expression of key elements involved in the activation of Spo0A, possibly the orphan histidine kinases that are thought to be involved in the phosphorylation of Spo0A.

The results observed in the SigB operon, along with the findings in the *agr* loci, revealed that the regulation of toxins and spores are not controlled by a single regulatory pathway and that different pathways work together in the regulation of virulence factors.

6.2 Further research

CRISPR/Cas9 genome editing systems are powerful newly developed tools that are currently being used in several species for the precise manipulation of the genome. The potential of this tool in the editing of the *C. difficile* genome presents a good opportunity for the creation and complementation of mutants. The plasmid pMTLCas-intron::SNPrsbw created in this study (Section 5.2.2.2) was not able to effectively cleave the desired sequence in the genome. The change of the P_{araE} promoter in this plasmid for a more suiTable promoter in *C. difficile* presents a good opportunity to improve this system. Indeed, since completion of this work a derivative in which the P_{araE} promoter was replaced with the P_{tcdB} promoter has been successfully used to generate a *pyrE* deletion mutant of *C. difficile* strain 630 (P. Ingle and NP Minton, personal communication). Moreover, a gRNA target directed against the Group II intron has been used to generate an in-frame deletion of a Type I restriction gene carrying a ClosTron insertion in *C. sporogenes* (R. Rodrigues and N. Minton, personal communication). Thus, the feasibility of the strategy attempted in this study has been vindicated. . Another potential route is illustrated by studies with *Clostridium pasteurianum*, where an endogenous CRISPR/Cas9 has been deployed to bring about in-frame deletions (Pyne *et al.*, 2016). One of the requirements for the use of an endogenous system is to identify the native PAM sequence. A study published by Boudry *et al.* (2015) identified putative PAM sequences in *C. difficile* 630 and *C. difficile* R20291, which present a good opportunity for the development of such a system. The successful creation of a CRISPR/Cas9 system that can cleave the group II intron offers opportunity to remove the intron from insertional mutants to complement the native gene back into the genome as well as the opportunity to insert or repair SNPs in the homology arm used for the restoration of the genome through homologous recombination.

6.3 Concluding remarks

The findings described in this thesis provide evidence that the virulence factors in *C. difficile* are regulated by a complex set of mechanisms. As expected, the *agr* QS loci play an important role in the production of toxin and sporulation in *C. difficile* R20291. These QS loci are, however, not the only regulatory systems that control this processes. Thus, the SigB operon is also intimately involved in the expression of toxins and sporulation. The interruption of the *agrBD1* operon, the *agr2* operon and the *sigB* operon all affect toxins and spore production but do not result in the total lack of these virulence factors, indicating that there are other mechanisms involved capable of inducing the expression of the PaLoc and the sporulation pathway. The development of new technologies such as the CRISPR/Cas9 system coupled with existing ones such as RNA.seq provide powerful options for the precise manipulation of the genome and further understanding of the genes involved in the regulatory pathways concerned with virulence in *C. difficile* R20291.

References

Abecasis, A. B., Serrano, M., Alves, R., Quintais, L., Pereira-Leal, J. B., & Henriques, A. O. (2013). A genomic signature and the identification of new sporulation genes. *Journal of bacteriology, 195*(9), 2101-2115.

Ainstein, V., & Bodey, G. (1981). Relapsing pseudomembranous colitis associated with cancer chemotherapy. *Journal of Infectious Diseases, 143*(6), 865-865.

Åkerlund, T., Persson, I., Unemo, M., Norén, T., Svenungsson, B., Wullt, M., & Burman, L. G. (2008). Increased sporulation rate of epidemic *Clostridium difficile* type 027/NAP1. *Journal of clinical microbiology, 46*(4), 1530-1533.

Aktories, K. (2011). Bacterial protein toxins that modify host regulatory GTPases. *Nature Reviews Microbiology, 9*(7), 487-498.

Aktories, K., Schwan, C., Papatheodorou, P., & Lang, A. E. (2012). Bidirectional attack on the actin cytoskeleton. Bacterial protein toxins causing polymerization or depolymerization of actin. *Toxicon, 60*(4), 572-581.

Alfa, M. J., Kabani, A., Lyerly, D., Moncrief, S., Neville, L. M., Al-Barrak, A., . . . Embil, J. M. (2000). Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile* responsible for a nosocomial outbreak of *Clostridium difficile*-associated diarrhea. *Journal of clinical microbiology, 38*(7), 2706-2714.

Ambrose, N., Johnson, M., Burdon, D., & Keighley, M. (1985). The influence of single dose intravenous antibiotics on faecal flora and emergence of *Clostridium difficile*. *Journal of Antimicrobial Chemotherapy, 15*(3), 319-326.

Andrey, D. O., Jousselin, A., Villanueva, M., Renzoni, A., Monod, A., Barras, C., . . . Kelley, W. L. (2015). Impact of the regulators sigB, rot, sarA and sarS on the toxic shock Tst promoter and TSST-1 expression in Staphylococcus aureus. *PloS one, 10*(8), e0135579.

Aronsson, B., Möllby, R., & Nord, C. (1981). Occurrence of toxin-producing*Clostridium difficile* in antibiotic-associated diarrhea in Sweden. *Medical microbiology and immunology, 170*(1), 27-35.

Asad, S., & Opal, S. M. (2008). Bench-to-bedside review: quorum sensing and the role of cell-to-cell communication during invasive bacterial infection. *Critical care, 12*(6), 1.

Atwood, D. N., Loughran, A. J., Courtney, A. P., Anthony, A. C., Meeker, D. G., Spencer, H. J., . . . Smeltzer, M. S. (2015). Comparative impact of diverse regulatory loci on Staphylococcus aureus biofilm formation. *Microbiologyopen, 4*(3), 436-451.

Aubry, A., Hussack, G., Chen, W., KuoLee, R., Twine, S. M., Fulton, K. M., . . . Logan, S. M. (2012). Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. *Infection and immunity, 80*(10), 3521-3532.

Baban, S. T., Kuehne, S. A., Barketi-Klai, A., Cartman, S. T., Kelly, M. L., Hardie, K. R., . . . Minton, N. P. (2013). The role of flagella in *Clostridium difficile* pathogenesis: comparison between a non-epidemic and an epidemic strain. *PloS one, 8*(9), e73026.

Bagdasarian, N., Rao, K., & Malani, P. N. (2015). Diagnosis and treatment of *Clostridium difficile* in adults: a systematic review. *Jama, 313*(4), 398-408.

Baker, S. S., Faden, H., Sayej, W., Patel, R., & Baker, R. D. (2010). Increasing incidence of community-associated atypical *Clostridium difficile* disease in children. *Clinical pediatrics, 49*(7), 644-647.

Bakker, D., Smits, W. K., Kuijper, E. J., & Corver, J. (2012). TcdC does not significantly repress toxin expression in *Clostridium difficile* 630ΔErm. *PloS one, 7*(8), e43247.

Bandow, J. E., Brötz, H., & Hecker, M. (2002). Bacillus subtilis tolerance of moderate concentrations of rifampin involves the σB-dependent general and multiple stress response. *Journal of bacteriology, 184*(2), 459-467.

Barketi-Klai, A., Hoys, S., Lambert-Bordes, S., Collignon, A., & Kansau, I. (2011). Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. *Journal of medical microbiology, 60*(8), 1155-1161.

Barketi-Klai, A., Monot, M., Hoys, S., Lambert-Bordes, S., Kuehne, S. A., Minton, N., . . . Kansau, I. (2014). The flagellin FliC of *Clostridium difficile* is responsible for pleiotropic gene regulation during in vivo infection. *PloS one, 9*(5), e96876.

Barroso, L., Wang, S.-Z., Phelps, C., Johnson, J., & Wilkins, T. (1990). Nucleotide sequence of *Clostridium difficile* toxin B gene. *Nucleic acids research, 18*(13), 4004-4004.

Bartlett, J. G. (1990). *Clostridium difficile*: clinical considerations. *Review of Infectious Diseases, 12*(Supplement 2), S243-S251.

Bartlett, J. G. (1992). Antibiotic-associated diarrhea. *Clinical Infectious Diseases, 15*(4), 573-581.

Bartlett, J. G., Moon, N., Chang, T. W., Taylor, N., & Onderdonk, A. B. (1978). Role of *Clostridium difficile* in antibiotic-associated pseudomembranous colitis. *Gastroenterology, 75*(5), 778-782.

Bartlett, J. G., Taylor, N. S., Chang, T., & Dzink, J. (1980). Clinical and laboratory observations in *Clostridium difficile* colitis. *American Journal of Clinical Nutrition, 33*(11 Suppl), 2521-2526.

Bartlett, J. G., Tedesco, F. J., Shull, S., Lowe, B., & Chang, T. (1980). Symptomatic relapse after oral vancomycin therapy of antibiotic-associated pseudomembranous colitis. *Gastroenterology, 78*(3), 431-434.

Bassler, B. L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. *Current opinion in microbiology, 2*(6), 582-587.

Bassler, B. L., Greenberg, E. P., & Stevens, A. M. (1997). Cross-species induction of luminescence in the quorum-sensing bacterium Vibrio harveyi. *Journal of bacteriology, 179*(12), 4043-4045.

Bassler, B. L., Wright, M., Showalter, R. E., & Silverman, M. R. (1993). Intercellular signalling in Vibrio harveyi: sequence and function of genes regulating expression of luminescence. *Molecular microbiology, 9*(4), 773-786.

Bassler, B. L., Wright, M., & Silverman, M. R. (1994). Multiple signalling systems controlling expression of luminescence in Vibrio harveyi: sequence and function of genes encoding a second sensory pathway. *Molecular microbiology, 13*(2), 273-286.

Benson, A. K., & Haldenwang, W. G. (1993). Bacillus subtilis sigma B is regulated by a binding protein (RsbW) that blocks its association with core RNA polymerase. *Proceedings of the National Academy of Sciences, 90*(6), 2330-2334.

Biazzo, M., Cioncada, R., Fiaschi, L., Tedde, V., Spigaglia, P., Mastrantonio, P., . . . Galeotti, C. L. (2013). Diversity of cwp loci in clinical isolates of *Clostridium difficile*. *Journal of medical microbiology, 62*(9), 1444-1452.

Bijlsma, J. J., & Groisman, E. A. (2003). Making informed decisions: regulatory interactions between two-component systems. *Trends in microbiology, 11*(8), 359-366.

Bischoff, M., Dunman, P., Kormanec, J., Macapagal, D., Murphy, E., Mounts, W., . . . Projan, S. (2004). Microarray-based analysis of the

Staphylococcus aureus σB regulon. *Journal of bacteriology, 186*(13), 4085-4099.

Bischoff, M., Entenza, J., & Giachino, P. (2001). Influence of a Functional sigB Operon on the Global Regulators sar and agr inStaphylococcus aureus. *Journal of bacteriology, 183*(17), 5171-5179.

Blöcker, D., Behlke, J., Aktories, K., & Barth, H. (2001). Cellular uptake of the Clostridium perfringens binary iota-toxin. *Infection and immunity, 69*(5), 2980-2987.

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., Jacquier, A. (2007). Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & development, 21*(11), 1353-1366.

Boone, J., Archbald-Pannone, L., Wickham, K., Carman, R., Guerrant, R., Franck, C., & Lyerly, D. (2014). Ribotype 027 *Clostridium difficile* infections with measurable stool toxin have increased lactoferrin and are associated with a higher mortality. *European Journal of Clinical Microbiology & Infectious Diseases, 33*(6), 1045-1051.

Bordeleau, E., Fortier, L.-C., Malouin, F., & Burrus, V. (2011). c-di-GMP turn-over in *Clostridium difficile* is controlled by a plethora of diguanylate cyclases and phosphodiesterases. *PLoS Genet, 7*(3), e1002039.

Bordeleau, E., Purcell, E. B., Lafontaine, D. A., Fortier, L.-C., Tamayo, R., & Burrus, V. (2015). Cyclic di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. *Journal of bacteriology, 197*(5), 819-832.

Borriello, S. (1979). *Clostridium difficile* and its toxin in the gastrointestinal tract in health and disease. *Res Clin Forums, 1*, 33-35.

Borriello, S., & Bhatt, R. (1995). Chemotaxis by *Clostridium difficile*. *Medical and Dental Aspects of Anaerobes (Duerden B, Wade JG, Brzier JG, Eley JS, Wren B, Hudson MJ. Eds). Sceintific Reviews Ltd Middlesex*, 241.

Borriello, S., Wren, B., Hyde, S., Seddon, S., Sibbons, P., Krishna, M., . . . Price, A. (1992). Molecular, immunological, and biological characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile*. *Infection and immunity, 60*(10), 4192-4199.

Boudry, P., Semenova, E., Monot, M., Datsenko, K. A., Lopatina, A., Sekulovic, O., . . . Dupuy, B. (2015). Function of the CRISPR-Cas System of the Human Pathogen *Clostridium difficile*. *MBio, 6*(5), e01112-01115.

Bouillaut, L., Perchat, S., Arold, S., Zorrilla, S., Slamti, L., Henry, C., . . . Lereclus, D. (2008). Molecular basis for group-specific activation of the virulence regulator PlcR by PapR heptapeptides. *Nucleic acids research, 36*(11), 3791-3801.

Brahmbhatt, T. N., Janes, B. K., Stibitz, E. S., Darnell, S. C., Sanz, P., Rasmussen, S. B., & O'Brien, A. D. (2007). Bacillus anthracis exosporium protein BclA affects spore germination, interaction with extracellular matrix proteins, and hydrophobicity. *Infection and immunity, 75*(11), 5233-5239.

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

Brazier, J., Patel, B., & Pearson, A. (2007). Distribution of *Clostridium difficile* PCR ribotype 027 in British hospitals. *Euro Surveill, 12*(4), E070426.

Brazier, J., Raybould, R., Patel, B., Duckworth, G., Pearson, A., Charlett, A., & Duerden, B. (2008). Distribution and antimicrobial susceptibility patterns of *Clostridium difficile* PCR ribotypes in English hospitals, 2007-08. *Euro surveillance: bulletin europeen sur les maladies transmissibles= European communicable disease bulletin, 13*(41), 541-544.

Brown, K. L., & Hughes, K. T. (1995). The role of anti-sigma factors in gene regulation. *Molecular microbiology, 16*(3), 397-404.

Brown, T. A. (2007). *Genomes 3* (3rd ed.): Garland Science Publishing.

Buckley, A. M., Spencer, J., Candlish, D., Irvine, J. J., & Douce, G. R. (2011). Infection of hamsters with the UK Clostridiumdifficile ribotype 027 outbreak strain R20291. *Journal of medical microbiology, 60*(8), 1174-1180.

Cafardi, V., Biagini, M., Martinelli, M., Leuzzi, R., Rubino, J. T., Cantini, F., . . . Unnikrishnan, M. (2013). Identification of a novel zinc metalloprotease through a global analysis of *Clostridium difficile* extracellular proteins. *PloS one, 8*(11), e81306.

Calabi, E., Calabi, F., Phillips, A. D., & Fairweather, N. F. (2002). Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. *Infection and immunity, 70*(10), 5770-5778.

Callahan, S. M., & Dunlap, P. V. (2000). LuxR-and Acyl-Homoserine-Lactone-Controlled Non-luxGenes Define a Quorum-Sensing Regulon in Vibrio fischeri. *Journal of bacteriology, 182*(10), 2811-2822.

Capra, E. J., & Laub, M. T. (2012). Evolution of two-component signal transduction systems. *Annual review of microbiology, 66*, 325-347.

Carignan, B. M., Brumfield, K. D., & Son, M. S. (2016). Single Nucleotide Polymorphisms in Regulator-Encoding Genes Have an Additive Effect on Virulence Gene Expression in a Vibrio cholerae Clinical Isolate. *mSphere, 1*(5), e00253-00216.

Carter, G. P., Rood, J. I., & Lyras, D. (2012). The role of toxin A and toxin B in the virulence of *Clostridium difficile*. *Trends in microbiology, 20*(1), 21-29.

Cartman, S. T., Kelly, M. L., Heeg, D., Heap, J. T., & Minton, N. P. (2012). Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdC genotype and toxin production. *Applied and environmental microbiology, 78*(13), 4683-4690.

Case, R. J., Labbate, M., & Kjelleberg, S. (2008). AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. *ISME journal, 2*(4), 345.

Cerca, N., Brooks, J. L., & Jefferson, K. K. (2008). Regulation of the intercellular adhesin locus regulator (icaR) by SarA, σB, and IcaR in Staphylococcus aureus. *Journal of bacteriology, 190*(19), 6530-6533.

Chambers, C. J., Roberts, A. K., Shone, C. C., & Acharya, K. R. (2015). Structure and function of a *Clostridium difficile* sortase enzyme. *Scientific reports, 5*, 9449.

Chen, J., & McClane, B. A. (2012). Role of the Agr-like quorum-sensing system in regulating toxin production by Clostridium perfringens type B strains CN1793 and CN1795. *Infection and immunity, 80*(9), 3008-3017.

Chylinski, K., Le Rhun, A., & Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA biology, 10*(5), 726-737.

Clancy, K. W., Melvin, J. A., & McCafferty, D. G. (2010). Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Peptide Science, 94*(4), 385-396.

Cohen, S. H., Gerding, D. N., Johnson, S., Kelly, C. P., Loo, V. G., McDonald, L. C., . . . Wilcox, M. H. (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, 31*(05), 431-455.

Colarian, J. (1988). *Clostridium difficile* colitis following antiviral therapy in the acquired immunodeficiency syndrome. *The American journal of medicine, 84*(6), 1081.

Collery, M. M., Kuehne, S. A., McBride, S. M., Kelly, M. L., Monot, M., Cockayne, A., . . . Minton, N. P. (2016). What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of *Clostridium difficile* strain 630 and derivatives. *Virulence*(just-accepted), 00-00.

Collins, D. A., Elliott, B., & Riley, T. V. (2015). Molecular methods for detecting and typing of *Clostridium difficile*. *Pathology-Journal of the RCPA, 47*(3), 211-218.

Cooksley, C. M., Davis, I. J., Winzer, K., Chan, W. C., Peck, M. W., & Minton, N. P. (2010). Regulation of neurotoxin production and sporulation by a putative agrBD signaling system in proteolytic Clostridium botulinum. *Applied and environmental microbiology, 76*(13), 4448-4460.

Đapa, T., Leuzzi, R., Ng, Y. K., Baban, S. T., Adamo, R., Kuehne, S. A., . . . Unnikrishnan, M. (2013). Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. *Journal of bacteriology, 195*(3), 545-555.

Darkoh, C., DuPont, H., Norris, S., & Kaplan, H. (2015). Toxin synthesis by *Clostridium difficile* is regulated through quorum signaling. mBio 6 (2): e02569-14.

Darkoh, C., Odo, C., & DuPont, H. L. (2016). Accessory Gene Regulator-1 Locus Is Essential for Virulence and Pathogenesis of *Clostridium difficile*. *MBio, 7*(4), e01237-01216.

Davies, D. G., Parsek, M. R., Pearson, J. P., Iglewski, B. H., Costerton, J., & Greenberg, E. (1998). The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science, 280*(5361), 295-298.

Dawson, L. F., Valiente, E., Faulds-Pain, A., Donahue, E. H., & Wren, B. W. (2012). Characterisation of *Clostridium difficile* biofilm formation, a role for Spo0A. *PloS one, 7*(12), e50527.

de Been, M., Francke, C., Siezen, R. J., & Abee, T. (2011). Novel σB regulation modules of Gram-positive bacteria involve the use of complex hybrid histidine kinases. *Microbiology, 157*(1), 3-12.

de la Riva, L., Willing, S. E., Tate, E. W., & Fairweather, N. F. (2011). Roles of cysteine proteases Cwp84 and Cwp13 in biogenesis of the cell wall of *Clostridium difficile*. *Journal of bacteriology, 193*(13), 3276-3285.

Deakin, L. J., Clare, S., Fagan, R. P., Dawson, L. F., Pickard, D. J., West, M. R., . . . Lawley, T. D. (2012). The *Clostridium difficile* spo0A gene is a persistence and transmission factor. *Infection and immunity, 80*(8), 2704-2711.

Declerck, N., Bouillaut, L., Chaix, D., Rugani, N., Slamti, L., Hoh, F., . . . Arold, S. T. (2007). Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. *Proceedings of the National Academy of Sciences, 104*(47), 18490-18495.

Delmée, M., & Avesani, V. (1990). Virulence of ten serogroups of *Clostridium difficile* in hamsters. *Journal of medical microbiology, 33*(2), 85-90.

Dineen, S. S., Villapakkam, A. C., Nordman, J. T., & Sonenshein, A. L. (2007). Repression of *Clostridium difficile* toxin gene expression by CodY. *Molecular microbiology, 66*(1), 206-219.

DON, C., & LAVIS, A. (1981). The association between antibiotic-associated diarrhoea and c. dlfflclle toxin in children. *Australian and New Zealand journal of medicine, 11*(3), 433-434.

Donahue, E. H., Dawson, L. F., Valiente, E., Firth-Clark, S., Major, M. R., Littler, E., . . . Wren, B. W. (2014). *Clostridium difficile*has a single sortase, SrtB, that can be inhibited by small-molecule inhibitors. *BMC Microbiology, 14*(1), 219. doi:10.1186/s12866-014-0219-1

Dove, C., Wang, S., Price, S., Phelps, C., Lyerly, D., Wilkins, T., & Johnson, J. (1990). Molecular characterization of the *Clostridium difficile* toxin A gene. *Infection and immunity, 58*(2), 480-488.

Driessen, A. J., & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem., 77*, 643-667.

Drudy, D., Fanning, S., & Kyne, L. (2007). Toxin A-negative, toxin B-positive *Clostridium difficile*. *International Journal of Infectious Diseases, 11*(1), 5-10.

Drudy, D., Harnedy, N., Fanning, S., O'mahony, R., & Kyne, L. (2007). Isolation and characterisation of toxin A-negative, toxin B-positive *Clostridium difficile* in Dublin, Ireland. *Clinical Microbiology and Infection, 13*(3), 298-304.

Dufour, A., & Haldenwang, W. G. (1994). Interactions between a Bacillus subtilis anti-sigma factor (RsbW) and its antagonist (RsbV). *Journal of bacteriology, 176*(7), 1813-1820.

Dusane, D. H., Zinjarde, S. S., Venugopalan, V. P., Mclean, R. J., Weber, M. M., & Rahman, P. K. (2010). Quorum sensing: implications on rhamnolipid biosurfactant production. *Biotechnology and Genetic Engineering Reviews, 27*(1), 159-184.

Eberhard, A., Burlingame, A., Eberhard, C., Kenyon, G., Nealson, K., & Oppenheimer, N. (1981). Structural identification of autoinducer of Photobacterium fischeri luciferase. *Biochemistry, 20*(9), 2444-2449.

Eckert, C., Emirian, A., Le Monnier, A., Cathala, L., De Montclos, H., Goret, J., . . . Jean-Pierre, H. (2015). Prevalence and pathogenicity of binary toxin–positive *Clostridium difficile* strains that do not produce toxins A and B. *New microbes and new infections, 3*, 12-17.

Edwards, A. N., & McBride, S. M. (2014). Initiation of sporulation in *Clostridium difficile*: a twist on the classic model. *FEMS Microbiology Letters, 358*(2), 110-118.

Egerer, M., Giesemann, T., Herrmann, C., & Aktories, K. (2009). Autocatalytic Processing of *Clostridium difficile* Toxin B binding of inositol hexakisphosphate. *Journal of Biological Chemistry, 284*(6), 3389-3395.

Ehsaan, M., Kuit, W., Zhang, Y., Cartman, S. T., Heap, J. T., Winzer, K., & Minton, N. P. (2016). Mutant generation by allelic exchange and genome resequencing of the biobutanol organism Clostridium acetobutylicum ATCC 824. *Biotechnology for biofuels, 9*(1), 4.

Engebrecht, J., Nealson, K., & Silverman, M. (1983). Bacterial bioluminescence: isolation and genetic analysis of functions from Vibrio fischeri. *Cell, 32*(3), 773-781.

Engelmann, S., & Hecker, M. (1996). Impaired oxidative stress resistance of Bacillus subtilis sigB mutants and the role of katA and katE. *FEMS Microbiology Letters, 145*(1), 63-69.

Erdem, A. L., Avelino, F., Xicohtencatl-Cortes, J., & Girón, J. A. (2007). Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing Escherichia coli. *Journal of bacteriology, 189*(20), 7426-7435.

Escobar-Cortés, K., Barra-Carrasco, J., & Paredes-Sabja, D. (2013). Proteases and sonication specifically remove the exosporium layer of spores of *Clostridium difficile* strain 630. *Journal of microbiological methods, 93*(1), 25-31.

Etienne-Manneville, S., & Hall, A. (2002). Rho GTPases in cell biology. *Nature, 420*(6916), 629-635.

Fagan, R. P., & Fairweather, N. F. (2011). *Clostridium difficile* has two parallel and essential Sec secretion systems. *Journal of Biological Chemistry, 286*(31), 27483-27493.

Fagan, R. P., & Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nature Reviews Microbiology*, *12*(3), 211-222.

Faulds-Pain, A., Twine, S. M., Vinogradov, E., Strong, P. C., Dell, A., Buckley, A. M., . . . Wren, B. W. (2014). The post-translational modification of the *Clostridium difficile* flagellin affects motility, cell surface properties and virulence. *Molecular microbiology, 94*(2), 272-289.

Feklístov, A., Sharon, B. D., Darst, S. A., & Gross, C. A. (2014). Bacterial sigma factors: a historical, structural, and genomic perspective. *Annual review of microbiology, 68*, 357-376.

Ferreira, A., O'Byrne, C. P., & Boor, K. J. (2001). Role of ςB in Heat, Ethanol, Acid, and Oxidative Stress Resistance and during Carbon Starvation inListeria monocytogenes. *Applied and environmental microbiology, 67*(10), 4454-4457.

Fimlaid, K. A., Bond, J. P., Schutz, K. C., Putnam, E. E., Leung, J. M., Lawley, T. D., & Shen, A. (2013). Global analysis of the sporulation pathway of *Clostridium difficile*. *PLoS Genet, 9*(8), e1003660.

Florin, I., & Thelestam, M. (1983). Internalization of *Clostridium difficile* cytotoxin into cultured human lung fibroblasts. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 763*(4), 383-392.

Francis, M. B., Allen, C. A., Shrestha, R., & Sorg, J. A. (2013). Bile acid recognition by the *Clostridium difficile* germinant receptor, CspC, is important for establishing infection. *PLoS Pathog*, *9*(5), e1003356.

Fu, Z., Donegan, N. P., Memmi, G., & Cheung, A. L. (2007). Characterization of MazFSa, an endoribonuclease from Staphylococcus aureus. *Journal of bacteriology, 189*(24), 8871-8879.

Fuqua, W. C., Winans, S. C., & Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of bacteriology, 176*(2), 269.

Gao, R., Mack, T. R., & Stock, A. M. (2007). Bacterial response regulators: versatile regulatory strategies from common domains. *Trends in biochemical sciences, 32*(5), 225-234.

George, A. P. W., Rolfe, A. P. R., Klein, A. P. R., & Finegold, S. (1982). *Clostridium difficile* and cytotoxin in feces of patients with antimicrobial agent-associated pseudomembranous colitis. *Infection, 10*(4), 205-207.

Gerhard, R., Nottrott, S., Schoentaube, J., Tatge, H., Olling, A., & Just, I. (2008). Glucosylation of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells. *Journal of medical microbiology, 57*(6), 765-770.

Geyer, M., Wilde, C., Selzer, J., Aktories, K., & Kalbitzer, H. R. (2003). Glucosylation of Ras by Clostridium sordellii lethal toxin: consequences for effector loop conformations observed by NMR spectroscopy. *Biochemistry, 42*(41), 11951-11959.

Giesemann, T., Egerer, M., Jank, T., & Aktories, K. (2008). Processing of *Clostridium difficile* toxins. *Journal of medical microbiology, 57*(6), 690-696.

Gohar, M., Faegri, K., Perchat, S., Ravnum, S., Økstad, O. A., Gominet, M., . . . Lereclus, D. (2008). The PlcR virulence regulon of Bacillus cereus. *PloS one, 3*(7), e2793.

Gominet, M., Slamti, L., Gilois, N., Rose, M., & Lereclus, D. (2001). Oligopeptide permease is required for expression of the Bacillus thuringiensis plcR regulon and for virulence. *Molecular microbiology, 40*(4), 963-975.

Goorhuis, A., Van der Kooi, T., Vaessen, N., Dekker, F., Van den Berg, R., Harmanus, C., . . . Kuijper, E. (2007). Spread and epidemiology of *Clostridium difficile* polymerase chain reaction ribotype 027/toxinotype III in The Netherlands. *Clinical Infectious Diseases, 45*(6), 695-703.

Goulding, D., Thompson, H., Emerson, J., Fairweather, N. F., Dougan, G., & Douce, G. R. (2009). Distinctive profiles of infection and pathology in hamsters infected with *Clostridium difficile* strains 630 and B1. *Infection and immunity, 77*(12), 5478-5485.

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

Govind, R., Fitzwater, L., & Nichols, R. (2015). Observations on the role of TcdE isoforms in *Clostridium difficile* toxin secretion. *Journal of bacteriology, 197*(15), 2600-2609.

Goymer, P. (2007). Synonymous mutations break their silence. *Nature Reviews Genetics, 8*(2), 92-92.

Gruber, T. M., & Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Reviews in Microbiology, 57*(1), 441-466.

Guldimann, C., Boor, K. J., Wiedmann, M., & Guariglia-Oropeza, V. (2016). Resilience in the Face of Uncertainty: Sigma Factor B Fine-Tunes Gene Expression To Support Homeostasis in Gram-Positive Bacteria. *Applied and environmental microbiology, 82*(15), 4456-4469.

Guttenplan, S. B., & Kearns, D. B. (2013). Regulation of flagellar motility during biofilm formation. *FEMS microbiology reviews, 37*(6), 849-871.

Hall, I. C., & O'TOOLE, E. (1935). Intestinal flora in new-born infants: with a description of a new pathogenic anaerobe, Bacillus difficilis. *American journal of diseases of children, 49*(2), 390-402.

Hamm, E. E., Voth, D. E., & Ballard, J. D. (2006). Identification of *Clostridium difficile* toxin B cardiotoxicity using a zebrafish embryo model of intoxication. *Proceedings of the National Academy of Sciences, 103*(38), 14176-14181.

Harry, K. H., Zhou, R., Kroos, L., & Melville, S. B. (2009). Sporulation and enterotoxin (CPE) synthesis are controlled by the sporulation-specific sigma factors SigE and SigK in Clostridium perfringens. *Journal of bacteriology, 191*(8), 2728-2742.

Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y.-K., Cartman, S. T., Winzer, K., & Minton, N. P. (2012). Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic acids research, 40*(8), e59-e59.

Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C., & Minton, N. P. (2010). The ClosTron: mutagenesis in Clostridium refined and streamlined. *Journal of microbiological methods, 80*(1), 49-55.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus Clostridium. *Journal of microbiological methods, 70*(3), 452-464.

Heap, J. T., Pennington, O. J., Cartman, S. T., & Minton, N. P. (2009). A modular system for Clostridium shuttle plasmids. *Journal of microbiological methods, 78*(1), 79-85.

Hecker, M., Pané-Farré, J., & Uwe, V. (2007). SigB-dependent general stress response in Bacillus subtilis and related gram-positive bacteria. *Annu. Rev. Microbiol., 61*, 215-236.

Hedrick, P. W. (2011). *Genetics of populations*: Jones & Bartlett Learning.

Hemmasi, S., Czulkies, B. A., Schorch, B., Veit, A., Aktories, K., & Papatheodorou, P. (2015). Interaction of the *Clostridium difficile* Binary Toxin CDT and Its Host Cell Receptor, Lipolysis-stimulated Lipoprotein Receptor (LSR). *Journal of Biological Chemistry, 290*(22), 14031-14044.

Hendrickx, A. P., Budzik, J. M., Oh, S.-Y., & Schneewind, O. (2011). Architects at the bacterial surface—sortases and the assembly of pili with isopeptide bonds. *Nature Reviews Microbiology, 9*(3), 166-176.

Henriques, A. O., & Moran, J., Charles P. (2007). Structure, assembly, and function of the spore surface layers. *Annu. Rev. Microbiol., 61*, 555-588.

Henriques, B., Florin, I., & Thelestam, M. (1987). Cellular internalisation of *Clostridium difficile* toxin A. *Microbial pathogenesis, 2*(6), 455-463.

Hensbergen, P. J., Klychnikov, O. I., Bakker, D., van Winden, V. J., Ras, N., Kemp, A. C., . . . Kuijper, E. J. (2014). A novel secreted metalloprotease (CD2830) from *Clostridium difficile* cleaves specific proline sequences in LPXTG cell surface proteins. *Molecular & Cellular Proteomics, 13*(5), 1231-1244.

Hensgens, M. P., Keessen, E. C., Squire, M. M., Riley, T. V., Koene, M. G., de Boer, E., . . . Kuijper, E. J. (2012). *Clostridium difficile* infection in the community: a zoonotic disease? *Clinical Microbiology and Infection, 18*(7), 635-645.

Hernández-Rocha, C., Barra-Carrasco, J., Pizarro-Guajardo, M., Ibanez, P., Bueno, S. M., Sarker, M. R., . . . Paredes-Sabja, D. (2012). Epidemic *Clostridium difficile* ribotype 027 in Chile. *Emerg Infect Dis, 18*(8), 1370-1372.

Hofmann, N., Wurm, R., & Wagner, R. (2011). The E. coli anti-sigma factor Rsd: studies on the specificity and regulation of its expression.

Honda, H., & Dubberke, E. R. (2014). The changing epidemiology of *Clostridium difficile* infection. *Current opinion in gastroenterology, 30*(1), 54-62.

Hope, J. N., Bell, A. W., Hermodson, M. A., & Groarke, J. (1986). Ribokinase from Escherichia coli K12. Nucleotide sequence and overexpression of the rbsK gene and purification of ribokinase. *Journal of Biological Chemistry, 261*(17), 7663-7668.

Horsburgh, M. J., Aish, J. L., White, I. J., Shaw, L., Lithgow, J. K., & Foster, S. J. (2002). σB modulates virulence determinant expression and stress resistance: characterization of a functional rsbU strain derived from Staphylococcus aureus 8325-4. *Journal of bacteriology, 184*(19), 5457-5467.

Huang, H., Chai, C., Li, N., Rowe, P., Minton, N. P., Yang, S., . . . Gu, Y. (2016). CRISPR/Cas9-based efficient genome editing in Clostridium ljungdahlii, an autotrophic gas-fermenting bacterium. *ACS Synthetic Biology*.

Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M., & von EICHEL-STREIBER, C. (1997). Transcription analysis of the genes tcdA-E of the pathogenicity locus of Clostridium dij-jficile. *Eur. J. Biochem*, *244*, 735-742.

Hunt, J. J., & Ballard, J. D. (2013). Variations in virulence and molecular biology among emerging strains of *Clostridium difficile*. *Microbiology and Molecular Biology Reviews, 77*(4), 567-581.

Ibberson, C. B., Jones, C. L., Singh, S., Wise, M. C., Hart, M. E., Zurawski, D. V., & Horswill, A. R. (2014). Staphylococcus aureus hyaluronidase is a CodY-regulated virulence factor. *Infection and immunity, 82*(10), 4253-4264.

Inouye, M., & Dutta, R. (2002). *Histidine kinases in signal transduction*: Academic Press.

Jank, T., & Aktories, K. (2008). Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends in microbiology, 16*(5), 222-229.

Janoir, C., Denève, C., Bouttier, S., Barbut, F., Hoys, S., Caleechum, L., . . . Collignon, A. (2013). Adaptive strategies and pathogenesis of *Clostridium difficile* from in vivo transcriptomics. *Infection and immunity, 81*(10), 3757-3769.

Janoir, C., Péchiné, S., Grosdidier, C., & Collignon, A. (2007). Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. *Journal of bacteriology, 189*(20), 7174-7180.

Jiang, M., Shao, W., Perego, M., & Hoch, J. A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in Bacillus subtilis. *Molecular microbiology, 38*(3), 535-542.

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology, 31*(3), 233-239.

Jimenez, P. N., Koch, G., Thompson, J. A., Xavier, K. B., Cool, R. H., & Quax, W. J. (2012). The multiple signaling systems regulating virulence in Pseudomonas aeruginosa. *Microbiology and Molecular Biology Reviews, 76*(1), 46-65.

Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M., & Aktories, K. (1995). Glucosylation of Rho proteins by *Clostridium difficile* toxin B.

Kahali, B., Basak, S., & Ghosh, T. C. (2007). Reinvestigating the codon and amino acid usage of S. cerevisiae genome: a new insight from protein secondary structure analysis. *Biochemical and biophysical research communications, 354*(3), 693-699.

Kaplan, H. B., & Greenberg, E. (1985). Diffusion of autoinducer is involved in regulation of the Vibrio fischeri luminescence system. *Journal of bacteriology, 163*(3), 1210-1214.

Kazmierczak, M. J., Wiedmann, M., & Boor, K. J. (2005). Alternative sigma factors and their roles in bacterial virulence. *Microbiology and Molecular Biology Reviews, 69*(4), 527-543.

Kelly, C. P., Pothoulakis, C., & LaMont, J. T. (1994). *Clostridium difficile* colitis. *New England Journal of Medicine, 330*(4), 257-262.

Kim, S., Lee, S., Kim, Y., Kim, J., Ryu, P., Chung, S., & Rhee, J. (2000). *Virulence regulatory role of luxS quorum sensing system in Vibrio vulmificus.* Paper presented at the American Society for Microbiology

General Meeting. Washington, DC: American Society for Microbiology Press, abstract B-248.

Kimchi-Sarfaty, C., Oh, J. M., Kim, I.-W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V., & Gottesman, M. M. (2007). A" silent" polymorphism in the MDR1 gene changes substrate specificity. *Science, 315*(5811), 525-528.

Kirby, J. M., Ahern, H., Roberts, A. K., Kumar, V., Freeman, Z., Acharya, K. R., & Shone, C. C. (2009). Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of *Clostridium difficile*. *Journal of Biological Chemistry, 284*(50), 34666-34673.

Kleerebezem, M., & Quadri, L. E. (2001). Peptide pheromone-dependent regulation of antimicrobial peptide production in Gram-positive bacteria: a case of multicellular behavior. *Peptides, 22*(10), 1579-1596.

Knoop, F. C., Owens, M., & Crocker, I. C. (1993). *Clostridium difficile*: clinical disease and diagnosis. *Clinical Microbiology Reviews, 6*(3), 251-265.

Koenig, R. L., Ray, J. L., Maleki, S. J., Smeltzer, M. S., & Hurlburt, B. K. (2004). Staphylococcus aureus AgrA binding to the RNAIII-agr regulatory region. *Journal of bacteriology, 186*(22), 7549-7555.

Köhler, T., Epp, S. F., Curty, L. K., & Pechère, J.-C. (1999). Characterization of MexT, the regulator of the MexE-MexF-OprN multidrug efflux system of Pseudomonas aeruginosa. *Journal of bacteriology, 181*(20), 6300-6305.

Kovacs-Simon, A., Leuzzi, R., Kasendra, M., Minton, N., Titball, R. W., & Michell, S. L. (2014). Lipoprotein CD0873 is a novel adhesin of *Clostridium difficile*. *Journal of Infectious Diseases*, *210*(2), 274-284.

Krukonis, E. S., Yu, R. R., & DiRita, V. J. (2000). The Vibrio cholerae ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. *Molecular microbiology, 38*(1), 67-84.

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A., & Minton, N. P. (2010). The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature, 467*(7316), 711-713.

Kuehne, S. A., Collery, M. M., Kelly, M. L., Cartman, S. T., Cockayne, A., & Minton, N. P. (2014). Importance of toxin A, toxin B, and CDT in virulence of an epidemic *Clostridium difficile* strain. *Journal of Infectious Diseases, 209*(1), 83-86.

Kuijper, E., Barbut, F., Brazier, J., Kleinkauf, N., Eckmanns, T., Lambert, M., . . . Brown, D. (2008). Update of *Clostridium difficile* infection due to PCR ribotype 027 in Europe, 2008. *Euro surveillance: bulletin europeen sur les maladies transmissibles= European communicable disease bulletin*.

Kuijper, E., Coignard, B., & Tüll, P. (2006). Emergence of *Clostridium difficile*-associated disease in North America and Europe. *Clinical Microbiology and Infection, 12*(s6), 2-18.

Kullik, I., Giachino, P., & Fuchs, T. (1998). Deletion of the alternative sigma factor ςB in Staphylococcus aureus reveals its function as a global regulator of virulence genes. *Journal of bacteriology, 180*(18), 4814-4820.

Kyne, L., Warny, M., Qamar, A., & Kelly, C. P. (2001). Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. *The Lancet, 357*(9251), 189-193.

Lagier, J.-C., Brouqui, P., Fenollar, F., & Raoult, D. (2016). A Regional Outbreak of *Clostridium difficile* PCR-Ribotype 027 Infections in Southeastern France from a Single Long-Term Care Facility. *Infect Control Hosp Epidemiol, 1*, 5.

LaSala, P. R., Svensson, A. M., Mohammad, A. A., & Perrotta, P. L. (2012). Comparison of analytical and clinical performance of three methods for detection of *Clostridium difficile*. *Archives of pathology & laboratory medicine, 136*(5), 527-531.

Laub, M. T., & Goulian, M. (2007). Specificity in two-component signal transduction pathways. *Annu. Rev. Genet., 41*, 121-145.

Lauderdale, K. J., Boles, B. R., Cheung, A. L., & Horswill, A. R. (2009). Interconnections between Sigma B, agr, and proteolytic activity in Staphylococcus aureus biofilm maturation. *Infection and immunity, 77*(4), 1623-1635.

Lawley, T. D., Clare, S., Deakin, L. J., Goulding, D., Yen, J. L., Raisen, C., . . . Clark, T. G. (2010). Use of purified *Clostridium difficile* spores to facilitate evaluation of health care disinfection regimens. *Applied and environmental microbiology, 76*(20), 6895-6900.

Lawley, T. D., Croucher, N. J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., . . . Dougan, G. (2009). Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *Journal of bacteriology, 191*(17), 5377-5386.

Lerat, E., & Moran, N. A. (2004). The evolutionary history of quorum-sensing systems in bacteria. *Molecular biology and evolution, 21*(5), 903-913.

Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., . . . Yang, J. (2016). CRISPR-based genome editing and expression control systems in Clostridium acetobutylicum and Clostridium beijerinckii. *Biotechnology journal*.

Lin, Y.-P., Kuo, C.-J., Koleci, X., McDonough, S. P., & Chang, Y.-F. (2011). Manganese binds to *Clostridium difficile* Fbp68 and is essential for fibronectin binding. *Journal of Biological Chemistry, 286*(5), 3957-3969.

Lina, G., Jarraud, S., Ji, G., Greenland, T., Pedraza, A., Etienne, J., . . . Vandenesch, F. (1998). Transmembrane topology and histidine protein kinase activity of AgrC, the agr signal receptor in Staphylococcus aureus. *Molecular microbiology, 28*(3), 655-662.

Lyerly, D. M., Saum, K. E., MacDonald, D. K., & Wilkins, T. D. (1985). Effects of *Clostridium difficile* toxins given intragastrically to animals. *Infection and immunity*, *47*(2), 349-352.

Lyon, G. J., & Novick, R. P. (2004). Peptide signaling in Staphylococcus aureus and other Gram-positive bacteria. *Peptides*, *25*(9), 1389-1403.

Lyon, W. R., Madden, J. C., Levin, J. C., Stein, J. L., & Caparon, M. G. (2001). Mutation of luxS affects growth and virulence factor expression in Streptococcus pyogenes. *Molecular microbiology, 42*(1), 145-157.

Lyras, D., O'Connor, J. R., Howarth, P. M., Sambol, S. P., Carter, G. P., Phumoonna, T., . . . Johnson, S. (2009). Toxin B is essential for virulence of *Clostridium difficile*. *Nature, 458*(7242), 1176-1179.

Mackin, K. E., Carter, G. P., Howarth, P., Rood, J. I., & Lyras, D. (2013). Spo0A differentially regulates toxin production in evolutionarily diverse strains of *Clostridium difficile*. *PloS one, 8*(11), e79666.

Maillard, J.-Y. (2011). Innate resistance to sporicides and potential failure to decontaminate. *Journal of Hospital Infection, 77*(3), 204-209.

Majerczyk, C. D., Sadykov, M. R., Luong, T. T., Lee, C., Somerville, G. A., & Sonenshein, A. L. (2008). Staphylococcus aureus CodY negatively regulates virulence gene expression. *Journal of bacteriology, 190*(7), 2257-2265.

Maldarelli, G. A., De Masi, L., von Rosenvinge, E. C., Carter, M., & Donnenberg, M. S. (2014). Identification, immunogenicity, and cross-reactivity of type IV pilin and pilin-like proteins from *Clostridium difficile*. *Pathogens and disease, 71*(3), 302-314.

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, 98*(10), 5844-5849.

Mani, N., Lyras, D., Barroso, L., Howarth, P., Wilkins, T., Rood, J. I., . . . Dupuy, B. (2002). Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. *Journal of bacteriology, 184*(21), 5971-5978.

Maresso, A. W., Chapa, T. J., & Schneewind, O. (2006). Surface protein IsdC and Sortase B are required for heme-iron scavenging of Bacillus anthracis. *Journal of bacteriology, 188*(23), 8145-8152.

Martin-Verstraete, I., Peltier, J., & Dupuy, B. (2016). The regulatory networks that control *Clostridium difficile* toxin synthesis. *Toxins, 8*(5), 153.

Martin, M. J., Clare, S., Goulding, D., Faulds-Pain, A., Barquist, L., Browne, H. P., . . . Wren, B. W. (2013). The agr locus regulates virulence and colonization genes in *Clostridium difficile* 027. *Journal of bacteriology, 195*(16), 3672-3681.

Matamouros, S., England, P., & Dupuy, B. (2007). *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Molecular microbiology, 64*(5), 1274-1288.

Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E., & Schneewind, O. (2000). Staphylococcus aureus sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proceedings of the National Academy of Sciences, 97*(10), 5510-5515.

Mazmanian, S. K., Liu, G., Ton-That, H., & Schneewind, O. (1999). Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science, 285*(5428), 760-763.

McDonald, L. C., Killgore, G. E., Thompson, A., Owens Jr, R. C., Kazakova, S. V., Sambol, S. P., . . . Gerding, D. N. (2005). An epidemic, toxin gene–variant strain of *Clostridium difficile*. *New England Journal of Medicine, 353*(23), 2433-2441.

Melville, S., & Craig, L. (2013). Type IV pili in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews, 77*(3), 323-341.

Mena, K. D., & Gerba, C. P. (2009). Risk assessment of Pseudomonas aeruginosa in water *Reviews of Environmental Contamination and Toxicology Vol 201* (pp. 71-115): Springer.

Méndez, M. B., Orsaria, L. M., Philippe, V., Pedrido, M. E., & Grau, R. R. (2004). Novel roles of the master transcription factors Spo0A and σB for survival and sporulation of Bacillus subtilis at low growth temperature. *Journal of bacteriology, 186*(4), 989-1000.

Merrigan, M. M., Venugopal, A., Roxas, J. L., Anwar, F., Mallozzi, M. J., Roxas, B. A., . . . Vedantam, G. (2013). Surface-layer protein A (SlpA) is a major contributor to host-cell adherence of *Clostridium difficile*. *PloS one, 8*(11), e78404.

Mitchell, G., Fugère, A., Gaudreau, K. P., Brouillette, E., Frost, E. H., Cantin, A. M., & Malouin, F. (2013). SigB is a dominant regulator of virulence in Staphylococcus aureus small-colony variants. *PloS one, 8*(5), e65018.

Mitchell, L. S., & Tuomanen, E. I. (2002). Molecular analysis of antibiotic tolerance in pneumococci. *International journal of medical microbiology, 292*(2), 75-79.

Mitchell, T., Ketley, J., Haslam, S., Stephen, J., Burdon, D., Candy, D., & Daniel, R. (1986). Effect of toxin A and B of *Clostridium difficile* on rabbit ileum and colon. *Gut, 27*(1), 78-85.

Moisan, H., Brouillette, E., Jacob, C. L., Langlois-Bégin, P., Michaud, S., & Malouin, F. (2006). Transcription of virulence factors in Staphylococcus aureus small-colony variants isolated from cystic fibrosis patients is influenced by SigB. *Journal of bacteriology, 188*(1), 64-76.

Monds, R. D., & O'Toole, G. A. (2009). The developmental model of microbial biofilms: ten years of a paradigm up for review. *Trends in microbiology, 17*(2), 73-87.

Mooyottu, S., Flock, G., Kollanoor-Johny, A., Upadhyaya, I., Jayarao, B., & Venkitanarayanan, K. (2015). Characterization of a multidrug resistant *C. difficile* meat isolate. *International journal of food microbiology, 192*, 111-116.

Nadon, C. A., Bowen, B. M., Wiedmann, M., & Boor, K. J. (2002). Sigma B contributes to PrfA-mediated virulence in Listeria monocytogenes. *Infection and immunity, 70*(7), 3948-3952.

Nakayama, J., Chen, S., Oyama, N., Nishiguchi, K., Azab, E. A., Tanaka, E., . . . Sonomoto, K. (2006). Revised model for Enterococcus faecalis fsr quorum-sensing system: the small open reading frame fsrD encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal AgrD. *Journal of bacteriology, 188*(23), 8321-8326.

Nannapaneni, P., Hertwig, F., Depke, M., Hecker, M., Mäder, U., Völker, U., . . . van Hijum, S. A. (2012). Defining the structure of the general stress regulon of Bacillus subtilis using targeted microarray analysis and random forest classification. *Microbiology, 158*(3), 696-707.

Nawrocki, K. L., Edwards, A. N., Daou, N., Bouillaut, L., & McBride, S. M. (2016). CodY-dependent regulation of sporulation in *Clostridium difficile*. *Journal of bacteriology, 198*(15), 2113-2130.

Nealson, K., & Hastings, J. W. (1979). Bacterial bioluminescence: its control and ecological significance. *Microbiological reviews, 43*(4), 496.

Nealson, K. H., Platt, T., & Hastings, J. W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *Journal of bacteriology, 104*(1), 313-322.

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., . . . Minton, N. P. (2013). Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using pyrE alleles. *PloS one*, *8*(2), e56051.

Nicholas, R., Li, T., McDevitt, D., Marra, A., Sucoloski, S., Demarsh, P., & Gentry, D. (1999). Isolation and Characterization of a sigBDeletion Mutant of Staphylococcus aureus. *Infection and immunity, 67*(7), 3667-3669.

Nikolskaya, A. N., & Galperin, M. Y. (2002). A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic acids research*, *30*(11), 2453-2459.

Novak, R., Henriques, B., Charpentier, E., Normark, S., & Tuomanen, E. (1999). Emergence of vancomycin tolerance in Streptococcus pneumoniae. *Nature, 399*(6736), 590-593.

Novick, R. P., & Geisinger, E. (2008). Quorum sensing in staphylococci. *Annual review of genetics, 42,* 541-564.

O'Neil, H. S., & Marquis, H. (2006). Listeria monocytogenes flagella are used for motility, not as adhesins, to increase host cell invasion. *Infection and immunity*, *74*(12), 6675-6681.

Oh, J.-H., & van Pijkeren, J.-P. (2014). CRISPR–Cas9-assisted recombineering in Lactobacillus reuteri. *Nucleic acids research*, gku623.

Ohtani, K., Yuan, Y., Hassan, S., Wang, R., Wang, Y., & Shimizu, T. (2009). Virulence gene regulation by the agr system in Clostridium perfringens. *Journal of bacteriology, 191*(12), 3919-3927.

Okada, M., Sato, I., Cho, S. J., Iwata, H., Nishio, T., Dubnau, D., & Sakagami, Y. (2005). Structure of the Bacillus subtilis quorum-sensing peptide pheromone ComX. *Nature chemical biology, 1*(1), 23-24.

Olivas, A. D., Shogan, B. D., Valuckaite, V., Zaborin, A., Belogortseva, N., Musch, M., . . . Gilbert, J. (2012). Intestinal tissues induce an SNP mutation in Pseudomonas aeruginosa that enhances its virulence: possible role in anastomotic leak. *PloS one*, *7*(8), e44326.

Olling, A., Seehase, S., Minton, N. P., Tatge, H., Schröter, S., Kohlscheen, S., . . . Gerhard, R. (2012). Release of TcdA and TcdB from *Clostridium difficile* cdi 630 is not affected by functional inactivation of the tcdE gene. *Microbial pathogenesis, 52*(1), 92-100.

Ormonde, P., Hörstedt, P., O'Toole, R., & Milton, D. L. (2000). Role of motility in adherence to and invasion of a fish cell line by Vibrio anguillarum. *Journal of bacteriology, 182*(8), 2326-2328.

Owens, R. C., Donskey, C. J., Gaynes, R. P., Loo, V. G., & Muto, C. A. (2008). Antimicrobial-associated risk factors for *Clostridium difficile* infection. *Clinical Infectious Diseases, 46*(Supplement 1), S19-S31.

Paget, M. S., & Helmann, J. D. (2003). The σ 70 family of sigma factors. *Genome biology, 4*(1), 203.

Palma, M., & Cheung, A. L. (2001). ςB activity in Staphylococcus aureus is controlled by RsbU and an additional factor (s) during bacterial growth. *Infection and immunity, 69*(12), 7858-7865.

Pané-Farré, J., Jonas, B., Hardwick, S. W., Gronau, K., Lewis, R. J., Hecker, M., & Engelmann, S. (2009). Role of RsbU in controlling SigB activity in Staphylococcus aureus following alkaline stress. *Journal of bacteriology, 191*(8), 2561-2573.

Papatheodorou, P., Carette, J. E., Bell, G. W., Schwan, C., Guttenberg, G., Brummelkamp, T. R., & Aktories, K. (2011). Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin *Clostridium difficile* transferase (CDT). *Proceedings of the National Academy of Sciences, 108*(39), 16422-16427.

Papatheodorou, P., Hornuss, D., Nölke, T., Hemmasi, S., Castonguay, J., Picchianti, M., & Aktories, K. (2013). *Clostridium difficile* binary toxin CDT induces clustering of the lipolysis-stimulated lipoprotein receptor into lipid rafts. *MBio, 4*(3), e00244-00213.

Papatheodorou, P., Zamboglou, C., Genisyuerek, S., Guttenberg, G., & Aktories, K. (2010). Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. *PloS one, 5*(5), e10673.

Paredes-Sabja, D., Shen, A., & Sorg, J. A. (2014). *Clostridium difficile* spore biology: sporulation, germination, and spore structural proteins. *Trends in microbiology, 22*(7), 406-416.

Paredes, C. J., Alsaker, K. V., & Papoutsakis, E. T. (2005). A comparative genomic view of clostridial sporulation and physiology. *Nature Reviews Microbiology, 3*(12), 969-978.

Park, P. J. (2009). ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics, 10*(10), 669-680.

Péchiné, S., Hennequin, C., Boursier, C., Hoys, S., & Collignon, A. (2013). Immunization using GroEL decreases *Clostridium difficile* intestinal colonization. *PloS one, 8*(11), e81112.

Peltier, J., Shaw, H. A., Couchman, E. C., Dawson, L. F., Yu, L., Choudhary, J. S., . . . Fairweather, N. F. (2015). Cyclic diGMP regulates production of sortase substrates of *Clostridium difficile* and their surface exposure through ZmpI protease-mediated cleavage. *Journal of Biological Chemistry, 290*(40), 24453-24469.

Perego, M. (2001). A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis. *Molecular microbiology, 42*(1), 133-143.

Pettit, L. J., Browne, H. P., Yu, L., Smits, W. K., Fagan, R. P., Barquist, L., . . . Flint, H. J. (2014). Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. *BMC genomics, 15*(1), 160.

Piepenbrink, K. H., & Sundberg, E. J. (2016). Motility and adhesion through type IV pili in Gram-positive bacteria. *Biochemical Society Transactions, 44*(6), 1659-1666.

Pituch, H., Van Belkum, A., Van Den Braak, N., Obuch-Woszczatyñski, P., Sawicka-Grzelak, A., Verbrugh, H., & Meisel-Mikołajczyk, F. (2003). Clindamycin-resistant, toxin A-negative, toxin B-positive *Clostridium difficile* strains cause antibiotic-associated diarrhea among children hospitalized in a hematology unit. *Clinical Microbiology and Infection, 9*(8), 903-904.

Pizarro-Guajardo, M., Olguín-Araneda, V., Barra-Carrasco, J., Brito-Silva, C., Sarker, M. R., & Paredes-Sabja, D. (2014). Characterization of the collagen-like exosporium protein, BclA1, of *Clostridium difficile* spores. *Anaerobe, 25*, 18-30.

Pohl, K., Francois, P., Stenz, L., Schlink, F., Geiger, T., Herbert, S., . . . Wolz, C. (2009). CodY in Staphylococcus aureus: a regulatory link between metabolism and virulence gene expression. *Journal of bacteriology, 191*(9), 2953-2963.

Pomerantsev, A. P., Pomerantseva, O. M., Camp, A. S., Mukkamala, R., Goldman, S., & Leppla, S. H. (2009). PapR peptide maturation: role of the NprB protease in Bacillus cereus 569 PlcR/PapR global gene regulation. *FEMS Immunology & Medical Microbiology, 55*(3), 361-377.

Popoff, M. R., Rubin, E., Gill, D., & Boquet, P. (1988). Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* strain. *Infection and immunity, 56*(9), 2299-2306.

Pruitt, R. N., Chagot, B., Cover, M., Chazin, W. J., Spiller, B., & Lacy, D. B. (2009). Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in *Clostridium difficile* toxin A. *Journal of Biological Chemistry, 284*(33), 21934-21940.

Pruitt, R. N., Chambers, M. G., Ng, K. K.-S., Ohi, M. D., & Lacy, D. B. (2010). Structural organization of the functional domains of *Clostridium difficile* toxins A and B. *Proceedings of the National Academy of Sciences, 107*(30), 13467-13472.

Pruitt, R. N., & Lacy, D. B. (2012). Toward a structural understanding of *Clostridium difficile* toxins A and B. *Front Cell Infect Microbiol, 2*(March), 28.

Purdy, D., O'Keeffe, T. A., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., . . . Minton, N. P. (2002). Conjugative transfer of clostridial shuttle vectors from Escherichia coli to *Clostridium difficile* through circumvention of the restriction barrier. *Molecular microbiology, 46*(2), 439-452.

Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A., & Chou, C. P. (2016). Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium. *Scientific reports, 6*.

Qa'Dan, M., Spyres, L. M., & Ballard, J. D. (2000). pH-Induced Conformational Changes in*Clostridium difficile* Toxin B. *Infection and immunity, 68*(5), 2470-2474.

Redelings, M. D., Sorvillo, F., & Mascola, L. (2007). Increase in *Clostridium difficile*–related mortality rates, United States, 1999–2004. *Emerging infectious diseases, 13*(9), 1417.

Reder, A., Albrecht, D., Gerth, U., & Hecker, M. (2012). Cross-talk between the general stress response and sporulation initiation in Bacillus subtilis–the σB promoter of spo0E represents an AND-gate. *Environmental microbiology, 14*(10), 2741-2756.

Reder, A., Gerth, U., & Hecker, M. (2012). Integration of σB activity into the decision-making process of sporulation initiation in Bacillus subtilis. *Journal of bacteriology, 194*(5), 1065-1074.

Reynolds, C. B., Emerson, J. E., de la Riva, L., Fagan, R. P., & Fairweather, N. F. (2011). The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. *PLoS Pathog, 7*(4), e1002024.

Rigel, N. W., & Braunstein, M. (2008). A new twist on an old pathway–accessory secretion systems. *Molecular microbiology, 69*(2), 291-302.

Riley, T., O'Neill, G., Bowman, R., & Golledge, C. (1994). *Clostridium difficile*-associated diarrhoea: epidemiological data from Western Australia. *Epidemiology and infection, 113*(01), 13-20.

Rosenbusch, K. E., Bakker, D., Kuijper, E. J., & Smits, W. K. (2012). *C. difficile* 630Δerm Spo0A regulates sporulation, but does not contribute to toxin production, by direct high-affinity binding to target DNA. *PloS one, 7*(10), e48608.

Roux, A., Todd, D. A., Velázquez, J. V., Cech, N. B., & Sonenshein, A. L. (2014). CodY-mediated regulation of the Staphylococcus aureus Agr system integrates nutritional and population density signals. *Journal of bacteriology, 196*(6), 1184-1196.

Rupnik, M., Avesani, V., Janc, M., von Eichel-Streiber, C., & Delmée, M. (1998). A Novel Toxinotyping Scheme and Correlation of Toxinotypes with Serogroups of *Clostridium difficile*Isolates. *Journal of clinical microbiology, 36*(8), 2240-2247.

Rutherford, S. T., & Bassler, B. L. (2012). Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine, 2*(11), a012427.

Said-Salim, B., Dunman, P., McAleese, F., Macapagal, D., Murphy, E., McNamara, P., . . . Kreiswirth, B. (2003). Global regulation of Staphylococcus aureus genes by Rot. *Journal of bacteriology, 185*(2), 610-619.

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*(13), 3186-3196.

Saujet, L., Pereira, F. C., Serrano, M., Soutourina, O., Monot, M., Shelyakin, P. V., . . . Martin-Verstraete, I. (2013). Genome-wide analysis of cell type-specific gene transcription during spore formation in *Clostridium difficile*. *PLoS Genet, 9*(10), e1003756.

Savariau-Lacomme, M.-P., Lebarbier, C., Karjalainen, T., Collignon, A., & Janoir, C. (2003). Transcription and analysis of polymorphism in a cluster of genes encoding surface-associated proteins of *Clostridium difficile*. *Journal of bacteriology, 185*(15), 4461-4470.

Savidge, T. C., Pan, W.-h., Newman, P., O'Brien, M., Anton, P. M., & Pothoulakis, C. (2003). *Clostridium difficile* toxin B is an inflammatory enterotoxin in human intestine. *Gastroenterology, 125*(2), 413-420.

Sayedy, L., Kothari, D., & Richards, R. J. (2010). Toxic megacolon associated *Clostridium difficile* colitis. *World J Gastrointest Endosc, 2*(8), 293-297.

Schaefer, A. L., Hanzelka, B. L., Eberhard, A., & Greenberg, E. (1996). Quorum sensing in Vibrio fischeri: probing autoinducer-LuxR interactions with autoinducer analogs. *Journal of bacteriology, 178*(10), 2897-2901.

Schauder, S., Shokat, K., Surette, M. G., & Bassler, B. L. (2001). The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Molecular microbiology, 41*(2), 463-476.

Schneewind, O., & Missiakas, D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1592), 1123-1139. doi:10.1098/rstb.2011.0210

Schuster, M., & Greenberg, E. P. (2006). A network of networks: quorum-sensing gene regulation in Pseudomonas aeruginosa. *International journal of medical microbiology, 296*(2), 73-81.

Schwab, U., Bowen, B., Nadon, C., Wiedmann, M., & Boor, K. J. (2005). The Listeria monocytogenes prfAP2 promoter is regulated by sigma B in a growth phase dependent manner. *FEMS Microbiology Letters, 245*(2), 329-336.

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., STabler, R., . . . Wang, H. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature genetics*, *38*(7), 779-786.

Seddon, S., Hemingway, I., & Borriello, S. (1990). Hydrolytic enzyme production by *Clostridium difficile* and its relationship to toxin production and virulence in the hamster model. *Journal of medical microbiology, 31*(3), 169-174.

Sekulovic, O., Ospina Bedoya, M., Fivian-Hughes, A. S., Fairweather, N. F., & Fortier, L. C. (2015). The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. *Molecular microbiology, 98*(2), 329-342.

Semenyuk, E. G., Laning, M. L., Foley, J., Johnston, P. F., Knight, K. L., Gerding, D. N., & Driks, A. (2014). Spore formation and toxin production in *Clostridium difficile* biofilms. *PloS one, 9*(1), e87757.

Semenyuk, E. G., Poroyko, V. A., Johnston, P. F., Jones, S. E., Knight, K. L., Gerding, D. N., & Driks, A. (2015). Analysis of bacterial communities during *Clostridium difficile* infection in the mouse. *Infection and immunity*, *83*(11), 4383-4391.

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. *Trends in microbiology*, *15*(4), 172-180.

Sheng, L., Kovács, K., Winzer, K., Zhang, Y., & Minton, N. P. (2017). Development and implementation of rapid metabolic engineering tools for chemical and fuel production in Geobacillus thermoglucosidasius NCIMB 11955. *Biotechnology for biofuels, 10*(1), 5.

Slamti, L., & Lereclus, D. (2002). A cell–cell signaling peptide activates the PlcR virulence regulon in bacteria of the Bacillus cereus group. *The EMBO journal, 21*(17), 4550-4559.

Soehn, F., Wagenknecht-Wiesner, A., Leukel, P., Kohl, M., Weidmann, M., von Eichel-Streiber, C., & Braun, V. (1998). Genetic rearrangements in the pathogenicity locus of *Clostridium difficile* strain 8864–implications for transcription, expression and enzymatic activity of toxins A and B. *Molecular and General Genetics MGG, 258*(3), 222-232.

Sonenshein, A. L. (2000). Control of sporulation initiation in Bacillus subtilis. *Current opinion in microbiology, 3*(6), 561-566.

Sonenshein, A. L. (2005). CodY, a global regulator of stationary phase and virulence in Gram-positive bacteria. *Current opinion in microbiology, 8*(2), 203-207.

Sonenshein, A. L., Hoch, J. A., & Losick, R. (2002). Bacillus subtilis: from cells to genes and from genes to cells *Bacillus subtilis and Its Closest Relatives* (pp. 3-5): American Society of Microbiology.

Sorek, R., Lawrence, C. M., & Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annual review of biochemistry, 82*, 237-266.

Sorg, J. A., & Sonenshein, A. L. (2008). Bile salts and glycine as cogerminants for *Clostridium difficile* spores. *Journal of bacteriology, 190*(7), 2505-2512.

Soutourina, O. A., Monot, M., Boudry, P., Saujet, L., Pichon, C., Sismeiro, O., . . . Coppée, J.-Y. (2013). Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet, 9*(5), e1003493.

Sperandio, V., Mellies, J. L., Nguyen, W., Shin, S., & Kaper, J. B. (1999). Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic Escherichia coli. *Proceedings of the National Academy of Sciences, 96*(26), 15196-15201.

Spigaglia, P., Barketi-Klai, A., Collignon, A., Mastrantonio, P., Barbanti, F., Rupnik, M., . . . Kansau, I. (2013). Surface-layer (S-layer) of human and animal *Clostridium difficile* strains and their behaviour in adherence to epithelial cells and intestinal colonization. *Journal of medical microbiology, 62*(9), 1386-1393.

Spirig, T., Weiner, E. M., & Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Molecular microbiology, 82*(5), 1044-1059.

STabler, R., Gerding, D., Songer, J., Drudy, D., Brazier, J., Trinh, H., . . . Wren, B. (2006). Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. *Journal of bacteriology, 188*(20), 7297-7305.

STabler, R. A., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., . . . Rose, G. (2009). Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol, 10*(9), R102.

Staneck, J. L., Weckbach, L. S., Allen, S. D., Siders, J. A., Gilligan, P. H., Coppitt, G., . . . Willis, D. H. (1996). Multicenter evaluation of four methods for *Clostridium difficile* detection: ImmunoCard *C. difficile*, cytotoxin assay, culture, and latex agglutination. *Journal of clinical microbiology, 34*(11), 2718-2721.

Steele, J., Mukherjee, J., Parry, N., & Tzipori, S. (2013). Antibody against TcdB, but not TcdA, prevents development of gastrointestinal and systemic *Clostridium difficile* disease. *Journal of Infectious Diseases, 207*(2), 323-330.

Steiner, E., Scott, J., Minton, N. P., & Winzer, K. (2012). An agr quorum sensing system that regulates granulose formation and sporulation in Clostridium acetobutylicum. *Applied and environmental microbiology, 78*(4), 1113-1122.

Stock, A. M., Robinson, V. L., & Goudreau, P. N. (2000). Two-component signal transduction. *Annual review of biochemistry, 69*(1), 183-215.

Stubbs, S., Rupnik, M., Gibert, M., Brazier, J., Duerden, B., & Popoff, M. (2000). Production of actin-specific ADP-ribosyltransferase (binary toxin) by strains of *Clostridium difficile*. *FEMS Microbiology Letters, 186*(2), 307-312.

Surette, M. G., Miller, M. B., & Bassler, B. L. (1999). Quorum sensing in Escherichia coli, Salmonella typhimurium, and Vibrio harveyi: a new family of genes responsible for autoinducer production. *Proceedings of the National Academy of Sciences, 96*(4), 1639-1644.

Sylvestre, P., Couture-Tosi, E., & Mock, M. (2002). A collagen-like surface glycoprotein is a structural component of the Bacillus anthracis exosporium. *Molecular microbiology, 45*(1), 169-178.

Tasteyre, A., Barc, M.-C., Karjalainen, T., Dodson, P., Hyde, S., Bourlioux, P., & Borriello, P. (2000). A *Clostridium difficile* gene encoding flagellin. *Microbiology, 146*(4), 957-966.

Tasteyre, A., Barc, M., Dodson, P., Avesani, V., Hyde, S., & Boriello, S. (1997). Isolation of a genetic determinant coding for *Clostridium difficile* flagellin and its relation to different serogroups. *Biosci Microflora, 16*(suppl 19), 19-20.

Tasteyre, A., Karjalainen, T., Avesani, V., Delmée, M., Collignon, A., Bourlioux, P., & Barc, M.-C. (2001). Molecular Characterization of fliD Gene Encoding Flagellar Cap and Its Expression among *Clostridium difficile*Isolates from Different Serogroups. *Journal of clinical microbiology, 39*(3), 1178-1183.

Totten, M., Gregg, J., Fremont-Smith, P., & Legg, M. (1978). Clinical and pathological spectrum of antibiotic-associated colitis. *The American journal of gastroenterology, 69*(3 Pt 1), 311.

Tucker, K. D., & Wilkins, T. D. (1991). Toxin A of *Clostridium difficile* binds to the human carbohydrate antigens I, X, and Y. *Infection and immunity, 59*(1), 73-78.

Tulli, L., Marchi, S., Petracca, R., Shaw, H. A., Fairweather, N. F., Scarselli, M., . . . Leuzzi, R. (2013). CbpA: a novel surface exposed adhesin of *Clostridium difficile* targeting human collagen. *Cellular microbiology, 15*(10), 1674-1687.

Underwood, S., Guan, S., Vijayasubhash, V., Baines, S. D., Graham, L., Lewis, R. J., . . . Stephenson, K. (2009). Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *Journal of bacteriology, 191*(23), 7296-7305.

van Schaik, W., van der Voort, M., Molenaar, D., Moezelaar, R., de Vos, W. M., & Abee, T. (2007). Identification of the σB regulon of Bacillus cereus and conservation of σB-regulated genes in low-GC-content gram-positive bacteria. *Journal of bacteriology, 189*(12), 4384-4390.

van Werkhoven, C. H., van der Tempel, J., Jajou, R., Thijsen, S. F., Diepersloot, R. J., Bonten, M. J., . . . Oosterheert, J. J. (2015). Identification of patients at high risk for *Clostridium difficile* infection: development and validation of a risk prediction model in hospitalized patients treated with antibiotics. *Clinical Microbiology and Infection, 21*(8), 786. e781-786. e788.

Vijay, K., Brody, M. S., Fredlund, E., & Price, C. W. (2000). A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the σB transcription factor of Bacillus subtilis. *Molecular microbiology, 35*(1), 180-188.

Voelker, U., Voelker, A., Maul, B., Hecker, M., Dufour, A., & Haldenwang, W. (1995). Separate mechanisms activate sigma B of Bacillus subtilis in response to environmental and metabolic stresses. *Journal of bacteriology, 177*(13), 3771-3780.

Voth, D. E., & Ballard, J. D. (2005). *Clostridium difficile* toxins: mechanism of action and role in disease. *Clinical Microbiology Reviews, 18*(2), 247-263.

Waligora, A.-J., Hennequin, C., Mullany, P., Bourlioux, P., Collignon, A., & Karjalainen, T. (2001). Characterization of a Cell Surface Protein of*Clostridium difficile* with Adhesive Properties. *Infection and immunity, 69*(4), 2144-2153.

Wang, Y., Zhang, Z.-T., Seo, S.-O., Choi, K., Lu, T., Jin, Y.-S., & Blaschek, H. P. (2015). Markerless chromosomal gene deletion in Clostridium beijerinckii using CRISPR/Cas9 system. *Journal of biotechnology, 200*, 1-5.

Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., . . . McDonald, L. C. (2005). Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *The Lancet, 366*(9491), 1079-1084.

Wei, Y., Perez, L. J., Ng, W.-L., Semmelhack, M. F., & Bassler, B. L. (2011). Mechanism of Vibrio cholerae autoinducer-1 biosynthesis. *ACS chemical biology, 6*(4), 356-365.

West, A. H., & Stock, A. M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in biochemical sciences, 26*(6), 369-376.

Wilcox, M., Shetty, N., Fawley, W., Shemko, M., Coen, P., Birtles, A., . . . Green, S. (2012). Changing epidemiology of *Clostridium difficile* infection following the introduction of a national ribotyping-based surveillance scheme in England. *Clinical Infectious Diseases*, cis614.

Williams, P., & Cámara, M. (2009). Quorum sensing and environmental adaptation in Pseudomonas aeruginosa: a tale of regulatory networks and multifunctional signal molecules. *Current opinion in microbiology, 12*(2), 182-191.

Winzer, K., Hardie, K. R., & Williams, P. (2003). LuxS and autoinducer-2: their contribution to quorum sensing and metabolism in bacteria. *Advances in applied microbiology, 53*, 291-396.

Wise, A. A., & Price, C. W. (1995). Four additional genes in the sigB operon of Bacillus subtilis that control activity of the general stress factor sigma B in response to environmental signals. *Journal of bacteriology, 177*(1), 123-133.

Wright, A., Wait, R., Begum, S., Crossett, B., Nagy, J., Brown, K., & Fairweather, N. (2005). Proteomic analysis of cell surface proteins from *Clostridium difficile*. *Proteomics, 5*(9), 2443-2452.

Yang, X., Kang, C. M., Brody, M. S., & Price, C. W. (1996). Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes & development, 10*(18), 2265-2275.

Young, J. W., Locke, J. C., & Elowitz, M. B. (2013). Rate of environmental change determines stress response specificity. *Proceedings of the National Academy of Sciences, 110*(10), 4140-4145.

Zemanick, E. T., Sagel, S. D., & Harris, J. K. (2011). The airway microbiome in cystic fibrosis and implications for treatment. *Current opinion in pediatrics, 23*(3), 319-324.

Zhang, L., & Ji, G. (2004). Identification of a staphylococcal AgrB segment (s) responsible for group-specific processing of AgrD by gene swapping. *Journal of bacteriology, 186*(20), 6706-6713.

Zhang, Z., Meng, Q., Qiao, J., Yang, L., Cai, X., Wang, G., . . . Zhang, L. (2013). RsbV of Listeria monocytogenes contributes to regulation of environmental stress and virulence. *Archives of microbiology, 195*(2), 113-120.

Ziebandt, A. K., Becher, D., Ohlsen, K., Hacker, J., Hecker, M., & Engelmann, S. (2004). The influence of agr and σB in growth phase dependent regulation of virulence factors in Staphylococcus aureus. *Proteomics, 4*(10), 3034-3047.