THE UNIVERSITY *of* EDINBURGH

Development of a novel genetic system for generation of markerless deletions in *Clostridium difficile*

Elena-Stella Theophilou

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2013

DECLARATION

I declare that this is my own work and that any contribution made by other parties is clearly stated.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Ian Poxton and Garry Blakely, for giving me the opportunity for the PhD project, for guidance, support and much appreciated help throughout. I wish to acknowledge the staff in the Darwin and Chancellor's buildings for their assistance, when required. Many thanks also to students in Darwin for their friendship and academic contribution. Last but not least, I would like to thank my family for always being there for me.

ABSTRACT

C. difficile is an obligate anaerobic, Gram-positive, rod-shaped and spore-forming bacterium. It is a well-recognised causative agent of antibiotic-associated diarrhoea and pseudomembranous colitis. *C. difficile* has emerged as an important nosocomial pathogen in recent years, associated with considerable morbidity, mortality and economic burden. Despite its importance, functional genomic studies have been lagging behind in comparison to other enteric pathogens. This is attributed to the fact that *C. difficile* is difficult to manipulate genetically and the lack of robust, reproducible mutagenesis systems for many years. The ideal mutation for robust functional genomic studies is a markerless, in-frame deletion of the gene of interest. All systems developed for *C. difficile*, up to the start of this study, involve insertional inactivation of the gene of interest. This study describes the development of a novel genetic system for *C. difficile*, to create precise and markerless chromosomal deletions, using the meganuclease I-SceI. For validation of the system, the *addBA* genes in *C. difficile* were deleted. The AddAB enzyme complex is important in the survival of many bacteria, since it maintains genome integrity, by the repair of double-strand breaks. Deletion of *addBA* in *C. difficile* did not significantly affect growth and viability, but the mutant strains were sensitive to DNA damaging agents. In addition, it was shown that *C. difficile* is capable of initiating the SOS response after DNA damage and that AddAB is not necessary for the induction of this response. The genetic system was further optimised to delete type IV pili (TFP)-associated genes, particularly *pilT* (CD3505) and *pilA* (CD3507), to investigate twitching motility. TFP are important in virulence and pathogenesis of many bacteria

and twitching motility is often involved. TFP in *C. difficile* may be expressed *in vivo* during infection and may be involved in biofilm formation and colonization. To study potential TFP-mediated motility, a non-flagellated *C. difficile* strain was first constructed by deleting the *fliC* gene. The *pilT* gene, predicted to encode a protein involved in TFP retraction, was then deleted in the Δ*fliC* strain. A Δ*pilT* strain was also generated. Preliminary experimental work using these strains did not show any evidence for twitching motility and no difference between the Δ*pilT* strains and the parental strains. Examination of cells from the Δ*fliC* strain, under various conditions, did not reveal any pili, which indicates that TFP are regulated in *C. difficile* and that the TFP locus might be repressed at the transcriptional level. Preliminary work to investigate an intergenic region located upstream of the TFP locus in *C. difficile*, that might be involved in regulation, suggested that transcription is being initiated within a 500 bp region upstream of the CD3513 gene.

CONTENTS

FIGURES

Figures Chapter 1

Figures Chapter 3

Figures Chapter 4

CHAPTER 1: Main introduction

1.1: *Clostridium difficile* overview

Microbiology:

C. difficile is an obligate anaerobic, Gram-positive, rod-shaped and spore-forming bacterium. It was first isolated in 1935 from the faecal microbiota of healthy neonates and was identified in the late 1970s as the aetiological agent of antibiotic-associated pseudomembranous colitis (Bartlett *et al.*, 1978; George *et al.*, 1978; Larson *et al.*, 1978).

Pathogenesis:

Clinical manifestations of *C. difficile* infection (CDI) can range from asymptomatic carriage to diarrhoea. Severity of diarrhoea varies, from self-limiting to serious diarrhoea, with or without formation of pseudomembranes in the colon (**FIG 1.1**), with the possibility of toxic megacolon or bowel perforation, sepsis, shock and death (Poxton *et al.*, 2001; Rupnik *et al.*, 2009; Freiler *et al.*, 2001; Gebhard *et al.*, 1985; Kawamoto *et al.*, 1999; Riggs *et al.*, 2007; Rubin *et al.*, 1995; Shim *et al.*, 1998; Triadafilopoulos and Hallstone, 1991).

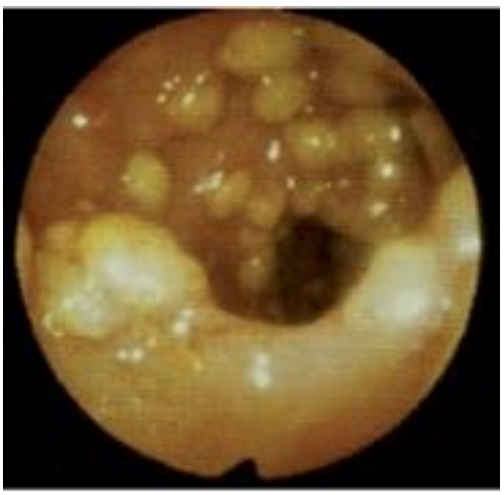


FIG 1.1 Pseudomembranous colitis

A typical endoscopic appearance of *C. difficile*-induced pseudomembranous colitis is shown. Pseudomembranes (yellow-white) formed are composed of destroyed intestinal cells and leukocytes. The figure is adapted from (Rupnik *et al.*, 2009).

C. difficile is acquired through ingestion of vegetative cells or spores and it is widely accepted that CDI is predominantly a toxin-mediated disease. Disruption of normal colonic microbiota and thus colonisation resistance allows spore germination (spores germinate in response to specific bile salts in the small intestine) and multiplication of vegetative cells (Vedantam *et al.*, 2012; Deneve *et al.*, 2009; Giel *et al.*, 2010; Sorg and Sonenshein, 2008). The organism then adheres to the mucus layer covering the epithelial surface of the gastrointestinal (GI) tract via multiple adhesins, penetrates the mucus and adheres to enterocytes marking the beginning of the first phase of the pathogenic process, colonization (Deneve *et al.*, 2009; Carroll and Bartlett, 2011). Several factors have been reported to be implicated in adhesion and colonization including, fibronectin-binding proteins Fbp68 and FbpA, surface layer and cell wall

proteins SlpA, Cwp84, Cw66, CwpV and heat shock protein GroEL (Vedantam *et al.*, 2012; Calabi *et al.*, 2002; Barketi-Klai *et al.*, 2011; Lin *et al.*, 2011; Waligora *et al.*, 2001; Hennequin *et al.*, 2001).

The second phase of the pathogenic process is toxin production (Carroll and Bartlett, 2011; Deneve *et al.*, 2009). Many *C. difficile* strains, but not all, produce two major toxins, termed toxin A (TcdA) and toxin B (TcdB), which are recognised as the main virulence factors. They belong to a group of large clostridial toxins (LCTs), due to their high molecular mass, TcdA being 308 kDa and TcdB being 270 kDa, and their production increases as cells enter the stationary phase of growth (Davies *et al.*, 2011; Rupnik *et al.*, 2009; Dupuy and Sonenshein, 1998; Hundsberger *et al.*, 1997; von Eichel-Streiber *et al.*, 1996). TcdA and TcdB are encoded by *tcdA* and *tcdB* genes, respectively and are both located in a 19.6 kb region known as the Pathogenicity Locus (PaLoc), (**FIG 1.2**) (Braun *et al.*, 1996; Cohen *et al.*, 2000). The PaLoc contains three additional genes, *tcdC* and *tcdR* which encode proteins involved in the regulation of toxin expression and *tcdE* which encodes a protein with pore-forming activity which allows release of toxins from the cell (Govind and Dupuy, 2012; Mani and Dupuy, 2001; Mani *et al.*, 2002; Matamouros *et al.*, 2007; Tan *et al.*, 2001). The PaLoc is conserved between toxigenic strains which encode one or both of the toxins. In non-toxigenic strains, the PaLoc locus is usually replaced by a 115-bp non-coding sequence and these strains are considered non-pathogenic (Rupnik *et al.*, 2009; Braun *et al.*, 1996; Rupnik, 2008).

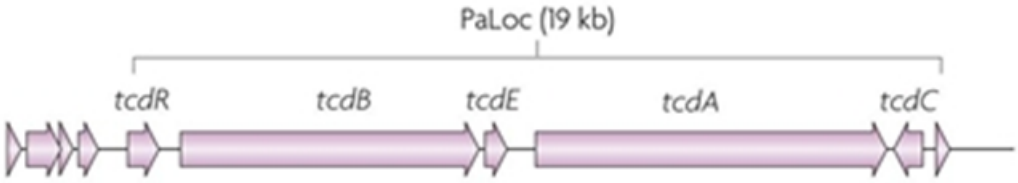


FIG 1.2 Schematic of Paloc in *C. difficile*

Paloc (19.6 kb) comprises of five genes *tcdA*, *tcdB*, *tcdR*, *tcdE* and *tcdC*. Toxins A and B are encoded from *tcdA* and *tcdB*, respectively. The genes *tcdR* and *tcdC* encode for proteins that positively and negatively regulate transcription of toxins, respectively. The *tcdE* gene encodes for a protein with pore-forming activity. The figure is adapted from (Rupnik *et al.*, 2009).

The actual role of toxin A or B in the disease pathogenesis of CDI remains a matter of debate. Two groups have constructed *C. difficile* mutant strains deficient in the production of TcdA, TcdB or both toxins, and tested their virulence in the hamster model (Lyras *et al.*, 2009; Kuehne *et al.*, 2010). One group concluded that TcdB and not TcdA is essential for virulence, while the other, concluded that both toxins are important. It was suggested that these differences may be attributed to inherent differences between the hamsters used in each study, or to differences between the parent *C. difficile* strains used to create the mutants (Kuehne *et al.*, 2010).

Toxins A and B are structurally similar and contain multiple domains including; i) N-terminal glycotransferase domain, ii) cysteine protease domain, iii) translocation hydrophobic domain and iv) C-terminal binding domain (**FIG 1.3**) (Davies *et al.*, 2011; Jank and Aktories, 2008). The interaction between the C-terminal binding domain of the toxin and host cell receptors initiates receptor-mediated endocytosis, although, the precise mechanism of toxin

uptake remains unclear (**FIG 1.3**) (Davies *et al.*, 2011; Oezguen *et al.*, 2012; Frisch *et al.*, 2003; Giesemann *et al.*, 2008; Greco *et al.*, 2006; Ho *et al.*, 2005; Jank *et al.*, 2007). Within the endosome, a decrease in pH causes conformational changes within the toxin, followed by pore formation in the endosomal membrane by the hydrophobic domain, and subsequent translocation of the N-terminal domain into the cytosol (Barth *et al.*, 2001; Qa'Dan *et al.*, 2000). Then, the toxin undergoes self-cleavage after a conformational change of the cysteine protease domain, a process that requires inositol hexakisphosphate ($InsP_6$), resulting in the release of the N-terminal domain into the cytosol (Egerer *et al.*, 2007; Giesemann *et al.*, 2008; Oezguen *et al.*, 2012; Pfeifer *et al.*, 2003; Pruitt *et al.*, 2009; Reineke *et al.*, 2007; Rupnik *et al.*, 2005). Once in the cytosol, the glycotransferase domain of the toxin, mono-O-glycosylates and inactivates Rho GTPases (Rho, Ras and Cdc42), the proteins that are important for maintaining cytoskeletal integrity (Davies *et al.*, 2011; Voth and Ballard, 2005; Oezguen *et al.*, 2012; Jank and Aktories, 2008; Jank *et al.*, 2007; Just *et al.*, 1995b; Just *et al.*, 1995a).

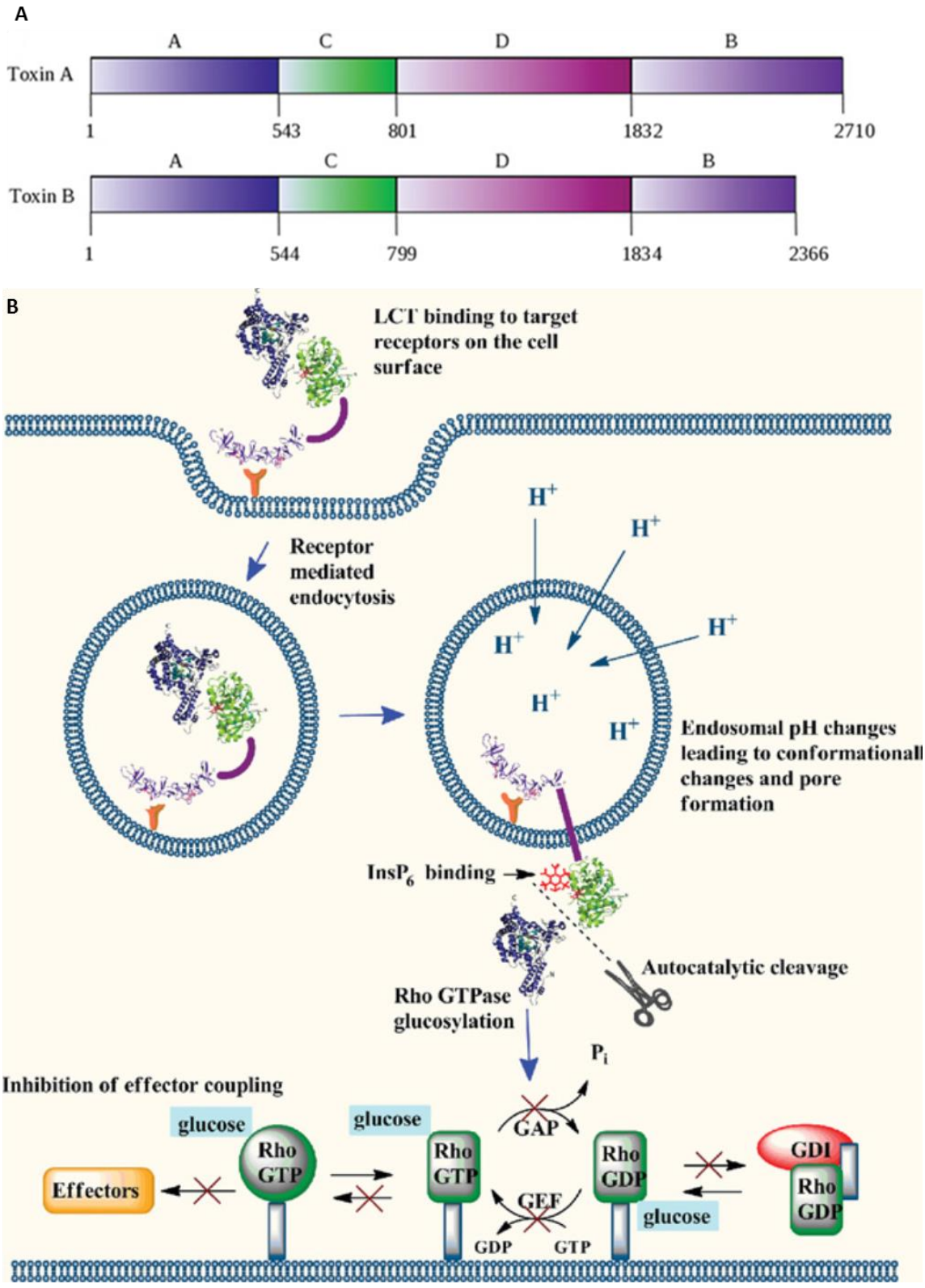


FIG 1.3 Schematic of functional domains of *C. difficile* toxin A and B and the processing machinery

A] The four functional domains of toxins A and B are shown. A (dark blue) refers to N- terminal glycotransferase domain. C (green) refers to cysteine protease domain. D (magenta) refers to the translocation hydrophobic domain. B (purple) refers to the C-terminal binding domain.

B] The processing machinery of toxins is shown. The colour scheme of the toxin domains is maintained in this schematic. The C-terminal binding domain binds to the host cell and the toxin is endocytosed. Whithin the endosome, a decrease in pH causes conformational changes within the toxin, followed by pore formation in the endosomal membrane by the hydrophobic domain, and subsequent translocation of the N-terminal domain into the cytosol. Then, the toxin undergoes self-cleavage after a conformational change of the cysteine protease domain, a process that requires inositol hexakisphosphate ($InsP_6$ binds to cystein protease domain), resulting in the release of the N-terminal domain into the cytosol. Once in the cytosol, the glycotransferase domain of the toxin, glycosylates and inactivates Rho GTPases. The figure is adapted from (Davies *et al.*, 2011).

Inactivation of Rho GTPases leads to loss of structural integrity, cell rounding and death of the intoxicated cell. The tight junctions between the intestinal epithelial cells are disturbed, increasing the permeability of colonic epithelial layers, allowing neutrophils to migrate to the intestines which contribute to the inflammatory response typical of colitis. Moreover, toxin exposure of intestinal epithelial cells leads to production of tumour necrosis factor and cytokines, which leads to fluid accumulation and further inflammatory response (Carroll and Bartlett, 2011; Voth and Ballard, 2005; Carter *et al.*, 2012). This intense inflammatory response is an important contributory factor in causing intestinal injury in CDI (Kelly and Kyne, 2011).

In addition to TcdA and TcdB, some *C. difficile* strains, produce the *C. difficile* transferase (CDT), otherwise named binary toxin. CDT is encoded by the Cdt locus (CdtLoc) which is 4.3 kb and comprised of the *cdtA*, *cdtB* and *cdtR* genes (**FIG 1.4**) (Carroll and Bartlett, 2011; Carter *et al.*, 2007; Carter *et al.*, 2012; Davies *et al.*, 2011; Perelle *et al.*, 1997; Popoff *et al.*, 1988; Rupnik *et al.*, 2009). The CdtLoc is found in 6-12.5% of strains overall and those that lack CDT have a 68-bp sequence in place of the CdtLoc

(Carroll and Bartlett, 2011). CDT is composed of two proteins, the enzymatic component CdtA encoded by *cdtA* and the transport and binding component CdtB encoded by *cdtB*. The *cdtR* gene encodes a positive regulator of CDT (Carter *et al.*, 2012; Rupnik *et al.*, 2009; Carter *et al.*, 2007).

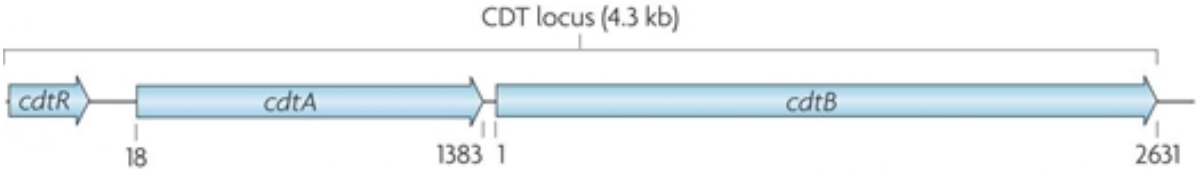


FIG 1.4 Schematic of CdtLoc in *C. difficile*

CdtLoc is 4.3 kb and comprises of the *cdtA*, *cdtB* and *cdtR* genes. The *cdtA* gene encodes the enzymatic component CdtA and the *cdtB* gene encodes the transport and binding component CdtB of CDT. The *cdtR* gene encodes for a positive regulator of CDT. The figure is adapted from (Rupnik *et al.*, 2009).

CdtB binds to the host cells and translocates CdtA into the cytosol where this protein ADP-ribosylates actin molecules, leading to disruption or rearrangement of the host cell cytoskeleton and subsequently excess fluid loss from the cells, rounding of the cell and cell death (Rupnik *et al.*, 2009; Vedantam *et al.*, 2012; Sundriyal *et al.*, 2010). The role of CDT in the pathogenesis of CDI has yet to be established. CDT has been shown to be cytotoxic for Vero cells *in vitro* (Sundriyal *et al.*, 2010). It has recently been shown that CDT induces the formation of microtubule structures that consist of long microtubule-based protrusions on the surface of epithelial cells, leading to increased adherence of *C. difficile* (Schwan *et al.*, 2009).

Risk factors, treatment, epidemiology, and economic impact:

Risk factors:

The most recognised risk factor associated with CDI is exposure to antibiotics. Almost every antibiotic has been associated with CDI, but some are linked with a higher risk of disease than others, including clindamycin, broad-spectrum cephalosporins and fluoroquinolones (Rupnik *et al.*, 2009; Ananthakrishnan, 2011; Freeman *et al.*, 2010; Kachrimanidou and Malisiovas, 2011; McFarland *et al.*, 1995; McFarland *et al.*, 1990). This higher risk relates to antibiotics that disrupt the normal microbiota allowing *C. difficile* to multiply and produce toxins (Ananthakrishnan, 2011; Owens *et al.*, 2008). In addition, *C. difficile* is resistant to a wide range of antibiotics allowing it to thrive and cause disease (Kachrimanidou and Malisiovas, 2011; Owens *et al.*, 2008).

Older age and hospitalization are other major risk factors. Patients older than 65 years of age are at 20-fold higher risk for disease than younger individuals (Bartlett and Gerding, 2008). In adults, the ability to produce adequate levels of IgG against toxin A defines whether the route to asymptomatic colonisation or CDI is followed (Kelly and Kyne, 2011; Kyne *et al.*, 2000, 2001). The increased susceptibility and mortality in elderly populations is possibly attributed to inadequate immune response to toxin exposure (Rupnik *et al.*, 2009; Kyne *et al.*, 2000). Hospitalization but also duration in a hospital or long-term care facility, increases the risk of CDI (Freeman *et al.*, 2010). This is attributed to a mixture of multiple risk factors in this setting including, exposure to antibiotics, exposure to a spore-contaminated environment, inefficient hand hygiene (e.g. alcohol-based hand gels are

ineffective against spores) and the presence of a susceptible elderly population (Rupnik *et al.*, 2009; Fawley and Wilcox, 2001; Debast *et al.*, 2009; Gerding *et al.*, 2008; Labbe *et al.*, 2008; Oughton *et al.*, 2009; Riggs *et al.*, 2007).

Other factors that have been linked to CDI include, immunosuppression, gastric acid suppressants such as H_2 blockers and proton pump inhibitors, gastrointestinal diseases such as inflammatory bowel disease and gastrointestinal surgery (Ananthakrishnan, 2011; Freeman *et al.*, 2010; Kachrimanidou and Malisiovas, 2011).

CDI can also occur in the community setting despite no recent exposure to antibiotics or hospitalization (Hensgens *et al.*, 2012). However, the incidence of CDI in the community is lower than the health care facilities. *C. difficile* is ubiquitous, it is found in soil, water, food products and it is a commensal or pathogen of the intestinal tracts of many animals. All can be possible community sources for CDI. However, direct transmission from these sources to humans has not been proven, although similar ribotypes have been found. As no outbreaks have been reported in the community, host factors may be more important for vulnerability to CDI than increased exposure to *C. difficile* (Hensgens *et al.*, 2012).

Treatment:

Treatment of CDI depends on the clinical presentation of disease (Kachrimanidou and Malisiovas, 2011; Ananthakrishnan, 2011). The first therapeutic step in most CDI cases is the immediate discontinuation of the inciting antibiotic. For mild or moderate CDI, metronidazole is used

as the first line drug and for severe CDI, vancomycin is the first choice. A combination of metronidazole and vancomycin is recommended in the most severe and complicated cases of CDI, i.e. those with hypotension, shock, ileus or megacolon. In the most severe cases, colectomy may be required (Ananthakrishnan, 2011; Kachrimanidou and Malisiovas, 2011). A novel drug, fidaxomicin, appears superior to vancomycin in treating CDI and preventing CDI relapses (Poxton, 2010). Faecal transfer from a healthy donor has also been shown to be effective in treating patients with recurrent CDI (Ananthakrishnan, 2011).

Epidemiology:

In the last decade, there has been a dramatic increase in the incidence and severity of CDI in the US, Canada and Europe (Kachrimanidou and Malisiovas, 2011; Carroll and Bartlett, 2011; Jones *et al.*, 2013; Freeman *et al.*, 2010). In 2003, the first reports were from Quebec, Canada which showed a 5-fold increase in incidence of CDI in the whole population from 1991 to 2003. In addition, an 8-fold increase in the incidence of CDI in the elderly was reported in North America (Pepin *et al.*, 2004). Similarly, in Europe the incidence rate of CDI increased with associated outbreaks, first in the UK from 2003 to 2004, then in the Netherlands and Belgium, then in France and other European countries (Deneve *et al.*, 2009; Kuijper *et al.*, 2007). In the United States, from 1999 to 2008, CDI-associated deaths increased 9-fold and from 2000 to 2008 CDI cases increased 2-fold (Lessa *et al.*, 2012).

The increase in incidence rate and severity of CDI has been partially attributed to the emergence of a hypervirulent

strain *C. difficile* PCR-ribotype 027. The hypervirulence of this strain is attributed to several factors: i) increased toxin A and B production in comparison to control strains, due to mutations in the *tcdC* gene that normally negatively regulates the *tcdA* and *tcdB* genes. There is an 18-bp deletion and single base pair deletion at position 117 relative to the *tcdC* ATG start codon, the latter resulting in formation of a stop codon and thus a truncated, inactive TcdC protein; ii) production of CDT in addition to toxins A and B; iii) altered surface proteins that increase adherence to the intestinal epithelium (Carroll and Bartlett, 2011; Deneve *et al.*, 2009; Vohra and Poxton, 2011; Curry *et al.*, 2007; MacCannell *et al.*, 2006; Warny *et al.*, 2005). A further feature that may be a contributing factor to the emergence of 027 strains is resistance to fluoroquinolones, associated with mutations in the *gyrA* and *gyrB* genes (Deneve *et al.*, 2009; McDonald *et al.*, 2005; Spigaglia *et al.*, 2008; Drudy *et al.*, 2006; Drudy *et al.*, 2007). Depending on the country, other *C. difficile* strains have been responsible for outbreaks and severe cases, including ribotypes 001, 018, 078 and 106 (Kachrimanidou and Malisiovas, 2011).

In the UK, several outbreaks have been reported, with a double outbreak at Stoke Mandeville Hospital (Buckinghamshire Hospitals NHS Trust) being prominent. The first phase, between October 2003 and June 2004, was caused by ribotype 027. This resulted in 174 cases and 19 (11%) deaths that were either definitely or probably due to *C. difficile*. The second phase occurred between October 2004 and June 2005, which involved 160 cases and 19 (12%) deaths (HCC, 2006; Smith, 2005). The outbreaks were a consequence of poor practice in infection control, poor environmental hygiene, and lack of facilities for patient isolation (HCC,

2006). Although, there has been a dramatic increase in the incidence of CDI in England and Wales from 1990 to 2006, the reported cases of CDI have been decreasing since 2007 (**FIG 1.5**) (Cartman *et al.*, 2010). In addition, the number of deaths involving *C. difficile*, increased from 2004 to 2007, from 2,238 to 8,324 and have decreased to 2,053 in 2011 (**FIG 1.6**) (Office for National Statistics, 2011).

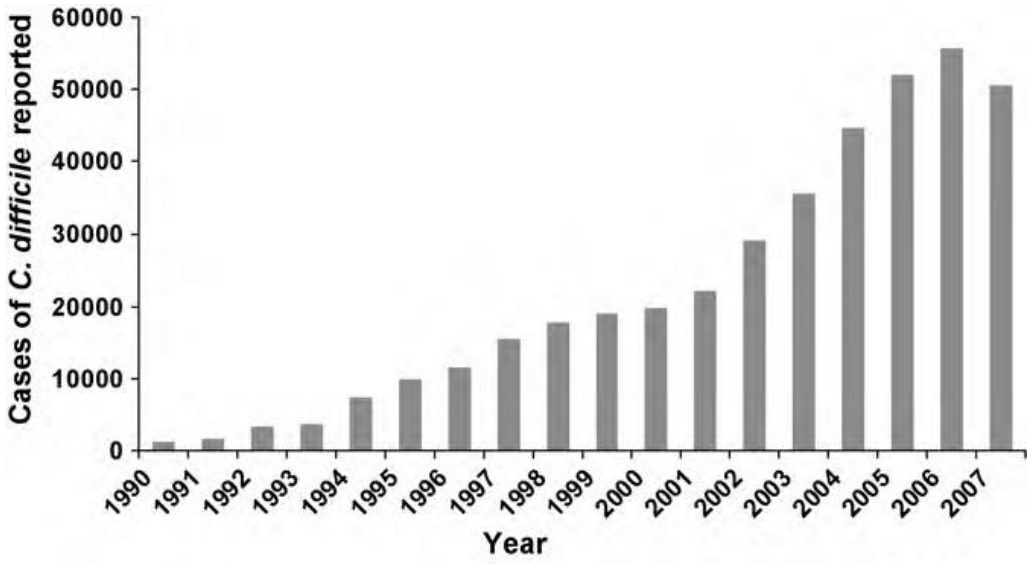


FIG 1.5 Cases of *C. difficile* in England and Wales

The increasing *C. difficile* incidence in England and Wales between 1990 and 2007 is shown. The figure was adapted from (Cartman *et al.*, 2010). Original source of data from The Health Protection Agency.

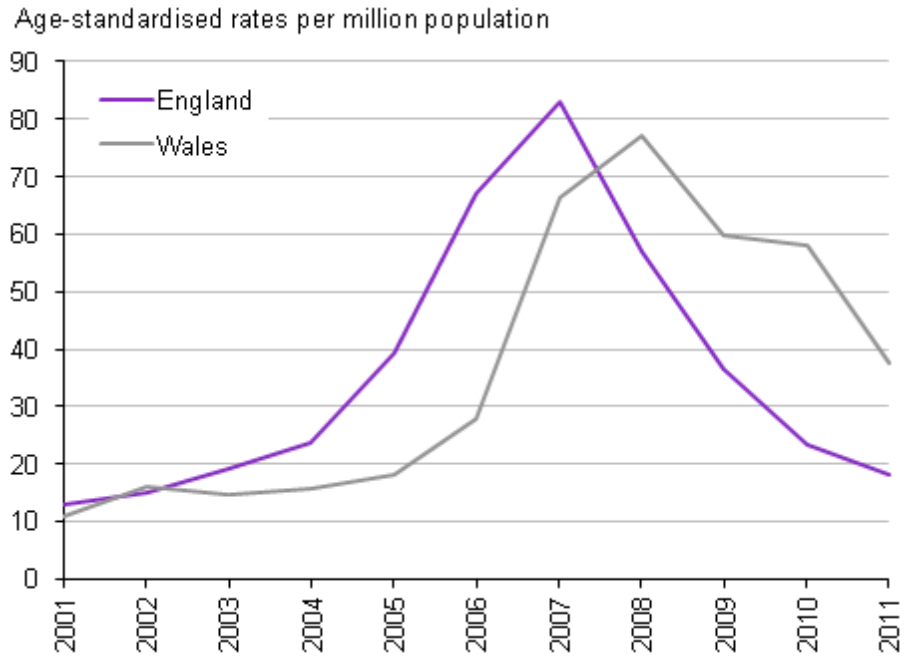


FIG 1.6 Mortality rates for deaths mentioning *C. difficile* in England and Wales

The number of death certificates in England and Wales mentioning *C. difficile* is shown between 2001 and 2011. Rates per million population are standardised to the European Standard population. The figure was adapted from the Office for National Statistics (http://www.ons.gov.uk/ons/dcp171778_276892.pdf).

The decrease in CDI cases is probably attributed to a combination of improved infection control, better environmental cleaning with chlorine-based disinfectants and good antibiotic stewardship, restricting the use of antibiotics that are known to precipitate CDI. As in England and Wales, since 2008, other European countries reported significant decrease in the overall number of CDI cases (Jones *et al.*, 2013). In Scotland, the predominant strains associated with recent outbreaks and severe cases have been ribotype 078 and to a lesser extent 027, and since 2009, CDI incidence rate and CDI-associated deaths have been decreasing (Health Protection Scotland, 2012; General Register Office for Scotland, 2001-2011). Similarly to ribotype 027, the ribotype 078 also contains mutations

in the *tcdC* gene of Paloc (39-bp deletion and single base pair deletion at position 184) and also contains binary toxin genes (Cartman *et al.*, 2010; Goorhuis *et al.*, 2008).

Cost:

The economic burden of CDI on the healthcare systems in Europe and US is high. The cost per case in England has been estimated to be €5,000-15,000. If the EU has 500 million in population, then a potential estimate for CDI per year could be €3000 million, a number that might double over the next four decades (Jones *et al.*, 2013). In the US the cost per year has been estimated to be $1.1 billion (Jones *et al.*, 2013). An economic computer simulation model suggested that the annual US economic burden for CDI would be, ≥$496 million from the hospital perspective, ≥$547 million from third-party payer perspective and ≥$796 million from the societal perspective (McGlone *et al.*, 2012).

1.2: Genetic manipulation of *C. difficile*

Historical overview of *C. difficile* genetic manipulation and mutagenesis systems:

Understanding the molecular basis of pathogenicity and virulence of a pathogen is of great importance as it is the way to provide knowledge for the development of measures for combatting disease. Despite the medical and economic importance of *C. difficile* and the availability of genomic and proteomic data, functional genomic studies have been lagging behind in comparison to other enteric pathogens

(Emerson *et al.*, 2008; Janvilisri *et al.*, 2009; Lawley *et al.*, 2009; Lemee *et al.*, 2005; Sebaihia *et al.*, 2006; Stabler *et al.*, 2006; Stabler *et al.*, 2009). This is attributed to the fact that *C. difficile* is difficult to manipulate genetically and the lack of robust and reproducible genetic systems for many years.

By 2007, only a handful of *C. difficile* mutants had been generated. For many years, until 2001, the only means of transfer of heterologous DNA into *C. difficile* was via conjugative transposons such as Tn916 with conjugative transfer involving *Bacillus subtilis* donors. In 1991, Mullany *et al.*, showed that the conjugative transposon Tn916 and a derivative Tn916ΔE can be transferred from *B. subtilis* into *C. difficile* CD37 by filter mating (Mullany *et al.*, 1991). The transposon appeared to insert at a preferred site, later designated *att*916, in all *C. difficile* CD37 transconjugants obtained and thus was not an appropriate candidate for random mutagenesis (Mullany *et al.*, 1991; Wang *et al.*, 2000a). However, it was suggested that it may be used for introduction of cloned DNA into *C. difficile* (Mullany *et al.*, 1991). In 1994, Mullany *et al.*, used the pC1195 vector containing a 4.2-kb region of the conjugative transposon Tn919, which is homologous to Tn916, to introduce heterologous DNA into *C. difficile* (Mullany *et al.*, 1994). A 1.1 kb fragment of *C. difficile* toxin B gene was inserted into pC1195, generating pPPM100, which was then introduced into *C. difficile* CD37 via filter mating from an isolated *B. subtilis* donor, into which the plasmid had integrated. The plasmid entered the genome of *C. difficile* CD37 strain at a specific site. The frequency of transfer was extremely low, at 10^{-8} per donor, and the delivery system involved was cumbersome. However, these

experiments were the first example of gene cloning in this organism.

In 2001, Liyanage *et al.*, were the first to introduce a replication deficient plasmid into *C. difficile* from an *Escherichia coli* donor. Importantly, they were the first to report a targeted gene disruption in *C. difficile* (Liyanage *et al.*, 2001). In their study, an internal fragment of the *C. difficile gldA* gene (encoding for glycerol dehydrogenase) was inserted into the *oriT*-based suicide vector pMTL31 and the resultant plasmid was conjugated from an *E. coli* donor into *C. difficile* CD37. Since the plasmid carried an erythromycin cassette, the integrants were selected by acquisition of erythromycin resistance. The plasmid was confirmed by PCR to have insertionally inactivated the *gldA* gene, via single cross-over integration. Interruption of the *gldA* gene, however, was lethal, as it gave rise to pinpoint colonies that could not be subcultured. This is the only report to date where a suicide vector has been used for the insertional inactivation of a gene in *C. difficile*. Subsequent, repeated efforts from other groups to inactivate genes using this method were unsuccessful, which raised questions about the reproducibility of this approach (Carter *et al.*, 2005; O'Connor *et al.*, 2006).

In 2003, to examine sporulation, Haraldsen and Sonenshein introduced a Tn916-based conjugative transposon carrying an intact copy of the *C. difficile sigK* gene, from *B. subtilis* into *C. difficile* CD196, that had a naturally inactive copy of *sigK* (Haraldsen and Sonenshein, 2003). Unexpectedly, the transposon did not insert at the preferred site, but integrated into the chromosome at the *sigK* gene location via homologous recombination. The reproducibility of this method, however, has not been established.

A more encouraging approach was described in 2006 by O'Connor *et al.*, that reported the first example of a reproducible method for the insertional inactivation of target genes in *C. difficile* JIR8094 (an erythromycin-sensitive derivative of *C. difficile* 630). In this case, the authors utilised the instability of *E. coli-Clostridium perfringens* shuttle vectors in *C. difficile*, such as pJIR1456 and pJIR2816, to inactivate two genes, *rgaR* and *rgbR*, via a homology-driven single cross-over event. However, in some cell populations of the *rgaR* mutant, there was reversion back to wild-type (O'Connor *et al.*, 2006). This is expected from mutants created via single cross-over recombination. Recently, the same group has used this method to insertionally inactivate the *tcdA* and *tcdB* genes in *C. difficile* JIR8094, in which case the mutants were stable (Lyras *et al.*, 2009). However, this method was also used by Dineen *et al.*, to inactivate the *C. difficile codY* gene, which emphasized the instability of the insertional inactivation by a single cross-over event (Dineen *et al.*, 2007). An internal region of the *C. difficile codY* gene was inserted into pJIR1456, generating pSD21. The mutant *C. difficile* JIR8094::pSD21 strain with interrupted *codY* gene was generated by the single cross-over integration of the recombinant plasmid pSD21 into the chromosome. This strain grew more slowly than the wild type and stable integration could only be maintained in the presence of antibiotic, since resistance was conferred by the integrated plasmid. In the absence of antibiotic selection, however, the *codY* mutant could revert back to wild type, with functional *codY* after excision of the integrated plasmid, and overpopulate the culture (Dineen *et al.*, 2007).

Heap *et al.* (2007), developed a mutagenesis system, named ClosTron, to overcome the instability observed with single

cross-over integration and to bypass the difficulty of isolating rare double cross-over events in the Clostridia; which had been only observed in *C. perfringens* (due to the high DNA transformation frequencies in this organism) (Heap *et al.*, 2007). This is to date the most widely used genetic system in *C. difficile* and since its appearance there has been a dramatic increase in the construction of *C. difficile* mutant strains. The ClosTron system was based on the TargeTron technology (Sigma-Aldrich) and has been modified for use in the Clostridia. In this approach, the mobile group II intron from the *ltrB* gene of *Lactococcus lactis* (Ll.ltrB) was utilised. Group II introns are catalytically active RNA molecules that can self-splice from an RNA transcript, via a lariat intermediate, and insert into a new target gene (Kuehne and Minton, 2012). These retro-homing ribozymes can be designed to target almost any gene of interest and thus interrupt the gene via insertion. The first ClosTron plasmid was named pMTL007 and was used to inactivate five genes in *C. difficile* and other Clostridia, including *C. acetobutylicum* and *C. sporogenes* (Heap *et al.*, 2007). Subsequently, the system was optimised to include features for a faster, less-labour intensive mutagenesis in the Clostridia. In addition, features were added for the possibility of interruption of multiple genes in a Clostridial strain of interest. The second generation ClosTron plasmids for mutagenesis in *C. difficile* strains were named pMTL007C-E2 and pMTL007C-E5 (**FIG 1.7**) (Heap *et al.*, 2010; Cartman *et al.*, 2010). The ClosTron plasmid pMTL007 and derivatives, pMTL007C-E2 and pMTL007C-E5, carry a group II intron which contains an *ermB* retrotransposition-activated selectable marker (RAM), which itself is interrupted by a group I intron that abolishes antibiotic resistance. Due to the relative orientations of the group II intron and RAM element, a functional *ermB* gene

is only produced if the group II intron inserts into the chromosomal target gene, following retrotransposition and self-splicing of the group I intron from the group II intron mRNA. The mutants are thus selected by resistance to erythromycin and then confirmed by PCR assay. The frequency of intron integration at the site of interest can range from 7% to 100%, depending on the gene of interest (Heap *et al.*, 2007; Heap *et al.*, 2010; Cartman *et al.*, 2010; Kuehne and Minton, 2012). The Clostron mutagenesis procedure is shown in **FIG 1.8**

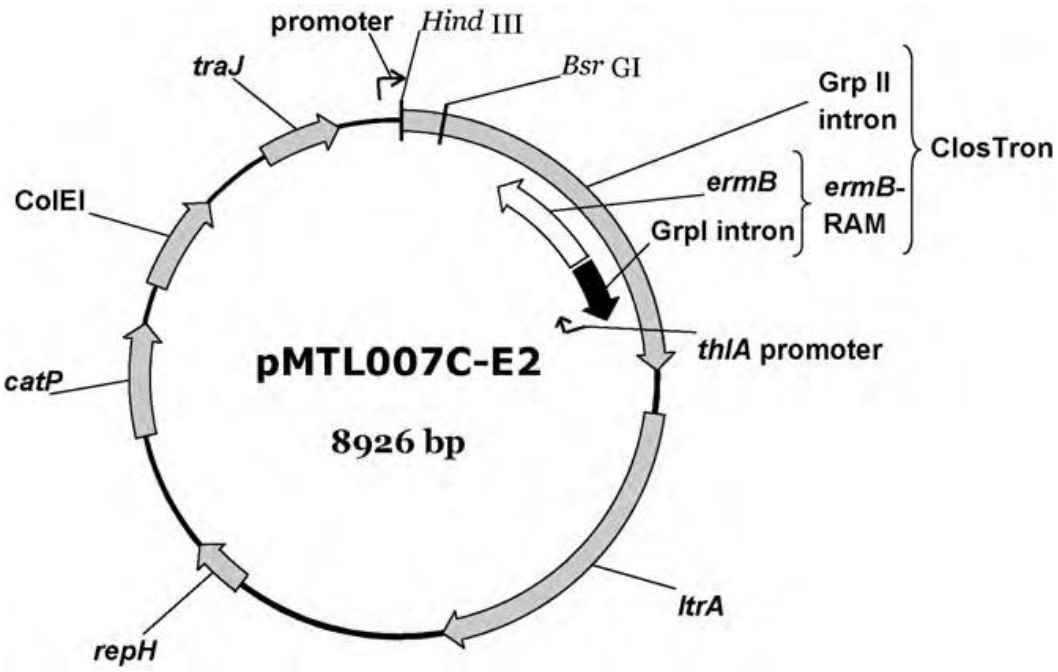


FIG 1.7 The second generation plasmid pMTL007C-EC2 of ClosTron

The Group II intron expression is directed by the constitutive promoter from the *fdx* gene of *Clostridium sporogenes*. The promoter of the *thlA* gene from *Clostridium acetobutylicum* is used for expression of the *ermB* element. The plasmid contains FRT sites flanking the RAM to facilitate FLP-mediated removal. It also features a *lacZ* stuffer sequence that is replaced during retargeting for the identification of re-targeted plasmids by blue/white screening, restriction analysis or PCR. The figure is adapted from (Cartman *et al.*, 2010).

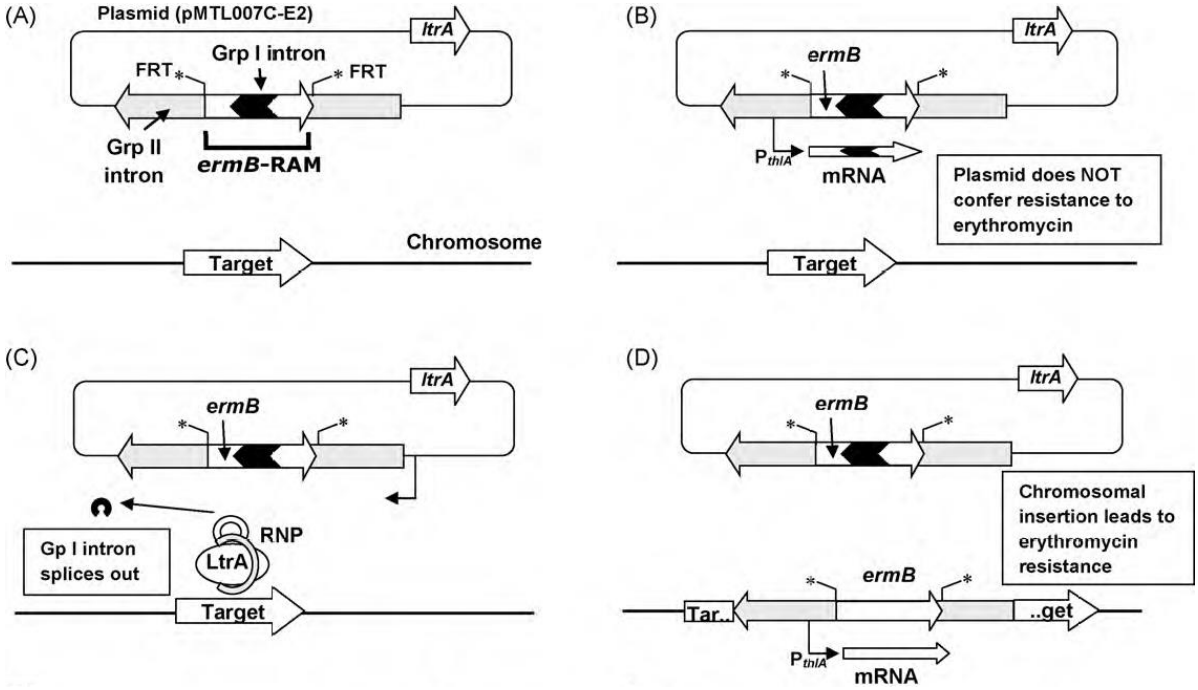


FIG 1.8 ClosTron mutagenesis procedure using pMTL007C-EC2

The ClosTron mutagenesis procedure involves inserting around a 350-bp fragment in an appropriate ClosTron plasmid (e.g pMTL007C-E2) which changes the specificity of the group II intron for insertion in the gene of interest. (A) The ClosTron plasmid is conjugated in *C. difficile* (B) Expression from *thlA* promoter, does not confer erythromycin resistance since the *ermB* gene is interrupted by the group I intron. Group I introns are self-catalytically spliced from RNA transcripts but the process is orientation specific. The group I intron cannot splice out of the *ermB* transcript, as it is in the reverse orientation (C) Expression of the group II intron gives a transcript that binds to the LtrA protein leading to the formation of a ribonuclear protein complex (RNP). In this case, the group I intron is in the correct orientation and it splices out of the transcript. The antibiotic resistance is still not conferred, because the gene is in the reverse orientation (D) The RNP complex recognises and binds to specific sequences in the target gene, the LtrA protein nicks the target DNA and inserts the group II intron RNA. LtrA reverse transcribes the complementary DNA strand. Host nucleases degrade the RNA strand and DNA polymerase synthesizes the opposite strand. Host ligases seal the insertion site leading to completion of the process, in which a functional *ermB* gene is present in the target gene. The mutants are selected by resistance to erythromycin and intron insertion into the target chromosomal site is confirmed by PCR assay. The ClosTron plasmid is then lost due to segregational instability. The figure is adapted from (Cartman *et al.*, 2010).

1.3: Introduction to the development of a novel markerless deletion mutagenesis system for *C. difficile*

Aim of the study:

The primary aim of the study was to develop a novel genetic system for *C. difficile* to produce markerless deletions. The mutagenesis systems developed, for *C. difficile*, including the widely used ClosTron system, up to the start of this study, produce mutant strains by the insertional inactivation of genes of interest. An insertion of DNA into any given gene can interfere with the expression of downstream genes. In addition to polar effects, gene interruption by insertion of an antibiotic marker can prevent multiple manipulations of the genome, particularly in multidrug resistant strains such as *C. difficile* 630 (Sebaihia *et al.*, 2006). The ideal mutation for robust functional genomic studies is a markerless, in-frame deletion of the gene of interest.

Strategy for the development of a novel markerless deletion mutagenesis system for *C. difficile*:

As exemplified above, the construction of mutants via single cross-over can be unstable. However, a spontaneous second cross-over to generate a deletion can be a rare event, laborious to isolate and it may not occur if the diploid is stable. Posfai *et al.* (1999), exploited the homing endonuclease, I-SceI, as a tool for the promotion of allele exchange in *E. coli*, stimulated by a double strand break (DSB) and developed a method that permits the targeted construction of markerless deletions in this bacterium (Posfai *et al.*, 1999). The basic principle of the

method is that cleavage of the chromosome at an artificially introduced I-SceI site stimulates intramolecular recombination by a DSB (known to stimulate homologous recombination in many microorganisms) (Posfai *et al.*, 1999). Subsequently, this method was successfully adapted by various groups to generate deletion mutants of other bacteria including the anaerobe *Bacteroides fragilis* (Patrick *et al.*, 2009).

Homing endonucleases cleave double-stranded DNA with high specificity in the presence of divalent metal ions and recognise long DNA sequences (12-40 bp) (Belfort and Roberts, 1997). I-SceI is encoded by the group I intron of the large rRNA gene of *Saccharomyces cerevisiae* and it is a member of the largest class of homing enzymes, characterised by the presence of either one or two conserved amino acid residue sequence motifs (LAGLIDADG motifs) (Belfort and Roberts, 1997; Beylot and Spassky, 2001). The I-SceI protein acts as a monomeric endonuclease of 235 amino acids and cleavage of its 18 bp recognition sequence (TAGGGATAACAGGGTAAT) generates 4 bp overhangs with 3'-hydroxyl termini (Beylot and Spassky, 2001; Perrin *et al.*, 1993; Monteilhet *et al.*, 1990). This sequence is absent from all eukaryotic (except the source *S. cerevisiae*) and prokaryotic sequenced genomes which abolishes the risk of genome fractionation by the meganuclease and this makes the enzyme a useful tool for genetic engineering (Lopez *et al.*, 2009).

The present study has adapted the method of Posfai *et al.* (1999), to generate markerless deletions in *C. difficile*. An overview of the method is shown in **FIG 1.9**. The method uses two vectors, one I-SceI site delivery plasmid, containing the recognition site of I-SceI and one I-SceI expression plasmid. The allele replacement vector, made

from the insertion of homologous sequences flanking the gene to be deleted into the I-SceI site delivery vector, is introduced into *C. difficile*. Homologous recombination between either flanking sequence on the plasmid and its chromosomal homologue, results in plasmid integration into the chromosome. The I-SceI expression vector is then introduced into confirmed plasmid integrants. I-SceI meganuclease recognises its integrated site and introduces a DSB into the chromosome, stimulating diploid resolution via homologous recombination. Depending on the site of the cross-over, diploid resolution results in either a deletion mutant, or the wild type genotype.

The method was developed and first applied to generate a deletion of the *addBA* genes in *C. difficile*, which is described in chapter 3. Subsequently, the method was optimised for the deletion of the *pilA* and *pilT* genes and is described in chapter 4.

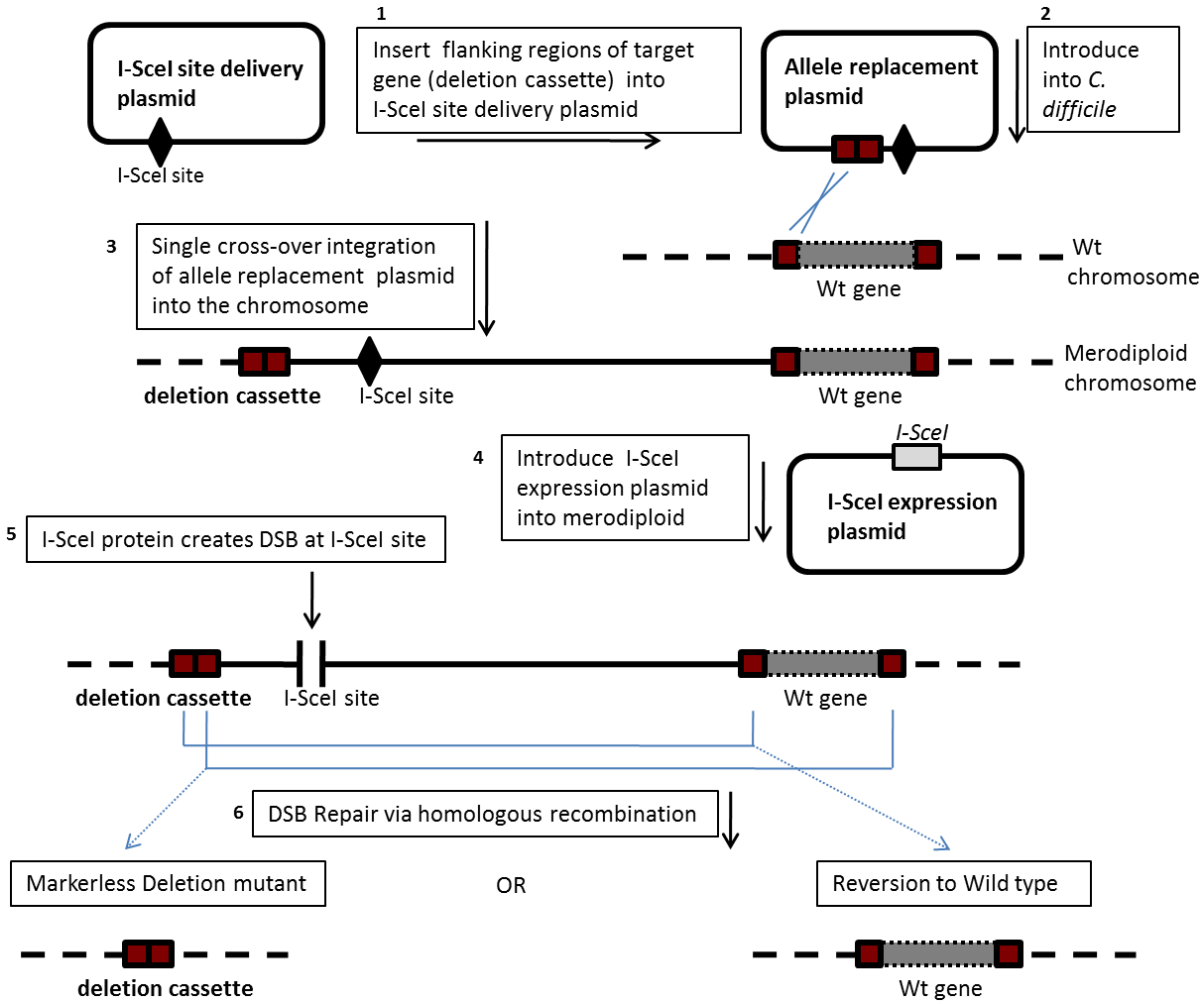


FIG 1.9 Strategy for generation of markerless deletion in *C. difficile*

(1)Sequences (~500 bp) flanking the gene to be deleted (red boxes) are inserted in the I-SceI site delivery vector, generating the allele replacement vector (2) The allele replacement vector is introduced into *C. difficile* (3) Homologous recombination between either flanking sequence on the allele replacement plasmid and its chromosomal homologue, results in plasmid integration into the chromosome (4) The I-SceI expression vector is then introduced into confirmed merodiploids (5) I-SceI meganuclease (expressed by an appropriate promoter) recognises its integrated site and introduces a DSB into the chromosome of the merodiploid (6) The DSB stimulates repair via homologous recombination. Depending on the site of the cross-over, diploid resolution results in either a deletion mutant, or wild type genotype.

CHAPTER 2: Materials and Methods

2.1: LIST OF STRAINS, PLASMIDS AND OLIGONUCLEOTIDES

List of strains:

Strain Name	Genotype	Reference
E. coli* S17-1λ*pir	*recA thi pro hsdR*$^{-}$ *hsdM*$^{+}$ RP4-2-Tc::Mu-Km::Tn7 λ*pir*	(Simon R., 1983)
***C. difficile* 630**	Epidemic type X, Ribotype 012, A^{+} B^{+}	(Sebaihia *et al.*, 2006)
C. difficile* 630Δ*erm	Erythromycin sensitive strain of *C. difficile* 630	(Hussain *et al.*, 2005)
***C. difficile* Δ*addBA* 24**	*C. difficile* 630Δ*erm* Δ*addBA*	This study
***C. difficile* Δ*addBA* 242**	*C. difficile* 630Δ*erm* Δ*addBA*	This study
***C. difficile* Δ*fliC* 88**	*C. difficile* 630Δ*erm* Δ*fliC*	This study
***C. difficile* Δ*fliC* 381**	*C. difficile* 630Δ*erm* Δ*fliC*	This study
***C. difficile* Δ*fliC* 383**	*C. difficile* 630Δ*erm* Δ*fliC*	This study
***C. difficile* Δ*fliC* 415**	*C. difficile* 630Δ*erm* Δ*fliC*	This study
***C. difficile* Δ*pilT* 271**	*C. difficile* 630Δ*erm* Δ*pilT*	This study
***C. difficile* Δ*pilT* 272**	*C. difficile* 630Δ*erm* Δ*pilT*	This study

***C. difficile* Δ*pilT* 288**	*C. difficile* 630Δ*erm* Δ*pilT*	This study
***C. difficile* Δ*pilT* 292**	*C. difficile* 630Δ*erm* Δ*pilT*	This study
***C. difficile* Δ*fliC* Δ*pilT* 234**	*C. difficile* 630Δ*erm* Δ*fliC* Δ*pilT*	This study
***C. difficile* Δ*fliC* Δ*pilT* 238**	*C. difficile* 630Δ*erm* Δ*fliC* Δ*pilT*	This study
***C. difficile* Δ*fliC* Δ*pilT* 250**	*C. difficile* 630Δ*erm* Δ*fliC* Δ*pilT*	This study

List of Plasmids:

Plasmid Name	Description	Reference
pJIR1456	*E. coli-C. perfringens* shuttle vector (pIP404 +ve replicon, *catP* marker), Tm^R/Cm^R	(Lyras and Rood, 1998)
pJIR2816	*E. coli-C. perfringens* shuttle vector (pIP404 +ve replicon, *catP* marker), Tm^R/Cm^R	(O'Connor *et al.*, 2006)
pES185	pJIR1456 + I-SceI recognition site (SacI site), Tm^R/Cm^R	This study
pES2761	pJIR2816 + I-SceI recognition site (PvuI site), Tm^R/Cm^R	This study
pMTL82254	*E. coli-Clostridium* shuttle vector (pBP1 +ve replicon, *ermB* marker), Erm^R	(Heap *et al.*, 2009)
pMTL83353	*E. coli-Clostridium* shuttle vector (pCB102 +ve replicon, *aad9* marker), Sp^R	(Heap *et al.*, 2009)
pGB920	pLYL01 with I-SceI under control of *B. fragilis fucR* promoter (*tetQ* marker), Tc^R	(Patrick *et al.*, 2009)

pES288	pMTL82254 + P*fdx*::*I-SceI* (SbfI site), Erm^R	This study
pES124	pMTL82254 + P*gdh*::*I-SceI* (FSpI site), Erm^R	This study
pES86	pMTL82254 + P*thl*::*I-SceI* (NotI/NdeI site), Erm^R	This study
pMTL84422	*E. coli-Clostridium* shuttle vector (pCD6 +ve replicon, *tetA* marker), Tc^R	(Heap *et al.*, 2009)
pES271	pES185 + *addBA* deletion cassette (SphI site), Tm^R/Cm^R	This study
pMTL83151	*E. coli-Clostridium* shuttle vector (pCB102 +ve replicon, *catP* marker), Tm^R/Cm^R	(Heap *et al.*, 2009)
pES242	pMTL83151 + I-SceI recognition site (SacI site), Tm^R/Cm^R	This study
pES2921	pES242 + *fliC* deletion cassette (FspI site), Tm^R/Cm^R	This study
pMTL84151	*E. coli-Clostridium* shuttle vector (pCD6 +ve replicon, *catP* marker), Tm^R/Cm^R	(Heap *et al.*, 2009)
pES196	pMTL84151 + *fliC* gene and native promoter (NotI/XhoI site), Tm^R/Cm^R	This study
pES2922	pES242 + *pilT* deletion cassette (FspI site), Tm^R/Cm^R	This study
pES2923	pES242 + *pilA* deletion cassette (FspI site), Tm^R/Cm^R	This study
pES2241	pMTL82254 + pilprom1 (NotI/NdeI site), Erm^R	This study
pES2242	pMTL82254 + pilprom2 (NotI/NdeI site), Erm^R	This study
pES2243	pMTL82254 + pilprom3 (NotI/NdeI site), Erm^R	This study
pES2244	pMTL82254 + pilprom4 (NotI/NdeI site), Erm^R	This study

pES2245	pMTL82254 + pilprom5 (NotI/NdeI site), Erm^R	This study
pES2246	pMTL82254 + pilprom6 (NotI/NdeI site), Erm^R	This study
pES2247	pMTL82254 + pilprom7 (NotI/NdeI site), Erm^R	This study

***** Tm^R -thiamphenicol resistance in *C. difficile*; Cm^R -chloramphenicol resistance in *E. coli*; Sp^R –spectinomycin resistance in *E. coli*; Tc^R –tetracycline resistance in *B. fragilis* and *E. coli*; Erm^R –erythromycin resistance in *C. difficile* and *E. coli*

List of Primers:

Primer Name	Sequence 5'- 3'
ISceI_F_SacI	CCCTCGAATTACCCTGTTATCCCTATCGAGGAGCTTGAGCT
ISceI_R_SacI	CAAGCTCCTCGATAGGGATAACAGGGTAATTCGAGGGAGCT
ISceI_F_PvuI	CGCCTCGAATTACCCTGTTATCCCTATCGAGGAGCTTCGAT
ISceI_R_PvuI	CGAAGCTCCTCGATAGGGATAACAGGGTAATTCGAGGCGAT
SbfI-PfdxFor	AAGTTTCCTGCAGGGTGTAGTAGCCTGTGAAATAAG
PfdxRev	TAACACACCTCCTTAAAAATTAC
PfdxIsceI	GTAATTTTTAAGGAGGTGTGTTAATGAAAAACATCAAAAAAAA CCAGG
SbfI-IsceIRev	AAGTTTCCTGCAGGTTATTTCAGGAAAGTTTCGGAG
SphI_addBA1	TTCCGCATGCTAAATGGGGATATAATACAGGC
addBA2	CCTAAGTCCCATAAATTTCCG
addBA2_addBA3	CGGAAATTTATGGGACTTAGGTGGAGTTGATGAAGCTGTTTG
SphI_addBA4	TTCCGCATGCTAGCAACCACAATATTTTCTCC
INTBAFor	GATGATGTTGGTAGTAGAGAG

INTBARev	GACCTAGTGCAGTATATTGAG
SphI-fliC1	TTCCGCATGCTTCAGCTTTAGAGTCTTTGTTG
fliC2	CTCCTTAGTATAGTTGACATCC
fliC3	GGATGTCAACTATACTAAGGAGAAAAGAAAGGATAAGGCTTTGC
SphI-fliC4	TTCCGCATGCTGGTTGTTCATGAACTTTCCC
INTfliC-For	TTGAATATTCTGGAAGTTCAGC
INTfliC-Rev	CAAGGTTTGTTATCAAGTATCC
cmFliCFor	CCCTGGCGGCCGCAACTTTATGATAGTATGGAGC
cmFliCRev	CCCTGCTCGAGCTATCCTAATAATTGTAAAACTC
Trans-1	TATCAGGAAACAGCTATGACC
Trans-2	AAAATACTCTTTTCTGTTCCAAC
Trans-3	AGTGCCAAGCTTGCATGTCTG
SphI-CD3505-1	ATATGCATGCAATCCAGACACTTCCTCCAC
CD3505-2	ATGATAGGTGGAGTAAACTAC
CD3505-3	GTAGTTTACTCCACCTATCATCATAATATTTACTCTCCATATCTTC
SphI-CD3505-4	ATATGCATGCGGCTTTGATAACAACTTTTGAC
INT3505For	TACACAGATTCCATAGACTGC
INT3505Rev	ATTTTACCAAATCAAACTCAAGG
SphI-CD3507-1	TCTCGCATGCTTCTAAGATAGAATCTCTATGTG
CD3507-2	GTGAGATTTTGAGGAAGTGG
CD3507-3	CCACTTCCTCAAAATCTCACATTCTTATTACCCTATTCTTGAC

SphI-CD3507-4	TCTCGCATGCTGTTCAATGGTATAACATCAGC
INT3507For	ATCTAATGCCTTATAAGTATGTG
INT3507Rev	TACAGATAATCCATCTTCATGG
NdeI-pilprom	TCTCTCATATGTTTTATTCCCCCTTAAATTTTTTAATTAATAC
NotI-pilprom1	TATCTGCGGCCGCTAGAAATAAATTTGGTTAGTACG
NotI-pilprom2	TATCTGCGGCCGCGATTTATGTTCTGTAATGTGGG
NotI-pilprom3	TATCTGCGGCCGCTAAACAGTGTTGCTAAAATATGG
NotI-pilprom4	TATCTGCGGCCGCATTAGCAAAGGATGGAAATGAC
NotI-pilprom5	TATCTGCGGCCGCTCTTTAGTGTTTGTGAGAGGG
NotI-pilprom6	TATCTGCGGCCGCAGAAAGAGGTGTTTCCAATGG
NotI-pilprom7	TATCTGCGGCCGCGAGTTGTTTAAAAGGATATAAACC

*****Underlined sequences indicate restriction enzyme sites. All primers were supplied by Eurofins MWG Operon, Germany. Primers were resuspended in 1 x TE at 100 pmol/µl. These stocks were then diluted in dH_2O to give a working stock of 10 pmol/µl.

2.2: GROWTH MEDIA

LB (L-broth): Difco Bacto-Tryptone (10 g/l), Difco Bacto-yeast extract (5 g/l), NaCl (5 g/l), dH_2O (1 l), pH adjusted to 7.2 with NaOH.

BHI (Brain Heart Infusion): Difco Bacto Brain Heart Infusion (37 g/l), yeast extract (Oxoid) (5 g/l), distilled H_2O (1 l).

AIM (Anaerobic Identification Medium): Proteose peptone (Oxoid) (20 g/l), yeast extract (Oxoid) (5 g/l), trypticase peptone (BBL) (5 g/l), NaCl (5 g/l), Na_2CO_3 (0.4 g/l), distilled H_2O (1 l).

BHIS or AIMS: BHI or AIM containing L-cysteine and Haemin.

L-cysteine: Stock solution was made at 10% w/v in dH2O and filter sterilised using a 0.22 µm filter. 1 ml of stock was added per 100 ml of medium.

Haemin: Stock solution was made with 50 mg Haemin, 1.74 g K_2HPO_4, 0.4 g NaOH and 100 ml dH_2O and then was autoclaved. 1 ml of stock was added per 100 ml of medium.

Blood agar: Columbia blood agar base (39 g/l), 5% defibrinated horse blood.

Agar medium concentration:

Unless otherwise stated, the agar concentration used for solid media was 1.5% w/v.

Antibiotics:

Antibiotics were supplied from Sigma and were prepared and stored according to the manufacturer's instructions. Antibiotics were added at the following final concentrations: erythromycin 500 µg/ml for *E. coli* and 10 µg/ml for *C. difficile*; chloraphenicol 30 µg/ml for *E. coli*; thiaphenicol 15 µg/ml for *C. difficile*; spectinomycin 250 µg/ml for *E. coli*; cefoxitin sodium salt 8 µg/ml; D-cycloserine 250 µg/ml.

2.3: BACTERIAL METHODS

Growth conditions:

E. coli strains were cultured aerobically using LB as appropriate, at 37˚C, with liquid cultures shaking at 160-200 rpm. *C. difficile* strains were grown in a Don Whitley Scientific (UK) MiniMacs anaerobic work station at 37˚C, with an anaerobic gas mix of 10% hydrogen, 10% carbon dioxide and 80% nitrogen. BHI, AIM (broth or agar) or Blood agar medium (supplemented when appropriate) were used to culture *C. difficile*. Unless otherwise stated, all media for *C. difficile* anaerobic use were reduced overnight.

Growth measurement of liquid culture:

The growth of liquid cultures was monitored by measuring the optical density at 600 nm (OD_{600}) with a Hitachi U-2000 spectrophotometer (Hitachi Instruments, Japan). Since spectrophotometer readings above an OD_{600} of 1 are not accurate, the sample was diluted in appropriate medium in the cuvette and mixed before a reading was taken. The reading was then multiplied by the dilution factor, to give the optical density.

Growth curves:

A single colony of *C. difficile* was used to inoculate 10 ml AIM broth and incubated anaerobically overnight. Fresh AIM broth was then inoculated with the overnight culture at 1/100 dilution and OD_{600} was measured at 1-h time intervals for 10-h. Growth curves were performed in triplicate.

Preparation of competent cells:

A single colony of the *E. coli* strain of interest was used to inoculate 5 ml LB. Pre-warmed LB (250 ml) containing 20 mM $MgSO_4$ was inoculated with the overnight culture at 1/100 dilution and was incubated at 37°C shaking, until an OD_{600} of ~0.4 was reached. The culture was centrifuged at 4,500 x g in a Sorvall RC-5B centrifuge, with GSA rotor (Sorvall Ltd, UK) for 5 min at 4°C. The supernatant was discarded and the pellet was resuspended in 100 ml ice-cold TFB1, followed by a 5 min incubation on ice. The cells were pelleted by centrifugation, as described above, the supernatant was discarded and the pellet was resuspended in 10 ml ice-cold TFB2. Cells were incubated for 1-h on ice.

Aliquots (100-200 µl/tube) of the competent cells were snap-frozen in dry ice containing isopropanol and stored at -80°C.

TFB1: 30 mM CH_3COOK, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15% v/v glycerol, pH adjusted to 5.8 with 1 M acetic acid and filter sterilised with a 0.22 µm filter.

TFB2: 10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% v/v glycerol, pH adjusted to 6.5 with 1 M potassium hydroxide and filter sterilised with a 0.22 µm filter.

Heat-shock transformation of competent cells:

Competent cells were thawed on ice. An appropriate amount of plasmid DNA was added to 20-100 µl of competent cells, followed by 30 min incubation on ice. The cells were heat-shocked at 37°C for 2 min, followed by 5 min incubation on ice. Pre-warmed LB was added to the cells, followed by incubation at 37°C for 1-h. The culture was centrifuged at 13000 rpm for 1 min, the supernatant was discarded and the pelleted cells were then resuspendend in appropriate volume of LB (usually 100 µl). The cells were then spread on media with appropriate antibiotics and incubated overnight at 37°C. The usual strain used for transformation was S17-1*λpir*. Occasionally, competent cells were purchased from various companies including, NEB 5-alpha competent *E. coli*, Invitrogen DH5a *E. coli* competent cells or Clontech Stellar competent cells that were transformed using the protocol described or according to the manufacturer's instructions, as appropriate.

Transfer of plasmid DNA into *C. difficile*:

Plasmids were introduced into *C. difficile* by conjugation from *E. coli* S17-1λ*pir*. Conjugations were performed using cultures of donor and recipient from either stationary or exponential phase of growth. A single colony of *E. coli*, containing the appropriate plasmid, was inoculated into 5 ml of LB with appropriate antibiotic and incubated aerobically overnight. A single colony of the *C. difficile* recipient strain was inoculated into 10 ml of appropriate broth and incubated overnight. For stationary mating, 1 ml of overnight *E. coli* culture was centrifuged at 5000 rpm for 2 min and then the cell pellet was washed twice in 1 ml of 1 x PBS (centrifuging as before), to remove traces of antibiotics. The pellet of donor cells was then transferred to the anaerobic workstation and resuspended in 150-200 µl of *C. difficile* overnight culture. For exponential mating, overnight culture of *E. coli*, prepared as described, was used to inoculate fresh LB without antibiotic and grown until an OD_{600} of 0.2-0.4. Then, 1 ml of *E. coli* culture was centrifuged at 5000 rpm for 1 min, followed by transfer of the cell pellet in the anaerobic workstation. The pellet of donor cells was resuspended in 150-200 µl of *C. difficile* culture, also grown to an OD_{600} of 0.2-0.4. Occasionally, different donor-recipient ratio was used, to optimize for a particular mating. The donor-recipient suspension (from stationary or exponential cultures) was spotted onto agar medium without antibiotic, as appropriate for the recipient, and incubated anaerobically at 37˚C for 16 to 24-h. The conjugation mixture was then harvested in 2 ml of 1 x PBS (reduced overnight) and spread on growth medium suitable for the recipient, containing appropriate antibiotic to select for the plasmid-encoded resistance and also containing cycloserine (250 µg/ml) and cefoxitin (8

µg/ml)to counter-select against the *E. coli* donor. Plates were incubated anaerobically at 37°C for 24 to 120-h, before colonies were picked and purified by re-streaking. Putative transconjugants were replica patched on LB and incubated aerobically to exclude *E. coli* contamination. Since background growth would appear over time, donor-only and recipient-only mock mating controls were used, to help judge when background growth should be expected.

Sporulation measurement:

A starter culture was prepared in AIM broth using a 1% inoculum of *C. difficile* overnight culture (grown from a single colony in AIM broth) and incubated until the optical density at OD_{600} was approximately 0.3. This culture was then diluted at 1/100 into fresh AIM broth and this was considered the start of sporulation T0. The sporulation culture was incubated for five days, where the optical density OD_{600}, heat-resistant colony formation and untreated colony formation, were measured every 24-h. To measure heat resistant CFU, 0.5 ml of sample was transferred out of the anaerobic workstation, heated at 60°C for 25 min, 10-fold serially diluted in 1 x PBS and 5 µl of each dilution spotted (4-6 times)on AIM agar containing 0.1% sodium taurocholate. The spots were left to dry and plates were returned to the anaerobic chamber. To measure untreated viable CFU, 0.5 ml of sample was serially diluted in 1 x PBS (reduced overnight) and 5 µl of each dilution was spotted (4-6 times) on AIM containing 0.1% sodium taurocholate. This was performed in the anaerobic workstation. Plates were incubated for 48-h before CFU were enumerated. All experiments were performed in triplicate.

Total toxin production:

Total toxin (A + B) was measured using the *C. difficile* TOX A/B II ELISA kit (Techlab), according to the manufacturer's instructions. AIM broth was inoculated with overnight culture of *C. difficile* (grown from a single colony in AIM broth) at 1/100 and was incubated for 24-h. Culture supernatants were collected by centrifugation at 13200 rpm for 2 min, when optical density OD_{600} was approximately 0.5, 0.8 and 1 and stored at -20^{o}C until required. Total toxin levels were determined by measuring A_{450}/OD_{600}.

Culture supernatants were given to Dr. Allison Wroe (MPRL) to measure toxin production.

Metronidazole sensitivity assay:

AIM broth was inoculated with overnight culture of *C. difficile* (grown in AIM from a single colony) at 1/100 dilution and was incubated until an OD_{600} of ~0.3 was reached. The culture was then diluted at 1/100 into fresh AIM broth and incubated until an OD_{600} of ~0.3 was reached. Then, 10-fold serial dilutions of this culture were prepared in 1 x PBS (reduced overnight) and 5 µl of each dilution was spotted on AIM without metronidazole or on AIM with metronidazole concentration of 0.25, 0.125 and 0.06 µg/ml. Plates were incubated for 48-h before the result was recorded. All experiments were performed in triplicate.

UV sensitivity assay:

AIM broth was inoculated with overnight culture of *C. difficile* (grown in AIM broth from a single colony) at 1/100 dilution and was incubated until an OD_{600} of ~0.3 was

reached. The culture was then diluted at 1/100 into fresh AIM broth and incubated until an OD_{600} of ~0.3 was reached. Then, 10-fold serial dilutions of this culture were prepared in 1 x PBS (reduced overnight) and 5 µl of each dilution was spotted on AIM. Four sets of AIM agar were spotted. All plates were briefly transferred out of the anaerobic workstation. Plates were exposed to either just atmospheric O_2 i.e. not UV irradiated, or exposed to UV irradiation of 10, 15 or 20 J/m^2, using the Stratalinker UV 2400 (Stratagene). All plates were returned to the anaerobic chamber and incubated for 48-h before results were recorded. All experiments were performed in triplicate.

SOS induction and cell morphology visualization:

AIM broth was inoculated with overnight culture of *C. difficile* (grown in AIM broth from a single colony) at 1/100 dilution and was incubated until an OD_{600} of ~0.3 was reached. This was then inoculated into fresh AIM broth at 1/100 dilution, containing metronidazole at a concentration of 0.25, 0.125 or 0.06 µg/ml and into fresh AIM broth without metronidazole. The culture was incubated for 24-h. At 0-h (T0, inoculation time), 4-h (T4), 6-h (T6) and 24-h (T24), the optical density at OD_{600} was measured and 1 ml sample was centrifuged at 5000 rpm for 1 min, followed by resuspension of the cell pellet in 100 µl of cell fixative. All samples were stored at 4 $^{\circ}C$ until they were required for microscopy. All experiments were performed in triplicate. The morphology of cells was visualised by phase-contrast microscopy using a 10 µl wet mount, prepared from a fixed sample. To visualise the septum, the membrane stain FM4-64 (Molecular Probes) was added to 1 ml of culture at

different time points and incubated in the anaerobic workstation for 20 min. The stained sample was centrifuged at 5000 rpm for 1 min, followed by resuspension of the cell pellet in 100 µl of cell fixative and stored at 4 °C in the dark until required. The FM4-64-stained sample (8–10 µl) was applied onto a poly-L-lysine microscope slide, a coverslip was placed on top and the coverslip edges were sealed with varnish. The stained cells were viewed by fluorescence microscopy.

The Leitz Wetzlar Metalux II microscope (Leitz Wetzlar now Leica-microsystems, Germany) was used and images were captured with a Hamamatsu digital c4742-95 camera (Hamamatsu Photonics K.K., Japan) and Improvision Openlab software (PerkinElmer Inc., USA).

Cell fixative: 20% v/v formaldehyde solution (Sigma) in bacterial buffer.

Preparation of poly-L-lysine microscope slide: A drop (15 µl) of poly-L-lysine (Sigma) was spread thinly on an ethanol-cleaned and dried microscope slide. Poly-L-lysine was left to dry before samples were applied.

Flagella-mediated motility stab assay:

A single colony of *C. difficile* was stab inoculated into motility agar tube containing appropriate medium with 0.175% agar. The swim agar tube was incubated anaerobically at 37°C for 24-h.

Transmission electron microscopy:

A 48-h grown colony of *C. difficile* from an appropriate medium was resuspended in 10 µl reduced 1 x PBS, in the anaerobic workstation. The suspension was transferred out of the anaerobic chamber and 100 µl of 3% glutaraldehyde made in 1 x PBS was added, followed by 30 min incubation at room temperature (RT), to fix the cells. Then, a drop of the fixed sample was placed onto carbon-coated grid, incubated at RT for 5 min and excess liquid was blotted away with filter paper. The cells were stained with 1% phosphotungstic acid for 1 min and excess liquid was blotted away with filter paper and then each grid was air dried. The grids were examined using a Philips CM120 Biotwin transmission electron microscope.

Surface twitching motility assay:

BHI broth was inoculated with overnight culture of *C. difficile* (grown in BHI broth from a single colony) at 1/100 dilution and was incubated until an OD_{600} of ~0.3 was reached. Next 1 ml of culture was concentrated two-fold by centrifugation at 5000 rpm for 1 min and resuspension of pellet in 0.5 ml 1 x PBS (reduced overnight). Then, 5 µl of concentrated culture was spotted onto appropriate, well-dried medium containing 0.7 % agar. Plates were incubated for 5 days anaerobically, with the growth medium facing upwards and in sealed bags to maintain humid conditions. The morphology and spread of the colony (whole diameter and translocation of inoculation spot) were recorded.

Subsurface twitching motility assay:

Appropriate growth medium containing 1% agar was poured very thinly (~3 mm) into a plastic petri dish and was dried and reduced overnight before inoculation. To inoculate the interface between the agar and the petri dish, a single 24-hr grown colony of *C. difficile*, was picked with a toothpick and was stabbed through the agar to the plastic dish at a single point, tapping lightly to ensure inoculation. The agar medium was incubated for 5 days, with the growth medium facing upwards. Plates were then transferred out of the anaerobic workstation, the agar was removed, leaving behind bacterial growth attached on the plastic dish, which was subsequently dried under a lamp. The dried bacteria were stained with Coomasie blue by flooding the dish with the stain for 1 min. The excess stain was removed with destaining solution by several washes, until clear. The staining solution contained 0.5% Coomassie Brilliant blue R250, 40% methanol and 10% acetic acid in water. The destaining solution contained 40% methanol and 10% acetic acid in water. The stained bacteria were then dried at 37°C and the diameter of colony expansion was measured.

2.4: DNA METHODS

Chromosomal DNA extraction:

C. difficile genomic DNA from liquid culture was extracted using the DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. Alternatively, a *C.*

difficile pellet from liquid culture or a loop-full of colonies were vortexed in a 100 µl of 5% suspension of Chelex 100 Molecular biology grade resin (Bio-Rad), prepared in sterile water. An incubation of 10-15 min at 100°C followed, to destroy the cell membranes and proteins and denature genomic DNA. The suspension was then centrifuged at 13200 rpm for 10 min and the supernatant containing genomic DNA was removed and stored at -20°C.

Plasmid DNA extraction:

Plasmid DNA was prepared using the QIAprep Spin Miniprep kit (Qiagen) according to manufacturer's instructions. An additional step was included for *C. difficile*. The *C. difficile* cell pellet (usually from 5-10 ml culture), was resuspended in 250 µl Buffer P1 containing lysozyme (Sigma) at a final concentration of 1 mg/ml and incubated at 37°C for 30-60 min before addition of the Buffer P2. Then the protocol continued as specified by the manufacturer. If a QIAfilter Midi kit (Qiagen) was used, when a large volume of plasmid was required from *E. coli*, a 100 ml of overnight culture was used and DNA was eluted in 1 ml of 1 x TE buffer. Isolated plasmid DNA was stored in -20°C.

1 x TE buffer: 10 mM Tris-HCl, 1 mM EDTA pH8

Quantification of DNA:

DNA concentration was quantified using a nanodrop ND-1000 or ND-2000 spectrophotometer (Thermo Scientific). The 260/280 (nm) and 260/230 (nm) absorbance ratio was recorded, to assess purity of DNA.

Annealing of complementary pairs of oligonucleotides:

Oligonucleotides were mixed in equal molar amounts in an Eppendorf tube and used the Eppendorf Mastercycler gradient cycler programmed to heat at 98°C for 5 min, followed by gradual decrease in temperature until 20°C was reached and incubated for 5 min. Alternatively, the Eppendorf tube containing the mixture of oligos was placed in a beaker with boiling water and left to cool until it reached room temperature. Annealed oligos were then placed on ice and stored at -20°C.

Polymerase Chain Reactions:

PCR reactions were performed in an Eppendorf Mastercycler gradient cycler and using Phusion High-Fidelity DNA polymerase (New England Biolabs). A typical PCR reaction contained 10 µl of 5x Phusion HF or GC buffer, 1 µl of 10 mM dNTPs mix, 1 µl of 10 pmol/µl forward primer, 1 µl of 10 pmol/µl reverse primer, appropriate amount of template DNA as required, 0.5 µl Phusion DNA polymerase and dH_2O to give a final volume of 50 µl. Initial denaturation was performed at 98°C for 30 sec. This was followed by 31 cycles of denaturation at 98°C for 10 sec, annealing at primer specific temperature for 30 sec, and extension at 72°C for the time appropriate for the template size. Final extension was at 72°C for 10 min. PCR products were purified using the QIAquick PCR purification kit (Qiagen), according to manufacturer's instructions.

Agarose gel electrophoresis:

DNA was separated by electrophoresis at 80 V for 60-80 min through 1% agarose gel (occasionally 0.8% for large-sized DNA) in 1 x TAE, using Bio-Rad electrophoresis tanks (small mini/wide mini subcell, subcell), powered by a BIO-RAD Power-Pak 300. DNA samples were loaded onto the gel, mixed with 6x blue gel loading dye (NEB) to final concentration of 1x. The DNA size was determined by loading 1 kb ladder or 100 bp ladder (NEB) mixed with loading dye as described. Following electrophoresis, the gel was stained in ethidium bromide (1 µg/ml) for 20 minutes and then destained in dH_2O for 10-20 minutes. The DNA was visualised under UV light.

50 x TAE Buffer: 2 M Tris, 0.95 M glacial acetic acid, 50 mM EDTA pH8

DNA extraction from agarose gel:

Agarose gel containing DNA of interest was excised under UV light, and DNA was extracted using the QIAquick Gel Extraction Kit (Qiagen) or the Wizard SV Gel and PCR clean-up system (Promega), according to manufacturer's instructions.

Digestion of DNA with restriction endonuclease:

DNA was cleaved using appropriate restriction endonucleases supplied by NEB, according to the manufacturer's instructions. The volume of enzyme was kept below 10% of the total reaction volume, to reduce star activity. Incubation times varied from 1 to 24-h. The enzymes were inactivated as necessary, according to the manufacturer's instructions. When appropriate, digested DNA was purified

using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions.

Dephosphorylation of DNA:

To remove 5′ phosphate groups from linearized plasmid DNA fragments, to prevent self-ligation, NEB Alkaline Phosphatase (CIP) or NEB Antarctic Phosphatase was used, according to the manufacturer's instructions. Antarctic Phosphatase was inactivated as instructed by the manufacturer. Before ligation, linear and dephosphorylated plasmid DNA was purified using the QIAquick PCR purification kit (Qiagen), according to the manufacturer's instructions.

Phosphorylation of DNA:

DNA products that required addition of 5'-phosphates, prior to ligation, the enzyme T4 Polynucleotide Kinase (NEB) was used, according to the manufacturer's instructions. The reaction was incubated at 37°C for 30 min, followed by heat inactivation of the enzyme at 65°C for 20 min. Phosphorylated DNA products were purified using the QIAquick PCR purification kit (Qiagen), according to the manufacturer's instructions.

Ethanol precipitation of DNA:

To concentrate DNA by ethanol precipitation, 1/10 volume of 3 M sodium acetate (pH 5.2) was added to the DNA solution, followed by addition of 2x volume of chilled 100% ethanol. The mixture was incubated either overnight at −20°C or at

-80°C for 1 to 2-h. The DNA was pelleted by centrifugation at 15000 rpm for 30 min at 4 °C, the supernatant was discarded and the pellet was resuspended in 0.5 ml 70% ethanol. The mix was centrifuged at 15000 rpm for 20 min at 4 °C, the supernatant was discarded, and the DNA pellet was dried by centrifuging at the highest drying rate in a Savant DNA Speed Vac DNA110. The dried DNA pellet was resuspended in an appropriate volume of dH_2O or Elution Buffer (Qiagen) and stored at -20°C.

Ligation of DNA:

DNA fragments were purified and quantified, as previously described, before ligation. The T4 DNA ligase (Promega) was used for ligation, according to the manufacturer's instructions. Depending on the ligation, various vector:insert molar ratios were used. Typically, a 10 µl reaction contained 3 Weiss units of T4 DNA ligase. A control ligation without insert was also set up to assess any re-ligation of the vector. Incubation time and temperature varied, as appropriate for each ligation. Incubation time ranged from 1 to 16-h and temperature from 8-22°C.

Single colony gel:

To determine whether cells contained the construct of interest, following transformation with the ligated mixture, single colony gels were run. Single colonies from the transformation plates were patched on fresh plates with appropriate selection and incubated at 37°C overnight. A toothpick was used to transfer and resuspend cells from a patched out colony into 150-200 µl of SCFS buffer. The

toothpick was left to stand in the buffer for 10-15 min at room temperature, it was then removed and the suspension was centrifuged at 13200 rpm for 30 min. 30 µl of the supernatant containing plasmid DNA was run on a 1% gel at 80 V for 80 min. Supercoiled plasmid DNA was used as a control. The DNA was stained and visualised as described previously.

SCFS Buffer: 2.5% v/v Ficoll, 1.25% w/v SDS, 0.015% w/v Bromophenol Blue, 10 µg/ml RNase A, made to volume in 1 x TAE buffer

DNA sequencing:

DNA was sequenced by the GenePool facility, at the University of Edinburgh, from appropriately prepared samples, as specified by their instructions.

CHAPTER 3: The development of a novel markerless deletion system in *C. difficile* and the deletion of *addBA* genes

3.1: INTRODUCTION

Aim and Objectives:

- To construct a markerless genetic system for *C. difficile* using I-SceI.
  To identify and use tools for the construction of the system including working strain, vectors for I-SceI site and allele delivery into the *C. difficile* chromosome and promoters and vectors for I-SceI expression.
- To validate the system via the targeted deletion of *addBA* genes using the I-SceI system.
- To analyse the function of AddAB in *C. difficile*.

Background:

It is essential for all organisms to ensure the integrity of genomic DNA and the accurate transmission of their genetic material to progeny. Bacterial genomic DNA is constantly faced with damage from endogenous and exogenous sources. It has been suggested that under normal aerobic conditions *E. coli* can suffer 3000 – 5000 DNA lesions per cell per generation, the majority being due to oxidative damage (Cox, 2001). One type of damage that is considered lethal, if not repaired, is a double-strand break (DSB). It has been estimated that in *E. coli* a spontaneous double-

strand break occurs once per chromosome replication cycle (Reuter *et al.*, 2010).

Double-strand breaks (DSBs) can be caused directly or indirectly from various endogenous and exogenous sources including, ionizing radiation (e.g. X rays and γ rays), UV light, DNA damaging chemicals, desiccation, mechanical stress, oxygen radicals and inappropriate cleavage of DNA by host restriction-modification systems (Dillingham and Kowalczykowski, 2008; Kuzminov, 1999). Ionizing radiation, for example, can directly break DNA with the production of two free DNA ends (Kuzminov, 1999; Dillingham and Kowalczykowski, 2008; Kowalczykowski, 2000).

In the absence of exogenous damaging agents, during normal cell growth, a DSB can result from an encounter of the replication fork with discontinuities or blockages in the DNA template (Kuzminov, 1999; Dillingham and Kowalczykowski, 2008; Kreuzer, 2005; Michel *et al.*, 1997; McGlynn and Lloyd, 2002; Cox, 2001; Michel *et al.*, 2007; Kuzminov, 2001; Rothstein *et al.*, 2000). These include gaps, secondary structures, tightly bound proteins, nicks which for example can occur from free radicals arising from intermediary metabolism or when a SbcCD endonuclease cuts a hairpin structure at palindromic DNA sequences, etc. If a replication fork encounters an unrepaired nick in the template strand which disturbs the continuity of the replication fork, this can result in replication fork collapse and generation of an intact chromosome and a detached double-strand end (Kuzminov, 1999; McGlynn and Lloyd, 2002). Blockage of replication due to a DNA lesion in the lagging strand can be bypassed if the replication complex stays intact, by restarting lagging strand synthesis with a newly primed Okazaki fragment downstream of the lesion. This leaves a gap which can eventually lead

to a DSB if left unrepaired prior to the arrival of the replisome in the next round of duplication (McGlynn and Lloyd, 2002; Kuzminov, 1999). Moreover, replication fork stalling due to a defective replisome, DNA damage, blockage due to secondary structures or tightly bound proteins can lead to replication fork reversal, where the ends of the newly synthesized leading and lagging strands anneal. This can lead to different outcomes including DSBs at stalled replication forks or resetting of the replication fork and reassembly of the replisome (aided by the RecBCD enzyme complex discussed below) (Dillingham and Kowalczykowski, 2008; De Septenville *et al.*, 2012).

Bacteria have evolved two primary mechanisms for the repair of DSBs: non-homologous end joining (NHEJ) and homologous recombination (HR) (Lenhart *et al.*, 2012; Ayora *et al.*, 2011; Bowater and Doherty, 2006; Shuman and Glickman, 2007; Kowalczykowski, 2000). In NHEJ the broken ends, after minor processing, are joined together by ligation using minimal or no sequence homology. Therefore, this is a low fidelity DSB repair pathway accompanied by nucleotide insertions or deletions (Lenhart *et al.*, 2012). In addition, NHEJ only exists in a subset of bacteria (Della *et al.*, 2004). In contrast, homologous recombination is a high fidelity DNA repair pathway that requires a homologous sequence to repair the break. It is highly conserved among viruses, bacteria, archaea and eukaryotes (Cromie, 2009; Lenhart *et al.*, 2012). NHEJ is used when HR is impaired or an intact homologous template is not available, for example during prolonged stationary phase in specialised cell stages such as sporulation (Yeeles and Dillingham, 2010; Bowater and Doherty, 2006; Shuman and Glickman, 2007).

HR is recognised to have three stages and the RecA protein is central in this (Yeeles and Dillingham, 2010; Cromie,

2009; Kowalczykowski, 2000). In the first stage, presynapsis, the DNA ends are processed to form suitable single-stranded 3’ overhangs to which RecA binds and forms a functional RecA-ssDNA complex. In the second stage, synapsis, the RecA nucleoprotein filament performs a homology search, pairs with its homologous DNA followed by strand exchange resulting in a joint molecule called a Holliday junction. The third stage, postsynapsis, involves resolution of the junction to form recombinant repaired DNA products.

In bacteria there are three distinct classes of helicase-nuclease enzyme complexes namely, RecBCD, AddAB (also called RexAB) and AdnAB that process DSBs to form a 3’-ssDNA overhang for the production of a RecA nucleoprotein filament (Rocha *et al.*, 2005; Cromie, 2009). The AddAB and RecBCD proteins are widely distributed across the bacterial kingdom, with a single RecBCD- or AddAB-type enzyme complex being present in the majority of bacterial species; although several members of the Actinobacteria contain both enzyme types (Cromie, 2009; Yeeles and Dillingham, 2010). Structural homologues of these enzyme complexes have not been found in eukaryotes (Yeeles *et al.*, 2009). The majority of the genes encoding the enzymes are generally clustered in an operon (Cromie, 2009). Recently, the AdnAB type enzyme has been identified in Mycobacteria species (Sinha *et al.*, 2009; Gupta *et al.*, 2011). The *E. coli* RecBCD and the *B. subtilis* AddAB have been the paradigms for these enzymes.

The RecBCD and AddAB enzyme complexes are structurally distinct but they are functionally analogous, supported by the observation that the heterologous expression of *Bacillus subtilis* AddAB or *Lactobacillus lactis* RexAB can

nearly fully compensate for the DNA repair defect in Δ*recBCD E. coli* strain (Yeeles and Dillingham, 2010; Dillingham and Kowalczykowski, 2008). Both complex types are regulated by a recombination hotspot sequence Chi (crossover hotspot instigator) (Yeeles and Dillingham, 2010; Chedin and Kowalczykowski, 2002; Wigley, 2013). RecBCD and AddAB essentially follow the same steps for the production of the RecA nucleoprotein filament (Wigley, 2013). The enzymes bind to blunt or nearly blunt duplex DNA ends, the DNA is unwound and digested at the same time until a Chi site is recognised which this changes the activity of the enzyme complex to produce a 3'-ssDNA overhang to which RecA binds with the aid of the enzymes. The length and sequence of Chi recognised by different enzyme complexes vary between different bacterial species but are conserved between different strains of the same species. For example, *E. coli* RecBCD recognises the octameric sequence 5'-GCTGGTGG-3', *B. subtilis* AddAB recognises the pentameric sequence 5'-AGCGG-3', *L. lactis* AddAB the heptameric sequence 5'-GCGCGTG-3', *Staphylococcus aureus* AddAB the heptameric sequence 5'-GAAGCGG-3' and *Haemophilus influenzae* RecBCD the octameric sequence 5'-GNTGGWGG-3' (Wigley, 2013; Yeeles and Dillingham, 2010; Chedin and Kowalczykowski, 2002). In addition Chi sequences are over-represented in the genome (e.g. in *E. coli* one site per 5 kb) and most are skewed towards the origin of replication to promote recombination of collapsed replication forks (Yeeles and Dillingham, 2010; Smith, 2012; Halpern *et al.*, 2007).

RecBCD is a 330 kDa heterotrimeric complex composed of the RecB, RecC and RecD proteins (Yeeles and Dillingham, 2010; Singleton *et al.*, 2004). The complex contains two helicase motors, RecB and RecD with opposite polarities that are

powered by ATP hydrolysis (Dillingham and Kowalczykowski, 2008; Lenhart *et al.*, 2012; Taylor and Smith, 2003; Dillingham *et al.*, 2003). The C-terminus of RecB also contains a nuclease domain that is responsible for cleavages made on both DNA strands (Yeeles and Dillingham, 2010; Wang *et al.*, 2000b). The N-terminus of RecB contains a Superfamily 1A (SF1A) helicase with 3′-5′directionality, while the C-terminus is a λ-family nuclease (Wigley, 2013; Singleton *et al.*, 2007). RecD is a SF1B helicase with 5′-3′ specificity (Wigley, 2013; Singleton *et al.*, 2007). The RecC subunit is responsible for recognizing Chi (Wigley, 2013; Singleton *et al.*, 2004). RecBCD binds to the duplex DNA end (**FIG 3.1**) and the complex travels along the DNA using energy from the ATP hydrolysis in the RecB and RecD helicase motors. RecB travels in 3′-5′ direction and RecD in 5′-3′ direction, with RecD being the faster motor and consequently a loop of ssDNA is formed ahead of the RecB motor. As the DNA is being unwound, the nuclease domain of RecB degrades both strands with the 3′ strand being more vigorously hydrolysed than the 5′ strand (Wigley, 2013; Smith, 2012; Dillingham and Kowalczykowski, 2008; Dixon and Kowalczykowski, 1993; Taylor and Smith, 2003; Dillingham *et al.*, 2003). During unwinding, the 3′ stand may pass through a tunnel in RecC where it is scanned for the Chi sequence (Smith, 2012; Singleton *et al.*, 2004). Upon recognition of the Chi sequence, the RecC subunit binds tightly to the Chi site and the complex pauses translocation and resumes at a reduced rate due to the complex being led by the slower RecB motor instead of the RecD motor. Moreover, following Chi recognition, cleavage of the 3′-5′ strand is attenuated whilst cleavage of the 5′-3′ strand is enhanced, resulting in the formation of a 3′ ssDNA overhang close to the Chi sequence. The single-strand-binding protein (SSB) binds rapidly and tightly to the 3′ssDNA tail produced and RecA

is unable to compete with this. However, following Chi recognition, the RecB nuclease domain is released from RecC and RecB actively loads RecA proteins onto ssDNA, displacing SSB and generating the RecA nucleoprotein filament (Dillingham and Kowalczykowski, 2008; Lenhart *et al.*, 2012; Wigley, 2013; Anderson and Kowalczykowski, 1997; Dixon and Kowalczykowski, 1993; Spies *et al.*, 2003; Spies *et al.*, 2007).

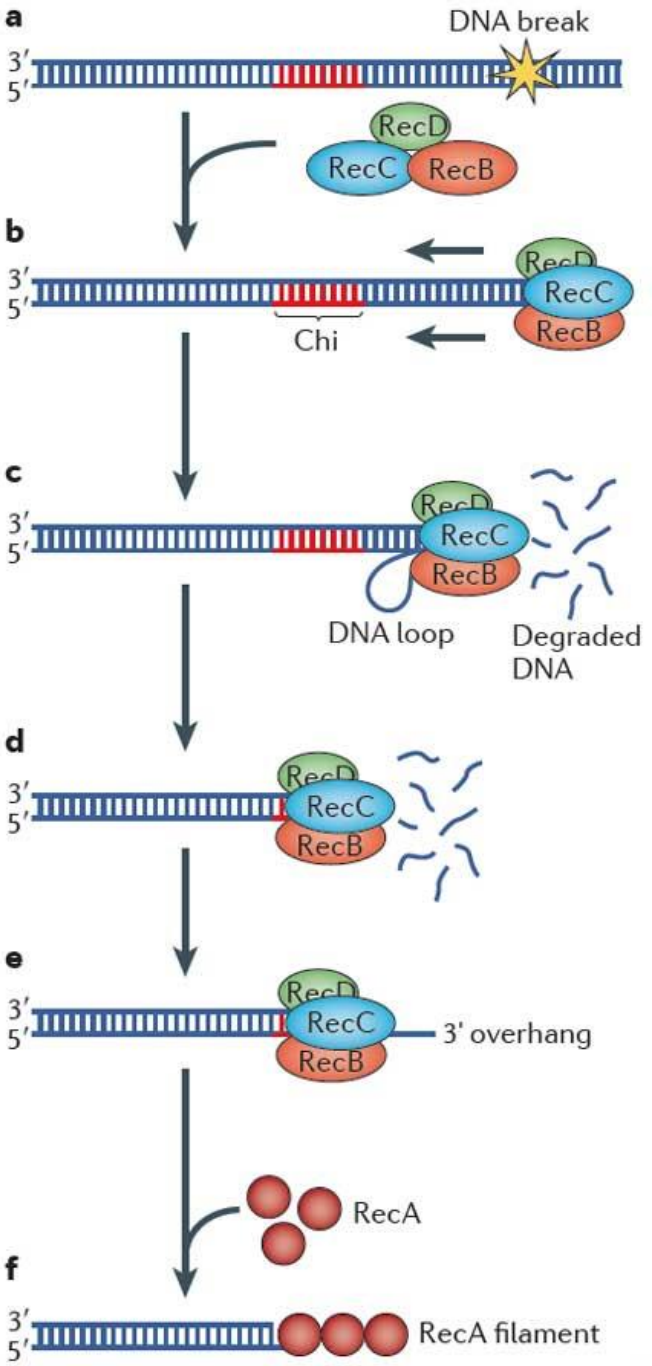


FIG 3.1 Schematic of processing of DNA ends by RecBCD

a) A double strand break on DNA is represented by a star. **b)** RecBCD binds to the DNA end and the complex travels along the DNA unwinding the duplex at the same time. A Chi site is shown in red. **c)** RecB travels in 3′-5′ direction and RecD in 5′-3′ direction. A loop of ssDNA is formed ahead of the RecB motor due to RecD being the faster motor. RecB nuclease domain degrades both strands with 3′ssDNA end being more vigorously hydrolysed than the 5′ ssDNA end. **d)** A Chi site is encountered **e)** At encounter, the 3′-5′ nuclease activity is attenuated and cleavage of the 5′-3′ strand is enhanced resulting in the formation of a 3′ ssDNA. **f)** RecBCD loads RecA onto the ssDNA resulting in the RecA nucleoprotein filament formation. The image is adapted from (Wigley, 2013).

The AddAB complex consists of two subunits, AddA (135 kDa) and AddB (141 kDa) (Chedin and Kowalczykowski, 2002). AddA is both a helicase and an endonuclease, the N-terminal region is a SF1A family helicase with 3′-5′ specificity and the C-terminal region is a RecB-type nuclease which digests DNA at 3′-5′ direction (Lenhart *et al.*, 2012; Wigley, 2013; Saikrishnan *et al.*, 2012; Yeeles and Dillingham, 2007). The AddB subunit is also a nuclease, with the C-terminus forming a RecB-like nuclease domain that cuts ssDNA with 5′-3′ polarity (Yeeles and Dillingham, 2007). In addition, the C-terminus of AddB contains a 4Fe-4S iron-sulphur cluster that seems to have a structural role by binding to duplex DNA ends and stabilizing the protein structure (Yeeles and Dillingham, 2010; Saikrishnan *et al.*, 2012). AddB is analogous to RecC facilitating Chi recognition (Wigley, 2013; Saikrishnan *et al.*, 2012). Similarly to RecBCD, the AddAB enzyme initiates end processing by binding to the dsDNA end, the DNA is unwound and cleaved as the enzyme complex translocates along the DNA. In contrast to RecBCD, unwinding of duplex DNA by AddAB requires SSB (Wigley, 2013; Yeeles *et al.*, 2011). Before a Chi site is encountered, AddA and AddB nucleases each cleave a single DNA strand at equal frequencies as the DNA is extruded from the translocating complex (Yeeles and Dillingham, 2007; Chedin *et al.*, 2000). When AddAB recognises a Chi site the complex binds tightly, unwinding continues and the cleavage of the 3′-5′ strand is attenuated resulting in the production of a 3′-ssDNA for RecA binding (Lenhart *et al.*, 2012; Chedin *et al.*, 2000; Chedin *et al.*, 2006; Yeeles and Dillingham, 2007).

AddAB and RecBCD are involved in a range of important processes in addition to DNA damage and DSB repair including, bacterial conjugation, transduction, bacterial

competence, degradation of host and foreign DNA and the SOS response (Chedin and Kowalczykowski, 2002; Dillingham and Kowalczykowski, 2008; Lenhart *et al.*, 2012). The importance of AddAB or RecBCD activity in bacteria is evident from the fact that bacteria deficient in these genes have reduced viability, are defective in DSB repair and have increased sensitivity to DNA damaging agents (Chedin and Kowalczykowski, 2002). During bacterial conjugation, the DNA enters the cell as a single-strand and is converted to linear dsDNA, by replication, which is a substrate for RecBCD to initiate HR with host homologous sequences (Dillingham and Kowalczykowski, 2008; Smith, 1991). Moreover, during transduction, the ends of injected linear DNA of a previous host can be a substrate for RecBCD/AddAB enzymes for recombination and integration of the linear DNA at homologous sites in the new host (Dillingham and Kowalczykowski, 2008; Smith, 1991). In *B. subtilis*, expression of *addAB* is directly linked to competence in this organism and AddAB is required for recombination of incoming donor DNA (Haijema *et al.*, 1995). Inactivation of *addA* or *addB* genes in *B. subtilis* and *L. lactis* showed reduced frequency of homologous recombination during conjugation and transduction (Chedin and Kowalczykowski, 2002; Kooistra *et al.*, 1997; el Karoui *et al.*, 1998). However, RecBCD can also contribute to resistance of intruding foreign DNA, e.g. phage DNA that exposes free DNA ends is subject to degradation by the enzyme (Dillingham and Kowalczykowski, 2008; Benzinger *et al.*, 1975). In addition, linear DNA fragments produced from cleavage of foreign DNA from restriction modification systems are further degraded by RecBCD (Dillingham and Kowalczykowski, 2008; Handa *et al.*, 2000; Simmon and Lederberg, 1972). In many bacteria DNA damage induces the SOS response (Simmons *et al.*, 2009; Lenhart *et al.*, 2012). In *E. coli*, a critical

step in this reaction is the formation of RecA-ssDNA nucleoprotein filament which is either produced by RecBCD if the DNA substrate is a DSB, or by RecFOR when the DNA substrate is a single-strand gap (Simmons *et al.*, 2009). The RecA-ssDNA filament stimulates the self-cleavage of the LexA repressor, inactivating the protein and subsequently activating transcription of SOS regulated genes (Lenhart *et al.*, 2012; Michel, 2005).

DSB-processing enzymes are also important in the adaptation of microorganisms in extreme environments and bacterial virulence. For example, RecD of the RecBCD complex is required for the normal growth and cell morphology of the marine bacterium *Photobaterium profundum* at high pressures (Bidle and Bartlett, 1999). The ability of *Salmonella enterica typhimurium* to repair DNA damage is essential for survival and virulence, as mutants lacking RecBCD function are avirulent in the murine typhoid model and are also sensitive to reactive oxygen species, nitric oxide, H_2O_2 (produced by phagocytes) or bile salts (Buchmeier *et al.*, 1993; Buchmeier *et al.*, 1995; Cano *et al.*, 2002; Prieto *et al.*, 2006). RecBCD contributes to the survival of *Neisseria gonorrhoeae* following oxidative damage produced by neutrophils and lactobacilli encountered during infection (Stohl and Seifert, 2006). Moreover, *E. coli recBCD* mutants are highly sensitive to nitric oxide (Spek *et al.*, 2001). *Coxiella burnetti* is able to survive in the hostile environment of a parasitophorous vacuole, with lysosomal characteristics, possibly aided by the strong upregulated expression of *addAB* under conditions of oxidative stress (Mertens *et al.*, 2008). *Helicobacter pylori* undergoes oxidative damage soon after infection due to infiltration of neutrophils and macrophages during the innate immune response (Amundsen *et al.*, 2008). DSB repair by AddAB in

this organism is required for successful *in vivo* colonization (Amundsen *et al.*, 2008). In addition, *H. pylori* AddAB promotes recombination-dependent variation of cell surface proteins that may allow evasion of adaptive immune response or, alternatively, may modify bacterial tropism (Amundsen *et al.*, 2008).

Taking into consideration the importance of RecBCD and AddAB enzymes in bacterial infection and in SOS-induced mutagenesis and their presence in most sequenced bacterial species, these enzymes could be possible targets for the development of novel broad spectrum antibacterial therapeutics.

In this study a novel markerless genetic system for *C. difficile* was constructed via the deletion of *addA* and *addB* genes. Furthermore, the function of AddAB enzyme complex in *C. difficile* was analysed.

3.2: RESULTS

3.2.1: Identification and construction of tools for the development of the I-SceI deletion system in *C. difficile*:

Choice of working strain:

The first task towards the development of the genetic system was the choice of a *C. difficile* strain. Ideally, a working strain had to have two characteristics: sensitivity to two different antibiotics, to allow selection of the plasmid constructs in the process; and an available genome

sequence to enable the design of gene deletion primers. In 2006, Sebaihia *et al.*, determined the first complete genome sequence of *C. difficile* 630, which was a multidrug resistant isolate from a hospital patient with pseudomembranous colitis in Zurich, Switzerland (Sebaihia *et al.*, 2006). This was the only sequenced strain at the start of this study. In addition, Hussain *et al.* (2005) generated an erythromycin-sensitive derivative of *C. difficile* 630, named 630Δ*erm*, which is also sensitive to thiamphenicol (Hussain *et al.*, 2005). This strain was later used by Heap *et al.* (2007) to develop the ClosTron genetic system in *C. difficile*, as it allowed the introduction of plasmids that encoded erythromycin and thiamphenicol resistance in the bacterium (Heap *et al.*, 2007). Moreover, a similar strain to 630Δ*erm*, designated JIR8094, was also generated by O'Connor *et al.* (2006)and was used for the construction of *C. difficile* chromosomal mutants via the integration of plasmids encoding thiamphenicol resistance (O'Connor *et al.*, 2006; Dineen *et al.*, 2007; Lyras *et al.*, 2009). In all, strain *C. difficile* 630Δ*erm* was chosen for the development of the system in this study. Despite the later availability of information on genome sequence and genetic modification of other *C. difficil*e strains, including the hypervirulent *C. difficile* 027, this study focused on 630Δ*erm*.

Construction of I-SceI site delivery vectors:

The next step in the development of the genetic system was the identification and construction of the I-SceI site delivery vectors to be used for allelic replacement in the chromosome of *C. difficile*. Ideally, the allele replacement vectors would be delivered and integrated into the

chromosome via suicide vectors to allow efficient selection of merodiploids through the antibiotic resistance conferred from the integrated plasmid. However, as mentioned, conjugative transfer of a suicide plasmid in *C. difficile* has been achieved only once, despite subsequent attempts from other groups (O'Connor *et al.*, 2006; Carter *et al.*, 2005; Liyanage *et al.*, 2001). Thus, replicative but unstable plasmids introduced via conjugation have been used instead to transfer DNA into *C. difficile* (O'Connor *et al.*, 2006; Cartman *et al.*, 2012).

In this study, plasmids pJIR1456 and pJIR2816 (**FIG 3.2**) were chosen for the construction of the I-SceI site delivery vectors pES185 and pES2761 (O'Connor *et al.*, 2006). These plasmids can be transferred into *C. difficile* by RP4-mediated conjugation from an *E. coli* donor and have been used for the insertional inactivation of *C. difficile* chromosomal genes due to their instability, stemming from the presence of a *Clostridium perfringens*-derived pIP404 replication region (O'Connor *et al.*, 2006; Purdy *et al.*, 2002). Both plasmids carry a thiamphenicol resistance gene *catP* allowing their selection at required stages in this genetic system. To construct the I-SceI site delivery vector pES185, oligos ISceI_F_SacI and ISceI_R_SacI carrying the I-SceI recognition site were annealed to each other, 5’-phosphates were added and the oligos ligated into the SacI site of pJIR1456 (**FIG 3.3**). Similarly, oligos ISceI_F_PvuI and ISceI_R_PvuI were phosphorylated at 5’ ends, annealed to each other and ligated into the PvuI site of pJIR2816 generating pES2761 (**FIG 3.4**).

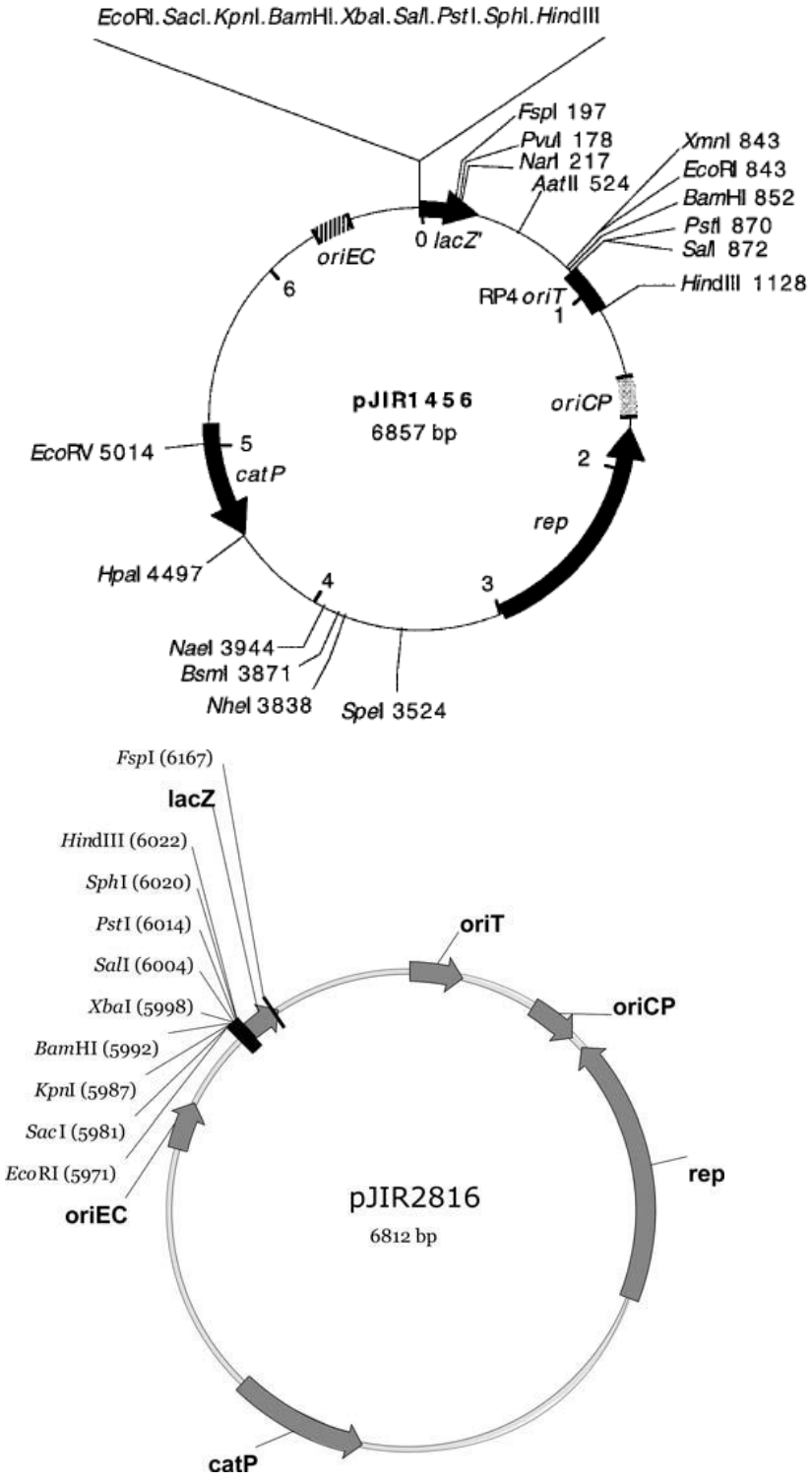
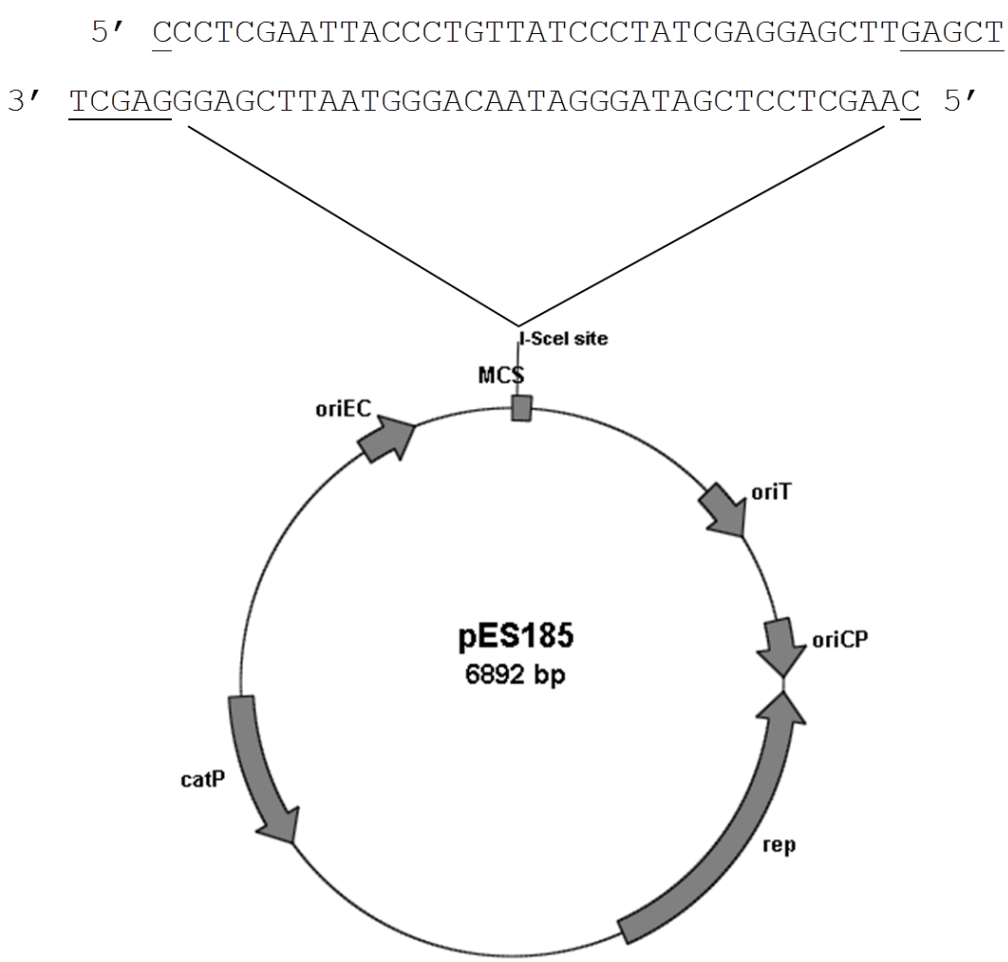


FIG 3.2 Schematic representation of pJIR1456 and pJIR2816

Both plasmids have *oriCP* and *oriEC*, origins of replication from pIP404 and pUC18, respectively; *rep*, replication gene from pIP404; *catP*, chloramphenicol resistance gene from *C. perfringens;* RP4*oriT*, origin of DNA transfer. pJIR1456 adapted from (Lyras and Rood, 1998) and pJIR2816 from (Carter *et al.*, 2010).

A



B

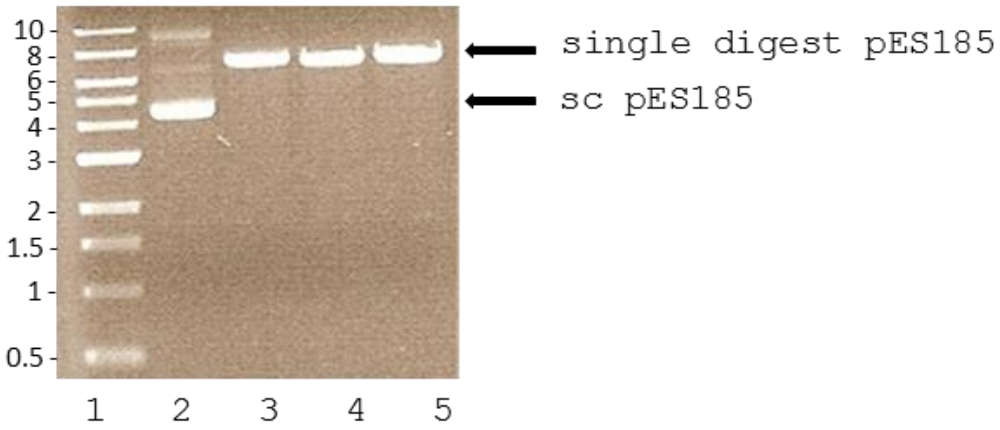
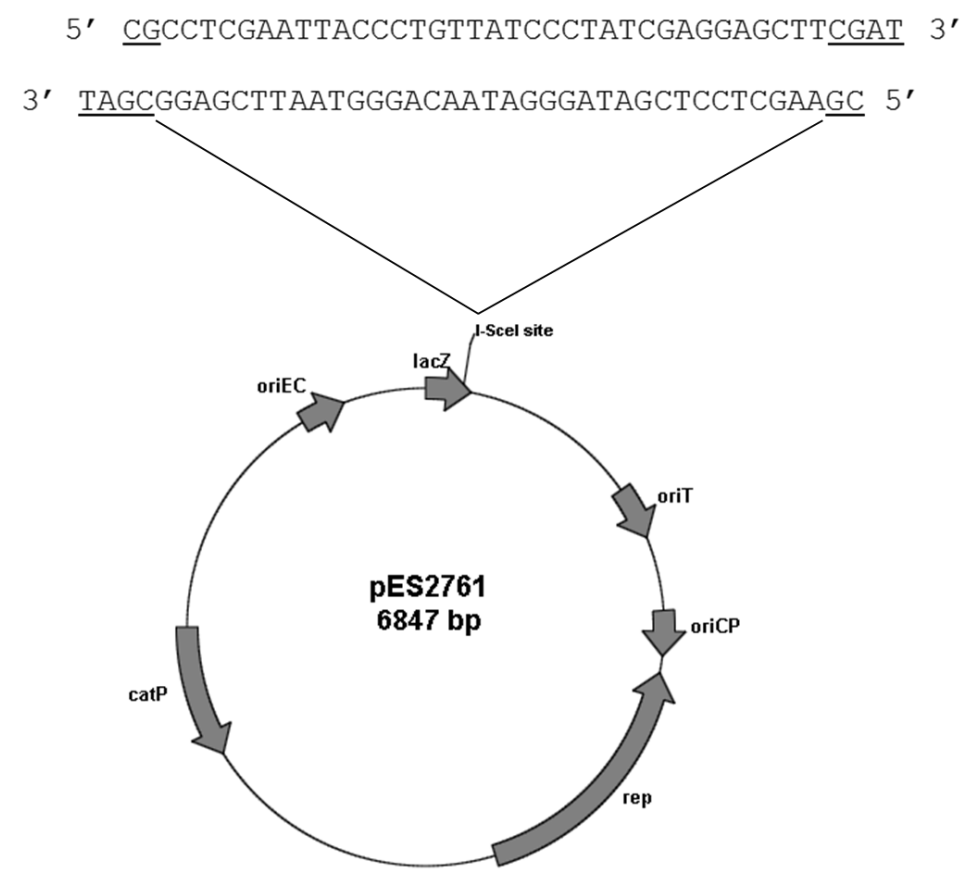


FIG 3.3 Schematic representation and agarose gel analysis of pES185

A] Oligos ISceI_F_SacI and ISceI_R_SacI carrying the I-SceI recognition site and a SacI restriction site (underlined) were ligated into the SacI site of pJIR1456 generating pES185 (6892 bp). **B]** Restriction digest analysis of pES185. Lane 1 refers to 1 kb ladder. Lanes 2, 3, 4 and 5 refer to supercoiled pES185, I-SceI digested, SacI digested and EcoRV digested pES185, respectively.

A



B

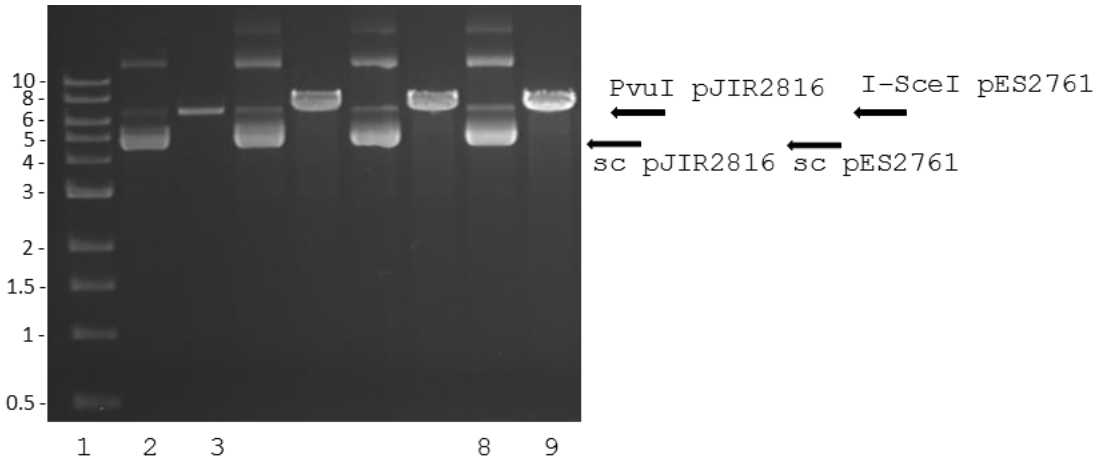


FIG 3.4 Schematic representation and agarose gel analysis of pES2761

A] Oligos ISceI_F_PvuI and ISceI_R_PvuI carrying the I-SceI recognition site and a PvuI restriction site (underlined) were ligated into the PvuI site of pJIR2816 generating pES2761. The plasmid features a *lacZ* gene to aid cloning. **B]** Restriction digest analysis of pES2761. Lane 1 corresponds to 1 kb ladder. Lanes 2 and 3 refer to supercoiled pJIR2816 and PvuI digested pJIR2816, respectively. Lanes 8 and 9 refer to supercoiled pES2761 and I-SceI digested pES2761, respectively.

Construction of I-SceI expression vectors:

To construct the I-SceI expression vector, an appropriate promoter to express I-SceI and an appropriate plasmid i.e., a replicative plasmid encoding an antibiotic resistance different from the I-SceI site delivery vector, were required. This study utilised the ferredoxin (*fdx*) promoter of the *fdx* gene from *Clostridium sporogenes* to express I-SceI. This promoter had previously been used to express the Group II intron in the second generation ClosTron plasmid pMTL007C-E2 (Heap *et al.*, 2010; Cartman *et al.*, 2010). The *E. coli-Clostridium* shuttle vector pMTL82254, that carries an erythromycin marker, was utilised for the construction of the I-SceI expression vector (Heap *et al.*, 2009). The *fdx* promoter (P*fdx*) and ribosome binding site (RBS) were amplified by PCR, from pMTL83353, using primers SbfI-PfdxFor and PfdxRev and the *I-SceI* encoding gene was amplified by PCR from pGB920 using primers PfdxIsceI and SbfI-IsceIRev (**FIG 3.5**) (Heap *et al.*, 2009; Patrick *et al.*, 2009). The P*fdx* and *I-sceI* PCR products were mixed and used as template for a cross-over PCR, using primers SbfI-PfdxFor and SbfI-IsceIRev, resulting in a fused product between the promoter and *I-SceI;* named P*fdx*::*I-SceI* (0.9 kb) (**FIG 3.5**).

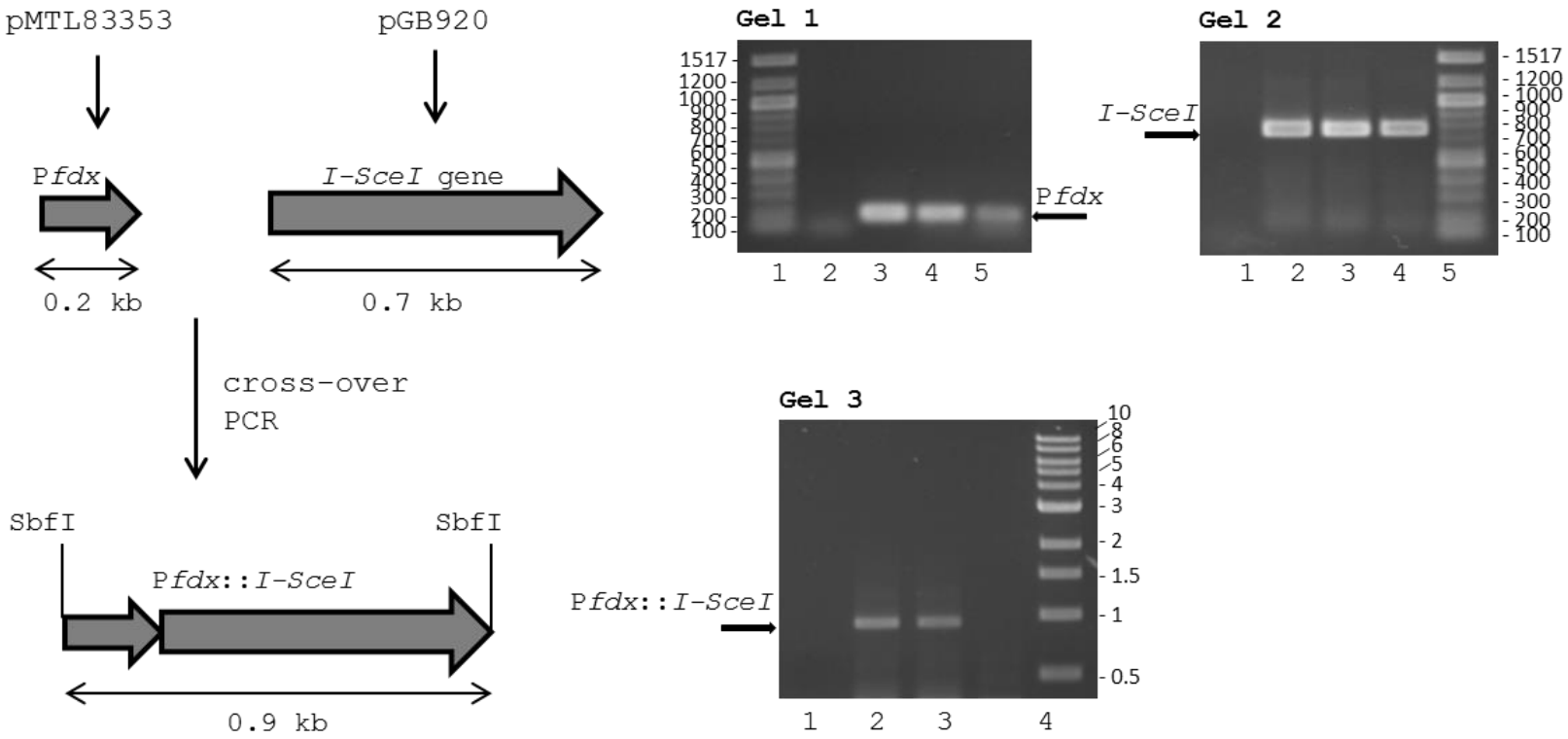


FIG 3.5 Construction of P*fdx*::*I-SceI*

P*fdx* (200 bp) and *I-SceI* (708 bp) were amplified by PCR from pMTL83353 and pGB920, respectively. P*fdx* was fused to *I-SceI* via cross-over PCR producing P*fdx*::*I-SceI* (908 bp) which carried SbfI sites at each end to aid subsequent cloning. Gel 1, gel 2 and gel 3 represent the PCR amplification of P*fdx*, *I-SceI* and P*fdx*::*I-SceI*, respectively. Gel 1 lanes: 1 and 2 correspond to 100 bp ladder and primers control PCR, respectively. Lanes 3, 4, and 5 correspond to P*fdx* product obtained using varying amount of template. Gel 2 lanes: 1 and 5 correspond to primers control PCR and 100 bp ladder, respectively. Lanes 2, 3, and 4 correspond to *I-SceI* product obtained from varying template concentration. Gel 3 lanes: 1 and 4 correspond to primers control PCR and 1 kb ladder, respectively. Lanes 2 and 3 correspond to P*fdx*::*I-SceI* obtained from varying template concentration.

Subsequently, P*fdx*::*I-SceI* was ligated into the SbfI site of pMTL82254 generating the I-SceI expression vector pES288 (**FIG 3.6**). In this scenario, selection of thiamphenicol resistant co-integrants containing pES288 would be via erythromycin resistance. This vector was used for merodiploid resolution and construction of various *C. difficile* deletion strains which will be described in the following chapters.

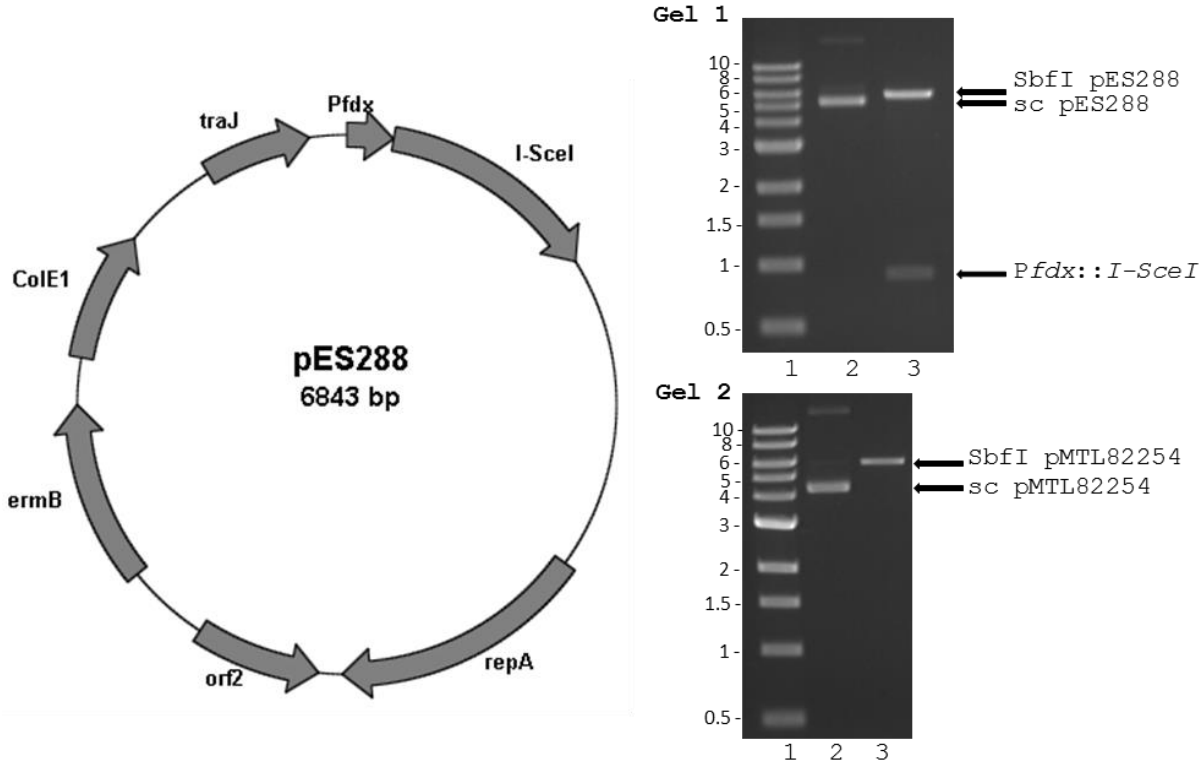


FIG 3.6 Schematic representation and agarose gel analysis of pES288

P*fdx*::*I-SceI* was ligated into the SbfI site of pMTL82254 generating pES288. Gel 1 lanes 1, 2 and 3 correspond to 1 kb ladder, supercoiled pES288 (6843 bp) and SbfI digested pES288 (releasing P*fdx*::*I-SceI*), respectively. Gel 2 lanes 1, 2 and 3 correspond to 1 kb ladder, supercoiled pMTL82254 (5935 bp) and SbfI digested pMTL82254, respectively. Plasmid pES288 features: plasmid replication gene *repA*, putative plasmid replication gene *orf2*, erythromycin resistance marker *ermB*, Gram –ve replicon ColE1, Gram +ve replicon pBP1 and plasmid conjugative transfer gene *traJ*.

Other I-SceI expression vectors were constructed in an attempt to increase expression of I-SceI; namely pES124 and pES86, where the *I-sceI* gene was fused to the glutamate dehydrogenase promoter (*gdh*)from *C. difficile* 630 or to the thiolase promoter (*thl*) from *Clostridium acetobutylicum* ATCC 824 (using pMTL84422 as PCR template), respectively (Mani *et al.*, 2006; Heap *et al.*, 2009). However, their use resulted in no expression of I-SceI enzyme, implied from the absence of merodiploid resolution i.e. no thiamphenicol sensitive strains were obtained from loss of the integrated allele replacement vector (data not shown). All I-SceI expression vectors were sequenced to validate their construction and I-SceI expression was tested via conjugation of each construct into thiamphenicol resistant co-integrants and screening for resolution via loss of thiamphenicol resistance (experiments with pES288 are described below). In addition, it was attempted to validate the I-SceI expression from pES124 or pES288 transconjugants via western blotting with anti-I-SceI antibodies, but these data were inconclusive because the I-SceI antibody used, despite being monoclonal, bound to multiple proteins from whole cell-lysates *C. difficile* 630Δ*erm* and derivates.

3.2.2: Construction of *C. difficile addBA* deletion strain

Arrangement of *addBA* genes and strategy for deletion:

Following the construction of the I-SceI genetic system, its effectiveness was initially tested in *C. difficile* 630Δ*erm* to delete the *addBA* genes. Using the available genomic sequence of *C. difficile* 630, suitable primers were designed to delete *addA* and *addB* genes. **Figure 3.7** shows the arrangement of these genes in *C. difficile* 630, the deletion primers and target deletion region. Upstream of the *addB* gene, there is a putative helicase (CD630_10380) of the UvrD family which is known to be involved in excision and mismatch repair and SOS induction (Janion, 2008; Ossanna and Mount, 1989; Selby and Sancar, 1994). Downstream of the *addA* gene, there is a predicted *sbcD* gene (CD630_10420) encoding the nuclease SbcCD subunit D, followed by the predicted nuclease SbcCD subunit C encoding gene (CD630_10430) which recognises and cleaves secondary DNA structures formed from long DNA palindromes (Connelly *et al.*, 1998; Eykelenboom *et al.*, 2008). Primer sets SphI_addBA1 / addBA2 and addBA2_addBA3 / SphI_addBA4 were designed such that their use in the I-SceI genetic system would result in the markerless deletion of a 7.256 kb region covering almost all of the *addB* and *addA* genes in *C. difficile* 630Δ*erm*.

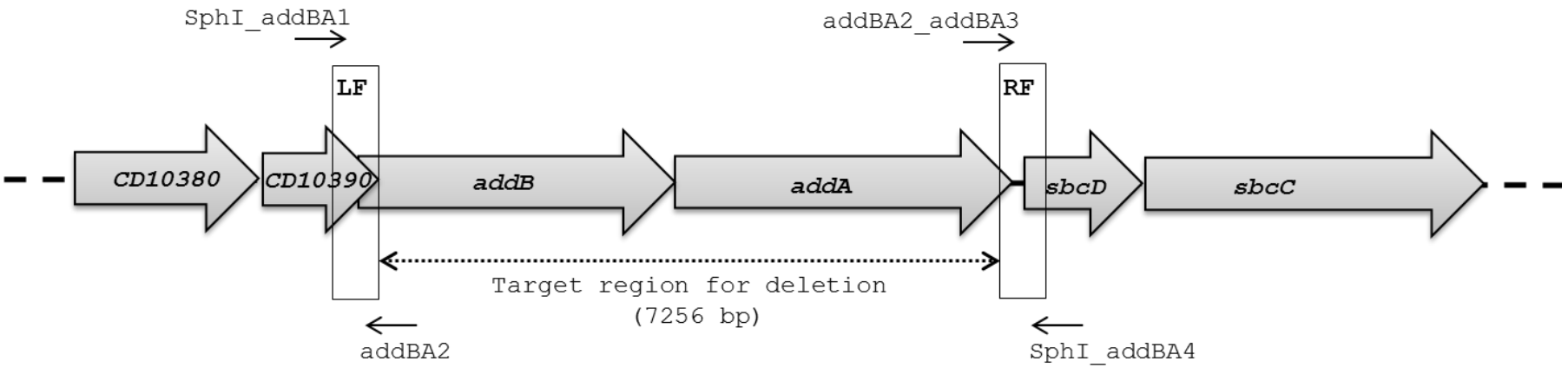


FIG 3.7 Schematic representation of *addA* and *addB* genes and targeted region for deletion

Arrangement of *addA*, *addB* and surrounding genes is shown according to the genome annotation of *C. difficile* 630 (Accesion No. AM180355). Primer sets SphI_addBA1 / addBA2 and addBA2_addBA3 / SphI_addBA4 were designed such that their use in the I-SceI genetic system would result in the markerless deletion of a 7.256 kb sized region which is shown with dotted arrow. The boxes represent flanking regions (left flank LF and right flank RF) of target genes to replace wt type alleles.

Construction of allele replacement vector pES271:

To deliver the Δ*addBA* recombinant allele into the chromosome of *C. difficile* 630Δ*erm*, the construction of allele replacement vector pES271 was required. The first step in achieving this was the construction of the *addBA* deletion cassette. Around 0.5 kb sized regions flanking the *addB* and *addA* genes were amplified by PCR using primer sets SphI_addBA1 / addBA2 and addBA2_addBA3 / SphI_addBA4, with genomic DNA from *C. difficile* 630 as template (**FIG 3.8**). The resulting products were mixed and used as template for subsequent cross-over PCR using primers SphI_addBA1 and SphI_addBA4, generating the *addBA* deletion cassette where the two flanking regions were fused together via their homologous region introduced by addBA2 and addBA2_addBA3 primers (**FIG 3.8**). The *addBA* cross-over PCR resulted into two products, possibly due to the presence of homologous regions in the PCR mixture. However the product of interest, the *addBA* deletion cassette of ~1 kb in size was detected. To generate the allele replacement vector pES271, the *addBA* deletion cassette was ligated into the SphI site of pES185 (**FIG 3.9**). This construct was also sequenced to validate its construction.

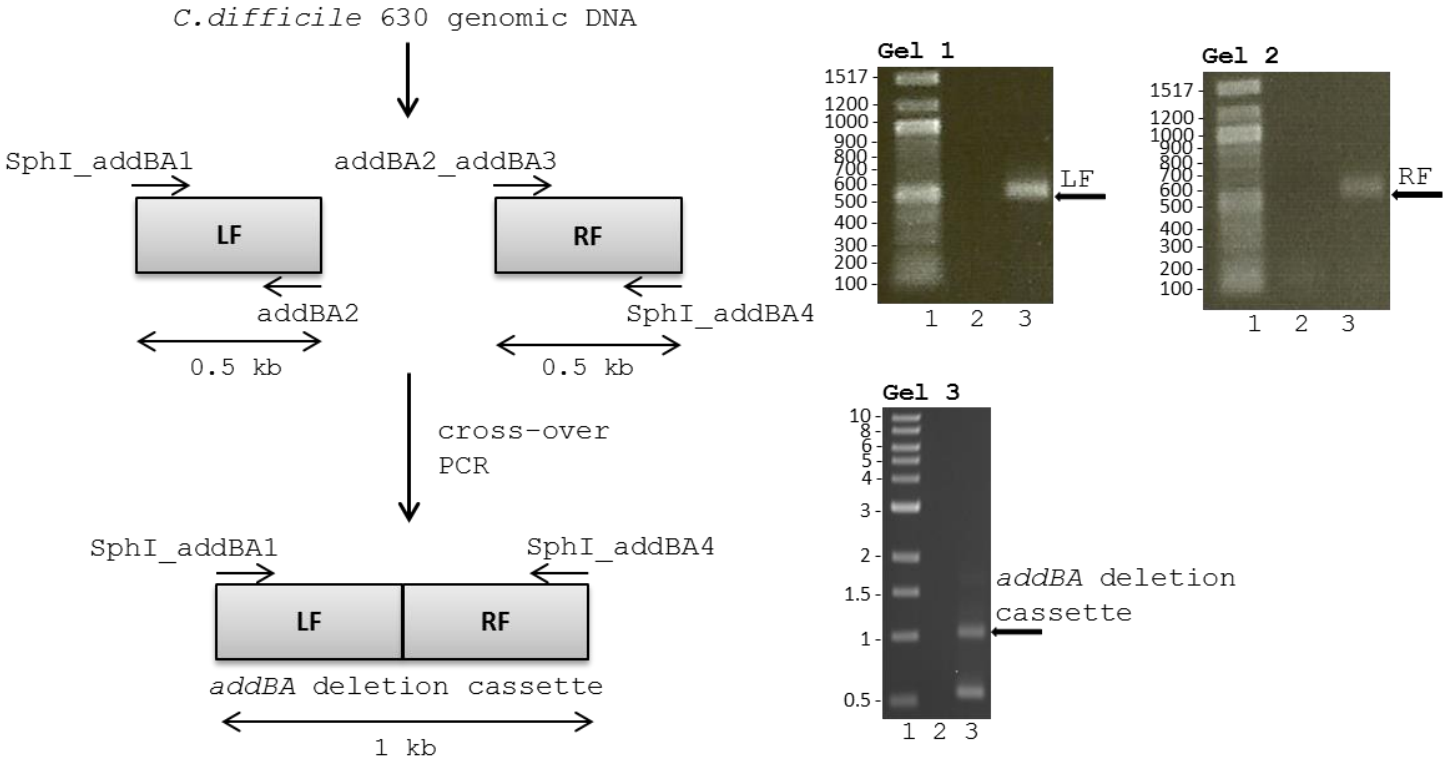


FIG 3.8 Schematic diagram and agarose gel image for *addBA* deletion cassette

Left flanking region (LF, 501 bp) of target genes was amplified using primers SphI_addBA1 and addBA2 whilst right flanking region (RF, 503 bp) of target genes was amplified using primers addBA2_addBA3 and SphI_addBA4. LF and RF were fused together in subsequent cross-over PCR using SphI_addBA1 and SphI_addBA4 primers to generate the *addBA* deletion cassette. Gel 1 lanes 1, 2 and 3 correspond to 100 bp ladder, primers control PCR and LF product, respectively. Gel 2 lanes 1, 2 and 3 correspond to 100 bp ladder, primers control PCR and RF product, respectively. Gel 3 lanes 1, 2 and 3 correspond to 1 kb ladder, primers control PCR and *addBA* deletion cassette, respectively.

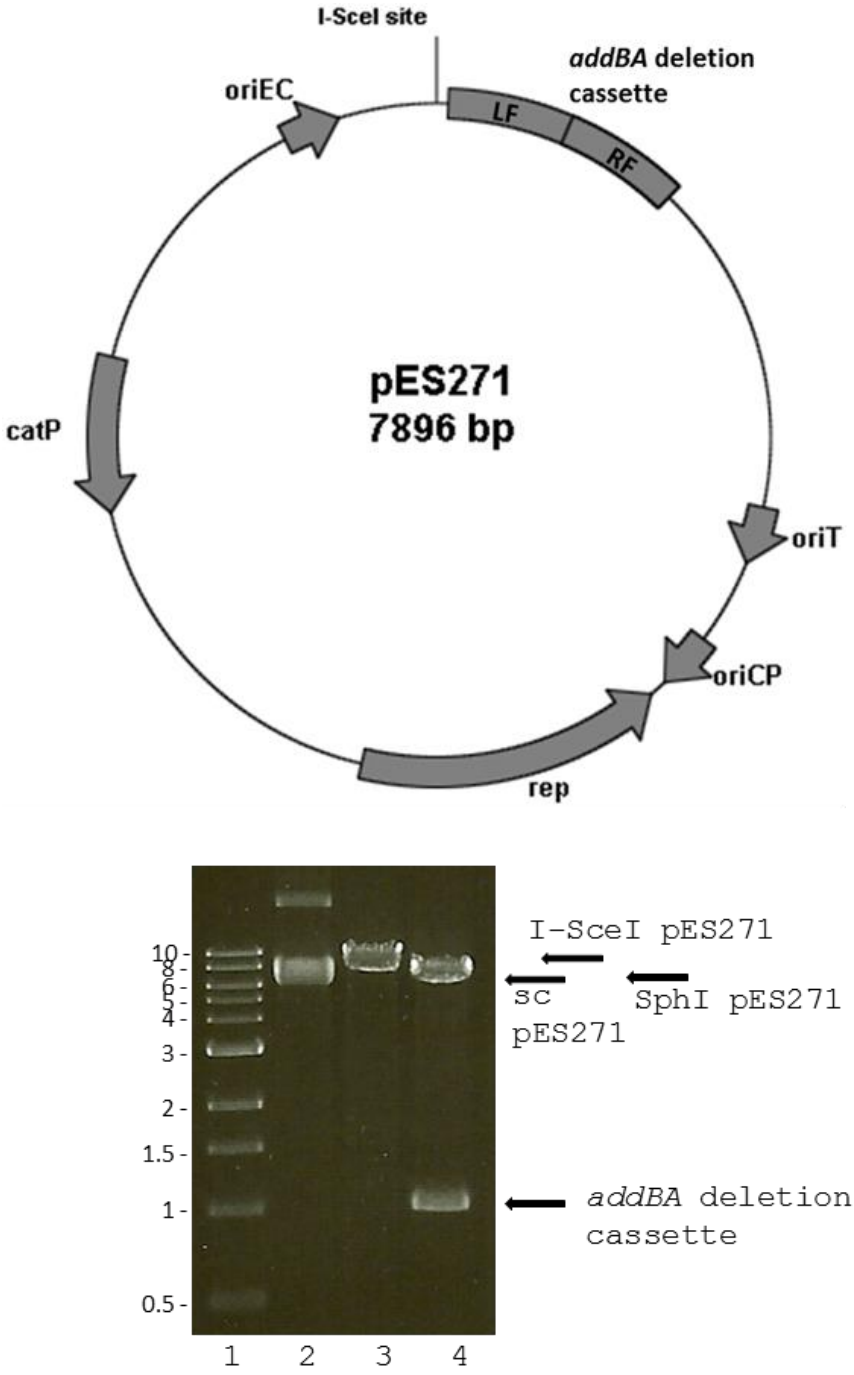


FIG 3.9 Schematic representation and agarose gel analysis of pES271 allele replacement vector

The *addBA* deletion cassette was ligated into the SphI site of pES185 generating the allele replacement vector pES271. Gel lanes 1, 2, 3 and 4 correspond to 1 kb ladder, supercoiled pES271, I-SceI digested pES271 and SphI digested pES271 (releasing *addBA* deletion cassette), respectively.

Conjugation and integration of pES271 in *C. difficile* 630Δ*erm:*

Following the construction of pES271 the next step was the introduction of the construct into *C. difficile* followed by integration into the chromosome. As the plasmid harbours a *catP* gene in its backbone, this allowed selection of thiamphenicol resistance in strains carrying the plasmid either in free replicative or chromosomal integrated form. The recombinant plasmid was conjugated from *E. coli* S17-1λ*pir* into *C. difficile* using exponential cultures grown in BHIS (BHI supplemented with haemin and L-cysteine). Ten putative transconjugants, No. 1-10, were recovered after 72-h and subsequently streaked twice on BHIS containing Tm15 µg/ml (BHISTm15) for purity and to allow for a single cross-over event. *E. coli* contaminants were excluded by replica screening the colonies via aerobic incubation on Luria-Bertani (LB) agar. The presence of pES271 in *C. difficile* was confirmed via amplification of the *addBA* deletion cassette using PCR with SphI_addBA1 and SphI_addBA4 primers and template chromosomal DNA extracted from 48-h grown putative transconjugants. The *addBA* deletion cassette was only amplified from transconjugant No. 1 which also showed better growth on BHISTm15 (**FIG 3.10**).

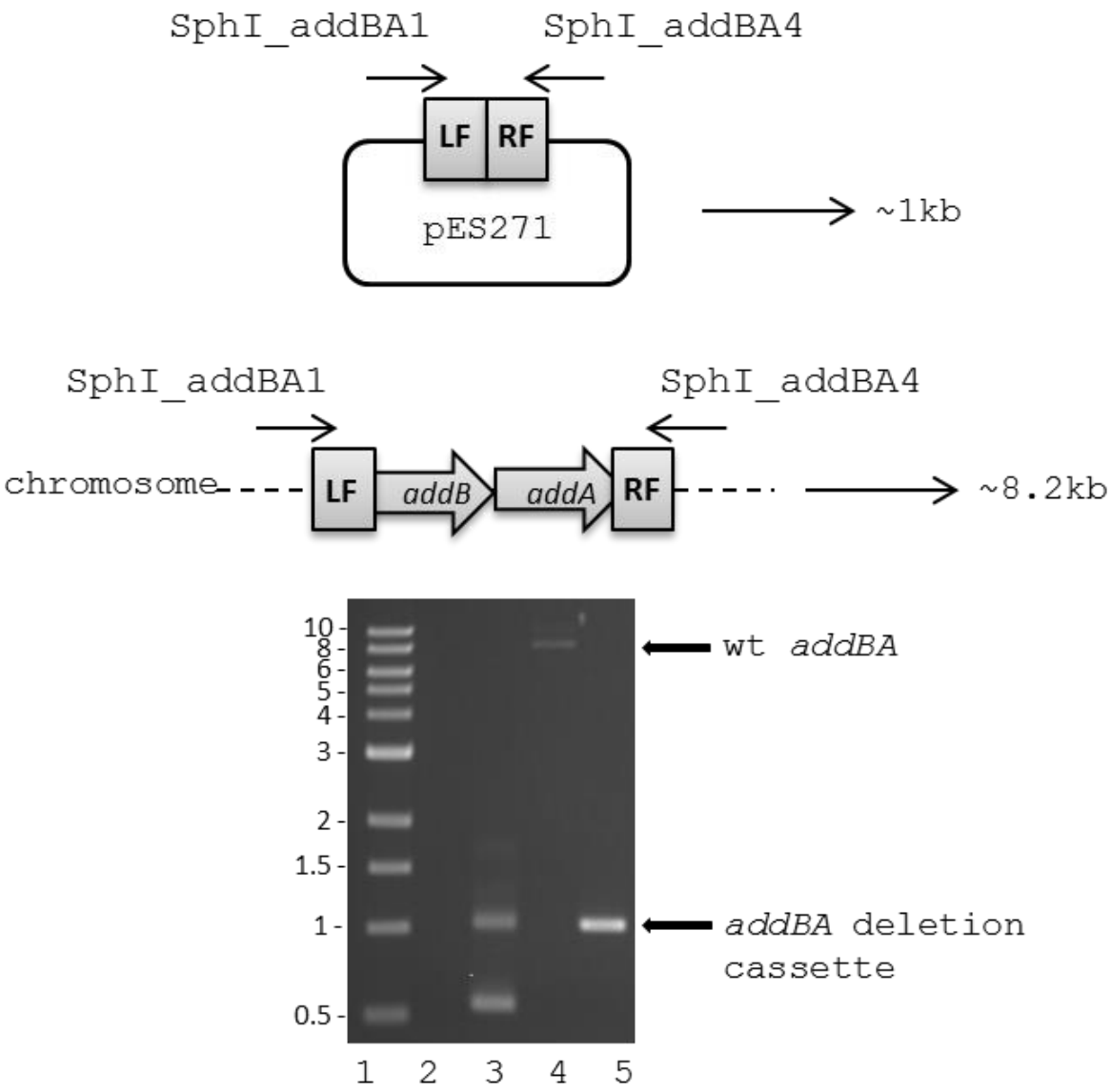


FIG 3.10 Schematic diagram and agarose gel image analysis for PCR screening of pES271 transconjugants

Primers SphI_addBA1 and SphI_addBA4 were used to confirm the presence of pES271 via the amplification of the *addBA* deletion cassette (~1 kb) by PCR from chromosomal extract of putative transconjugants. Wild type strains not containing pES271, result in the amplification of wt *addBA* (8 kb) using the same primers. Gel lanes 1, 2, and 3 correspond to 1 kb ladder, primers control PCR and *addBA* deletion cassette (undigested), respectively. Lane 4 corresponds to wt *addBA* product from wt *C. difficile* 630Δ*erm* and lane 5 to *addBA* deletion cassette product from pES271/630Δ*erm*/transconjugant No. 1.

To allow for the loss of non-integrated pES271 and for the selection of putative merodiploid strains, transconjugant strains were passaged through BHIS lacking thiamphenicol multiple times and then plated on BHISTm15. From this process, 11 possible thiamphenicol resistant integrants were selected. Confirmation of pES271 integration into the chromosome of these strains was achieved by PCR amplification of a 1.2 kb integration product using SphI_addBA1 and INTBARev or SphI-addBA4 and INTBAFor primer sets and chromosomal DNA of putative merodiploids. **Figures 3.11 & 3.12** show the integration of pES271 in No. 1-10 merodiploids via the region downstream of the *addA* gene and the integration of pES271 in No. 11 merodiploid via the region upstream of the *addB* gene. For experiments aiming to test for the expression of I-SceI, resolution and deletion of target genes in the recombinant strains, merodiploids pES271/630Δ*erm*/No. 8 and No. 11 were selected.

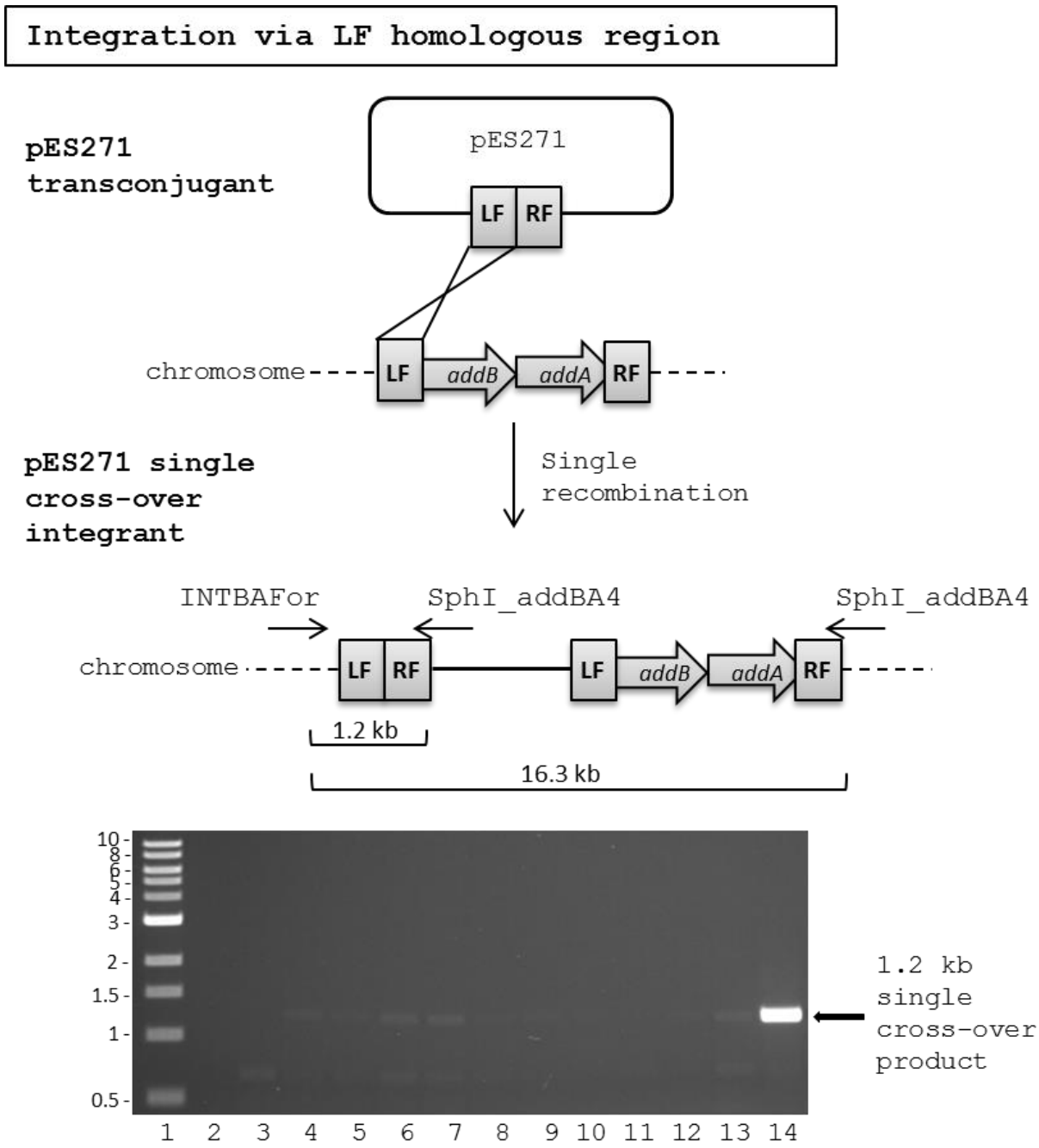


FIG 3.11 Schematic diagram and agarose gel image analysis for PCR screening of pES271 single cross-over integration via the left flanking homologous region

Primers INTBAFor and SphI_addBA4 were used to confirm the event of single cross-over chromosomal integration of pES271 though the LF homologous region via the amplification of a 1.2 kb product by PCR from chromosomal extractions. Note that INTBAFor anneals 220 bp upstream of LF onto the chromosome. Solid line represents plasmid sequence. Dashed line represents chromosomal sequence. Gel lanes 1, 2 and 3 correspond to 1 kb ladder, primers control PCR, and wt *C. difficile* 630Δ*erm* control PCR, respectively. Lanes 4 to 14 correspond to merodiploids pES271/630Δ*erm*/No.1-11. Single cross-over integration of pES271 via LF was confirmed in merodiploid pES271/630Δ*erm*/No. 11.

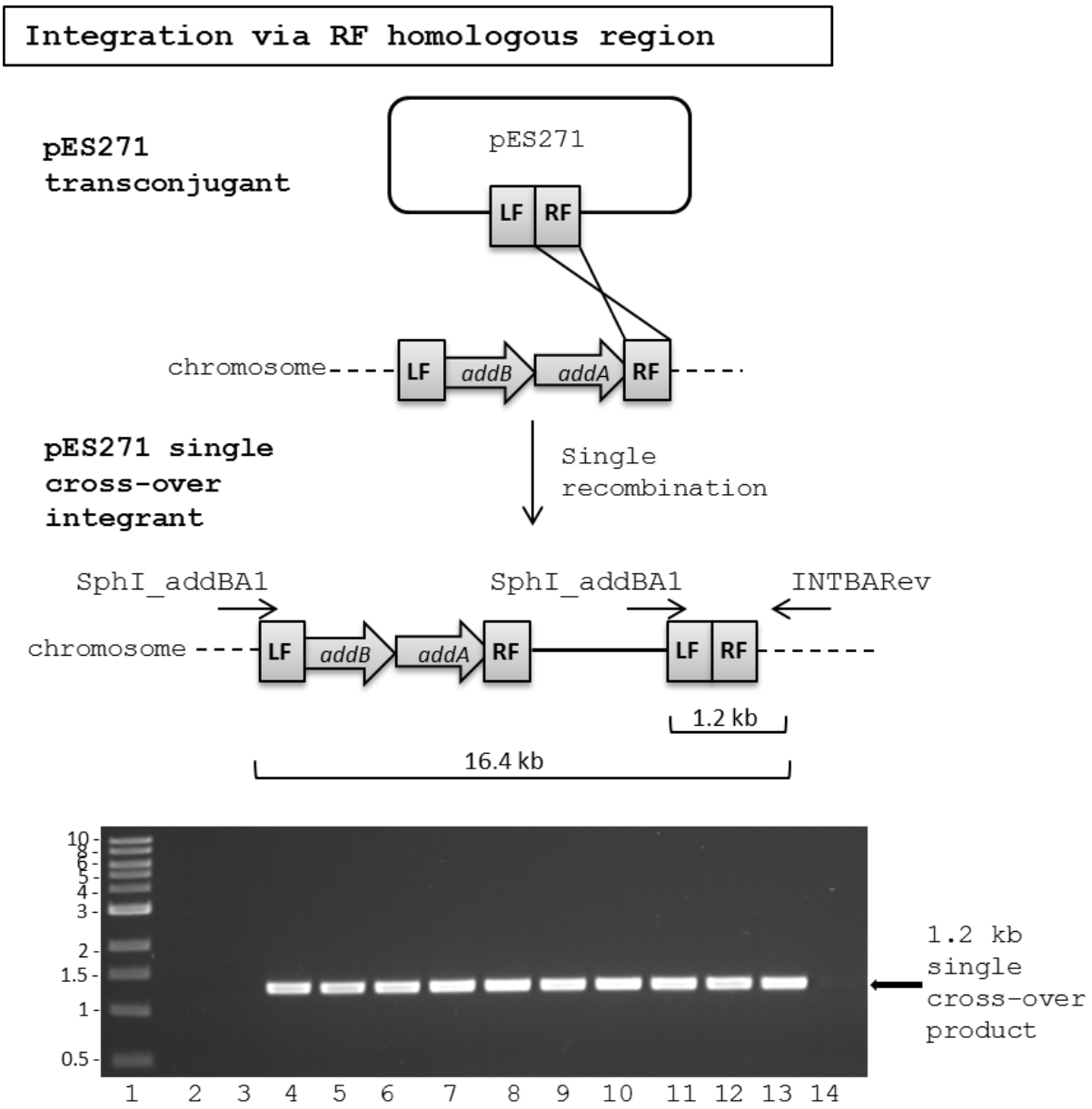


FIG 3.12 Schematic diagram and agarose gel image analysis for PCR screening of pES271 single cross-over integration via the right flanking homologous region

Primers SphI_addBA1 and INTBARev were used to confirm the event of single cross-over chromosomal integration of pES271 though the RF homologous region via the amplification of a 1.2 kb product by PCR from chromosomal extractions. Note that INTBARev anneals 270 bp downstream of RF onto the chromosome. Solid line represents plasmid sequence. Dashed line represents chromosomal sequence. Gel lanes 1, 2 and 3 correspond to 1 kb ladder, primers control PCR, and wt *C. difficile* 630Δ*erm* control PCR, respectively. Lanes 4 to 14 correspond to merodiploids pES271/630Δ*erm*/No.1-11. Single cross-over integration of pES271 via RF was confirmed in merodiploids pES271/630Δ*erm*/No. 1-10.

Introduction of I-SceI expression vector and induction of a double-strand break:

Following the construction of the Δ*addBA* plasmid integrants, the next step involved the introduction of the I-SceI expression vector pES288 into these strains and generation of a DSB at the integrated I-SceI site. Since the *fdx* promoter had been used for the constitutive expression of the Group II intron in the ClosTron system it was assumed that I-SceI under the control of this promoter would result in constitutive expression, giving rise to a DSB as soon as the plasmid entered the integrants (Heap *et al.*, 2010; Cartman *et al.*, 2010). Identification of resolved strains would therefore be easily achieved via screening for loss of thiamphenicol resistance following loss of integrated plasmid which carried a *catP* marker.

To test for the expression of I-SceI under P*fdx*, initial experiments involved the introduction of I-SceI expression vector pES288, into the merodiploid pES271/630Δerm/No.8 strain via stationary mating. From this mating, the resulting 176 putative transconjugants were screened for loss of thiamphenicol resistance to detect for resolution, screened for erythromycin resistance to confirm the presence of pES288, and screened for aerobic growth to exclude *E. coli* contamination. All transconjugants were erythromycin resistant and 97% were anaerobes. Moreover, all were thiamphenicol resistant suggesting that, in this experiment, resolution did not occur when the I-SceI expressing vector entered the cells or there was a low I-SceI concentration in cells. Therefore, to enrich for a population of cells expressing I-SceI with an increased probability of obtaining a DSB and for purity, two erythromycin resistant pES288 transconjugants were subcultured two to four times on BHISErm10 or AIMSErm10

(BHIS or AIMS supplemented with erythromycin at 10 µg/ml) agar medium to keep the selective pressure for pES288. Colonies were selected at random from second or fourth subcultures and were screened for thiamphenicol resistance and aerobic growth. All the resolved colonies (28.6%) i.e. thiamphenicol sensitive colonies isolated (54 sensitive colonies from total 189 screened) were confirmed via subsequent tests (not described here) to have reverted to wt, which suggested that an *addBA* deletion might be lethal to *C. difficile* or might reduce the growth rate of the strain to an extent where it was hard to isolate because the wild type might dominate the population.

Up to this point, isolation of an *addBA* deletion strain had been attempted via resolution of the merodiploid pES271/630Δ*erm*/No.8 strain where integration had occurred via the downstream region of the *addA* gene. The next attempt to isolate an *addBA* deletion strain involved the introduction of the I-SceI expression vector pES288 into the merodiploid pES271/630Δ*erm*/No.11, where integration of allele replacement vector occurred via the upstream region of *addB* gene. To isolate two independent *addBA* deletion mutants, the vector pES288 was conjugated into the merodiploid pES271/630Δ*erm*/No.11 strain via stationary mating at two independent occasions using AIMS medium. On both occasions, putative transconjugants were recovered and subcultured twice (sub1 & sub2) on AIMSErm10, then from each subculture (sub1 & sub2) colonies were picked and screened for thiamphenicol sensitivity to test for resolution. It was postulated that an *addBA* deletion would delay the growth of *C. difficile* and so smaller sized colonies were preferentially picked.

From the first mating, 55 putative transconjugants were obtained, then screened for thiamphenicol sensitivity and

their resistance implied no resolution, which agreed with the previous result from pES271/630Δerm/No.8 – pES288 transconjugants that further sub-streaking may be required to increase resolution. Therefore, any subsequent resolution screening tests were performed after subculturing putative pES288 transconjugants at least once, i.e. after enrichment for a resolved population. Following this reasoning, putative transconjugants were streaked twice on AIMSErm10 (sub1 and sub2) and colonies from each subculture were screened for thiamphenicol sensitivity, indicative of resolution. Screening for thiamphenicol sensitive strains from AIMSErm10 (sub1) resulted in 33.3% resolved colonies (10 sensitive colonies from total 30 screened) and screening of thiamphenicol sensitivity from sub2 resulted in 11.8% resolved colonies (4 sensitive colonies from total 34 screened). From the second mating, 13 putative transconjugants were obtained which were streaked on AIMSErm10 twice (sub1 & sub2). Screening for thiamphenicol sensitivity from sub1 gave 7% resolved colonies (9 sensitive colonies from total 135 screened) and from sub2 gave 11.9% resolved colonies (14 sensitive colonies from total 118 screened). All resolved strains obtained, were further tested to confirm reversion to wt or deletion of *addB*A genes (described below).

Screening for double recombination events and confirmation of resolved wt or Δ*addBA* strains:

It was anticipated that deletion of *addB*A genes would result in sensitivity to DNA damaging agents. This allowed the initial screening for putative deletion strains via metronidazole sensitivity. All resolved strains isolated above, together with wt 630Δ*erm* and the merodiploid

pES271/630Δ*erm*/No.11 strain were streaked on AIMS containing metronidazole at 0.5 µg/ml (AIMSMZ0.5). This resulted in the identification of two metronidazole sensitive strains, each isolated from the two independent matings. The efficiency of obtaining an *addBA* deletion strain was 7% in the first occasion and 4.3% in the second. To investigate whether the phenotype of metronidazole sensitivity of the strains was due to resolution and deletion of *addBA*, the next step involved screening of the strains via PCR. Chromosomal DNA extracted from all isolated thiamphenicol sensitive strains and genomic DNA from wt 630Δ*erm*, was used as template for the amplification of the *addBA* region by PCR using primers SphI_addBA1 and SphI_addBA4. The *addBA* deletion cassette (~1 kb) was amplified from the two metronidazole sensitive strains, while the wt *addBA* region (~8.2 kb) was amplified from the wt 630Δ*erm* and from the thiamphenicol sensitive but metronidazole resistant strains (**FIG 3.13**). This result confirmed the isolation of strains that resolved to wt and two resolved *addBA* deletion strains, from two independent events, which were named Δ*addBA* 24 and Δ*addBA* 242. To provide further confirmation of the genotype of the Δ*addBA* strains, the *addBA* deletion cassette PCR products from above were sequenced. The data confirmed that the *addBA* deletion cassette, that replaced the wt allele in Δ*addBA* 24 and Δ*addBA* 242, and the cross-over point between the two flanking regions were both correct (**FIG 3.13**).

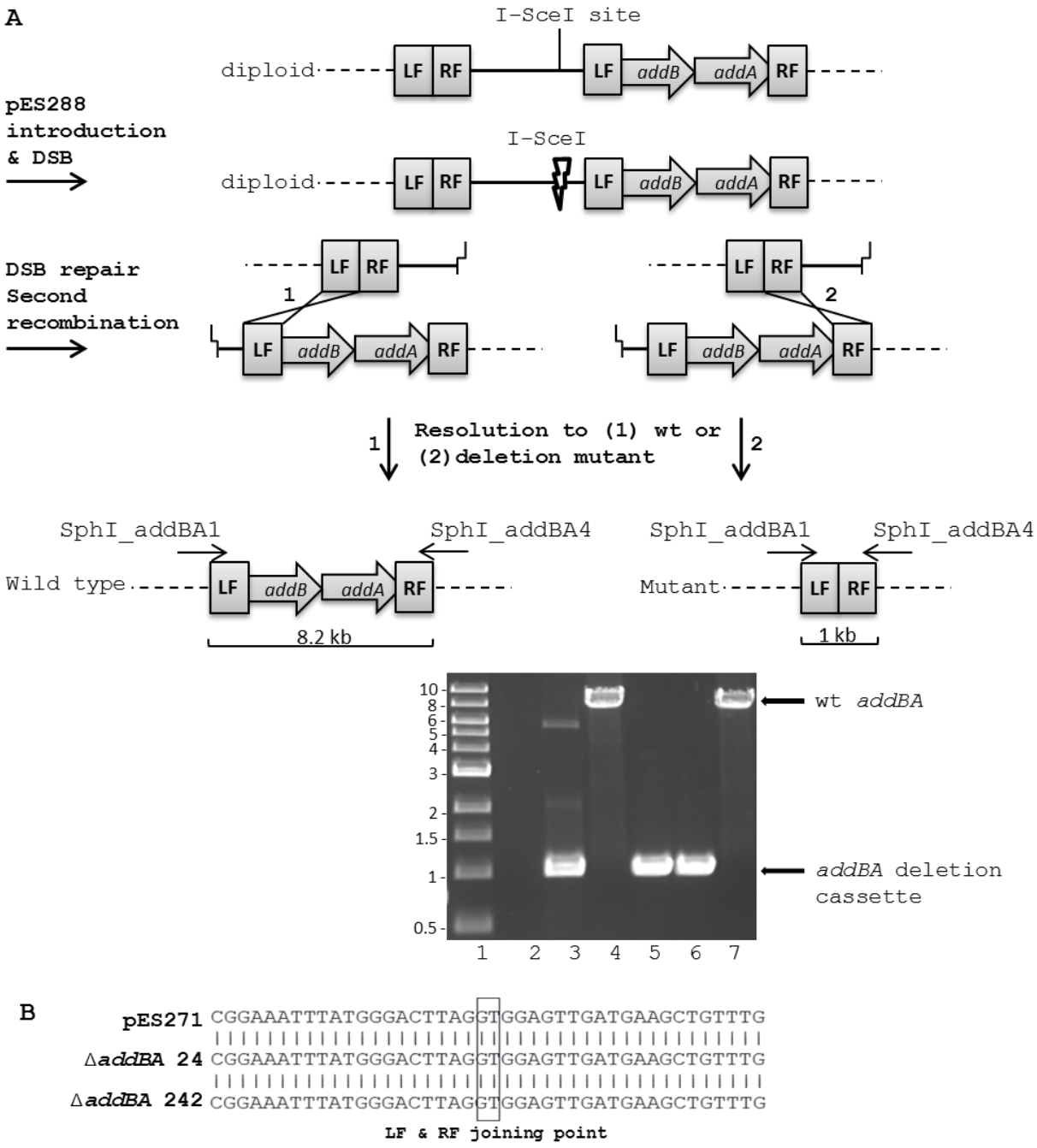


FIG 3.13 Schematic diagram and agarose gel image analysis for screening of second cross-over recombination event

A] Upon introduction of pES288 into the merodiploid, expressed I-SceI creates a DSB resulting in a second recombination event and resolution to wild type (1) or *addBA* deletion strain (2). Solid line represents plasmid sequence. Dashed line represents chromosomal sequence. Primers SphI_addBA1 and SphI_addBA4 were used to screen for a second cross-over event via the amplification of an 8 kb or 1 kb PCR product confirming resolution to wt or *addBA* deletion strain, respectively. Gel image represents the amplification by PCR of the wt *addBA* gene from wt strains or the *addBA* deletion cassette from the deletion strains. Gel lane 1 and 2 correspond to 1 kb ladder and primers control PCR, respectively. Lanes 3, 5 and 6 correspond to the amplification of *addBA*

deletion cassette from pES271, Δ*addBA* 24 and Δ*addBA* 242, respectively. Lanes 4 and 7 correspond to amplification of wt *addBA* genes from *C. difficile* 630Δ*erm* and resolved wt R20, respectively. **B]** A partial sequence alignment is shown of the sequenced *addBA* deletion cassettes amplified from Δ*addBA* 24 and Δ*addBA* 242 above with the sequenced *addBA* deletion cassette from pES271. The sequences show the joining point between the left and right flanking regions of the deleted *addB* and *addA* genes.

Confirmation of construction of *addBA* deletion strain due to I-SceI:

The next step towards the validation of the I-SceI genetic system was to confirm that the isolation of the *addBA* deletion strains was due to I-SceI induced DSB repair and not due to rare spontaneous second cross-over event. To test this, the shuttle vector pMTL82254 from which pES288 was made, was conjugated into the merodiploid pES271/630Δ*erm*/No.11 via stationary mating. Forty of the transconjugants obtained were streaked on AIMErm10 (sub 1) and in total 80 colonies (two from each sub 1 culture) were screened for thiamphenicol sensitivity. Subsequently, ten randomly selected transconjugants from sub 1 were further subcultured on AIMErm10 (sub 2) and eight colonies from each sub 2 culture were screened for thiamphenicol sensitivity. All colonies screened were thiamphenicol resistant, which indicated that no resolution had occurred. This result was in contrast to the resolution observed following the introduction of pES288 into merodiploid pES271/630Δerm/No.11. It was therefore concluded that, resolution was not due to spontaneous secondary recombination events, but due to expression of I-SceI from pES288.

3.2.3: The function of AddAB in *C. difficile*

Sporulation and toxin production:

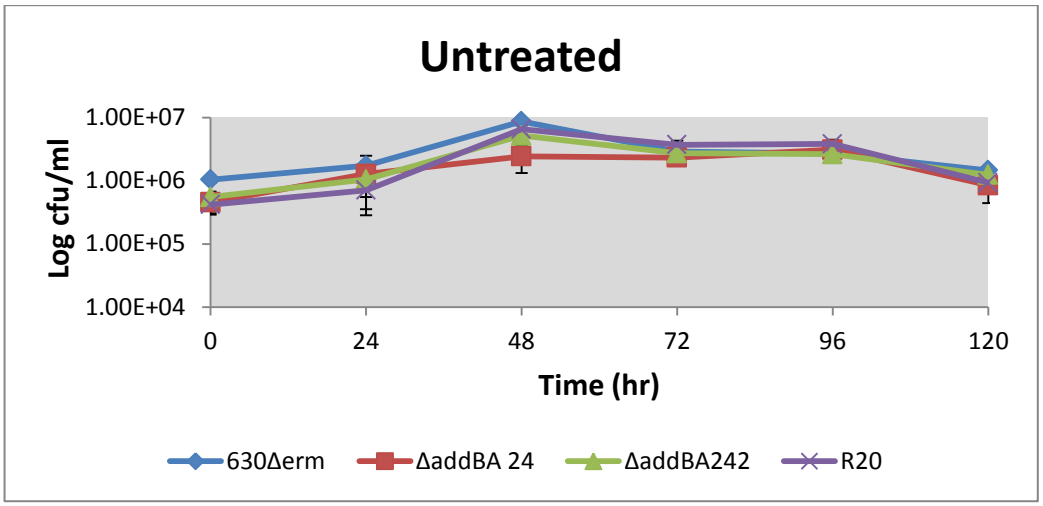
In *C. difficile* infection, as mentioned, sporulation is the aetiologic agent and toxins are well-established virulence determinants. Therefore, sporulation and toxin production were measured in the Δ*addBA* strains, the wt 630Δ*erm* and the resolved wt R20 (generated from resolution of pES271/630Δ*erm*/No.11 from the first mating above).

Spore measurement:

To compare sporulation characteristics between the strains wt 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and resolved wt R20, the development of heat resistant spores was measured over a five day period at 24-h intervals, during which time the growth (OD_{600}) and viability of non-heat-treated samples were also measured. To determine heat resistant and untreated viable cfu/ml counts, samples were plated onto AIM agar supplemented with 0.1% sodium taurocholate(AIMTA), which had been previously shown to enhance *C. difficile* spore recovery when used in rich medium (Burns *et al.*, 2010; Wilson *et al.*, 1982). The data from the heat-treated cfu/ml counts (**FIG 3.14**) showed that all strains, except wt 630Δ*erm*, displayed a similar sporulation pattern through the 120-h period with an increase in spore numbers until 24-h and a slower increase thereafter. In contrast, the wt 630Δ*erm* strain yielded sporulation numbers at 72-h similar to those produced at 24-h by the other strains (at 24-h 630Δ*erm* produced 100-fold less spores in comparison to the rest), but showed a steady gradual increase in heat resistant spore development throughout the assay period,

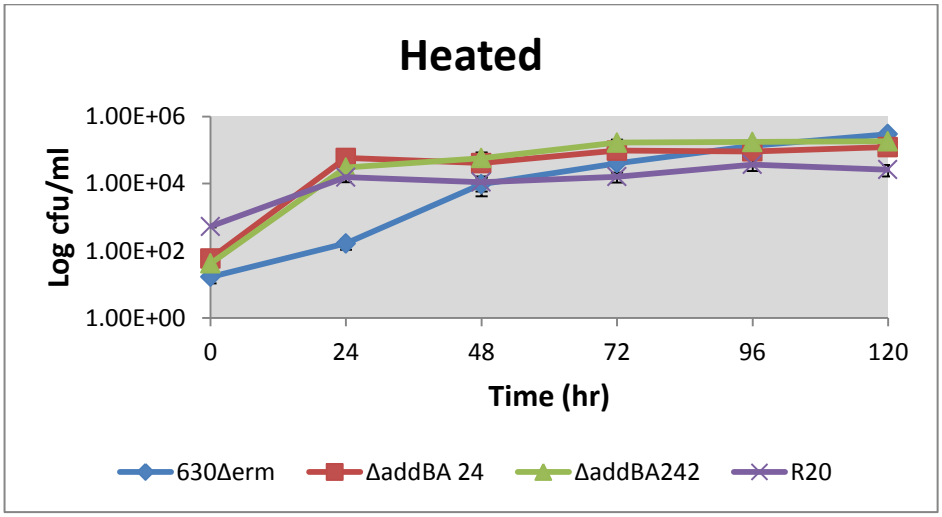
reaching final numbers similar to Δ*addBA* 24 and Δ*addBA* 242 at 120-h. The wt strain R20 yielded an average of 10-fold more heat resistant spores at 0-h but the final spore yield at 120-h was an average of 10-fold lower than the rest of the strains. In addition, the data from the non-heat-treated cfu/ml counts showed that all strains had similar CFU and viability patterns through the 120-h period, with an increase in viable CFU until 48-h at which time numbers peaked and decreased thereafter (**FIG 3.14**). The strain Δ*addBA* 24 showed slightly lower viability at 48-h, but a triplicate experiment from another occasion resulted in similar CFU at 48-h to the rest of the strains, suggesting an erroneous result. The measurement of growth by OD_{600} showed the same pattern in all strains, with a maximum optical density at 24-h and a decline in growth thereafter (**FIG 3.14**). There was little difference in growth measurement between the strains during the sporulation experiment, although, Δ*addBA* 24 and Δ*addBA* 242 displayed lower optical density at 24 and 48-h compared to the wt strains.

A



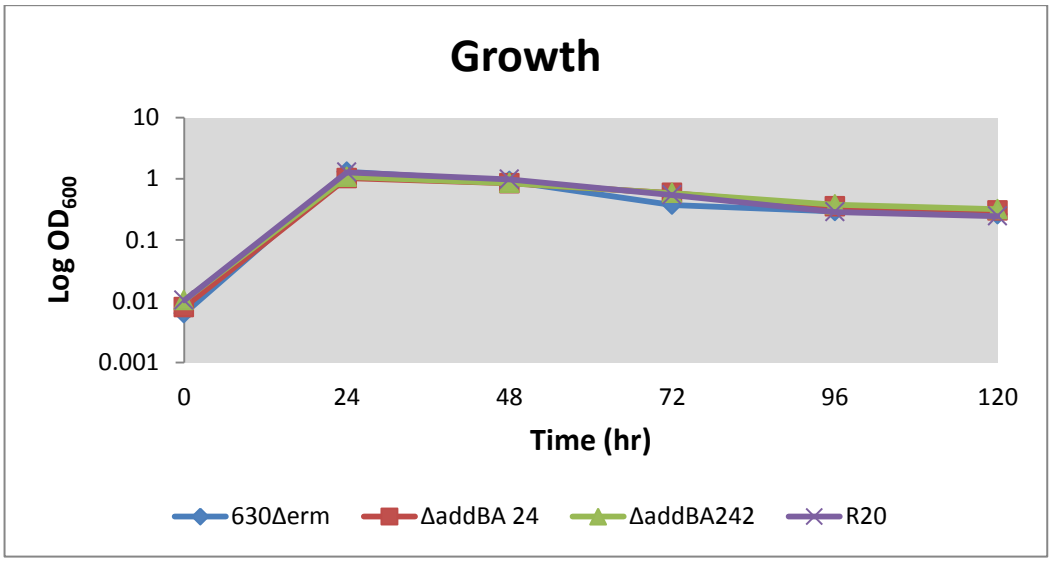
Untreated cfu/ml						
	0	24	48	72	96	120
630Δ*erm*	1.04×10^6	1.73×10^6	8.66×10^6	2.89×10^6	2.71×10^6	1.46×10^6
Δ*addBA* 24	4.53×10^5	1.27×10^6	2.44×10^6	2.33×10^6	3.09×10^6	8.48×10^5
Δ*addBA*242	5.52×10^5	1.04×10^6	5.16×10^6	2.73×10^6	2.63×10^6	1.24×10^6
R20	4.20×10^5	7.09×10^5	6.54×10^6	3.70×10^6	3.83×10^6	9.33×10^5

B



Heated cfu/ml						
	0	24	48	72	96	120
630Δ*erm*	1.67×10^1	1.64×10^2	9.62×10^3	4.03×10^4	1.36×10^5	2.97×10^5
Δ*addBA* 24	5.87×10^1	5.81×10^4	4.07×10^4	9.62×10^4	9.12×10^4	1.22×10^5
Δ*addBA*242	4.17×10^1	3.01×10^4	5.73×10^4	1.67×10^5	1.73×10^5	1.82×10^5
R20	5.30×10^2	1.56×10^4	1.11×10^4	1.60×10^4	3.71×10^4	2.61×10^4

C



Growth OD_{600}						
	0	24	48	72	96	120
630Δ*erm*	0.006	1.297	0.912	0.374	0.294	0.262
Δ*addBA* 24	0.008	1.024	0.841	0.582	0.352	0.302
Δ*addBA*242	0.010	1.069	0.849	0.584	0.377	0.318
R20	0.010	1.272	0.976	0.544	0.286	0.245

FIG 3.14 Sporulation assay of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 over a five day period

The development of heat resistant spores of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 was measured over five days. Sporulation cultures were incubated anaerobically at 37 °C in AIM broth and at 24-h intervals the colony forming units per millimetre (cfu/ml) of **(A)** untreated and **(B)** heat-treated samples were determined on AIMTA agar medium (AIM supplemented with 0.1 % sodium taurocholate) after 48-h anaerobic incubation period. The **(C)** growth of the same samples was also measured over five days at 24-h intervals. The graphs and accompanying tables represent the averages of three experiments and the error bars indicate the standard error of the means. At the end of the experiment all strains tested were confirmed for their genotype, colonial morphology and against contamination.

In addition, other sporulation measurement experiments revealed an over-sporulating phenotype at 48-h for the merodiploid pES271/630Δerm/No.11 strain and derivatives, including Δ*addBA* 24 and Δ*addBA* 242 and resolved wt strains R20 and R215 (generated from resolution of pES271/630Δ*erm*/No.11 from the second mating) compared to wt 630Δ*erm* and other pES271/630Δ*erm* merodiploids (**FIG 3.15**). Therefore, in subsequent experiments, particularly those measuring UV and metronidazole sensitivity and SOS response, cultures were maintained in exponential growth with the presence of a minimum number of spores. As a control, resolved wt R20 was also tested in all experiments.

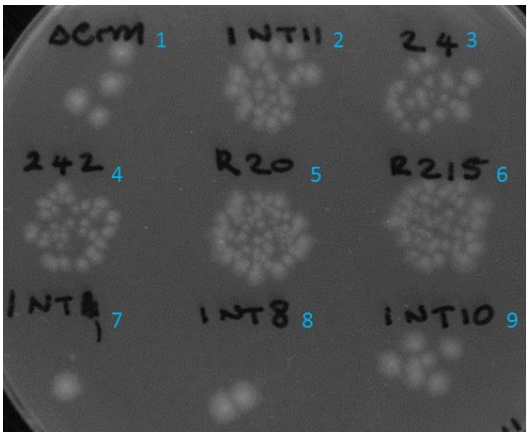
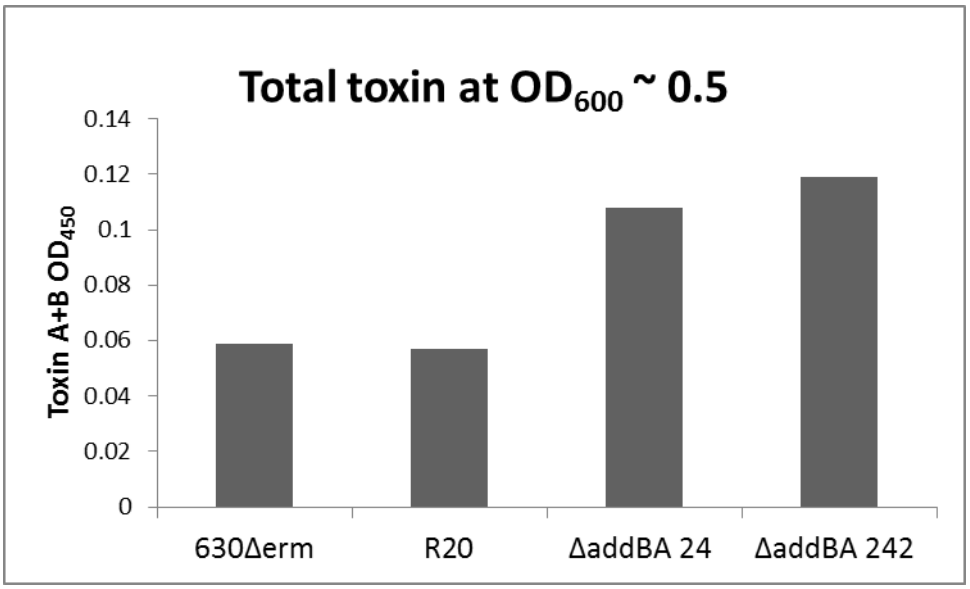


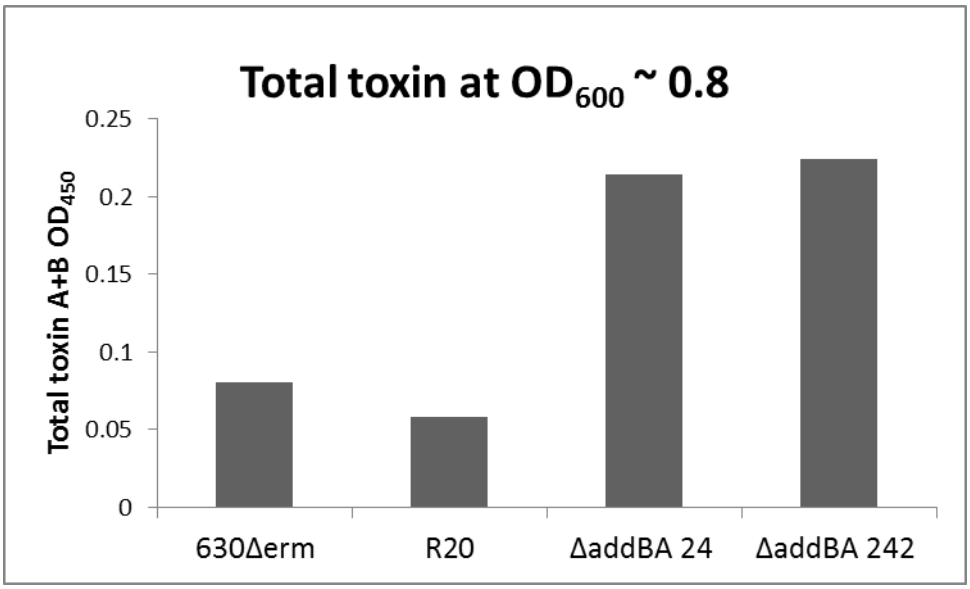
FIG 3.15 Sporulation of merodiploid pES271/630Δerm/No.11 and derivatives

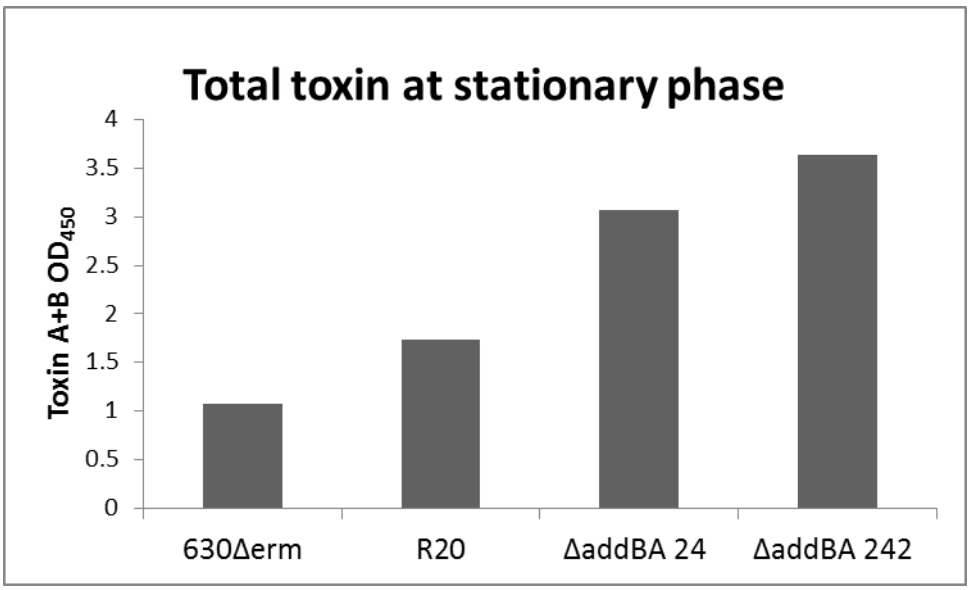
Sporulation cultures were started by inoculation of 10 ml AIM broths with a single colony of each strain. At 48-h incubation period, 0.5 ml of each culture were incubated at 65 °C for 25 min, diluted in 1 x PBS and then 10 µl of each was spotted onto AIMTA agar medium. The image shows heat resistant colonies of each strain after 48-h incubation period. Strains *C. difficile* 630Δ*erm* (1), merodiploid pES271/630Δerm/No.1 (7), merodiploid pES271/630Δerm/No.8 (8) and merodiploid pES271/630Δerm/No.10 (9) have similar sporulation phenotype. In contrast, merodiploid pES271/630Δerm/No.11 (2) and derivatives including; Δ*addBA* 24 (3), Δ*addBA* 242 (4) and resolved wt strains R20 (5) and R215 (6) overproduce heat resistant spores at 48-h compared to the rest of the strains.

Toxin measurement:

Total toxin (A+B) production was measured from cultures of an optical density OD_{600} of ~0.5, ~0.8 and ~1 (**FIG 3.16**). Both *addBA* deletion strains showed a 2-3- fold higher toxin production compared to wt strains taken from different growth stages. Total toxin (A+B) was measured once in these strains and further measurements are required to confirm the result.







SAMPLE	OPTICAL DENSITY (OD_{600})	TOXIN A+B (OD_{450})
630Δ*erm*	0.593	0.059
R20	0.584	0.057
Δ*addBA* 24	0.54	0.108
Δ*addBA* 242	0.585	0.119
630Δ*erm*	0.877	0.08
R20	0.855	0.058
Δ*addBA* 24	0.773	0.214
Δ*addBA* 242	0.804	0.224
630Δ*erm*	1.452	1.071
R20	1.368	1.735
Δ*addBA* 24	1.008	3.07
Δ*addBA* 242	1.104	3.633

Total toxin A+B (OD_{450}) control values:

positive control: 1.683

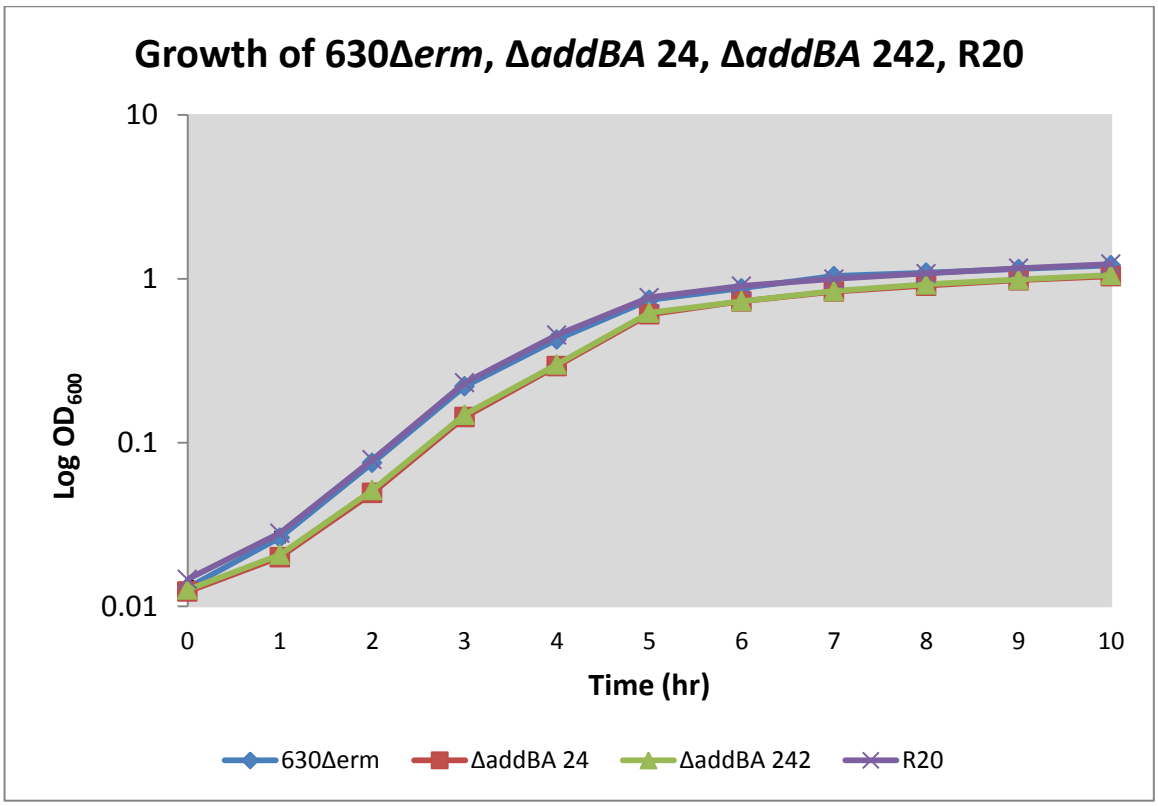
negative control: 0.047

FIG 3.16 Total toxin production in *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20

Total toxin production (A+B) in *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 was measured from culture supernatants at optical densities OD_{600} of around 0.5, 0.6 and 1. The bars and accompanying table indicate the result of one experiment. In this experiment the *addBA* deletion strains yielded 2-3 fold higher combined toxin titre compared to both wt strains.

Growth measurement and colony morphology:

It was predicted that deletion of *addBA* genes in *C. difficile* would result in a growth defect (i.e. reduced growth rate and small colony morphology) due to the requirement to repair DSBs that occur during normal cell growth (Dillingham and Kowalczykowski, 2008; Kuzminov, 1999). The growth of 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and resolved wt R20 was measured in triplicate for 10-h at 1-h intervals (**FIG 3.17**). All strains had similar growth rates with doubling times of 40 min for both wt strains and 45 min for both Δ*addBA* strains. It was also noted that the wt strains had shorter lag phases in comparison to the deletion strains. From the average optical density data obtained, it was shown that the wt strains reached an OD_{600} of around 1 in 7-h, while the Δ*addBA* strains reached this density in 10-h of incubation. Additionally, all strains showed a similar colony size and morphology on AIM agar medium.



Growth OD_{600}

	0 hr	1 hr	2 hr	3 hr	4 hr	5 hr	6 hr	7 hr	8 hr	9 hr	10 hr
630Δ*erm*	0.013	0.026	0.075	0.22	0.426	0.744	0.877	1.041	1.09	1.148	1.208
Δ*addBA* 24	0.012	0.02	0.049	0.143	0.293	0.607	0.728	0.835	0.908	0.982	1.037
Δ*addBA* 242	0.012	0.02	0.051	0.148	0.3	0.622	0.731	0.842	0.924	0.989	1.057
R20	0.014	0.028	0.078	0.232	0.454	0.769	0.905	0.998	1.078	1.16	1.231

FIG 3.17 Growth curves of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20

The growth of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 was measured over 10 hours at 1-h intervals. Combined growth curves of all strains with accompanying table are shown. All strains had similar doubling time of 40 min and 45 min for wt and *addBA* deletion strains, respectively. The graph and accompanying table represent the averages of three experiments and the error bars indicate the standard error of the means.

Sensitivity to DNA damaging agents:

UV sensitivity:

Previous studies have shown that DNA damage caused by UV irradiation can eventually lead to double-strand breaks as a result of attempts to repair UV-induced photoproducts via nucleotide excision repair (Bonura and Smith, 1975; Wang and Smith, 1983). It was therefore hypothesised that UV irradiation would greatly affect the viability of *C. difficile* Δ*addBA* strains due to lethality from accumulated unrepaired DSBs. To test this, exponential cultures of similar optical density (OD_{600}=~0.3) of wt 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and wt R20 were diluted and spotted onto AIM agar followed by irradiation with UV at 10, 15 or 20 J/m^2. To compare survival from UV irradiation, non-UV irradiated cultures (No UV) were carried through the experiment. This experiment was performed in triplicate and a representative example is shown (**FIG 3.18**). In the absence of UV irradiation all strains showed similar viability. However, there was a 3 $\log_{10}$ to more than a 4 $\log_{10}$ reduction in cell survival of the UV irradiated Δ*addBA* strains (3 $\log_{10}$ at 10 J/m^2 and ≥ 4 $\log_{10}$ at 15 and 20 J/m^2). UV irradiation had very little effect on the survival of the wt strains (1 $\log_{10}$ at 20 J/m^2).

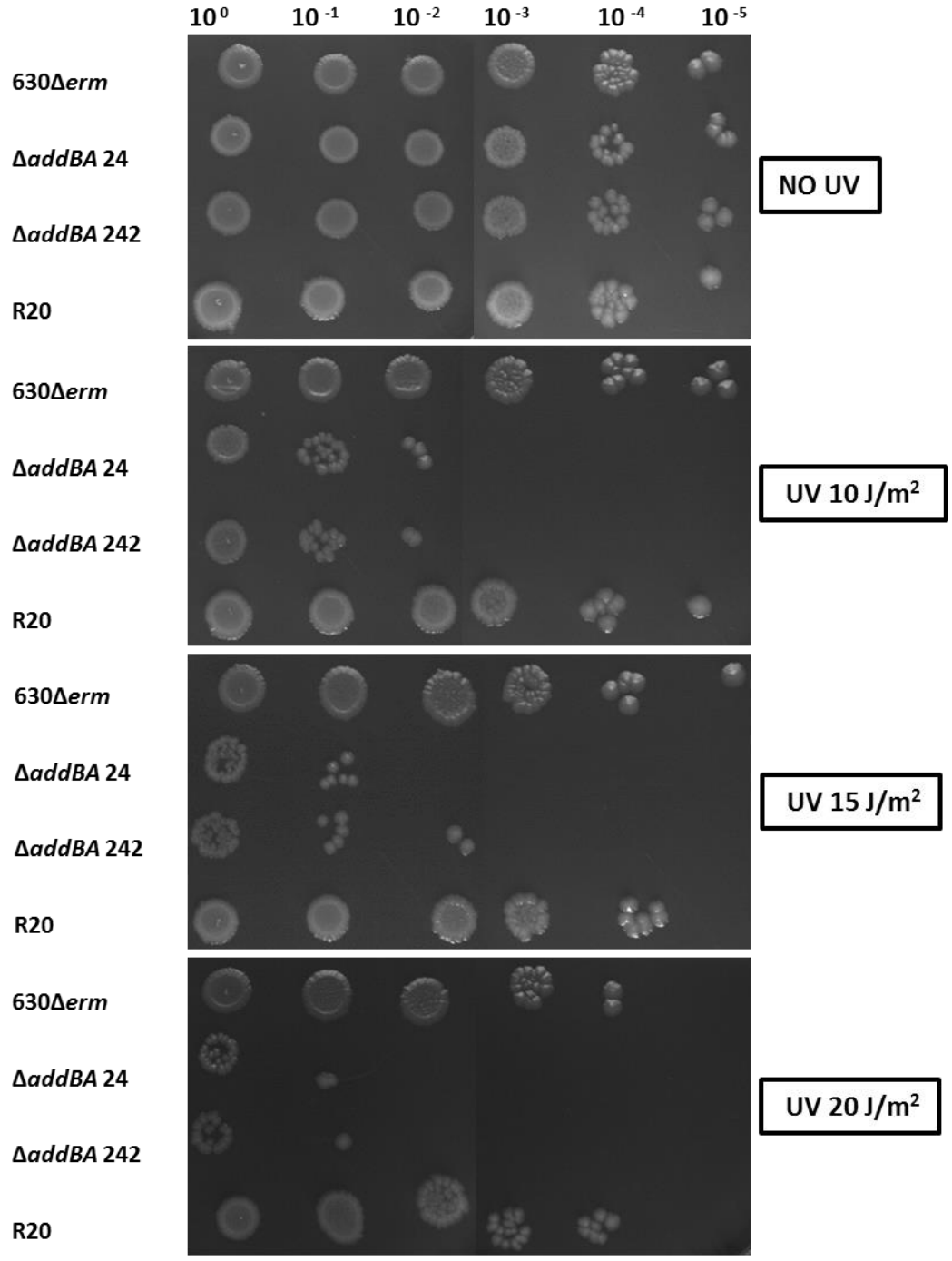


FIG 3.18 UV sensitivity assay of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20

Exponential cultures (OD_{600} of ~0.3) of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 were serially diluted and 5 µl of each was spotted onto AIM agar medium. Replica plates were made which were either irradiated with UV at 10, 15 or 20 J/m^2 or not exposed to UV irradiation. The results were taken after a 48-h incubation period. Representative photos are shown from a triplicate experiment. UV irradiated plates show that the *addBA* deletion strains are more sensitive to UV in comparison to wt strains as all strains show similar viability on non-irradiated replica plates.

Metronidazole sensitivity:

The mode of action of metronidazole is not fully understood but it is believed to cause chromosomal fragmentation including single- and double-strand breaks (Dachs *et al.*, 1995; Lofmark *et al.*, 2010; Sisson *et al.*, 2000; Steffens *et al.*, 2010). It was therefore expected that deletion of *addBA* genes in *C. difficile* would lead to enhanced sensitivity to metronidazole. To test this, exponential cultures of similar optical density (OD_{600}=~0.3)of wt 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and wt R20 were spotted onto AIM agar containing either no metronidazole, or different concentrations of metronidazole including 0.25 µg/ml, 0.125 µg/ml and 0.06 µg/ml. This experiment was performed in triplicate and a representative example is shown (**FIG 3.19**). Deletion of *addBA* genes in *C. difficile* resulted in increased sensitivity to all metronidazole concentrations tested, compared to the wt strains. Survival of the Δ*addBA* strains was reduced by 4 $\log_{10}$ at MZ 0.06 µg/ml, by 5 $\log_{10}$ at MZ 0.125 µg/ml and resulted to total death at MZ 0.25 µg/ml, compared to the survival of the population without exposure to metronidazole. In contrast, there was no effect on the survival of the wt strains at MZ 0.06 µg/ml, there was a 3 $\log_{10}$ survival reduction at MZ 0.125 µg/ml and more than 4 $\log_{10}$ reduction at MZ 0.25 µg/ml, compared to the survival of the population without metronidazole.

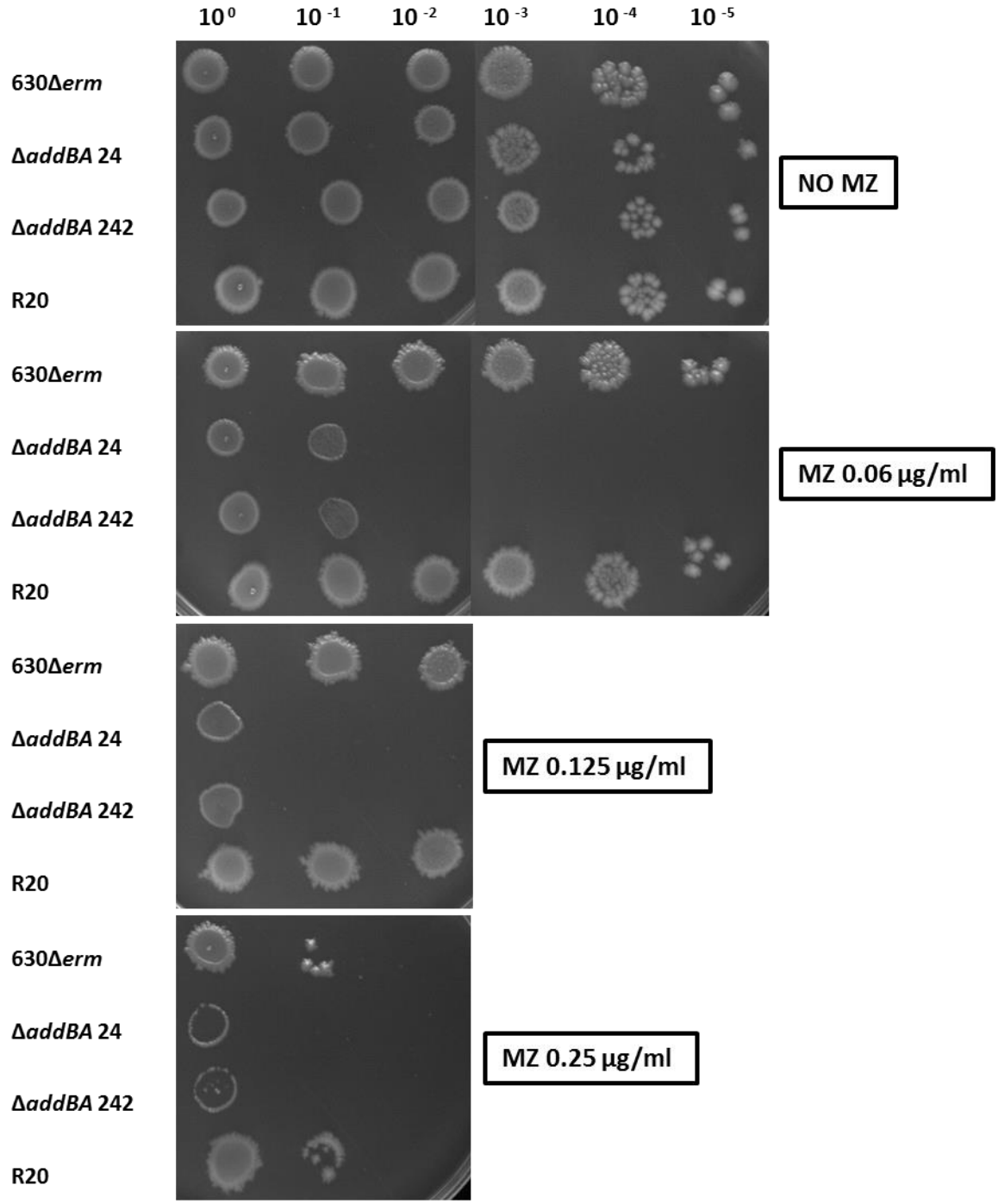


FIG 3.19 Metronidazole sensitivity assay of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20

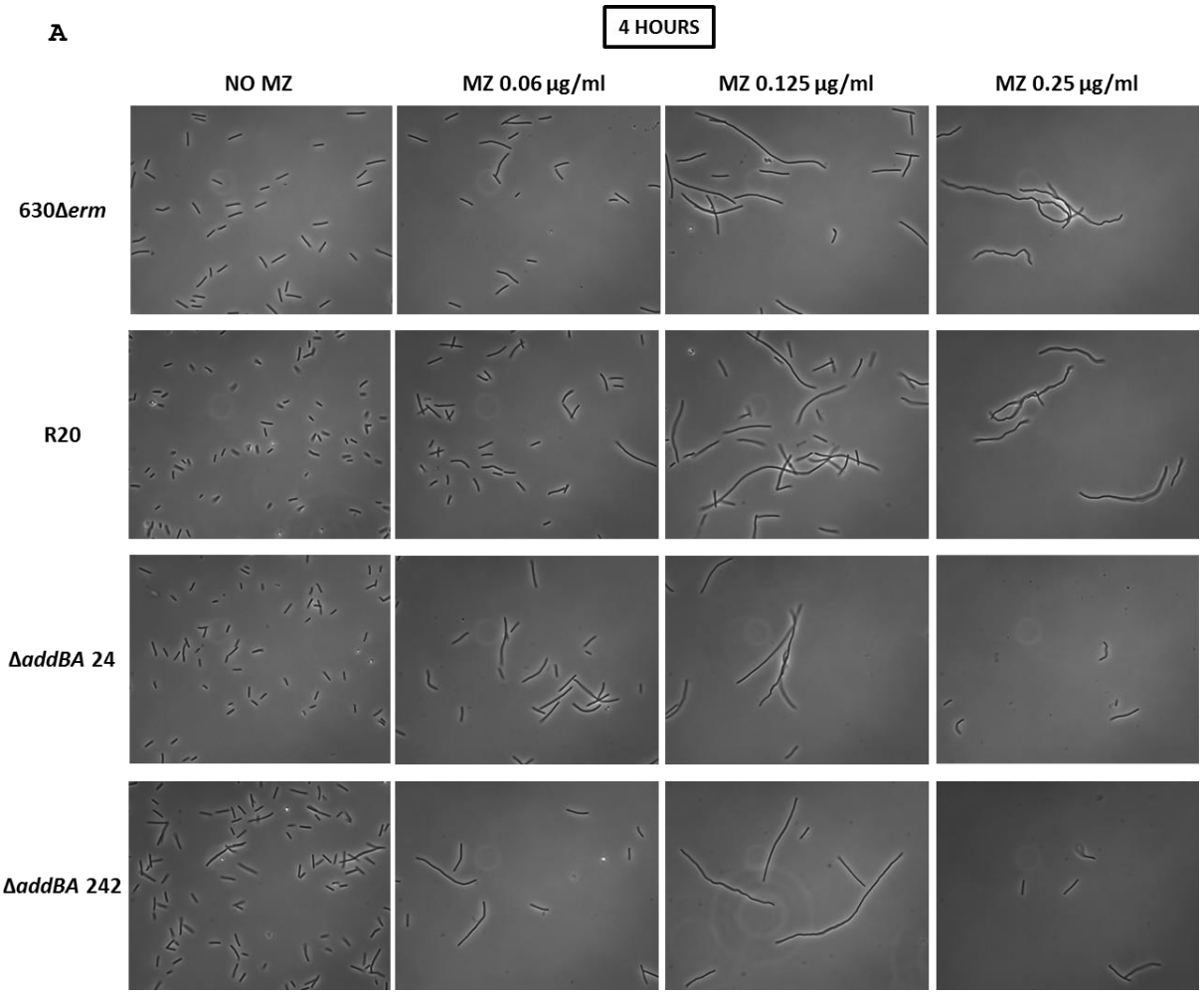
Exponential cultures (OD_{600} of ~0.3) of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 were serially diluted and 5 µl of each was spotted onto AIM agar medium containing varying concentrations of metronidazole including 0.06 µg/ml, 0.125 µg/ml, 0.25 µg/ml or no metronidazole. The results were taken after a 48-h incubation period. Representative photos are shown from a triplicate experiment. The *addBA* deletion strains were more sensitive to metronidazole compared to the wt strains.

Tests to complement the sensitivity phenotype of Δ*addBA* strains to DNA damaging agents were not possible, as multiple attempts to construct a complementation vector were unsuccessful, possibly due to the size of the required insert.

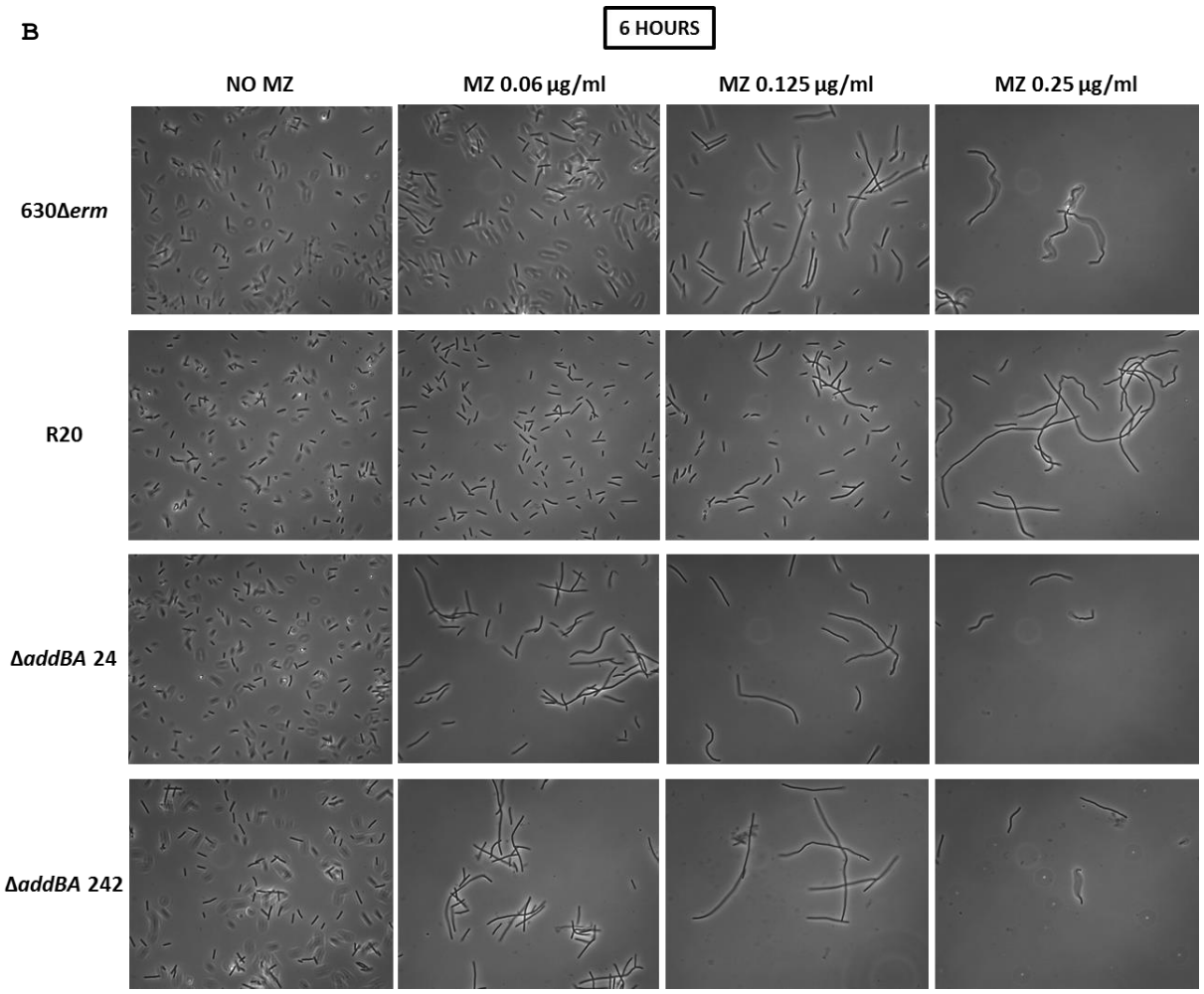
SOS response and *addBA* genes in *C. difficile*:

In most bacteria, an SOS response is induced when cells are treated with DNA damaging agents (Fernandez De Henestrosa *et al.*, 2000; O'Reilly and Kreuzer, 2004). To explore whether *C. difficile* has an SOS response system and whether *addBA* genes are involved, cells were treated with metronidazole. A filamenting phenotype is indicative of an SOS response considering that during an SOS response cell division is inhibited to allow cells to repair DNA damage (Dajkovic *et al.*, 2008; Hill *et al.*, 1997; Huisman *et al.*, 1984). In addition, in *E. coli* RecBCD is required for SOS induction because it is responsible for the formation of the RecA/ssDNA filament and subsequent autodigestion of LexA repressor allowing transcription of SOS genes (Janion, 2008; Michel, 2005). Taking this into consideration and in keeping with the results above, it was expected that in the presence of metronidazole the wt strains, if capable of SOS response, would filament while the *addBA* deletion strains would show a non-filamenting lethal phenotype. Exposure of exponential cultures of 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 to metronidazole resulted in the filamentation of cells compared to the cell morphology of non-exposed cultures (**FIG 3.20**). The wt strains showed long filaments at higher metronidazole concentrations of 0.25 µg/ml and 0.125 µg/ml while filamented cells were rare and shorter in length at 0.06 µg/ml. The *addBA* deletion strains showed long

filaments at metronidazole concentrations of 0.06 µg/ml and 0.125 µg/ml while the 0.25 µg/ml concentration was lethal, which is also indicated by the OD_{600} measurements taken at 24-h (T24 average No MZ OD_{600}= 1.548 & MZ 0.25 µg/ml OD_{600}= 0.006)which also correlates with the metronidazole spot testing (**FIG 3.19**). Also, filaments from the wt and deletion strains were more wavy shaped at MZ 0.025 µg/ml and 0.0125 µg/ml, respectively, compared to other concentrations suggesting a direct or indirect effect of MZ on peptidoglycan.



T4 OD_{600}	630Δ*erm*	Δ*addBA* 24	Δ*addBA* 242	R20
No MZ	0.414	0.351	0.358	0.443
MZ 0.25	0.024	0.013	0.013	0.04
MZ 0.125	0.146	0.049	0.05	0.168
MZ 0.06	0.282	0.109	0.104	0.32



T6 OD_{600}	630Δ*erm*	Δ*addBA* 24	Δ*addBA* 242	R20
No MZ	0.735	0.656	0.659	0.752
MZ 0.25	0.033	0.015	0.014	0.059
MZ 0.125	0.402	0.055	0.053	0.435
MZ 0.06	0.624	0.165	0.159	0.647

FIG 3.20 Morphology of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 cells in the presence of metronidazole at 4 and 6 hours of SOS induction

Exponential cultures (OD_{600} of ~0.3) of *C. difficile* strains 630Δ*erm*, Δ*addBA* 24, Δ*addBA* 242 and R20 were diluted at 1:100 into AIM broths without metronidazole or with metronidazole at a concentration of 0.06 µg/ml, 0.125 µg/ml or 0.25 µg/ml. The optical density was measured at inoculation time T0. At T4 **(A)** and T6 **(B)** hours of incubation, the optical density was measured and the morphology of fixed cells was visualised by wet mount microscopy. Representative photomicrographs are shown from triplicate experiments. The wt strains showed filamenting cell morphology at metronidazole concentrations of 0.125 µg/ml and 0.25 µg/ml with a wavy phenotype particularly at 0.25 µg/ml. The *addBA* deletion strains showed filamenting cell morphology at metronidazole concentrations of 0.06 µg/ml and 0.125 µg/ml with a wavy phenotype particularly at 0.125 µg/ml. Wet mount microscopy of the *addBA* deletion

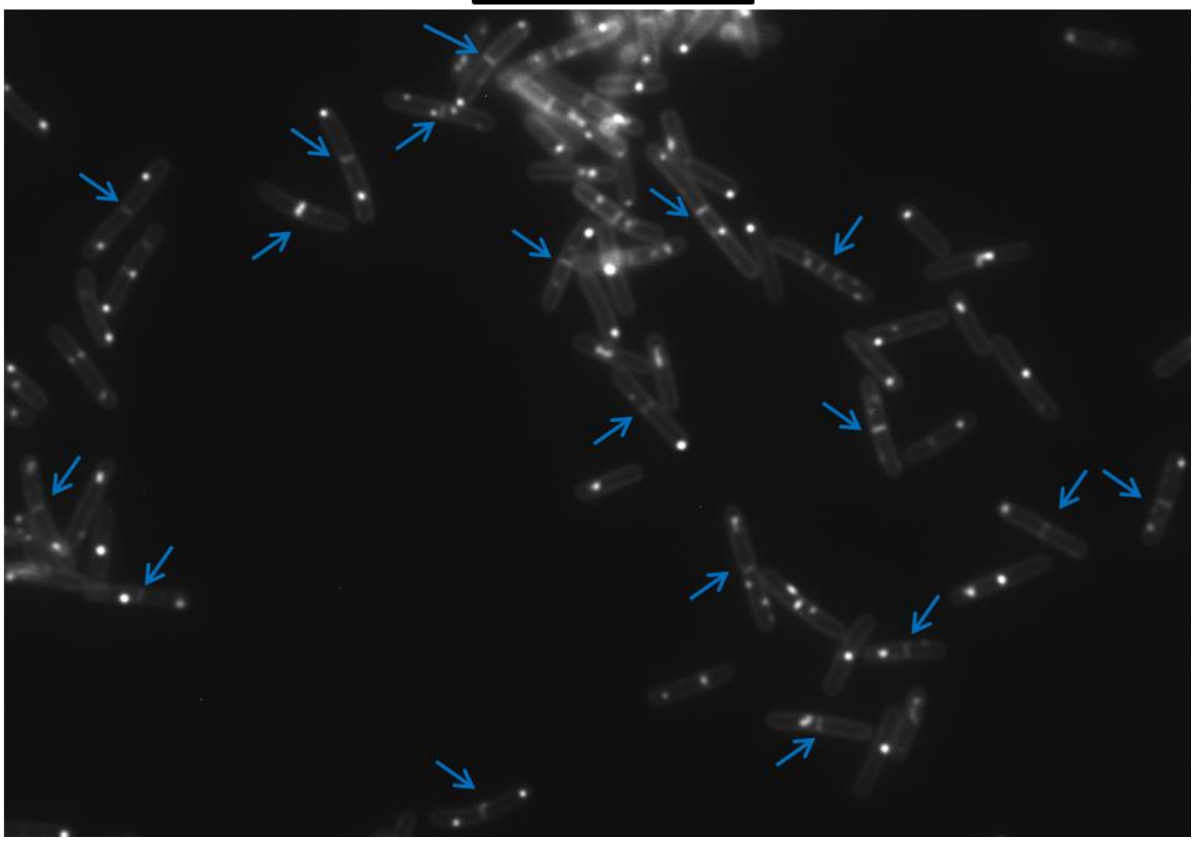
strains at 0.25 µg/ml showed fewer cells due to reduced survival which agreed with the optical density recorded in the tables above. Samples that had not been subjected to metronidazole showed normal morphology. All strains tested were confirmed for genotype, colonial morphology and against contamination.

To confirm that the filamenting cells of wt 630Δ*erm* had one division septum, samples from time point T6 (6-h after MZ addition) from cultures containing MZ 0.125 µg/ml or without MZ, were stained with the lipophilic membrane dye FM4-64. Fluorescence microscopy of FM4-64 stained cells showed the presence of one septum in induced and non-induced wt 630Δ*erm* cells but unexpectedly, it has also showed the presence of bright foci (**FIG 3.21**). In most cells that had not been exposed to metronidazole there were usually two bright foci, one on each side of the septum. Similarly, filamenting cells had either bright spots on each side of the septum or multiple less bright spots throughout the cytoplasm. However, due to the shape of the filamenting cells microscope focusing was not ideal and possibly the use of confocal microscopy would be more informative. The nature of these foci is not clear. They could simply be an artefact of the experimental procure used or alternatively, considering FM4-64 stains lipids, they could be lipophilic inclusions or granules.

A

630Δ*erm*

NO MZ – 6 HOURS



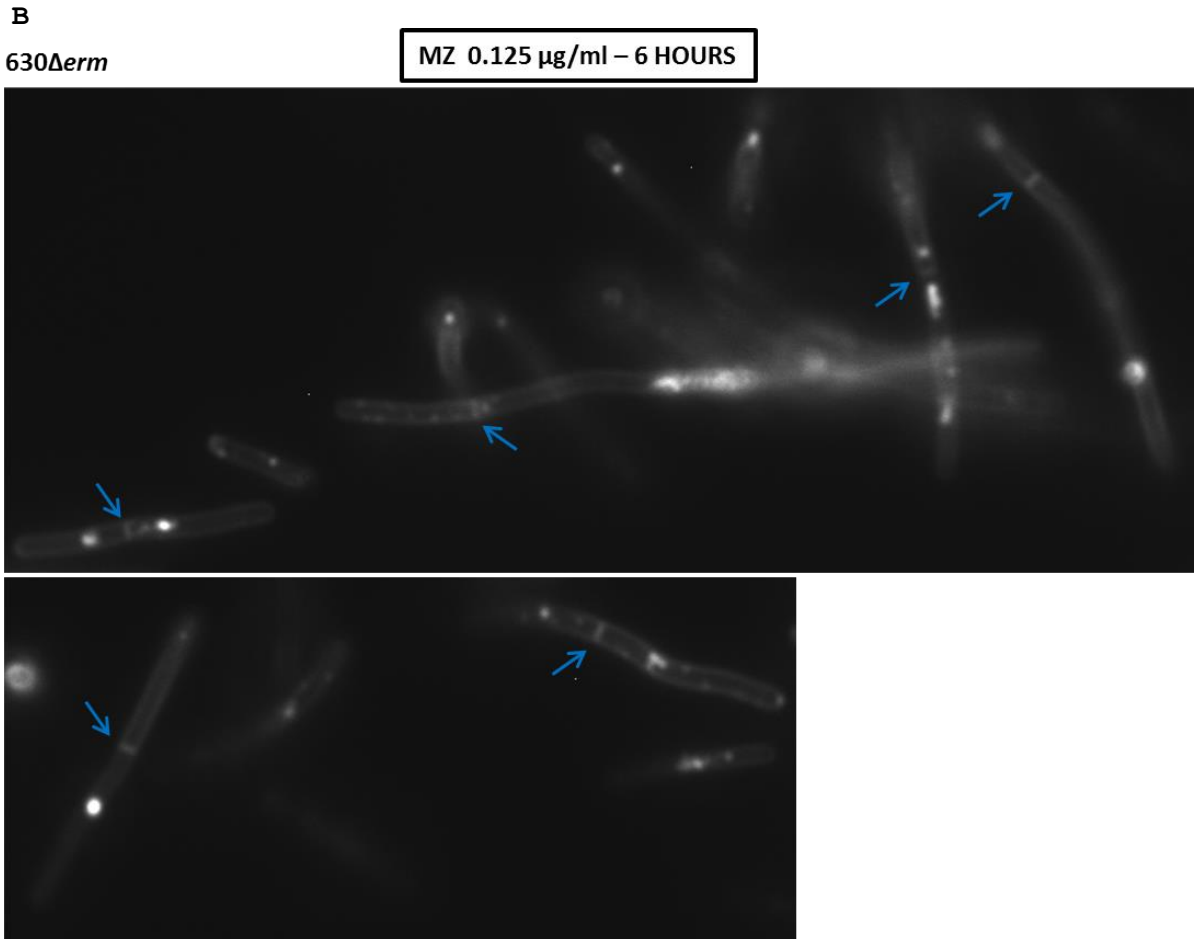


FIG 3.21 Septal ring in *C. difficile* 630Δ*erm* in the absence and presence of metronidazole

Representative photomicrograph from fluorescence microscopy of FM4-64 stained *C. difficile* 630Δ*erm* in the absence of metronidazole **(A)** and in the presence of metronidazole at 0.125 µg/ml **(B)** at T6 of the SOS induction experiment. Blue arrows point to the septal ring. Also, note the presence of bright foci that have been also stained with FM4-64.

3.3: DISCUSSION

Development of *C. difficile* markerless I-SceI genetic system:

C. difficile is known to be notoriously resistant to genetic manipulation. In addition, *C. difficile* is a multidrug resistant organism and particularly 630Δ*erm* strain which is the most widely used lab strain. This poses a limitation in the use of antibiotic resistance markers for the manipulation of this organism. Most developed methods (including the most widely used ClosTron system) of *C. difficile*, result in the interruption of the target gene via the integration of an antibiotic resistance marker or a whole plasmid carrying antibiotic resistance, which in addition to a possible polar effect on neighbouring genes, can prevent multiple manipulations of the genome. This study has described the development of a novel genetic system for *C. difficile* to create targeted, precise, markerless chromosomal deletions using the meganuclease I-SceI. For validation of this system, the *addBA* genes in *C. difficile* 630Δ*erm* were deleted. The method developed here, results in the deletion of target genes without any foreign DNA remaining in the host genome, including resistance genes. The system requires the use of two plasmids; one that carries the recognition site of I-SceI (I-SceI site delivery vector), into which the deletion cassette is inserted (allele replacement vector) for single cross-over integration, and one that expresses the I-SceI (I-SceI expression vector) under an appropriate promoter (ferredoxin promoter, P*fdx*) to stimulate a double cross-over via DSB repair. Here, the I-SceI recognition vectors pES185 or pES2761 carry a thiamphenicol resistance marker and the I-SceI expression vector pES288 constitutively

expresses I-SceI and carries an erythromycin resistance gene in its backbone. With the use of this system, the deletion of a target gene is identified by; 1) the loss of thiamphenicol resistance resulting from deletion of the integrated plasmid after DSB and resolution, 2) the PCR amplification of the deletion cassette (~1 kb) that replaces the wt allele and 3) by confirming the new sequence joints. In theory, the system could be used to delete genes in other *C. difficile* strains of interest, however, the genome sequence is required to aid the design of primers for the precise targeted chromosomal deletion and provided the strain is sensitive to erythromycin and thiamphenicol. In the case of the hypervirulent *C. difficile* R20291 (027) the erythromycin resistance marker *ermB* present in pES288 would confer resistance to lincomycin at 20 µg/ml (Heap *et al.*, 2009). It should be noted that, the transfer of plasmids into BI/NAP1/027 strains relying on RP4-mediated conjugation from *E. coli*, was reported to be unsuccessful, although, the reported efficiency of RP4-mediated plasmid transfer in *C. difficile* JIR8094 and CD37 strains is $2.3 \times 10^2 - 3 \times 10^3$ and $4 \times 10^1 - 2.4 \times 10^3$, respectively (Carter *et al.*, 2011; O'Connor *et al.*, 2006; Mullany *et al.*, 1990). In this study multiple attempts were required to transfer pES271 into *C. difficile* 630Δ*erm*, with only one transconjugant obtained. In addition, O' Connor *et al.* (2006), reported that 'some difficulty was experienced' in the introduction of pJIR2816 (similar to pJIR1456 from which pES271 was generated) into *C. difficile*.

With the genetic system constructed in this study, any gene can be deleted, irrespective of size. In contrast, a limitation of genetic systems based on the TargeTron technology might have, is that DNA sequences less that 400

bp may not contain target sites for mobile group II intron recognition (TargeTron user guide, Sigma-Aldrich).

The low resolution rate and the increasing resolution percentage with sub-streaking of co-integrates in the presence of pES288, suggested that a low I-SceI concentration was present in cells. The effect of I-SceI levels on resolution has been previously demonstrated by Posfai *et al*. (1999). These authors showed that in *E. coli* stimulation of intramolecular recombination by I-SceI is dependent on the concentration of I-SceI in the cell and how long I-SceI is allowed to exert its effect (Posfai *et al*., 1999). It was shown that for a 3.4 kb deletion, when high I-SceI levels were expressed in the cells, only 8-10% of potential transformants survived, but 100% resolution of the co-integrates was obtained. When a low I-SceI level was expressed, 100% of potential transformants survived but a mixed population of resolved and unresolved strains was obtained and colonies had to be restreaked to increase resolution (> 90%). When extremely low I-SceI levels were expressed in the cell, for the same deletion, growth of up to 60 generations was required to achieve resolution close to 100% (Posfai *et al*., 1999). If, in the current study, the constitutive expression of I-SceI is considered, it might be expected that a high concentration of I-SceI was present and a high resolution rate should be observed. However, the low resolution rate observed may suggest inefficient translation of I-SceI or instability of messenger RNA or protein. Despite this, it was confirmed that the resolution observed was due to I-SceI expression and not due to spontaneous recombination. The efficiency of obtaining a Δ*addBA* was an average of 5.7%. This rare isolation of Δ*addBA* strains may have been attributed to the moderately large size (7.2 kb) of the targeted deletion,

where reversion to wt is favoured with the efficiency of obtaining a deletion mutant decreasing with increasing size of deleted region (Posfai *et al.*, 1999; Suzuki *et al.*, 2005).

Function of AddAB in *C. difficile*:

A further aspect of this study was to analyse the function of AddAB in *C. difficile* using the two isolated *addBA* deletion strains, the parental *C. difficile* 630Δ*erm* and one resolved wt strain R20. Here, the sporulation-frequency, toxin production and growth were measured. Sensitivity to DNA damaging agents was determined and the involvement of *addA* and *addB* genes in the SOS response in *C. difficile* was questioned.

Spore production:

The measurement of sporulation over a five day period showed that Δ*addBA* 24, Δ*addBA* 242 and wt R20 produce at least 100-fold more heat resistant spores at least 24-h earlier than the parental strain 630Δ*erm*. This was not due to a difference in growth rate or due to a difference in viability, and this characteristic exists only in derivative strains of merodiploid pES271/630Δ*erm*/No.11 where integration occurred via the upstream region of *addB* gene. An explanation for this phenomenon cannot be given at this point and perhaps chromosomal sequencing of these strains would shed light on any genomic alteration in these strains that provides them with such a phenotype. However, deletion of *addB*A did not affect sporulation since reduced numbers were not observed in the mutant strains when compared to wt R20, as one would expect for two reasons.

First, wet heat does not kill spores by DNA damage but instead it can damage spore proteins (Coleman *et al.*, 2007; Zhang *et al.*, 2010; Setlow, 2007). Secondly, the fact that spores have a single chromosome means that HR would not be used as the DSB repair pathway at spore germination, since this pathway requires a second homologous chromosome (Moeller *et al.*, 2007). In *B. subtilis*, NHEJ is responsible for the DSB repair during spore germination (Moeller *et al.*, 2007). Bioinformatic analysis of the *C. difficile* genome did not reveal a Ku homologue. In addition, it was noted from the heat-treated data that both *addBA* deletion strains displayed a 10-fold lower spore number compared to wt R20 at time 0-h of the sporulation assay. Considering that the amount of cells used to inoculate the sporulation culture at 0-h (according to OD_{600} measurement) and that viable CFU were similar between the strains, it is therefore suggested that either DNA damage inhibits and delays sporulation or that wt R20 simply sporulates more rapidly, even though this is not observed from 24-h onwards. Multiple studies have shown that DNA damage and disruption of DNA replication can delay sporulation, which might explain the presence of 10-fold reduced spore numbers of exponential cultures of *addBA* deletion strains, where an intact replicated chromosome may not be present to initiate sporulation (Bejerano-Sagie *et al.*, 2006; Michael, 2001; Lemon *et al.*, 2000; Burkholder *et al.*, 2001; Ireton and Grossman, 1992; Ireton *et al.*, 1993).

Toxin production:

Preliminary data for toxin measurement showed that both Δ*addBA* strains produced 2-3 times more total toxin (A+B) compared to the wt strains at all cell growth stages. There

are no previous reports of increased toxin production due to loss of *addAB* genes in other organisms. Evidence suggests an association between toxin production and sporulation in *C. difficile* (Underwood *et al.*, 2009; Karlsson *et al.*, 2008; Akerlund *et al.*, 2006; Kamiya *et al.*, 1992). Although, the resolved wt R20 strain and both deletion mutant strains have an altered sporulation phenotype, the wt R20 does not produce more total toxin compared to the parental strain. It is unclear thus, at the moment, why increased toxin production in the mutant strains is observed. However, these tests need to be repeated.

Growth measurement and Viability:

Deletion of *addB*A in *C. difficile* had little effect on the exponential phase of bacterial growth and there was not a defect in colony formation. This was in contrast to our expectation, taking into account the importance of RecBCD/AddAB complexes during normal cell growth (e.g. DNA replication, replication fork stalling/collapse etc.). Previous studies have shown reduced growth rate and impaired colony growth, for example, compared to wt, *E. coli recB*$^{-}$ *recC*$^{-}$ had generation time increased by 16 min, *rexA* and *rexB* mutants of *S. pneumoniae* were severely impaired for growth and *B. fragilis* Δ*addAB* had increased doubling time by 31 min and impaired colony growth (Capaldo *et al.*, 1974; Halpern *et al.*, 2004; Parry, 2010). It may be suggested that in *C. difficile* replication does not stop as frequently during exponential growth. It was noted that the mutant strains displayed a longer lag time. This observation may suggest a requirement for AddAB for DNA maintenance during stationary phase, as recently

demonstrated in *E. coli* for the RecBCD counterpart, and this may be reflected in the starting inoculum population, or it may be accumulated DNA damage that has to be repaired before exponential growth commences (Williams *et al.*, 2011).

In the absence of exogenous DNA damaging agent (e.g. UV or MZ), both *addBA* deletion strains showed similar viability to the wt strains. This also agreed with the viability measurement over 120-h period during the sporulation assay. The viability of the *addBA* mutant strains observed here is in contrast to multiple studies that report reduced viability of *addAB*/*recBCD* mutants in other species. Reduced survival in these studies may mostly reflect cells where DSBs occurred from disrupted replication forks that were not repaired and did not complete replication. For example, viability of *E. coli recB recC* mutant is reduced by 73%, *S. pneumoniae rexA* and *rexB* mutants by 80%, *B. subtilis addB* mutant by 50% and *B. fragilis* Δ*addAB* has 1-2 $\log_{10}$ reduction in viable cells (Dillingham and Kowalczykowski, 2008; Capaldo *et al.*, 1974; Halpern *et al.*, 2004; Sciochetti *et al.*, 2001; Parry, 2010).

The fact that the deletion of *addBA* in *C. difficile* has not significantly affected the growth and viability of *C. difficile* suggests that, in contrast to the *E. coli* paradigm, the replication fork does not stop as frequently or if a replication fork stops, a regressed fork is not formed or is processed by an enzyme other than AddAB.

Sensitivity to DNA damaging agents:

The deletion of *addBA* in *C. difficile* resulted in reduced survival when exposed to agents known to produce DSBs

indirectly or directly, such as UV radiation or MZ (Bonura and Smith, 1975; Dachs *et al.*, 1995; Steffens *et al.*, 2010). This confirmed our expectation that in *C. difficile* the AddAB complex is required for the repair of DSB produced by these DNA damaging agents. Similarly, for example, *E. coli* Δ*recBCD*, *Pseudomonas syringae* Δ*recCBD*, *Bacteroides fragilis* Δ*addAB* and *Streptococcus pneumoniae rexAB* mutants are sensitive to UV irradiation (Reuter *et al.*, 2010; Halpern *et al.*, 2004; Pavankumar *et al.*, 2010). Moreover, *B. fragilis* Δ*addAB* shows increased sensitivity to MZ compared to wt (Parry, 2010). Other examples of *recBCD* or *addAB* mutants with increased sensitivity to antibiotics known to lead to DSBs include *H. pylori addB* and *addA* mutants with increased sensitivity to mitocycin C and ciprofloxacin, *E. coli recD* mutant and *P. syringae* Δ*recCBD* strains with mitomycin C sensitivity (Amundsen *et al.*, 2000; Amundsen *et al.*, 2008; Pavankumar *et al.*, 2010).

SOS response:

The presence of an SOS system in *C. difficile* had not previously been reported in the literature, so the involvement of the AddAB complex was investigated by the exposure of cells to metronidazole. During the SOS response cell division is inhibited, for example by SfiA in *E. coli* and by YneA in *B. subtilis*, resulting in a filamenting cell morphology reflecting the continuation of cell growth in the absence of cell division (Mo and Burkholder, 2010; Huisman *et al.*, 1984). Therefore, the filamenting morphology in the presence of metronidazole indicated an SOS induction in all strains including the *addBA* deletion strains. FM4-64 staining confirmed that these were indeed elongated cells with one septum. These data lead to the

conclusion that AddAB is not necessary for induction of the SOS response in *C. difficile* and another exonuclease is generating ssDNA. In *B. subtilis* the SOS response is strongly reduced in *addAB* mutant cells and not abolished as in *recA* cells (Kidane *et al.*, 2004; Sanchez *et al.*, 2006). In this organism two end processing avenues have been shown to operate, the AddAB- and the RecJ-dependant, both equally contributing to the generation of 3'-tailed duplex DNA required for RecA filament formation (Sanchez *et al.*, 2006; Kidane and Graumann, 2005). The deletion of *recJ* in *B. subtilis addAB* mutant cells resulted in a strong synergistic effect where the double mutants were as sensitive to DSBs as *recA* mutant cells (Sanchez *et al.*, 2006).

To summarise, this study has demonstrated the development of a new genetic system for the construction of markerless deletions in *C. difficile*, using *addBA* as an example. The use of the I-sceI meganuclease in the system increases intramolecular recombination and without this protein resolution is not observed. The next step in development involved the construction of different deletions and a strain with multiple deletions to prove reproducibility and robustness of the genetic system. In the *C. difficile* Δ*addBA* mutants generated here, growth was reduced slightly, spore formation was unaffected and viability was not impaired in the absence of exogenous DNA damage. However, when DNA damage was imposed e.g. MZ, cell survival was reduced. This may lead to a new avenue for therapeutics to increase effectiveness of treatment i.e. a combination drug that causes DNA damage, particularly DSB, together with inhibition of AddAB.

CHAPTER 4: Optimization of the I-SceI genetic system for the deletion of *fliC*, *pilT* and *pilA* genes and investigation of the type IV pilus in *C. difficile*

4.1: INTRODUCTION

Aims and Objectives:

- To construct a non-flagellated *C. difficile* strain, impaired in flagella-mediated motility.
- To delete the type IV pilus-associated genes *pilT* and *pilA* in *C. difficile*.
- To optimise the genetic system for the construction of the *fliC*, *pilT* and *pilA* deletions.
- To characterise the deletion strains.
- To investigate the intergenic region upstream of type IV pilus-associated genes in *C. difficile*.

Background:

Type IV pili (TFP) are unique appendages on the surface of non-pathogenic and pathogenic bacteria and are peritrichus or polarly located (Jarrell, 2009). They are extremely flexible, strong, thin (<8 nm in diameter) and ranging between 1000–2500 nm in length (Strom and Lory, 1993; Craig and Li, 2008). These surface structures are widespread across β-, γ- and δ- proteobacteria and cyanobacteria (Nudleman and Kaiser, 2004; Proft and Baker, 2009; Bhaya *et al.*, 2000). They have been mostly found in Gram-negative

bacteria including enteropathogenic *E. coli* (EPEC), *Salmonella enterica*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Neisseria gonorrhoeae*, *Neisseria meningitidis* and *Vibrio cholerae* (Proft and Baker, 2009). TFP have also been recently found in Gram-positive bacteria particularly *Clostridium perfringens* and *Ruminococcus albus* (Varga *et al.*, 2006; Rakotoarivonina *et al.*, 2002).

TFP structure and function have been extensively studied in *Neisseria* and *Pseudomonas* species (particularly in *P. aeruginosa*) and in the soil bacterium *Myxococcus xanthus* (Burrows, 2012; Mignot and Kirby, 2008). **FIG 4.1** shows a model of the TFP apparatus. The major structural subunit of TFP is a small (15-20 kDa) pilin protein called PilA in *P. aeruginosa* and PilE in *Neisseria* (Burrows, 2005; Proft and Baker, 2009). Pilins are arranged in a helical fashion but the number of subunits per turn has not been established (Burrows, 2005; Craig *et al.*, 2004). The pilin subunits are extremely variable in sequence and length, but are conserved in the amino acid sequence in the N-terminal segment (Pelicic, 2008). Pilins are initially synthesized as prepilins, the leader sequence is cleaved and the N-terminal amino acid created is methylated by a prepilin peptidase (PilD) resulting in a mature protein (Mattick, 2002; Nudleman and Kaiser, 2004; Lory and Strom, 1997; Strom and Lory, 1993). Type IV pilins can be classified into two groups, Type IVa and Type IVb, based on the lengths of the leader peptide and the mature protein (Craig and Li, 2008; Pelicic, 2008; Proft and Baker, 2009; Burrows, 2012; Giltner *et al.*, 2012). The Type IVa pilins are short, have short leader peptides, are relatively homogeneous and are found in a variety of bacteria with broad host ranges. The Type IVb pilins are more diverse, they are either short or long, they possess long leader

peptides and are commonly found on enteric pathogens. *C. perfringens* PilA2 has an intermediate length of both mature sequence and leader peptide and does not fall clearly into either class (Jarrell, 2009). Type IVb pilins are further divided into tight adherence pili (Tad or Flp) and are found in a variety of Gram-positive and Gram-negative species (Giltner *et al.*, 2012; Burrows, 2012). Considering the absence of a narrow channel in the TFP fibre the pilus must therefore be assembled from its base (in contrast to flagella), and minor pilin-like proteins (in *P. aeruginosa* PilE, PilV, PilW, PilX, FimT and FimU) are involved in pilus assembly (Mattick, 2002; Nudleman and Kaiser, 2004; Burrows, 2005; Russell and Darzins, 1994; Alm and Mattick, 1995, 1996; Winther-Larsen *et al.*, 2005). Pilus assembly/extension (addition of pilin subunits to the base) and/or pilus disassembly/retraction (removal of pilin subunits from the base) is powered by ATP hydrolysis from two ATPases, PilB and PilT, which have opposing activities (Burrows, 2005; Mattick, 2002; Nudleman and Kaiser, 2004; Proft and Baker, 2009; Merz *et al.*, 2000; Maier *et al.*, 2002; Herdendorf *et al.*, 2002; Jakovljevic *et al.*, 2008). PilB is responsible for pilus extension and PilT is responsible for pilus retraction. ATP hydrolysis by these proteins results in a piston-like movement of a cytoplasmic membrane protein (possibly PilC) which leads to the addition or removal of pilin subunits during extension and retraction, respectively (Jarrell and McBride, 2008; Craig *et al.*, 2006). In Gram-negative organisms, the translocation of the pilus through the outer membrane is achieved via a large oligomeric pore made by the secretin PilQ which appears to be stabilized by the lipoprotein PilP (Jarrell, 2009; Mattick, 2002; Nudleman and Kaiser, 2004; Wall *et al.*, 1999; Collins *et al.*, 2001; Collins *et al.*, 2004; Drake and Koomey, 1995; Drake *et al.*, 1997; Bitter *et*

al., 1998). In Gram-positive bacteria, including the Clostridia, the secretin PilQ and the associated lipoprotein PilP are absent (Jarrell and McBride, 2008; Jarrell, 2009; Varga *et al.*, 2006). The tip of the pilus is thought to be involved in the attachment to surfaces, possibly via the primary pilin *pilA* and/or additional proteins such as PilC1 (*in N. gonorrhoeae*) and PilY1 (in *P. aeruginosa*) (Jarrell and McBride, 2008; Mattick, 2002; Proft and Baker, 2009; Skerker and Berg, 2001).

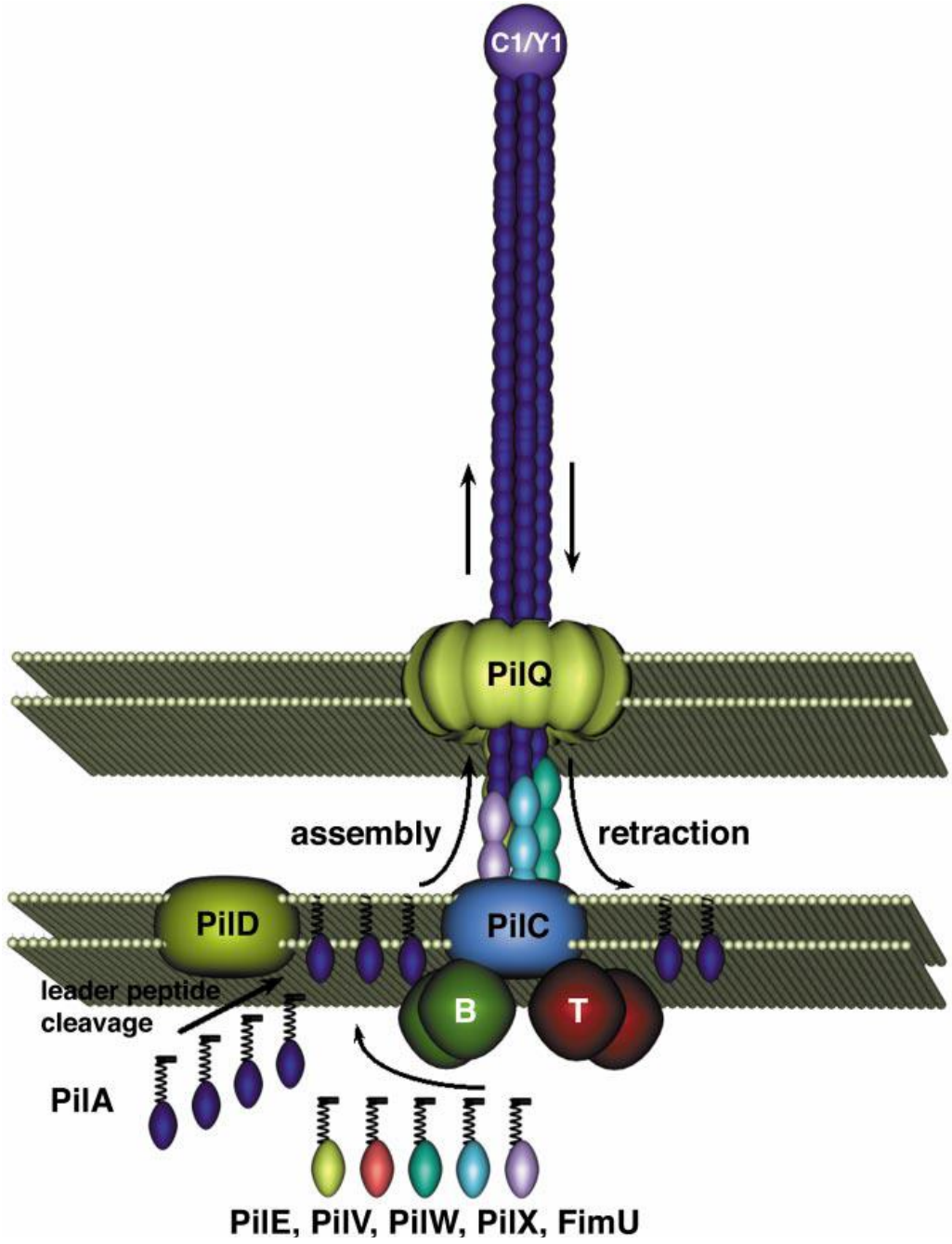


FIG 4.1 Model of the type IV pilus apparatus in a Gram-negative bacterium

Pilins (PilA) are synthesised as prepilins and the prepilin leader sequence is cleaved and the N-terminal amino acid created is methylated by a prepilin peptidase (PilD) resulting in a mature protein. Pilus is assembled from its base and minor pilin-like proteins (*in P. aeruginosa* PilE, PilV, PilW, PilX, FimT and FimU) are involved in this. PilB is responsible for pilus extension and PilT is responsible of pilus retraction and ATP hydrolysis from these proteins results in a piston-like movement of a cytoplasmic membrane protein (possibly PilC) which leads to the addition or removal of pilin subunits during extension and retraction, respectively. In Gram-negative bacteria the pilus is extruded through the outer membrane via PilQ, stabilized by PilP and both proteins are absent in Gram-positive bacteria. Attachment of the

pilus to surfaces is aided through the tip. The image is adapted from (Mattick, 2002).

TFP are involved in a wide range of functions in bacteria that are important to their survival and to their success as pathogens. These functions include adherence to biological (including other bacteria) and non-biological surfaces (including plastic, metal, glass etc.), DNA uptake, bacteriophage binding, cell aggregation, microcolony formation, biofilm production, fruiting body formation, host cell invasion, secretion of exoproteins and flagella-independent surface translocation (which is unique to TFP and has been called twitching motility) (Shi and Sun, 2002; Nudleman and Kaiser, 2004; Burrows, 2005; Craig and Li, 2008; Jarrell and McBride, 2008; Proft and Baker, 2009; Filloux, 2010; Burrows, 2012; Giltner *et al.*, 2012; Henrichsen, 1972). A lot of these functions are attributed to the retraction of the pilus in which PilT is involved (Burrows, 2012). Considering the importance of TFP as virulence factors and their surface location, these structures are important targets of potential vaccine development (Proft and Baker, 2009).

Multiple TFP functions exemplify their importance in virulence and pathogenesis, some of which are described here. *N. gonorrhoeae* and *N. meningitidis* initiate colonisation of human mucosal surfaces by binding of their TFP, named GC/MC pili, to non-ciliated host cells and possibly the receptor for both pili is the human membrane cofactor protein CD46 (Swanson, 1973; Kallstrom *et al.*, 1997; Proft and Baker, 2009). *N. meningitidis* PilC1 and PilC2 adhere to the uropod of polymorphonuclear neutrophils, thus escaping phagocytosis and increasing their ability to spread at the epithelial cell layer

(Soderholm *et al.*, 2011). *N. meningitidis* PilT is required for intimate attachment to epithelial cells and for attaching and effacing lesions (Pujol *et al.*, 1999). In *N. meningitidis* the minor pilin PilX is essential for bacterial aggregation (Helaine *et al.*, 2005). In *N. gonorrhoeae* TFP retraction in a PilT-dependent manner, can enhance activation of the nuclear factor NF-kB in infected cells and gonococcal microcolony formation (Dietrich *et al.*, 2011). In *P. aeruginosa*, PilA binds to glycosphingolipid asialo-GM1 on epithelial cells and the importance of pili is demonstrated by the inability of non-piliated strains to bind to human A549 type II pneumonocytes (Lee *et al.*, 1994; Chi *et al.*, 1991; Proft and Baker, 2009). *P. aeruginosa* PilT and PilU (homologue of PilT) are required for epithelial cell cytotoxicity and for full virulence in the mouse model of acute pneumonia (Comolli *et al.*, 1999). Moreover, *P. aeruginosa pilT* and *pilU* mutants are hyperpiliated, defective in twitching motility and impaired in the corneal infection model, while the *pilA* mutant is non-piliated and impaired in its ability to invade corneal epithelial cells (Zolfaghar *et al.*, 2003). Twitching-mediated biofilm structure and development is also affected in *P. aeruginosa pilT* and *pilU* mutants, while *pilB* and *pilC* mutants do not form microcolonies due to loss of twitching-mediated cell aggregation (Chiang and Burrows, 2003; O'Toole and Kolter, 1998). In *V. cholerae*, the type IV toxin co-regulated pilus (TCP) is required for success of the pathogen during infection and TCP has multiple functions including: attachment to Caco-2 human intestinal epithelial cells; secretion of colonization factor TcpF; microcolony formation by mediating bacterium-bacterium interactions; and protection from the antimicrobial component of bile by engulfing the bacterium during infection (Kirn *et al.*,

2003; Krebs and Taylor, 2011). In EPEC the TFP called the bundle-forming pilus (BFP) initiates adhesion with the intestinal epithelium and recruits EPEC cells into aggregates resulting in microcolonies. These bacterial aggregates can disaggregate and then disperse, which is facilitated by pilus retraction driven by BfpF (putative ATPase) (Bieber *et al.*, 1998; Zahavi *et al.*, 2011). BFP retraction is suggested to contribute to pathogenesis by bringing the bacterium to close proximity with the host cell, allowing prompt and efficient translocation of bacterial protein effectors through the type III secretion apparatus in the host cell, resulting in the disruption of the epithelial barrier and generation of actin-rich pedestals (Zahavi *et al.*, 2011). *Salmonella enterica* serovar Typhi adheres to the intestinal tract by binding to the cellular receptor cystic fibrosis transmembrane conductance regulator (an epithelial chloride channel) via the TFP structural pilin, PilS. The pili are required for epithelial cell invasion (Balakrishna *et al.*, 2009).

In *C. perfringens*, PilT and PilC enable the bacterium to move on surfaces via gliding motility. Mutations introduced into the *pilC* and *pilT* genes abolish both surface localization of TFP and motility in this organism (Varga *et al.*, 2006). The observation that the *pilT* mutation abolishes detectable TFP expression in *C. perfringens* is a critical difference with all other instances studied, as loss or absence of *pilT* usually leads to hyperpiliation (Jarrell, 2009). In contrast to the extensively studied paradigms of *N. gonorrhoeae*, *P. aeruginosa* and *M. xanthus*, *pilT* in *C. perfringens* may thus also be involved in pilus biogenesis in addition to pilus retraction (Jarrell, 2009). Moreover, TFP in *C. perfirgens* are required for efficient biofilm formation (Varga *et al.*, 2008). Expression of *C.*

perfringens pilA2 in a non-piliated *N. gonorrhoeae* strains provides the organism with the ability to attach to mouse and rat myoblast cell lines, while a *C. perfringens pilT* mutant has reduced ability to adhere to mouse myoblasts under anaerobic conditions (Rodgers *et al.*, 2011).

In *C. difficile*, two TFP-associated loci (CD3294-CD3297 and CD3503-CD3513) are present in the genome and are conserved across many strains of human and animal origin, together with low DNA sequence divergence (Fig. 5, Janvilisri *et al.*, 2009) this would suggest a function of importance to the organism (Sebaihia *et al.*, 2006; Stabler *et al.*, 2006; Janvilisri *et al.*, 2009). These TFP loci contain genes predicted to encode pilin subunits, prepilin peptidases and proteins involved in TFP assembly. Previously, Borriello *et al.* (1988) demonstrated the presence of multiple polar fimbriae which were 4-9 nm in diameter and up to 6 µm long on the surface of 5 out 15 *C. difficile* strains examined (Borriello *et al.*, 1988). A recent *in vivo* study of *C. difficile* 630 infection in the Golden Syrian hamster model found that high numbers of bacteria were present at the base of the crypts of the gut tissue, and were observed to interact with the gut microvillus via one pole of the bacterium. Pilus-like structures recognised by antisera against PilA (CD3507) were detected on the surface of bacterial cells, suggesting that pili may be expressed in the hamster model (Goulding *et al.*, 2009). In addition, Purcell *et al.* (2012) demonstrated that elevated levels of cyclic diguanylate (c-di-GMP) enhance aggregate formation in *C. difficile* and these aggregates were visibly bound by long, thin fibres (10 nm thick) which are consistent with the size of pili, although their composition is yet to be determined (Purcell *et al.*, 2012). These authors hypothesized that c-di-GMP may upregulate production of

pili to enhance aggregation of *C. difficile* cells and biofilm formation, thus aiding in colonization by the pathogen (Purcell *et al.*, 2012). It is unknown at the moment when TFP is expressed, how TFP genes are regulated in this organism and whether it is important in the survival and infection of *C. difficile*.

This study describes further optimisations of the ISce-I system generated in chapter 3. The system was used to delete the *pilT* (CD3505) and *pilA* (CD3507) genes in *C. difficile*, predicted to encode a protein involved in TFP retraction and for the type IV pilin subunit, respectively. To enable the study of potential twitching motility in *C. difficile*, a non-flagellated *C. difficile* strain was first constructed via deletion of the *fliC* gene. Furthermore, an intergenic region hypothesised to be involved in the regulation of TFP genes in *C. difficile* was also investigated.

4.2: RESULTS

4.2.1: Optimization of the *C. difficile* I-SceI genetic system and construction of *C. difficile* Δ*fliC*

Construction of a new I-SceI site delivery vector:

Using the I-SceI site vectors pES185 and pES2761, that were generated in chapter 3, allele replacement vectors were constructed (not described here) to delete the *fliC*, *pilT* (CD3505) and *pilA* (CD3507) genes in *C. difficile*. However, multiple attempts to conjugate and integrate these vectors into *C. difficile* were unsuccessful (data not shown). Therefore, a new candidate plasmid was sought to generate an I-SceI site delivery vector. Recently, Heap *et al.* (2012) used the *E. coli – Clostridium* replicative, but segregationally unstable, shuttle plasmid pMTL83151 (i.e. the plasmid replicates at a rate slower than that of the host chromosome) to construct an integration vector for use in *C. difficile* (Heap *et al.*, 2012; Heap *et al.*, 2009). The plasmid pMTL83151 carries a thiamphenicol resistance marker, *catP*, for selection. The authors observed that following introduction of an integration vector based on pMTL83151 into *C. difficile* 630Δ*erm*, thiamphenicol selection resulted in two populations of cells with different growth rates; i) faster growing cells, that represented cells in which single cross-over occurred and ii) slower growing cells, that corresponded to cells in which the vector remained in a free replicative form. This growth difference was attributed to the segregational instability of the plasmid (due to the pCB102 replicon) in *C. difficile* 630 and was considered advantageous since

integrants could be enriched from a population of transconjugants by culturing on antibiotic containing medium twice. Therefore, the plasmid pMTL83151 was chosen to construct a new I-SceI site delivery vector. Using a construction strategy similar to that for the pES185 I-SceI site delivery vector (chapter 3), oligos ISceI_F_SacI and ISceI_R_SacI representing the I-SceI recognition site were phosphorylated at 5' ends, annealed to each other and ligated into the SacI site of pMTL83151, generating the new I-SceI site delivery vector pES242 (**FIG 4.2**). The use of pES242 in the I-SceI genetic system would aid the selection for a single cross-over event via thiamphenicol resistance and the screening for a second cross-over event via the loss of thiamphenicol resistance.

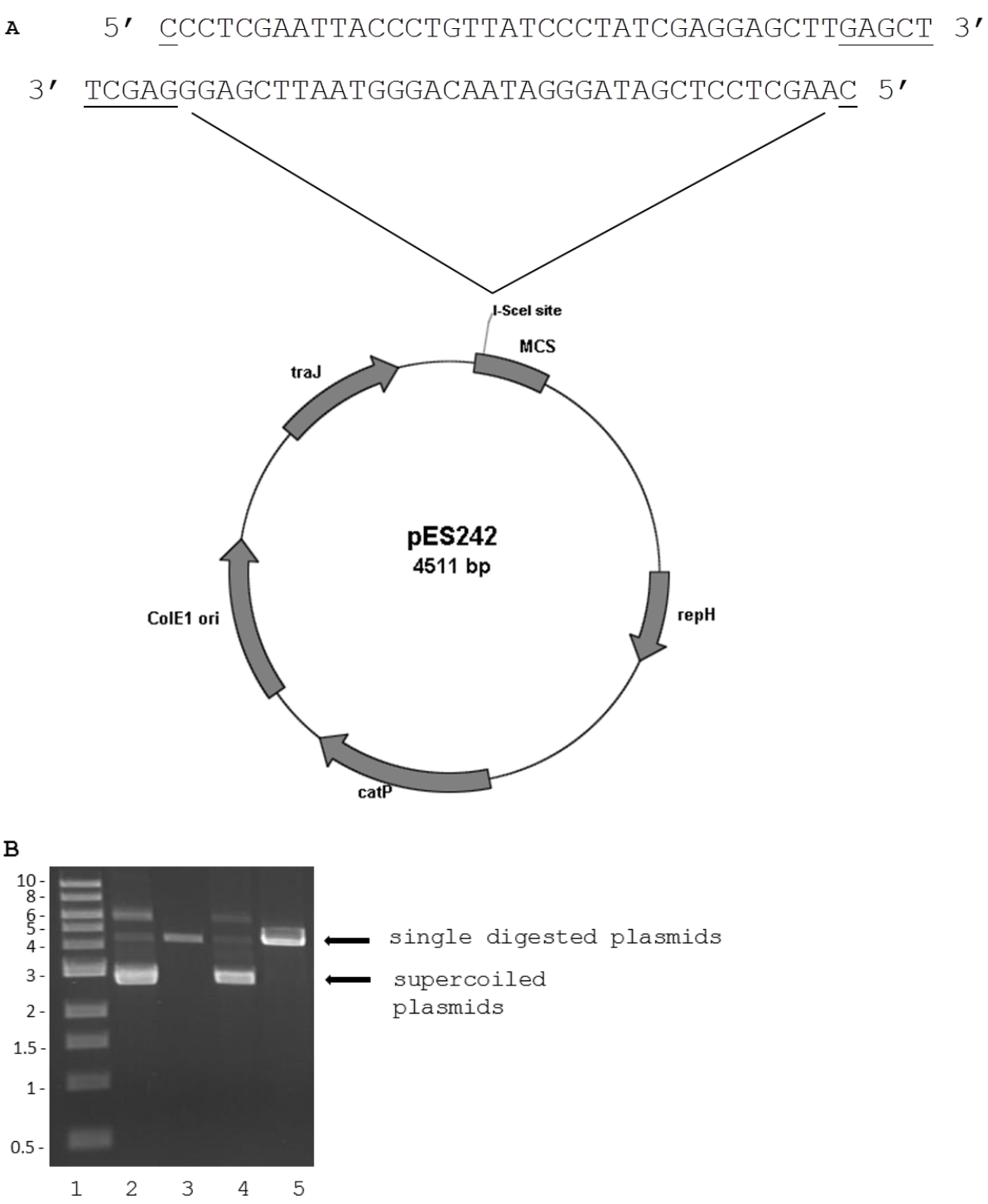


FIG 4.2 Schematic representation and agarose gel analysis of pES242

A] Oligos ISceI_F_SacI and ISceI_R_SacI carrying the I-SceI recognition site and SacI restriction site (underlined)were ligated into the SacI site of pMTL83151 generating pES242 (4511 bp). **B]** Restriction digest analysis of pES242. Lane 1 corresponds to 1 kb ladder. Lanes 2, 3, 4 and 5 correspond to supercoiled pMTL83151, SacI digested pMTL83151, supercoiled pES242 and I-SceI digested pES242, respectively. The vector pES242 contains a Gram-positive replicon pCB102, a Gram-negative replicon ColE1, a conjugative transfer *traJ* gene and a *catP* thiamphenicol resistance marker.

Construction of *C. difficile* Δ*fliC*:

C. difficile produces peritrichous flagella (17-22 nm in diameter consistent with the size on other enteric bacteria) that are uniformly distributed on the surface of the bacterium and allow swimming or swarming motility (Delmee *et al.*, 1990; Vedantam *et al.*, 2012). To investigate twitching motility and aid visualisation of type IV pilus surface production in *C. difficile*, a non-flagellated bacterium was constructed. In the genome of *C. difficile* 630 two flagella-associated loci (CD0226-CD0240 and CD0245-CD0271), separated by an intervening locus, are present with one of these loci CD0226-CD0240 containing the *fliC* gene (Stabler *et al.*, 2009). The FliC protein is a 39 kDa flagellin protein which is the major structural component of the flagellar filament (Tasteyre *et al.*, 2000). It was previously shown that interruption of this gene results in the loss of flagellum production and flagella-mediated motility (Twine *et al.*, 2009; Dingle *et al.*, 2011). To delete the *fliC* gene (**FIG 4.3**) the I-SceI vector pES242 and the I-SceI expression vector pES288 were used.

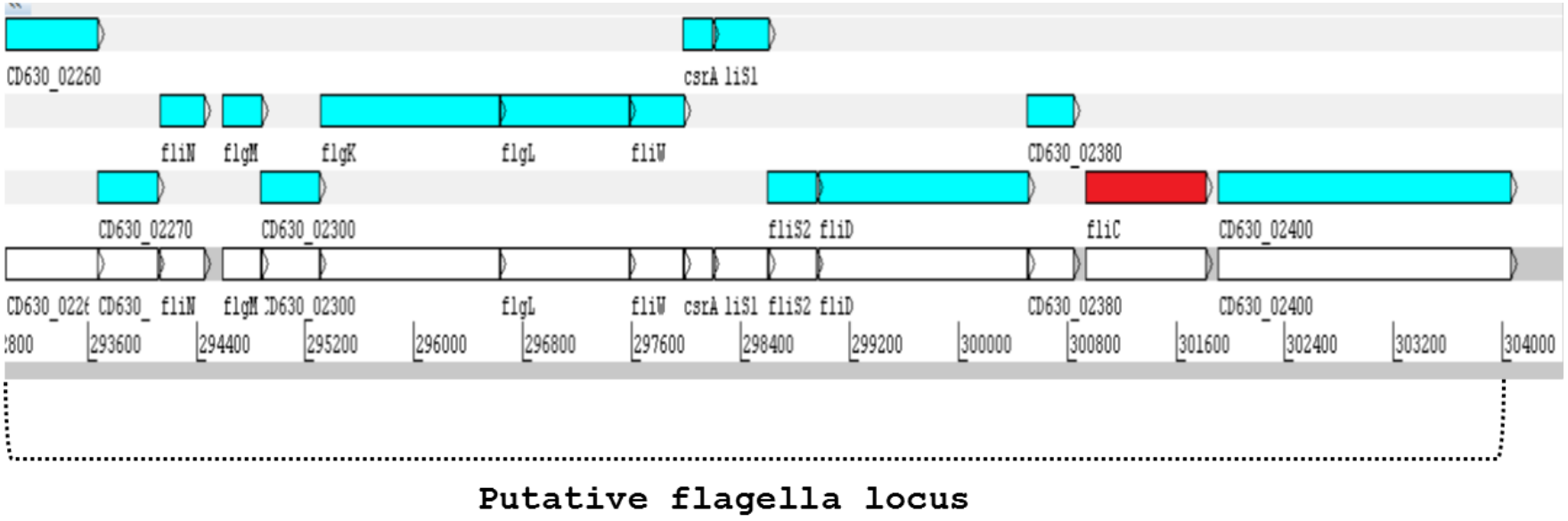


FIG 4.3 *C. difficile* flagella locus and *fliC* gene for deletion

The putative flagella locus CD0226 – CD0240 in the *C. difficile* 630 genome is shown and the *fliC* gene for deletion is highlighted in red.

Construction of the fliC allele replacement vector pES2921:

Regions(~500 bp) flanking the *fliC* gene were amplified by PCR, using primers SphI-fliC1 and fliC2 to amplify the left flank (LF) and primers fliC3 and SphI-fliC4 to amplify the right flank (RF), with *C. difficile* 630Δ*erm* genomic DNA as template (**FIG 4.4**). The LF and RF PCR products were mixed and fused by cross-over PCR using primers SphI-fliC1 and SphI-fliC4, generating the *fliC* deletion cassette (**FIG 4.4**). The deletion cassette was subsequently ligated into the FspI site of pES242, resulting in the allele replacement vector pES2921 (**FIG 4.4**). The use of this allele replacement vector in the I-SceI system would result in the replacement of the *fliC* gene (873 bp) by the *fliC* deletion cassette (~1 kb). Sequencing of plasmid pES2921 validated its construction.

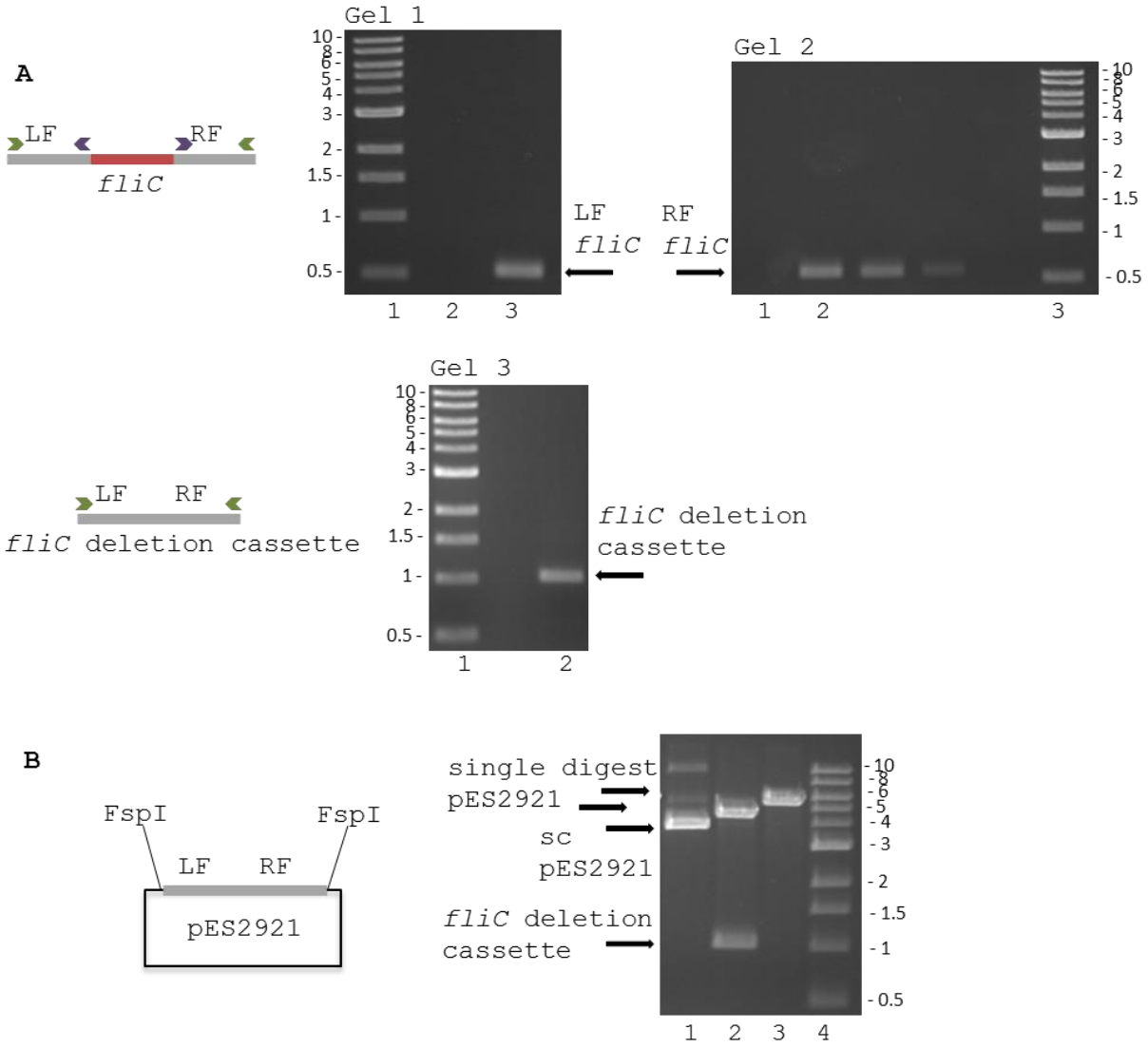


FIG 4.4 Construction of the allele replacement vector pES2921

A] Construction of the *fliC* deletion cassette. Primers SphI-fliC1 (➤) and fliC2 (➤) were used to amplify by PCR the left flanking region (LF, 502 bp) of the *fliC* gene. Primers fliC3 (➤) and SphI-fliC4 (➤) were used to amplify by PCR the right flanking region (RF, 499 bp) of the *fliC* gene. The LF and RF products were fused together by subsequent cross-over PCR using primers SphI-fliC1 (➤) and SphI-fliC4 (➤) generating the *fliC* deletion cassette (~1 Kb). Gel 1 lanes 1, 2 and 3 correspond to 1 kb ladder, primers control PCR and LF PCR product, respectively. Gel 2 lanes 1, 2 and 3 correspond to primers control PCR, RF PCR product and 1 Kb ladder, respectively. Gel 3 lanes 1 and 2 correspond to 1 Kb ladder and *fliC* deletion cassette, respectively. **B]** Construction and gel analysis of pES2921. The *fliC* deletion cassette was ligated into the FspI site of pES242 generating pES2921. Gel lanes 1, 2, 3 and 4 correspond to supercoiled pES2921, SphI digested pES2921, I-SceI digested pES2921 and 1 kb ladder, respectively.

Introduction of pES2921 in C. difficile 630Δerm and screening for a single cross-over event:

The allele replacement vector pES2921 was introduced into *C. difficile* from *E. coli* via an exponential mating. From this conjugation, three putative transconjugants were obtained (named 10/1, 10/2, 10/3) which were inoculated into liquid BHITm15 (BHI containing thiamphenicol at 15 µg/ml) and were also streaked on BHITm15cfx8cls250 agar (BHI containing thiamphenicol 15 µg/ml, cefoxitin 8 µg/ml, D-cycloserine 250 µg/ml). The presence of pES2921 in *C. difficile* was confirmed by amplification, via PCR, of the *fliC* deletion cassette together with the wt *fliC*, including flanking regions, using plasmid DNA preparations from the overnight cultures grown in BHITm15 as template and primers SphI-fliC1 and SphI-fliC4 (**FIG 4.5**). In addition, the agar medium streaked with the transconjugant strains showed the presence of a mixed sized colony population (as expected with the use of pMTL83151-based plasmids), indicative of cells containing single cross-overs and cells containing free plasmid (**FIG 4.6**).

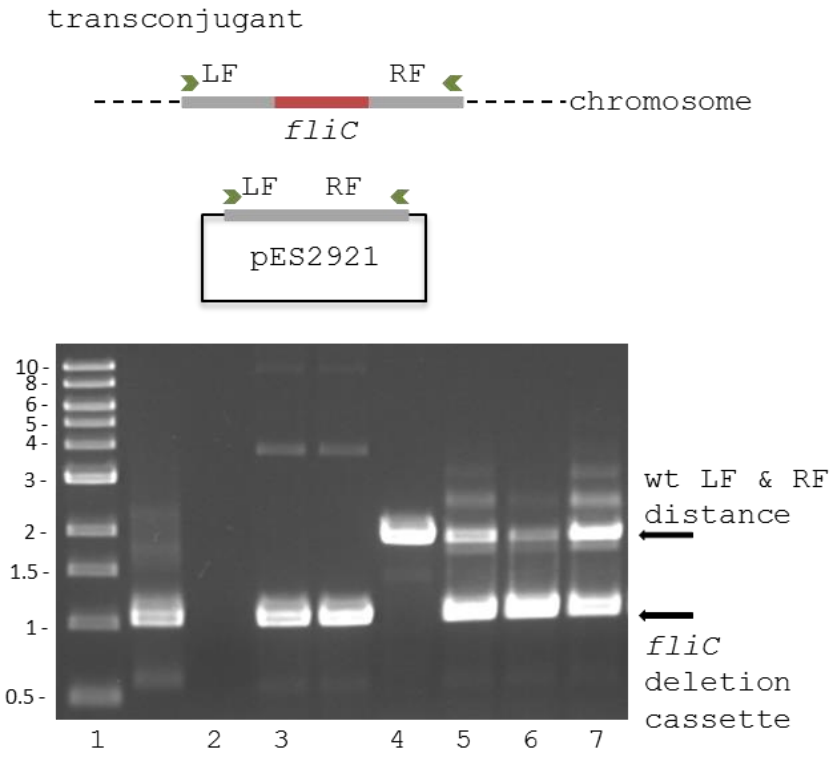


FIG 4.5 PCR screening for pES2921 transconjugants

Primers SphI-fliC1 (➤) and SphI-fliC4 (◄) were used to confirm the presence of pES2921 in putative transconjugant strains 10/1, 10/2 and 10/3 by PCR amplification of the the *fliC* deletion cassette together with the wt *fliC* including flanking regions. From the wt strain only the wt *fliC* including flanking regions is amplified. Gel lane 1 and 2 correspond to 1 kb ladder and primer control PCR, respectively. Lane 3 corresponds to control PCR using pES2921 as template. Lane 4, 5, 6 and 7 corresponds to PCR using plasmid extraction from wt *C. difficile* 630Δ*erm*, from 10/1, from 10/2 and from 10/3, respectively.

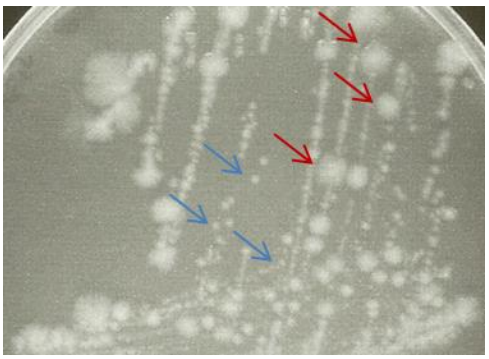


FIG 4.6 Colony sizes after the introduction of pES2921 into *C. difficile* 630Δ*erm*

BHITm15cfx8cls250 agar streaked with pES2921 transconjugant 10/1 displaying mixed sized colony population. Red arrows indicate big colonies corresponding to putative merodiploids and blue arrows indicate small colonies corresponding to putative strains with free plasmid.

Enrichment for a single cross-over population, via a further sub-culture in liquid BHITm15, passaging in BHI broth without thiamphenicol and streaking back on BHITm15 agar, resulted in the loss of the small sized population indicating the presence of putative co-integrants only. To screen, by PCR, for the integration of pES2921 into *C. difficile* 630Δ*erm*, chromosomal DNA was used as template with primer sets INTfliC-For and SphI-fliC4 or SphI-fliC1 and INTfliC-Rev to amplify a ~1.2 kb fragment (**FIG 4.7**). The INTfliC-For and INTfliC-Rev primers were designed to anneal to regions upstream (170 bp) or downstream (278 bp) of the *fliC* LF or *fliC* RF in the targeted chromosome. This confirmed the integration of pES2921 in three merodiploid strains (INT10/1, INT10/2 and INT10/3) via the *fliC* LF (**FIG 4.7**). Two co-integrants (INT10/1 and INT10/3) were chosen for subsequent resolution.

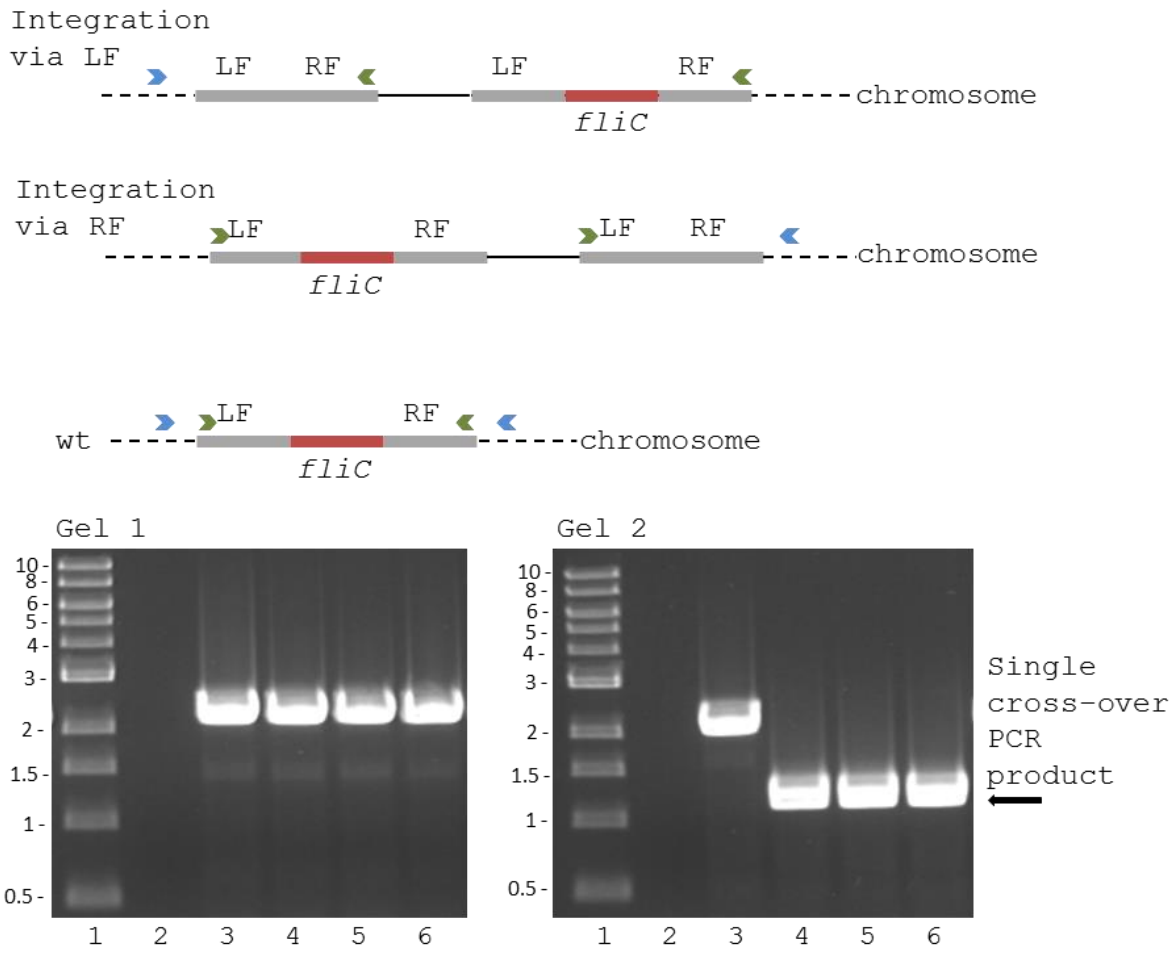


FIG 4.7 PCR screening for single cross-over integration of pES2921

A single cross-over integration of pES2921 into *C. difficile* 630Δ*erm* chromosome was screened by PCR using primer sets INTfliC-For (➤) and SphI-fliC4 (◀) or SphI-fliC1 (➤) and INTfliC-Rev (◀) with chromosomal DNA of putative merodiploids. Note that primers INTfliC-For and INTfliC-Rev anneal outwith of the *fliC* flanking regions on the chromosome. Solid line represents plasmid sequence while dashed line represents chromosomal sequence. The amplification of a ~1.2 kb product with primers INTfliC-For and SphI-fliC4 confirmed three merodiploid strains where integration occurred via the *fliC* LF. Gel 1 represents PCR products with primers SphI-fliC1 and INTfliC-Rev and Gel 2 products with primers INTfliC-For and SphI-fliC4. Gel 1 lane 1 and gel 2 lane 1 correspond to 1 kb ladder. In both gels 1 and 2, lane 2, 3, 4, 5 and 6 correspond to primers control PCR, to PCR product from wt 630Δ*erm*, from 10/1, from 10/2 and from 10/3 strains, respectively.

Introduction of pES288 into merodiploids INT10/1 and INT10/3 and screening for a second cross-over event:

The I-SceI expression vector pES288 was introduced into the partial diploids INT10/1 and INT10/3, via exponential matings, and the resulting putative transconjugants (16 and 8 obtained from INT10/1 and INT10/3 mating, respectively) were streaked twice (sub1 and sub2) on AIM containing erythromycin at 10 µg/ml (AIMErm10). Screening for thiamphenicol sensitive strains from the first subculture showed resolution of 0% (total 65 colonies screened) and 2.8% (1 sensitive colony from total 35 screened) from INT10/1 and INT10/3, respectively. Screening for a second cross-over event via thiamphenicol sensitivity from the strains streaked twice resulted in resolution of 2% (3 sensitive colonies from total 144 screened) and 3.6% (4 sensitive colonies form total 111 screened) from 10/1 and 10/3, respectively. However, this low resolution rate may reflect the visual screening for less rhizoid, smaller-sized colony population. To confirm that the double cross-over event in the merodiploids INT10/1 and 10/3 was due to I-SceI induced DSB repair, and not due to a spontaneous event, plasmid pMTL82254 was introduced into these strains and subsequently screened for resolution. Transconjugant populations streaked once or twice on AIMErm10 did not become sensitive to thiamphenicol (total 50 once streaked and 133 twice streaked colonies screened from each 10/1 and 10/3 transconjugant population).

Verification of a double cross-over and phenotypic characterization of ΔfliC recombinant strains:

The double cross-over events were confirmed via PCR using primers SphI-fliC1 and SphI-fliC4 and genomic DNA from the thiamphenicol-sensitive strains isolated above. The *fliC* deletion cassette was amplified from putative Δ*fliC* strains No. 88, 381, 383 and 415, while the wt *fliC* region was amplified from the parental strain 630Δ*erm* and from the putative resolved wt strains No.382, 485 and 489 (**FIG 4.8**). Subsequent sequencing of the amplified *fliC* deletion cassette from strains No. 88, 381, 383 and 415 confirmed the replacement of the wt *fliC* allele by the deletion cassette in the chromosome of these strains with the predicted sequence joints (**FIG 4.8**). The Δ*fliC* strains No. 381 and 383, and the resolved wt No. 382, were isolated from co-integrant INT10/1. The Δ*fliC* strains No. 88 and 415, and the resolved wt No. 485 and 489, were isolated from the resolution of the INT10/3 partial diploid. It should also be noted that the pES288 vector was easily eliminated from all resolved strains simply by streaking twice on medium without erythromycin. This was important for subsequent experiments, particularly for the construction of strains with multiple deletions.

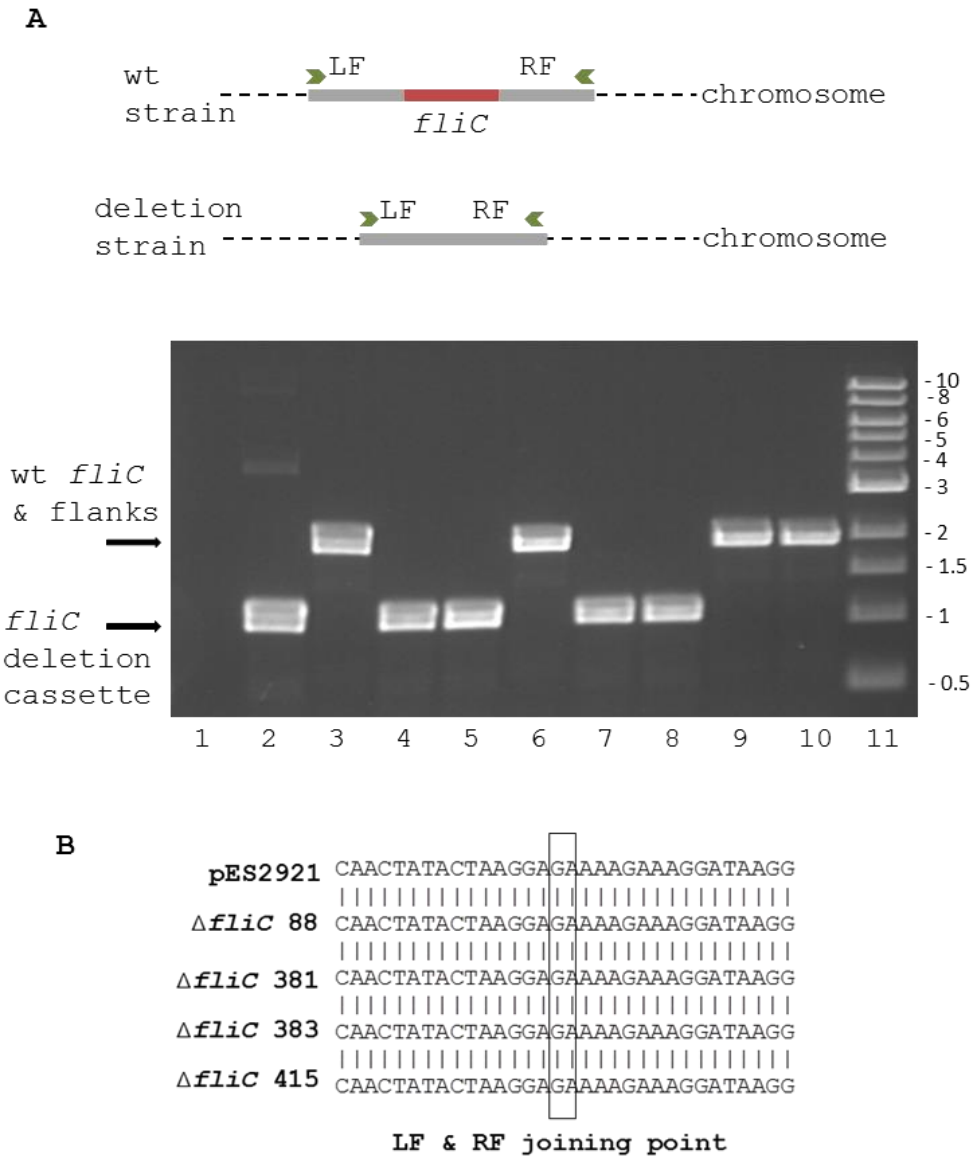


FIG 4.8 Screening and confirmation of second cross-over recombinants and Δ*fliC* strains

A] A second cross-over event was identified by PCR, using primers SphI-fliC1 (➤) and SphI-fliC4 (◄) with chromosomal DNA of putative recombinant strains. The amplification of the *fliC* deletion cassette (~1 kb) identified putative Δ*fliC* recombinant strains, while the amplification of the wt *fliC* including flanking regions (~1.9 kb) identified strains that resolved to wt. Gel Lane 11 corresponds to 1 kb ladder. Lane 1 and lane 2 correspond to primers control PCR and pES2921 control PCR. Lanes 4, 5, 7 and 8 correspond to the amplification of the *fliC* deletion cassette from Δ*fliC* No. 88, 381, 383 and 415 strains, respectively. Lanes 3, 6, 9 and 10 correspond to the amplification of the wt *fliC* and flanks from wt 630Δ*erm*, from resolved wt strains No.382, 485 and 489, respectively. **B]** The *fliC* deletion cassettes amplified from the Δ*fliC* strains were sequenced and a partial sequence alignment of the *fliC* LF and RF together with the joining point is shown.

As mentioned above, previously published studies demonstrated that interruption of the *fliC* gene in *C. difficile* resulted in non-flagellated strains that were impaired in flagella-mediated motility (Dingle *et al.*, 2011; Twine *et al.*, 2009). Motility tests were therefore used, to characterize and confirm the predicted phenotype of the markerless deletion strains, such as soft agar swimming, motility examination by wet mount microscopy, colony morphology and electron microscopy to visualize flagella. The assessment of flagella-mediated motility by stab inoculation of the parental strain 630Δ*erm*, together with wt recombinant strains No.382, 485 and 489, showed a spreading and diffuse growth away from the inoculation point after 24-h, which indicated a motile phenotype (**FIG 4.9**). Microscopic examination of wet mount preparations of the resolved wt strains No.382, 485 and 489 showed the same motile behaviour as the wt 630Δ*erm* strain. In contrast, strains Δ*fliC* No. 88, 381, 383 and 415 did not display a spreading phenotype in the motility agar after 24-h and growth was concentrated near the inoculum stab, which indicated a non-motile phenotype (**FIG 4.9**). In addition, these strains were non-motile as examined by wet mount microscopy. The Δ*fliC* 383 strain displayed typical *C. difficile* colony morphology, similar to the wt, with spreading rhizoids when streaked on BHI, AIM and Blood agar. Flagella production in Δ*fliC* No. 383 and in Δ*fliC* No. 88 was examined by electron microscopy. This showed that the Δ*fliC* mutant strains did not produce any flagella filaments, whereas both wt 630Δ*erm* and resolved wt No. 489 produced visible flagella (**FIG 4.10**).

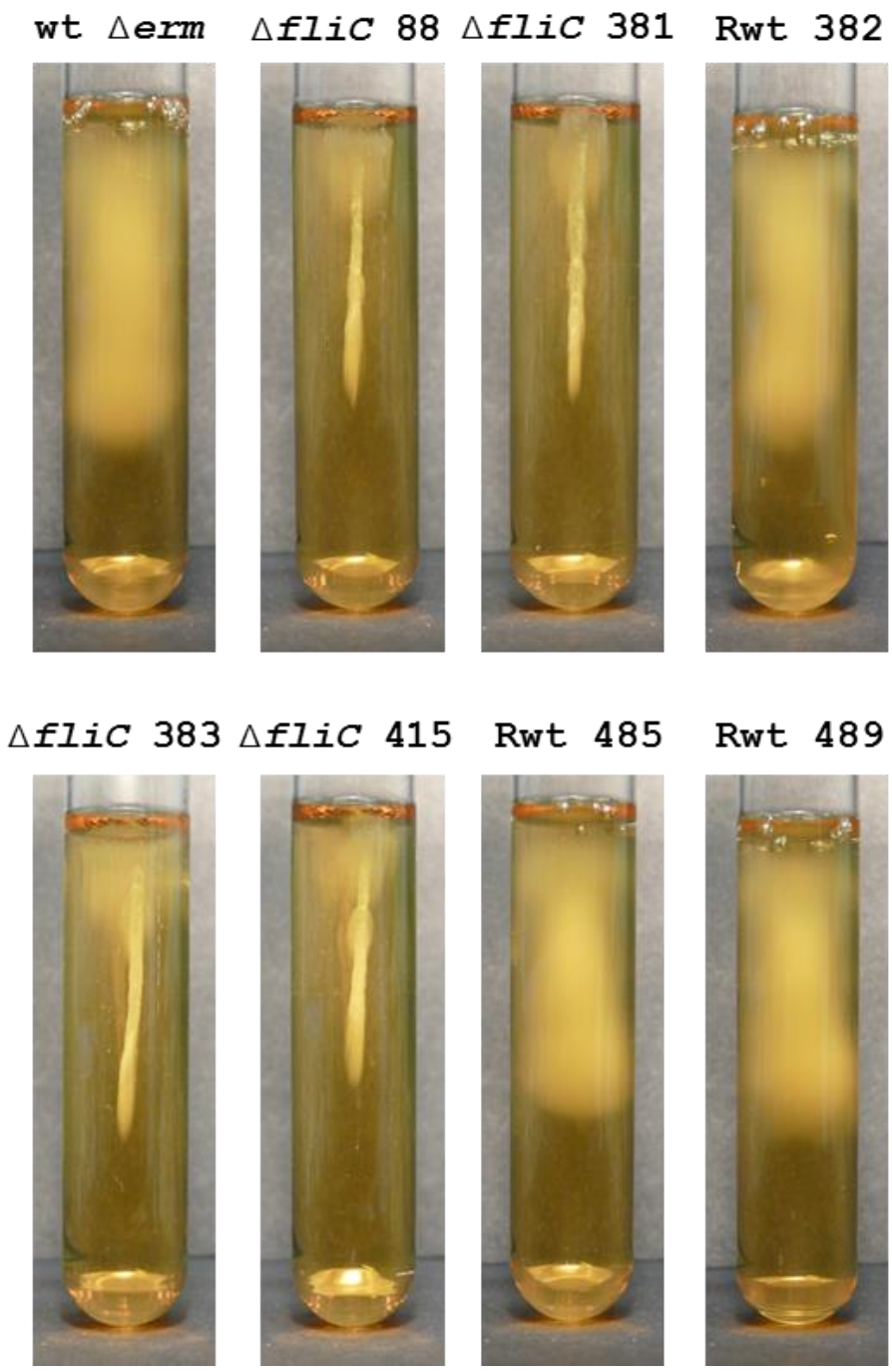


FIG 4.9 Flagella mediated motility stab assay of Δ*fliC* and wt recombinant strains and wt 630Δ*erm*

The result of swimming motility assay at 24-h incubation in AIM medium (0.175% agar) is shown. The Δ*fliC* strains display a non-spreading phenotype in contrast to the wt 630Δ*erm* and to the resolved wt strains where motility is clearly visible.

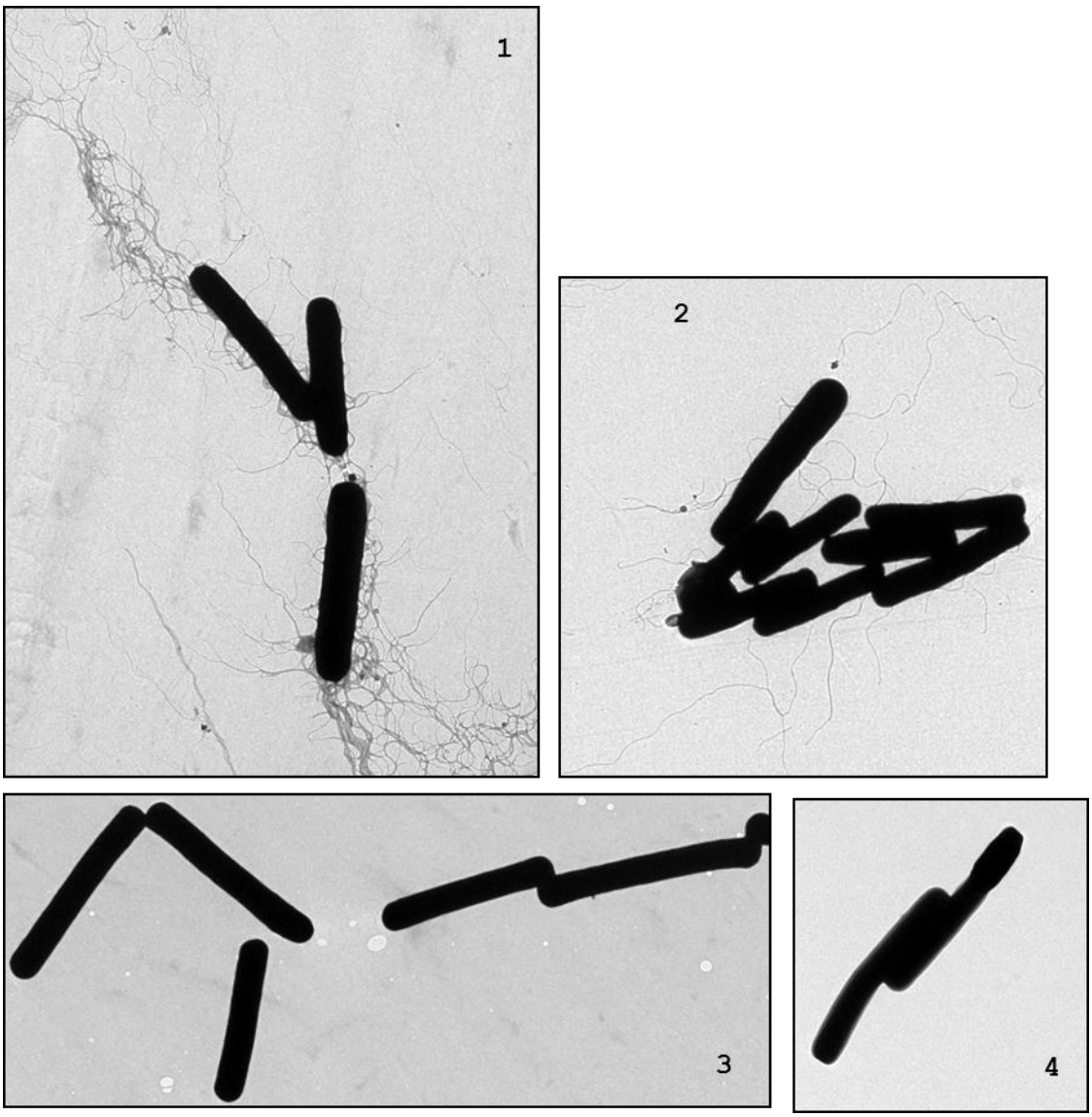


FIG 4.10 Electron microscopy of flagella production in wt and Δ*fliC* strains

Transmission electron micrographs of 48-h single colonies grown on AIM agar medium revealing the lack of flagella production in Δ*fliC* 88 (3) and Δ*fliC* 383 (4) mutant strains in contrast to the wt 630Δ*erm* (1) and resolved wt 489 (2) strains where visible flagella is produced.

Complementation of the ΔfliC mutants:

To complement the Δ*fliC* mutants, the recombinant vector pES196 was constructed. To generate this plasmid, the intact wt *fliC* gene and its native promoter, previously documented by Dingle *et al.* (2011), were amplified by PCR using primers cmFliCFor and cmFliCRev and genomic DNA from 630Δ*erm* as template. The resulting PCR product was ligated into NotI/XhoI- digested *E. coli – Clostridium* shuttle vector pMTL84151 (**FIG 4.11**) (Heap *et al.*, 2009). The plasmid pES196 was sequenced to validate construction.

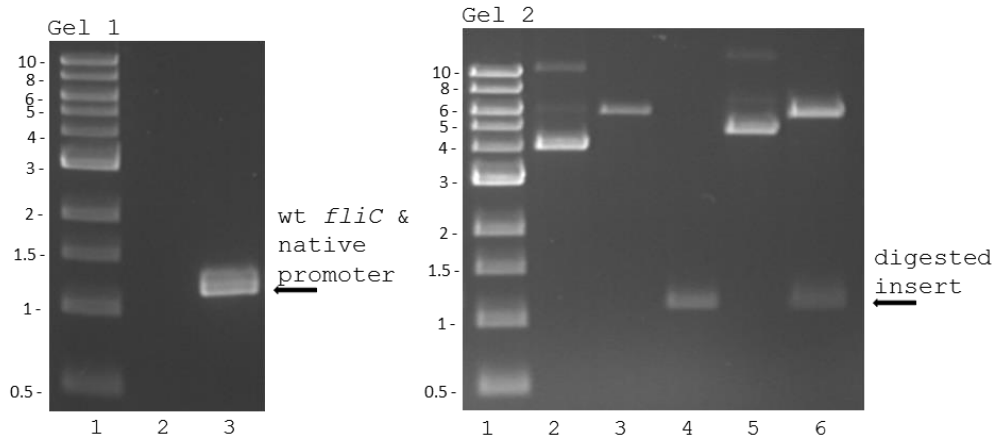


FIG 4.11 Construction of complementation plasmid pES196

The complementation plasmid was generated be the ligation of the wt *fliC* gene and its respective promoter in the NotI/XhoI site of pMTL84151. Gel 1 lane 3 represents the PCR amplification of the wt *fliC* and promoter from genomic DNA of 630Δ*erm* using primers cmFliCFor and cmFliCRev. Gel 1 lane 1 and 2 correspond to 1 kb ladder and primers control PCR. Gel 2 represents pES196 restriction digest analysis for insert confirmation. Gel 2 lane 1 corresponds to 1 kb ladder. Gel 2 lane 2 and 3 correspond to supercoiled pMTL84151 and NotI/XhoI digested pMTL84151, respectively. Gel 2 lane 5 and 6 correspond to supercoiled pES196 and NotI/XhoI digested pES196, respectively. Gel 2 lane 4 corresponds to NotI/XhoI digested insert (1149 bp) used for construction of pES196.

The complementation vector was conjugated into wt 630Δ*erm*, Δ*fliC* 88 and Δ*fliC* 383 strains using exponential matings. For control experiments, the plasmid pMTL84151 was also conjugated into these strains. The presence of pMTL84151 or pES196 in the strains was confirmed by PCR amplification of plasmid specific products using primers Trans-1 and Trans-3 with chromosomal DNA from transconjugant strains as template (**FIG 4.12**).

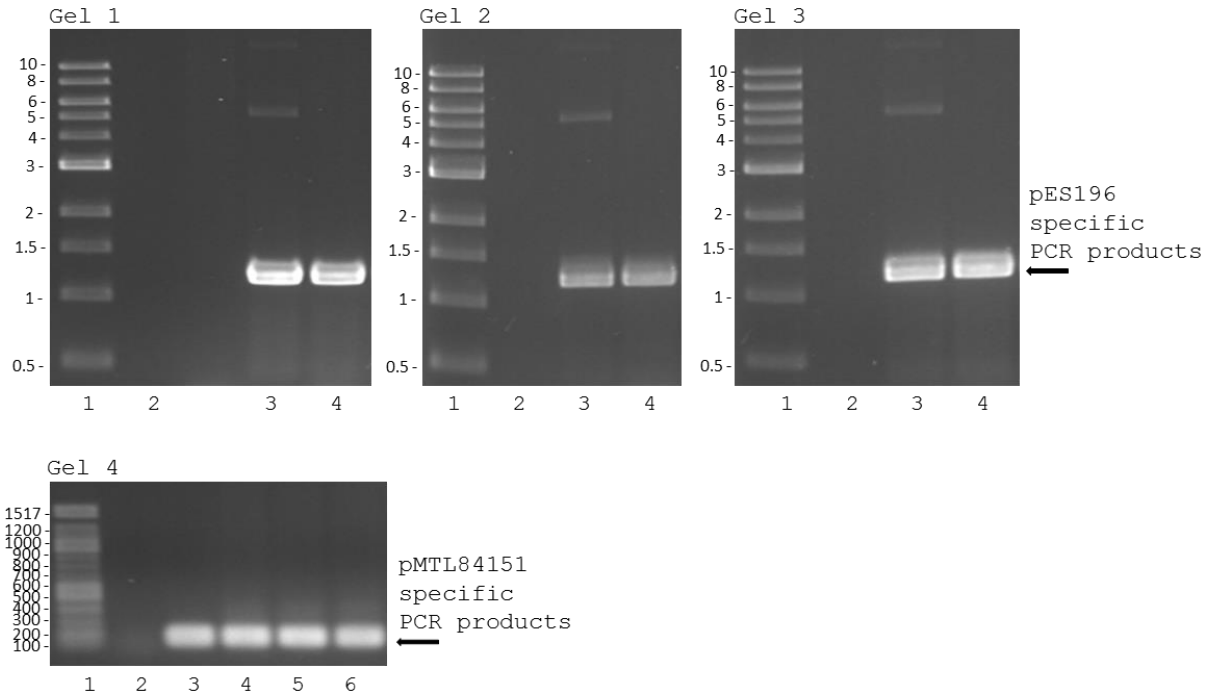


FIG 4.12 PCR confirmation of pES196 and pMTL84151 transconjugant strains

Transconjugant strains were confirmed to harbour pMTL84151 or pES196 by PCR amplification of plasmid specific products, using chromosomal DNA of putative transconjugant as template and primers Trans-1 and Trans-3, which were designed to anneal outwith of the NotI and XhoI sites of pMTL84151. Lane 1 of gel 1, gel 2 and gel 3 refer to 1 kb ladder and lane 1 of gel 4 refers to 100 bp ladder. Lane 2 of all gels refers to primers control PCR. Lane 3 of gels 1, 2 and 3 refer to control PCR using plasmid pES196 as template. Lane 4 of gel 1, 2 and 3 refer to PCR product produced from Δ*fliC* 383, Δ*fliC* 88 and wt 630Δ*erm* chromosomal DNA, respectively. Lane 3 of gel 4 refers to control PCR using plasmid pMTL84151 as template. Lane 4, 5 and 6 of gel 4 refer to pMTL84151 specific PCR product produced from chromosomal extract of wt 630Δ*erm*, Δ*fliC* 383 and Δ*fliC* 88, respectively.

The strains wt 630Δ*erm*, Δ*fliC* 383 and Δ*fliC* 88 containing pES196 or pMTL84151 were stab inoculated into soft agar medium. After 24-h incubation, only strains of 630Δ*erm* containing either plasmid showed a spreading phenotype, while the Δ*fliC* strains, containing either plasmid, showed a non-spreading phenotype (**FIG 4.13**). Wet mount microscopic examination of the wt 630Δ*erm* strain, containing either plasmid, displayed typical *C. difficile* motility with vigorous back and forth movements. Wet mount microscopy of both Δ*fliC* 383 and Δ*fliC* 88 strains containing pES196 showed some motile cells, singly or in pairs, that displayed a tumbling motion. In addition, both Δ*fliC* 383 and Δ*fliC* 88 strains containing the control vector pMTL84151 were non-motile when examined by wet mount microscopy. Further examination by electron microscopy of the Δ*fliC* 383 and Δ*fliC* 88 strains containing the complementation vector, pES196, revealed production of flagellar filaments, while the Δ*fliC* 383 and Δ*fliC* 88 strains containing pMTL84151 did not produce any visible flagella (**FIG 4.14**). Electron microscopy of the parental strain 630Δ*erm* containing pES196 or pMTL84151 showed production of flagella (**FIG 4.14**).

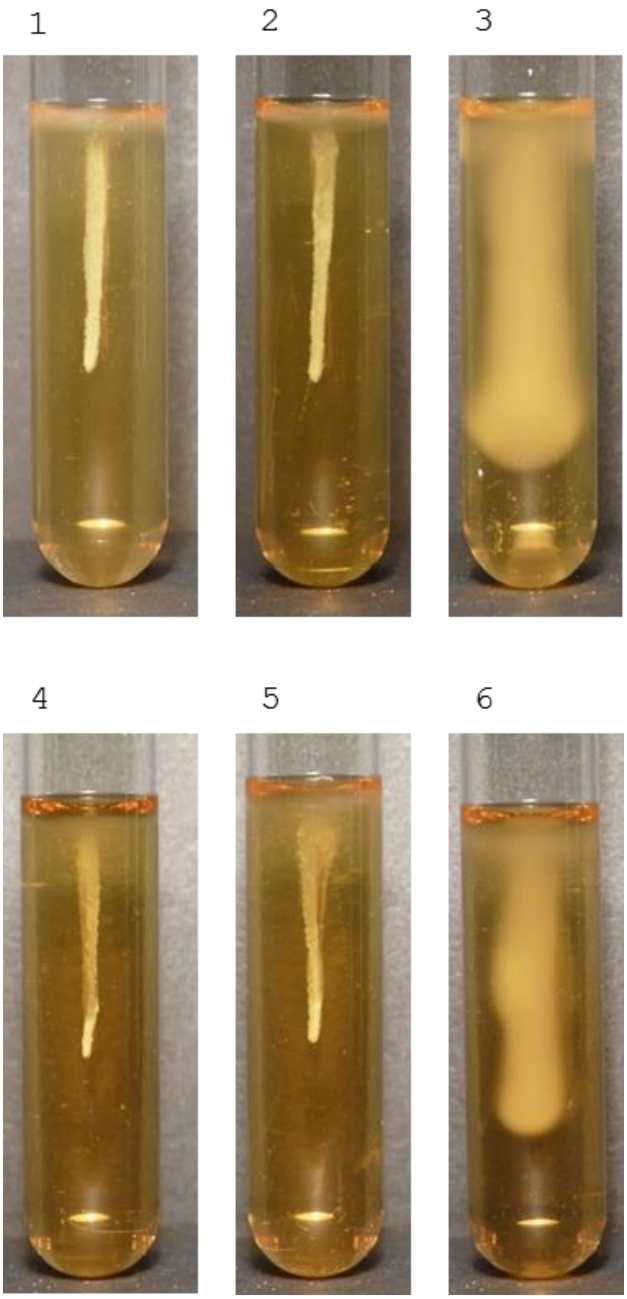


FIG 4.13 Flagella motility stab assay of complemented Δ*fliC* strains

The result of swimming motility assay at 24-h incubation in BHI medium (0.175% agar) containing thiamphenicol at 15 µg/ml is shown. Photos 1, 2 and 3 refer to Δ*fliC* 383, Δ*fliC* 88 and wt 630Δ*erm* containing pES196, respectively. Photos 4, 5 and 6 refer to Δ*fliC* 383, Δ*fliC* 88 and wt 630Δ*erm* containing pMTL84151, respectively. Only, the wt 630Δ*erm* containing either plasmid displays a spreading phenotype in the motility agar medium.

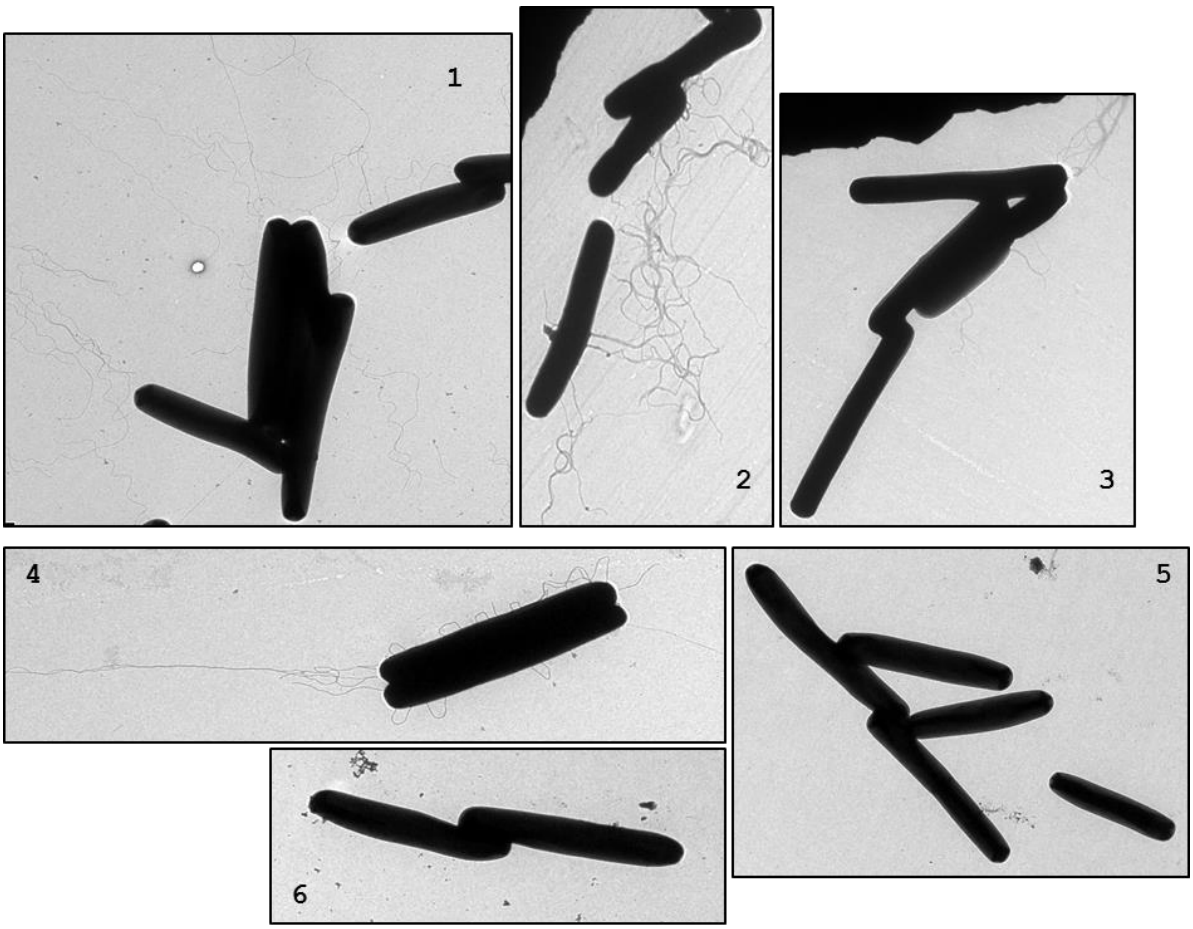


FIG 4.14 Electron microscopy of complemented Δ*fliC* strains for flagella production

Transmission electron micrographs of 48-h single colonies grown on AIM agar medium revealed flagella production in *fliC* complemented Δ*fliC* 88 and Δ*fliC* 383 mutant strains. Photos 1, 2 and 3 refer to wt 630Δ*erm*, Δ*fliC* 383 and Δ*fliC* 88 containing pES196, respectively. Photos 4, 5 and 6 refer to wt 630Δ*erm*, Δ*fliC* 383 and Δ*fliC* 88 containing control plasmid pMTL84151, respectively.

4.2.2: Construction of *C. difficile* Δ*pilT*, Δ*fliC*Δ*pilT*, attempted construction of *C. difficile* Δ*pilA* and characterization of the mutant strains

Arrangement of putative type IV pili genes in *C. difficile* and targeted *pilT* and *pilA* genes for deletion:

FIG 4.15 shows one of the type IV pilus-associated loci, CD3503–CD3513, in the *C. difficile* 630 genome. The *pilT* (CD3505) and *pilA* (CD3507) genes which are predicted to encode proteins for pilus retraction and type IV pilin subunit, respectively, are shown. The I-SceI site vector pES242 and the I-SceI expression vector pES288 were used to delete these genes.

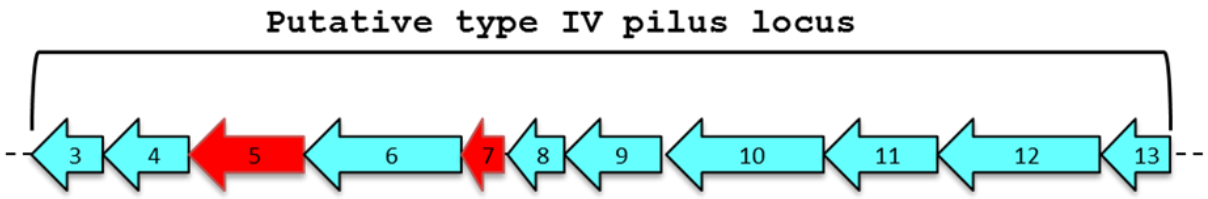


FIG 4.15 Arrangement of type IV pilus locus in *C. difficile* 630 and genes for deletion

The putative type IV pilus biosynthesis locus (CD3503 – CD3513) in the genome of *C. difficile* 630 is shown. Putative gene and putative protein nomenclature: 3- CD3503 type IV prepilin leader peptidase, 4- CD3504 type IV prepilin peptidase, 5- CD3505 twitching motility protein PilT, 6- CD3506 hypothetical protein, 7- CD3507 type IV pilin (PilA, nomenclature according to Goulding *et al.*, 2009), 8- CD3508 type IV pilin, 9- CD3509 type IV pilus assembly protein, 10- CD3510 membrane protein, 11- CD3511 type IV pilus secretion protein, 12- CD3512 type IV pilus transporter and 13- CD3513 pilin protein. The genes *pilT* (CD3505) and *pilA* (CD3507) for deletion are coloured red.

It was predicted that *C. difficile pilA* or *pilT* mutants would be impaired in twitching motility. The deletion of *pilT* was generated in both *C. difficile* 630Δ*erm* and in *C.*

difficile Δ*fliC* 383 strains. The double deletion mutant *C. difficile* Δ*fliC* Δ*pilT* was constructed to: i) aid the investigation of TFP-mediated motility by exclusion of flagella-mediated motility and to aid in the investigation of other future TFP function tests e.g. adhesion assays, biofilms in which in other species flagella are known to be involved; and, ii) test the robustness of the genetic system by the construction of multiple deletions.

Construction of a *C. difficile pilT* deletion strain:

Construction of allele replacement vector pES2922:

Regions flanking (~500 bp) the *pilT* gene were amplified by PCR using chromosomal DNA from 630Δ*erm* as template, with primers SphI-CD3505-1 and CD3505-2 used to amplify the *pilT* LF and primers CD3505-3 and SphI-CD3505-4 used to amplify the *pilT* RF (**FIG 4.16**). The *pilT* flanking regions were fused together by cross-over PCR using primers SphI-CD3505-1 and SphI-CD3505-4, resulting in the *pilT* deletion cassette (**FIG 4.16**), which was subsequently ligated into the FspI site of pES242, generating the allele replacement vector pES2922 (**FIG 4.16**). Plasmid pES2922 was sequenced to validate construction.

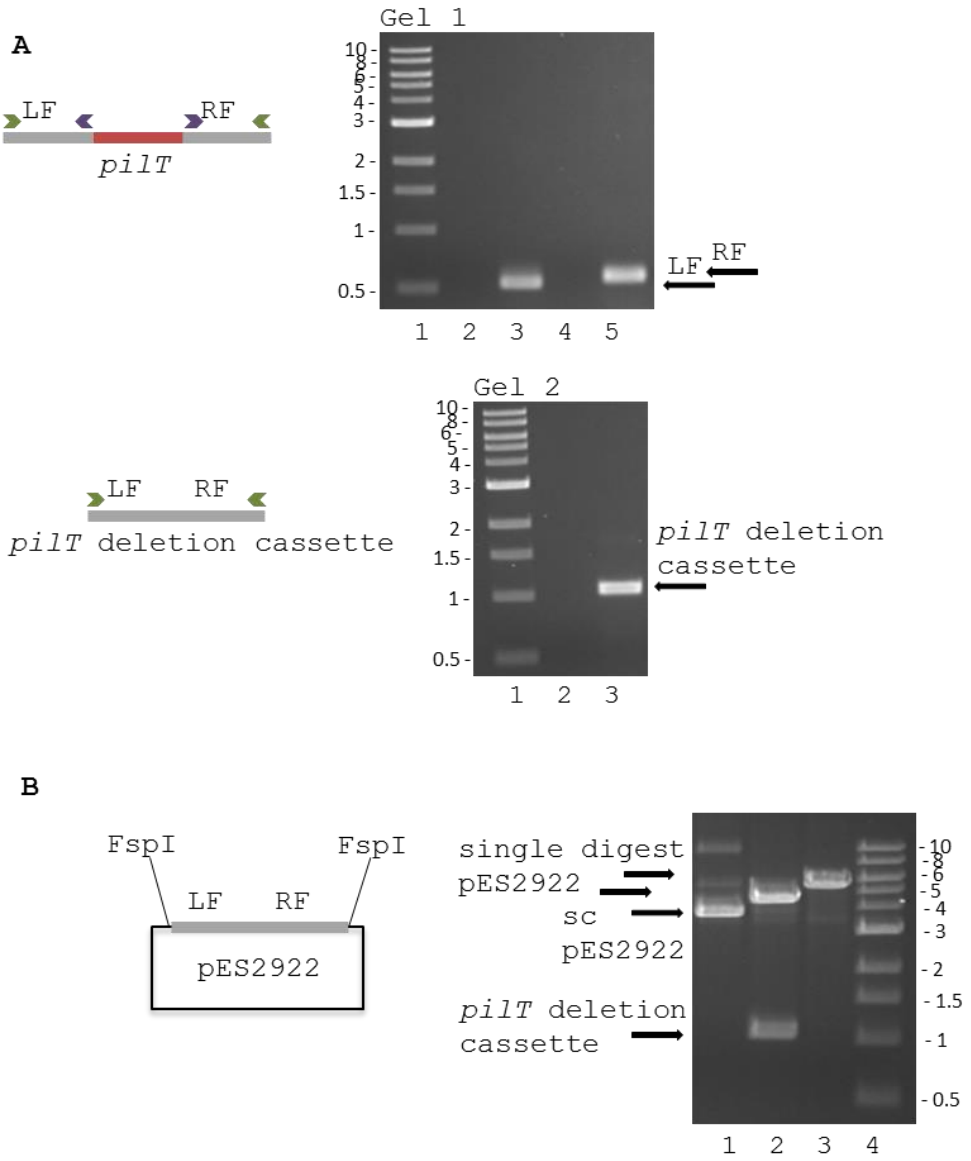


FIG 4.16 Construction of the allele replacement vector pES2922

A] Construction of the *pilT* deletion cassette. Primers SphI-CD3505-1(➤) and CD3505-2 (◀) were used to amplify by PCR the left flanking region (LF, 514 bp) of the *pilT* gene from 630Δ*erm* genomic DNA. Primers CD3505-3 (➤) and SphI-CD3505-4(◀) were used to amplify by PCR the right flanking region (RF, 522 bp) of the *pilT* gene. The LF and RF products were fused together by subsequent cross-over PCR using primers SphI-CD3505-1(➤) and SphI-CD3505-4(◀) generating the *pilT* deletion cassette (~1 Kb). Gel 1 lane 1 refers to 1 kb ladder. Gel 1 lane 2 and 4 refer to primers SphI-CD3505-1/CD3505-2 and CD3505-3/SphI-CD3505-4 control PCR, respectively. Gel 1 lane 3 and 5 refer to LF and RF PCR products, respectively. Gel 2 lanes 1, 2 and 3 refer to 1 kb ladder, primers SphI-CD3505-1/SphI-CD3505-4 control PCR and *pilT* deletion cassette, respectively. **B]** Construction and gel analysis of pES2922. The *pilT* deletion cassette was ligated into the FspI site of pES242 generating pES2922. Gel lanes 1, 2, 3 and 4 correspond to supercoiled pES2922, SphI digested pES2922, I-SceI digested pES2922 and 1 kb ladder, respectively.

Introduction of pES2922 into C. difficile 630Δerm and screening for a single cross-over event:

The Δ*pilT* allele replacement vector, pES2922, was introduced into *C. difficile 630Δerm* by exponential mating. From this mating, six putative transconjugants were obtained (named 35/1, 35/3, 35/5, 35/7, 35/8 and 35/10) which were inoculated into liquid BHITm15 and also streaked on BHITm15cfx8cls250 agar. All putative transconjugants obtained were confirmed to contain the plasmid pES2922 by PCR amplification of the *pilT* deletion cassette and the wt *pilT* gene, including flanking regions, using primers SphI-CD3505-1 and SphI-CD3505-4 with plasmid DNA prepared from overnight cultures as template (**FIG 4.17**). After overnight incubation of the streaked transconjugants strains on BHITm15cfx8cls250 agar, a mixed sized population of big and small colonies was observed. A single cross-over event was enriched as previously described which resulted in the loss of the small sized population. The integration of pES2922, by a single cross-over event, was confirmed by PCR amplification of a ~1.2 kb sized fragment using chromosomal DNA of putative merodiploids with primer set SphI-CD3505-4 and INT3505For or SphI-CD3505-1 and INT3505Rev (**FIG 4.18**). The INT3505For and INT3505Rev primers were designed to anneal outwith the *pilT* LF and RF regions in the *C. difficile* 630Δ*erm* chromosome. This PCR assay confirmed the integration of pES2922 via the *pilT* LF in merodiploids INT35/7 and INT35/8, and the integration via the *pilT* RF in merodiploids INT35/1, INT35/5 and INT35/10 (**FIG 4.18**). In the 35/3 transconjugant strain, pES2922 remained in its replicative form. The partial diploids INT35/5 and INT35/8 were chosen for subsequent resolution by induction of a second cross-over.

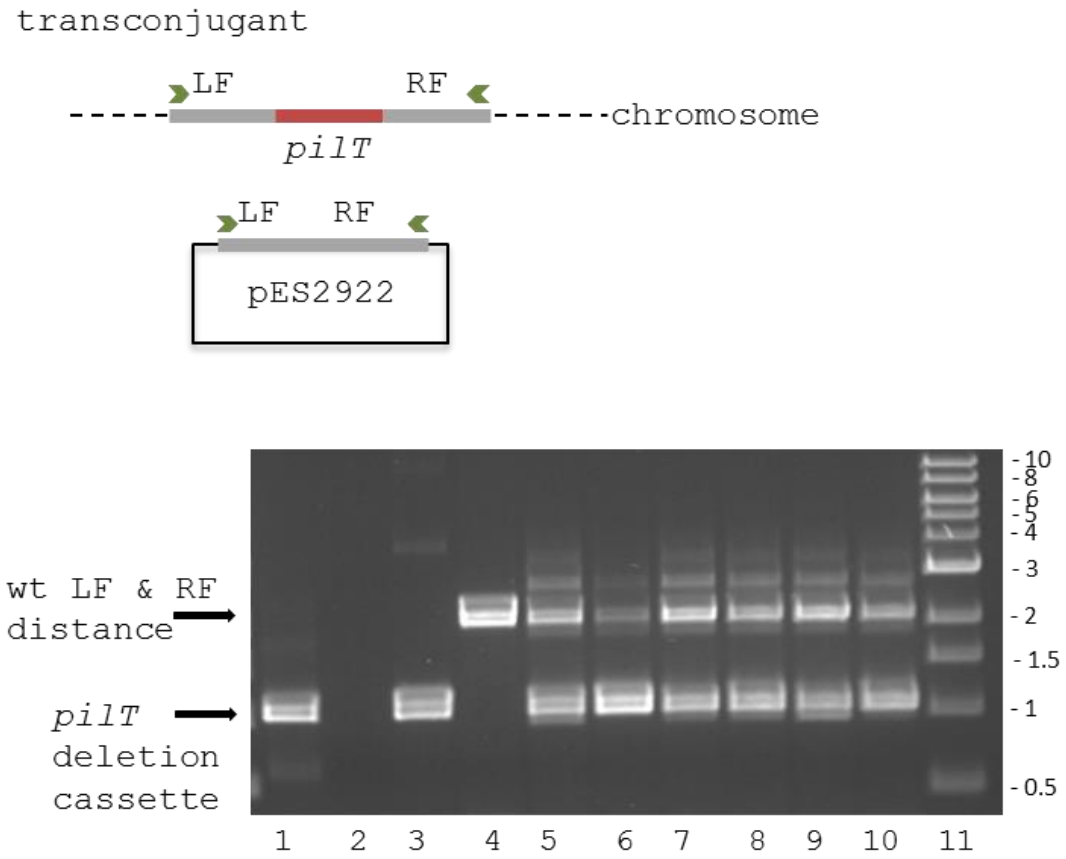


FIG 4.17 PCR screening for pES2922 transconjugants

PCR amplification of the *pilT* deletion cassette together with the wt *pilT* including flanking regions, using primers SphI-CD3505-1 (➤) and SphI-CD3505-4 (➤) with plasmid preparation as template, confirmed the presence of pES2922 in transconjugant strains 35/1, 35/3, 35/5, 35/7, 35/8 and 35/10. From the wt strain 630Δ*erm* only the wt *pilT* including flanking regions is amplified. Gel lanes 1 and 2 refer to the *pilT* deletion cassette used for ligation and primers control PCR, respectively. Lane 11 refers to 1 kb ladder. Lanes 3, 4, 5, 6, 7, 8, 9, and 10 refer to PCR product from pES2922, 630Δ*erm*, 35/1, 35/3, 35/5, 35/7, 35/8 and 35/10, respectively.

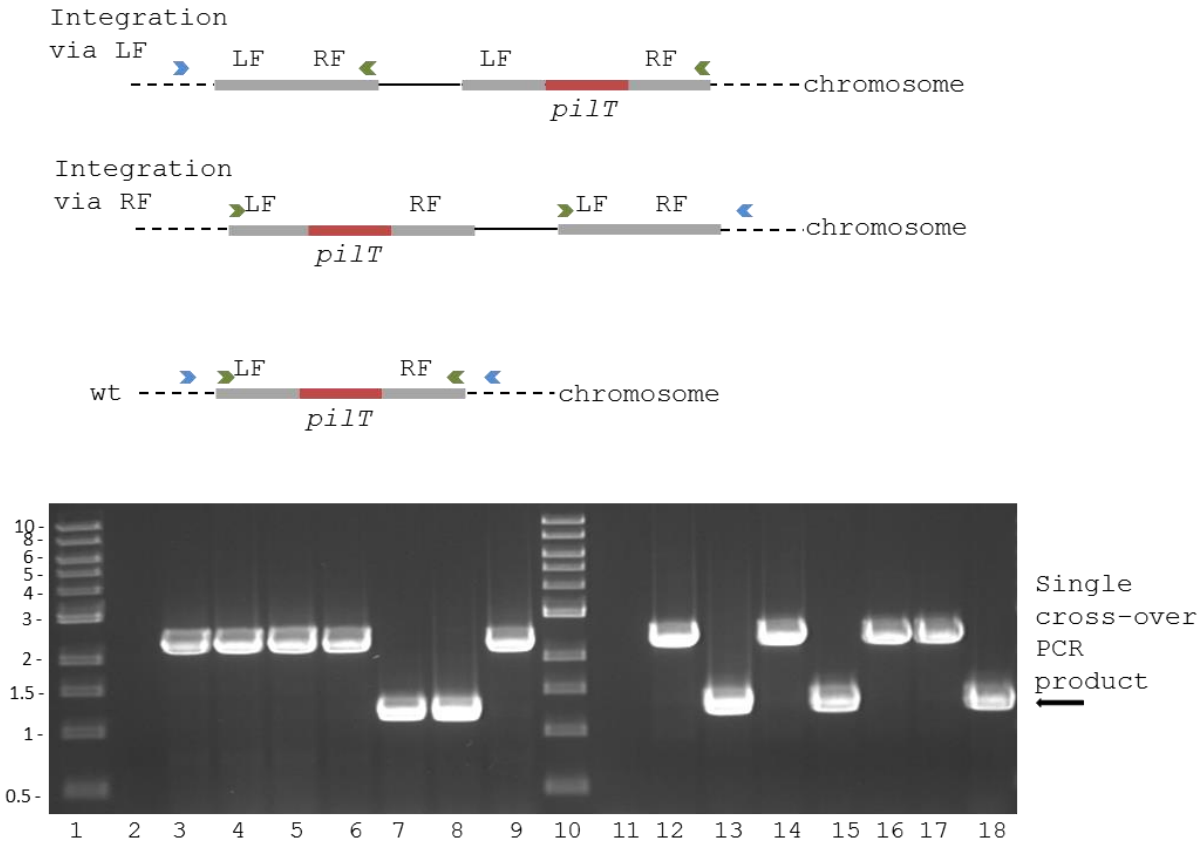


FIG 4.18 PCR screening for single cross-over integration of pES2922

A single cross-over integration of pES2922 into *C. difficile* 630Δ*erm* chromosome was screened by PCR, using primer sets INT3505For (➤) and SphI-CD3505-4 (◀) or SphI-CD3505-1 (➤) and INT3505Rev (◀) with chromosomal DNA of putative merodiploids as template. Note that primers INT3505For and INT3505Rev anneal outwith of the *pilT* flanking regions on the chromosome. Solid line represents plasmid sequence while dashed line represents chromosomal sequence. The amplification of a ~1.2 kb product using either primer set confirmed an integration event. Gel lanes 1 and 10 refer to 1 kb ladder. Lanes 2 and 11 refer to primers SphI-CD3505-4/INT3505For and SphI-CD3505-1/INT3505Rev control PCR, respectively. Lanes 3, 4, 5, 6, 7, 8 and 9 refer to PCR product obtained from 630Δ*erm*, 35/1, 35/3, 35/5, 35/7, 35/8 and 35/10 using primers SphI-CD3505-4/INT3505For, respectively. Lanes 12, 13, 14, 15, 16, 17 and 18 refer to PCR product obtained from 630Δ*erm*, 35/1, 35/3, 35/5, 35/7, 35/8 and 35/10 using primers SphI-CD3505-1/INT3505Rev, respectively.

Introduction of pES288 into merodiploids INT35/5 and INT35/8, screening and confirmation of a second cross-over event:

The I-SceI expression vector pES288 was introduced into the merodiploids INT35/5 and INT35/8 via exponential mating. The putative transconjugants obtained (6 and 7 transconjugant colonies from INT35/5 and INT35/8 matings, respectively) were streaked twice on BHI containing erythromycin at 10 µg/ml (BHIErm10) and colonies from each subculture were then screened for thiamphenicol sensitivity, which is indicative of resolution. Screening for thiamphenicol sensitive strains from the first subculture showed no resolution for either merodiploid strain INT35/5 or INT35/8 (total 12 and 14 colonies screened from INT35/5 and INT35/8, respectively). Screening for thiamphenicol sensitive colonies from the second subculture showed a resolution of 29.2% for merodiploid INT35/5 (19 sensitive colonies from total 65 screened) and a 0% for merodiploid INT35/8 (total 66 colonies screened).

The double cross-over events were confirmed by PCR, using primers SphI-CD3505-1 and SphI-CD3505-4 and genomic DNA from the thiamphenicol sensitive strains isolated above (**FIG 4.19**). The *pilT* deletion cassette was amplified from strains No. 271, 272, 288 and 292. The wt *pilT* region was amplified from resolved strains No. 273, 274, 275, 276, 279, 281, 282, 376, 378, 379, 380, 391, 430, 432 and 434. Subsequent sequencing of the amplified *pilT* deletion cassette, from Δ*pilT* strains No. 271, 272, 288 and 292, confirmed the in-frame replacement of the wt *pilT* allele by the deletion cassette in the chromosome of these strains with the predicted sequence joints (**FIG 4.19**). The I-SceI expression vector pES288 was eliminated from these strains by streaking twice on agar BHI without erythromycin.

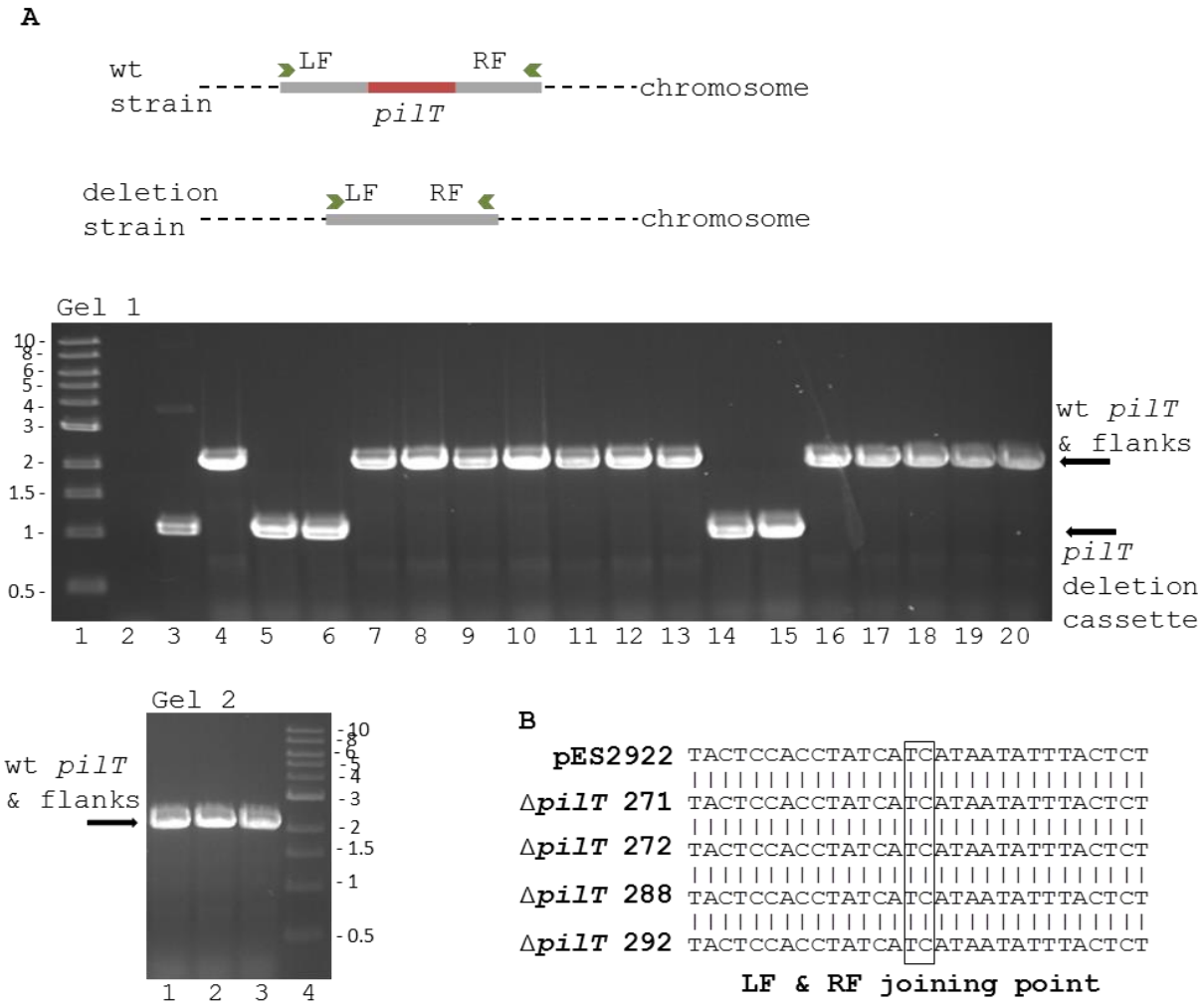


FIG 4.19 Screening and confirmation of second cross-over in Δ*pilT* recombinant strains

A] A second cross-over event was identified by PCR using primers SphI-CD3505-1 (➤) and SphI-CD3505-4 (◄) with chromosomal DNA of putative recombinant strains. The amplification of the *pilT* deletion cassette (~1 kb) identified putative Δ*pilT* recombinant strains, while the amplification of the wt *pilT* including flanking regions (~2 kb) identified strains that resolved to wt. Gel 1 lanes 1, 2 and 3 refer to 1 kb ladder, primer control PCR and product obtained from pES2922, respectively. Gel 1 lanes 5, 6, 14 and 15 refer to PCR product obtained from Δ*pilT* strains No. 271, 272, 288 and 292, respectively. Gel 1 lanes 4, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19 and 20 refer to PCR product obtained from wt strains 630Δ*erm*, 273, 274, 275, 276, 279, 281, 282, 376, 378, 379, 380 and 391, respectively. Gel 2 lane 4 refers to 1 kb ladder. Gel 2 lanes 1, 2 and 3 refer to PCR product obtained from resolved wt strains No. 430, 432 and 434, respectively. **B]** The *pilT* deletion cassettes amplified from the Δ*pilT* strains were sequenced and a partial sequence alignment of the *pilT* LF and RF together with the joining point is shown.

Attempted construction of *C. difficile pilA* deletion strain:

Construction of allele replacement vector pES2923:

Around 500 bp regions flanking the putative *pilA* gene (CD3507) in *C. difficile* 630Δ*erm* were amplified by PCR using genomic DNA as template, with primers SphI-CD3507-1 and CD3507-2 to amplify the *pilA* LF and primers CD3507-3 and SphI-CD3507-4 to amplify the *pilA* RF (**FIG 4.20**). The *pilA* LF and RF were fused together by subsequent cross-over PCR using primers SphI-CD3507-1 and SphI-CD3507-4 to generate the *pilA* deletion cassette (**FIG 4.20**). The deletion cassette was then ligated into the FspI site of pES242 resulting in the allele replacement vector pES2923 (**FIG 4.20**). The vector was sequenced to verify construction.

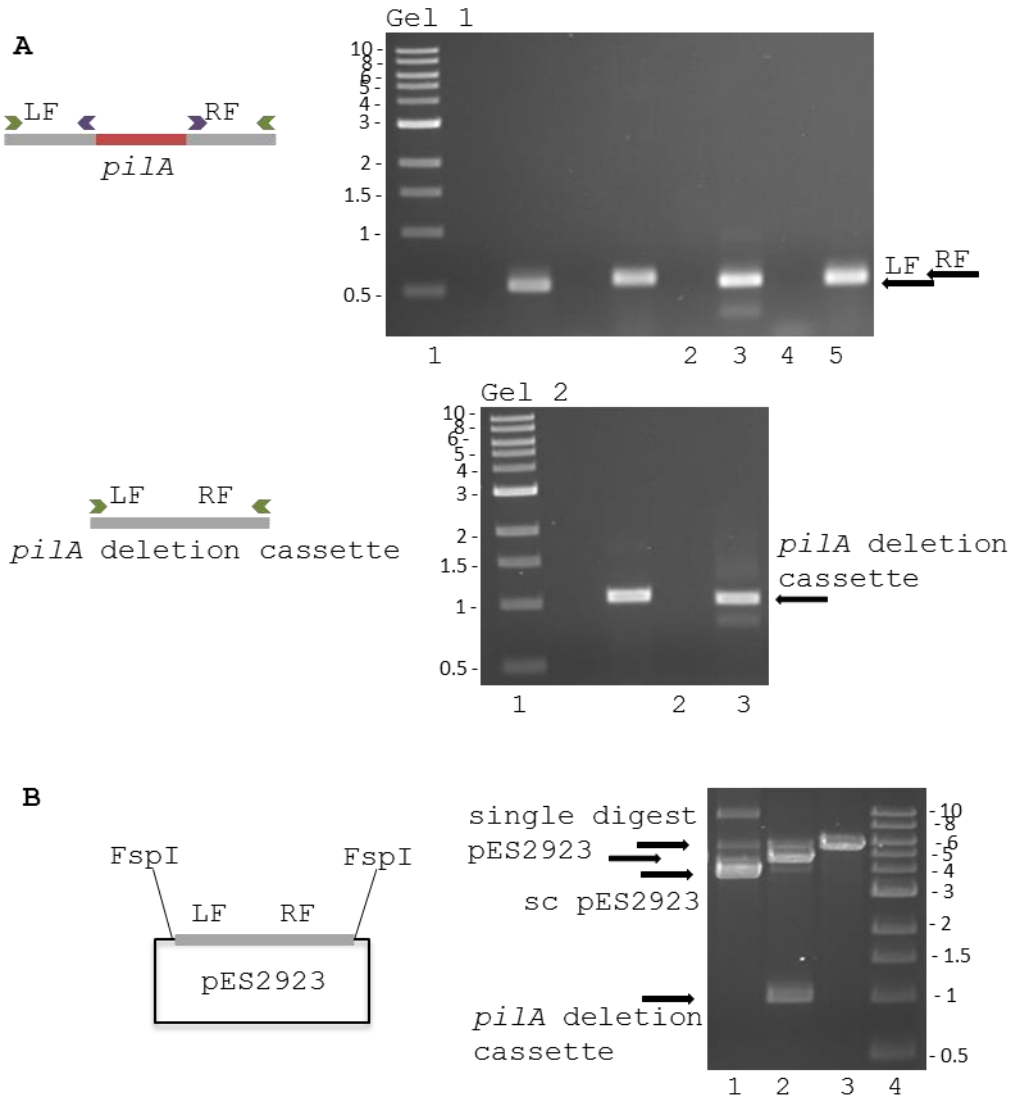


FIG 4.20 Construction of the allele replacement vector pES2923

A] Construction of the *pilA* deletion cassette. Primers SphI-CD3507-1 (➤) and CD3507-2 (◀) were used to amplify by PCR the left flanking region (LF, 521 bp) of the *pilA* gene from 630Δ*erm* genomic DNA. Primers CD3507-3 (➤) and SphI-CD3507-4 (◀) were used to amplify by PCR the right flanking region (RF, 477 bp) of the *pilA* gene. The LF and RF products were fused together by subsequent cross-over PCR using primers SphI-CD3507-1 (➤) and SphI-CD3507-4 (◀) generating the *pilA* deletion cassette (~1 Kb). Gel 1 lane 1 refers to 1 kb ladder. Gel 1 lane 2 and 4 refer to primers SphI-CD3507-1/CD3507-2 and CD3507-3/SphI-CD3507-4 control PCR, respectively. Gel 1 lanes 3 and 5 refer to LF and RF PCR product, respectively. Gel 2 lanes 1, 2 and 3 refer to 1 kb ladder, primers SphI-CD3507-1/SphI-CD3507-4 control PCR and cross-over PCR product, respectively. **B]** Construction and gel analysis of pES2923. The *pilA* deletion cassette was ligated into the FspI site of pES242 generating pES2923. Gel lanes 1, 2, 3 and 4 correspond to supercoiled pES2923, SphI digested pES2923, I-SceI digested pES2923 and 1 kb ladder, respectively.

Introduction of pES2923 into C. difficile 630Δerm, single cross-over and double cross-over event:

The allele replacement vector pES2923 was introduced into *C. difficile 630Δerm* from *E. coli* via stationary mating. One putative transconjugant was obtained (named 41/2), which was inoculated into liquid BHITm15 and also streaked on BHITm15cfx8cls250 agar. The presence of pES2923 in 630Δ*erm* was confirmed by PCR amplification of the *pilA* deletion cassette together with the amplification of the wt *pilA*, including flanking regions, using plasmid DNA extracted from overnight culture of the putative transconjugant with primers SphI-CD3507-1 and SphI-CD3507-4 (**FIG 4.21**). In addition, as observed previously, the overnight incubated agar streaked with the putative pES2923 transconjugant strain 41/2 showed a mixed sized population.

A single cross-over event was enriched and selected as previously, by a second subculture on thiamphenicol containing medium, passage through medium without thiamphenicol and selection of putative single cross-over integrants by plating on thiamphenicol. This resulted in the loss of small sized colonies. Integration of pES2923 into the *C. difficile* 630Δ*erm* chromosome was confirmed via the amplification of a ~1.2 kb sized fragment, by PCR, using genomic DNA of putative partial diploids with primer set INT3507For and SphI-CD3507-4 or SphI-CD3507-1 and INT3507Rev(**FIG 4.22**). The INT3507For and INT3507Rev primers were designed to anneal to regions outwith of the *pilA* flanking regions on the chromosome. This PCR screening confirmed the integration of pES2923 in three merodiploids, named INT41/2, INT41/2/1/2 and INT41/2/2/2 via the *pilA* RF.

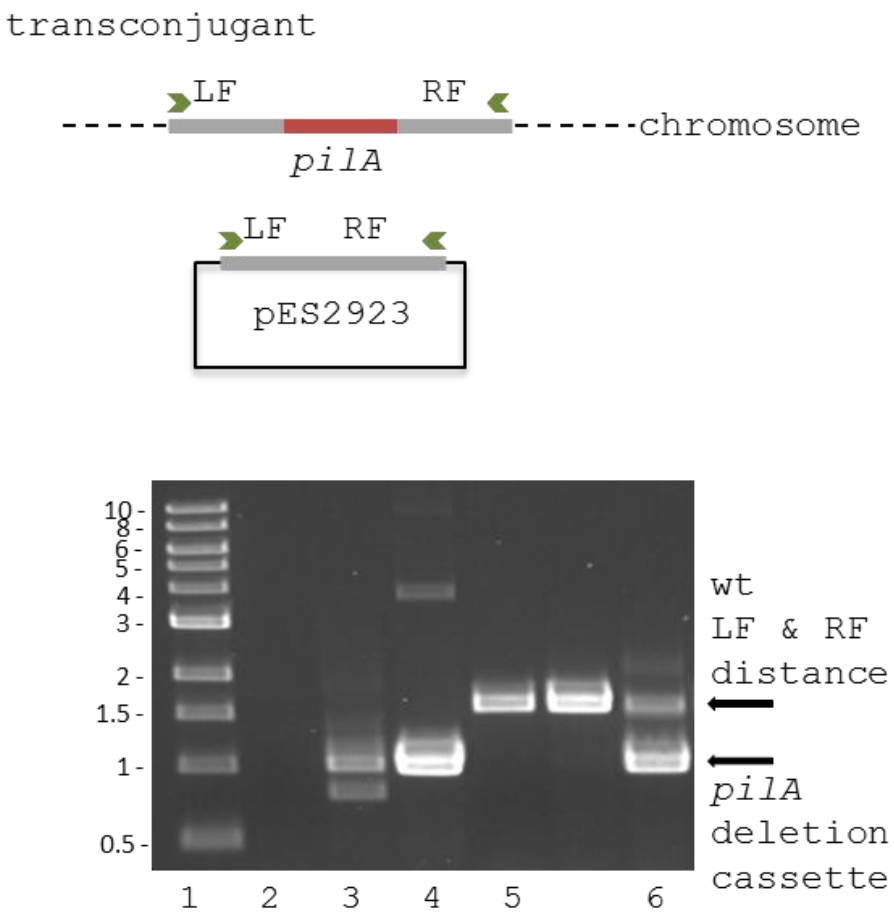


FIG 4.21 PCR screening for pES2923 transconjugants

PCR amplification of the *pilA* deletion cassette together with the wt *pilA* including flanking regions using primers SphI-CD3507-1 (➤) and SphI-CD3507-4 (➤) with plasmid extract as template confirmed the presence of pES2923 in transconjugant strain 41/2. From the wt strain 630Δ*erm* only the wt *pilA* including flanking regions is amplified. Gel lanes 1 and 2 refer to 1 kb ladder and primers control PCR. Lane 3 refers to *pilA* deletion cassette used for ligation. Lane 4, 5 and 6 refer to PCR product obtained from pES2923, 630Δ*erm* and transconjugant 41/2, respectively.

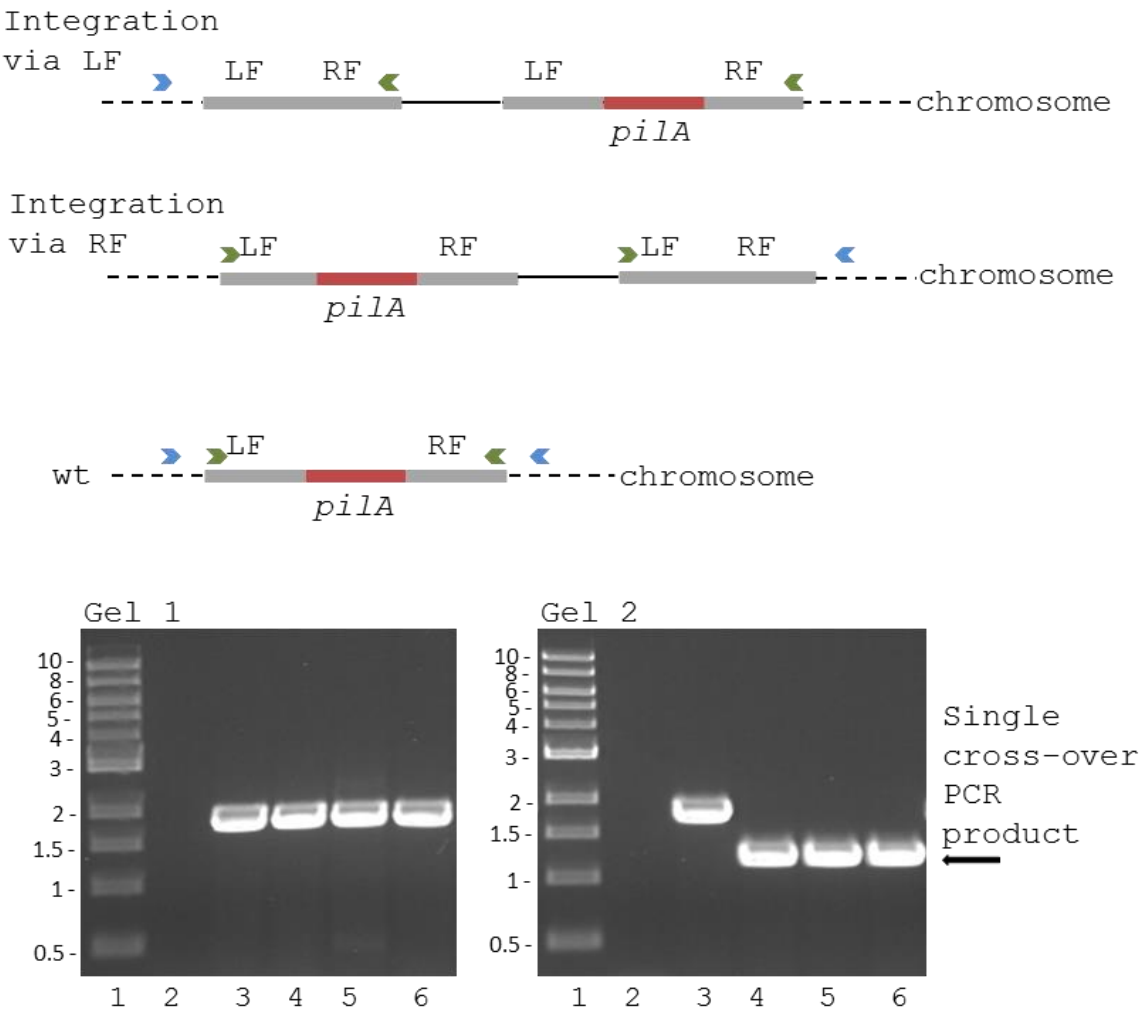


FIG 4.22 PCR screening for single cross-over integration of pES2923

A single cross-over integration of pES2923 into *C. difficile* 630Δ*erm* chromosome was screened by PCR using primer sets INT3507For (➤) and SphI-CD3507-4 (◄) or SphI-CD3507-1 (➤) and INT3507Rev (◄) with chromosomal DNA of putative merodiploids as template. Note that primers INT3507For and INT3507Rev anneal outwith of the *pilA* flanking regions on the chromosome. Solid line represents plasmid sequence while dashed line represents chromosomal sequence. The amplification of a ~1.2 kb product using either primer set confirmed an integration event. Gel 1 refers to PCR products obtained using primers INT3507For and SphI-CD3507-4. Gel 1 lane 1 and 2 refer to 1 kb ladder and primers control PCR, respectively. Gel 1 lanes 3, 4, 5 and 6 refer to PCR products obtained from 630Δ*erm*, INT41/2, INT41/2/1/2 and INT41/2/2/2, respectively. Gel 2 refers to PCR products obtained using primers SphI-CD3507-1 and INT3507Rev. Gel 2 lane 1 and 2 refer to 1 kb ladder and primers control PCR, respectively. Gel 2 lanes 3, 4, 5 and 6 refer to PCR product obtained from 630Δ*erm*, INT41/2, INT41/2/1/2 and INT41/2/2/2, respectively.

The I-SceI expression vector pES288 was introduced into all three merodiploids by exponential mating with *E. coli*. The transconjugants obtained were streaked twice on AIMErm10 and BHIErm10 to enrich for a double cross-over event. However, screening for thiamphenicol sensitive strains from both subcultures gave no resolved colonies. Total colonies screened from sub1 transconjugant cultures of INT41/2, INT41/2/1/2 and INT41/2/2/2 were 35, 17 and 5, respectively. Total colonies screened from sub2 transconjugant cultures of INT41/2, INT41/2/1/2 and INT41/2/2/2 were 154, 110 and 23, respectively. Therefore, in these merodiploids containing pES288, a second cross-over did not occur.

Construction of *C. difficile* Δ*fliC* Δ*pilT*:

Introduction of pES2922 into C. difficile ΔfliC and screening for a single cross-over event:

The allele replacement vector pES2922 which carries the *pilT* deletion cassette was introduced via exponential mating from *E. coli* into the *C. difficile* Δ*fliC* 383 strain isolated previously. From this mating one putative transconjugant was obtained (named 383/35/1) which was inoculated into liquid BHITm15 and also streaked on BHITm15 agar. It was confirmed that this strain contained pES2922 by amplification of the *pilT* deletion cassette, together with the amplification of the wt *pilT* and flanking regions by PCR using primers SphI-CD3505-1 and SphI-CD3505-4 with DNA extracted from the overnight culture as template (**FIG 4.23**). In addition, the overnight incubated BHITm15 agar streaked with the putative transconjugant strain showed

growth of mixed sized colonies, similar to all previously described cases following the introduction of the allele replacement vectors (**FIG 4.24**).

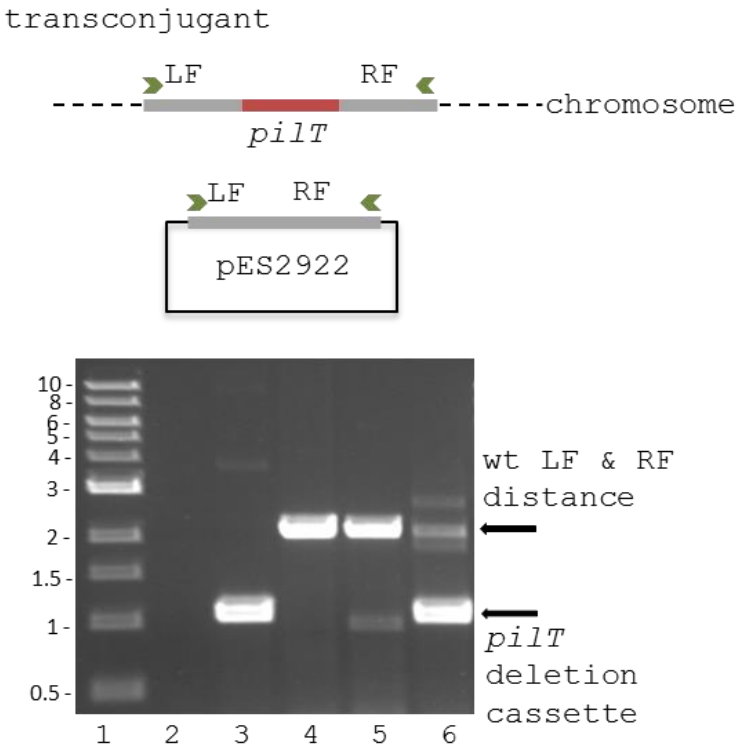


FIG 4.23 PCR screening for pES2922 transconjugants of Δ*fliC* 383

PCR amplification of the *pilT* deletion cassette together with the wt *pilT* including flanking regions using primers SphI-CD3505-1 (➤) and SphI-CD3505-4 (➤) with plasmid extract as template, confirmed the presence of pES2922 in transconjugant strain 383/35/1. From the wt strain 630Δ*erm* only the wt *pilT* including flanking regions is amplified. Gel lanes 1 and 2 refer to 1 kb ladder and primers control PCR, respectively. Lanes 3, 4, 5, and 6 refer to PCR product obtained from pES2922, 630Δ*erm*, Δ*fliC* 383 and 383/35/1, respectively.

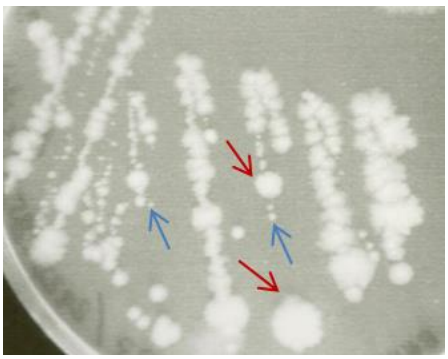


FIG 4.24 Colony size after introduction of pES2922 into *C. difficile* ΔfliC 383

BHITm15 agar streaked with the putative transconjugant strain 383/35/1 showed growth of a mixed sized population. Big colonies possibly represent putative merodiploids and small colonies possibly represent strains harbouring free plasmid. Examples of big and small colonies are indicated with red and blue arrows, respectively.

A single cross-over event was enriched as before, via subculture on thiamphenicol containing medium, passaging through medium lacking thiamphenicol and streaking back on thiamphenicol containing medium. This resulted in the loss of small sized colonies. Single cross-over events were confirmed in four selected thiamphenicol resistant strains via the amplification of a ~1.2 kb sized fragment using primer set SphI-CD3505-4 and INT3505For or SphI-CD3505-1 and INT3505Rev with chromosomal DNA of putative merodiploids as template (**FIG 4.25**). This confirmed the integration of pES2922 via the *pilT* LF into the genome of three merodiploids, named 383/35/INT2, 383/35/INT3 and 383/35/INT4, and also confirmed the integration via the *pilT* RF into the genome of one merodiploid named 383/35/INT1. The merodiploids 383/35/INT1 and 383/35/INT2 were chosen for subsequent resolution.

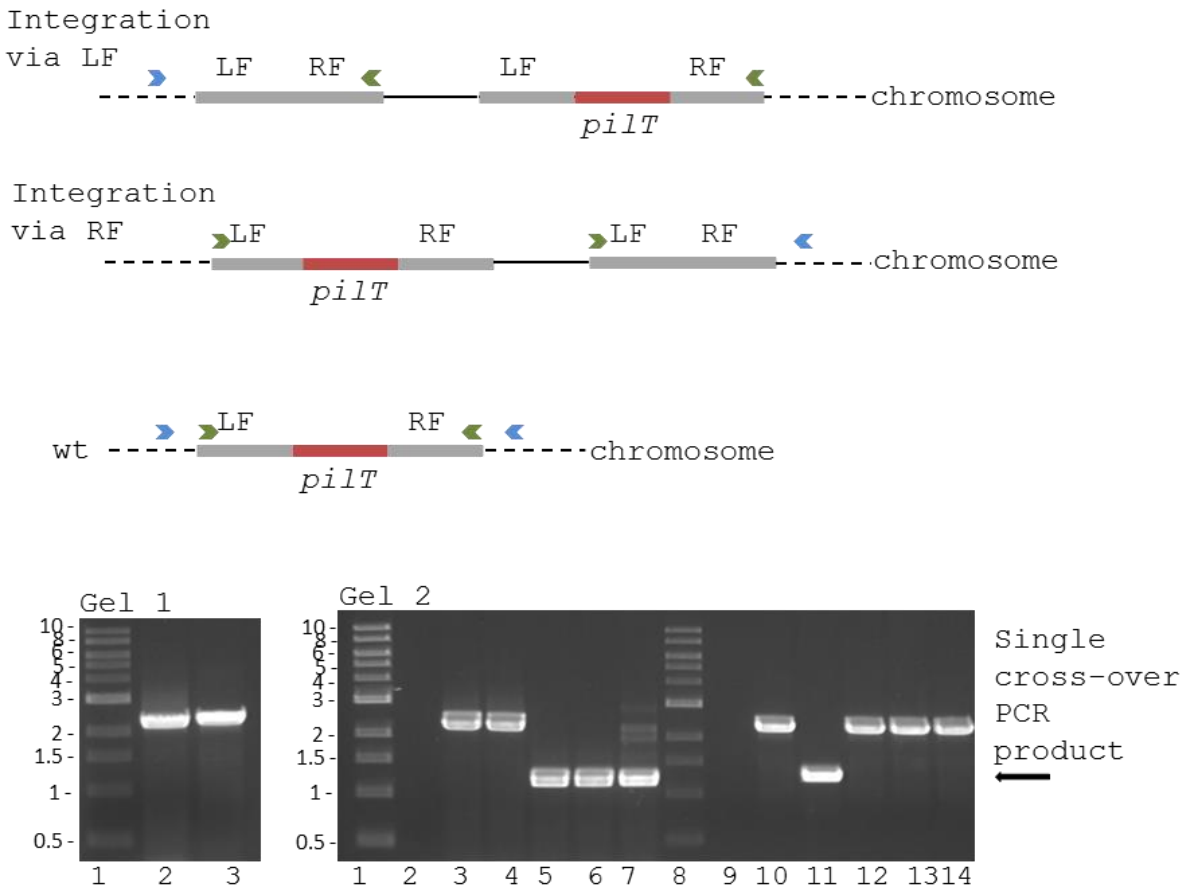


FIG 4.25 PCR screening for single cross-over integration of pES2922 in *C. difficile* Δ*fliC* 383

A single cross-over integration of pES2922 into Δ*fliC* 383 chromosome was screened by PCR using primer sets INT3505For (➤) and SphI-CD3505-4 (◄) or SphI-CD3505-1 (➤) and INT3505Rev (◄) with chromosomal DNA of putative merodiploids as template. Note that primers INT3505For and INT3505Rev anneal outwith of the *pilT* flanking regions on the chromosome. Solid line represents plasmid sequence while dashed line represents chromosomal sequence. The amplification of a ~1.2 kb product using either primer set confirmed an integration event. Gel 1 lane 1 refers to 1 kb ladder. Gel 1 lanes 2 and 3 refer to PCR product obtained from Δ*fliC* 383 using primers SphI-CD3505-4/INT3505For and SphI-CD3505-1/INT3505Rev, respectively. Gel 2 lanes 1 and 8 refer to 1 kb ladder. Gel 2 lanes 2 and 9 refer to primers SphI-CD3505-4/INT3505For and SphI-CD3505-1/INT3505Rev control PCR, respectively. Gel 2 lanes 3, 4, 5, 6 and 7 refer to PCR product obtained from 630Δ*erm*, 383/35/INT1, 383/35/INT2, 383/35/INT3 and 383/35/INT4 using primers SphI-CD3505-4/INT3505For, respectively. Gel 2 lanes 10, 11, 12, 13 and 14 refer to PCR product obtained from 630Δ*erm*, 383/35/INT1, 383/35/INT2, 383/35/INT3 and 383/35/INT4 using primers SphI-CD3505-1/INT3505Rev, respectively.

Introduction of pES288 into merodiploids 383/35/INT1 and 383/35/INT2, screening and confirmation of a second cross-over event:

The I-SceI expression vector pES288 was introduced into the merodiploids 383/35/INT1 and 383/35/INT2 from *E. coli* by exponential mating. Putative transconjugants obtained from the matings were streaked twice on BHIErm10 agar to enrich and screen for double cross-over events. Screening for resolution, via loss of thiamphenicol resistance, from the first subculture yielded no recombinants for 383/35/INT1 (total 80 colonies screened) but 4% resolved colonies from 383/35/INT2 (3 sensitive colonies from total 75 screened). Screening of thiamphenicol sensitive colonies from the second subculture revealed a resolution of 0% from 383/35/INT1 (total 40 colonies screened) and 23.8% from 383/35/INT2 (20 sensitive colonies from total 84 screened).

Double cross-over events, in the isolated thiamphenicol sensitive strains, were confirmed by PCR using primers SphI-CD3505-1 and SphI-CD3505-4 and genomic DNA from the candidate strains (**FIG 4.26**). The *pilT* deletion cassette was amplified from thirteen *C. difficile* putative Δ*flic*Δ*pilT* strains. The *fliC* deletion cassette was also amplified, by PCR, from these strains using SphI-fliC1 and SphI-fliC4 primers and genomic DNA as template. Together these data indicated successful construction of double deletion mutant. Subsequent sequencing of the amplified *pilT* deletion cassette from three chosen strains namely Δ*fliC* Δ*pilT* No. 234, 238 and 250, confirmed the replacement of the wt *pilT* allele by the *pilT* deletion cassette with the predicted sequence joints in the chromosome of these strains (**FIG 4.26**). The wt *pilT* region was amplified from ten resolved strains from which the *fliC* deletion was also amplified, indicating strains that had resolved to the parental genotype. **FIG 4.26** shows the amplified *pilT* and

fliC deletion cassettes from *C. difficile* Δ*fliC* Δ*pilT* No. 234, 238 and 250 and from three resolved strains representing parental *C. difficile* Δ*fliC* 383, No.235, 236 and 249.

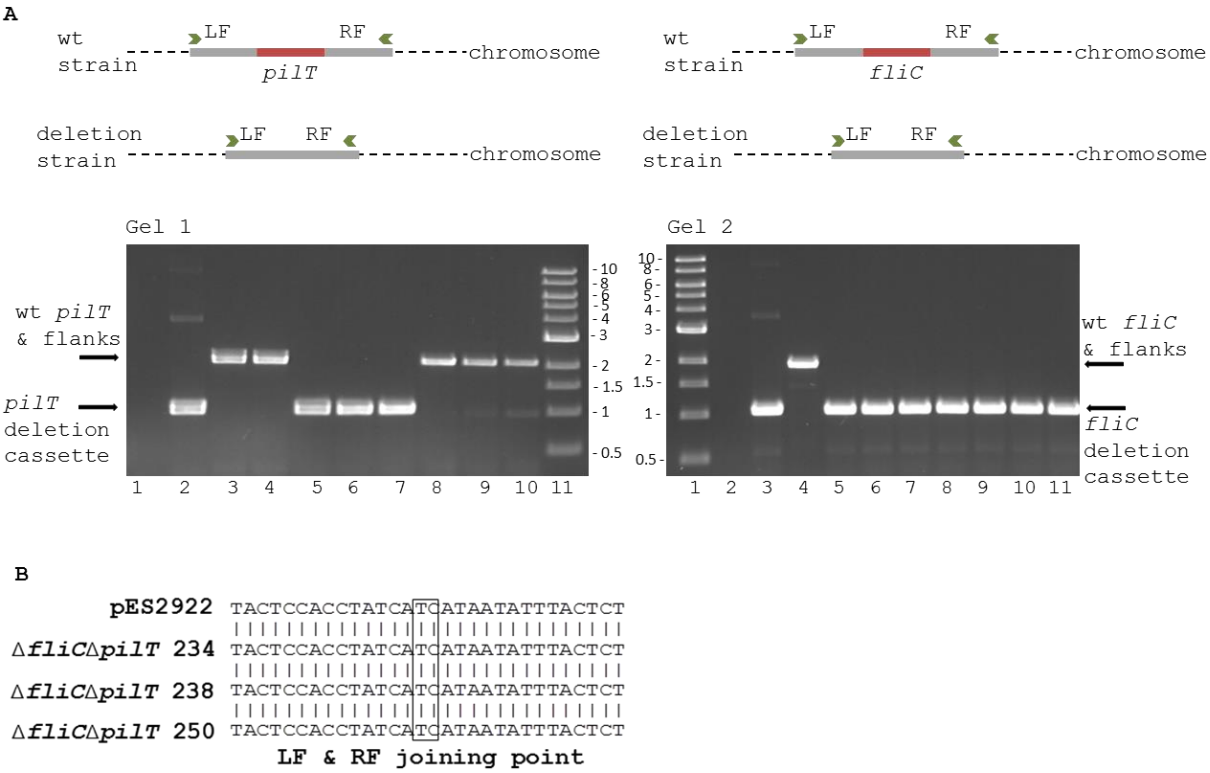


FIG 4.26 Screening and confirmation of second cross-over Δ*fliC* Δ*pilT* recombinants

A] A second cross-over event was identified by PCR using primers SphI-CD3505-1 (➤) and SphI-CD3505-4 (◀) with chromosomal DNA of putative recombinant strains. The amplification of the *pilT* deletion cassette (~1 kb) identified putative Δ*pilT* recombinant strains while the amplification of the wt *pilT* including flanking regions (~2 kb) identified strains that resolved to parental. Gel 1 lanes 1 and 11 refer to primers SphI-CD3505-1/SphI-CD3505-4 control PCR and 1 kb ladder, respectively. Gel 1 lanes 2, 5, 6 and 7 refer to amplification of the *pilT* deletion cassette from pES2922, from putative Δ*fliC* Δ*pilT* No. 234, 238 and 250, respectively. Gel 1 lanes 3, 4, 8, 9 and 10 refer to the amplification of the wt *pilT* including flanking regions from 630Δ*erm*, Δ*fliC* 383, resolved parental No. 235, 236 and 249, respectively. Double deletion mutant strains Δ*fliC* Δ*pilT* and resolved strains back to Δ*fliC* 383 were indicated by the PCR amplification of the *fliC* deletion cassette using primers SphI-fliC1 and SphI-fliC4. Gel 2 lanes 1 and 2 refer to 1 kb ladder and primers SphI-fliC1/SphI-fliC4 control PCR, respectively. Gel 2 lanes 3, 4, 5, 6, 7, 8, 9, 10 and 11

refer to PCR product using primers SphI-fliC1 and SphI-fliC4 obtained from pES2922, 630Δ*erm*, Δ*fliC* 383, Δ*fliC* Δ*pilT* 234, Δ*fliC* Δ*pilT* 238, Δ*fliC* Δ*pilT* 250, resolved Δ*fliC* 383 No. 235, resolved Δ*fliC* 383 No. 236 and resolved Δ*fliC* 383 No. 249, respectively. **B]** The *pilT* deletion cassettes amplified from the Δ*fliC* Δ*pilT* strains were sequenced and a partial sequence alignment of the *pilT* LF and RF together with the joining point is shown.

Phenotypic characterization of the *C. difficile pilT* mutant strains:

Twitching motility:

Bacteria can swim in a liquid medium or swarm across a surface, powered by flagella rotation. They can translocate on solid and semi-solid surfaces via a flagellum-independent form, but type-IV pilus-dependent motion, called twitching motility (Mattick, 2002; Burrows, 2005; Jarrell and McBride, 2008). Twitching motility is unique to TFP and the mechanism is attributed to repeated pilus extension, tethering and retraction resulting in a jerky movement (Henrichsen, 1972; Burrows, 2012; Giltner *et al.*, 2012). Multiple pieces of evidence have demonstrated that PilT is responsible for pilus retraction (Nudleman and Kaiser, 2004; Burrows, 2005). Therefore, twitching motility was investigated in the *pilT* mutant strains and it was hypothesised that these mutants would be impaired in this type of motility. Preliminary tests for twitching motility involved two methods; 1) colony expansion on the surface of semi-solid agar medium, and 2) colony expansion at the interstitial surface between agar and a plastic surface. It has been documented that different factors can influence twitching motility, including nutrients of the growth medium (Henrichsen, 1983; Semmler *et al.*, 1999; Huang *et al.*, 2003; Pamp and Tolker-Nielsen, 2007; Burrows, 2012).

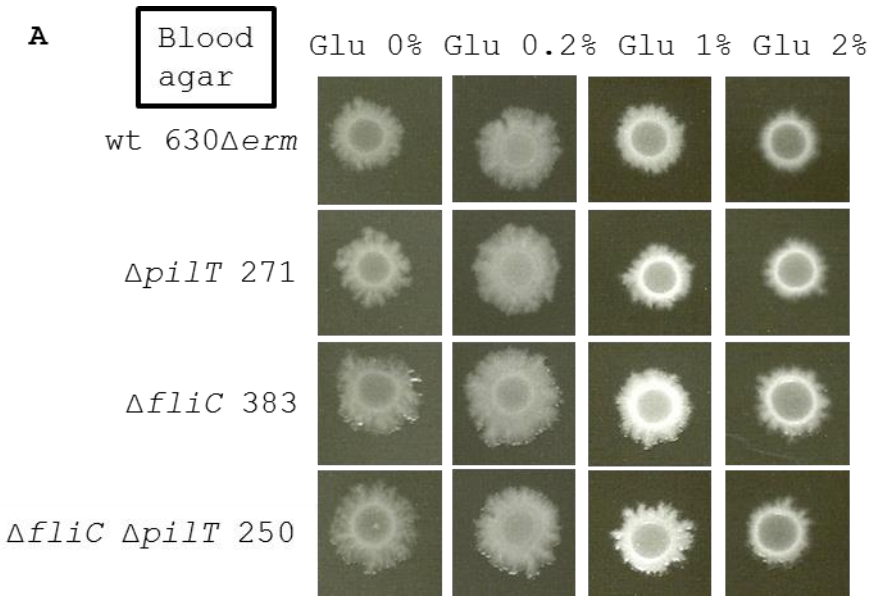
In *C. perfringens*, gliding motility on a semi-solid agar surface is abolished in a *pilT* mutant and *pilT* expression is depressed in the presence of glucose with minimum inhibition observed at 0.5% and maximum inhibition observed at 1% glucose concentration (Varga *et al.*, 2006; Mendez *et al.*, 2008). Therefore, in this study, twitching motility was investigated using AIM (does not contain glucose), BHI (contains 0.2% dextrose) and Blood agar medium without glucose addition or with glucose added at concentrations of 0.2%, 1% or 2%. Twitching motility was investigated in the wt C*. difficile* 630Δ*erm* but also in strains containing Δ*pilT* (Δ*pilT* 271), Δ*fliC* Δ*pilT* (Δ*fliC* Δ*pilT* 250) or Δ*fliC* (Δ*fliC* 383) mutations.

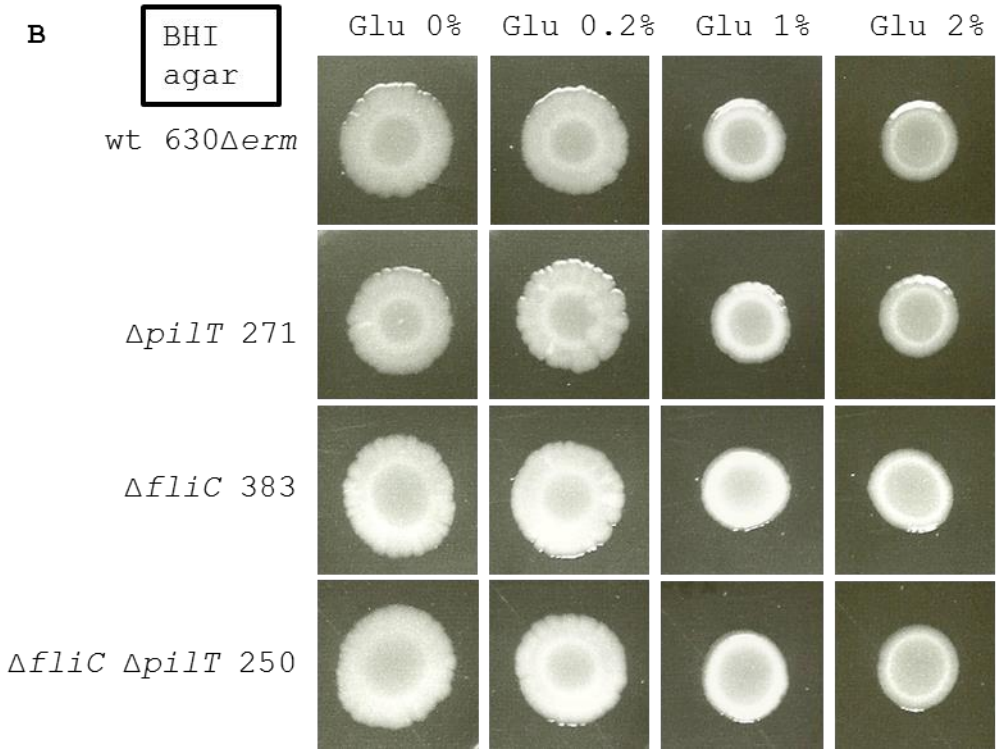
Translocation and morphology on semi-solid surface:

Exponential cultures of the strains were spotted onto AIM, BHI and Blood agar medium with or without glucose and colony morphology and surface expansion of the cultures were recorded after five days of incubation. This was tested at least twice for each growth medium used.

All strains tested (630Δ*erm*, Δ*pilT* 271, Δ*fliC* Δ*pilT* 250, Δ*fliC* 383), behaved similarly between them regarding culture expansion on the different media used (**FIG 4.27**). Under the conditions used in this study, the strains lacking the *pilT* gene did not show any obvious impairment in the surface translocation compared to the parental strain. Addition of glucose to the growth medium had very little effect on the culture surface expansion with reduction of 1 mm on AIM and Blood agar and 2 mm on BHI agar compared to medium without glucose.

The morphology of the colony spots of all strains, although different on different media, displayed the same pattern (**FIG 4.27**). All strains were irregularly shaped and more rhizoid or more undulate shaped at glucose concentrations less than 1%, and had a more entire margin and displayed a more concentrated colony mass at glucose concentrations of 1% and 2%. On Blood agar without glucose the *fliC* deficient strains, Δ*fliC* Δ*pilT* 250 and Δ*fliC* 383, displayed thinner rhizoids than the wt 630Δ*erm* or the Δ*pilT* 271 mutant. On Blood agar with 1% and 2% glucose, all strains displayed a thin outline of cells. On BHI with glucose at 0% and 0.2%, all strains were more undulate shaped than on BHI with 1% glucose, which also displayed a thin outline of cells. All strains on BHI with 2% glucose were not undulate but had an entire margin. Similar to BHI, all strains displayed a more undulate shape on AIM at glucose concentrations of 0% and 0.2%, compared to those containing glucose at 1% and 2%. On AIM with 1% glucose concentration, the wt 630Δ*erm* and Δ*pilT* 271 mutant were more undulate shaped than the *fliC* deficient strains Δ*fliC* Δ*pilT* 250 and Δ*fliC* 383. All strains on this medium displayed a thin outline of cells. Together these data indicate that any differences observed between the strains was due to absence or presence of flagella. Representative examples of all strains on different media are shown in **FIG 4.27**


A
Blood agar
Glu 0%
Glu 0.2%
Glu 1%
Glu 2%
wt 630Δerm
ΔpilT 271
ΔfliC 383
ΔfliC ΔpilT 250


B
BHI agar
Glu 0%
Glu 0.2%
Glu 1%
Glu 2%
wt 630Δerm
ΔpilT 271
ΔfliC 383
ΔfliC ΔpilT 250

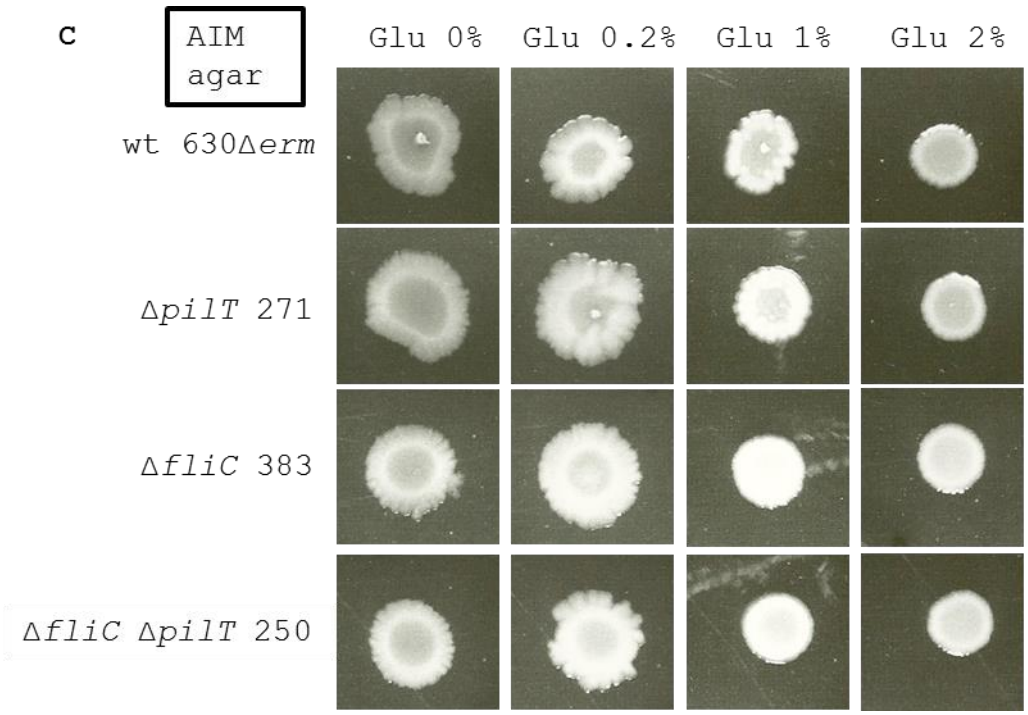


FIG 4.27 Twitching surface motility assay

Photos are shown of five day incubated Blood **(A)**, BHI **(B)** and AIM **(C)** agar(0.7% agar) medium with increasing glucose concentration spotted with exponential cultures of wt 630Δ*erm*, Δ*pilT* 271, Δ*fliC* 383, and Δ*fliC*Δ*pilT* 250. All strains behaved similarly under the conditions tested. Colony expansion was the same between *pilT* deletion and *pilT* proficient strains. Less rhizoid or less wavy shaped and smaller colony spots are produced at glucose concentrations of 1% and 2%.

Subsurface translocation:

Colony expansion at the interstitial surface between agar and a plastic petri dish was measured for strains Δ*pilT* 271, Δ*fliC*Δ*pilT* 250, Δ*fliC* 383 and wt 630Δ*erm*. Measurements were made in triplicate and used Blood or BHI agar with glucose omitted or with glucose at a concentration of 0.2%, 1% and 2%. Colony expansion was the same for all strains with each medium used. As with surface translocation, the *pilT* mutant strains did not show an obvious phenotype.

Moreover, as described above, the addition of glucose at 1% and 2% reduced (1–2mm) the colony expansion of all strains (data not shown).

Colony morphology:

At the macroscopic level all strains, Δ*pilT* 271, Δ*fliC* Δ*pilT* 250, Δ*fliC* 383 and wt 630Δ*erm*, displayed the same colony morphology and colony size when streaked to single colonies on BHI, AIM and Blood agar medium. In addition, when these strains were streaked on BHI or AIM containing 1% glucose they produced the same size colonies as growth without glucose. While the gross morphology was the same, there was a greater biomass present.

Electron microscopy for type IV pilus:

The strain Δ*fliC* 383 was examined for the presence of pili. This strain was chosen for visualisation of pili surface production since it is non-flagellated. Single colonies of the Δ*fliC* 383 strain grown on BHI, Blood, AIM and AIM supplemented with 0.2% fucose (fucose was used to mimic the types of sugars present on the surface of epithelial cells in the GI tract), were examined by EM for the presence of pili. No visible pili were observed on the surface of any of the cells examined, despite a large number of cells being examined. A representative image obtained from the EM of Δ*fliC* 383, grown on AIM, is shown in **FIG 4.10**

4.2.3: Gene regulation of *C. difficile* type IV pilus biosynthesis locus

Introduction and bioinformatics analysis of an intergenic region upstream of the TFP biosynthesis locus:

Under the conditions used in this study the Δ*fliC* 383 strain revealed lack of pili on the surface of the cells, as demonstrated by electron microscopy, which indicates that TFP are regulated and the TFP locus may be repressed at the transcriptional level. In the genome of *C. difficile* 630 there is an intergenic region, with no annotated genes, immediately upstream of the TFP locus. Bioinformatics analysis of this sequence demonstrated no potential open reading frames with potentially significant translation products. **FIG 4.28** shows the TFP biosynthesis locus in the genome of 630 and the 1087 bp upstream intergenic region which is located between the putative pilin protein encoding CD3513 locus and the gene *prs* (ribose-phosphate pyrophosphokinase).

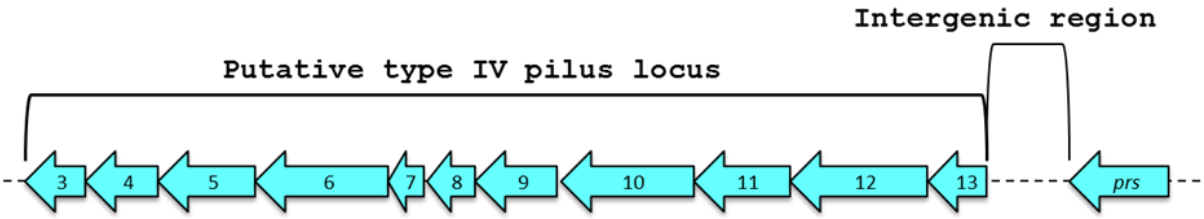


FIG 4.28 ***C. difficile*** **type IV pilus locus and location of the intergenic region**

The eleven type IV pilus associated genes in the CD3503 – CD3513 locus in the *C. difficile* 630 genome are shown. Nomenclature of genes is given in FIG 4.15. The intergenic region (1087 bp) located between the CD3513 locus (13, putative pilin protein) and the *prs* gene (putative ribose-phosphate pyrophosphokinase) is also shown.

Intergenic regions in bacteria are rare and may contribute to gene regulation of downstream genes. Considering the location of this region, we hypothesised that it may be involved in regulation of TFP genes in *C. difficile* 630. Comparison of the intergenic region to the genomes of eight other *C. difficile* strains (R2091, CD196, M68, Bl1, Bl9, M120, 2007855 and CF5) revealed a high percentage of sequence similarity (97% - 99%). Conservation of this region in other strains implies importance. In addition, further computational analysis showed that, the intergenic region in two *C. difficile* strains, R20291 and CD196, is also located upstream of the genes predicted to encode type IV pilus proteins. This further supports the hypothesis that this region may be associated with the regulation of TFP biosynthesis in *C. difficile*.

To identify potential binding sites for regulatory proteins in the intergenic region, computational analysis, using DNA fold, was performed which revealed the presence of inverted repeats (**FIG 4.29**). Downstream of the *prs* gene a putative intrinsic terminator was identified, therefore transcription of the TFP locus would most likely start downstream of this terminator. In addition the region was scanned for putative promoters, using as a reference point the promoters described in *C. difficile* for the rRNA and glutamate dehydrogenase genes (Mani *et al.*, 2006). However, a promoter was not easily identifiable because the intergenic region is AT rich, with a total AT content of 78%. The genome of *C. difficile* 630 features a low CG content of 29.06% (Sebaihia *et al.*, 2006). The aim of this study was to look for promoter activity in the intergenic region. This involved the use of a thiamphenicol expression *catP* reporter system and is described below.

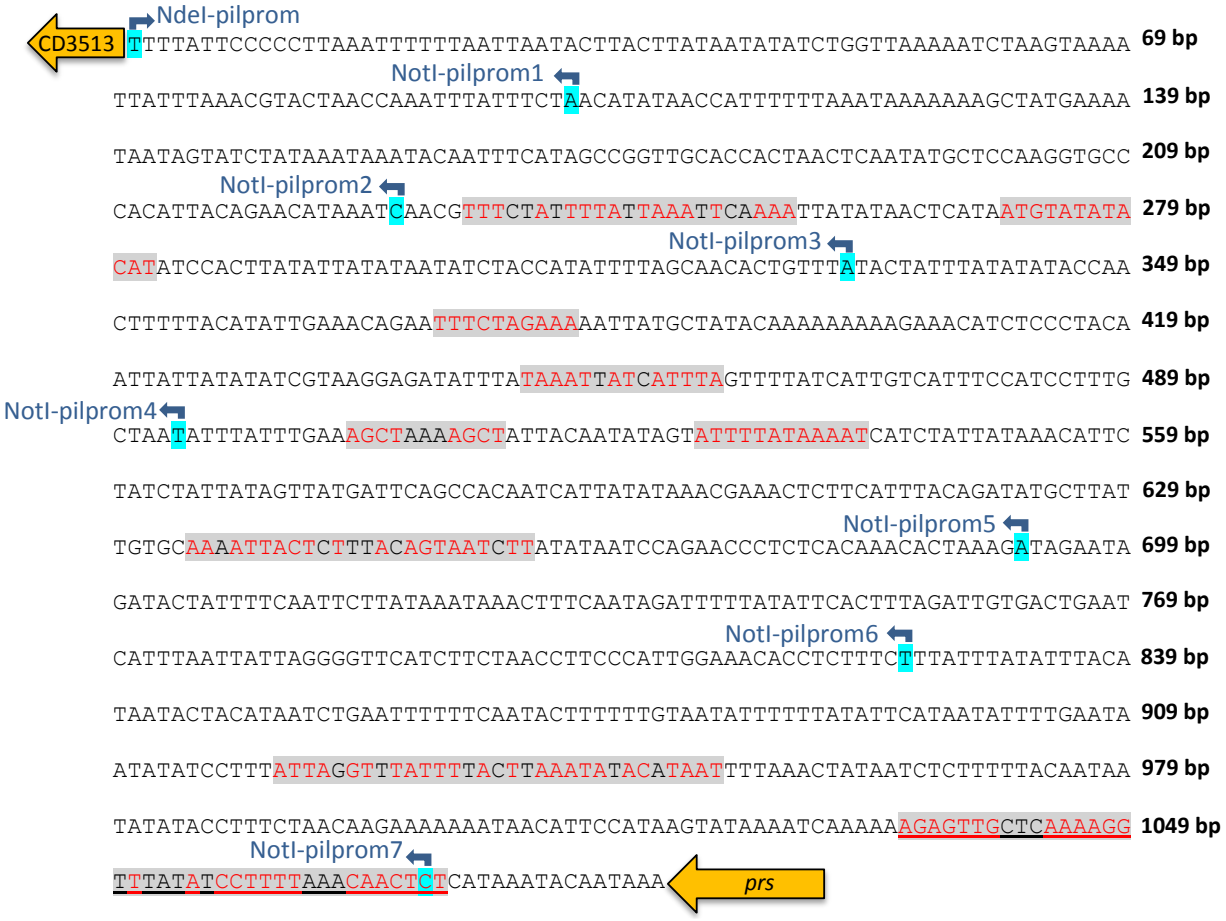


FIG 4.29 Intergenic region sequence

The sequence of the intergenic region between the CD3513 locus and the *prs* gene on the forward strand from the *C. difficile* 630 genome is shown. The direction of transcription of the genes is shown in yellow arrows. The inverted repeats identified in this region are highlighted in grey and complementary sequences are coloured red. The terminator sequence downstream of the *prs* gene is underlined. The 5'-end (blue highlighted bases) and direction of primers (blue arrows) used to amplify fragments of the intergenic region described in FIG 4.30 are indicated. Promoters predicted in the intergenic region using the BPROM (Softberry) program are listed in Appendix I. Nucleotide content of the sequence; A content: 424 (39.0%), C content: 166 (15.27%), G content: 72 (6.62%) and T content: 425 (39.09%).

Construction of the *catP* reporter vectors pES2241-7:

A series of plasmids containing fragments of the intergenic region were constructed using the *catP* reporter plasmid pMTL82254 (Heap *et al.*, 2009). As described in chapter 3, this plasmid is an *E. coli – Clostridium* shuttle vector that carries an erythromycin resistance gene and also contains a promoter-less *catP* gene (Heap *et al.*, 2009).

As shown in **FIG 4.30**, primers were designed to amplify fragments of the intergenic region, starting with the largest fragment at 1071 bp upstream of the methionine codon of the CD3513 gene and creating consecutive deletion PCR fragments ending at 101 bp upstream of the CD3513 gene start codon. The fragments were designated pilprom1 (excluding added restriction sites) (101 bp), pilprom2 (229 bp), pilprom3 (330 bp), pilprom4 (494 bp), pilrpom5 (692 bp), pilrom6 (824 bp) and pilprom7 (1071 bp), and were amplified by PCR using genomic DNA from *C. difficile* 630Δ*erm* as template with primer NdeI-pilprom in combination with primer NotI-pilprom1, NotI-pilprom2, NotI-pilprom3, NotI-pilprom4, NotI-pilprom5, NotI-pilprom6 or NotI-pilprom7 (**FIG 4.30**). These fragments were subsequently ligated into the NotI/NdeI site of pMTL82254 such that each fragment was directly upstream to the promoter-less *catP* gene in this vector. Ligation of fragments pilprom 1 to 7 generated the vectors pES2241 to pES2247 (**FIG 4.30**). The constructs were sequenced to validate construction.

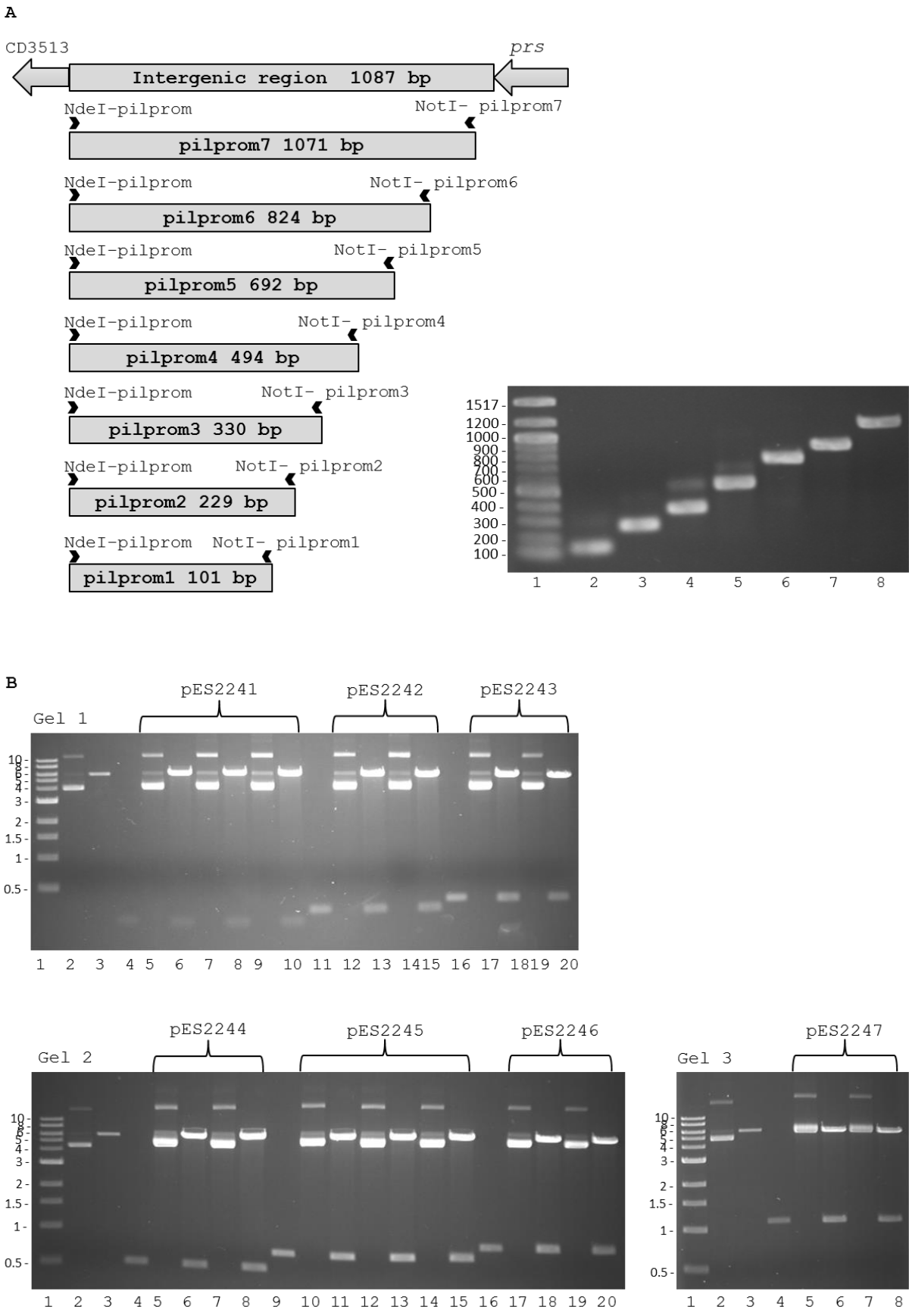


FIG 4.30 Construction of pES2241-7 *catP* reporter recombinant plasmids

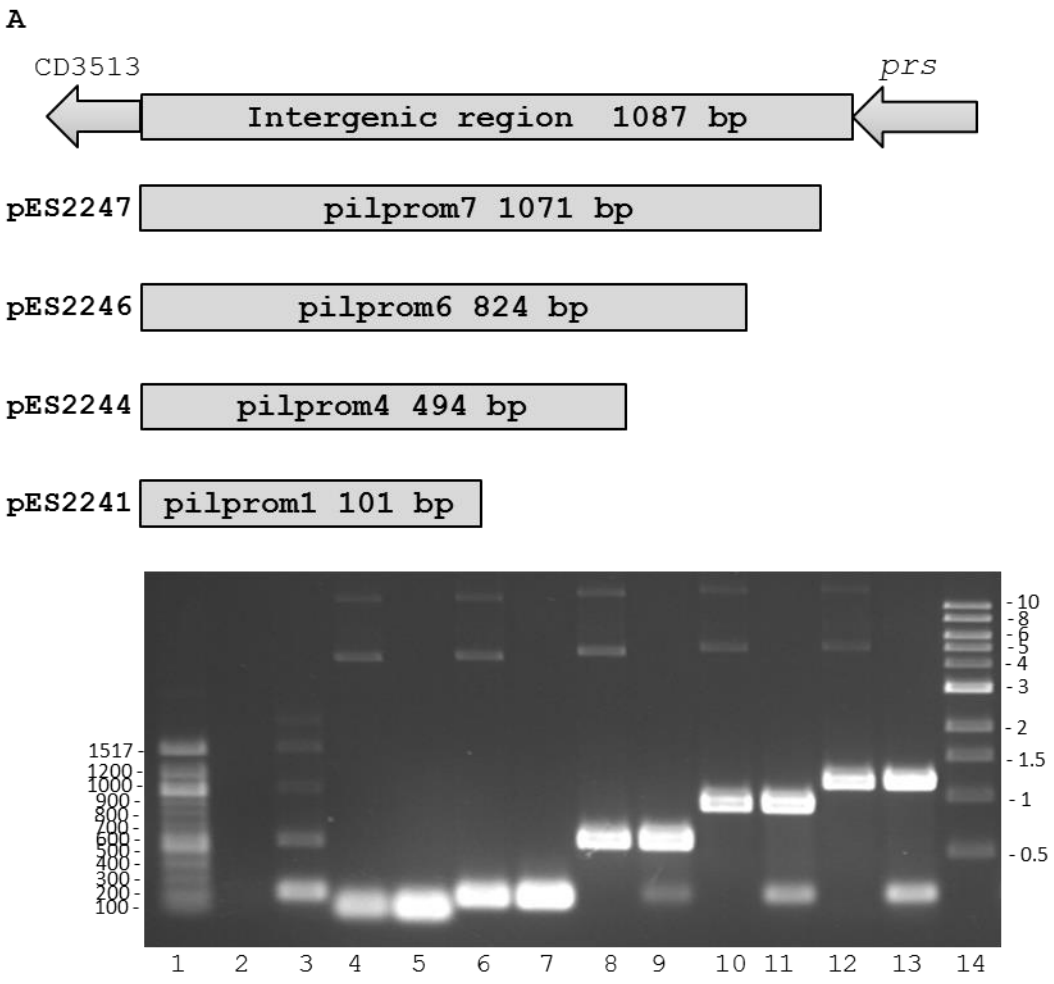
A] Seven deletion fragments pilprom1-7 were amplified by PCR using genomic DNA of *C. difficile* 630Δ*erm* as template and primer NdeI-pilprom in combination with primer NotI-pilprom1, NotI-pilprom2, NotI-pilprom3, NotI-pilprom4, NotI-pilprom5, NotI-pilprom6 or NotI-pilprom7. The sizes of fragments are shown without restriction sites added from primers. Gel lane 1 refers to 100 bp ladder. Gel lanes 2, 3, 4, 5, 6, 7 and 8 refer to PCR fragment pilrpom1, pilprom2, pilprom3, pilprom4, pilrpom5, pilprom6 and pilrom7 obtained, respectively.
B] NotI and NdeI digested pilprom1-7 fragments were ligated into the NotI/NdeI sites of pMTL82254 generating pES2241-7. The gels refer to restriction digest analysis of pES2241-7. Lanes 1, 2 and 3 of all gels refer to 1 kb ladder, supercoiled pMTL82254 and NotI/NdeI digested pMTL82254, respectively. Gel 1 lane 4, 11 and 16 refer to NotI/NdeI digested pilprom1, pilprom2 and pilprom3 fragments, respectively. Gel 1 lanes 5-10 refer to pES2241 from three clones with lanes 5, 7 & 9 representing supercoiled pES2241 and lanes 6, 8 & 10 representing NotI/NdeI digested pES2241. Gel 1 lanes 12-15 refer to pES2242 from two clones with lanes 12 & 14 representing supercoiled pES2242 and lanes 13 & 15 representing NotI/NdeI digested pES2242. Gel 1 lanes 17-20 refer to pES2243 from two clones with lanes 17 & 19 representing supercoiled pES2243 and lanes 18 & 20 representing NotI/NdeI digested pES2243. Gel 2 lanes 4, 9 and 16 refer to NotI/NdeI digested pilprom4, pilprom5 and pilprom6 fragments, respectively. Gel 2 lanes 5-8 refer to pES2244 from two clones with lanes 5 & 7 representing supercoiled pES2244 and lanes 6 & 8 representing NotI/NdeI digested pES2244. Gel 2 lanes 10-15 refer to pES2245 from three clones with lanes 10, 12 & 14 representing supercoiled pES2245 and lanes 11, 13 & 15 representing NotI/NdeI digested pES2245. Gel 2 lanes 17-20 refer to pES2246 from two clones with lanes 17 & 19 representing supercoiled pES2246 and lanes 18 & 20 representing NotI/NdeI digested pES2246. Gel 3 lane 4 refers to NotI/NdeI digested pilprom7 fragment. Gel 3 lanes 5-8 refer to pES2247 from two clones with lanes 5 & 7 representing supercoiled pES2247 and lanes 6 & 8 representing NotI/NdeI digested pES2247.

Conjugation of vectors and expression of the *catP* gene:

From the seven constructs made, four plasmids pES2241, pES2244, pES2246 and pES2247 were successfully introduced into *C. difficile* 630Δ*erm* from *E. coli* by exponential conjugation. For control experiments, the pMTL82254 vector was also conjugated into *C. difficile* 630Δ*erm*. Transconjugants were confirmed via PCR using pMTL82254 specific primers Trans-1 and Trans-2 with plasmid extracts from putative transconjugants (**FIG 4.31**).

An initial qualitative test for promoter activity in the intergenic region involved streaking the strains harbouring pES2241, pES2244, pES2246 and pES2247 to single colonies on growth medium containing thiamphenicol. It was assumed that strains harbouring the recombinant plasmids with a promoter would be thiamphenicol resistant. Single colonies of the *C. difficile* 630Δ*erm* strains containing pES2241, pES2244, pES2246 or pES2247, were streaked on AIMErm10 agar (AIM agar containing erythromycin at 10 µg/ml to select for the plasmid), on AIMErm10Tm15 agar (AIM agar containing erythromycin at 10 µg/ml and thiamphenicol at 15 µg/ml) and on AIMTm15 (AIM agar containing thiamphenicol at 15 µg/ml). As a control, the strains *C. difficile* 630Δ*erm* and *C. difficile* 630Δ*erm* containing pMTL82254 were also streaked to single colonies on the same media. **FIG 4.31** shows photos of 48-h incubated plates. As expected, *C. difficile* 630Δ*erm* containing pMTL82254 grew on the medium containing erythromycin but not on the medium containing thiamphenicol, while *C. difficile* 630Δ*erm* did not grow on any antibiotic containing medium. All *C. difficile* 630Δ*erm* strains containing pES2241, pES2244, pES2246 or pES2247 grew on all media used. The transconjugant strain harbouring pES2241, containing the first 101 bp upstream of the CD3513 locus grew less well and showed patchy colony

growth on thiamphenicol containing medium, particularly on AIMTm15, compared to the rest of the strains. These data suggest that transcription is being initiated within the first 500 bp upstream of the CD3513 locus. The reduced thiamphenicol resistance in the strain harbouring pES2241 may either suggest that a promoter exists 101 bp upstream of the CD3513 codon, but is being repressed or that a promoter does not exist within this region. The latter suggestion would agree with the observation that a promoter sequence was not easily defined within the first 100 bp of the intergenic region when compared to other described *C. difficile* promoters.



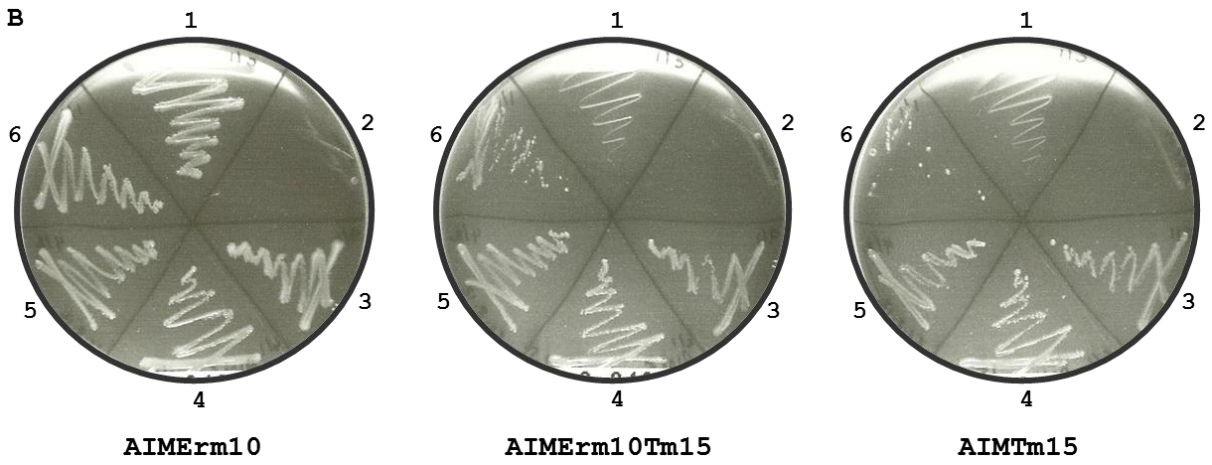


FIG 4.31 Thiamphenicol resistance of pES224 transconjugant strains

A] *C. difficile* 630Δ*erm* strains were confirmed to harbour pMTL82254, pE2241, pES2244, pES2246 or pES2247 by PCR, using pMTL82254 specific primers Trans-1 and Trans-2 which anneal outwith of the NotI/NdeI sites of pMTL82254 and as a template plasmid extracts of putative transconjugants. Gel lanes 1 and 14 refer to 100 bp ladder and 1 kb ladder, respectively. Lane 2 refers to primers control PCR. Lane 3 refers to PCR product obtained from wt *C. difficile* 630Δ*erm* plasmid extract. Lanes 5, 7, 9, 11 and 13 refer to PCR product obtained from plasmid extracts from transconjugant strains containing pMTL82254, pES2241, pES2244, pES2246 and pES22547, respectively. Lanes 4, 6, 8, 10 and 12 refer to pML82254, pES2241, pES2244, pES2246 and pES2247 control PCR product obtained from *E. coli* plasmid extracts, respectively.
B] Photos of 48-h incubated AIMErm10, AIMErm10Tm15 and AIMTm15 agar media streaked with strains wt *C. difficile* 630Δ*erm* (2), *C. difficile* 630Δ*erm* containing pMTL82254 (1), *C. difficile* 630Δ*erm* containing pES2247 (3), *C. difficile* 630Δ*erm* containing pES2246 (4), *C. difficile* 630Δ*erm* containing pES2244 (5) and *C. difficile* 630Δ*erm* containing pES2241 (6). All strains containing pMTL82254-based plasmid grow on AIMErm10. Only strains containing pES224 plasmids grow on AIMErm10Tm15 and AIMTm15. The strain *C. difficile* 630Δ*erm* containing pES2241 shows reduced thiamphenicol resistance with sporadic colony growth particularly on AIMTm15 agar medium.

4.3: DISCUSSION

In this study the optimization of the I-SceI genetic system, particularly the construction of a new I-SceI site delivery vector for the construction of various deletions in *C. difficile*, was described. The I-SceI delivery vector pES242 (chapter 4) and the I-SceI expression vector pES288 (chapter 3) were used to delete the *fliC*, *pilT* and *pilA* genes in *C. difficile*. In this system, as with the genetic system in chapter 3, single cross-over integration was selected via thiamphenicol resistance and second cross-over resolution was screened via loss of thiamphenicol resistance. The use of the optimized system has shown its robustness and reproducibility by construction of the *C. difficile* Δ*fliC* and *C. difficile* Δ*pilT* strains, and construction of the double deletion strain *C. difficile* Δ*fliC* Δ*pilT*. The ability to construct multiple deletions, using this system, is attributed to the fact that resolution results in the complete loss of thiaphenicol resistance and elimination of I-SceI expression vector pES288 is easily achieved by streaking twice on medium without erythromycin.

Despite the recent development of other markerless deletion systems in *C. difficile*, it is important to note that a markerless double gene deletion of *C. difficile* has not been reported previously, supporting the uniqueness of this system over currently used genetic systems in *C. difficile* (Cartman *et al.*, 2012; Ng *et al.*, 2013). In addition, one of the systems uses a *pyrE*-based allelic exchange that requires the construction of an auxotrophic mutant strain before use of the mutagenesis system (Ng *et al.*, 2013). Resulting strains therefore, carry additional mutations

that are not representative of the wild type and may have unforeseen consequences. The system can be used to convert the *pyrE* mutant strain with the modified target gene back to a *pyrE*$^{+}$ strain, but this could add to laboratory work load. Although the widely used ClosTron system includes FRT sites to allow excision of *ermB* via the action of FLP recombinase for the re-use of the system for multiple gene interruptions, the *ermB* gene has never been excised in *C. difficile*, despite multiple attempts. This can be bypassed by using different ClosTron plasmids containing a retrotransposition activated marker (RAM) carrying different antibiotic resistance genes, with the end result being interruption of different genes of interest by different antibiotic genes (Kuehne and Minton, 2012). Nevertheless, this approach still produces inactivation of genes of interest by interruption and thus has potential polar effects and limits the number of genes that can be knocked out.

Introduction of the I-SceI delivery vector into *C. difficile* in all cases, whether *C. difficile* 630Δ*erm* or *C. difficile* Δ*fliC* 383, and in the presence of thiamphenicol selection resulted in a mixed sized colony population (2-3 mm and pinpoint size) where larger colonies represented single cross-over strains and smaller colonies represented strains with free plasmid. The loss of the small sized population in all cases, after single cross-over enrichment, supports the hypothesis that bigger colonies may represent a merodiploid population. This was the predicted phenotype with the use of an allele replacement vector based on a plasmid carrying the pCB102 replicon, as previously observed by two different studies (Heap *et al.*, 2012; Cartman *et al.*, 2012). According to observations from other studies, to enrich for single cross-over by growth

rate under antibiotic selection in other Clostridia, such as *C. difficile* R20291 and *C. sporogenes*, the Gram-positive replicons pBP1 (for *C. difficile* R20291) and pIM13 (for *C. sporogenes*) would be appropriate instead of pCB102 (Cartman *et al.*, 2012; Heap *et al.*, 2012).

Introduction of the I-SceI expression vector pES288 in the co-integrant strains displayed a low resolution, with a maximum of 29%, which increased with subculturing, as previously discussed in chapter 3. However, the resolution observed in this study, as well as described in chapter 3, was due to I-SceI expression and not due to spontaneous second cross-over event as demonstrated by the lack of resolution after the introduction of pMTL82254 in the merodiploid strains INT10/1 and INT10/3 (pES2921 integrated in *C. difficile* 630Δ*erm* chromosome). An intriguing observation was that introduction of pES288 did not always result in resolution of the merodiploid and it seemed that, for a particular gene deletion, a second cross-over event occurred in only one of the merodiploids that resulted from a right flank (RF) or left flank (LF) integration but not necessarily both. For example, when pES288 was introduced into the merodiploids, resolution occurred in both *fliC* INT10/1 and INT10/3 both generated from LF integration, resolution occurred in *pilT* INT35/5 generated from RF integration but not in *pilT* INT35/8 generated from LF integration, and resolution occurred in *fliCpilT* 383/35/INT2 generated by LF integration but not in *fliCpilT* 383/35/INT1 generated by RF integration. Introduction of pES288 into the three *pilA* integrants (INT41/2, 41/2/1/2 and 41/2/2/2) all generated by RF integration, did not result in resolution. It would be intriguing to test if a *pilA* integrant that had recombined via the LF could generate a second cross-over event. The resolution observed

according to integration site may be attributable to the positioning of Chi sites in relation to the integrated construct, and/or the distance that AddAB must translocate before encountering a Chi site.

Deletion of the *fliC* gene in *C. difficile* resulted in a strain incapable of flagella-mediated motility. This phenotype agreed with the phenotype observed previously by two other groups when the *fliC* gene was insertionally inactivated (Twine *et al.*, 2009; Dingle *et al.*, 2011). Introduction of the complementation vector, pES196, into two Δ*fliC* strains, produced cells with a tumbling motility phenotype, as observed by wet mount microscopy. This non-directional motility presumably prevented the strains from penetrating the soft agar medium, despite the EM observation that flagella were produced. The lack of a spreading phenotype in the soft agar was in contrast to the previously reported complemented strain containing an interrupted *fliC* gene were the spreading phenotype was restored (Dingle *et al.*, 2011). The parental *C. difficile* strain containing pES196 showed a typical vigorous motility and a spreading phenotype in the soft agar medium, which excludes the possibility of adverse effects from flagellin overproduction. The observed phenotype of the complemented Δ*fliC* strains may be attributed either to the out of frame deletion of the *fliC* gene or, alternatively, the deletion of the *fliC* gene may have also deleted a region that may be involved in the regulation of flagella-associated genes downstream, resulting in improper flagella function. Maybe a precise amount of FliC or other proteins is required for appropriate flagellar function. Immediately downstream of *fliC* is the CD0240 locus which encodes a glycosyltransferase enzyme. It has been shown that interruption of CD0240 results in cells that display a non-

spreading phenotype in soft motility agar and produce limited amounts of flagellar filaments that are truncated in length (Twine *et al.*, 2009). In addition, it is thought that the rotational force of the flagella motor is generated by the interaction between the rotor component FliG and the stator component MotA (Terashima *et al.*, 2008). In the genome of *C. difficile* 630, genes predicted to encode for FliG and MotA are located downstream of the *fliC* gene and an effect on their transcription and subsequent stoichiometry may be critical for flagellar rotation. Nevertheless, the non-flagellated Δ*fliC* 383 strain was further used to search for presence of pili and twitching motility experiments.

To characterize the isolated Δ*pilT* strains, this study performed preliminary tests to investigate twitching motility. The measurement of colony expansion on the surface of agar or the interstitial surface revealed that all strains, whether Δ*pilT* or *pilT* $^{+}$, expanded to the same size. Moreover, there was no obvious difference in colony morphology between the Δ*pilT* or *pilT* $^{+}$ strains. These observations were in contrast to a previous study on TFP-mediated motility in *C. perfringens*, where surface motility was abolished in a *pilT* mutant (Varga *et al.*, 2006). It should be noted that these tests were only preliminary due to time constrains, and certain factors should be taken into account before the interpretation of results and future work is suggested. First, it is important to note that twitching motility at the macroscopic level varies between different species and *in vitro* culture conditions need to be optimal for twitching to occur (Semmler *et al.*, 1999; Mattick, 2002). For example, in *N. gonorrhoeae* and *N. meningitidis* twitching zones formed from surface-grown colonies are so thin they can only be visualized

microscopically (Wolfgang *et al.*, 1998; Mattick, 2002). *In vitro* culture conditions that have been reported to influence (stimulate or inhibit) twitching motility, particularly the composition of the growth medium include; nutrients or other compounds that influence the osmolality conditions, viscosity, hydrophobicity, tension of the medium surface i.e. tryptone, yeast extract, sucrose, glucose, mucin, bovine serum albumin, salts (e.g. NaCl, KCL), agar concentration, surfactants (e.g. Tween 20)and polyvinylpyrrolidone (Henrichsen, 1983; Semmler *et al.*, 1999; Huang *et al.*, 2003; Pamp and Tolker-Nielsen, 2007; Burrows, 2012). In addition, available humidity and smoothness of the surface are also influencing factors, with the latter being important for the agar-air surface translocation experiments (Henrichsen, 1983; Semmler *et al.*, 1999).

In the current study, both surface and subsurface translocation were tested using various growth media, including Blood, BHI and AIM agar without glucose or with varying glucose concentrations. Plates were all kept in a sealed bag to maintain humid conditions, although for the agar-air surface translocation motility assay the agar was set against air possibly resulting in an unsmooth surface. However, the edges of the expanded colonies, whether grown on the agar surface or subsurface, were not examined microscopically in search of cell protrusions or raft arrangement indicative of twitching cells, or for any other subtle differences that may not be visible to the naked eye. Spotted cultures on the surface of agar medium containing high glucose concentration (1-2%) produced a thin outer cell zone and it would be intriguing to see the arrangement of these cells microscopically. In *P. aeruginosa* (where twitching motility has been extensively

studied), microscopic examination of the edges of the twitching zones at the interstitial surface between a glass coverslip and growth medium, the wt PAK strain which is capable of twitching shows typical twitching zone with large rafts, while the PAK *pilA* mutant strain not capable of twitching does not produce any rafts (Semmler *et al.*, 1999; Huang *et al.*, 2003). In *C. perfringens*, observation of the edges of surface grown colonies by time lapse microscopy revealed the migration of cells as multicellular curvilinear shaped flares which were similar to those produced in *M. xanthus* via TFP-dependent social motility (Varga *et al.*, 2006). These flares are not formed in a *C. perfringens pilT* mutant, but they are formed in a complemented *C. perfringens pilT* mutant (Varga *et al.*, 2006). In a recent study, microscopic visualization of the edges of colonies of *C. difficile* 630, grown between a glass surface and BHI agar medium (0.7% agar)in a glass bottom dish, revealed diffuse edges with directional protrusions which resembled those produced by *C. perfringens* TFP-mediated motility (Reynolds *et al.*, 2011). This may suggest that twitching motility does occur in *C. difficile* and exclude the possibility that even if TFP retraction occurs in *C. difficile* via the putative retraction protein PilT, it may not be involved in twitching motility. This is the case with EPEC where TFP bundle-forming pili can retract but are not involved in twitching motility (Giltner *et al.*, 2012). Perhaps an experiment similar to the one used by Reynolds *et al.* (2011) with the same growth medium and glass bottom dish could be performed with the strains isolated in this study before any conclusions on twitching motility are formed. Further to this, it has been demonstrated that TFP are produced mainly by cells in the outermost zone of the expanding colonies, where twitching motility is actively

taking place, than across the whole colony (Semmler *et al.*, 1999; Henrichsen, 1983). Therefore, in future searches for pilus production on the surface of Δ*pilT* or *pilT* $^{+}$ strains via electron microscopy, samples would be best prepared from the leading edge of the bacterial growth, by flooding the colony edge with a drop of negative stain or other liquid and floating the grid on top of the drop, allowing the attachment and the selection of the most actively twitching cells on the grid. It would be interesting to see if electron microscopy of the Δ*pilT* strains isolated in this study would reveal hyperpiliation or lack of TFP production on the surface of cells. Electron microscopy examination of the strain Δ*fliC* 383, grown on different media, for the presence of pili used samples prepared from the whole colony. Although many cells were examined, none were observed to produce pili, which might indicate that TFP expression is regulated and may only be expressed in certain cells under certain conditions. Previous studies reporting possible TFP expression in *C. difficile* and the presence of type IV pilus-associated genes in the *C. difficile* genome encoding all genes required for TFP biosynthesis and assembly, support this hypothesis (Sebaihia *et al.*, 2006; Stabler *et al.*, 2006; Janvilisri *et al.*, 2009; Goulding *et al.*, 2009; Borriello *et al.*, 1988). It would be intriguing to test whether pili are observed on Δ*fliC* cells at the colony outer edge. Perhaps a future experiment may involve fusing the *pilT* gene to a reporter gene and investigating its expression under different conditions, which may shed light on the optimal conditions required for TFP gene expression.

The last aspect of this study involved the investigation of an intergenic region located upstream of the *C. difficile* TFP-associated genes, which might potentially be involved

in regulation. Preliminary qualitative results suggested that a promoter may exist between 100 bp and 500 bp upstream of the CD3513 locus. The thiamphenicol resistance displayed by the *C. difficile* strains containing pES2244, pES2246 and pES2247 suggested that transcription is initiated in the 500 bp region upstream of the CD3513 locus. This observation contrasts with the lack of pili obtained from the Δ*fliC* 383 electron microscopy data. An explanation for this may be that low levels of *catP* transcription are sufficient for resistance, while a high level of transcription of TFP associated genes may be required for TFP production. In *M. xanthus*, a minimum level of pilin expression is necessary (more than 50% of the maximum wt level) for efficient pilus assembly and motility (Jelsbak and Kaiser, 2005). Further research is required for any conclusions to be formed which includes, the quantification of the *catP* expression under appropriate conditions in the strains containing the *catP* reporter recombinant plasmids and the identification of the 5' transcription start site for the TFP operon.

In *C. perfringens* carbon catabolite repression (CCR), a regulatory mechanism that responds to carbohydrate availability, regulates TFP-dependent gliding motility, where the catabolite control protein (CcpA) plays a central role acting as a repressor or activator (Mendez *et al.*, 2008). In the presence of glucose (1%) gliding motility is inhibited through the CcpA-mediated repression of *pilT* and *pilD* genes, while in the absence of glucose CcpA positively regulates gliding motility. Moreover, in *C. perfringens* TFP-mediated gliding motility is necessary for optimal biofilm formation that is dependent on a functional CcpA protein (Varga *et al.*, 2008). *C. difficile* has all genes involved in CCR and recently, it was shown that CcpA is

involved in the glucose-dependent repression of *C. difficile* toxin genes, *tcdA* and *tcdB*, by binding directly to their regulatory regions (Antunes *et al.*, 2011). Therefore, questions that arise are, whether the intergenic region is subject to glucose catabolism and whether CcpA has a regulatory role in the expression of TFP genes in *C. difficile*.

A molecule that is also considered as a candidate of future investigation for the regulation of the intergenic region and TFP genes is the second messenger c-di-GMP. C-di-GMP is synthesized by diguanylate cyclases (DGCs) that contain a GGDEF domain and is degraded by phoshodiesterases (PDEs) that contain an EAL or HD-GYP domain (Hengge, 2009). *C. difficile* encodes a large number of proteins, predicted to be involved in c-di-GMP metabolism. The enzymatic functionality of many of the homologues was confirmed in *V. cholerae* (Bordeleau *et al.*, 2011). The high number of functional DGCs and PDEs, encoded by conserved genes in the genomes of different *C. difficile* strains, indicates that c-di-GMP signalling is important in this organism and might play a role in regulating diverse phenotypes (Bordeleau *et al.*, 2011). It has been shown that c-di-GMP regulates flagella-mediated motility in *C. difficile* (Purcell *et al.*, 2012; Sudarsan *et al.*, 2008). A c-di-GMP riboswitch is located upstream of the large *flgB* flagella operon and this riboswitch functions as an 'off switch' in response to c-di-GMP (Sudarsan *et al.*, 2008). C-di-GMP decreases flagella-mediated motility and represses production of flagella by reducing transcription of flagellar genes (Sudarsan *et al.*, 2008; Purcell *et al.*, 2012). As mentioned, c-di-GMP regulates aggregate formation in *C. difficile* where aggregates are bound by structures that resemble pili (Purcell *et al.*, 2012). Two c-di-GMP-related

proteins, namely FimX and PilZ, involved in TFP biogenesis have been identified (Burrows, 2012; Guzzo *et al.*, 2009; Huang *et al.*, 2003; Kazmierczak *et al.*, 2006). In *P. aeruginosa*, FimX possesses both a GGEF and EAL domain and is implicated in TFP assembly and twitching motility (Kazmierczak *et al.*, 2006; Huang *et al.*, 2003). PilZ is required for TFP biogenesis and has various TFP-associated functions in different bacterial species including *P. aeruginosa*, *N. meningitides* and *Xanthomonas campestris* (Guzzo *et al.*, 2009). In *Xanthomonas* spp. PilZ binds to PilB (a pilin polymerase) and to the EAL domain of XAC2398 (a homolog of the TFP regulatory protein FimX of *P. aeruginosa*) and has been suggested that PilZ and FimX might regulate the function of PilB in a c-di-GMP-dependent manner (Guzzo *et al.*, 2009).

To conclude, TFP are important to the survival and success of many pathogens in infection and their surface location makes them targets for vaccine development. TFP in *C. difficile* may be expressed *in vivo* during infection and may be involved in biofilm formation and colonization (Goulding *et al.*, 2009; Purcell *et al.*, 2012). Therefore, pursuing further research to investigate the function of these pili in *C. difficile* (e.g. adhesion assays, biofilm formation) could be important in shedding light on the disease pathogenesis and treatment.

REFERENCES

Akerlund T, Svenungsson B, Lagergren A and Burman L G 2006 Correlation of disease severity with fecal toxin levels in patients with Clostridium difficile-associated diarrhea and distribution of PCR ribotypes and toxin yields in vitro of corresponding isolates *Journal of clinical microbiology* **44** 353-8

Alm R A and Mattick J S 1995 Identification of a gene, pilV, required for type 4 fimbrial biogenesis in Pseudomonas aeruginosa, whose product possesses a pre-pilin-like leader sequence *Molecular microbiology* **16** 485-96

Alm R A and Mattick J S 1996 Identification of two genes with prepilin-like leader sequences involved in type 4 fimbrial biogenesis in Pseudomonas aeruginosa *J Bacteriol* **178** 3809-17

Amundsen S K, Fero J, Hansen L M, Cromie G A, Solnick J V, Smith G R and Salama N R 2008 Helicobacter pylori AddAB helicase-nuclease and RecA promote recombination-related DNA repair and survival during stomach colonization *Molecular microbiology* **69** 994-1007

Amundsen S K, Taylor A F and Smith G R 2000 The RecD subunit of the Escherichia coli RecBCD enzyme inhibits RecA loading, homologous recombination, and DNA repair *Proceedings of the National Academy of Sciences of the United States of America* **97** 7399-404

Ananthakrishnan A N 2011 Clostridium difficile infection: epidemiology, risk factors and management *Nature reviews. Gastroenterology & hepatology* **8** 17-26

Anderson D G and Kowalczykowski S C 1997 The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner *Cell* **90** 77-86

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

Ayora S, Carrasco B, Cardenas P P, Cesar C E, Canas C, Yadav T, Marchisone C and Alonso J C 2011 Double-strand break repair in bacteria: a view from Bacillus subtilis *FEMS microbiology reviews* **35** 1055-81

Balakrishna A M, Saxena A M, Mok H Y and Swaminathan K 2009 Structural basis of typhoid: Salmonella typhi type IVb pilin (PilS) and cystic fibrosis transmembrane conductance regulator interaction *Proteins* **77** 253-61

Barketi-Klai A, Hoys S, Lambert-Bordes S, Collignon A and Kansau I 2011 Role of fibronectin-binding protein A in Clostridium difficile intestinal colonization *Journal of medical microbiology* **60** 1155-61

Barth H, Pfeifer G, Hofmann F, Maier E, Benz R and Aktories K 2001 Low pH-induced formation of ion channels by clostridium difficile toxin B in target cells *The Journal of biological chemistry* **276** 10670-6

Bartlett J G and Gerding D N 2008 Clinical recognition and diagnosis of Clostridium difficile infection *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **46 Suppl 1** S12-8

Bartlett J G, Moon N, Chang T W, Taylor N and Onderdonk A B 1978 Role of Clostridium difficile in antibiotic-associated pseudomembranous colitis *Gastroenterology* **75** 778-82

Bejerano-Sagie M, Oppenheimer-Shaanan Y, Berlatzky I, Rouvinski A, Meyerovich M and Ben-Yehuda S 2006 A checkpoint protein that scans the chromosome for damage at the start of sporulation in Bacillus subtilis *Cell* **125** 679-90

Belfort M and Roberts R J 1997 Homing endonucleases: keeping the house in order *Nucleic acids research* **25** 3379-88

Benzinger R, Enquist L W and Skalka A 1975 Transfection of Escherichia coli spheroplasts. V. Activity of recBC nuclease in rec+ and rec minus spheroplasts measured with different forms of bacteriophage DNA *Journal of virology* **15** 861-71

Beylot B and Spassky A 2001 Chemical probing shows that the intron-encoded endonuclease I-SceI distorts DNA through binding in monomeric form to its homing site *The Journal of biological chemistry* **276** 25243-53

Bhaya D, Bianco N R, Bryant D and Grossman A 2000 Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp. PCC6803 *Molecular microbiology* **37** 941-51

Bidle K A and Bartlett D H 1999 RecD function is required for high-pressure growth of a deep-sea bacterium *J Bacteriol* **181** 2330-7

Bieber D, Ramer S W, Wu C Y, Murray W J, Tobe T, Fernandez R and Schoolnik G K 1998 Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic Escherichia coli *Science* **280** 2114-8

Bitter W, Koster M, Latijnhouwers M, de Cock H and Tommassen J 1998 Formation of oligomeric rings by XcpQ and PilQ, which are involved in protein transport across the outer membrane of Pseudomonas aeruginosa *Molecular microbiology* **27** 209-19

Bonura T and Smith K C 1975 Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of Escherichia coli K-12 *J Bacteriol* **121** 511-7

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

Borriello S P, Davies H A and Barclay F E 1988 Detection of fimbriae amongst strains of *Clostridium difficile FEMS Microbiol Lett* **49** 65-7

Bowater R and Doherty A J 2006 Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining *PLoS genetics* **2** e8

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

Buchmeier N A, Libby S J, Xu Y, Loewen P C, Switala J, Guiney D G and Fang F C 1995 DNA repair is more important than catalase for Salmonella virulence in mice *The Journal of clinical investigation* **95** 1047-53

Buchmeier N A, Lipps C J, So M Y and Heffron F 1993 Recombination-deficient mutants of Salmonella typhimurium are avirulent and sensitive to the oxidative burst of macrophages *Molecular microbiology* **7** 933-6

Burkholder W F, Kurtser I and Grossman A D 2001 Replication initiation proteins regulate a developmental checkpoint in Bacillus subtilis *Cell* **104** 269-79

Burns D A, Heap J T and Minton N P 2010 SleC is essential for germination of Clostridium difficile spores in nutrient-rich medium supplemented with the bile salt taurocholate *J Bacteriol* **192** 657-64

Burrows L L 2005 Weapons of mass retraction *Molecular microbiology* **57** 878-88

Burrows L L 2012 Pseudomonas aeruginosa twitching motility: type IV pili in action *Annual review of microbiology* **66** 493-520

Calabi E, Calabi F, Phillips A D and Fairweather N F 2002 Binding of Clostridium difficile surface layer proteins to gastrointestinal tissues *Infect Immun* **70** 5770-8

Cano D A, Pucciarelli M G, Garcia-del Portillo F and Casadesus J 2002 Role of the RecBCD Recombination Pathway in Salmonella Virulence *Journal of Bacteriology* **184** 592-5

Capaldo F N, Ramsey G and Barbour S D 1974 Analysis of the growth of recombination-deficient strains of Escherichia coli K-12 *J Bacteriol* **118** 242-9

Carroll K C and Bartlett J G 2011 Biology of Clostridium difficile: implications for epidemiology and diagnosis *Annual review of microbiology* **65** 501-21

Carter G P, Douce G R, Govind R, Howarth P M, Mackin K E, Spencer J, Buckley A M, Antunes A, Kotsanas D, Jenkin G A, Dupuy B, Rood J I and Lyras D 2011 The anti-sigma factor TcdC modulates hypervirulence in an epidemic BI/NAP1/027 clinical isolate of Clostridium difficile *PLoS pathogens* **7** e1002317

Carter G P, Lyras D, Allen D L, Mackin K E, Howarth P M, O'Connor J R and Rood J I 2007 Binary toxin production in Clostridium difficile is regulated by CdtR, a LytTR family response regulator *J Bacteriol* **189** 7290-301

Carter G P, Lyras D, Poon R, Howarth P M and Rood J I 2010 Methods for gene cloning and targeted mutagenesis *Methods Mol Biol* **646** 183-201

Carter G P, Purdy D, Williams P and Minton N P 2005 Quorum sensing in Clostridium difficile: analysis of a luxS-type signalling system *Journal of medical microbiology* **54** 119-27

Carter G P, Rood J I and Lyras D 2012 The role of toxin A and toxin B in the virulence of Clostridium difficile *Trends Microbiol* **20** 21-9

Cartman S T, Heap J T, Kuehne S A, Cockayne A and Minton N P 2010 The emergence of 'hypervirulence' in Clostridium difficile *International journal of medical microbiology : IJMM* **300** 387-95

Cartman S T, Kelly M L, Heeg D, Heap J T and 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** 4683-90

Chedin F, Ehrlich S D and Kowalczykowski S C 2000 The Bacillus subtilis AddAB helicase/nuclease is regulated by its cognate Chi sequence in vitro *Journal of molecular biology* **298** 7-20

Chedin F, Handa N, Dillingham M S and Kowalczykowski S C 2006 The AddAB helicase/nuclease forms a stable complex with its cognate chi sequence during translocation *The Journal of biological chemistry* **281** 18610-7

Chedin F and Kowalczykowski S C 2002 A novel family of regulated helicases/nucleases from Gram-positive bacteria: insights into the initiation of DNA recombination *Molecular microbiology* **43** 823-34

Chi E, Mehl T, Nunn D and Lory S 1991 Interaction of Pseudomonas aeruginosa with A549 pneumocyte cells *Infect Immun* **59** 822-8

Chiang P and Burrows L L 2003 Biofilm formation by hyperpiliated mutants of Pseudomonas aeruginosa *J Bacteriol* **185** 2374-8

Cohen S H, Tang Y J and Silva J, Jr. 2000 Analysis of the pathogenicity locus in Clostridium difficile strains *The Journal of infectious diseases* **181** 659-63

Coleman W H, Chen D, Li Y Q, Cowan A E and Setlow P 2007 How moist heat kills spores of Bacillus subtilis *J Bacteriol* **189** 8458-66

Collins R F, Davidsen L, Derrick J P, Ford R C and Tonjum T 2001 Analysis of the PilQ secretin from Neisseria meningitidis by transmission electron microscopy reveals a dodecameric quaternary structure *J Bacteriol* **183** 3825-32

Collins R F, Frye S A, Kitmitto A, Ford R C, Tonjum T and Derrick J P 2004 Structure of the Neisseria meningitidis outer membrane PilQ secretin complex at 12 A resolution *The Journal of biological chemistry* **279** 39750-6

Comolli J C, Hauser A R, Waite L, Whitchurch C B, Mattick J S and Engel J N 1999 Pseudomonas aeruginosa gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia *Infect Immun* **67** 3625-30

Connelly J C, Kirkham L A and Leach D R 1998 The SbcCD nuclease of Escherichia coli is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA *Proceedings of the National Academy of Sciences of the United States of America* **95** 7969-74

Cox M M 2001 Recombinational DNA repair of damaged replication forks in Escherichia coli: questions *Annual review of genetics* **35** 53-82

Craig L and Li J 2008 Type IV pili: paradoxes in form and function *Current opinion in structural biology* **18** 267-77

Craig L, Pique M E and Tainer J A 2004 Type IV pilus structure and bacterial pathogenicity *Nature reviews. Microbiology* **2** 363-78

Craig L, Volkmann N, Arvai A S, Pique M E, Yeager M, Egelman E H and Tainer J A 2006 Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions *Molecular cell* **23** 651-62

Cromie G A 2009 Phylogenetic ubiquity and shuffling of the bacterial RecBCD and AddAB recombination complexes *J Bacteriol* **191** 5076-84

Curry S R, Marsh J W, Muto C A, O'Leary M M, Pasculle A W and Harrison L H 2007 tcdC genotypes associated with severe TcdC truncation in an epidemic clone and other strains of Clostridium difficile *Journal of clinical microbiology* **45** 215-21

Dachs G U, Abratt V R and Woods D R 1995 Mode of action of metronidazole and a Bacteroides fragilis metA resistance gene in Escherichia coli *The Journal of antimicrobial chemotherapy* **35** 483-96

Dajkovic A, Mukherjee A and Lutkenhaus J 2008 Investigation of regulation of FtsZ assembly by SulA and development of a model for FtsZ polymerization *J Bacteriol* **190** 2513-26

Davies A H, Roberts A K, Shone C C and Acharya K R 2011 Super toxins from a super bug: structure and function of Clostridium difficile toxins *The Biochemical journal* **436** 517-26

De Septenville A L, Duigou S, Boubakri H and Michel B 2012 Replication fork reversal after replication-transcription collision *PLoS genetics* **8** e1002622

Debast S B, Vaessen N, Choudry A, Wiegers-Ligtvoet E A, van den Berg R J and Kuijper E J 2009 Successful combat of an outbreak due to Clostridium difficile PCR ribotype 027 and recognition of specific risk factors *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **15** 427-34

Della M, Palmbos P L, Tseng H M, Tonkin L M, Daley J M, Topper L M, Pitcher R S, Tomkinson A E, Wilson T E and Doherty A J 2004 Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine *Science* **306** 683-5

Delmee M, Avesani V, Delferriere N and Burtonboy G 1990 Characterization of flagella of Clostridium difficile and their role in serogrouping reactions *Journal of clinical microbiology* **28** 2210-4

Deneve C, Janoir C, Poilane I, Fantinato C and Collignon A 2009 New trends in Clostridium difficile virulence and pathogenesis *International journal of antimicrobial agents* **33 Suppl 1** S24-8

Dietrich M, Bartfeld S, Munke R, Lange C, Ogilvie L A, Friedrich A and Meyer T F 2011 Activation of NF-kappaB by Neisseria gonorrhoeae is associated with microcolony formation and type IV pilus retraction *Cellular microbiology* **13** 1168-82

Dillingham M S and Kowalczykowski S C 2008 RecBCD enzyme and the repair of double-stranded DNA breaks *Microbiology and molecular biology reviews : MMBR* **72** 642-71, Table of Contents

Dillingham M S, Spies M and Kowalczykowski S C 2003 RecBCD enzyme is a bipolar DNA helicase *Nature* **423** 893-7

Dineen S S, Villapakkam A C, Nordman J T and Sonenshein A L 2007 Repression of Clostridium difficile toxin gene expression by CodY *Molecular microbiology* **66** 206-19

Dingle T C, Mulvey G L and Armstrong G D 2011 Mutagenic analysis of the Clostridium difficile flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters *Infect Immun* **79** 4061-7

Dixon D A and Kowalczykowski S C 1993 The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the E. coli RecBCD enzyme *Cell* **73** 87-96

Drake S L and Koomey M 1995 The product of the pilQ gene is essential for the biogenesis of type IV pili in Neisseria gonorrhoeae *Molecular microbiology* **18** 975-86

Drake S L, Sandstedt S A and Koomey M 1997 PilP, a pilus biogenesis lipoprotein in Neisseria gonorrhoeae, affects expression of PilQ as a high-molecular-mass multimer *Molecular microbiology* **23** 657-68

Drudy D, Kyne L, O'Mahony R and Fanning S 2007 gyrA mutations in fluoroquinolone-resistant Clostridium difficile PCR-027 *Emerging infectious diseases* **13** 504-5

Drudy D, Quinn T, O'Mahony R, Kyne L, O'Gaora P and Fanning S 2006 High-level resistance to moxifloxacin and gatifloxacin associated with a novel mutation in gyrB in toxin-A-negative, toxin-B-positive Clostridium difficile *The Journal of antimicrobial chemotherapy* **58** 1264-7

Dupuy B and Sonenshein A L 1998 Regulated transcription of Clostridium difficile toxin genes *Molecular microbiology* **27** 107-20

Egerer M, Giesemann T, Jank T, Satchell K J and Aktories K 2007 Auto-catalytic cleavage of Clostridium difficile toxins A and B depends on cysteine protease activity *The Journal of biological chemistry* **282** 25314-21

el Karoui M, Ehrlich D and Gruss A 1998 Identification of the lactococcal exonuclease/recombinase and its modulation by the putative Chi sequence *Proceedings of the National Academy of Sciences of the United States of America* **95** 626-31

Emerson J E, Stabler R A, Wren B W and Fairweather N F 2008 Microarray analysis of the transcriptional responses of Clostridium difficile to environmental and antibiotic stress *Journal of medical microbiology* **57** 757-64

Eykelenboom J K, Blackwood J K, Okely E and Leach D R 2008 SbcCD causes a double-strand break at a DNA palindrome in the Escherichia coli chromosome *Molecular cell* **29** 644-51

Fawley W N and Wilcox M H 2001 Molecular epidemiology of endemic Clostridium difficile infection *Epidemiology and infection* **126** 343-50

Fernandez De Henestrosa A R, Ogi T, Aoyagi S, Chafin D, Hayes J J, Ohmori H and Woodgate R 2000 Identification of additional genes belonging to the LexA regulon in Escherichia coli *Molecular microbiology* **35** 1560-72

Filloux A 2010 A variety of bacterial pili involved in horizontal gene transfer *J Bacteriol* **192** 3243-5

Freeman J, Bauer M P, Baines S D, Corver J, Fawley W N, Goorhuis B, Kuijper E J and Wilcox M H 2010 The changing epidemiology of Clostridium difficile infections *Clinical microbiology reviews* **23** 529-49

Freiler J F, Durning S J and Ender P T 2001 Clostridium difficile small bowel enteritis occurring after total colectomy *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **33** 1429-31; discussion 32

Frisch C, Gerhard R, Aktories K, Hofmann F and Just I 2003 The complete receptor-binding domain of Clostridium difficile toxin A is required for endocytosis *Biochemical and biophysical research communications* **300** 706-11

Gebhard R L, Gerding D N, Olson M M, Peterson L R, McClain C J, Ansel H J, Shaw M J and Schwartz M L 1985 Clinical and endoscopic findings in patients early in the course of clostridium difficile-associated pseudomembranous colitis *The American journal of medicine* **78** 45-8

George R H, Symonds J M, Dimock F, Brown J D, Arabi Y, Shinagawa N, Keighley M R, Alexander-Williams J and Burdon D W 1978 Identification of Clostridium difficile as a cause of pseudomembranous colitis *British medical journal* **1** 695

Gerding D N, Muto C A and Owens R C, Jr. 2008 Measures to control and prevent Clostridium difficile infection *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **46 Suppl 1** S43-9

Giel J L, Sorg J A, Sonenshein A L and Zhu J 2010 Metabolism of bile salts in mice influences spore germination in Clostridium difficile *PloS one* **5** e8740

Giesemann T, Egerer M, Jank T and Aktories K 2008 Processing of Clostridium difficile toxins *Journal of medical microbiology* **57** 690-6

Giltner C L, Nguyen Y and Burrows L L 2012 Type IV pilin proteins: versatile molecular modules *Microbiology and molecular biology reviews : MMBR* **76** 740-72

Goorhuis A, Bakker D, Corver J, Debast S B, Harmanus C, Notermans D W, Bergwerff A A, Dekker F W and Kuijper E J 2008 Emergence of Clostridium difficile infection due to a new hypervirulent strain, polymerase chain reaction ribotype 078 *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **47** 1162-70

Goulding D, Thompson H, Emerson J, Fairweather N F, Dougan G and Douce G R 2009 Distinctive profiles of infection and pathology in hamsters infected with Clostridium difficile strains 630 and B1 *Infect Immun* **77** 5478-85

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

Greco A, Ho J G, Lin S J, Palcic M M, Rupnik M and Ng K K 2006 Carbohydrate recognition by Clostridium difficile toxin A *Nature structural & molecular biology* **13** 460-1

Gupta R, Barkan D, Redelman-Sidi G, Shuman S and Glickman M S 2011 Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways *Molecular microbiology* **79** 316-30

Guzzo C R, Salinas R K, Andrade M O and Farah C S 2009 PILZ protein structure and interactions with PILB and the FIMX EAL domain: implications for control of type IV pilus biogenesis *Journal of molecular biology* **393** 848-66

Haijema B J, Hamoen L W, Kooistra J, Venema G and van Sinderen D 1995 Expression of the ATP-dependent deoxyribonuclease of Bacillus subtilis is under competence-mediated control *Molecular microbiology* **15** 203-11

Halpern D, Chiapello H, Schbath S, Robin S, Hennequet-Antier C, Gruss A and El Karoui M 2007 Identification of DNA motifs implicated in maintenance of bacterial core genomes by predictive modeling *PLoS genetics* **3** 1614-21

Halpern D, Gruss A, Claverys J P and El-Karoui M 2004 rexAB mutants in Streptococcus pneumoniae *Microbiology* **150** 2409-14

Handa N, Ichige A, Kusano K and Kobayashi I 2000 Cellular responses to postsegregational killing by restriction-modification genes *J Bacteriol* **182** 2218-29

Haraldsen J D and Sonenshein A L 2003 Efficient sporulation in Clostridium difficile requires disruption of the sigmaK gene *Molecular microbiology* **48** 811-21

HCC 2006 Investigation into outbreaks of *Clostridium difficile* at Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust. Commission for Healthcare Audit and Inspection, London, United Kingdom. http://www.buckshealthcare.nhs.uk/Downloads/healthcarecommision/HCC-Investigation-into-the-Outbreak-of-Clostridium-Difficile.pdf.

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

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

Heap J T, Pennington O J, Cartman S T, Carter G P and Minton N P 2007 The ClosTron: a universal gene knock-out system for the genus Clostridium *Journal of microbiological methods* **70** 452-64

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

Helaine S, Carbonnelle E, Prouvensier L, Beretti J L, Nassif X and Pelicic V 2005 PilX, a pilus-associated protein essential for bacterial aggregation, is a key to pilus-facilitated attachment of Neisseria meningitidis to human cells *Molecular microbiology* **55** 65-77

Hengge R 2009 Principles of c-di-GMP signalling in bacteria *Nature reviews. Microbiology* **7** 263-73

Hennequin C, Porcheray F, Waligora-Dupriet A, Collignon A, Barc M, Bourlioux P and Karjalainen T 2001 GroEL (Hsp60) of Clostridium difficile is involved in cell adherence *Microbiology* **147** 87-96

Henrichsen J 1972 Bacterial surface translocation: a survey and a classification *Bacteriological reviews* **36** 478-503

Henrichsen J 1983 Twitching motility *Annual review of microbiology* **37** 81-93

Hensgens M P, Keessen E C, Squire M M, Riley T V, Koene M G, de Boer E, Lipman L J, Kuijper E J, European Society of Clinical M and Infectious Diseases Study Group for Clostridium d 2012 Clostridium difficile infection in the community: a zoonotic disease? *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **18** 635-45

Herdendorf T J, McCaslin D R and Forest K T 2002 Aquifex aeolicus PilT, homologue of a surface motility protein, is a thermostable oligomeric NTPase *J Bacteriol* **184** 6465-71

Hill T M, Sharma B, Valjavec-Gratian M and Smith J 1997 sfi-independent filamentation in Escherichia coli Is lexA dependent and requires DNA damage for induction *J Bacteriol* **179** 1931-9

Ho J G, Greco A, Rupnik M and Ng K K 2005 Crystal structure of receptor-binding C-terminal repeats from Clostridium difficile toxin A *Proceedings of the National Academy of Sciences of the United States of America* **102** 18373-8

Huang B, Whitchurch C B and Mattick J S 2003 FimX, a multidomain protein connecting environmental signals to twitching motility in Pseudomonas aeruginosa *J Bacteriol* **185** 7068-76

Huisman O, D'Ari R and Gottesman S 1984 Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation *Proceedings of the National Academy of Sciences of the United States of America* **81** 4490-4

Hundsberger T, Braun V, Weidmann M, Leukel P, Sauerborn M and von Eichel-Streiber C 1997 Transcription analysis of the genes tcdA-E of the pathogenicity locus of Clostridium difficile *European journal of biochemistry / FEBS* **244** 735-42

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

Ireton K and Grossman A D 1992 Coupling between gene expression and DNA synthesis early during development in Bacillus subtilis *Proceedings of the National Academy of Sciences of the United States of America* **89** 8808-12

Ireton K, Rudner D Z, Siranosian K J and Grossman A D 1993 Integration of multiple developmental signals in Bacillus subtilis through the Spo0A transcription factor *Genes & development* **7** 283-94

Jakovljevic V, Leonardy S, Hoppert M and Sogaard-Andersen L 2008 PilB and PilT are ATPases acting antagonistically in type IV pilus function in Myxococcus xanthus *J Bacteriol* **190** 2411-21

Janion C 2008 Inducible SOS response system of DNA repair and mutagenesis in Escherichia coli *International journal of biological sciences* **4** 338-44

Jank T and Aktories K 2008 Structure and mode of action of clostridial glucosylating toxins: the ABCD model *Trends Microbiol* **16** 222-9

Jank T, Giesemann T and Aktories K 2007 Rho-glucosylating Clostridium difficile toxins A and B: new insights into structure and function *Glycobiology* **17** 15R-22R

Janvilisri T, Scaria J, Thompson A D, Nicholson A, Limbago B M, Arroyo L G, Songer J G, Grohn Y T and Chang Y F 2009 Microarray identification of Clostridium difficile core components and divergent regions associated with host origin *J Bacteriol* **191** 3881-91

Jarrell K F 2009 *Pili and flagella : current research and future trends* (Wymondham: Caister Academic)

Jarrell K F and McBride M J 2008 The surprisingly diverse ways that prokaryotes move *Nature reviews. Microbiology* **6** 466-76

Jelsbak L and Kaiser D 2005 Regulating pilin expression reveals a threshold for S motility in Myxococcus xanthus *J Bacteriol* **187** 2105-12

Jones A M, Kuijper E J and Wilcox M H 2013 Clostridium difficile: a European perspective *The Journal of infection* **66** 115-28

Just I, Selzer J, Wilm M, von Eichel-Streiber C, Mann M and Aktories K 1995a Glucosylation of Rho proteins by Clostridium difficile toxin B *Nature* **375** 500-3

Just I, Wilm M, Selzer J, Rex G, von Eichel-Streiber C, Mann M and Aktories K 1995b The enterotoxin from Clostridium difficile (ToxA) monoglucosylates the Rho proteins *The Journal of biological chemistry* **270** 13932-6

Kachrimanidou M and Malisiovas N 2011 Clostridium difficile infection: a comprehensive review *Critical reviews in microbiology* **37** 178-87

Kallstrom H, Liszewski M K, Atkinson J P and Jonsson A B 1997 Membrane cofactor protein (MCP or CD46) is a cellular pilus receptor for pathogenic Neisseria *Molecular microbiology* **25** 639-47

Kamiya S, Ogura H, Meng X Q and Nakamura S 1992 Correlation between cytotoxin production and sporulation in Clostridium difficile *Journal of medical microbiology* **37** 206-10

Karlsson S, Burman L G and Akerlund T 2008 Induction of toxins in Clostridium difficile is associated with dramatic changes of its metabolism *Microbiology* **154** 3430-6

Kawamoto S, Horton K M and Fishman E K 1999 Pseudomembranous colitis: spectrum of imaging findings with clinical and pathologic correlation *Radiographics : a review publication of the Radiological Society of North America, Inc* **19** 887-97

Kazmierczak B I, Lebron M B and Murray T S 2006 Analysis of FimX, a phosphodiesterase that governs twitching motility in Pseudomonas aeruginosa *Molecular microbiology* **60** 1026-43

Kelly C P and Kyne L 2011 The host immune response to Clostridium difficile *Journal of medical microbiology* **60** 1070-9

Kidane D and Graumann P L 2005 Dynamic formation of RecA filaments at DNA double strand break repair centers in live cells *The Journal of cell biology* **170** 357-66

Kidane D, Sanchez H, Alonso J C and Graumann P L 2004 Visualization of DNA double-strand break repair in live bacteria reveals dynamic recruitment of Bacillus subtilis RecF, RecO and RecN proteins to distinct sites on the nucleoids *Molecular microbiology* **52** 1627-39

Kirn T J, Bose N and Taylor R K 2003 Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in Vibrio cholerae *Molecular microbiology* **49** 81-92

Kooistra J, Haijema B J, Hesseling-Meinders A and Venema G 1997 A conserved helicase motif of the AddA subunit of the Bacillus subtilis ATP-dependent nuclease (AddAB) is essential for DNA repair and recombination *Molecular microbiology* **23** 137-49

Kowalczykowski S C 2000 Initiation of genetic recombination and recombination-dependent replication *Trends in biochemical sciences* **25** 156-65

Krebs S J and Taylor R K 2011 Protection and attachment of Vibrio cholerae mediated by the toxin-coregulated pilus in the infant mouse model *J Bacteriol* **193** 5260-70

Kreuzer K N 2005 Interplay between DNA replication and recombination in prokaryotes *Annual review of microbiology* **59** 43-67

Kuehne S A, Cartman S T, Heap J T, Kelly M L, Cockayne A and Minton N P 2010 The role of toxin A and toxin B in Clostridium difficile infection *Nature* **467** 711-3

Kuehne S A and Minton N P 2012 ClosTron-mediated engineering of Clostridium *Bioengineered* **3** 247-54

Kuijper E J, Coignard B, Brazier J S, Suetens C, Drudy D, Wiuff C, Pituch H, Reichert P, Schneider F, Widmer A F, Olsen K E, Allerberger F, Notermans D W, Barbut F, Delmee M, Wilcox M, Pearson A, Patel B C, Brown D J, Frei R, Akerlund T, Poxton I R and Tull P 2007 Update of Clostridium difficile-associated disease due to PCR ribotype 027 in Europe *Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin* **12** E1-2

Kuzminov A 1999 Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda *Microbiology and molecular biology reviews : MMBR* **63** 751-813, table of contents

Kuzminov A 2001 Single-strand interruptions in replicating chromosomes cause double-strand breaks *Proceedings of the National Academy of Sciences of the United States of America* **98** 8241-6

Kyne L, Warny M, Qamar A and Kelly C P 2000 Asymptomatic carriage of Clostridium difficile and serum levels of IgG antibody against toxin A *The New England journal of medicine* **342** 390-7

Kyne L, Warny M, Qamar A and Kelly C P 2001 Association between antibody response to toxin A and protection against recurrent Clostridium difficile diarrhoea *Lancet* **357** 189-93

Labbe A C, Poirier L, Maccannell D, Louie T, Savoie M, Beliveau C, Laverdiere M and Pepin J 2008 Clostridium difficile infections in a Canadian tertiary care hospital before and during a regional epidemic associated with the BI/NAP1/027 strain *Antimicrobial agents and chemotherapy* **52** 3180-7

Larson H E, Price A B, Honour P and Borriello S P 1978 Clostridium difficile and the aetiology of pseudomembranous colitis *Lancet* **1** 1063-6

Lawley T D, Croucher N J, Yu L, Clare S, Sebaihia M, Goulding D, Pickard D J, Parkhill J, Choudhary J and Dougan G 2009 Proteomic and genomic characterization of highly infectious Clostridium difficile 630 spores *J Bacteriol* **191** 5377-86

Lee K K, Sheth H B, Wong W Y, Sherburne R, Paranchych W, Hodges R S, Lingwood C A, Krivan H and Irvin R T 1994 The binding of Pseudomonas aeruginosa pili to glycosphingolipids is a tip-associated event involving the C-terminal region of the structural pilin subunit *Molecular microbiology* **11** 705-13

Lemee L, Bourgeois I, Ruffin E, Collignon A, Lemeland J F and Pons J L 2005 Multilocus sequence analysis and comparative evolution of virulence-associated genes and housekeeping genes of Clostridium difficile *Microbiology* **151** 3171-80

Lemon K P, Kurtser I, Wu J and Grossman A D 2000 Control of initiation of sporulation by replication initiation genes in Bacillus subtilis *J Bacteriol* **182** 2989-91

Lenhart J S, Schroeder J W, Walsh B W and Simmons L A 2012 DNA repair and genome maintenance in Bacillus subtilis *Microbiology and molecular biology reviews : MMBR* **76** 530-64

Lessa F C, Gould C V and McDonald L C 2012 Current status of Clostridium difficile infection epidemiology *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **55 Suppl 2** S65-70

Lin Y P, Kuo C J, Koleci X, McDonough S P and Chang Y F 2011 Manganese binds to Clostridium difficile Fbp68 and is essential for fibronectin binding *The Journal of biological chemistry* **286** 3957-69

Liyanage H, Kashket S, Young M and Kashket E R 2001 Clostridium beijerinckii and Clostridium difficile detoxify methylglyoxal by a novel mechanism involving glycerol dehydrogenase *Applied and environmental microbiology* **67** 2004-10

Lofmark S, Edlund C and Nord C E 2010 Metronidazole is still the drug of choice for treatment of anaerobic infections *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **50 Suppl 1** S16-23

Lopez C M, Rholl D A, Trunck L A and Schweizer H P 2009 Versatile dual-technology system for markerless allele replacement in Burkholderia pseudomallei *Applied and environmental microbiology* **75** 6496-503

Lory S and Strom M S 1997 Structure-function relationship of type-IV prepilin peptidase of Pseudomonas aeruginosa--a review *Gene* **192** 117-21

Lyras D, O'Connor J R, Howarth P M, Sambol S P, Carter G P, Phumoonna T, Poon R, Adams V, Vedantam G, Johnson S, Gerding D N and Rood J I 2009 Toxin B is essential for virulence of Clostridium difficile *Nature* **458** 1176-9

Lyras D and Rood J I 1998 Conjugative transfer of RP4-oriT shuttle vectors from Escherichia coli to Clostridium perfringens *Plasmid* **39** 160-4

MacCannell D R, Louie T J, Gregson D B, Laverdiere M, Labbe A C, Laing F and Henwick S 2006 Molecular analysis of Clostridium difficile PCR ribotype 027 isolates from Eastern and Western Canada *Journal of clinical microbiology* **44** 2147-52

Maier B, Potter L, So M, Long C D, Seifert H S and Sheetz M P 2002 Single pilus motor forces exceed 100 pN *Proceedings of the National Academy of Sciences of the United States of America* **99** 16012-7

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

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

Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood J I, Sonenshein A L and Dupuy B 2002 Environmental response and autoregulation of Clostridium difficile TxeR, a sigma factor for toxin gene expression *J Bacteriol* **184** 5971-8

Matamouros S, England P and Dupuy B 2007 Clostridium difficile toxin expression is inhibited by the novel regulator TcdC *Molecular microbiology* **64** 1274-88

Mattick J S 2002 Type IV pili and twitching motility *Annual review of microbiology* **56** 289-314

McDonald L C, Killgore G E, Thompson A, Owens R C, Jr., Kazakova S V, Sambol S P, Johnson S and Gerding D N 2005 An epidemic, toxin gene-variant strain of Clostridium difficile *The New England journal of medicine* **353** 2433-41

McFarland L V, Bauwens J E, Melcher S A, Surawicz C M, Greenberg R N and Elmer G W 1995 Ciprofloxacin-associated Clostridium difficile disease *Lancet* **346** 977-8

McFarland L V, Surawicz C M and Stamm W E 1990 Risk factors for Clostridium difficile carriage and C. difficile-associated diarrhea in a cohort of hospitalized patients *The Journal of infectious diseases* **162** 678-84

McGlone S M, Bailey R R, Zimmer S M, Popovich M J, Tian Y, Ufberg P, Muder R R and Lee B Y 2012 The economic burden of Clostridium difficile *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **18** 282-9

McGlynn P and Lloyd R G 2002 Recombinational repair and restart of damaged replication forks *Nature reviews. Molecular cell biology* **3** 859-70

Mendez M, Huang I H, Ohtani K, Grau R, Shimizu T and Sarker M R 2008 Carbon catabolite repression of type IV pilus-dependent gliding motility in the anaerobic pathogen Clostridium perfringens *J Bacteriol* **190** 48-60

Mertens K, Lantsheer L, Ennis D G and Samuel J E 2008 Constitutive SOS expression and damage-inducible AddAB-mediated recombinational repair systems for Coxiella burnetii as potential adaptations for survival within macrophages *Molecular microbiology* **69** 1411-26

Merz A J, So M and Sheetz M P 2000 Pilus retraction powers bacterial twitching motility *Nature* **407** 98-102

Michael W M 2001 Cell cycle: connecting DNA replication to sporulation in Bacillus *Current biology : CB* **11** R443-5

Michel B 2005 After 30 years of study, the bacterial SOS response still surprises us *PLoS Biol* **3** e255

Michel B, Boubakri H, Baharoglu Z, LeMasson M and Lestini R 2007 Recombination proteins and rescue of arrested replication forks *DNA repair* **6** 967-80

Michel B, Ehrlich S D and Uzest M 1997 DNA double-strand breaks caused by replication arrest *The EMBO journal* **16** 430-8

Mignot T and Kirby J R 2008 Genetic circuitry controlling motility behaviors of Myxococcus xanthus *BioEssays : news and reviews in molecular, cellular and developmental biology* **30** 733-43

Mo A H and Burkholder W F 2010 YneA, an SOS-induced inhibitor of cell division in Bacillus subtilis, is regulated posttranslationally and requires the transmembrane region for activity *J Bacteriol* **192** 3159-73

Moeller R, Stackebrandt E, Reitz G, Berger T, Rettberg P, Doherty A J, Horneck G and Nicholson W L 2007 Role of DNA repair by nonhomologous-end joining in Bacillus subtilis spore resistance to extreme dryness, mono- and polychromatic UV, and ionizing radiation *J Bacteriol* **189** 3306-11

Monteilhet C, Perrin A, Thierry A, Colleaux L and Dujon B 1990 Purification and characterization of the in vitro activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron *Nucleic acids research* **18** 1407-13

Mullany P, Wilks M, Lamb I, Clayton C, Wren B and Tabaqchali S 1990 Genetic analysis of a tetracycline resistance element from Clostridium difficile and its conjugal transfer to and from Bacillus subtilis *Journal of general microbiology* **136** 1343-9

Mullany P, Wilks M, Puckey L and Tabaqchali S 1994 Gene cloning in Clostridium difficile using Tn916 as a shuttle conjugative transposon *Plasmid* **31** 320-3

Mullany P, Wilks M and Tabaqchali S 1991 Transfer of Tn916 and Tn916 delta E into Clostridium difficile: demonstration of a hot-spot for these elements in the C. difficile genome *FEMS Microbiol Lett* **63** 191-4

Ng Y K, Ehsaan M, Philip S, Collery M M, Janoir C, Collignon A, Cartman S T and 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** e56051

Nudleman E and Kaiser D 2004 Pulling together with type IV pili *Journal of molecular microbiology and biotechnology* **7** 52-62

O'Connor J R, Lyras D, Farrow K A, Adams V, Powell D R, Hinds J, Cheung J K and Rood J I 2006 Construction and analysis of chromosomal Clostridium difficile mutants *Molecular microbiology* **61** 1335-51

O'Reilly E K and Kreuzer K N 2004 Isolation of SOS constitutive mutants of Escherichia coli *J Bacteriol* **186** 7149-60

O'Toole G A and Kolter R 1998 Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development *Molecular microbiology* **30** 295-304

Oezguen N, Power T D, Urvil P, Feng H, Pothoulakis C, Stamler J S, Braun W and Savidge T C 2012 Clostridial toxins: sensing a target in a hostile gut environment *Gut microbes* **3** 35-41

Ossanna N and Mount D W 1989 Mutations in uvrD induce the SOS response in Escherichia coli *J Bacteriol* **171** 303-7

Oughton M T, Loo V G, Dendukuri N, Fenn S and Libman M D 2009 Hand hygiene with soap and water is superior to alcohol rub and antiseptic wipes for removal of Clostridium difficile *Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America* **30** 939-44

Owens R C, Jr., Donskey C J, Gaynes R P, Loo V G and Muto C A 2008 Antimicrobial-associated risk factors for Clostridium difficile infection *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **46 Suppl 1** S19-31

Pamp S J and Tolker-Nielsen T 2007 Multiple roles of biosurfactants in structural biofilm development by Pseudomonas aeruginosa *J Bacteriol* **189** 2531-9

Parry F L 2010 Identification of pre-synaptic processing proteins from *Bacteroides fragilis*. The University of Edinburgh)

Patrick S, Houston S, Thacker Z and Blakely G W 2009 Mutational analysis of genes implicated in LPS and capsular polysaccharide biosynthesis in the opportunistic pathogen Bacteroides fragilis *Microbiology* **155** 1039-49

Pavankumar T L, Centre for Cellular and Molecular Biology C o S a I R, Hyderabad, India, Sinha A K, Centre for Cellular and Molecular Biology C o S a I R, Hyderabad, India, Ray M K and Centre for Cellular and Molecular Biology C o S a I R, Hyderabad, India 2010 All Three Subunits of RecBCD Enzyme Are Essential for DNA Repair and Low-Temperature Growth in the Antarctic Pseudomonas syringae Lz4W *PloS one* **5**

Pelicic V 2008 Type IV pili: e pluribus unum? *Molecular microbiology* **68** 827-37

Pepin J, Valiquette L, Alary M E, Villemure P, Pelletier A, Forget K, Pepin K and Chouinard D 2004 Clostridium difficile-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* **171** 466-72

Perelle S, Gibert M, Bourlioux P, Corthier G and Popoff M R 1997 Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by Clostridium difficile CD196 *Infect Immun* **65** 1402-7

Perrin A, Buckle M and Dujon B 1993 Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions *The EMBO journal* **12** 2939-47

Pfeifer G, Schirmer J, Leemhuis J, Busch C, Meyer D K, Aktories K and Barth H 2003 Cellular uptake of Clostridium difficile toxin B. Translocation of the N-terminal catalytic domain into the cytosol of eukaryotic cells *The Journal of biological chemistry* **278** 44535-41

Popoff M R, Rubin E J, Gill D M and Boquet P 1988 Actin-specific ADP-ribosyltransferase produced by a Clostridium difficile strain *Infect Immun* **56** 2299-306

Posfai G, Kolisnychenko V, Bereczki Z and Blattner F R 1999 Markerless gene replacement in Escherichia coli stimulated by a double-strand break in the chromosome *Nucleic acids research* **27** 4409-15

Poxton I R 2010 Fidaxomicin: a new macrocyclic, RNA polymerase-inhibiting antibiotic for the treatment of Clostridium difficile infections *Future microbiology* **5** 539-48

Poxton I R, McCoubrey J and Blair G 2001 The pathogenicity of Clostridium difficile *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **7** 421-7

Prieto A I, Ramos-Morales F and Casadesus J 2006 Repair of DNA damage induced by bile salts in Salmonella enterica *Genetics* **174** 575-84

Proft T and Baker E N 2009 Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease *Cellular and molecular life sciences : CMLS* **66** 613-35

Pruitt R N, Chagot B, Cover M, Chazin W J, Spiller B and Lacy D B 2009 Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in Clostridium difficile toxin A *The Journal of biological chemistry* **284** 21934-40

Pujol C, Eugene E, Marceau M and Nassif X 1999 The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion *Proceedings of the National Academy of Sciences of the United States of America* **96** 4017-22

Purcell E B, McKee R W, McBride S M, Waters C M and Tamayo R 2012 Cyclic diguanylate inversely regulates motility and aggregation in Clostridium difficile *J Bacteriol* **194** 3307-16

Purdy D, O'Keeffe T A, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A and 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** 439-52

Qa'Dan M, Spyres L M and Ballard J D 2000 pH-induced conformational changes in Clostridium difficile toxin B *Infect Immun* **68** 2470-4

Rakotoarivonina H, Jubelin G, Hebraud M, Gaillard-Martinie B, Forano E and Mosoni P 2002 Adhesion to cellulose of the Gram-positive bacterium Ruminococcus albus involves type IV pili *Microbiology* **148** 1871-80

Reineke J, Tenzer S, Rupnik M, Koschinski A, Hasselmayer O, Schrattenholz A, Schild H and von Eichel-Streiber C 2007 Autocatalytic cleavage of Clostridium difficile toxin B *Nature* **446** 415-9

Reuter M, Parry F, Dryden D T and Blakely G W 2010 Single-molecule imaging of Bacteroides fragilis AddAB reveals the highly processive translocation of a single motor helicase *Nucleic acids research* **38** 3721-31

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

Riggs M M, Sethi A K, Zabarsky T F, Eckstein E C, Jump R L and Donskey C J 2007 Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic Clostridium difficile strains among long-term care facility residents *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **45** 992-8

Rocha E P, Cornet E and Michel B 2005 Comparative and evolutionary analysis of the bacterial homologous recombination systems *PLoS genetics* **1** e15

Rodgers K, Arvidson C G and Melville S 2011 Expression of a Clostridium perfringens type IV pilin by Neisseria gonorrhoeae mediates adherence to muscle cells *Infect Immun* **79** 3096-105

Rothstein R, Michel B and Gangloff S 2000 Replication fork pausing and recombination or "gimme a break" *Genes & development* **14** 1-10

Rubin M S, Bodenstein L E and Kent K C 1995 Severe Clostridium difficile colitis *Diseases of the colon and rectum* **38** 350-4

Rupnik M 2008 Heterogeneity of large clostridial toxins: importance of Clostridium difficile toxinotypes *FEMS microbiology reviews* **32** 541-55

Rupnik M, Pabst S, Rupnik M, von Eichel-Streiber C, Urlaub H and Soling H D 2005 Characterization of the cleavage site and function of resulting cleavage fragments after limited proteolysis of Clostridium difficile toxin B (TcdB) by host cells *Microbiology* **151** 199-208

Rupnik M, Wilcox M H and Gerding D N 2009 Clostridium difficile infection: new developments in epidemiology and pathogenesis *Nature reviews. Microbiology* **7** 526-36

Russell M A and Darzins A 1994 The pilE gene product of Pseudomonas aeruginosa, required for pilus biogenesis, shares amino acid sequence identity with the N-termini of type 4 prepilin proteins *Molecular microbiology* **13** 973-85

Saikrishnan K, Yeeles J T, Gilhooly N S, Krajewski W W, Dillingham M S and Wigley D B 2012 Insights into Chi recognition from the structure of an AddAB-type helicase-nuclease complex *The EMBO journal* **31** 1568-78

Sanchez H, Kidane D, Castillo Cozar M, Graumann P L and Alonso J C 2006 Recruitment of Bacillus subtilis RecN to DNA double-strand breaks in the absence of DNA end processing *J Bacteriol* **188** 353-60

Schwan C, Stecher B, Tzivelekidis T, van Ham M, Rohde M, Hardt W D, Wehland J and Aktories K 2009 Clostridium difficile toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria *PLoS pathogens* **5** e1000626

Sciochetti S A, Blakely G W and Piggot P J 2001 Growth phase variation in cell and nucleoid morphology in a Bacillus subtilis recA mutant *J Bacteriol* **183** 2963-8

Sebaihia M, Wren B W, Mullany P, Fairweather N F, Minton N, Stabler R, Thomson N R, Roberts A P, Cerdeño-Tárraga A M, Wang H, Holden M T, Wright A, Churcher C, Quail M A, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B and Parkhill J 2006 The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome *Nat Genet* **38** 779-86

Selby C P and Sancar A 1994 Mechanisms of transcription-repair coupling and mutation frequency decline *Microbiological reviews* **58** 317-29

Semmler A B, Whitchurch C B and Mattick J S 1999 A re-examination of twitching motility in Pseudomonas aeruginosa *Microbiology* **145 (Pt 10)** 2863-73

Setlow P 2007 I will survive: DNA protection in bacterial spores *Trends Microbiol* **15** 172-80

Shi W and Sun H 2002 Type IV pilus-dependent motility and its possible role in bacterial pathogenesis *Infect Immun* **70** 1-4

Shim J K, Johnson S, Samore M H, Bliss D Z and Gerding D N 1998 Primary symptomless colonisation by Clostridium difficile and decreased risk of subsequent diarrhoea *Lancet* **351** 633-6

Shuman S and Glickman M S 2007 Bacterial DNA repair by non-homologous end joining *Nature reviews. Microbiology* **5** 852-61

Simmon V F and Lederberg S 1972 Degradation of bacteriophage lambda deoxyribonucleic acid after restriction by Escherichia coli K-12 *J Bacteriol* **112** 161-9

Simmons L A, Goranov A I, Kobayashi H, Davies B W, Yuan D S, Grossman A D and Walker G C 2009 Comparison of responses to double-strand breaks between Escherichia coli and Bacillus subtilis reveals different requirements for SOS induction *J Bacteriol* **191** 1152-61

Simon R. P U, Puhler A. 1983 A Broad Host Range Mobilization System for *In Vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria *Nature Biotechnology* **1** 784 - 91

Singleton M R, Dillingham M S, Gaudier M, Kowalczykowski S C and Wigley D B 2004 Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks *Nature* **432** 187-93

Singleton M R, Dillingham M S and Wigley D B 2007 Structure and mechanism of helicases and nucleic acid translocases *Annual review of biochemistry* **76** 23-50

Sinha K M, Unciuleac M C, Glickman M S and Shuman S 2009 AdnAB: a new DSB-resecting motor-nuclease from mycobacteria *Genes & development* **23** 1423-37

Sisson G, Jeong J Y, Goodwin A, Bryden L, Rossler N, Lim-Morrison S, Raudonikiene A, Berg D E and Hoffman P S 2000 Metronidazole activation is mutagenic and causes DNA fragmentation in

Helicobacter pylori and in Escherichia coli containing a cloned H. pylori RdxA(+) (Nitroreductase) gene *J Bacteriol* **182** 5091-6

Skerker J M and Berg H C 2001 Direct observation of extension and retraction of type IV pili *Proceedings of the National Academy of Sciences of the United States of America* **98** 6901-4

Smith A 2005 Outbreak of Clostridium difficile infection in an English hospital linked to hypertoxin-producing strains in Canada and the US *Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin* **10** E050630 2

Smith G R 1991 Conjugational recombination in E. coli: myths and mechanisms *Cell* **64** 19-27

Smith G R 2012 How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view *Microbiology and molecular biology reviews : MMBR* **76** 217-28

Soderholm N, Vielfort K, Hultenby K and Aro H 2011 Pathogenic Neisseria hitchhike on the uropod of human neutrophils *PloS one* **6** e24353

Sorg J A and Sonenshein A L 2008 Bile salts and glycine as cogerminants for Clostridium difficile spores *J Bacteriol* **190** 2505-12

Spek E J, Wright T L, Stitt M S, Taghizadeh N R, Tannenbaum S R, Marinus M G and Engelward B P 2001 Recombinational repair is critical for survival of Escherichia coli exposed to nitric oxide *J Bacteriol* **183** 131-8

Spies M, Amitani I, Baskin R J and Kowalczykowski S C 2007 RecBCD enzyme switches lead motor subunits in response to chi recognition *Cell* **131** 694-705

Spies M, Bianco P R, Dillingham M S, Handa N, Baskin R J and Kowalczykowski S C 2003 A molecular throttle: the recombination hotspot chi controls DNA translocation by the RecBCD helicase *Cell* **114** 647-54

Spigaglia P, Barbanti F, Mastrantonio P, Brazier J S, Barbut F, Delmee M, Kuijper E, Poxton I R and European Study Group on Clostridium d 2008 Fluoroquinolone resistance in Clostridium difficile isolates from a prospective study of C. difficile infections in Europe *Journal of medical microbiology* **57** 784-9

Stabler R A, Gerding D N, Songer J G, Drudy D, Brazier J S, Trinh H T, Witney A A, Hinds J and Wren B W 2006 Comparative phylogenomics of Clostridium difficile reveals clade specificity and microevolution of hypervirulent strains *J Bacteriol* **188** 7297-305

Stabler R A, He M, Dawson L, Martin M, Valiente E, Corton C, Lawley T D, Sebaihia M, Quail M A, Rose G, Gerding D N, Gibert M, Popoff M R, Parkhill J, Dougan G and Wren B W 2009 Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium *Genome biology* **10** R102

Steffens L S, Nicholson S, Paul L V, Nord C E, Patrick S and Abratt V R 2010 Bacteroides fragilis RecA protein overexpression causes resistance to metronidazole *Res Microbiol* **161** 346-54

Stohl E A and Seifert H S 2006 Neisseria gonorrhoeae DNA recombination and repair enzymes protect against oxidative damage caused by hydrogen peroxide *J Bacteriol* **188** 7645-51

Strom M S and Lory S 1993 Structure-function and biogenesis of the type IV pili *Annual review of microbiology* **47** 565-96

Sudarsan N, Lee E R, Weinberg Z, Moy R H, Kim J N, Link K H and Breaker R R 2008 Riboswitches in eubacteria sense the second messenger cyclic di-GMP *Science* **321** 411-3

Sundriyal A, Roberts A K, Ling R, McGlashan J, Shone C C and Acharya K R 2010 Expression, purification and cell cytotoxicity of actin-modifying binary toxin from Clostridium difficile *Protein expression and purification* **74** 42-8

Suzuki N, Nonaka H, Tsuge Y, Okayama S, Inui M and Yukawa H 2005 Multiple large segment deletion method for Corynebacterium glutamicum *Applied microbiology and biotechnology* **69** 151-61

Swanson J 1973 Studies on gonococcus infection. IV. Pili: their role in attachment of gonococci to tissue culture cells *The Journal of experimental medicine* **137** 571-89

Tan K S, Wee B Y and Song K P 2001 Evidence for holin function of tcdE gene in the pathogenicity of Clostridium difficile *Journal of medical microbiology* **50** 613-9

Tasteyre A, Barc M C, Karjalainen T, Dodson P, Hyde S, Bourlioux P and Borriello P 2000 A Clostridium difficile gene encoding flagellin *Microbiology* **146 (Pt 4)** 957-66

Taylor A F and Smith G R 2003 RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity *Nature* **423** 889-93

Terashima H, Kojima S and Homma M 2008 Flagellar motility in bacteria structure and function of flagellar motor *International review of cell and molecular biology* **270** 39-85

Triadafilopoulos G and Hallstone A E 1991 Acute abdomen as the first presentation of pseudomembranous colitis *Gastroenterology* **101** 685-91

Twine S M, Reid C W, Aubry A, McMullin D R, Fulton K M, Austin J and Logan S M 2009 Motility and flagellar glycosylation in Clostridium difficile *J Bacteriol* **191** 7050-62

Underwood S, Guan S, Vijayasubhash V, Baines S D, Graham L, Lewis R J, Wilcox M H and Stephenson K 2009 Characterization of the sporulation initiation pathway of Clostridium difficile and its role in toxin production *J Bacteriol* **191** 7296-305

Varga J J, Nguyen V, O'Brien D K, Rodgers K, Walker R A and Melville S B 2006 Type IV pili-dependent gliding motility in the Gram-positive pathogen Clostridium perfringens and other Clostridia *Molecular microbiology* **62** 680-94

Varga J J, Therit B and Melville S B 2008 Type IV pili and the CcpA protein are needed for maximal biofilm formation by the gram-positive anaerobic pathogen Clostridium perfringens *Infect Immun* **76** 4944-51

Vedantam G, Clark A, Chu M, McQuade R, Mallozzi M and Viswanathan V K 2012 Clostridium difficile infection: toxins and non-toxin virulence factors, and their contributions to disease establishment and host response *Gut microbes* **3** 121-34

Vohra P and Poxton I R 2011 Comparison of toxin and spore production in clinically relevant strains of Clostridium difficile *Microbiology* **157** 1343-53

von Eichel-Streiber C, Boquet P, Sauerborn M and Thelestam M 1996 Large clostridial cytotoxins--a family of glycosyltransferases modifying small GTP-binding proteins *Trends Microbiol* **4** 375-82

Voth D E and Ballard J D 2005 Clostridium difficile toxins: mechanism of action and role in disease *Clinical microbiology reviews* **18** 247-63

Waligora A J, Hennequin C, Mullany P, Bourlioux P, Collignon A and Karjalainen T 2001 Characterization of a cell surface protein of Clostridium difficile with adhesive properties *Infect Immun* **69** 2144-53

Wall D, Kolenbrander P E and Kaiser D 1999 The Myxococcus xanthus pilQ (sglA) gene encodes a secretin homolog required for type IV pilus biogenesis, social motility, and development *J Bacteriol* **181** 24-33

Wang H, Roberts A P and Mullany P 2000a DNA sequence of the insertional hot spot of Tn916 in the Clostridium difficile genome and discovery of a Tn916-like element in an environmental isolate integrated in the same hot spot *FEMS Microbiol Lett* **192** 15-20

Wang J, Chen R and Julin D A 2000b A single nuclease active site of the Escherichia coli RecBCD enzyme catalyzes single-stranded DNA degradation in both directions *The Journal of biological chemistry* **275** 507-13

Wang T C and Smith K C 1983 Mechanisms for recF-dependent and recB-dependent pathways of postreplication repair in UV-irradiated Escherichia coli uvrB *J Bacteriol* **156** 1093-8

Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, Frost E and McDonald L C 2005 Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe *Lancet* **366** 1079-84

Wigley D B 2013 Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB *Nature reviews. Microbiology* **11** 9-13

Williams A B, Hetrick K M and Foster P L 2011 Double-Strand Break Repair and Holliday Junction Processing Are Required for Chromosome Processing in Stationary-Phase Escherichia coli Cells *G3* **1** 417-26

Wilson K H, Kennedy M J and Fekety F R 1982 Use of sodium taurocholate to enhance spore recovery on a medium selective for Clostridium difficile *Journal of clinical microbiology* **15** 443-6

Winther-Larsen H C, Wolfgang M, Dunham S, van Putten J P, Dorward D, Lovold C, Aas F E and Koomey M 2005 A conserved set of pilin-like molecules controls type IV pilus dynamics and organelle-associated functions in Neisseria gonorrhoeae *Molecular microbiology* **56** 903-17

Wolfgang M, Park H S, Hayes S F, van Putten J P and Koomey M 1998 Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function mutations in pilT, a twitching motility gene in Neisseria gonorrhoeae *Proceedings of the National Academy of Sciences of the United States of America* **95** 14973-8

Yeeles J T, Cammack R and Dillingham M S 2009 An iron-sulfur cluster is essential for the binding of broken DNA by AddAB-type helicase-nucleases *The Journal of biological chemistry* **284** 7746-55

Yeeles J T and Dillingham M S 2007 A dual-nuclease mechanism for DNA break processing by AddAB-type helicase-nucleases *Journal of molecular biology* **371** 66-78

Yeeles J T and Dillingham M S 2010 The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes *DNA repair* **9** 276-85

Yeeles J T, van Aelst K, Dillingham M S and Moreno-Herrero F 2011 Recombination hotspots and single-stranded DNA binding proteins couple DNA translocation to DNA unwinding by the AddAB helicase-nuclease *Molecular cell* **42** 806-16

Zahavi E E, Lieberman J A, Donnenberg M S, Nitzan M, Baruch K, Rosenshine I, Turner J R, Melamed-Book N, Feinstein N, Zlotkin-Rivkin E and Aroeti B 2011 Bundle-forming pilus retraction enhances enteropathogenic Escherichia coli infectivity *Molecular biology of the cell* **22** 2436-47

Zhang P, Kong L, Setlow P and Li Y Q 2010 Characterization of wet-heat inactivation of single spores of bacillus species by dual-trap Raman spectroscopy and elastic light scattering *Applied and environmental microbiology* **76** 1796-805

Zolfaghar I, Evans D J and Fleiszig S M 2003 Twitching motility contributes to the role of pili in corneal infection caused by Pseudomonas aeruginosa *Infect Immun* **71** 5389-93

Appendix I. Promoters predicted in the intergenic region located between the CD3513 locus and the *prs* gene using the BPROM (Softberry) program

```
Length of sequence-      1087
 Threshold for promoters -  0.20
 Number of predicted promoters -      3
 Promoter Pos:    968 LDF- 10.46
 -10 box at pos.    951 ttttaaact Score    69
 -35 box at pos.    932 tttact    Score    42
 Promoter Pos:    309 LDF-  9.02
 -10 box at pos.    294 ttatataat Score    75
 -35 box at pos.    271 atgtat    Score    12
 Promoter Pos:    646 LDF-  4.26
 -10 box at pos.    631 gtgcaaaat Score    51
 -35 box at pos.    613 tttaca    Score    47

 Oligonucleotides from known TF binding sites:

 For promoter at    968:
       phoB:  TAATATAT at position     908 Score -  15
      argR2:  TTTATTTT at position     927 Score -   7
     rpoD17:  TTTTACTT at position     931 Score -   8
     rpoD19:  TACTTAAA at position     934 Score -   8
        crp:  TATACATA at position     942 Score -  13
       tyrR:  TTTACAAT at position     970 Score -   8
       phoB:  TAATATAT at position     977 Score -  15
       tyrR:  AATATATA at position     978 Score -  15
       lexA:  ATATATAC at position     979 Score -  12
 For promoter at    309:
       glpR:  TTCAAAAT at position     250 Score -   6
       cysB:  TGTATATA at position     272 Score -  12
       lexA:  TATATACA at position     274 Score -  13
        crp:  TATACATA at position     276 Score -  13
     rpoS17:  TTATATTA at position     289 Score -  14
      argR2:  CATATTTT at position     309 Score -   8
       nagC:  ATATTTTA at position     310 Score -   7
     rpoD17:  TAGCAACA at position     316 Score -   7
 For promoter at    646:
       lexA:  TATATAAA at position     593 Score -  14
       trpR:  TACTCTTT at position     640 Score -   7
```

*Please note that this software is designed to identify promoters based on the *E.coli* sigma70 promoter consensus. Therefore may not be accurate for use in *C. difficile*.