Development of transposon based tools for the investigation of virulence factors in *Clostridium difficile*

Katherine Lucie Anne McCurrie

University College London

Thesis submitted for the degree of PhD

June 2011

Declaration

I, Katherine Lucie Anne McCurrie confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Foreword

I would like to begin by thanking Professor Mullany for providing me with the opportunity to conduct a PhD at UCL and for all his support over the last 4 years. I would also like to that Dr. Elaine Allan for her input and helpful suggestions throughout the project.

I thank Dr. Haitham Hussain who has been absolutely fantastic throughout my studies, and even through the endless question he maintained a friendly and helpful attitude. I am so grateful for your help support. I thank Dr. Richard Stabler for all his help with the transcriptomics work and for reviewing chapter 3; Dr. Lena Ciric for her help with PT-PCR and Dr. Florent Chan-Pi-Hin for allowing me to steal restriction enzymes and for answering my various PCR related questions. I also thank Dr. Tom Morgan for helping me to remain positive during the long days and evenings (aided by Gin O'clock) and for the various bizarre 'lab creations' left on my bench. Finally, I thank Linda, Sarah, Mike, and Zoie for making my time at the Eastman so much fun (especially the Christmas parties!) without whom I am not sure I would have made it this far.

Although, this has been one of the most overwhelming and stressful experiences of my life so far, I have really enjoyed working in the laboratory with some truly great colleagues and friends. I am going to miss working in the laboratory especially working with *C. difficile*, although I wont miss the smell! I hope that I will always be able to work in Science.

Table of Contents

List of Figures

List of Tables

Abbreviations

^{o}C	Degrees centigrade
1^{o}	Primary
2^{o}	Secondary
AAD	Antibiotic associated diarrhoea
ANOVA	Analysis of variance
Ap	Ampicillin
B & O bile	Bile of Bovine and Ovine origin
BHI	Brain Heart Infusion
BHIB	Brain Heart Infusion with 5% horse blood
BHIBS	Brain Heart Infusion with 5% horse blood and selective supplement CDSS
BHIS	Brain Heart Infusion with selective supplement CDSS
bp	Base pair
BSA	Bovine serum albumin
BSH	Bile salt hydrolase
CA	Cholic acid
CCFA	Cycloserine cefoxitin fructose agar
CDA	Chenodeoxycholic acid
CDI	*Clostridium difficile* infection
cDNA	Deoxyribonucleic acid complementary to ribonucleic acid
CDS	Coding sequence
CDSS	*Clostridium difficile* selective supplement
Cdt	*Clostridium difficile* binary toxin
CdtLoc	*Clostridium difficile* binary toxin locus
Cfu	Colony forming units
CIP	Calf intestinal phosphatase
Cli	Clindamycin
Cm	Chloramphenicol
CMC	Critical micellar concentration
CPE	Cytopathic effect
CTn(s)	Conjugative transposon(s)
dATP	Deoxyadenosine triphosphate
DCA	Deoxycholic acid
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide-triphosphates
dTTP	Deoxy-thymidine tri-phosphate
EDTA	Ethylenediaminetetraacetic acid
EIA	Immunoabsorbant assay
Em	Erythromycin
FCS	Bovine/foetal calf serum
g	Grams
GCA	Sodium Glycocholate
GCDA	Sodium Glycochenodeoxycholate
GI	Gastrointestinal
GIT	Gastrointestinal tract
GRE	Glycopeptide resistant Enterococci
H_2O	Distilled water
HCAI	Healthcare associated infection(s)
H-NS	Histone-like nuecleoid structuring protein
HU	Histone-like DNA binging protein
ICE	Integrative conjugative element(s)
IE	Inside end
IEP	Intron-encoded protein
IHF	Integration host factor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IS	Insertion sequence

Kan	Kanamycin
Kb	Kilobase(s)
KDa	Kilodaltons
kV	killovolts
l	Litre
LB	Luria-Bertani
LPS	Lipopolysaccharide
Lm	Lincomycin
M	Mole(s) per litre
Mb	Megabase
MCS	Multiple cloning site
mg	Milligrams
MIC	Minimum inhibitory concentration
min	Minute
ml	Millilitre(s)
MLS_B	Macrolide-lincosamide-streptogramin B antibiotics
mM	Millimole
molH_2O	Molecular grade filtered DNase/RNase free water
MRSA	Meticillin-resistant *Staphylococcus aureus*
MW	Molecular weight
ng	Nanograms
nm	Nanometers
nM	nanomole
OD	Optical density
OE	Outside end
ORF	Open reading frame
PaLoc	Pathogenicity locus
PB	Porcine bile
Pb-	Probe
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
Pen/Strep	Penicillin G (5000 U/ml) and Streptomycin sulphate (5000 mcg/ml)
pg	Picogram
PMC	Pseudomembranous colitis
PPI	Proton pump inhibitor
RAM	Retrotranspositional-activated marker
RBS	Ribosome binding site
RCM	Reinforced clostridial medium
Rif	Rifampicin
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulphate
SLP	Surface layer protein
SOC	Super Optimal Broth with glucose
SSC	Saline-sodium citrate buffer
TAE	Tris-Acetic acid-EDTA
Taq	*Thermus aquaticus* polymerase
Tc	Tetracycline
TCA	Sodium taurocholate
Tcd A	*Clostridium difficile* cytotoxin A
Tcd B	*Clostridium difficile* cytotoxin B
TCDA	Sodium taurochenodeoxycholate
Tm	Thiamphenicol
Tn(s)	Transposon(s)
U	Units
µF	microfarads
µg	Microgram
µl	Microliters
UV	Ultraviolet
V	Volts

vol/vol	Volume per volume
wt/vol	Weight per volume
x *g*	Times gravity
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Abstract

C. difficile is the leading cause of antibiotic associated diarrhoea worldwide. A lack of available tools has hindered the investigation of the molecular pathogenicity of this important nosocomial organism. A library of Tn*916* transconjugants has been generated and screened for phenotypes associated with virulence and colonisation. Investigation of the insertion sites of Tn*916* in the genomes of *C. difficile* strains R20291 and 630Δerm revealed an insertion bias to intergenic regions. A 15 bp consensus sequence has been identified as the preferential site of insertion in strain 630Δerm. Analysis of the genome of strain 630 revealed that this motif is found predominantly in intergenic regions. These results suggest that Tn*916* is not suitable as a random mutagen in these strains, however it may be used as a saturating mutagen to facilitate investigation of these poorly described genomic regions.

The tolerance of *C. difficile* strains to bile salts was investigated revealing differences in the tolerance profile of the hypervirulent strain R20291 comparred with strain 630Δerm. R20291 appears to be inherently more resistant to the actions of bile than 630Δerm. The transcriptional response of these strains was also investigated. The results of this study lay the foundation for further investigation of the role of bile in the colonisation and virulence of *C. difficile*.

A novel potential tool for random mutagenesis of *C. difficile* was developed using the mini transposon, mini-Tn*10*. Although it was possible to introduce the element into *C. difficile*, insertion of the element into the genome of several *C. difficile* strains was not demonstrated. This study highlights the potential of this type of tool, which, with further refinement, may be suitable for random mutagenesis of *C. difficile*. The development of a mini derivative of Tn*916*, mini-Tn*916* is also described.

Chapter 1

1. **Introduction**

1.1. Clostridium difficile

1.1.1. **The bacterium**

Clostridium difficile is a motile Gram-positive spore forming anaerobic bacterium identified in the 1930's by Hall and O'Toole and was originally known as *Bacillus difficilis* (Figure 1.1) [1, 2]. The bacterium is commonly isolated from faecal samples of neonates and the elderly, its carriage is often asymptomatic and is found in 2-3% of healthy adults and a high proportion of neonates (up to 64%) [3, 4]. In the 1970's an association of colitis and broad-spectrum antibiotics was observed; characterization of cytotoxin production identified *C. difficile* as the primary cause of pseudomembranous colitis (PMC) and as an important emerging pathogen [2, 5].

The genome of the pathogenic multidrug-resistant *C. difficile* strain 630 isolated from a hospital patient with severe PMC in Zurich, Switzerland has been fully sequenced [6]. The genome consists of a circular chromosome of approximately 4.3 Mb in length and a plasmid, pCD630 of 7.8 Kb (Table 1.1 & Figure 1.2) [6]. It encodes 3,776 predicted coding sequences (CDSs) with low G+C content with 82.1% of CDSs on the leading strand. It shares 15% of the CDSs with other sequenced clostridial genomes including *Clostridium acetobutylicum*, *Clostridium botulinum*, *Clostridium perfringens* and *Clostridium tetani*, encoding mainly essential functions [6-10]. Over 50% of *C. difficile's* genome is unique encoding many accessory functions and mobile elements. The genome contains several, often C+G rich, mobile elements that account for over 11% of the genome. These include six putative

conjugative elements and one proven conjugative transposon (CTn) (Tn*5397*), one mobilisable transposon (Tn*5398*) two prophages, a prophage-like sigK intervening sequence element and IStrons [6, 11-13].

Image unavailable due to copyright restrictions

Figure 1.1 Coloured transmission electron micrograph of *C. difficile* forming an endospore (red). Taken from Aslam *et al*., 2005 [14]

Clostridium difficile is a Gram-positive rod-shaped bacillus approximately 4-5μm in length. In this transmission electron micrograph image the spore can be seen clearly highlighted in red.

Feature	Chromosome	Plasmid
Size (bp)	4,290,252	7,881
G + C content (mole %)	29.06	27.9
Coding sequences	3,776	11
Coding density	0.087	1.39
Average gene size (bp)	943	563
Pseudogenes	32	0
rRNA operons	11	0
tRNA	87	0
Stable RNA	54	0

Table 1.1 General features of the genome of *C. difficile* adapted from Sebaihia *et al*., 2006 [6]

The table displays the features of the chromosome and plasmid, pCD630, of strain 630.

Figure 1.2 Circular representation of the genome of *C. difficile* taken from Sebaihia *et al*., 2006 [6]

The entire chromosome is represented by a series of concentric circles with features annotated using specific colours and position. Reading externally to internally the circles represent the following (adapted from Sebaihia *et al*., 2006): **1+2** = all CDSs (transcribed clockwise & anticlockwise); **3** = CDSs shared with other sequenced clostridia (blue); **4** = CDSs unique to *C. difficile* (red); **5** = mobile elements (red = conjugative transposons, pink = prophage, blue = skin element and unknown mobile element; magenta = genomic island carrying ethanolamine degradation genes, pale red = Tn*5398*, brown = partial prophage and CTns); **6 =** RNA genes (blue = rRNAs, red = tRNAs, purple = stable RNAs); **7** = G+C content (plotted using a 10-kb window); **8** = GC deviation ((G–C)/(G+C) plotted using a 10-kb window; khaki indicates values >1, purple <1). Genes are colour coded as follows: dark blue = pathogenicity/adaptation, black = energy metabolism, red = information transfer, dark green = surface-associated, cyan = degradation of large molecules, magenta = degradation of small molecules, yellow = central/intermediary metabolism, pale green = unknown, pale blue = regulators, orange = conserved hypothetical, brown = pseudogenes, pink = phage & IS elements, gray = miscellaneous.

1.1.2. *C. difficile* infection (CDI)

C. difficile symptomatic infection is almost always associated with the prior use of antimicrobial therapy which leads to disruption of the normal intestinal flora [15, 16]. In healthy individuals the normal gut flora provides an effective barrier to invading pathogens in a process known as colonisation resistance [15]. When this barrier is disrupted by administration of antibiotics, *C. difficile* is able to colonise the gut, specifically in the colon [17]. The organism is generally acquired from an exogenous source directly from an infected individual or indirectly from a contaminated environment due to its ability to produce highly resistant endospores [18]. The ability of *C. difficile* to persist in the environment has allowed it to become common in healthcare facilities and is now the leading cause of infectious diarrhoea in hospitals in developed countries [19, 20].

Infection with *C. difficile* leads to a spectrum of clinical outcomes ranging from asymptomatic colonisation to mild self-limiting diarrhoea and severe symptoms including abdominal pain, fever and leukocytosis [21]. In fulminant or severe disease, patients experience severe diarrhoea with inflammatory lesions and the formation of pseudomembranes in the colon (Figure 1.3) [17, 21]. Collectively known as *C. difficile* infections or CDI, there is the potential for life-threatening complications such as toxic mega colon or perforation, sepsis shock and death [16-18, 21, 22].

Image unavailable due to

copyright restrictions

Figure 1.3 Pseudomembranes formed on the internal wall of the colon. Taken from Rupnik *et al*., 2009 [17].

An endoscopic view of the pseudomembranes on the lining of the colon composed of destroyed intestinal cells and leukocytes.

Symptoms appear abruptly with production of explosive watery diarrhoea accompanied by abdominal pain [18]. *C. difficile* is widely regarded as the aetiological agent of PMC with 95 - 100% of cases testing positive for *C. difficile* toxinogenic strains in stool samples [2, 23]. It is unclear why one patient will develop PMC while others merely suffer from uncomplicated mild diarrhoea. It is suggested that host factors are more important in determining disease outcome than bacterial virulence factors since infants and asymptomatic carriers can be colonized with toxinogenic strains [2].

CDI or asymptomatic carriage of *C. difficile* has been identified in several animals including pigs, horses, monkeys, cats, and dogs [24]. At present there is no direct evidence of zoonotic transmission, however the situation is being closely monitored [25, 26]. *C. difficile* has also been isolated from a range of foodstuffs from America, Canada and Europe [27-33]. However, there is no direct epidemiological evidence to support human infections resulting from ingestion of contaminated foodstuffs [27].

1.2. CDI epidemiology, treatment and prevention

After its description in 1978, *C. difficile* was mainly associated with the development of antimicrobial-associated diarrhoea, colitis and PMC in the elderly with interest centred on its impact on morbidity and mortality in this group [5]. Over the last 10 years the incidence, clinical presentation, severity and epidemiology of CDI has changed significantly and *C. difficile* is now regarded as a major clinical problem [34, 35]. Healthcare associated infections (HCAIs) are a significant financial and social burden [36]. Methicillin-resistant *Staphylococcus aureus* (MRSA) was regarded as one of the most significant causes of HCAI-associated infections with mandatory surveillance introduced in 2001, shortly followed by glycopeptide resistant enterococci (GRE) in 2004 [37]. As a result of the increasing incidence of CDI and associated mortality, surveillance was declared mandatory in 2004 in the UK to facilitate epidemiological monitoring and analysis [38]. The number of deaths attributed to *C. difficile* over the last 10 years are approximately 5-fold higher than for MRSA (Figure 1.4) [39].

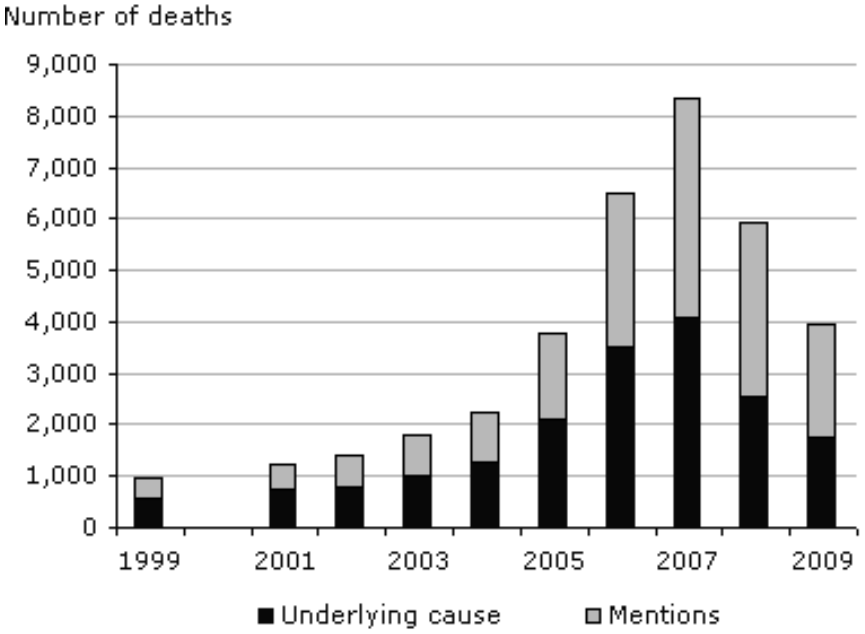
Although the epidemiology of this important nosocomial pathogen has changed dramatically, it is important not to over interpret the data. Serious outbreaks including two at Stoke Mandeville hospital in Buckinghamshire between 2003-6 have led to increased media attention and publicity of CDI [40]. Consequently, the education of health professionals has improved and brought *C. difficile* to the forefront of HCAI monitoring and prevention. The differing laboratory diagnosis methods and increased awareness may have

led to ascertainment bias during reporting [35]. Fortunately, the statistics issued for 2008-9 show that the incidence of CDI associated mortality is decreasing, although the rates are still relatively high compared to those of the early part of the century (Figure 1.4) [41].

1.2.1. **Emergence of hyper-epidemic strains**

The changing epidemiology of CDI highlighted the emergence of a new seemingly hypervirulent strain characterized as PCR ribotype 027, toxinotype III, REA group 1 (BI) and North American PFGE type 1 (NAP1) often referred to as 027/NAP1 isolates [42] [43]. Severe outbreaks of CDI between 2000 and 2005 in North America, in previously low-risk populations, have been attributed to these isolates, which prior to 2000 accounted for less than 1% of the isolates in the USA [20].

A



B

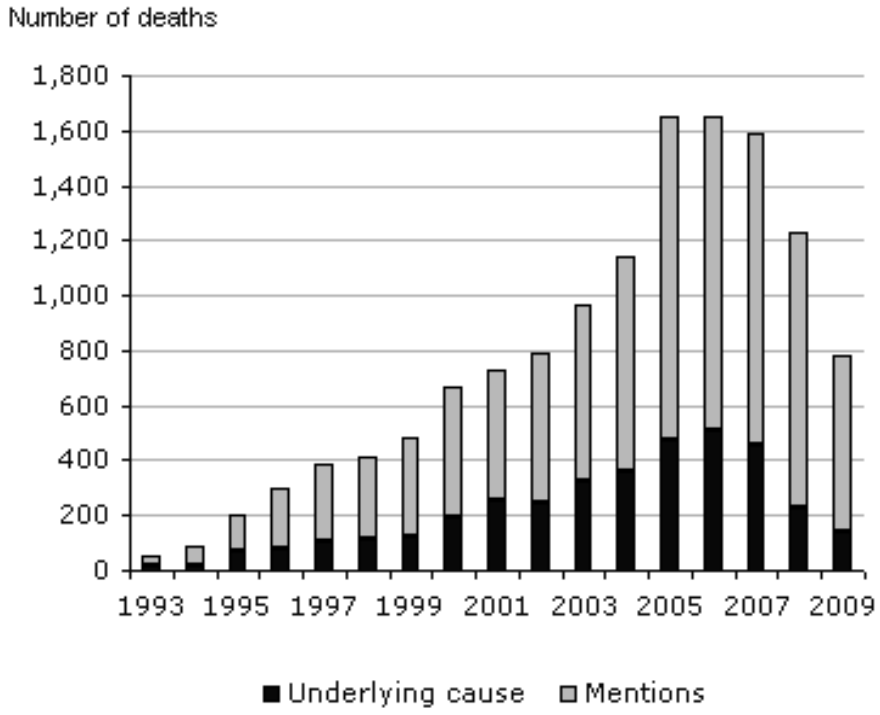


Figure 1.4 Deaths attributed to *C. difficile* (A) and MRSA (B).

The graphs show the total number of deaths reported to be due to or associated with *C. difficile* (A) and MRSA (B) in England and Wales. The reported rates for *C. difficile* are approximately 5-fold greater than those attributed to MRSA. Graphs taken from the Office for National Statistics [44].

Molecular characterisation of 027/NAP1 isolates revealed an 18 bp deletion (position 330-347) in the gene *tcdC*, a putative negative regulator of toxin gene expression (*tcdA* & *tcdB*) [43, 45]. This leads to an in-frame deletion of the *tcdC* transcript, however the protein encoded from this transcript is active *in vitro* and *in vivo* [46]. In addition to this in-frame deletion, a 1-base-pair deletion (Δ117) exists in the 5' end of *tcdC* leading to a frameshift and a premature stop in the early part of the gene [45, 46]. The resulting 65-residue fragment is unlikely to be functional disrupting the regulatory functions of TcdC [45]. These strains have been shown to produce high levels of TcdA and TcdB *in vitro* and it has been suggested that this may account for the increased severity of CDI arising from infection with 027/NAP1 isolates [43]. This theory was not supported by assessment of 027/NAP1 strains in a complex gut model; in these experiments the strain was shown to produce toxin for a longer period of time rather than producing more toxin per unit time [35, 47]. Consequently, some researchers term these strains as hyper-epidemic rather than hypervirulent strains. Furthermore, various deletions in *tcdC* have been observed in non-epidemic PCR-ribotypes, suggesting that disrupted toxin regulation is not solely responsible for the increased predominance of these strains [48].

027/NAP1 strains also produce a binary toxin (CDT), whose role in pathogenesis is unclear, and possess greater resistance to fluoroquinolones and third generation cephalosporins than older isolates of the same type and contemporary non-outbreak strains [35, 45, 49, 50]. It has been observed that 027/NAP1 strains have increased sporulation capacity, possibly aiding their

dissemination and persistence in the environment [51, 52]. 027/NAP1 isolates have spread worldwide, with significant outbreaks in North America and Europe [35, 43, 53-55]. The features of these strains described above, seem to have contributed to their rapid and widespread dissemination, however, other, as yet unidentified, factors may also be involved.

1.2.2. **Risk factors for CDI**

Several risk factors have been associated with the development of antibiotic-associated diarrhoea (AAD) (Figure 1.5), the foremost risk factor being disruption of the gastrointestinal (GI) flora. This is generally attributed to antimicrobial therapy, which destroys the colonization resistance of the bowel by killing the normal gut flora.[15, 56, 57]. Approximately 2-3% of healthy adults are asymptomatic carriers of *C. difficile* and this figure rises to 5-15% post antibiotic exposure, the subjects show no physical symptoms yet are positive for *C. difficile* toxins [42, 58].

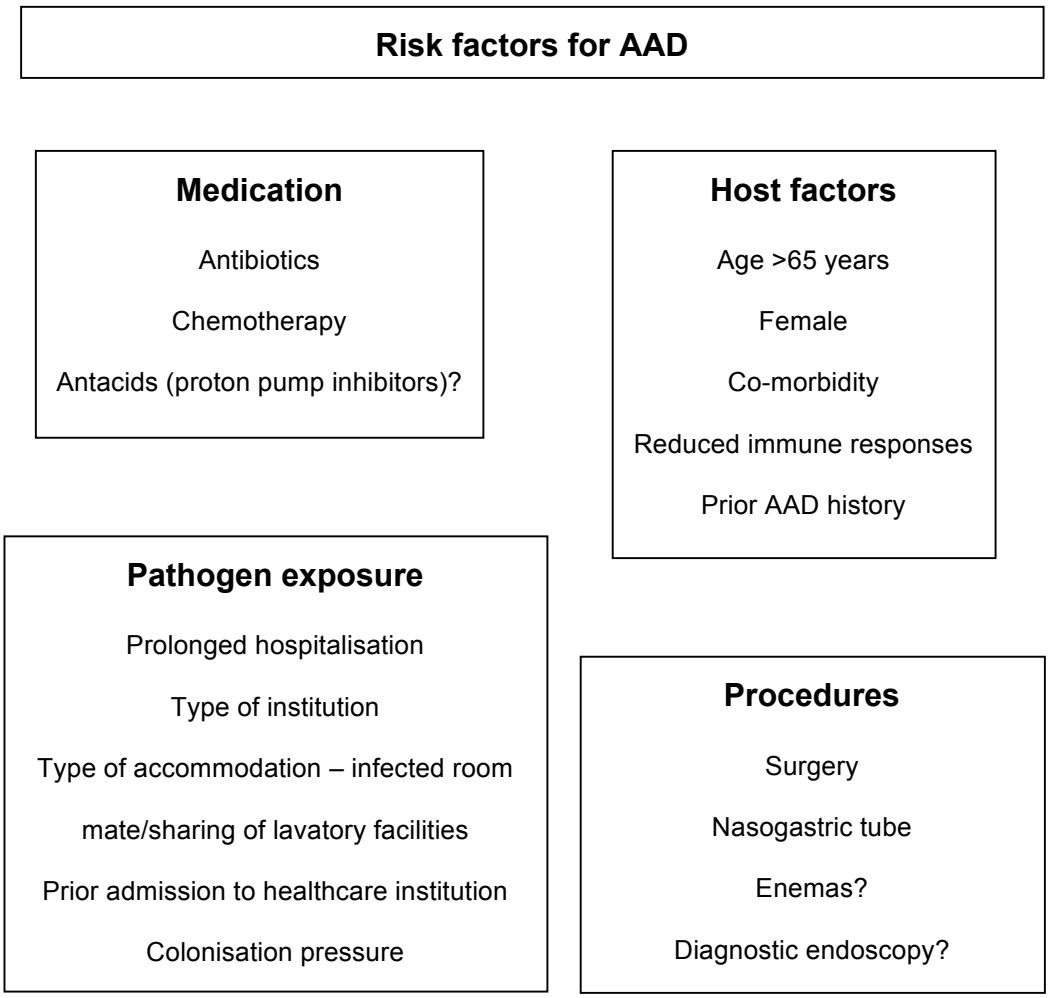


Figure 1.5 Risk factors associated with development of antibiotic-associated diarrhoea (AAD), adapted from McFarland, 2008 [59].

The major risk factors for the development of AAD, of which *C. difficile* is the leading cause, can be divided into 4 broad areas: medication, host factors, pathogen exposure and procedures which disrupt the natural colonic flora. The diagram lists some of the factors that have been associated with increasing the risk of developing AAD.

Almost all antibiotics have been associated with AAD and even short courses of antibiotics including perioperative prophylactic antibiotics, routinely given to patients undergoing surgery to prevent surgical site infections, can be associated with an increased risk of developing AAD [59-61]. Longer durations of antibiotic therapy may increase the relative risk of developing AAD [62]. Higher rates of AAD are associated with broad-spectrum antibiotics including penicillins (particularly ampicillin/amoxicillin), cephalosporins (particularly third generation), and clindamycin [56, 57, 63-66]. Recently, there has been increased interest in the use of proton pump inhibitors (PPIs) and the risk of CDI. However, there have been conflicting reports on whether there is a significant association [67-71].

Host factors are also very important when evaluating the risk of developing CDI. The predominant factor is age, CDI is more commonly associated with the elderly (>60 years) than with any other age group and the highest infection rates are seen in the over 65's [56, 63]. Co-morbidity has been shown to be significantly associated with CDI and this is not limited to the treatment of underlying diseases with antibiotics [59]. It is likely that there is an association with the immune response of the patients, which may be reduced due to the underlying disease and poor general health [59].

A strong association exists between hospital admission and *C. difficile* colonisation, with the duration of stay and stay in intensive care units increasing the chance of exposure to *C. difficile* [14]. It was thought that CDI occurred almost exclusively in healthcare associated environments, which

serve as a reservoir of infection due to the high density of at risk individuals. Recently, there have been increasing reports of community-acquired CDI, although the definition of such cases is not clear limiting the epidemiological information available [35]. Finally healthcare-associated procedures affecting the intestines can lead to disruption of the colonic flora and possibly introduction of C. *difficile* spores [59].

One of the major issues surrounding CDI infections is ease with which spores disseminate in the healthcare environment. Efficient and rapid isolation methods are imperative in containing the spread of the infection, however the eradication of spores and patient containment is a major problem in healthcare facilities. It is crutial that the infection is identified as early as possible to isolate the patient, introduce barrier-nursing techniques and change the treatment regime accordingly [2, 21].

1.2.3. **Treatment and Recurrence**

Promiscuous mobile genetic elements have contributed to the dynamic nature of many bacterial genomes, facilitating the generation of multi-drug resistant strains [72, 73]. *C. difficile* is a prime example of a bacterium which contains a plethora of mobile genetic elements originating from a variety of bacterial strains and genera [6, 74]. Elements frequently contain antimicrobial resistance genes with clinical isolates often showing resistance to tetracycline, macrolide–lincosamide–streptogramin B (MLS_B) and chloramphenicol [75]. Resistance to tetracycline in strain 630 is encoded by *tet*(M) found on CTn*3*/Tn*5397* and resistance to erythromycin results form the

presence of two copies of *ermB* on Tn*5398* [6]. Metronidazole and oral vancomycin are both effective agents for treatment of *C. difficile* infection, however there are reports of treatment failure, particularly with metronidazole [76]. Initially, the cumulative failure rates of these two drugs was almost identical (2.5% and 3.5% respectively), however since 2000 the cumulative failure rates for metronidazole have been reported as high as 18.2% [14, 34, 76, 77]. The 027/NAP1 strains in particular show a significant level of resistance to metronidazole [34, 77]. However, the current recommendations are to administer metronidazole as a first-line agent for mild infection due to its low cost and the concern over the increase in vancomycin-resistant nosocomial bacteria, with vancomycin only recommended for severe CDI [34].

An alternative treatment called tolevamer, a novel non-antibiotic polymer that binds and neutralises TcdA and TcdB, was compared with vancomycin and metronidazole treatment in two controlled phase 3 trials. The results of only one of the trials are available and it concluded that tolevamer was inferior to antibiotic treatment although a reduction in recurrence was observed [78].

A further complication of CDI pathology is the incidence of recurrence or relapse with reported rates between 8-25% of patients, although it has been suggested this may be as high as 50% [14, 21, 42]. Antibiotic treatment destroys the colonisation resistance of the colon facilitating recurrent infection, typically around 4 weeks after treatment [34]. The recurrence can result from reinfection or through persistence of the initial strain and may involve repeated episodes. This has also been confounded by the increasing CDI

incidence in previously low risk populations including children and pregnant women [79, 80]. Host immune response is thought to be an important factor in recurrence with higher levels of antibodies against TcdA and TcdB present in individuals who do not develop recurrent CDI [81].

Treatment with intravenous immunoglobulin has shown some success, although this has not been thoroughly investigated, and has prompted the development of a vaccine [14, 82]. Several vaccine types are being investigated including a toxoid vaccine, live vaccines with *Salmonella typhimurium* and *Vibrio cholerae* acting as vector strains expressing an attenuated TcdA, and conjugate vaccines combining the non-toxic peptide of TcdA with polysaccharides from other bacteria [83-86]. Finally, the use of probiotics as a preventative measure for CDI has also been investigated however, there is currently insufficient evidence to support their use, particularly due to the lack of standardization between the studies [87-89].

1.2.4. **Financial burden of CDI**

The increasing incidence of CDI along with other HCAI is a major burden worldwide both socially and economically [90]. It has been predicted that HCAI's cost the UK over £1 billion in 2005 and led to a number of likely preventable deaths [91]. In the UK the National Health Service has taken steps to minimise and effectively control HCAI's, including dedicated infection control staff to monitor and identify infections/risks, routine screening of patients for MRSA, the introduction of antimicrobial pharmacists, implementing intelligent hospital design and by the provision of

specialist isolation wards where possible [92, 93]. Attention has also been given to the further education of healthcare professionals to allow them to identify infections quickly and minimise risks.

1.3. Molecular pathogenesis and virulence

Bacterial pathogenesis relies on complex interactions between the host and the bacterium, with factors that aid the pathogenicity of an organism known as virulence factors [94]. Those factors that act directly to contribute to pathology are termed aggressins while others promote colonization thus enabling the production of aggressins [15]. Not all strains of *C. difficile* have the same pathogenic potential and it is estimated that 25% of strains are non-pathogenic and of those that produce toxins, a proportion do not cause disease [42].

1.3.1. **Cytotoxins and the PaLoc**

After colonisation, *C. difficile* can produce potent toxins, the enterotoxin TcdA and the cytotoxin TcdB, which are considered to be the bacterium's foremost virulence factors [95]. These toxins are largely responsible for the intestinal damage seen in symptomatic cases of CDI [96, 97]. They belong to a family of large toxins found in various members of the genus *Clostridium*; including lethal toxin (TcsL) and hemorrhagic toxin (TcsH) from *Clostridium sordellii*, alpha toxin (TcnA) from *Clostridium novyi*; and TpeL from *C. perfringens* [98-100]. Members are grouped due to their large size, between 250-308 KDa, and sequence identity ranging from 36-90% [101-104]. Both TcdA and TcdB are

monoglucosyltransferases that act intracellularly leading to the inhibition of members of the Rho family of GTPases, including Rho, Rac and Cdc42. These small regulatory proteins of the eukaryotic actin cell cytoskeleton are inhibited by glucosylation of a threonine residue which locks these proteins in an inactive conformation blocking all downstream signalling pathways (Figure 1.6) [101, 105, 106]. This leads to disruption of the actin cytoskeleton and tight junctions resulting in decreased transepithelial resistance, fluid accumulation cell rounding and eventually apoptosis of the exposed cell [97]. The toxins also cause the release of inflammatory mediators from intestinal epithelial cells, macrophages and mast cells. The cytokines affect enteric nerves and sensory neurones promoting an influx of inflammatory cells adding to fluid secretion and inflammation [17, 107].

Structurally, the toxins are formed of a C-terminal receptor binding domain, a central translocation domain and a catalytic N-terminal domain responsible for the toxic activity [97, 101, 105]. Originally this structure-function relationship was accepted due to similarity with the model of other AB-type toxins including diphtheria toxin, which included a biologically active domain and a binding or translocation domain [108]. However, Jank and Aktories (2008) propose that a multi-modular structure more accurately describes the structure function relationship of these toxins known as the ABCD domain structure model (Figure 1.7) [101]. The biologically active glucosyltransferase A-domain is located at the N terminus and includes the DXD motif involved in Mn^{2+} coordination. The cysteine protease C-domain, similar to *V. cholerae* RTX toxin, is located adjacent to the glucosyltransferase domain (DHC).

Image unavailable due to coryright restrictions

Figure 1.6 Cartoon of the mechanism of action of TcdA and TcdB taken from Carter *et al*., 2010 [101, 106].

Rho GTPases that control a number of signalling pathways in the cell. They are believed to be particularly important during host-pathogen interaction and considered essential for epithelial barrier functions, immune cell migration, adhesion, phagocytosis, superoxide production, cytokine secretion and immune cell signalling. Thus the presence of TcdA and TcdB prevent the normal functions of Rho GTPases leading to a breakdown in cell signalling and eventually death of the cell.

Image unavailable due to coryright restrictions

Figure 1.7 Proposed ABCD domain structure model of the clostridial glucosylating toxin TcdB taken from Jank and Aktories, 2008 [101].

The ABCD domain structure model of clostridial glucosylating toxin is shown for TcdB. The biologically active glucosyltransferase A-domain is located at the N terminus and includes the DXD motif involved in Mn^{2+} coordination. The C terminus (domain B) consists of polypeptide repeats and is involved in receptor binding. It is suggested that the DXG motif is possibly involved in processing of the toxins and that the hydrophobic region may be involved in pore formation and delivery of the catalytic domain into the cytosol of the target cell.

The central region (D-domain) contains a DXG motif, possibly involved in processing of the toxins, and a hydrophobic region that may be involved in pore formation and delivery of the catalytic domain into the cytosol of the target cell. The C terminus (B-domain) consisting of polypeptide repeats is involved in receptor binding, processing and cutting of the toxin. Crystallization of a fragment of the C-terminal repeats revealed a solenoid-like structure allowing the entire C-terminal structure to be deduced [101, 109]. The toxins have different tropism for host cells, TcdA has been shown to use glycoprotein gp96 as a co-receptor and binds more effectively to the apical side, whereas TcdB binds to an unknown receptor on the basolateral side of the host cell [17, 102, 110]. After binding, the toxins are endocytosed into the cell where proteolytic cleavage of the catalytic domain from the holotoxin occurs aided by the host derived co-substrate inositol phosphate; the catalytic domain is then transferred to the cytoplasm through a toxin-mediated pore [17, 111, 112].

The toxin genes and their regulators are located within a 19.6 Kb chromosomal pathogenicity locus known as the PaLoc which is found in the same chromosomal integration site in all toxinogenic strains. In non-toxinogenic strains (A^-B^-) this locus is completely absent and replaced by 115 bp of non-coding sequence [113, 114]. The PaLoc contains five genes, *tcdA* encoding toxin A, *tcdB* encoding toxin B, two regulatory genes *tcdC* and *tcdR* and, *tcdE* encoding a holin-like protein (Figure 1.8) [115]. TcdC is an acidic membrane associated protein thought to act as a negative regulator of toxin production [46]. It shares no homology with other previously described

regulatory proteins and recent investigations reveal a potentially unique mechanism of action [46]. TcdR, formerly known as TcdD, similar to the TetR and BotR regulators of tetanus and botulinum toxin respectively, is thought to be a positive regulator of toxin expression and activates toxin gene expression as an alternative RNA polymerase sigma factor [115, 116]. Expression of TcdR is also controlled by environmental conditions that can trigger or repress toxin gene expression [6, 117]. The *tcdE* gene encodes a protein with similarity to bacteriophage holins, which are cytolytic proteins that lyse bacterial hosts to allow the release of progeny phages; it is suggested that TcdE may function as a lytic protein facilitating the release of TcdA and TcdB to the extracellular environment [118].

Image unavailable due to copyright restrictions

Figure 1.8 Cartoon of the PaLoc of *C. difficile* taken from Rupnik *et al*., 2009 [17].

The PaLoc of *C. difficile* contains 5 ORFs encoding the cytotoxins TcdA and TcdB, along with the two regulatory proteins TcdC and TcdR. TcdE is thought to function as a lytic protein facilitating the release of TcdA and TcdB to the extracellular environment.

The sequence of the PaLoc is variable and strains displaying changes in this region are defined as toxinotypes. This method, developed by Rupnik *et al*. (1998) divides the PaLoc into ten regions of 2 Kb, PCR is used to amplify the regions which are subjected to restriction digest. The restriction profile length polymorphisms for two of the ten fragments are then inspected and strains assigned to the appropriate toxinotype [119, 120]. There are currently 10 groups (I to X) in addition to the reference strain VPI 10463 classified as type 0. It is not known whether this classification relates to disease outcome, with the exception of type III, which includes the 027/NAP1 isolates showing increased virulence [106, 121].

The majority of strains that cause disease in humans produce both toxins, however CDI can also be caused by strains producing TcdB alone (A^-B^+) [49, 122, 123]. Such strains account for a very small percentage of the *C. difficile* population, yet have been responsible for several hospital outbreaks in Europe and Japan [49]. They are able to cause the same spectrum of diseases, however, severe disease is more likely with high mortality rates (14-66%) [123]. It is not clear why these strains lead to a more severe disease although it has been postulated that this may be due to a lack of a TcdA-specific immune response, a key determinant in controlling disease severity [123]. TcdB has been shown to be more potent than TcdA in experiments on human tissue causing mucosal necrosis and decreasing barrier function [124]. Early studies suggested that TcdA alone was able to cause an enterotoxic effect when applied intragastrically in the hamster model; conversely TcdB had no effect when applied alone, symptoms were only evident when TcdB

was supplemented with TcdA or applied to previously damaged gut mucosa [125]. This suggested that TcdA was required as a co-factor for disease with the toxins acting synergistically [106].

The ability of either TcdA or TcdB alone to cause disease is a contentious issue. The Rood laboratory recently published a study suggesting that TcdB was essential for *C. difficile* virulence in the Golden Syrian hamster model, generally accepted as the most appropriate animal model for *C. difficile* pathogenesis [126-128]. Lyras *et al*. (2009) generated three isogenic mutants in an erythromycin sensitive derivative of strain 630, JIR8094, using a gene specific disruption system [126]. An unstable plasmid, designated as the recombination vector was introduced into *C. difficile* carrying homologous DNA sequence to a specific part of the genome. Disruption of a gene is achieved through insertion, via homologous recombination, leading to specific inactivation. Truncated, and inactive toxins were generated by insertion of the plasmid in the catalytic region of each of the toxin genes in multiple independent experiments. Each TcdA and TcdB negative mutant was assessed *in vitro* and *in vivo* using the Golden Syrian hamster model. The TcdA mutants (A^-B^+) were found to be highly toxic producing the same cytotoxic phenotype as the wild type strain *in vitro*. Challenge of the hamster model with this mutant led to colonisation and death (94% mortality) similar to that observed for the wild type strain, confirming that toxin B alone was able to cause disease and mortality. The TcdB mutants (A^+B^-) were shown to be significantly less cytotoxic *in vitro*, comparable to the avirulent strain phenotype, although a low level of cytotoxicity attributed to TcdA was

observed. Challenge of the hamster model with the TcdB negative mutants led to colonisation similar to that seen for the wild type strain however, colonisation did not lead to death in almost 80% of the animals. Together the data presented in this study suggest that TcdB is required for virulence in *C. difficile* [126].

A similar study by Kuehne *et al*. (2010) contradicts these results suggesting that both A^+B^- and A^-B^+ strains are able to cause colonisation and mortality *in vivo* [129]. In this study, the ClosTron, a targeted gene knock out system, was used to generate three isogenic mutants containing a single intron insertion in the catalytic domain of *tcdA* and/or *tcdB* rendering the toxins non-functional. Each mutant was examined *in vitro* using the cytotoxicity assay and *in vivo* using the Golden Syrian hamster model. Both the A^-B^+ and A^+B^- strains produced a cytopathic effect *in vitro* and led to colonisation and mortality *in vivo* confirming that TcdA and TcdB both have the potential to cause significant damage to intestinal epithelium leading to mortality in the hamster model. The double mutant (A^-B^-), with two intron insertions, showed a complete loss of toxicity *in vitro* and *in vivo,* as anticipated. Both authors thoroughly investigated the results of the *in vivo* work confirming, with an exception in the Lyras *et al*. study, the presence of only the inoculated strain post mortem [126, 129]. Additionally, both groups state that gene inactivation is stable, although hamster deaths in the Lyras *et al*. study for the A^+B^- test group have been attributed to revertants where the plasmid insertion has been lost [126, 130]. The authors of the later study suggest that inherent differences between the hamsters used and the origin of the erythromycin

sensitive A^+B^+ strain investigated may account for the discrepancies [129]. Both 630Δerm and JIR8094 were generated through serial sub-culture of strain 630 meaning that either one could have obtained one or more secondary mutations in the toxin genes [129, 131, 132]. Neither of these groups has complemented the mutants meaning that definitive conclusions cannot be drawn from these conflicting reports.

Interestingly both studies found that the TcdB mutants (A^+B^-) produced between 2.2-3.3-fold more TcdA than the wild type strains (A^+B^+), confirmed by cytotoxicity assays and quantitative real-time reverse transcriptase PCR (qRT-PCR). Although the Kuehne *et al*. study suggested that expression of both toxins was required to kill hamsters *in vivo*, the authors point out that the DNA sequence of the toxin genes, particularly TcdB, is highly variable in naturally occurring strains and that as yet no naturally occurring A^+B^- strains have been isolated from clinical samples [129]. It is not clear whether this is because such strains have not yet evolved or, that due to the routine diagnostic techniques used to identify *C. difficile* in clinical samples, these cases have been missed [129].

Toxin production varies from strain to strain and is influenced by environmental conditions [133, 134]. Expression is significantly increased during entry into stationary phase and can be regulated by environmental conditions; for example, an increase in temperture from 22°C to 37°C has been shown to increase toxin production [135]. It is postulated that toxin expression is autoregulated in response to environmental signals such as the

presence of nutrients in the growth medium including glucose, certain amino acids, butyric acids and butanol [97, 136]. The global regulator of gene expression, CodY, has been shown to negatively regulate toxin gene expression in *C. difficile* [117]. Inactivation of CodY led to repression of all five genes in the PaLoc during exponential growth and stationary phase in tryptone-yeast medium. It is suggested that *in vivo* CodY acts as a direct repressor of *tcdR* expression leading to downstream effects on *tcdB*, *tcdE* and *tcdA* expression. A link between toxin gene expression and nutrient limitation was also observed, GTP and branched-chain amino acids were shown to be co-repressors for CodY in *C. difficile*, as they are for CodY in *B. subtilis* [117]. Thus in nutrient limitation during colonisation, toxin expression may increase, leading to lysis of epithelial cells which release nutrients to support *C. difficile* growth [117]. Expression of CodY is high during the early exponential phase but declines during the stationary phase, a point at which expression of both toxin A and B increases [46, 134].

1.3.2. **Binary toxin**

In addition to the cytotoxins, a third toxin, binary toxin also known as *C. difficile* transferase (CDT), which is independent of the PaLoc, is produced by approximately 5-6% of identified strains, including the 027/NAP1 epidemic isolates [20, 43]. CDT is a two-subunit toxin which belongs to a family of binary actin-specific ADP-ribosyltransferases produced by a range of pathogenic clostridia including *C. perfringens* (iota toxin), *Clostridium spiroforme* (CST), and *C. botulinum* (C2 toxin) [137, 138]. The genes *cdtA* and *cdtB* encode the enzymatic and binding components respectively and together with the

regulatory gene *cdtR* are located chromosomally at the binary toxin locus (CdtLoc) (Figure 1.9) [17, 139]. The toxin acts through the binding of CdtB to host cells allowing the translocation of CdtA into the cytosol. Here CdtA ADP-ribosylates actin molecules, blocking polymerisation, inhibiting elongation and favouring depolymerisation of F-actin resulting in disruption of the cytoskeleton [17, 140].

Although the role of CDT in disease severity and clinical outcome is unclear, there is frequent co-expression of CDT in hypervirulent strains of *C. difficile* and CDT is usually produced by strains with variant yet functional *tcdA* and *tcdB* genes [140, 141]. Purified CDT has been shown to cause fluid accumulation and haemorrhage in ligated ileal loops and is cytotoxic to Vero cells *in vitro* [50, 142]. Additionally, the majority of isolates from animals have functional CDT, suggesting a role in the disease of animals [143]. However, naturally occurring CDT producing A^-B^- strains can colonise but are avirulent in the hamster model suggesting that CDT alone is unable to cause disease [50]. Recent investigation by the Aktories group demonstrated that CDT, iota toxin and C2 toxin could induce the formation of a novel type of microtubule structure; these long, thin and dynamic microtubule-based structures protrude from epithelial cells aiding the adherence of clostridial cells [140].

Image unavailable due to copyright restrictions

Figure 1.9 Binary toxin locus (CdtLoc) and toxin structure taken from Rupnik *et al*., 2009 [17].

CDT is encoded on a chromosomal locus CdtLoc, which is not connected to the PaLoc. CDT is composed of two unlinked proteins, CdtA the catalytic domain encoded by *cdtA* and CdtB the translocation and binding domain encoded by *cdtB*. Expression of these genes is regulated by cdtR [139].

1.3.3. **Surface layer proteins and adhesion**

Adhesion to host tissue is an important step in virulence for many pathogens. Successful colonisation by *C. difficile* requires adhesion of the cells to the gut mucosa to overcome host defence mechanism such as peristaltic movement and the associated flushing affects [15, 144].

Some strains of *C. difficile* have peritrichous flagella, considered an important adhesin, with flagellated strains showing a 10-fold higher level of adherence to mouse caecum compared to non-flagellated strains [145, 146]

C. difficile cells also produce an S-layer, a paracrystalline array of proteins outside the cell wall, possibly involved in attachment and phenotypic variation required to evade host defences [18, 147, 148]. There is significant variability between the surface proteins of different strains, particularly in surface layer protein A (SlpA) [17]. The layer is composed of two protein subunits (one high- and one low-molecular weight) encoded by a single gene *slpA*; 28 paralogues of this gene exist in the genome including those encoding Cwp66, a heat-shock inducible adhesin and Cwp84, a cysteine protease [6, 149, 150]. Cwp66 has been shown to mediate binding to Vero cells (kidney epithelial) *in vitro* after heat-shock [151, 152]. It is hypothesised that SLPs may be involved in pathogen-host interactions critical to pathogenesis due to their abundance and localisation on the outer surface of the cell [153]. Evidence for their role in the adhesion of *C. difficile* has been demonstrated *in vitro* and *in vivo* in the Golden Syrian hamster model [152, 154].

1.3.4. **Sporulation and germination**

C. difficile produces highly resistant spores that are able to withstand desiccation, chemicals including alcohol and certain disinfectants, oxygen, and temperature extremes [155-158]. The metabolically dormant spores are responsible for person-to-person transmission due to their inherent ability to persist in the environment. Sporulation occurs when vegetative growth cannot be maintained. The process of sporulation in *C. difficile* is poorly understood, although some parallels have been drawn with the process in *B. subtilis* [155]. Asymmetric cell division leads to the generation of thick concentric external layers - the spore coat and cortex, which protect the cytoplasm and allow the bacteria to survive in harsh environments (Figure 1.10) [155]. The initiation of sporulation may be based on a range of environmental and physiological signals including nutrient limitation, cell density and DNA damage and synthesis [159].

Figure 1.10 Structure of the *C. difficile* spore taken from Lawley *et al*., 2009 [155].

A transmission electron micrograph of a sectioned *C. difficile* spore. The spore ultrastructure including the exosporium, coat, cortex, core, membrane, and ribosomes can be seen. The black bar represents 100 nm.

In *B. subtilis* two regulatory systems control spore formation, the transcription factor Spo0A and a series of alternative sigma factors of RNA polymerase, which appear to be conserved in all sequenced species of *Bacillus* and *Clostridium* [159-161]. True Spo0F and Spo0B homologues, based on protein sequence homology, have not been found in the genomes of sequenced clostridia [162]. It is believed that in the clostridia the sporulation initiation pathway has remained like a two-component signal transduction system where signals are sensed by sporulation associated sensor kinases that phosphorylate Spo0A directly, and has not evolved into a multicomponent phosphorelay present in *Bacillus* [162-164]. Spo0A is thought to be used by spore-forming bacterial pathogens to modulate their virulence and survival responses, for example in *C. perfringens* Spo0A is involved in sporulation and toxin production [163]. In *C. difficile*, however, the link between sporulation and toxin production has not been fully investigated [133, 165-167]. A recent study by Underwood and colleagues (2009) found that inactivation of Spo0A using the ClosTron led to a number of effects including a marked reduction in toxin production [162].

Similarly, the process of spore germination is also poorly understood. Once ingested, germination occurs in the anaerobic environment of the large intestine in response to external stimuli. In *Bacillus* and many *Clostridium* species spores germinate in response to amino acids, nucleosides, carbohydrates or potassium ions. Interestingly, *C. difficile* spores have shown an increased level of germination in response to cholate derivatives

found in bile, although this has been disputed (see chapter 3 for more details) [168-171].

Persistence is a major issue in healthcare environments, limiting effective control and confinement of infection [172]. Interestingly, sub-inhibitory concentrations of non-chlorine based cleaning agents actually increased sporulation capacity [51, 172]. Evidence suggests that epidemic strains of *C. difficile* including ribotype 027 and 001 have an increased inherent sporulation capacity which is likely to have contributed to the efficient dissemination and spread of these strains [51, 52, 162]. However, researchers at the Minton laboratory question these results, they did not observe a difference in sporulation of the epidemic strains compared to standard laboratory strains [173, 174].

1.3.5. **Additional putative virulence factors**

The role of other factors in *C. difficile* virulence is less well understood. These may include chemotaxis, presence of flagella and capsule, and processes involved in iron sequestration, hyaluronidase (hydrolytic activity), proteases and collagenase activity [15, 18]. Genomic analysis has revealed that *C. difficile* has the potential to utilize a wide range of carbohydrates due to the presence of various genes required for metabolism and transport [6]. This may be an important feature enabling the bacteria to survive intestinal passage. It has also been observed that *C. difficile* contains two orthologues of genes found in *Listeria monocytogenes* that are thought to be responsible for bile tolerance; these genes have not been identified in other sequenced

clostridia [6]. Resistance to bile is an essential feature of GI flora and similarly of enteric pathogens allowing colonization and survival [175]. Bile is a potent detergent enabling the solubilisation and emulsification of lipids. It acts as an antimicrobial agent causing cell lysis by dissolving phospholipids, cholesterol and proteins of the cell membrane [175]. Therefore, the ability of *C. difficile* to avoid lysis and damage to the cell wall is an important colonisation factor. Although the mechanism of this resistance is not fully understood, comparisons and suggested mechanisms can be inferred from other Gram-positive bacteria. It has also been postulated that the presence of bile and bile salts may trigger the germination of spores when entering the small intestine, as they have been shown to enhance germination *in vitro* [176].

1.4. Tools for genetic analysis and manipulation

Progress in understanding the molecular pathogenicity of CDI has been hindered by the lack of genetic tools for effective gene transfer [18]. Historically, the genetic manipulation of *C. difficile* has been extremely difficult, it has not been possible to introduce DNA through electroporation or chemical transformation [177]. However, recent advances including genome sequencing projects and the development of genetic tools such as shuttle vectors and the ClosTron have resulted in some significant advances in *C. difficile* research [130, 132, 177].

Advances have been made in the genomics field, facilitating rapid and cost effective sequencing of entire bacterial genomes. Sequence information has

provided crucial insights into the make-up and organisation of the genome of *C. difficile*. This has allowed comparisons with other well-studied organisms to identify putative virulence-associated genes and investigation of the evolution of *C. difficile* strains [6, 178].

Like the genomes of many other bacteria, the clostridial genome exists in a fluid and dynamic state due to the presence of extra-chromosomal elements such as plasmids and the presence of mobile genetic elements such as transposons (Tns). Mobile genetic elements, mainly conjugative transposons (CTns) make up almost 11% of the genome of *C. difficile* strain 630 and are predominantly responsible for the varied and wide ranging antibiotic resistance phenotypes seen in this strain [6].

1.4.1. **Transposons (Tns)**

Tns are mobile genetic elements made from discrete fragments of DNA that are able to relocate from one genomic location to another [179]. Transposable elements, first observed by Barbara McClintock, are ubiquitous from bacteria to human cells and have significantly contributed to genome content, plasticity and evolution [180]. Their movement is tightly regulated and their presence can significantly influence gene expression. Host organisms have developed a range of defence mechanisms to reduce their movement, which can be potentially harmful, such as DNA methylation and RNA interference to reduce the expression of transposon encoded proteins [181, 182].

Transposable elements and their transposition processes have been well studied in genetically amenable bacteria. They have been used to develop molecular tools for random insertion mutagenesis, facilitating the study of specific genes and genome organisation [179]. The use of Tns as mutagens has been focused on well-characterised elements from Gram-negative bacteria, with those that insert at random or near random being preferred. Other mobile elements play a role in genome plasticity and evolution including plasmids, mutator bacteriophage, integrons, mobile introns, shufflons, conjugative transposons, integrative conjugative elements and pathogenicity islands.

Transposable elements range in complexity; simple elements are known as insertion sequence (IS) elements consisting of a gene(s) for transposition bounded by inverted repeat sequences such as IS*1*. Composite transposons, such as Tn*5* and Tn*10*, contain additional genetic information such as antibiotic resistance genes, which are flanked by IS elements (Figure 1.11). More complex elements, including Tn*916*, have hybrid qualities of transposons, plasmids, and bacteriophages [179].

The composite Tn, Tn*5*, contains genes for resistance to kanamycin, bleomycin, and streptomycin and two inversely orientated IS*50* elements. The IS*50* elements are not identical, with IS*50*R encoding functional proteins for transposition, while IS*50*L encodes non-functional truncated proteins [183].

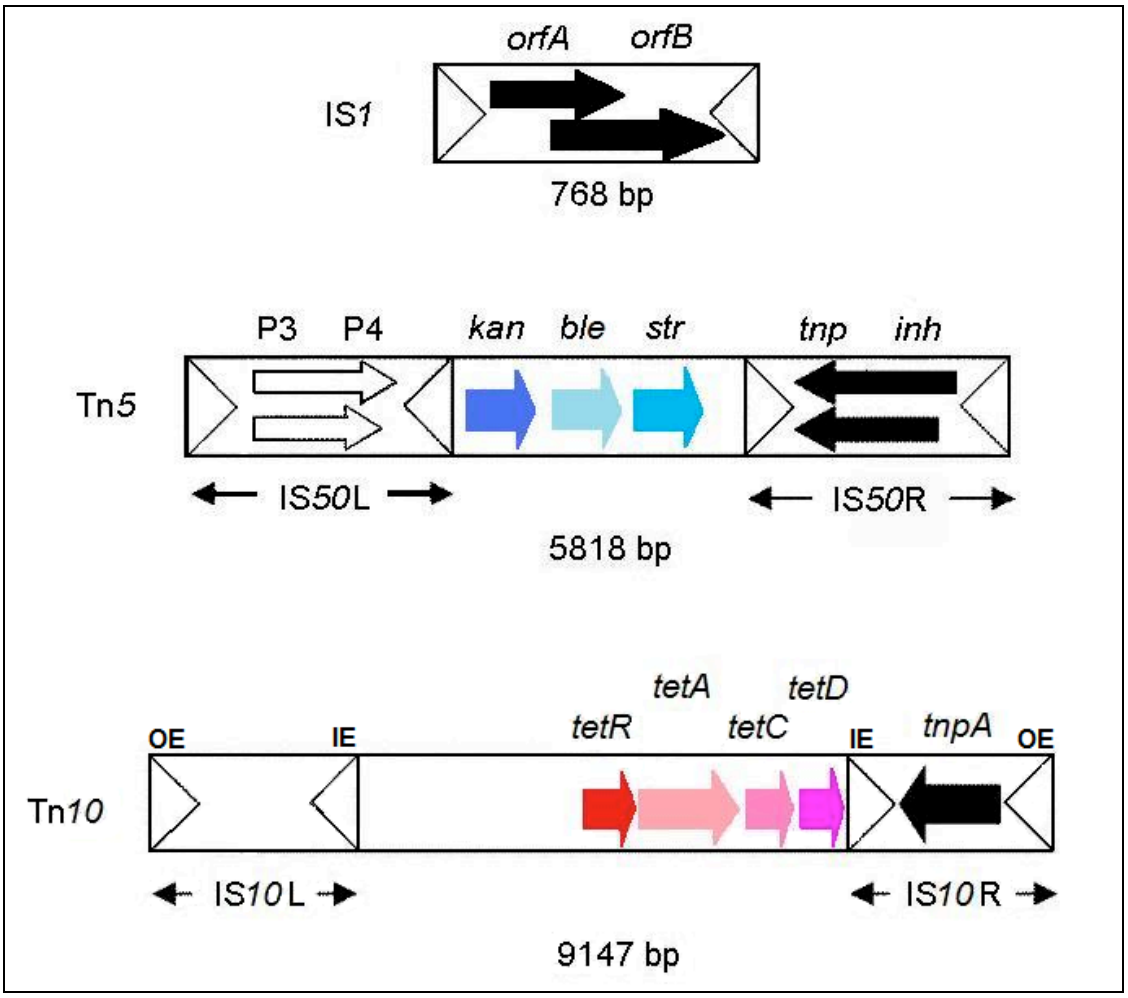


Figure 1.11 Schematic diagram of the genetic organisation of IS*1*, Tn*5* and Tn*10* adapted from Hayes, 2003 [179]

IS*1* is an example of a simple transposable element. The triangles represent repeat sequences and the black arrows represent genes involved in the transposition of the element. Tn*5* and Tn*10* are composite transposons, which have additional genes including antibiotic resistance genes (shown by coloured arrows) flanked by two inversely orientated copies of IS*50* and IS*10* elements, respectively.

The overlapping genes on IS*50*R produce transposase (Tnp) and a truncated version of Tnp (Inh) which downregulates Tnp, thus inhibiting transposition to promote movement of the transposon in the genome. Each of the IS*50* elements contains a 19 bp repeat sequence at its termini, which are crucial Tnp binding sites. Tn*5* is an excellent model for understanding cut and paste transposition where the element is precisely excised from its current position and inserted into a new target location via a stepwise process [184, 185]. The composite transposon Tn*10* will be discussed in detail in chapter 4.

1.4.1. **Conjugative transposons (CTns)**

CTns are a family of mobile genetic elements that are able to transfer themselves from the genome of a donor cell to another (recipient) via intimate contact using a conjugation-like mechanism. The mechanism of transposition is unique to this group, moving by excision from the donor DNA molecule, circularization and insertion into the target site of the recipient [186-188]. These elements are found extensively within bacteria and are highly promiscuous, spanning large phylogenetic distances [187]. CTns are considered to be one of the major vectors responsible for the spread of antibiotic resistance, by direct transfer and mobilization of other plasmids and Tns [189-191]. *C. difficile* strain 630 has six putative CTns, and the extensively studied Tn*5397* which mediates tetracycline resistance [6].

Tn*916* (18.3 Kb) was the first CTn to be identified, it was found in the chromosome of haemolytic multidrug-resistant *Enterococcus faecalis* DS16; the transposon has a broad host range and is responsible for the transfer of the tetracycline-resistance determinant *tet*(M) [189]. A closely related CTn,

Tn*1545* from *Streptococcus pneumoniae*, has been shown to behave similarly to Tn*916,* carrying the *tet*(M) determinant and additionally erythromycin (*ermAM*) and kanamycin (*aphA*-3) resistance determinants [191, 192].

Due to its broad host range Tn*916* has been used as a mutagen in several different species and for the introduction of heterologous DNA [193]. Through filter-mating with *Bacillus subtilis* it is possible to introduce Tn*916* and possibly other transposons into the *C. difficile* genome, although transfer frequencies are relatively low (10^{-8} per donor) [194]. It has been discovered that Tn*916* preferentially inserts into AT-rich sites that have limited sequence homology with the ends of the element [195]. In *C. difficile* Tn*916* was shown to insert into a single site in strain CD37 and multiple sites in strain 630 and CD79-685 [193]. Due to the insertion site preferance, particularly in CD37, its suitability for random mutagenesis is limited [131, 195].

1.4.2. **Other transposon-based mutagenesis systems**

A number of Tns have been investigated for their use as mutagens of Gram-positive bacteria, two examples are miniTn*10* (a derivative of Tn*10*) and *mariner*-based Tns [196]. MiniTn*10* is a composite transposon harbouring an antibiotic resistance cassette flanked by IS*10* elements in inverted orientation (discussed in detail in chapter 4) [197]. The transposase gene is located outside the Tn ends allowing stable integration and they have been used to create insertion libraries in *B. subtilis* and *Bacillus anthracis* with high frequency random insertions [197, 198]. *Mariner* Tns are part of the *mariner*/Tc1

super family, the most widespread and diverse class II transposable elements, they have broad host range and are found extensively in eukaryotes including humans [199]. *Mariner* Tns such as *Himar1* and their derivatives undergo *in vitro* and *in vivo* transposition into several bacterial chromosomes. Random and disperse insertion libraries have been created in *Escherichia coli*, *Campylobacter jejuni* and *Leptospira interrogans* [199-202].

1.4.3. **Transfer of replication-deficient and replication-proficient plasmids**

Methods have also been developed to introduce heterologous DNA and recombinant DNA into *C. difficile* using Tn*916.* The fact that Tn*916* enters the genome at a hot spot in strain CD37 can be exploited for gene cloning, although the process is cumbersome and relatively inefficient [193]. This strategy can also be used to introduce antisense RNA in order to down-regulate the expression of specific genes, however initial experiments have been disappointing [203]. Historically, conjugal transfer of DNA has been conducted using *B. subtilis* donors. Research in other clostridial species has shown this strategy of using conjugative mobilization of *oriT* based vectors from *E. coli* donors to be extremely effective [204]. A number of shuttle vectors have been constructed by cloning clostridial replicons into the replication-deficient ColE1-based vector pMTL31, which can act as suicide vectors and facilitate integration of the plasmid into the genome by single cross over. By cloning a fragment of the recipients chromosome in the plasmid polylinker it is possible to obtain gene knockouts [132, 177]. The use of autonomously replicating (replication-proficient) plasmids has also been investigated. An

efficient replicon has been identified on an indigenous *C. difficile* 7 Kb plasmid pCD6 originating from strain CD6. This replicon contains an extensive repeat region thought to represent the origin of replication and allows replication of constructed plasmids within *C. difficile* [177, 205]. Original experiments to introduce such plasmids revealed that as with other clostridia, it is necessary to elucidate and overcome restriction modification systems. This may be achieved through methylation of the vector DNA prior to transfer, however both strain CD37 and 630 seem to show no restriction barriers. Overcoming these barriers in other strains is, however, proving difficult due to the presence of unique restriction systems, hindering the progression of this system in *C. difficile*. Other aspects also warrant further investigation, including the development of efficient plasmid delivery systems to make this an effective strategy for gene-knockout [205].

1.4.4. **Use of a group II intron derivatives for targeted mutagenesis – The ClosTron**

A recent major development in the manipulation of clostridia is the ClosTron. This system, based on the TargeTron™, enables directed inactivation of specific genes. The system exploits mobile group II introns from the *ltrB* gene of *Lactococcus lactis* (L1.ltrB) that are able to integrate at specific sites independently of host factors. *In-silico* analysis of the potential target sites allows the ClosTron to be targeted to a specific gene of interest. To enable screening of mutants the group II intron contains an erythromycin resistance gene itself disrupted with a group I intron. The specific arrangement of the three elements ensures that expression of erythromycin resistance only

occurs when the group II intron is in the target site at which point the group I intron will have spliced out leaving an intact antibiotic resistance gene. This is known as a Retrotransposition-Activated Marker (RAM) and is an important feature of the ClosTron. This is a relatively rapid process allowing mutants to be generated in 10 to 14 days and has proved effective in a number of clostridia including *C. difficile* [130, 206].

1.5. Future perspectives

The changing epidemiology and emergence of hyper-epidemic strains has brought CDI to the forefront of nosocomial infection monitoring and prevention in the UK. The increasing burden of infections has spurred increased interest in the study of the molecular basis of pathogenesis of *C. difficile*. Recently, there have been significant breakthroughs in the generation of tools to allow the genetic manipulation of *C. difficile*. The ClosTron and the development of recombination plasmids allow targeted gene disruption in clostridia, although these methods may still require refinement. There is, however, a clear need for the development of techniques to enable random mutagenesis to allow the discovery of novel virulence factors and to elucidate the mechanisms of gene regulation. This following study describes the investigation of existing and novel tools for random mutagenesis of *C. difficile*.

Chapter 2

2. Mutagenesis of *C. difficile* with Tn*916*

2.1. Abstract

Transposon mutagenesis is an effective method for random mutagenesis of bacterial genomes. In *C. difficile* the use of this routine technique has been hindered by the lack of genetic amenability shown by this important nosocomial pathogen. The development of a CTn based system has, to some extent, overcome these difficulties. Although not ideal, this tool has proved successful in the generation of an insertion library in strain 630Δerm using Tn*916*ΔE. This chapter describes the generation and analysis of a Tn*916* insertion library in the hyper-epidemic, ribotype 027 strain, R20291.

2.2. Introduction

2.2.1. **Transposon mutagenesis**

Mutational analysis has been a popular and successful method for the identification of microbial virulence determinants. The approach can be either directed or random. Directed mutagenesis seeks to confirm the role of a putative virulence determinant identified as being important in pathogenesis using a reverse genetic approach. In contrast, random mutagenesis is known as a forward genetic approach as no assumptions are made about the role of genes in pathogenesis. Large libraries of mutants are generated using a mutagen that can then be screened for differential phenotypes associated with virulence. Further characterisation of mutated genes can allow the identification of novel virulence factors [207]. Both approaches have advantages and disadvantages and the appropriate method should be determined based on the tools and facilities available to yield the best possible results (Table 2.1).

Mutation approach	Advantages	Disadvantages
Directed	Specific	Relies on assumptions about the role of the gene in virulence, based on circumstantial evidence
	Can enable identification of downstream effects	Can be time consuming to obtain desired insertion
	Can be designed to ensure minimal downstream effects	Tools may not be readily available
	Potential to allow mutations in specific regulatory regions affecting the expression of multiple genes	Unlikely to identify novel factors
		Requires sequence information
Random	Numerous mutagens available e.g. Chemical (e.g. Base-Analog Mutagens, deamination), Physical (e.g. UV), Biological (e.g. transposons)	Large libraries may need to be constructed - labour intensive, and may lead to time and storage issues
	Potential to identify multiple genes associated with virulence	Can lead to downstream effects particularly when transposons are used
	Potential to identify novel factors and genes involved in regulation of virulence factors	Multiple mutations can be made leading to difficulties in characterisation
	Transposon mutagens can be tagged to allow easy identification	The choice of mutagen is important as some mutations may not be stable, or frequency of mutation can be low
	Can be relatively inexpensive and quick	Is a 'hit and miss' approach - there is no guarantee that the library will contain mutants associated with virulence

Table 2.1 Examples of the possible advantages and disadvantages of directed and random mutagenesis approaches [94].

The table highlights some of the advantages and disadvantages, which should be considered when designing a mutagenesis strategy.

Tn mutagenesis is a random (forward genetic) approach based on the ability of transposons to insert into the genomes of many bacterial species, often at random. A wealth of knowledge about bacterial virulence factors, particularly in important human pathogens, has been generated through transposon mutagenesis [179, 207]. Tn mutagenesis offers the potential to generate random single-site mutations relatively quickly and cheaply, with the added benefit that each mutation is tagged by the Tn facilitating mapping and subcloning [207].

2.2.2. **Conjugative transposons and *C. difficile***

For *C. difficile*, the ability to use standard Tn-based random mutagenesis has been hindered by the apparent fragility of this bacterium. Introduction of DNA through electroporation or natural competence has not been demonstrated in this organism; however, utilisation of the self-mobilising properties of CTns has been exploited to allow mutagenesis in *C. difficile*. Mullany *et al*. (1991) developed a method for the introduction of the conjugative transposon Tn*916* into *C. difficile* through filter mating with a *B. subtilis* donor cell containing a Tn*916* [194].

CTns are important determinants of antibiotic resistance, these promiscuous elements are able to conjugate between bacteria of different species and genus [187]. This ability to disseminate among bacteria, with carriage of antibiotic resistance genes, make CTns a major clinical problem.

CTns also referred to as integrative conjugative elements (ICEs), encode proteins required to promote their own transposition [208]. Unlike other common transposons including Tn*5* and Tn*10*, their mechanism of excision and integration involves the generation of a covalently closed circular transposition intermediate before insertion at the target site (Figure 2.1) [188].

2.2.2.1. Tn*916*

Tn*916* is an 18 Kb conjugative transposon originally discovered in the late 1970's when tetracycline resistance was found to be transferable from the haemolytic multidrug-resistant bacteria *E. faecalis* DS16 to *E. faecalis* JH2-2 [189, 209]. This well studied element is made up of four functional modules: conjugation, regulation, recombination and accessory genes, the last of which is not involved in conjugative transposition (Figure 2.2) [210]. Tn*916* encodes tetracycline-minocycline (*tet*(M)) encoding tetracycline resistance [191]. The host range of this element is exceptionally broad and it is widespread among many commensal and pathogenic bacteria [189, 191]. Sequence analysis of many bacterial species has resulted in the identification of numerous Tn*916*-like elements with a common core genome containing the conjugation and regulatory regions but with numerous variations in the surrounding regions [210].

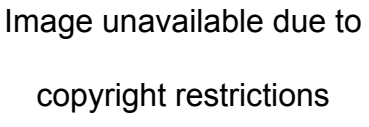


Figure 2.1 Mechanism of intercellular transposition of conjugative transposons. Taken from Salyers *et al*., 1995 [188].

The integrated form excises to form a covalently closed circle – the transfer intermediate, by staggered nicks at each end of the transposon. A single strand is nicked at the *oriT* site and is transferred to the recipient cell via a conjugation mechanism similar to plasmid transfer. A double-stranded form is then regenerated in each cell. The transposon is then able to integrate into the recipient cell's genome.

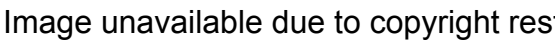


Figure 2.2 Schematic representation of Tn*916* taken from Roberts and Mullany, 2009 [210].

Tn*916* contains 4 functional modules: recombination, regulation, conjugation and accessory genes.

Excision of Tn*916* from its insertion site is mediated by staggered nicks at each end of the transposon leaving single stranded non-complementary 6-nucleotide stretches called coupling sequences at each end. This leads to the generation of a nonreplicative circular intermediate with mismatches at its joint (Figure 2.3) [187, 188, 211]. A single strand of the circular intermediate is transferred to a suitable recipient by a mechanism analogous to plasmid transfer [187, 188]. Cell to cell contact is required for transfer of the element, however the mechanism is not fully understood [208]. A second strand is synthesized in both the donor and recipient cell generating a double stranded circular molecule capable of insertion in both cells [210]. The latter process is mediated by the recombination module located at one end of the element, this contains a gene encoding a tyrosine recombinase, Int (required for integration and excision), and the accessory protein Xis (required for the directionality of the recombination reaction, usually promoting excision). The element can insert into a large number of target sites, yet in general has preference for AT rich targets [210, 212]. In some hosts it has a highly preferred site, for example in *C. difficile* strain CD37 insertion occurs at a single particular site [194, 195]. Insertion site sequence is not the only determining factor, experiments in other *C. difficile* strains reveal no apparent specific insertion site preference even though the insertion site from CD37 is present [210].

Figure 2.3 Proposed mechanism for excision and integration of Tn*916*. Taken from Salyers *et al*., 1995 [188, 211].

Staggered cuts occur 6 nucleotides from each end of the element producing single stranded overhangs known as coupling sequences. These ends are joined covalently to form a circular transposition intermediate. The bulged area represents areas of non - base pairing. Staggered cuts then open the target site and the circular intermediate before ligation of the ends into the target site. This generates regions of non-base pairing at each end, which are repaired by mismatch repair or by replication through the region.

2.2.2.2. Tn*1545*

The conjugative transposon Tn*1545*, originally identified in *S. pneumoniae* has ends identical to Tn*916* and behaves in a similar manner [189, 191, 213, 214]. Tn*1545* is larger than Tn*916* (25 Kb vs 18 Kb) and in addition to *tet*(M) contains erythromycin (*ermAM*) and kanamycin (*aphA-3*) resistance determinants [188, 189, 215]. It has been used successfully in *Clostridium cellulolyticum* although a low frequency of insertion was observed (transconjugants were obtained at frequencies of between 1.7 x 10^{-6} and 3.8 x 10^{-8} per recipient) [216].

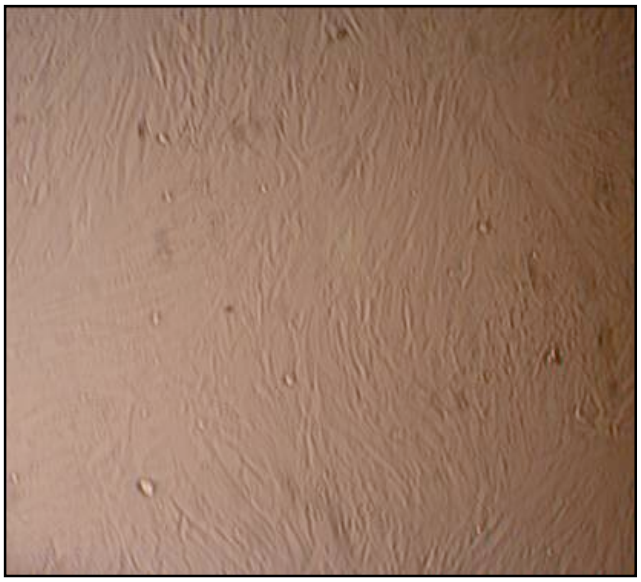
2.2.3. **Screening assays for *C. difficile* transposon libraries**

In *C. difficile* the production of toxins TcdA and TcdB is known to be important for virulence [18, 97]. Less definitive information is available to support the importance of other putative virulence factors involved in pathogenesis and survival in the gut. A library of *C. difficile* Tn*916* transconjugants can be screened for a myriad of phenotypes *in vitro* and potentially assessed *in vivo* in the Golden Syrian hamster model. High throughput and relatively simple assays are an effective method for screening large numbers of samples in a relatively short period of time. Ideally, the assays should provide definitive results quickly and reproducibly to avoid the requirement for duplication or speculative conclusions.

2.2.3.1. Detection of *C. difficlle* toxins

Assessment of each transconjugant for alteration in toxicity is a useful and informative assay, potentially providing insights into the genes involved in toxin production and regulation. The cytotoxicity assay is reputed to be the gold standard assay for the identification of *C. difficile* in clinical samples using cell lines including fibroblasts, CHO, Vero, HeLa, and other epithelioid lines [217-220]. The presence of the cytotoxins TcdA and TcdB can be easily visualised using an inverted microscope, confluent cell monolayers exposed to the cytotoxins produce a cytopathic effect (CPE) where cells round-up, detach and shrink loosing adherence (Figure 2.4) [221]. This relatively straightforward, yet labour intensive assay allows the clear identification of the presence of toxin, and with dilution series' estimations of relative toxin quantities can be determined between samples. Specific *C. difficile* or *C. sordellii* antitoxin is used to validate the assays specificity through neutralisation [218]. The presence of *C. difficile* toxins can also be assessed using an enzyme immunoassay (EIA). These rapid and easy tests allow diagnosis more rapidly than with the cytotoxicity assay. EIA also removes the need for tissue culture facilities, which are expensive and the results are also easier to standardise [222]. There are several commercial kits, proven to be both sensitive and specific, for the detection of TcdA and/or TcdB and current NHS recommendations are to use either of these assays in the clinical diagnosis of *C. difficile* [38, 222, 223]

A



B

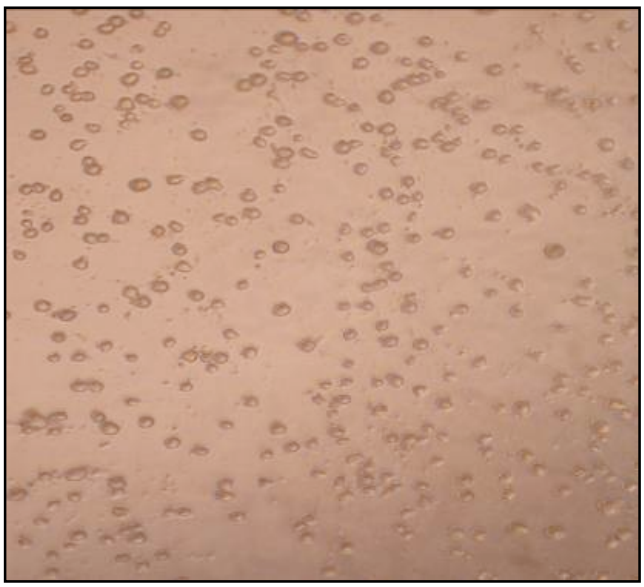


Figure 2.4 Visualisation of the cytopathic effect.

Photos courtesy of H. Hussain, UCL Eastman Dental Institute.

A – Healthy human foreskin fibroblasts (HFF-1) cells grown to a confluent monolayer viewed by a light microscope.

B – HFF-1 cells treated with a 30-fold dilution of cell free growth medium taken from 630Δerm cultures grown in rich growth medium for 24 hours. Cells are displaying the typical cell rounding and detachment attributed to the cytopathic effect (CPE) of TcdB and TcdA of *C. difficile*.

2.2.3.2. Sporulation/germination

The metabolically dormant *C. difficile* spore is responsible for person-to-person transmission due to its high level of resistance to oxygen, desiccation, temperature extremes, alcohol and certain disinfectants [155-157]. It has been suggested that spores may also facilitate persistence in the host during antibiotic therapy and subsequent recurrence after cessation of treatment [224]. Thus the ability of *C. difficile* to form spores is an essential component of the infectious cycle and therefore an important colonisation factor. However, the processes of sporulation and germination are poorly understood. Therefore, a screening assay to identify insertion mutants with non-wild type phenotypes may lead to important insights into these processes and their regulation.

2.2.3.3. Bile tolerance

During passage through the gastrointestinal tract (GIT) *C. difficile* comes into contact with the potent digestive secretion known as bile. Bile is a complex mixture of bile acids required for the emulsification of fats during digestion. Microorganisms that colonise the GIT must overcome the deleterious effects of bile, which can lead to lysis of the cell membrane. It is not clear how *C. difficile* counteracts the effects of bile and this may represent an additional colonisation factor (see chapter 3 for more details).

2.2.3.4. *p*-cresol tolerance

The phenolic compound *p*-cresol (4-methylphenol) is an end product of protein breakdown. *p*-cresol is one of the metabolites of tyrosine and to a

small extent phenylalanine and under normal physiological conditions it is excreted in the urine. Generation of *p*-cresol in the gut is primarily by aerobic bacteria, particularly enterobacteria, although anaerobes also play a role [225].

C. difficile has been shown to produce *p*-cresol via the degradation of tyrosine [226]. Early investigations found that the bacteria were able to tolerate up to 0.5% *p*-cresol which inhibits other anaerobic bacteria [227]. It has been hypothesised that the production and tolerance of *p*-cresol may offer a competitive advantage to *C. difficile* over other intestinal microflora and consequently contribute to *C. difficile*-associated disease [228]. Sequencing of the *C. difficile* strain 630 genome revealed the genetic pathway to *p*-cresol production; it is produced from tyrosine via the intermediate *p*-hydroxyphenylacetate (*p*HPA) by genes including those encoding components of pHPA decarboxylase (CD0153–CD0155, *hpdBCA*), which catalyses the decarboxylation of pHPA to yield *p*-cresol [6, 229]. Dawson *et al*. (2008) have further explored the ability of *C. difficile* to tolerate *p*-cresol [228]. They found that two ribotype 027 strains the hyper-epidemic strain R20291, isolated from a severe outbreak at Stoke Mandeville Hospital in 2005-2006, and strain BI-16 were significantly more tolerant to *p*-cresol than historical isolates including 630. They suggest that this increased level of tolerance may be a contributing virulence factor to the rapidly spreading epidemic 027 strains [228]. This interesting observation suggests that there may be a link between *p*-cresol tolerance and virulence.

This chapter describes investigation of the behaviour of Tn*1545* and Tn*916* in the hyper-epidemic strain R20291 and the generation of a library of Tn*916* transconjugants in R20291. The library was screened to identify mutants defective in cytotoxicity, sporulation/germination and tolerance of bile salts and *p*-cresol.

2.3. Aims and Objectives

This study aims to generate an insertion library in the hyper-epidemic strain R20291 using a conjugative transposon. The library can be screened for mutations affecting virulence-associated phenotypes to elucidate the molecular pathogenicity of *C. difficile*.

Objectives:

1. Determine whether Tn*916* and Tn*1545* can be introduced into R20291 using a *B. subtilis* donor by filter mating.
2. Investigate and characterise the insertions based on insertion site and frequency per cell.
3. Generate a Tn*916* insertion library in R20291.
4. Identify the insertion sites of Tn*916* in a sample of the library.
5. Develop several assays to screen the library for defects in cytotoxicity, sporulation/germination and tolerance of bile salts and *p*-cresol.

2.4. Materials & methods

2.4.1. **Bacterial strains and growth conditions**

C. difficile was cultured on Brain Heart Infusion (BHI) agar (Oxoid, UK) supplemented with 5% defibrinated horse blood (BHIB) (E & O laboratories, UK) or in pre-reduced BHI broth at 37^{o}C in an anaerobic chamber (Don Whitley Scientific Ltd, UK) with an atmosphere of 80% N_2, 10% H_2, 10% CO_2. Where appropriate, medium was supplemented with *C. difficile* selective supplement containing D-cycloserine 250 mg lt^{-1} and Cefoxitin 8 mg lt^{-1} (CDSS) (Oxoid, UK) and 10 µg ml^{-1} tetracycline (Tc). *C. acetobutylicum* was cultured on pre-reduced Reinforced Clostridial Medium (RCM) with addition of 5% agar for solid medium and grown under the conditions described for *C. difficile*. *B. subtilis* was cultured on BHIB agar or in BHI broth with 10 µg ml^{-1} Tc at 37^{o}C in aerobic conditions. Table 2.2 describes the properties of the bacterial strains.

2.4.2. **Resuscitation of 630Δerm::Tn*916*ΔE library**

Clones stored in Robertsons cooked meat medium (Oxoid, UK) were streaked onto BHIB agar containing CDSS and 10 µg ml^{-1} erythromycin (Em). Plates were incubated anaerobically for 48 hours at 37^{o}C. Cultures were streaked onto selective agar as described above and onto BHIB agar without supplements. The selective plate was incubated anaerobically while the non-selective plate was incubated aerobically at 37^{o}C to detect contaminants.

Bacterial strains, plasmids and transposons		Relevant characteristics	Reference / Source
Strains			
	Clostridium difficile R20291	Em^R, Tc^S ribotype 012	Stoke Mandeville Hospital outbreak isolate, NCTC 13366. Jon Brazier, Anaerobe Reference Laboratory, Cardiff.
	Clostridium difficile 630Δerm	Tc^R, Em^S, ribotype 012	H. Hussain, UCL Eastman Dental Institute. Hussain *et al*., 2005 [131]
	Clostridium difficile 630Δerm::Tn*916*ΔE	Tc^R, Em^R	H. Hussain, UCL Eastman Dental Institute. Hussain *et al*., 2005 [131]
	Clostridium difficile 37	Rif^R, toxin negative	P. Mullany, Eastman Dental Institute, UCL Hächler, *et al*., 1987 [230]
	Clostridium acetobutylicum		ATCC 824. N Minton, University of Nottingham. Weyer and Rettger, 1927 [231]
	Bacillus subtilis BS34A	*B. subtilis*::Tn*916* (CU2189 x FM12A), Tc^R	A. Roberts, Eastman Dental Institute, UCL. Roberts *et al*., 2003 [203]
	Bacillus subtilis SA315	Tn*1545*, Tc^R, Em^R, Kan^R	M. Young, Aberystwyth, UK.
Plasmid			
	pAM120	pGL101::Tn*916*	Gawron-Burke & Clewell, 1984 [232]
Conjugative transposons			
	Tn*916*	Tc^R	Franke *et al*. 1981 [209]
	Tn*1545*	Tc^R, Em^R, Kan^R	Courvalin & Catlier, 1987; Caillaud *et al*. 1987 [213, 214]

Table 2.2 Bacterial strains, plasmids and transposons.

Table Key:

Tc = tetracycline

Em = erythromycin

Kan = kanamycin

Rif = rifampicin

R = resistant

S = sensitive

2.4.3. **Filter mating and transconjugant selection**

BHI broth was inoculated with an 18-24 hour culture (~ 2%) of *C. difficile* R20291 (recipient) and incubated for 16-20 hours in anaerobic conditions at 37^{o}C. *C. difficile* cultures were centrifuged at 3000 x *g* for 15 min in anaerobic conditions to pellet the cells. A single BS43A or SA315 *B. subtilis* colony (donor) was inoculated into BHI broth supplemented with 10 µg ml^{-1} Tc and incubated for 4 hours at 37^{o}C with shaking at 200 rpm. The optical density (OD) of the culture was measured at 600 nm using a spectrophotometer (Thermo Fisher Scientific, UK), cultures with a reading of between 0.6-0.8 were considered to be in mid to late log phase growth and were centrifuged at 5000 x *g* for 5 min. Cells were washed twice with 20 mls BHI broth. The donor and recipient cell pellets were combined in anaerobic conditions and mixed gently but thoroughly with addition of 0.6 ml fresh BHI broth. The mating mixture was spread over cellulose nitrate filters (0.45 µm pore size, Sartorius) placed on pre-reduced BHIB agar. Plates were incubated for 24 hours at 37^{o}C in anaerobic conditions. Filters were transferred to sterile universals and thoroughly washed with 0.6 ml of BHI broth. Cells were spread onto 10-20 BHIB agar plates containing CDSS (BHIBS) and 10 µg ml^{-1} Tc. Plates were incubated for 72 to 96 hours before being examined for transconjugants. Resulting transconjugants were restreaked onto BHIBS agar plus 10 µg ml^{-1} Tc.

Control plates were also included in initial experiments containing either *B. subtilis* or *C. difficile* alone and incubated alongside experimental plates. They were processed in the same manner as experimental plates post

incubation to ensure the selection was appropriate and to test for the frequency at which spontaneous mutants appeared.

2.4.4. **Growth curve**

15 ml of BHI broth in a 20 ml universal was inoculated with R20291 or 630Δerm to an initial OD of 0.05-0.07 measured at 600 nm using a spectrophotometer (Thermo-Fisher, UK). Triplicate cultures were incubated without shaking at 37^{o}C in an anaerobic chamber. At time intervals 0.8 ml samples were taken under anaerobic conditions and the absorbance measured. The assay was repeated 3 times and the data was plotted graphically using Prism 5 (GraphPad Software Inc, v5.01).

2.4.5. **Storage of library**

Each individual transconjugant was inoculated into pre-reduced BHI broth containing 10 μg ml^{-1} Tc and CDSS, and incubated for 72 hours in anaerobic conditions in deep well microtiter plates (VWR International, UK). Sterile glycerol was added to a final concentration of 30% and plates were frozen at -70^{o}C.

2.4.6. **Characterisation of insertions**

2.4.6.1. DNA extraction

DNA from cultured cells was extracted using the Gentra Puregene Yeast Bact. Kit (Qiagen, UK) with minor modifications to the supplied protocol. All

additional reagents were from Sigma-Aldrich UK. Approximately 3-4 ml of 18 -20 hour culture was briefly placed on ice before centrifuging for 45 s at 15,000 x *g* to recover the cells. Cells were resuspended in 300 µl Cell Suspension Solution before addition of 5 µl Lytic Enzyme Solution and mixed by inversion. Samples were incubated at 37°C for 30 min before centrifuging at 15,000 x *g* for 1 min. Supernatant was discarded and the pellet was re-suspended in 300 µl Cell Lysis Solution. Samples were incubated at 80°C for 5 minutes and cooled before addition of 1.5 µl RNase A Solution. Samples were incubated at 37°C for 45-60 minutes. Samples were cooled on ice, 100 µl ice cold Protein Precipitation Solution was added and samples were vortexed vigorously for 20 s. Samples were incubated on ice for 45–60 minutes before centrifuging at 15,000 x *g*. for 3 minutes. Supernatant was added to 300 µl isopropanol and mixed by inversion before centrifuging at 16,000 x *g* for 1 minute to recover the precipitated DNA. The DNA pellet was washed with 70% ethanol and allowed to air dry. DNA was suspended in molecular grade water (molH_2O) (Sigma-Aldrich, UK) by incubation at 65°C for 1 h followed by 37^{o}C for 12-18 hours.

2.4.6.2. PCR

PCR reactions were carried out using *Taq* DNA polymerase and standard buffer according to the manufacturer's guidelines under the following conditions: [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 55^{o}C 1.5 min, 72^{o}C 1 min] x 25-35 cycles [72^{o}C 10 min] x1, hold 4^{o}C. Primers specific for the *xis* and *int* genes of Tn*916* were used to confirm the presence of Tn*916* as described by Roberts *et al*., 2001 [233]:

xis/int1 5’ CGCCAAAGGATCCTGTATATG 3’

xis/int2 5’ GCTGTAGGTTTTATCAGCTTTTGC 3’

PCR products were separated by electrophoresis in 0.7-2% agarose gels made with 1X TAE buffer (Table 2.3) (Sigma-Aldrich, UK). Gels were soaked in TAE buffer containing ethidium bromide for 5 min before visualisation of DNA using UV lamp.

2.4.6.3. Southern blotting

Wild type and transconjugants were grown in BHI broth or BHI supplemented with 10 µg ml^{-1} Tc, respectively, for 20-24 hours at 37^{o}C under anaerobic conditions. Southern blotting was carried out according to the manufacturer’s instructions using an ECL kit (GE Healthcare, UK). Briefly, genomic DNA was extracted from Tn*916* transconjugants and wild type R20291 using the PureGene Yeast/Bact DNA isolation kit. Approximately 300-500 ng of genomic DNA was digested with 200 U *Hin*dIII for Tn*916* transconjugants and R20291 at 37^{o}C for 16 hours (unless stated otherwise, all enzymes are from New England Biolabs, UK). Plasmid DNA containing Tn*916* was extracted from pAM120 using a Mini-prep kit according to the manufacturer’s guidelines (Qiagen, UK) and digested with *Hin*cII at 37^{o}C for 16 hours. Digested DNA was separated on a 0.8 % agarose gel with a size marker at 4 V cm^{-1} for approximately 90 min. The gel was processed according to the manufacturer’s instructions using depurination, denaturation, and neutralization buffers (Table 2.3). The gel was placed on blotting paper atop a ridged scaffold soaked in excess of 20X SSC (Table 2.3). Hybond-N

nitrocellulose membrane (GE Healthcare, UK) was placed on the gel before addition of paper towels, a glass plate and a suitable weight. The DNA was allowed to transfer for 12-15 hours before the membrane was carefully removed from the gel and fixed in a UV cross linker (UV Stratalinker 1800, Stratagene, UK) for 10 seconds.

2.4.6.4. Probe generation

The 0.7 Kb xis/int probe has been designed to hybridise to the xis/int region of Tn*916* [232]. PCR was used to amplify the probe sequence from plasmid pAM120 using previously described primer intxis1 and intxis2 (Figure 2.5) [233]. The conditions were as follows: [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 56^{o}C 1.5 min, 72^{o}C 1 min] x 30 cycles [72^{o}C 10 min] x1, hold 4^{o}C. The PCR product was separated on a 0.8% agarose gel before extraction and purification using a Gel Extraction Kit according to the manufacturer's guidelines (Qiagen, UK).

2.4.6.5. Probe labelling

Each probe was diluted to approximately 10 ng μl^{-1} and the ladder diluted to a final concentration of approximately 2-5 ng ul^{-1} before labelling with the Amersham ECL direct nucleic acid labelling and detection system according to the manufacturer's instructions. Briefly, the probe or marker DNA was denatured by boiling for 5 minutes in a water bath before cooling on ice. An equivalent volume of DNA labelling reagent (for example 10 µl) was added to the DNA and mixed gently but thoroughly. A volume equivalent to the

labelling reagent (10 µl) of glutaraldehyde was added and mixed thoroughly. The mixture was incubated for 10-15 minutes at 37^{o}C and used immediately.

2.4.6.6. Hybridisation and detection

Membranes were soaked in 5X SSC and placed into rotisserie tubes, the SSC was removed and membranes prehybridised in pre-heated hybridisation buffer for 1 hour at 42^{o}C in a rotisserie oven (Table 2.3). Freshly labelled probe and marker were added to the pre-hybridisation buffer and incubated with continual rotation for 16-18 hours at 42^{o}C in a rotisserie oven. Following incubation the membrane was washed with 5X SSC for 5 minutes, the SSC was removed and replaced with pre-heated primary wash buffer and incubated for 20 minutes at 42^{o}C, the wash was repeated with fresh pre-heated primary wash buffer for 10 minutes. The membrane was transferred to a clean container and incubated in secondary wash buffer at room temperature (20-22^{o}C) with gentle agitation for 5 minutes. The wash was repeated with fresh secondary wash buffer for 5-10 minutes.

ECL detection reagents were mixed in equal quantities and applied to the membrane 1 min before removal. The membrane was placed in a protective plastic wallet inside a photographic cassette. In the absence of light, a high performance chemiluminescence film (Amersham, UK) was exposed to the labelled membrane, incubated for between 30 seconds and 12 hours and developed using an AGFA CURIX 260 (AGFA, UK). Membranes were stored in a dust free environment at room temperature for at least 24 hours before being re-probed.

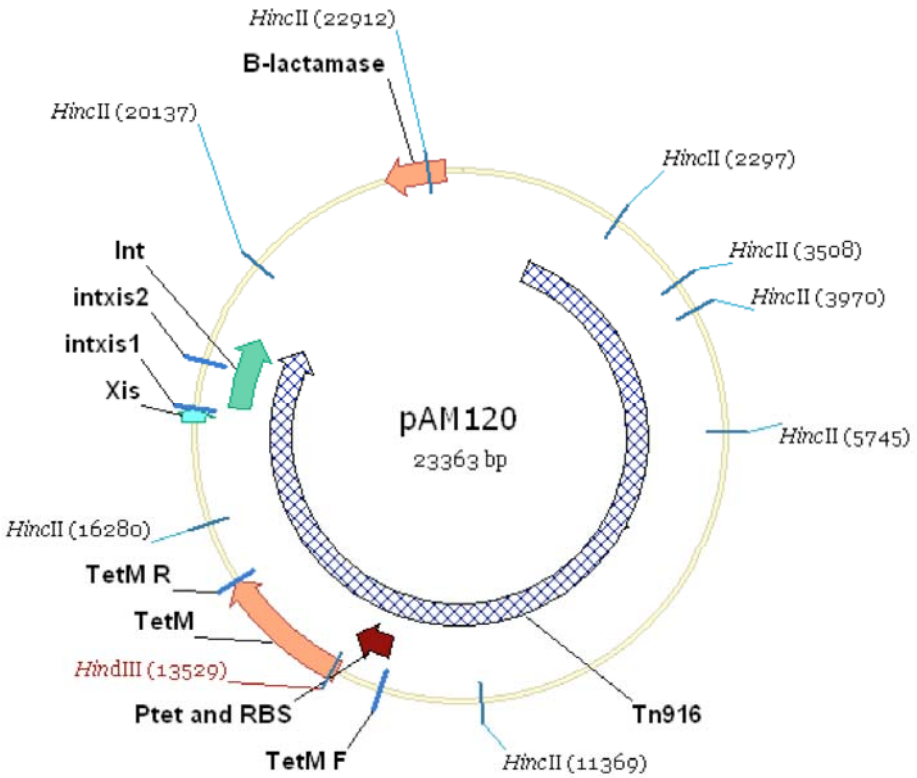


Figure 2.5 Schematic diagram of pAM120

pAM120 contains a complete copy of Tn*916*. The features of this CTn are labelled including P_{tetM} (red arrow), *tet*(M) (orange arrow), *xis* (turquoise arrow) and *int* (green arrow). *Hin*cII and *Hin*dIII restriction sites and the location of intxis1 and intxis2 are also labelled.

Solution	Ingredients	
50x TAE buffer	Tris (base)	2 M
Adjusted to pH8	EDTA (disodium salt)	0.05 M
Depurination solution	HCl	250 mM
Denaturation solution	NaCl	1.5 M
	NaOH	0.5 M
Neutralization solution	NaCl	1.5 M
pH adjusted to 7.5	Tris HCl	0.5 M
20x SSC	Na_3 citrate	0.3 M
pH 7.0	NaCl	3 M
Hybridisation buffer	Hybridisation solution (Amersham, UK)	50 ml
	NaCl	3% (w/v)
	Blocking agent (milk powder)	5% (w/v)
Primary wash buffer	Urea	6 M
	SDS	0.40%
	20X SSC	0.1-0.5X SSC
Secondary wash buffer	20XSSC	10%

Table 2.3 Solutions used for Southern blotting and hybridisation.

2.4.6.7. Sanger sequencing (Pooled samples)

DNA extraction

100 individual transconjugants, picked at random from the frozen R20291::Tn*916* library were plated on BHIB agar with 10 μg ml^{-1} Tc plus CDSS and incubated for 48 hrs at 37^{o}C under anaerobic conditions. Each transconjugant was inoculated into BHI broth with 10 μg ml^{-1} Tc and CDSS in 96 deep-well plates and incubated for 24 hrs at 37^{o}C under anaerobic conditions. A 96-pin replicator was used to transfer a sample of the growth suspension to 2 non-selective BHIB agar plates, with one plate incubated in anaerobic conditions and one plate incubated under aerobic conditions at 37^{o}C for 24 hrs to check for contamination with aerobes. 20 μl of cell suspension from the original growth culture was inoculated into 1 ml of fresh BHI broth with 10 μg ml^{-1} Tc and CDSS in 96 deep well plates and incubated for 24 hrs at 37^{o}C under anaerobic conditions. The transconjugant broth suspensions were pooled and split into manageable volumes before genomic DNA extraction (described in section 2.4.6.1). Purified DNA was stored at -20^{o}C and given to Dr Tim Perkins at the Wellcome Trust Sanger Institute, UK for sequencing.

The following methods (1-4) are those of Dr Tim Perkins adapted from Langridge *et al*. (2009) and all sequencing was performed by Dr Tim Perkins and Dr Gemma Langridge at the Wellcome Trust Sanger Institute [234]. They describe the transposon directed insertion-site sequencing (TraDIS) method. Data were then returned to the Eastman Dental Institute, UCL for further analysis.

1. Nucleotide sequencing (Sequencing-by-synthesis)

Five micrograms of genomic DNA was fragmented to an average size of 300 bp by using a Covaris AFA (adaptive focused acoustics) Ultra Sonicator and Illumina library preparation was performed following the manufacturer's instructions, but using 1.5x the recommended reagent volumes in each step (Covatis Inc; Illumina Inc, USA) [235]. Breifly, after shearing the DNA ends were repaired and A's were added to the ends to facilitate ligation of the Illumina adapters.

To amplify the transposon insertion site, 22 cycles of PCR were performed using a transposon-specific forward primer and a custom Illumina reverse primers (adapter-specific) and 100 ng of library DNA per reaction.

Forward primer

5' AATGATACGGCGACCACCGAGATCTACACATAAGTCCAGTTTTTATGCGGATAAC 3',

Reverse primer

5' CAAGCAGAAGACGGCATACGAGATCGGTACACTCTTTCCCTACACGACGCTCTT CCGATCT 3'.

Amplified libraries were cleaned up with a QiaQuick PCR product purification column following the manufacturer's instructions and eluted in 30 µl elution buffer, and the libraries were quantified by qPCR [235].

Ligated fragments (DNA frgments ligated to adapters) were separated on a 12 cm 2 % agarose gel in 1 x TBE buffer (45 mM Tris-borate and 1 mM EDTA), at 6 V cm^{-1} without the preceding column clean up step. After 45

minutes, fragments corresponding to an insert size of 250-350 bp were excised, and DNA was extracted from the gel slice without heating [235]. The DNA was quantified on an Agilent DNA1000 chip, using a bioanalyzer, following the manufacturer's instructions (Agilent technologies, Germany). The libraries were sequenced on a single end Illumina flowcell using an Illumina GAII, 54 cycles of sequencing, using a custom sequencing primer 5' TCTACACATAAGTCCAGTTTTTATGCGGATAACTAGAT 3' and 2x Hybridization Buffer (0.2% Tween 20 in 10x SSC). This primer was designed such that the first 14 bp of each read was Tn*916* transposon sequence (TTTTATGCTATTTT).

2. PCR

Fragments were amplified using Jumpstart *Taq* with 100 ng DNA template using the following conditions 94°C 2 minutes; [94°C 30 seconds; 65°C 20 seconds; 72°C 30 seconds] x 22 cycles; 72°C 10 minutes.

3. Analysis of nucleotide sequence data

Custom Perl scripts were written to identify sequence reads with a 100% identical match to the 14 bp transposon tag. All qualifying reads had this tag removed, along with a further 10 bp to take into account any possible coupling sequence. Maq-0.6.8 using the easyrun.pl option with all the default parameters was used to map the remaining 30 bp to the relevant *C. difficile* genome sequence; either strain 630 (Accession: AM180355) or R20291 [6, 121, 236]. Mapped reads were filtered for mapping quality and those with a score of >20 (630) or 10 (R20291) were taken forward. Each read mapped to a particular nucleotide position in the genome, named a Tn*916* insertion site.

All insertion sites were inspected for frequency of reads and sites with <20 reads were excluded from the strain 630 list as potentially spurious. Read depth of the R20291 library was much lower, hence sites with <10 reads were excluded as potentially spurious. Due to the potential presence of a coupling sequence, these insertion sites are up to 10 bp away from the true insertion site.

4. Insertion site motif in strain 630

Using the reference genome sequence, twenty base pairs upstream and downstream of each mapped Tn*916* insertion site in the 630Δerm pool was retrieved. All the 40bp regions were submitted to MEME specifying a distribution of 1 motif per sequence [237, 238]. The output from MEME was loaded into FIMO (also part of the MEME suite) to identify all occurrences of the predicted motif using *C. difficile* strain 630 as the reference genome sequence.

Analysis of insertion sites

Genes were assigned functional classes based on the *C. difficile* 630 genome annotation and using the JGI Integrated Microbial Genomics database [6, 239].

2.4.7. **Screening assays**

2.4.7.1. Tissue culture maintenance

HFF-1 fibroblasts (ATCC SCRC-1041) were grown in Dulbecco's Modified Eagle's Medium (D-MEM low glucose, Gibco-Invitrogen, UK) supplemented with 10% foetal calf serum (FCS) (Hyclone, USA), 0.1% glutamine (Cambrex, Belgium), penicillin/streptomycin mix (100 U ml^{-1}) (Lonza, Belgium) in T75 vented flasks (Sarstedt, Germany) and incubated at 37^{o}C, 5% CO_2. Cells were trypsinized (Cambrex, Belgium) and subcultured two to three times per week depending on confluence. Frozen stocks were maintained in DMEM with 45% FCS and 5% Dimethyl sulfoxide (DMSO, Sigma-Aldrich, UK) in liquid nitrogen.

2.4.7.2. Cytotoxicity assay

Sample preparation

Transconjugants and control cultures (wild type R20291, 630Δerm and CD37) were inoculated into 1 ml BHI broth containing 10 µg ml^{-1} Tc and CDSS, in deep well microtiter plates to an initial OD of 0.05 at 590 nm. Cultures were incubated at 37^{o}C without shaking for 48-60 hours under anaerobic conditions. Cultures were centrifuged at 2500 x *g* for 10 minutes to pellet the cells. A sample of cell free supernatant was diluted 50-fold in sterile PBS (Lonza, Belgium) under sterile conditions.

Tissue culture assay

HFF-1 fibroblasts were seeded at a density of 1 x 10^5 cells per well in DMEM medium in flat-bottomed 96-well microtiter plates. Plates were incubated for 16 hours at 37°C, in 5% (vol/vol) CO_2 to achieve confluence. Cell free supernatant from bacterial cultures was diluted in PBS and added to cell monolayers. BHI broth and PBS were included as negative controls. Plates were incubated at 37°C in 5% CO_2 for 24 hours. Cells were assessed under an inverted microscope for a cytopathic effect (CPE) i.e. rounding of at least 50% and loss of adherence.

2.4.7.3. Sporulation/germination assay

Bacterial cultures were grown for 48 hours in 1 ml BHI broth supplemented with 10 µg ml^{-1} Tc in deep well microtiter plates (without shaking) at 37°C under anaerobic conditions. 100 µl samples were transferred to round bottom microtiter plates and cultures were heated to 60°C in a water bath for 25 min to kill vegetative cells. Using a 96-pin replicator approximately 2 µl of heat-treated culture was plated onto BHIB agar. Heated and unheated samples of wild type R20291 were included as positive controls. Plates were incubated at 37°C in anaerobic conditions for 48-60 hours.

2.4.7.4. Bile tolerance assay

Bacterial cultures were grown for 24 hours in 1 ml BHI broth supplemented with 10 µg ml^{-1} Tc in deep well microtiter plates (without shaking) at 37°C under anaerobic conditions. Using a 96-pin replicator approximately 2 µl of culture was spotted onto BHIB containing 5% bovine and ovine bile (B & O

bile) (Sigma-Aldrich, UK). Wild type R20291 and *C. acetobutylicum* were spotted onto each plate as positive and negative controls, respectively. Plates were incubated at 37^{o}C in anaerobic conditions for up to 72 hours before analysis.

2.4.7.5. *p*-cresol tolerance assay

Bacterial cultures were grown for 24 hours in 1 ml BHI broth supplemented with 10 µg ml^{-1} Tc in deep well microtiter plates (without shaking) at 37^{o}C under anaerobic conditions. Using a 96-pin replicator approximately 2 µl of culture was spotted onto BHIB agar containing 0.1% *p*-cresol (4-methylphenol) (Sigma-Aldrich, UK). Wild type R20291 and *C. acetobutylicum* were spotted on to each plate as positive and negative controls, respectively. Plates were incubated at 37^{o}C in anaerobic conditions for up to 72 hours before analysis.

2.4.7.6. Dot blot assay

Bacterial cultures were grown for 48 hours in 10 ml BHI broth in 20 ml universals (without shaking) at 37^{o}C under anaerobic conditions. Cultures were centrifuged at 3000 x *g* for 5 min to pellet the bacteria. 10-30 µl of cell free supernatant was spotted onto Hybond C nitrocellulose and left to dry for 10 min at 20-22^{o}C (GE Healthcare, UK). Membranes were blocked with blocking buffer (5% skimmed milk powder (Marvel, UK) in PBS) for 2 hours at 20-22^{o}C. The membrane was washed for 5 min with wash buffer (PBS with 0.2% Tween; both Sigma-Aldrich, UK) and washing was repeated 3 times.

The membrane was incubated with primary antibody: mouse IgG anti-*C. difficile* TcdB (Patricell Ltd, Nottingham UK) diluted 1:250 – 1:50000 for 2 hours at 20-22^{o}C on an orbital shaker. The membrane was washed as described above and incubated with secondary antibody: rabbit anti-mouse IgG peroxidase conjugate (Sigma-Aldrich, UK) diluted to 1:500-1:50000 for 2 hours on an orbital shaker. The membrane was washed as described above before addition of the detection reagent Sigma Fast DAB peroxidase substrate (3,3'-diaminobenzidine tetrahydrochloride with metal enhancer, Sigma-Aldrich, UK). The reaction was stopped by flooding the membrane with dH_2O or PBS.

2.5. Results

2.5.1. **Library construction**

2.5.1.1. Tn*916*

Filter mating is a standard method to facilitate the transfer of conjugative elements between bacteria. This method has proved successful for the transfer of Tn*916* from *B. subtilis* to *C. difficile* strains 630Δerm and CD37 [131, 193]. *B. subtilis* containing Tn*916* (donor) was mated with the hyper-epidemic *C. difficile* strain R20291 (recipient) to investigate its behaviour and potential as a random mutagen in this strain. Tn*916* entered the genome of 630Δerm on multiple occasions with preference for AT rich sites; however it entered the genome of CD37 at a specific site (att916) [131, 194]. Mating of the donor with R20291 led to the generation of *C. difficile* transconjugants containing Tn*916*. Transconjugants arose at an average frequency of 10^{-10} per donor cell, which in comparison with the frequencies observed for strains 630Δerm and CD37 (10^{-7}), is extremely low [195]. Generation of the library was therefore very inefficient; in total 1000 transconjugants were generated for the library from circa 100 independent mating experiments. Attempts to optimise the mating procedure by sequential variation of experimental parameters, including time, growth phase and addition of cysteine failed to improve the conjugation frequency. One parameter was changed for each mating and repeated three times. Different ratios of donor to recipient cells were tested including ratios of approximately 1:1, 1:2, 1:3, 1:5, 2:1, 3:1, and 5:1. This had no obvious effect on the efficiency of the mating and the frequency with

which transconjugants were produced remained at approximately 10^{-10} per donor. The mating time was investigated with inoculated filters incubated for 15, 20, 24 or 40 hours. Again, no obvious increase in transconjugant generation was observed, the frequency of transconjugants production remained 10^{-10} per donor. Cells were also tested at three growth phases – mid-log (~12 hours), late-log (~17 hours) and stationary phase (~24 hours) as determined from a growth curve (Figure 2.6). The growth stage did not appear to have any obvious impact on the mating efficiency. The addition of cysteine to the growth medium resulted in the production of no detectable transconjugants in four independent experiments with an average of 10^{10} *C. difficile* recipients.

Each transconjugant was authenticated by its ability to grow in the presence of 10 μg ml^{-1} Tc, confirming the presence of *tet(*M) on Tn*916*. Six randomly selected transconjugants were subjected to PCR analysis that confirmed the presence of the *xis* and *int* genes of Tn*916* in all the transconjugants. Hybridisation analysis also confirmed the presence of Tn*916* in 30 transconjugants picked from the library (discussed in detail in section 2.5.3).

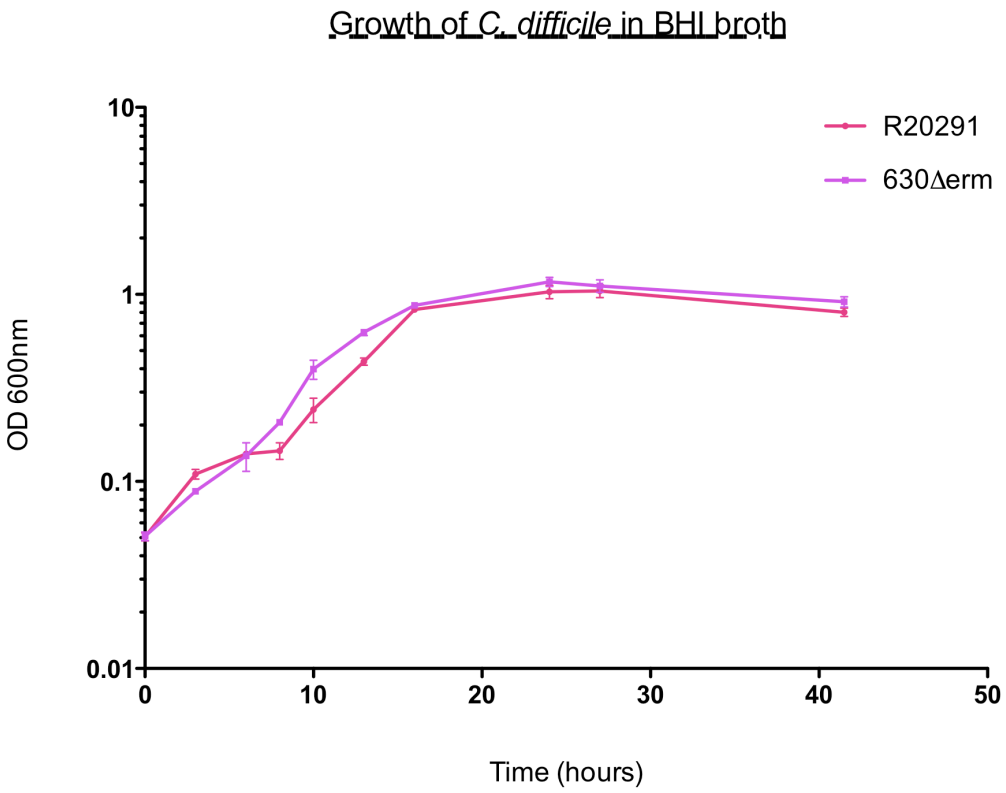


Figure 2.6 Growth of *C. difficile* strains R20291 and 630Δerm in BHI broth over time.

The growth of R20291 and 630Δerm was monitored over time by measurement of optical density (OD) at 600 nm. The results represent the mean of three experiments and error bars show the standard deviation.

2.5.1.2. Tn*1545*

The presence of Tn*1545* in *B. subtilis* strain SA315 was confirmed by PCR using primers specific for the Xis and Int region of the transposon. In total six matings, each with 5-10 filters, were performed between R20291 (recipient) and SA315 (donor). Mating of R20291 with SA315 failed to lead to the generation of tetracycline resistant transconjugants from the mating of approximately 4.2 x 10^9-7.7 x 10^{10} SA315 donors with 3.2 x 10^9-8.3 x 10^9 R20291 recipients in six independent experiments. Due to these results Tn*1545* was not selected for the generation of a library in R20291.

2.5.2. **Screening assays**

Before storage as frozen aliquots each R20291::Tn*916* transconjugant was subjected to four screening assays. These included three simple agar-based assays assessing: 1) sporulation/germination potential, 2) bile tolerance, and 3) *p*-cresol tolerance. A fourth assay, the cytotoxicity assay, was carried out to identify transconjugants with insertions that affect the production of TcdA and TcdB. Each member (approximately 1000) of the library was screened in each of the four assays. No mutants with defects in any of the screened phenotypes were identified, while the use of negative controls such as *C. acetobutylicum* and wild type strains R20291 and CD37 provided validation for the assays.

As an alternative to the cytotoxicity assay an antibody based dot-blot assay was investigated, this may have represented a relatively quick and less

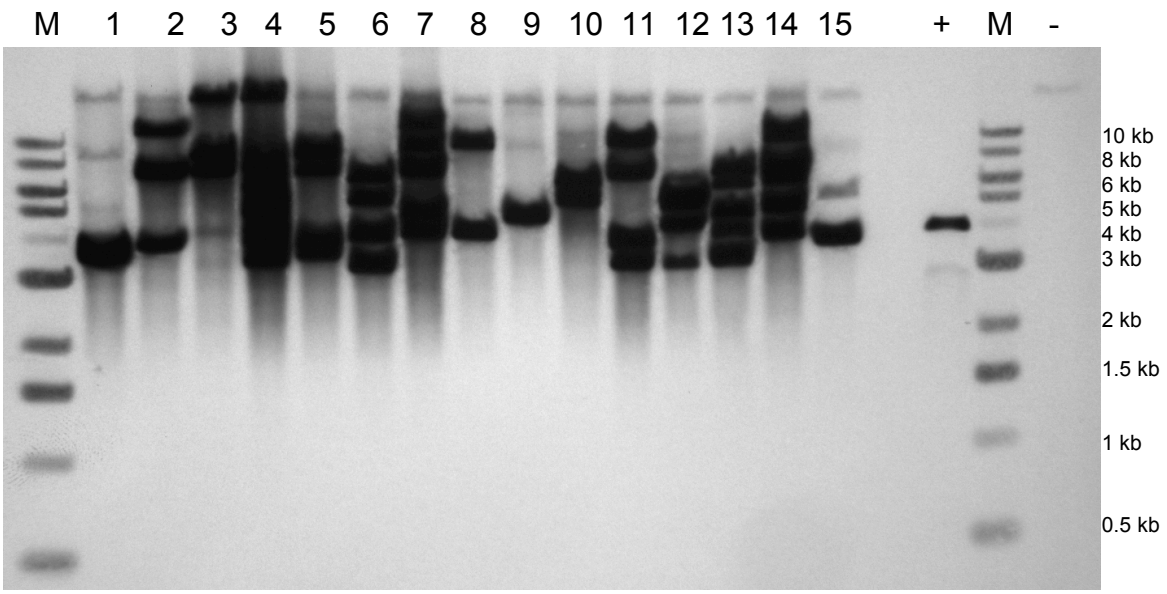
labour-intensive method for identification of toxin in growth media. Unfortunately, this assay failed to distinguish samples known to contain toxin and those that did not using a specific antibody for TcdB (data not shown).

2.5.3. **Characterisation of insertions**

2.5.3.1. Hybridisation analysis

To determine if Tn*916* entered the R20291 genome at multiple sites, thirty transconjugants were subjected to hybridisation analysis. Digested DNA was probed with a xis/int fragment of Tn*916*. Since there is one *Hin*dIII site in the transposon the probe will bind to only one of these fragments (Figure 2.5). Of the thirty transconjugants, fifteen represent colonies arising from a single filter (Figure 2.7 A). The hybridisation pattern obtained from these transconjugants reveals that each individual is unique indicating that they are not siblings. The remaining fifteen were derived from different filters and again show unique patterns of hybridisation. The majority show multiple insertions, with only two possessing a single insertion (Figure 2.7 B lane 5 & 11).

A



B

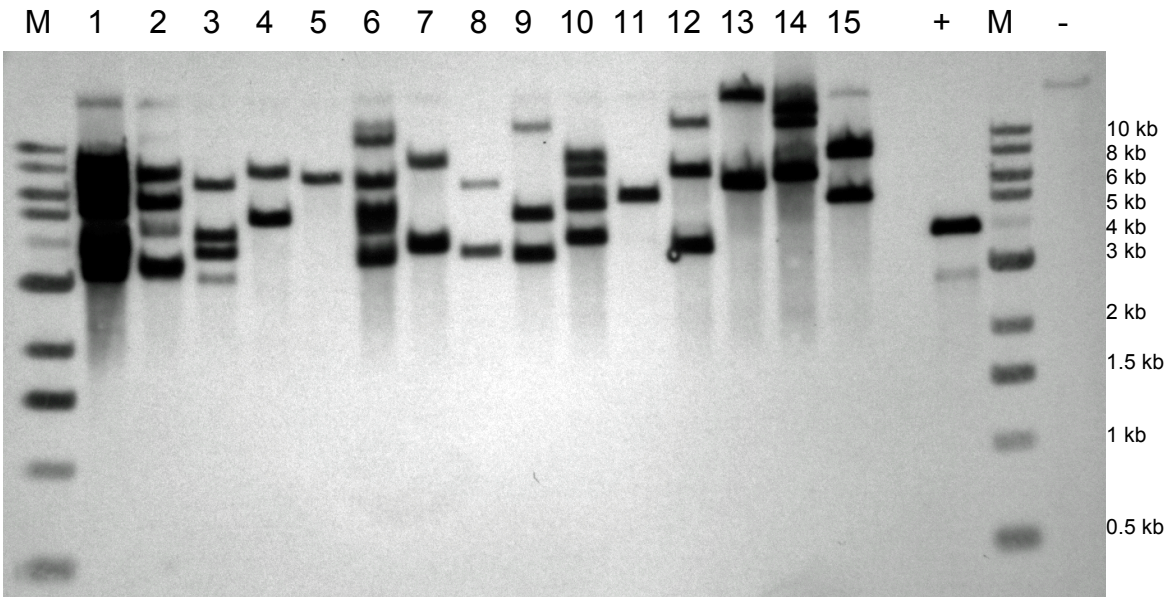


Figure 2.7 Hybridisation analysis of 30 Tn*916* transconjugants.

DNA from 30 transconjugants was digested with *Hin*dIII and probed with a xis/int probe to identify Tn*916* insertions. Blot A shows 15 transconjugants arising from the same filter with each one showing a unique hybridisation pattern. Blot B shows 15 transconjugants taken from different filters, again each one gives a unique hybridisation pattern. All of the transconjugants except two (blot B lane 5 & 11) show multiple insertions. M = 1kb marker, + = pAM120 *Hin*cII digested DNA, - = wild type R20291 *Hin*dIII digested DNA.

2.5.3.2. Sequencing of the Tn*916* insertion site

To identify the insertion site of Tn*916* in R20291, DNA from 100 randomly picked transconjugants (10% of the library) was extracted, pooled and sent to The Wellcome Trust Sanger Institute for sequencing. As the sequence of Tn*916* is known, insertion sites can be easily identified. Analysis of the sequence surrounding the insertion can identify the location of the insertion in the genome. From hybridisation analysis it was known that each transconjugant is likely to contain multiple insertion sites. Sequencing analysis revealed 102 unique insertion sites in R20291, with twenty-seven in CDSs, equating to approximately 25% of the total number of insertion sites (Table 2.4). Therefore, approximately 75% of the insertions occurred in non-coding intergenic regions of the genome even though only 18% of the genome is believed to be non-coding (Dr Gemma Langridge, personal communication). Closer inspection of the listed insertion sites reveals that Tn*916* appears to have inserted into the same gene on multiple occasions. In total five genes have more than one insertion; in one of these genes (CDR20291_2368) the two insertion sites have a 1 bp difference. In the other three other genes (CDR20291_2033, CDR20291_2642 & CDR20291_2855) the difference between the two insertion sites is 14 bp. Finally, one gene (CDR20291_1737) appears to have three Tn*916* insertions, the difference in insertion site ranges from 1 to 15 bp.

Insert position	Gene	630 orthologue	Start position	End position	Gene Length (bp)	Strand	Gene Product
413027	CDR20291_0344	CD0339	412623	413360	737	+	two-component response regulator
487710	CDR20291_0404	CD0463	487466	487999	533	-	TetR-family transcriptional regulator
612747	CDR20291_0506	CD0579	612337	612915	578	-	TetR-family transcriptional regulator
620016	CDR20291_0512	CD0588	619824	620111	287	+	hypothetical protein
777896	CDR20291_0628	CD0702	777396	778079	683	+	putative membrane protein
944576	CDR20291_0773	CD0843	943977	945212	1235	+	putative glycosyl transferase
1080945	CDR20291_0886	CD1030	1080844	1082100	1256	+	putative glycosyl transferase
1433968	CDR20291_1206	CD1363	1432968	1434032	1064	+	putative phage protein
1645610	CDR20291_1389	CD1540	1643765	1645645	1880	+	ABC transporter, permease protein
1919051	CDR20291_1634	CD1737	1918970	1920235	1265	+	putative gluconate permease
1934900	CDR20291_1645	CD1751	1932770	1935157	2387	+	cell surface protein (putative cell surface-associated cysteine protease)
2037397	CDR20291_1737	CD1842	2037161	2037907	746	+	putative membrane protein
2037411	CDR20291_1737	CD1842	2037161	2037907	746	+	putative membrane protein

Insert position	Gene	630 orthologue	Start position	End position	Gene Length (bp)	Strand	Gene Product
2037412	CDR20291_1737	CD1842	2037161	2037907	746	+	putative membrane protein
2339670	CDR20291_2001	CD2094	2339541	2341475	1934	-	putative restriction enzyme
2382637	CDR20291_2033	CD2126	2382398	2382655	257	-	putative membrane protein
2382651	CDR20291_2033	CD2126	2382398	2382655	257	-	putative membrane protein
2771369	CDR20291_2368	CD2475	2771181	2772483	1302	-	putative competence membrane protein
2771370	CDR20291_2368	CD2475	2771181	2772483	1302	-	putative competence membrane protein
2850388	CDR20291_2431	CD2544	2849896	2851011	1115	-	putative membrane protein
2990751	CDR20291_2551	CD2663	2989423	2991702	2279	-	putative signalling protein
3110019	CDR20291_2642	CD2752	3109477	3110142	665	+	hypothetical protein
3110033	CDR20291_2642	CD2752	3109477	3110142	665	+	hypothetical protein
3379082	CDR20291_2855	CD3019	3378774	3379706	932	-	putative transporter
3379096	CDR20291_2855	CD3019	3378774	3379706	932	-	putative transporter
3677207	CDR20291_3079	CD3220	3677190	3678047	857	-	putative methyltransferase
3802783	CDR20291_3188	-	3802387	3803703	1316	-	sensor histidine kinase VirS

Table 2.4 Location of Tn*916* insertions in the genome of R20291

The insertion sites of Tn*916* in 100 transconjugants was investigated by the Sanger Institute. 27 insertion sites were found to be in the coding regions of R20291. Details of the location of the insertions are listed in the table alongside the putative products of each gene and the orthologue in strain 630.

From the 102 characterised insertion sites in R20291, there was evidence of some bias of insertion into intergenic regions. This was also seen in the results obtained for 630Δerm::Tn*916*ΔE library where only 19 out of 112 unique insertion sites were found in genes. To investigate further Dr. Gemma Langridge at the Wellcome Trust Sanger Institute used MEME (Multiple Em for Motif Elicitation), an on-line motif-based sequence analysis tool, to investigate the 40 bp sequence surrounding the insertion sites in the 630Δerm::Tn*916*ΔE sample pool to identify a common motif [237, 238]. She found that a 15 bp motif, TTTTTATATTAAAAA, was associated with all of the insertion sites (Dr Gemma Langridge, personal communication). She used the MEME search results to find all occurrences of the 15 bp motif in the 630 reference genome sequence using FIMO (Find Individual Motif Occurrences). Using a *P*-value cut-off of 0.01 (which finds all the insertion sites indentified by the TraDIS method), there were a total of 100,987 occurrences of the 15 bp Tn*916* insertion motif in the genome sequence of strain 630. Approximately 63,000 of these are genic and approximately 35,000 are intergenic (the remainder cross gene boundaries). This shows a bias towards intergenic regions, as approximately 18,000 occurrences would be expected for randomly distributed motifs. This suggests that Tn*916* preferentially inserts into a consensus sequence, TTTTTATATTAAAAA, a sequence found predominantly within the intergenic regions of strain 630.

2.6. Discussion

The conjugative transposon Tn*916* has been used to generate an insertion library in the epidemic strain R20291. The frequency with which transconjugants arose was extremely low, meaning that the construction of the library was very time consuming. In many mating experiments no transconjugants were generated from any of the filters. In other experiments 1-15 were produced using the same experimental parameters. Attempts were made to optimise the filter mating procedure to promote transconjugant formation. Almost all of the variables were assessed individually including the ratio of donor to recipient cells, the growth phase of the bacteria, the mating time, and the effect of addition of cysteine to the growth medium. Some parameters remained constant, for example the media, filters and cell strains, additionally *C. difficile* was not exposed to oxygen during mating at any point. None of the changes improved the efficiency of the matings although these experiments were only repeated three times. Given the low frequency of conjugation, further replicate experiments may have produced transconjugants, allowing comparison of the efficiencies. The number of variables and the time required to carry out these different experiments precluded such investigation.

The role of cysteine was however investigated in a further three experiments. In all six of the experiments performed transconjugants failed to be generated indicating that cysteine in the growth medium has a negative impact on

conjugation. Despite attempts to increase the frequency of conjugation, the final size of the R20291::Tn*916* library was 1000 members.

Each of the 1000 transconjugants was assessed by each of the four screening assays before storage. Three of the assays were agar-based and therefore a negative result was indicated by growth. All of the samples tested were able to tolerate the same concentrations of *p*-cresol and bile as the wild type strain suggesting that insertion of the transposon had no effect on the tolerance mechanism of *C. difficile* to these two stresses. No differences were seen in the ability of the transconjugants to sporulate or germinate; this relatively simple assay revealed that all of the samples were able to produce spores and to germinate efficiently. Should colonies not have been observed the sample would have been investigated further to characterise the location of the insertion and which process was affected.

The cytotoxicity assay is relatively time consuming and expensive, and is not specific, as both TcdA and TcdB can cause a CPE. In general this assay will allow identification of strains able to produce toxins to a similar level as the wild type strain. It would only be possible to see differences in the level of toxin produced if these differences were fairly large due to the fact that only one dilution of growth medium was examined. All of the transconjugants tested were capable of causing a CPE at a 50-fold dilution. Should time and resources permit, further dilutions of the growth medium and assessment at various time points may have revealed differences in the amount of the toxins produced or in their regulation. Toxin production is believed to occur in

response to starvation [133]. By growing the bacteria in defined media it is possible to control the nutrients available allowing further monitoring of toxin production and potentially the identification of novel factors involved in production and regulation.

In general none of these assays would be able to reveal subtle differences between the transconjugants and wild type samples, but as transposon libraries often contain thousands of members these high throughput and relatively cheap assays are invaluable for the identification of mutations in key genes involved in these processes.

Other screening assays have also been devised to investigate this and the 630Δerm::Tn*916*ΔE library, and are to be conducted by a member of the Mullany group. These included a bacitracin tolerance assay, this antimicrobial attacks the Gram-positive cell wall and has been investigated as an alternative to vancomycin in the treatment of CDI; an agar based motility assay and potentially, an adhesion assay [240, 241]. The library represents an important resource which can be screened numerous times with various assays to identify novel genes involved in a range of processes, particularly those associated with virulence.

Insertion of Tn*916* in the hyper-epidemic strain R20291 has been shown to be analogous to insertion into strain 630Δerm. Multiple insertions were seen in twenty eight (93%) of the transconjugants investigated by Southern hybridisation, again similar to the results observed for 630Δerm with no

obvious 'hot-spot' [131]. On average, there were approximately three insertions of Tn*916* per R20291 transconjugant based on hybridisation analysis of 30 transconjugants.

To further characterise the insertion sites of Tn*916*, DNA from a random sample of 100 R20291::Tn*916* transconjugants was pooled and sent to the Sanger Institute for sequencing. Sequencing revealed that the majority of the insertion sites were within intergenic regions, with 27 insertions in CDSs. A similar result was observed for the 630Δerm::Tn*916*ΔE library where 100 random transconjugants were also sequenced to determine the site of insertion. From 112 unique insertion sites nineteen genes had a Tn*916*ΔE insertion, again the majority of insertions occurred in intergenic regions. Approximately 82% of the *C. difficile* genome is coding (personal communication; Dr. Gemma Langridge, The Wellcome Trust Sanger Institute). If insertion of Tn*916* occurs at random then one would expect an equivalent proportion of insertions in CDSs from this random sample of transconjugants. The insertion bias was confirmed by the identification of a 15 bp motif found at all of the Tn*916* insertion sites which is found predominantly within intergenic regions of the genome of strain 630. Thus, the usefulness of Tn*916* as a random mutagen in *C. difficile* may be limited. As the insertion site motif is found at low frequency in coding regions only a small percentage of library insertion mutants will have disrupted CDSs. However, disruption of intergenic regions may provide vital clues about the regulation of genes and genome organisation.

Comparative analysis of several *C. difficile* strains (including R20291) by Stabler and colleagues revealed that the insertion of other transposons in strain 630 occurred mainly within CDSs [121]. Eight full-length functional transposons copies had inserted in CDSs. Conversely, in both of the ribotype 027 strains examined, of the seventeen full-length transposon copies identified, only six had inserted into CDSs. Clearly, each transposon has a different mechanism of integration and insertion site preference, which may vary between strains and genera. Insertions in CDS and intergenic regions are likely to affect gene expression by direct inactivation or indirectly through downstream polar effects. Although it is likely that insertions in coding sequences are less well tolerated, particularly those in essential genes, intergenic regions are important for gene regulation, disruption of these regions could therefore lead to effects on downstream genes [242].

It may be possible to use Tn*916* as a saturating mutagen to investigate the remaining 737,320 bp (18.21% of 4.049 Mb) of sequence whose function is, at present, not well understood. Due to the preference of insertion into the 15 bp motif this suggests that a maximum of 100,000 insertions are possible with Tn*916*. Although limited, this may provide key insights into these poorly described regions, possibly elucidating the process of gene regulation and characterisation of genome architecture.

Closer inspection of the Tn*916* insertion sites in R20291 revealed that the transposon may have entered four genes on two occasions, and one on three occasions at seemingly unique sites with differences in the insertion

position of between 1 and 15 bp. Limitations of the sequencing assay mean that the insertion site can only be determined to within approximately 10 bp. Therefore, it is possible that each of these multiple insertions actually represents one insertion, but due to slight errors in the sequencing they are annotated as separate insertions. Thus, if this is indeed the case, the number of insertion sites in coding sequence is reduced to just twenty one. Screening of the library with the four screening assays failed to identify any putative mutants through phenotypic assays. Thus it would appear that the twenty one genes with Tn*916* insertions are not associated with toxin production, bile or *p*-cresol tolerance, or sporulation/germination.

The functions of these genes could be further investigated through comparative analysis with other bacterial species and genera to develop assays to identify whether these insertions affect a measurable phenotype. For example three genes are annotated as being putative membrane proteins, therefore assays designed to assess the functionality of the membrane and its integrity may reveal the functions of these genes. Assays may include the investigation of pH, osmolarity, heat/cold treatment and antibiotic susceptibility to assess whether phenotypic changes can be observed for these transconjugants. However, the annotation of R20291's genome is not definitive, the majority of gene functions have been inferred from other bacteria using *in silico* analysis, with putative protein function often assigned without experimental evidence [6, 121]. Consequently, the genes in this list may not be accurate, thus the benefits of developing further assays would need to be assessed. Interestingly, one of the genes disrupted by

Tn*916* was CDR20291_3188, which is not present in strain 630. This gene is annotated as sensor histidine kinase VirS which is one member of a two component signal transduction system shown to control toxin production in *C. perfringens* [243]. No toxin deficient mutants were found during the cytotoxicity screening assay; therefore it is unclear whether VirS is involved in toxin production in *C. difficile*. The screening assays investigated only one dilution of growth supernatant, therefore it is possible that the insertion in this gene could have affected the amount of toxin produced. Since the cytotoxicity assay is not specific for either toxin A or B, detailed *in vitro* assays with a full dilution series, and the use of specific antitoxins could elucidate the effects of this insertion on toxin production.

Finally, the ability of Tn*1545* to act as a random mutagen in R20291 was assessed. In these experiments, transconjugants containing this transposon and the associated tetracycline resistance were not observed. In total six independent assays were performed in parallel with Tn*916* mating. As no transconjugants were seen in these experiments work with this transposon was not continued. It is possible that there may be activity towards the transposon from host restriction systems or poor maintenance in this strain. This requires further investigation.

2.7. Conclusions & future work

Conjugation of Tn*916* into the hyper epidemic strain R20291 was successfully achieved, although the frequency of transfer was extremely low. This low conjugation efficiency for R20291 had a severe impact on the generation of an insertion library in this strain. 1000 transconjugants were screened successfully with four different assays, although no mutant phenotypes were observed. This library has been stored to allow further assessment using a range of screening assays. Further assays could include motility and adhesion assays, auxotroph screens and investigation of resistance to antimicrobial agents.

Hybridisation analysis of members of this library indicated that there was no insertion hot spot. Investigation of the insertion site of Tn*916* in approximately 100 members of the library revealed that insertion occurred predominantly in intergenic regions. Investigation of the insertion site of Tn*916*ΔE in 100 members of a 630Δerm::Tn*916*ΔE library revealed a 15 bp consensus sequence at the transposon insertion site. A search of the 630 genome revealed that this motif was found predominantly in intergenic regions therefore Tn*916* is more likely to insert into these regions. Thus this insertion bias prevents the use of Tn*916* or Tn*916*ΔE as random mutagens in both of these strains. However, production of a saturating library in R20291 could allow investigation of these poorly characterised regions and provide important information about gene regulation.

Chapter 3

3. Investigation of the effects of bile on *C. difficile* growth and gene expression

3.1. Abstract

C. difficile is GI pathogen which under certain conditions colonises the colon leading CDI. During passage through the GI tract the bacterium in either vegetative or spore form encounters innate antimicrobial defence mechanisms evolved to prevent colonisation by microorganisms. One of these natural defences is the production of a potent digestive secretion known as bile. The major role of bile is to emulsify fats from ingested material to aid digestion. However, it is also an effective antimicrobial. The tolerance of *C. difficile* and its interaction with bile has not yet been fully investigated. The results obtained here describe the tolerance levels of two pathogenic strains, 630Δerm and R20291, to bovine and ovine bile. Significant differences were seen in the tolerance of R20291 compared to 630Δerm, with R20291 showing tolerance of a much higher concentration of bile. Analysis of genome-wide transcription levels using a microarray revealed differences in the expression of genes in 630Δerm in response to bile exposure, although these were not statistically significant. These preliminary investigations do however provide a foundation for further elucidation of the interaction of *C. difficile* with bile.

3.2. Introduction

Microorganisms which colonise the GIT are subjected to a variety of stressful conditions and environmental extremes upon entering the body. Natural antimicrobial defence mechanisms exist to control colonisation through changes in pH, limitation of nutrients, reduced oxygen levels and increased osmolarity. Organisms that colonise these environments must have evolved mechanisms for surviving these conditions, which can be considered as colonisation factors.

One host defence mechanism is the production and release of bile from the liver, this highly potent digestive secretion plays a major role in the emulsification and solubilisation of lipids [244]. Thus, evolution of mechanisms to evade the actions of this digestive secretion are essential to allow survival and colonisation of the GIT.

3.2.1. **Physiological role of bile and bile salts**

The main role of bile is to act as a biological detergent to emulsify and solubilise ingested fats to aid digestion by lipases. This property also confers potent antimicrobial activity forming part of the body's physiochemical defence system [245]. Bile facilitates the excretion of insoluble or protein bound substances, particularly cholesterol, which cannot be excreted in the urine. Bile acids can also act as regulatory and signalling molecules, including the regulation of hepatic lipid and glucose metabolism [246]. Bile has numerous

important physiological and pathological roles, many of which are not fully understood.

Bile is an aqueous solution of organic and inorganic compounds including bile acids, cholesterol, phospholipids and the pigment biliverdin which provides the distinctive yellow/green appearance [244]. Bile has a pH between 7.5-8 with an osmolality of 300 mOsm/kg, and contains immunoglobulin A and mucus to prevent growth and adhesion of bacteria [245]. Approximately 1 litre of bile is synthesised by liver pericentral hepatocytes per day. After secretion, bile is transported along the bile ducts for storage in the gallbladder [247]. Here, the bile is concentrated by 5–10-fold, water and electrolytes are removed and the bile is acidified by Na^+/H^+ exchange. Not all bile is processed and stored in the gallbladder; approximately half of the bile is released directly into the duodenum and undergoes continuous enterohepatic cycling where bile acids are recycled in the liver [245]. Bile also facilitates enterohepatic cycling of many endogenous substances such as certain vitamins, hormones and trace metals. Some exogenous substances are also secreted into bile for a degree of enterohepatic cycling including antimicrobial agents, and drugs [245]. Bile salts range in concentration from 3-45 mM l^{-1} (0.2 -2% wt/vol) depending on the individual, and the type and amount of food ingested [248, 249]. High concentrations are found in the duodenum, jejunum and proximal ileum to facilitate fat digestion and absorption [250].

Bile salts are synthesised in the liver from cholesterol via a multienzyme process and represent approximately 50% of the organic components of bile (Figure 3.1). Synthesis occurs through two main pathways, the "classical" and the "alternative" pathway. The classical pathway occurs in the liver and leads to the generation of chenodeoxycholic acid (CDCA) and cholic acid (CA), two important human primary bile acids (Table 3.2) [245, 251, 252].

Prior to secretion, the bile salts must first be conjugated by N-acyl amidates (peptide linkage) with glycine (glycoconjugated) or taurine (tauroconjugated), these are present in a ratio of 3:1 respectively in average human bile (Figure 3.2). Conjugation increases the water-solubility of the bile acids by lowering the pK_a of the terminal acidic group [244]. The amidation also decreases the passive absorption of bile acids in the biliary tract and small intestine, ensuring high levels of bile are available to aid fat digestion and adsorption. When conjugated, bile acids are generally referred to as bile salts due to the fact that they are completely ionised at physiological pH [245]. The bile salts are amphipathic and thus can self-associate in water to form polymolecular aggregates known as micelles. Micelles form when bile salts reach a sufficient concentration known as the critical micellar concentration (CMC) [244]. Approximately 5% of bile salts evade absorption by epithelium and are extensively modified through specialised biotransformation reactions by the indigenous intestinal microflora [244, 245].

Bile acids	Abbreviation	Conjugates	Abbreviation	Site of production
Primary bile acids				
Cholic	CA	Taurocholic	TCA	Liver
		Glycocholic	GCA	
Chenodeoxycholic	CDCA	Taurochenodeoxycholic	TCDCA	Liver
		Glycochenodeoxycholic	GCDCA	
Secondary bile acids				
Deoxycholic	DCA	Taurodeoxycholic	TDCA	Intestine
		Glycodeoxycholic	GDCA	
Lithocholic	LCA	Taurolithocholic	TLCA	Intestine
		Glycolithocholic	GLCA	
Ursodeoxycholic	UDCA	Tauroursodeoxycholic	TUDCA	Intestine
		Glycoursodeoxycholic	GUDCA	
Sulpholithocholic	SLCA	Taurosulpholithocholic	TSLCA	Liver
		Glycosulpholithocholic	GSLCA	

Table 3.2 Bile acids and their conjugates. Adapted from Floch, 2002 [248].

The common bile acids and their conjugates are shown in the table alongside common abbreviations. Synthesis of primary bile acids occurs exclusively in the liver as does conjugation. Deconjugation occurs in the intestine.

Table 3.1 Major constituents of human hepatic bile.

Table adapted from Begley *et al*., 2005 [245].

Constituent	mmol/l
Sodium	145
Chloride	90
Bile salts	40
Phospholipids	7
Potassium	4
Cholesterol	3

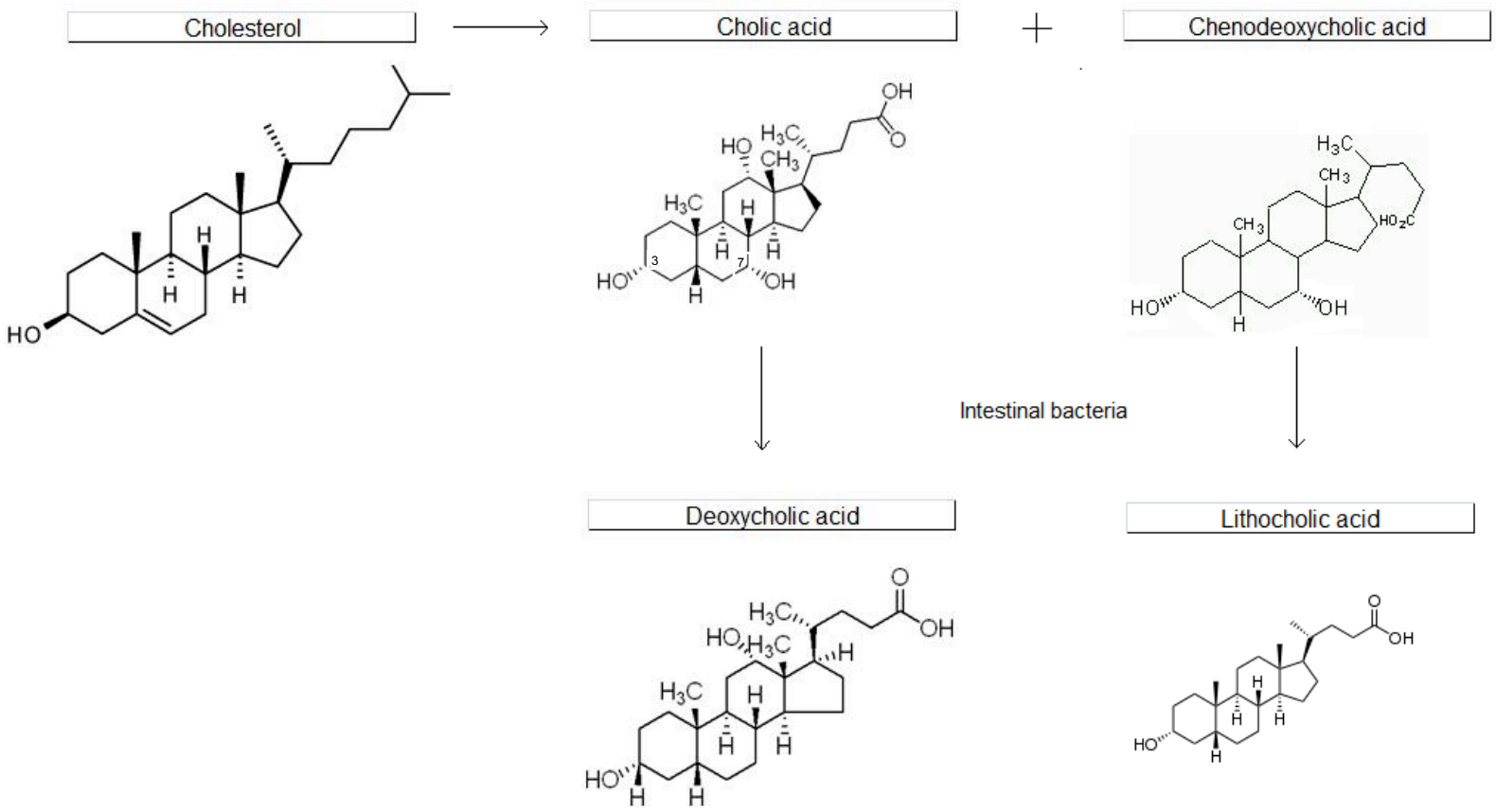


Figure 3.1 Chemical structure of cholesterol and the major primary and secondary bile acids found in human bile.

Cholic acid and chenodeoxycholic acid are synthesised in the liver from cholesterol via a multienzyme process. Indigenous intestinal microbiota carry out various biotransformation reactions to generate the secondary bile acids deoxycholic acid and lithocholic acid, both of which can be excreted in the faeces.

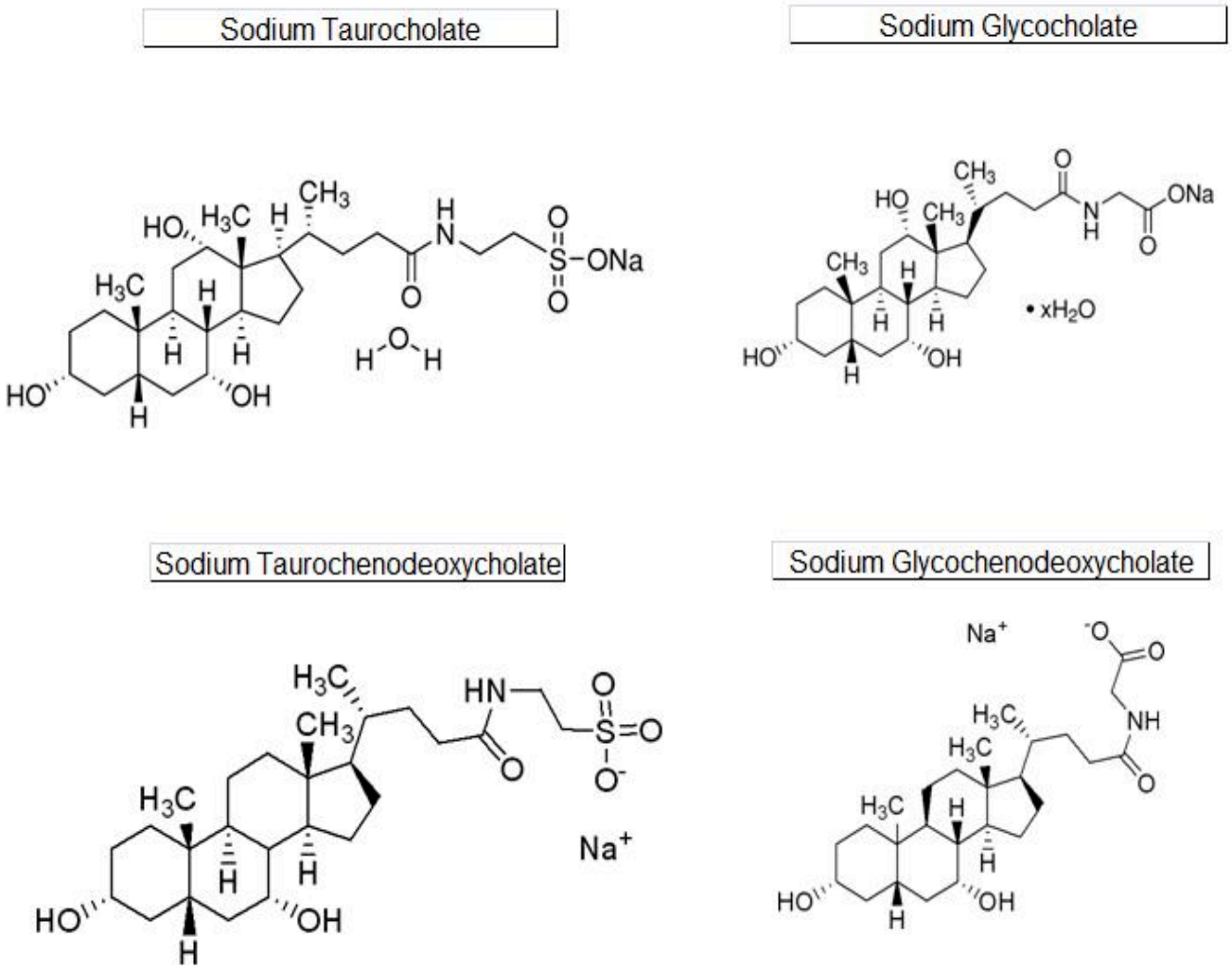


Figure 3.2 Structure of the common conjugated bile salts found in human bile primarily tauro and glyco conjugates of cholate and chenodeoxycholate.

3.2.1. **Bacterial biotransformation/modification of bile salts**

During intestinal transit, particularly in the caecum and colon, bile salts undergo further modifications by the indigenous intestinal bacteria. There are three main types of biotransformation i) deconjugation, ii) 7 α/β–dehydroxylation, and iii) 7 α-dehydrogenation (Figure 3.3) [252].

Deconjugation removes taurine or glycine by enzymatic hydrolysis of the C-24 N-acyl amide bond which links the amino acid conjugate to the bile salts [250]. This hydrolysis is catalysed by bile salt hydrolases (BSHs) produced by a broad spectrum of anaerobic intestinal bacteria including *Clostridium*, *Bacteroides*, *Lactobacilus*, *Bifidobacterium* and *Enterococcus* [245, 250, 251, 253]. BSHs are thought to improve bacterial colonisation of the lower GIT in higher mammals possibly by detoxification of the bile salts [250].

Dehydroxylation (replacement of a hydroxyl group with a hydrogen) occurs at the 7 position in the nucleus converting the primary bile salt CA to the secondary bile salt DCA and similarly CDCA to LCA (Figure 3.1) [251]. This transformation is quantitatively the most important bacterial bile salt biotransformation in the colon [252]. Execution of this rapid multistep transformation is restricted to specific members of the *Clostridium* and *Eubacterium* (approximately 0.0001% of colonic microbiota) [245, 250]. Dehydroxylation is limited to free bile acids and thus removal of glycine/taurine by BSHs is a prerequisite for 7 α/β–dehydroxylation by intestinal bacteria, with 7 β–dehydroxylation appearing to be nonessential [250].

Image unavailable due to copyright restrictions

Figure 3.3 Biotransformation reactions of bile salts performed by bacteria in the human intestinal tract. Taken from Ridlon *et al*., 2006 [250].

Schematic diagrams show the bacterial bile salt-biotransforming reactions in the human intestinal tract. The hydroxy group carbons of cholate are numbered and the AB rings are identified. Nomenclature is that of Hofmann *et al*. (1992) [254]. (BSH = bile salt hydrolase; HSDH = hydroxysteroid dehydrogenase).

Deconjugation and dehydroxylation increase the hydrophobicity and pK_a of the bile salts enabling recovery via passive absorption across the colonic epithelium [250]. The increased hydrophobicity is also thought to be associated with increased toxicity of the bile acids; high levels of secondary bile acids have been linked to cholesterol gallstone disease and colon cancer [245, 251, 255]

7 α-dehydrogenation involves oxidation and epimerisation of bile acid hydroxy groups. Oxidation of the A-ring of the steroid nucleus leads to conversion of CDCA to 7-oxolithocholic acid which then undergoes epimerisation by bacterial enzymes to form urodeoxycholic acid [251]. This reaction is carried out by hydroxysteroid dehydrogenase (HSDH) expressed by intestinal bacteria. Performing 7 α-dehydrogenation leads to energy generation and lowers the hydrophobicity and concentration of potentially toxic bile acids [250]. The physiochemical properties of bile acids can also be affected by changes to the hydroxy group stereochemistry, notably epimerisation of the 7α–hydroxy group of CDCA decreases its hydrophobicity and toxicity [251].

3.2.1. **Antimicrobial action of bile**

Three important agents produced by the GIT, whose primary role is to aid digestion, also play a significant role in the prevention of infection; including gastric secretions, hydrochloric acid and bile. Together gastric secretions and hydrochloric acid lower the pH in the stomach to between pH 1.5-2 which destroys the majority of ingested bacteria [247]. Bacteria which fail to be eradicated in the stomach then face high concentrations of bile in the small

intestine, which alongside rapid transit times, antimicrobial peptides (e.g. defensins), IgA and proteolytic enzymes, help defend the GIT against many invading pathogenic organisms [256].

The most significant antimicrobial action of bile is to cause membrane damage. Observations by electron microscopy, reveal that upon exposure to bile, cells become shrunken and empty, with leakage of intracellular material, this leakage has been confirmed by enzyme assays [245, 257]. Alteration of the membrane through environmental stimuli such as acid adaptation or increased osmolarity, and entry of the cell into stationary phase can increase a bacterium's resistance to bile-associated damage. Furthermore, mutants which are resistant to the deleterious effects of bile are in the majority of cases associated with mutations in loci linked directly or indirectly to membrane integrity [245].

The effects of bile salts on cell membranes vary according to several parameters, including concentration, the type and structure of the bile salt and the membrane architecture of the affected cell [245]. Membranes are rapidly dissolved and integral membrane proteins dissociated by high bile concentrations, leading to leakage of cell contents and ultimately cell death [245, 258]. Low or sub-micellar bile concentrations can subtly alter membrane integrity by leading to changes in fluidity and permeability [259]. Low concentrations can also alter the activity of membrane bound enzymes and affect the physical chemical properties of the cell surfaces such as hydrophobicity and zeta potentials (electrokinetic potential) [245, 260].

Binding of bile acids to membrane lipids is correlated to their hydrophobicity and therefore, the type of bile salt and its structure is a factor in the severity of membrane damage. Conjugated bile acids are strong acids and generally fully ionised at physiological pH [244]. They are therefore maintained in the outer hemileaflet of the bilayer unless actively transported into the cell. Unconjugated bile acids are able to 'flip-flop' passively across the lipid bilayer into the cell at a rate dependent on the number of hydroxy groups; dihydroxy flip more rapidly than trihydroxy bile acids [245]. Porcine bile, closely related to human bile in composition, is composed of dihydroxy bile acids and is significantly more inhibitory to bacterial membranes than bovine bile, which is composed of trihydroxy bile acids [245, 261]. Membrane architecture and composition including alteration of membrane characteristics such as charge, hydrophobicity and lipid fluidity, also play a role in the resistance of membranes to bile-associated damage. Additionally, structural damage to lipopolysaccharides caused by freezing has been shown to increase the sensitivity of *E. coli* to bile acids, and the presence of carbon dioxide can increase the permeability of membranes to bile acids [245, 262].

Bile acids have been shown to have a number of other effects on bacterial cells by disturbing macromolecule stability. Evidence suggests that they can induce DNA damage and activate DNA repair enzymes in bacterial and mammalian cells, and lead to secondary structure formation in RNA [245, 263, 264]. Bile may also induce oxidative stress by generation of oxygen free radicals and intracellular bile salts may lead to pH stress and osmotic effects. Furthermore, the detergent actions of bile can alter protein conformation

leading to misfolding or denaturation [245]. Bile salts may also act indirectly to prevent bacterial growth. Inagaki *et al*. (2006) present evidence that conjugated bile acids can regulate the expression of host genes involved in innate immune defence mechanisms in cells of the small intestine [265].

3.2.2. **Bacterial mechanisms for bile tolerance**

The human intestines contain a plethora of indigenous microflora which play a role in digestion, carrying out a wide range of metabolic actions and offering protection against invading pathogenic organisms through colonisation resistance (Figure 3.4) [266]. The small intestine is relatively poorly colonised with populations up to 10^5 bacteria, primarily aerobes [248]. Conversely, the colon harbours over 10^{11} bacteria per gram of wet weight of faeces which are predominately anaerobes [250, 267-269]. High levels of deconjugated bile acids exist in the small intestines which require the colonising bacteria to possess specialised adaptations for survival [175, 244].

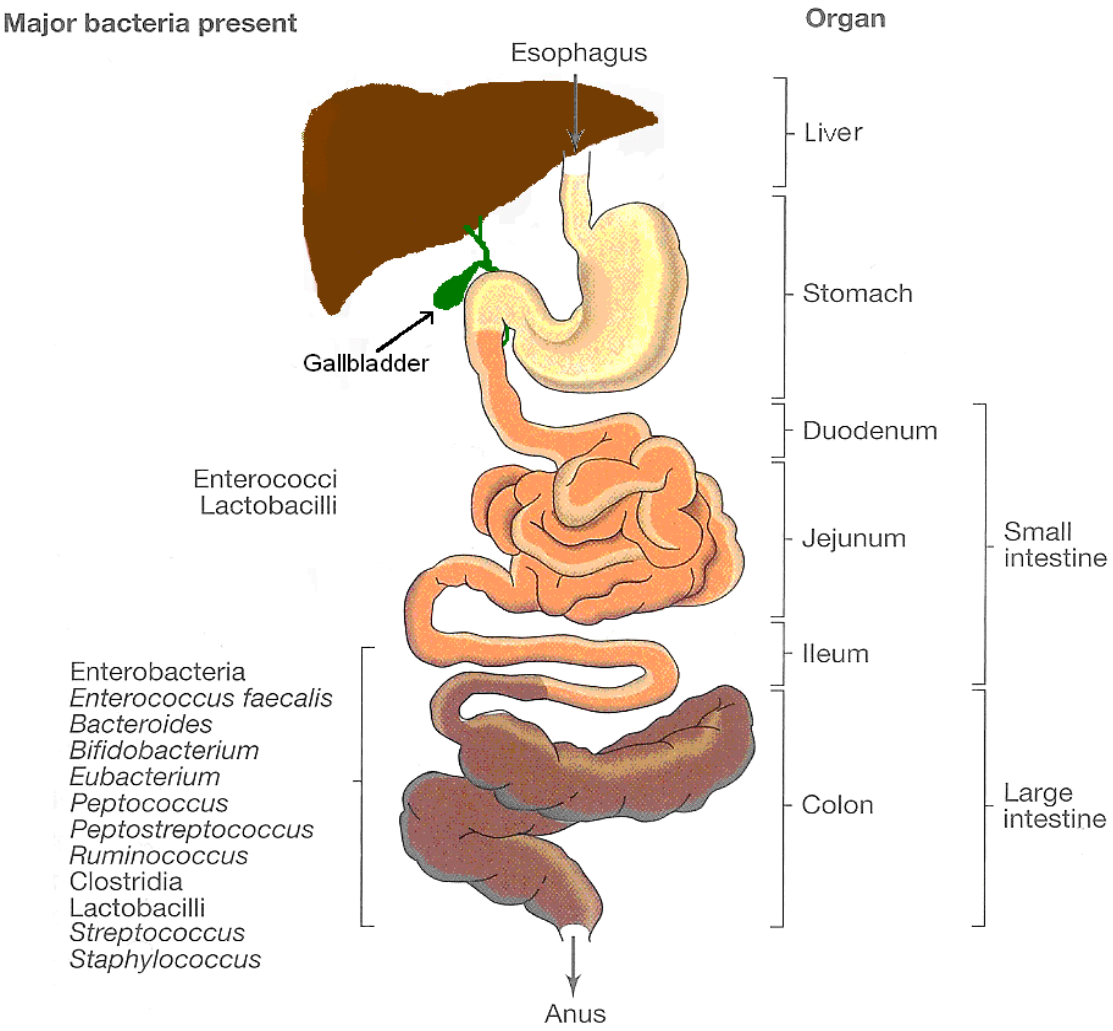


Figure 3.4 Cartoon of the anatomy of the human gastrointestinal tract. Adapted from Brock Biology of Microorganisms [269].

The distribution of non-pathogenic bacterial classes usually present in the small and large intestine of normal individuals is indicated.

Due to the potentially catastrophic actions of bile acids on bacterial membranes, enteric bacteria have developed several, often specific, tolerance mechanisms in response to the presence of bile [175]. Experiments performed with *E. faecalis* and *L. monocytogenes* reveal that exposure to sub-lethal concentrations of bile acids has a protective effect, dependent on growth phase and bile concentration, leading to rapid adaptation of the cell allowing it to withstand typically lethal concentrations of bile [270-272]. Assessment of protein expression by *Salmonella* spp. has also revealed significant differences in expression after exposure to bile [175]. How bacteria sense bile is not yet fully understood, it has been suggested that sensor kinases of two component regulatory systems may initiate a phosphorylation cascade with a cognate response regulator; or that bile may be sensed by 'AraC-like' transcriptional regulators which are able to regulate a diverse range of virulence factors; or finally that the sensory mechanism is an indirect process where the actions of bile on the cell are sensed rather than the bile itself [175, 273]. Clearly, further investigation is required to fully understand the mechanisms behind the resistance of enteric bacteria to bile. Such advances have the potential to aid the development of effective treatment strategies and the design of novel infection prevention methods [169].

3.2.2.1. LPS

Gram-negative bacteria are known to be inherently more resistant to the actions of bile, with bile salts often used for the selective enrichment of growth media [175, 245, 248]. *Salmonella* spp. show a high level of resistance to bile acids and are able to colonise the gallbladder. This is thought to allow

the chronic asymptomatic carrier state which occurs in 2-5% of infected individuals [274]. Investigation of the interaction of *Salmonella* spp. with bile has revealed increased biofilm formation and invasive potential [245]. Similarly, *E. coli* and *Campylobacter* spp. are often isolated from the gallbladder and bile of animals and humans [175, 245]. The Gram-negative cell wall provides an excellent hydrophobic barrier to bile salts (Figure 3.5). The lipopolysaccharide (LPS) represents the major portion of the outer leaflet of the outer membrane with the O-antigen providing a major barrier to external compounds [175].

3.2.2.2. Efflux pumps

Efflux of bile from the bacterial cytoplasm is the best characterised form of bile salt resistance. These systems are also responsible for efflux of other toxic compounds including antibiotics, organic solvents and oxidative stress agents [175, 275]. The intrinsic resistance shown by *E. coli* to solvents, detergents, lipophilic antibiotics (for example erythromycin) and bile salts is mediated by the *acrAB* encoded efflux system. In this well characterised system, AcrB is a proton motive force-dependent drug efflux transporter; AcrA bridges the inner and outer membrane to aid efflux directly from the cytoplasm to the environment [175, 275, 276].

Image unavailable due to copyright restrictions

Figure 3.5 Schematic diagram of the structure of the Gram-negative and Gram-positive bacterial cell wall taken from Cabeen and Jacobs-Wagner, 2005 [277].

a) Gram-positive cell wall – a thick multilayered peptidoglycan sheath covers the cytoplasmic membrane, with teichoic acids embedded in the peptidoglycan, and lipoteichoic acids which extend into the cytoplasmic membrane.

b) Gram-negative cell wall - the outer membrane is linked by lipoproteins to a thin, primarily single-layered peptidoglycan sheath, located within the periplasmic space. The outer membrane contains porins to allow passage of small hydrophilic molecules across the membrane, and lipopolysaccharide molecules that extend into extracellular space.

So-called Omp proteins are found on the outer membrane of Gram-negative bacteria and are important in the host-pathogen interaction; they are involved in a wide range of different functions including transportation of various molecules (porins), adhesion (adhesins), signalling pathways, cell structure and invasion of host cells and bacterial defence [278]. In *E. coli*, disruption of *ompF* and *ompC* leads to a bile sensitive phenotype, a phenomenon mirrored by disruption of *ompU* in *V. cholerae* [279, 280]. TolC is an outer membrane protein thought to aid the Acr system [276, 281]. Mutations in *tol* genes can have a direct impact on bile resistance by destabilising the outer membrane allowing greater access to bile salts [245]. Mutations of *tolC* in *V. cholerae* significantly reduced the ability of the organism to grow and colonise in the presence of bile [282].

In general Gram-positive bacteria are more sensitive to the effects of bile than Gram-negative bacteria. Numerous studies have been carried out on Gram-positive bacteria which may be used as probiotics in the human GIT, notably *Bifidobacteria* and *Lactobacillus*, as bile salt tolerance is an essential criteria for their suitability [283-285]. Unfortunately, the majority of these studies are not comparable due to technical differences; therefore it is difficult to draw accurate conclusions from these data. It is clear however, that the level of resistance to bile differs extensively between genus, species and strain [245].

Biotransformation reactions carried out by the microflora of the GIT are postulated to lead to the detoxification of bile acids which would otherwise exert deleterious effects on certain bacterial species [250]. Several Gram-positive indigenous bacteria have the ability to modify bile using BSHs to

render conjugated bile acids less toxic by removing the taurine/glycine conjugate [250, 286, 287].

3.2.3. **The interaction of bile with pathogenic bacteria**

Environmental conditions such as osmolarity, acidity and temperature are constantly sensed by bacteria during colonisation or infection of the host. These signals allow the organism to adapt to its environment by altering gene expression and regulation to aid survival, which in pathogenic bacteria includes regulation of virulence factors.

Vibrio parahaemolyticus is a marine and estuarine bacterium that causes acute gastroenteritis in humans. Several of this bacterium's virulence factors are affected by bile including thermostable direct haemolysin (TDH) and capsule production, adherence to epithelial cells and Congo red binding (associated with increased adherence and invasive potential) [245, 273, 288, 289]. Bacteria also show increased motility when grown in bile allowing them to reach the more favourable environment of the intestinal mucosal layer [245, 290]. Finally, as bile acids are calcium chelators they may trigger a low calcium response, which has been shown to regulate virulence factors in *Yersinia* spp. [245, 273].

The Gram-positive intracellular pathogen *L. monocytogenes* has been studied extensively with respect to bile tolerance. Like *S. typhimurium*, *L. monocytogenes* is able to colonise the gallbladder (known as cholecystis) and thus has a very high tolerance to bile. Begley *et al*. (2002) observed

growth of *L. monocytogenes* strain LO28 on agar supplemented with high bile concentrations of human (15%), porcine (2%), oxgall (15%) and bovine (15%) bile [245, 270]. Olier *et al.* (2004) demonstrated that clinical isolates from asymptomatic carriers, and food and environmental isolates could grow in broth supplemented with 5% porcine bile supporting the idea that *L. monocytogenes* is inherently resistant to bile salts [291]. Other studies have generated different results, and as seen with probiotic bacteria, there were major differences between strains. Experimental differences may also support the differing findings, Begley *et al*. (2002) note that pH can be a significant factor in the potential toxicity of glycoconjugated bile salts [270]. *L. monocytogenes* was the first pathogenic species not considered to be part of the normal enteric flora to be shown to have BSH. BSH is present in all pathogenic *Listeria* species and its expression *in vitro* in strain EGDe is positively regulated by PrfA, a transcriptional activator and the principal regulator of virulence genes in *L. monocytogenes* [287]. A PrfA-regulated bile exclusion system has also been identified as a virulence factor in *L. monocytogenes* [292]. The BilE operon contains two genes *bilEA* and *bilEB*, with significant homology to members of the betaine carnitine choline transporter family. Sleator *et al*. (2005) demonstrated that functional inactivation of BilE leads to a significant reduction in the bacterium's ability to withstand lethal concentrations of bovine bile (oxgall) and physiological concentrations of human bile. Additionally, cells with this phenotype were significantly less virulent in murine models [292].

E. faecalis is an important multidrug resistant nosocomial pathogen and is one of the most common lactic acid producing bacteria found in the GIT [293].

The bacterium has been isolated from bile, biliary drain devices used in biliary surgery, and, directly from the gallbladder, highlighting its high tolerance of bile [245, 294]. Alteration of the physiochemical surface properties of *E. faecalis* occurs after growth in bile (oxbile) resulting in increased invasion of biliary drain materials [295]. This property may enhance colonisation through increased adhesion [245, 295]. However *in vitro* analysis of *E. faecalis* by tolerance to bile showed the bacterium is rapidly killed after challenge with unconjugated bile acids [271]. This differential susceptibility to conjugated verses unconjugated bile acids highlights the potential drawback of *in vitro* analysis; the composition of the bile may not accurately reflect the *in vivo* conditions faced by the organism.

3.2.4. **The interaction of *C. difficile* with bile**

Clostridium spp., notably *C. perfringens,* are often found in the GIT and have been isolated from bile. However, little is known about their interaction with and tolerance mechanisms to bile salts. The ability of the clostridia to produce metabolically dormant highly resistant endospores enables them to avoid oxygen toxicity and to pass freely through the extremely acidic environment of the stomach [16]. Spores are therefore widely considered to be the aetiological agent of infection, however these spores must germinate in order to cause disease [296]. The process of germination and sporulation in clostridia has not been well studied, although parallels can be drawn from the model organism *B. subtilis* [169]. In particular, investigation of these processes in *C. difficile* has been limited by the lack of genetic tools.

A role for bile salts in the germination of *C. difficile* spores has been proposed suggesting that certain bile salts may act as cogerminants [169, 176]. Germination is initiated after exposure to specific environmental cues termed germinants, these include nutrients, muropeptides, cationic surfactants, high pressures, salts and CaDPA (a 1:1 chelate of dipicolinic acid and Ca^{2+}) [297]. In *B. subtilis*, germination is initiated by nutrients which are sensed by three sensory receptors GerA, GerB and GerK, leading to a cascade of reactions culminating in the commencement of metabolism [169, 297]. Homologues of GerA, GerB and GerK identified in many *Clostridium* species are absent from *C. difficile* [169].

After ingestion and passage through the acid stomach environment, *C. difficile* spores encounter high concentrations of primary bile salts such as CA in the small intestine. Upon entry to the anaerobic environment of the large intestine, the spores will encounter bile that evaded absorption and recycling. Under normal conditions, the indigenous microflora will convert these bile salts to secondary bile acids such as DCA through biotransformation reactions [244]. As observed for other bacteria that colonise the GIT, *C. difficile* may use bile acids as a signal to indicate its position in the GIT. The anaerobic environment and presence of deconjugated bile acids in the colon may signal germination allowing colonisation [169]. If the spore were to germinate in the small intestine the high concentrations of bile are likely to destroy vegetative cells, preventing colonisation. Sorg and Sonenshein (2008) have investigated the role of various bile salts in germination of *C. difficile* spores [169]. They confirmed previous research by Wilson *et al*. (1982 & 1983) who showed that spore recovery was enhanced

by the addition of the primary bile salt taurocholate to rich growth medium [168, 176]. Several studies have also shown that taurocholate (TCA) can enhance colony formation by *C. difficile* spores isolated from environmental surfaces [168, 298, 299]. Other cholate derivatives may act as germinants for *C. difficile* including cholate and glycocholate, although the authors failed to show that they were able to stimulate germination using their assays.[169]. Glycine was also identified as a co-germinant; as the bacterial microbiota deconjugate primary bile salts, free glycine is released into the lower bowel, and therefore available to stimulate germination. Sorg and Sonenshein (2009) also noted that the secondary bile salt deoxycholate could induce germination but that it prevented vegetative cell growth. Additionally chenodeoxycholate was shown to be an inhibitor of both germination and growth [170]. The authors hypothesise that should the normal gut flora be disrupted, say after administration of antibiotics, then the relative concentrations of deoxycholate will be reduced, creating a favourable environment for growth of vegetative cells and colonisation of the colon [169, 300]. However, the role of bile as an environmental cue for germination has been disputed by Paredes-Sabja *et al*. (2008) who failed to observe an increase in spore germination following exposure to bile salts, either alone or in nutrient-rich medium [171]. There are major differences in the methods used in these two studies, particularly those used to generate spore preparations and their storage and activation, which could account for the differences in the results.

Gene CD0065, from the 630 genome, is predicted by genome annotation to encode a bile acid inducible enzyme NADP-dependent 7-alpha-hydroxysteroid dehydrogenase which converts the primary bile acid,

chenodeoxycholic acid, into 7-keto-lithocholic acid, a potentially toxic secondary bile acid, although this has not been confirmed experimentally [6]. It is not clear how *C. difficile* might protect itself from lithocholic acid, and its role in germination has not been tested due to its poor solubility in water [169]. Genes CD3215 and CD3216 have been identified as orthologs of *bilEA* and *bilEB*, which encodes a bile exclusion system in *L. monocytogenes* [6, 292]. *C. perfringens* is known to express BSH on its cell surface enabling deconjugation of bile salts, however open reading frames with homology to these enzymes have not been found in *C. difficile* [169]. Investigation of the interaction of bacteria with bile is being increasingly assessed by expression analysis [283, 301-306]. Microarray technology allows quantitative assessment of the global response to certain conditions such as bile stress [307]. Such analysis can produce large quantities of data and through appropriate scrutiny, can lead to the identification of novel genes and pathways involved in the stress response, potentially identifying virulence factors suitable for further investigation [308].

Clearly, additional exploration of the interaction of bile with *C. difficile* is required. Enhanced understanding of these interactions may aid treatment and control strategies through the development of bile acid based therapeutics or environmental cleansing agents [169].

3.3. Aims and Objectives

To survive the GIT *C. difficile* must possess mechanisms to tolerate bile acid stress however, little research has been conducted to explore this important interaction. Although *in vitro* analysis cannot recreate the conditions of the complex *in vivo* environment, assessment of the tolerance profile of *C. difficile* to bile represents an interesting avenue of research. As *C. difficile* spores and vegetative cells interact with bile during passage in the human and animal GIT, there is a requirement to gain an understanding of how bile affects *C. difficile* and how *C. difficile* responds to this stress. The genetics of bile resistance in Gram-positive organisms is also poorly understood. Expression analysis of the bacterium's response to bile and the screening of a random transposon insertion library for changes in bile tolerance may identify novel virulence factors in *C. difficile*.

Objectives:

1. Develop assays to investigate the tolerance of *C. difficile* to bile salts in solid and liquid media using bovine and ovine, and porcine bile preparations.
2. Evaluate the tolerance shown by various *C. difficile* strains and by the soil dwelling non-pathogenic organism, *Clostridium acetobutylicum*.
3. Investigate the effects of sub-lethal bile concentrations on *C. difficile* prior to exposure to high bile concentrations.
4. Investigate the expression profile of *C. difficile* strain 630Δerm after growth in the presence of bile.

3.4. Materials & methods

3.4.1. **Bile tolerance assays**

3.4.1.1. Preparation of bile mix and individual salts

Dehydrated bile of bovine and ovine origin (B & O) (Oxoid, UK) was suspended in sterile distilled H_2O or sterile RCM broth (Oxoid) at 50% wt/vol and incubated for a minimum of 12 hours at 2-8^{o}C before use. Dehydrated bile of porcine origin (PB) (Oxoid, UK) was ground to a fine powder in a pestle and mortar before being suspended in sterile RCM broth at 10% wt/vol and incubated for a minimum of 12 hours at 2-8^{o}C before use. For each experiment, as the bile preparations were viscous they were warmed to 37^{o}C before use to aid pipetting.

3.4.1.2. Bovine & Ovine bile broth assay

50% B & O bile was added to RCM or BHI broth to generate the required percentage. 1 ml of each B & O bile solution was plated in to 96 deep well plates (ABgene, UK) in triplicate with appropriate controls and pre-reduced for a minimum of 3 hours (Figure 3.1). *C. difficile* strains R20291, 630Δerm and CD37 were grown in 10 ml BHI broth with or without addition of 0.05% cysteine in 20 ml universals for 18-20 hours under anaerobic conditions at 37^{o}C. *C. acetobutylicum* was grown in 10 ml RCM broth in 20 ml universals for 18-20 hours in anaerobic conditions at 37^{o}C. Cultures were centrifuged to pellet the bacteria; 15 min at 3000 x *g*, pellets were re-suspended in 0.25 ml fresh RCM medium. Wells were inoculated with *C. difficile, C.*

acetobutylicum or RCM broth as appropriate to establish a final OD at 600 nm (OD600 nm) of 0.1 at (except broth control) and mixed thoroughly. Plates were covered with an adhesive gas-permeable seal (ABgene, UK) and incubated for 24 hours in anaerobic conditions at 37^{o}C. After incubation a 100 µl sample from each well was transferred to a 96 well flat bottomed plate and optical density was measured using a microplate reader (MRXTC II, Dynex Technologies Inc., UK) at 590 nm with a 3 second shake prior to measurement. Appropriate (broth only) blank measurements were subtracted from each OD reading for data analysis. Prism 5 (GraphPad Software Inc, v5.01) and SPSS v19 were used to plot the data and for statistical analysis. Data were analysed for normality using a Q-Q plot and assessment of the Shapiro-Wilk statistic. A Mann-Whitney test for nonparametric data was performed to assess the significance of the observed growth differences.

Extended range assay

Starting cultures were grown as stated above and inoculated into BHI broth containing 0.1%, 0.3%, 0.5% increasing in increments of 0.5% to 12% B & O bile. Samples were taken and evaluated after 24 hours as described above. Data were analysed for normality as described above and found not to be normally distributed. A Kruskal-Wallis test for nonparametric data, with a post-hoc Dunn's multiple testing correction, was used to analyse the differences in growth under each condition.

3.4.1.3. Bovine & Ovine bile agar assay

50% B & O bile was added to BHI or RCM agar to generate the required bile percentage (0, 0.1, 0.3, 0.5, 1, 3, 5 and 10%). Cultures of *C. difficile* strains R20291, 630, 630Δerm and CD37 were grown in 10 ml BHI broth in 20 ml universals for 18-20 hours at 37^{o}C in anaerobic conditions. Cultures were standardized to an OD600 nm of 0.8 and spotted in quadruplicate onto pre-reduced agar using a multipoint replicator (Mast, Liverpool UK). *C. acetobutylicum* cultures, grown in 10 ml RCM broth in 20 ml universals for 18-20 hours, were standardised to an OD600 nm of 0.9 was plated alongside *C. difficile* on each test plate with the addition of a broth control using the multipoint replicator. Inoculated plates were incubated for 48 hours in anaerobic conditions at 37^{o}C before analysis.

3.4.1.4. Porcine bile agar assay

PB was added to RCM agar to generate the required bile percentage (0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 1, 2, 3, 4 and 5 %). *C. difficile* strains R20291, 630, 630Δerm and CD37, grown in 10 ml BHI broth in 20 ml universals for 18-20 hours, were standardized to an OD600 nm of 0.8, and spotted in quadruplicate onto pre-reduced agar using the multipoint replicator. *C. acetobutylicum* cultures, grown in 10 ml RCM broth in 20 ml universals for 18-20 hours, were standardised to an OD600 nm of 0.9. Cultures were plated alongside *C. difficile* on each test plate with the addition of a broth control using the multipoint replicator. Inoculated plates were incubated for 48 hours in anaerobic conditions at 37^{o}C before analysis.

Control plates without bile were also inoculated and incubated anaerobically and aerobically at 37^{o}C.

3.4.1.5. Growth curve

10 ml BHI broth was supplemented with 0.05% cysteine and 0.1%, 1%, 2.4%, or 5% B & O bile. Bile broths, and matched broth only controls, were inoculated in triplicate to a starting OD600 nm of 0.1 with 630Δerm culture grown in 10 ml BHI broth in a 20 ml universal for 18-20 hours. At regular time intervals 1 ml samples were taken and the absorbance measured OD600 nm using a spectrophotometer (Thermo-Fisher, UK). The data was plotted using Prism 5.

3.4.1.6. Bile adaptation assay

630Δerm and R20291 cultures were grown in 10 ml BHI broth in 20 ml universals for 16-18 hours to an OD600 nm of between 0.9-1.2. Cultures were supplemented with 0.1%, 1%, 2.4%, or 4% B & O bile in triplicate or with an equivalent volume of BHI broth. Cultures were incubated anaerobically for 15 minutes at 37^{o}C. Pre-reduced and warmed BHI broth supplemented with 0.1%, 1%, 2.4%, 4%, 8%, 10% and 12% B & O bile or BHI broth only controls, were inoculated with each of the pre-exposed samples in triplicate to an approximate starting OD590 nm of 0.05 in 96 deep well microtiter plates. Plates were incubated anaerobically for 24 hours without shaking at 37 ^{o}C. After incubation absorbance was measured at 590 nm using a microplate reader. Prism 5 and SPSS v19 were used to plot the

data and for statistical analysis. Data were analysed for normality as described previously and found not to be normally distributed. A Mann Whitney test was used to analyse the growth under each condition.

3.4.2. **Expression profiling**

3.4.2.1. Bile exposure assay

Vented 25 cm^3 tissue culture flasks containing 10 ml of BHI broth with 0.05% cysteine were inoculated with *C. difficile* strain 630Δerm to an initial OD600 nm of 0.1. Cultures were incubated with shaking (50 rpm) for 4 hours at 37^oC in anaerobic conditions (OD600 nm ~ 0.6-0.7) (cells determined to be at mid-log phase by growth curves performed by Nazilla Jafari, UCL Eastman Dental Institute). Under anaerobic conditions B & O bile (50% vol/vol solution) was added to a final concentration of 0.1, 2.4 or 4% in triplicate. For control cultures, also in triplicate, an equivalent volume of BHI broth was added. Cultures were incubated for a further 15 min (long exposure) or 1 min (short exposure) before growth was stopped by addition of RNA protect bacterial reagent (Qiagen, UK).

3.4.2.2. RNA extraction and assessment

Under anaerobic conditions 2 volumes of RNAprotect bacterial reagent (Qiagen, UK) was added to 1 volume of culture and vortexed immediately for 5 seconds. Samples were removed from the anaerobic chamber and incubated for 5 min at 15-25^oC and then centrifuged at 5000 x *g* for 10 min to pellet the bacteria. The supernatant was removed and the tubes were

drained on absorbent paper prior to freezing at -70^{o}C for 24-48 hours. All further processing and handling of RNA was conducted in controlled environments, using disposable plastic ware, filter tips, molecular grade RNase and DNase free water (Sigma-Aldrich, UK), and by the use of an RNaseZap solution (Ambion, UK) to minimize contamination. Two methods were used to extract the RNA, the first using an RNeasy kit (Qiagen, UK) involving enzymatic lysis and proteinase K digestion, and secondly, using a Fast RNA pro blue extraction kit (MP Biomedicals, USA) involving a phenol/chloroform based extraction method.

Extraction of RNA using an RNeasy kit (Qiagen, UK)

Extraction and purification of RNA were carried out according to the manufacturer's protocol 4 – Enzymatic lysis and proteinase K digestion and protocol 8 - Purification of total RNA from bacterial lysate using the RNeasy Midi kit. Briefly, frozen cell pellets were thawed on ice. 400 µg of Proteinase K (Qiagen, UK) was added to TE buffer, which contained 15 mg ml^{-1} of lysozyme (Sigma-Aldrich, UK). The cell pellets were carefully re-suspended by pipetting and vortexing for 10 sec and incubated for 10 min at 15-25^{o}C with periodical vortexing. Samples were processed according to the manufacturer's protocol without modification. RNA was eluted in 200 µl of RNase-free molH_2O passed though the column twice. The concentration of RNA was estimated using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, UK).

RNA extraction using a Fast RNA pro blue extraction kit (MP Biomedicals, USA)

1 ml of RNA pro solution was added to each bacterial cell pellet thawed on wet ice. Samples were processed according to the manufacturer's protocol without modification. The resulting RNA pellet was re-suspended in 100 µl of RNase-free molH_2O and the concentration estimated using a NanoDrop 1000 spectrophotometer.

3.4.2.3. DNase treatment

RNase-Free DNase Set (Qiagen, UK):

Before re-suspension of the RNA pellet in molH_2O, 2 µl of 10X DNase buffer was added to the pellet (RNase-Free DNase Set Qiagen, UK). 10 U of the RNase inhibitor and 0.5 Kunitz U of DNase I was added, the volume was adjusted to 20 µl with molH_2O. The sample was incubated for 30 min at 37^{o}C before addition of 2 µl 0.5 mM EDTA. The DNase I was inactivated by incubation at 65^{o}C for 5 min. RNA was stored at -70^{o}C.

DNA-free™ (Ambion, UK):

Aqueous RNA was transfer to a 0.5 µl microcentrifuge tube, 0.1 volumes of 10X DNase I buffer and 3 U of rDNase I was added and mixed gently. The sample was incubated at 37^{o}C for 30 min before addition of a further 3 U of rDNase I and 30 min incubation for rigorous treatment of the sample. 0.2 volumes of re-suspended DNase inactivation reagent was added, samples were vortexed and incubated at room temperature for 2 min with occasional

mixing. Sample was centrifuged at 10000 x *g* for 1.5 min and RNA transferred to a fresh tube. RNA was stored at -80^{o}C.

3.4.2.4. RNA quantification and quality assessment

Before use, RNA was quantified and assessed for quality on a Nanodrop 1000 spectrophotometer (Thermo Fisher, UK). The purity of the RNA was assessed using the ratio of sample absorbance at 260 and 280nm, samples giving a ratio between 1.9 and 2.1 were considered to be pure (based on the extinction coefficient of nucleic acids at 260 nm and 280 nm) [309]. A secondary measure of purity, using the ratio of absorbance of 260 nm and 230 nm was also taken, samples that gave a reading between 1.8 and 1.2 were considered to be pure. Some RNA samples were assayed using an Agilent 2100 Bioanalyzer using the RNA 6000 Nano Lab-Chip kit (Agilent Technologies, UK) to assess RNA integrity. This was performed by Dr. R. Stabler at the London School of Hygiene and Tropical Medicine, UK according to the manufacturer's guidelines.

3.4.2.5. Microarray hybridization

Sample labelling and hybridisation were carried out according to the BµG@S RNA vs DNA Labelling & hybridisation protocol (Bacterial Microarray Group at St. George's, London UK) [310]. All reagents used for labelling were purchased from Invitrogen life technologies, UK, except Cy3/Cy5 dCTP Fluorolink which was purchased from GE Healthcare UK. Microarray slides CDv1 were designed to include 3688 chromosomal predicted CDSs from strain 630. Each CDS was represented by a PCR product designed to

maximise sensitivity and minimise cross hybridisation, multiple reporters were included for certain genes with each one spotted in duplicate by split pin onto amino-silane coated glass slides (GAPS II, Corning) (Figure 3.6) [310].

Labelling:

For DNA, 1 µg of wild type 630Δerm DNA was added to 3 µg of random primers, H_2O was then added to a final volume of 41.5 µl. Samples were heated to 95°C for 5 min, snap cooled and briefly centrifuged. To each sample the following were added; 10X React 2 buffer, 5 mM dA/G/TTP, 2 mM dCTP, 25 nmol Cy5 dCTP and 3-9 U DNA polymerase 1 large fragment (Klenow) to a final volume of 50 µl. Samples were incubated in the dark at 37°C for 90 min.

For RNA, 8-10 µg of RNA was added to 3 µg of random primers, H_2O was then added to achieve a final volume of 11 µl. Samples were heated to 95°C for 5 min, snap cooled and briefly centrifuged. To each sample the following were added; 5X First Strand Buffer, 100 mM DTT, 5 mM dA/G/TTP, 2 mM dCTP, 25 nmol Cy3 dCTP and 500 U SuperScript II to a final volume of 50 µl. Samples were incubated in the dark at 25°C for 10 min followed by incubation at 42°C in the dark for 90 min. After incubation DNA and cDNA samples were combined and purified using the MinElute PCR purification kit (Qiagen UK). Samples were eluted with H_2O to a final volume of approximately 29 µl.

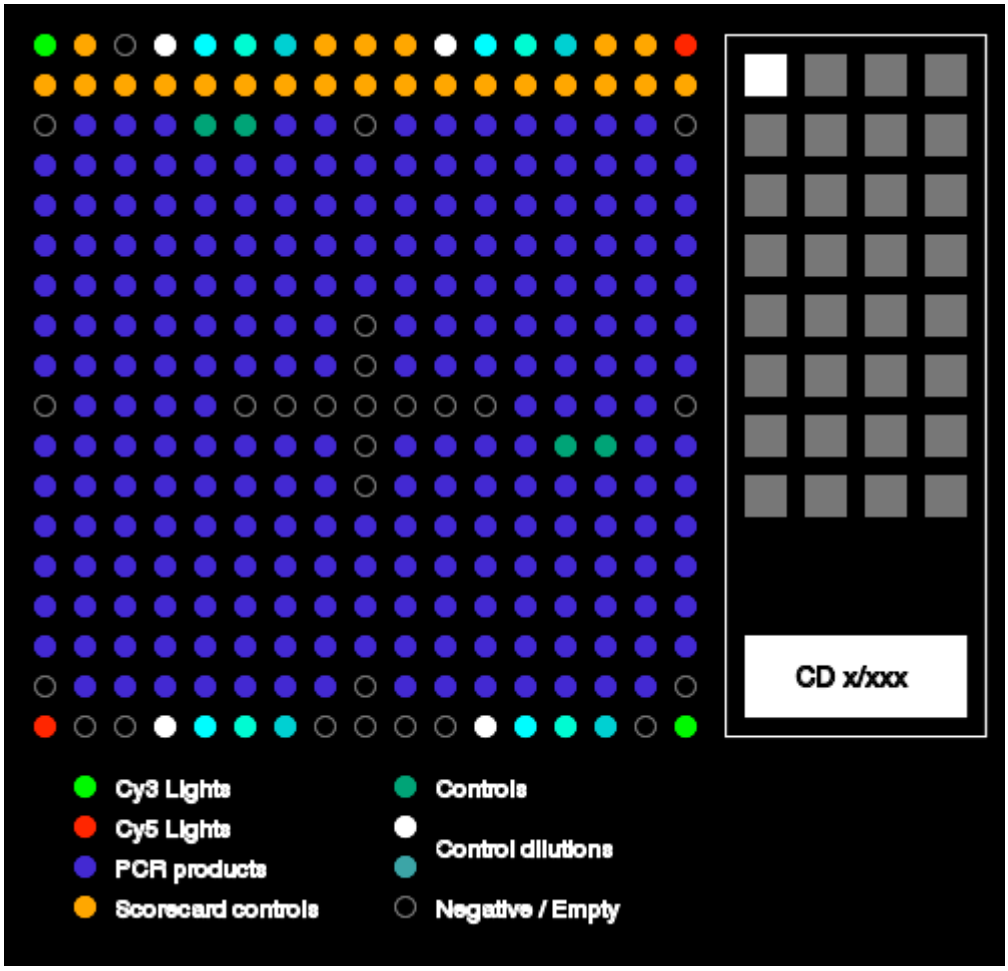


Figure 3.6 Strain 630 version 1_0_0 microarray slide design

The diagram represents the typical design of a grid found on the 630 microarray slide. Each blue spot represents a unique PCR product specific to genes found in *C. difficile* strain 630. Various control spots are included on the microarray to allow validation of the fluorescence data obtained after hybridisation.

Hybridisation:

Each microarray slide (Bμg@s, UK) was prehybridised in prehybridisation solution warmed to 65^{o}C containing 3.5X SSC, 0.1% SDS, 10 mg ml $^{-1}$ BSA Fraction V (all Sigma-Aldrich, UK) and H_2O. Slides were incubated in the prehybridisation solution for 20 min at 65^{o}C. After incubation, slides were rinsed thoroughly for 1 min with H_2O followed by a rinse with propan-2-ol (Sigma-Aldrich, UK) for 1 min. Slides were dried by centrifuging for 5 min at 433 x *g* and stored in the dark before addition of labelled samples. Purified Cy3/Cy5 labelled cDNA/DNA was mixed with 4X filtered SSC and 0.3% filtered SDS to a final volume of 46 μl before heating to 95^{o}C for 2 min. 15 μl H_2O was added to the wells in a hybridisation cassette (Telechem International, UK) and the prehybridised slide loaded. Two 22 x 22 mm LifterSlips (Erie Scientific, Portsmouth, USA) were carefully placed over the microarray area. The cooled sample/hybridisation solution mix was dispensed under one corner of each LifterSlip to cover each area by capillary action. The hybridisation cassette was sealed and placed in a water bath at 65^{o}C in the dark for 16-20 hours. After incubation, the slides were washed in a solution of 1X SSC & 0.05% SDS pre-warmed to 65^{o}C with continual agitation for 2 min. The slides were then transferred to a second wash solution of 0.06X SSC and washed with continual agitation for 2 min. This second wash was repeated. Slides were dried by centrifugation, 5 min at 433 x *g* and stored in the dark before scanning.

3.4.2.6. Microarray analysis/statistical analysis

Microarrays were scanned as soon as possible after washing by a GMS 418 microarray Scanner (Genetic Microsystems) using the microarray scanner software. Slides were initially scanned to ensure appropriate positioning and to identify the minimum gain required for maximum intensity but without signal intensity saturation. Scanning was repeated with gains increasing by increments of 15 points until the gain reached close or equal to 100 points for the RNA channel, and three times only for the DNA channel. 16 bit TIFF images were saved after each scan. Images for the Cy5 and Cy3 channels were uploaded into ImaGene 7.5 (Biodiscovery) for quantification and quality control. Images were aligned to the grid template, each grid and spot was checked systematically to ensure correct positioning by the software, unusable spots were flagged manually to be questioned or ignored during analysis. Raw data were combined using MAVI Pro v2.6.0 (MWG Biotech, UK) and paired datasets for Cy3 (RNA) and Cy5 (DNA) microarray fluorescence intensity quantifications were uploaded to GeneSpring GX 7.3.1 (Agilent Technologies, UK). Data were interpreted using the $\log_2$ ratios of the intensity of the RNA channel (Cy3) to that of the control DNA channel (Cy5). Data were initially normalised to each microarrays median (50^{th} percentile) intensity ratio (of present and marginal microarray elements) using standard Bµg@s parameters. Control and test samples were compared for statistical significance by parametric analysis using a one-way ANOVA welsh t-test with or without a Benjamini-Hochberg false discovery rate multiple testing correction at 5% confidence level.

3.5. Results

3.5.1. **Bovine and Ovine bile broth assay**

The growth of two pathogenic *C. difficile* strains (630Δerm & R20291) and *C. acetobutylicum* in bovine and ovine (B & O) bile was assessed after growth in a rich medium containing 0.08% to 16% bile (Figure 3.7). *C. acetobutylicum* was found to be unable to grow in all of the concentrations tested, with an average optical density (OD) for these cultures of 0.015 (range = 0.009 – 0.06, standard deviation = 0.008). The average OD value observed for growth in the absence of bile was 0.606 (range = 0.567-0.677, standard deviation = 0.043). The reduction in growth was found to be statistically significant in each of the concentrations tested with $P \leq 0.005$ (Mann-Whitney test).

The average growth, measured by OD, of R20291 in rich medium without the addition of bile was 0.371 (range = 0.341-0.387, standard deviation = 0.016), growth was observed in all of the bile concentrations tested except 16% where the average OD was 0.002. There was significant inhibition of growth in all bile containing broths compared to growth in broth without bile ($P \leq 0.005$, Mann-Whitney test). The growth of R20291 in medium containing bile can be seen in Figure 3.8 where the growth of each culture is expressed as a percentage of the total growth observed in broth without bile.

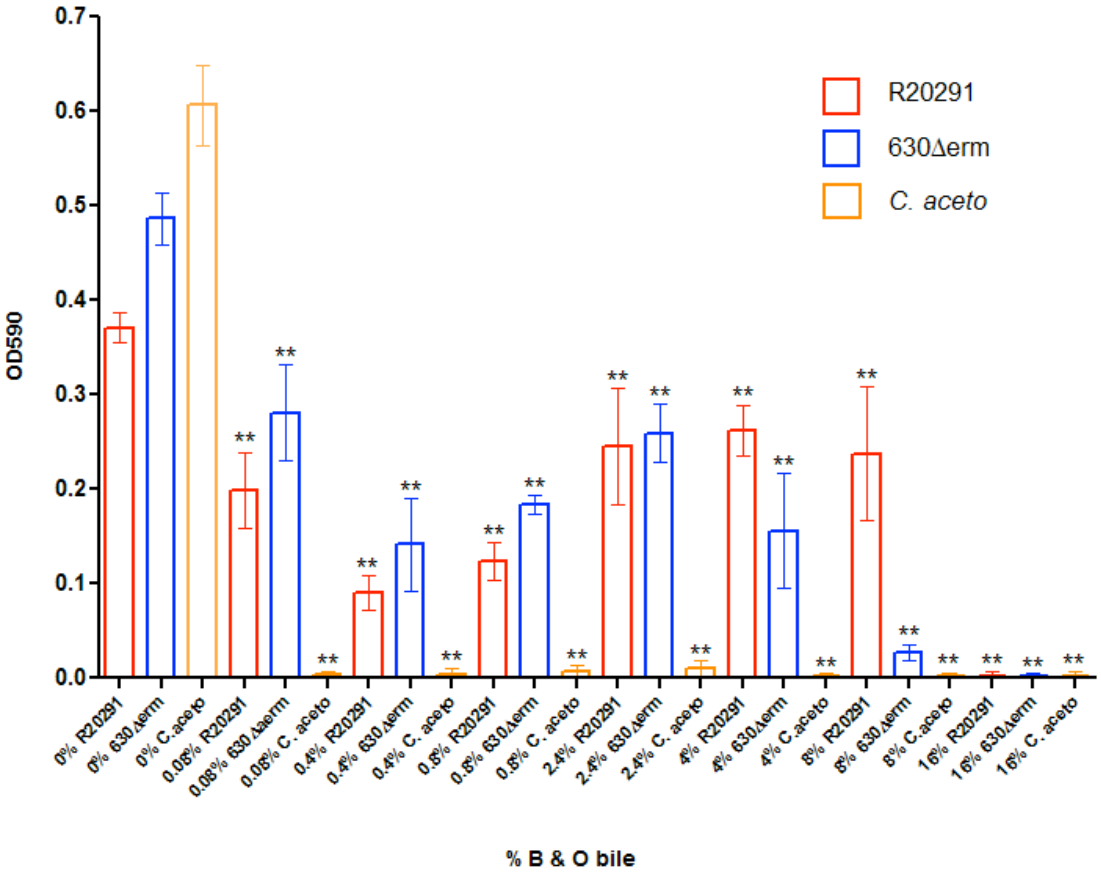
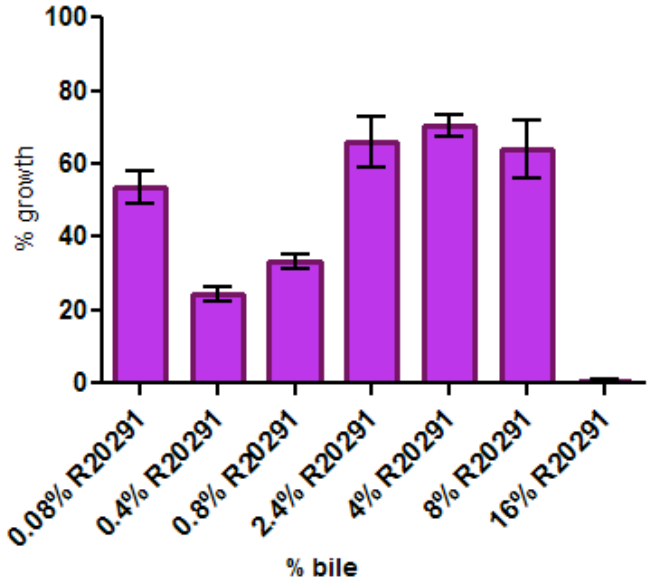


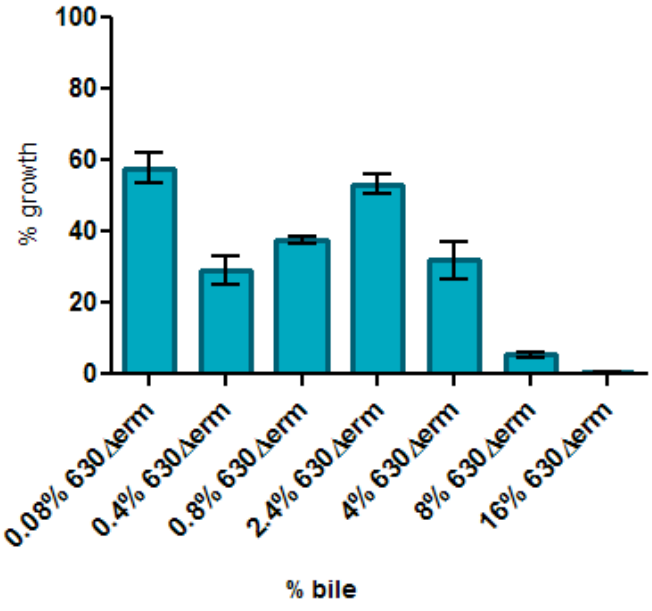
Figure 3.7 Growth of R20291 and 630Δerm in B & O bile

The growth of 630Δerm and R20291 in 0-16% bovine and ovine bile concentrations was measured using optical density (OD) at 590 nm 24 hours after inoculation. The OD of 630Δerm is represented in blue and the OD of R20291 in red. Bidirectional error bars represent the standard deviation of three biological and three independent experimental replicates. *C. acetobutylicum* (negative control) was included in each experiment; growth was observed in the no bile control and not observed in any of the bile test wells. Data was analysed using a Mann-Whitney test significance is denoted by ** ($P \leq 0.005$).



R20291 growth in B & O bile

A



630Δerm growth in B & O bile

B

Figure 3.8 Percent growth of R20291 and 630Δerm in B & O bile compared to unexposed culture growth.

The optical density (OD) values obtained for each test sample were compared to the OD obtained in the control samples and expressed as a percentage of total growth for both strains. Error bars represent the standard deviation obtained from three independent experiments.

The greatest inhibition of growth occurred for cultures grown in 0.4% bile with bacteria showing a four-fold reduction in growth compared to those grown in the absence of bile, with average OD values of 0.09 and 0.371, respectively (Figure 3.8). 0.4% bile is significantly more inhibitory than 0.08% and 0.8% bile with significance of P=0.002 and P=0.044, respectively. Growth in 2.4% to 8% bile ranged from 63-69% of the total possible growth observed for cultures grown in broth without bile, no significant differences were observed between the growth of 630Δerm vs R20291 at these concentrations.

The average growth of 630Δerm, measured using OD, in the absence of bile was 0.487 (Figure 3.7). Growth of 630Δerm was observed in broth containing 0.08 to 4% bile, however there is significant inhibition of growth at all of the concentrations of bile tested compared to the cultures grown without bile (P≤0.005). There is evidence of some growth of 630Δerm in 8% bile, an average OD of 0.027 was observed equating to 5.5% of the total growth observed in the unexposed control (Figure 3.8). 4% bile was significantly more inhibitory compared to 2.4% bile (P=0.015) and 0.8% bile was significantly more inhibitory than 2.4% bile (P=0.002).

Both strains appear better able to adapt to growth in high bile concentrations than in low concentrations. Growth in 0.4% bile was more inhibitory compared to 2.4% bile with an approximately three-fold reduction in growth for R20291 and a two-fold reduction for 630Δerm (significance P=0.002).

To characterise the inhibition of bile on the growth of R20291 and 630Δerm, particularly in low concentrations, an extended range bile broth assay was performed. Both strains were grown in broths containing 0.1 to 12% bile alongside control samples with growth measured by optical density. In general, the results of this assay support those observed in the initial experiment, revealing that growth in low bile concentrations (0.3-2% for R20291 and 0.3-1% for 630Δerm) is significantly more inhibitory than in medium to high concentrations (2.5% to 7.5% for R20291 and 1.3% to 5% for 630Δerm) (Figure 3.9 & Figure 3.10). However, growth in 0.1% bile was not found to be significantly inhibitory for either strain in this assay (see Appendix section 9.1).

For R20291 as the bile concentration increased from 0.1% to 0.3 and 0.5% growth decreases significantly (P<0.001) (Figure 3.9). Further increases in bile concentration are associated with increased growth, although this growth is significantly reduced compared to the no bile control up to bile concentrations of 2%. When the concentration of bile exceeds 2% there is no significant difference in the growth of these cultures compared to the no bile control. Bile concentrations over 7.5% led to significant inhibition of growth (P=0.05). R20291 was shown to grow reproducibly at bile concentrations of 12%, growing to approximately 25% of the total growth of the no bile control. The level of growth at 12% bile is comparable to the growth observed at 0.3% bile, suggesting that 0.3% and 12% B & O bile have a similar inhibitory effect on the growth of R20291 in broth culture.

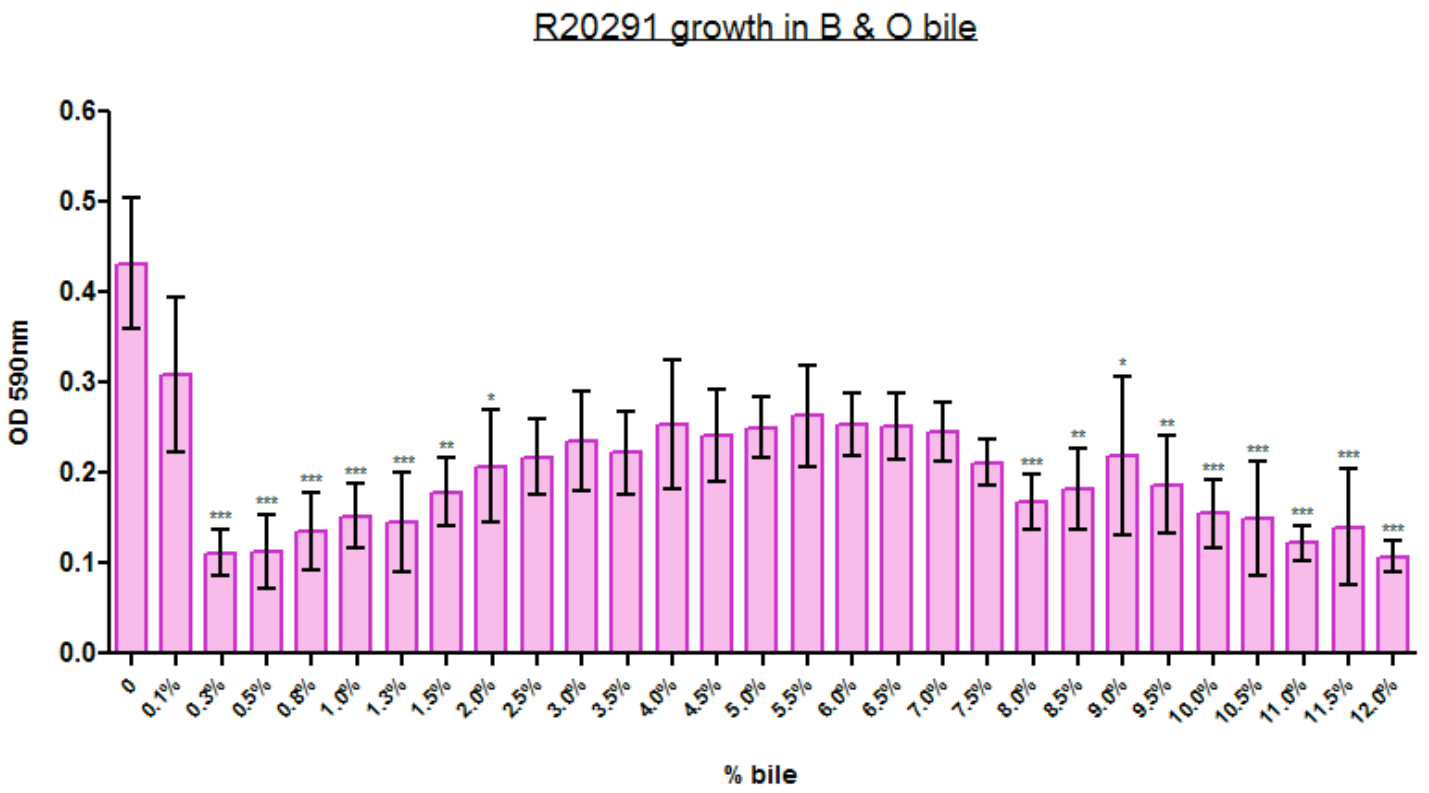


Figure 3.9 Extended range B & O bile broth assay for R20291.

A refined range bile assay was conducted to assess the growth of R20291 in bile concentrations ranging from 0.1 to 12%, with particular emphasis at the lower range of 0.1 to 1.5%. OD590 nm was measured at 24 hours post inoculation; error bars represent data from three biological and three independent experimental replicates. Significance is denoted by asterisks (* = $P<0.05$, ** = $P<0.01$, *** = $P<0.001$).

The growth of 630Δerm in the extended range B & O bile assay showed some similarities with the results obtained for R20291. Addition of 0.1% bile did not significantly inhibit growth, however significant inhibition of growth was observed in low bile concentrations of 0.3% to 1% ($P \leq 0.05$) (Figure 3.10). As the concentration of bile increased the growth of 630Δerm appeared to increase up to bile concentrations of 2%. After this point, there is an apparent reduction in growth; however, there is no significant difference between growth observed in each of the bile concentrations from 2% to 8.5% (see Appendix section 9.1). Statistically significant inhibition of growth is seen for cultures grown in bile concentrations of 5.5% and above when compared to growth in broth without bile.

Comparison of the tolerance of the two strains revealed that both strains were significantly inhibited at low bile concentration (<1.3% for 630Δerm and <2.5% for R20291). As the concentration of bile increased, both strains were able to grow in the bile broths with no significant inhibition. For 630Δerm this ability to tolerate increasing bile levels without a significant effect on growth, ceased as bile concentration reached 5.5%, for R20291 this point occurred at 8%. R20291's growth in 12% bile was 25% of the growth observed in the no bile control culture whereas 630Δerm was unable to grow at this concentration (2.2% growth).

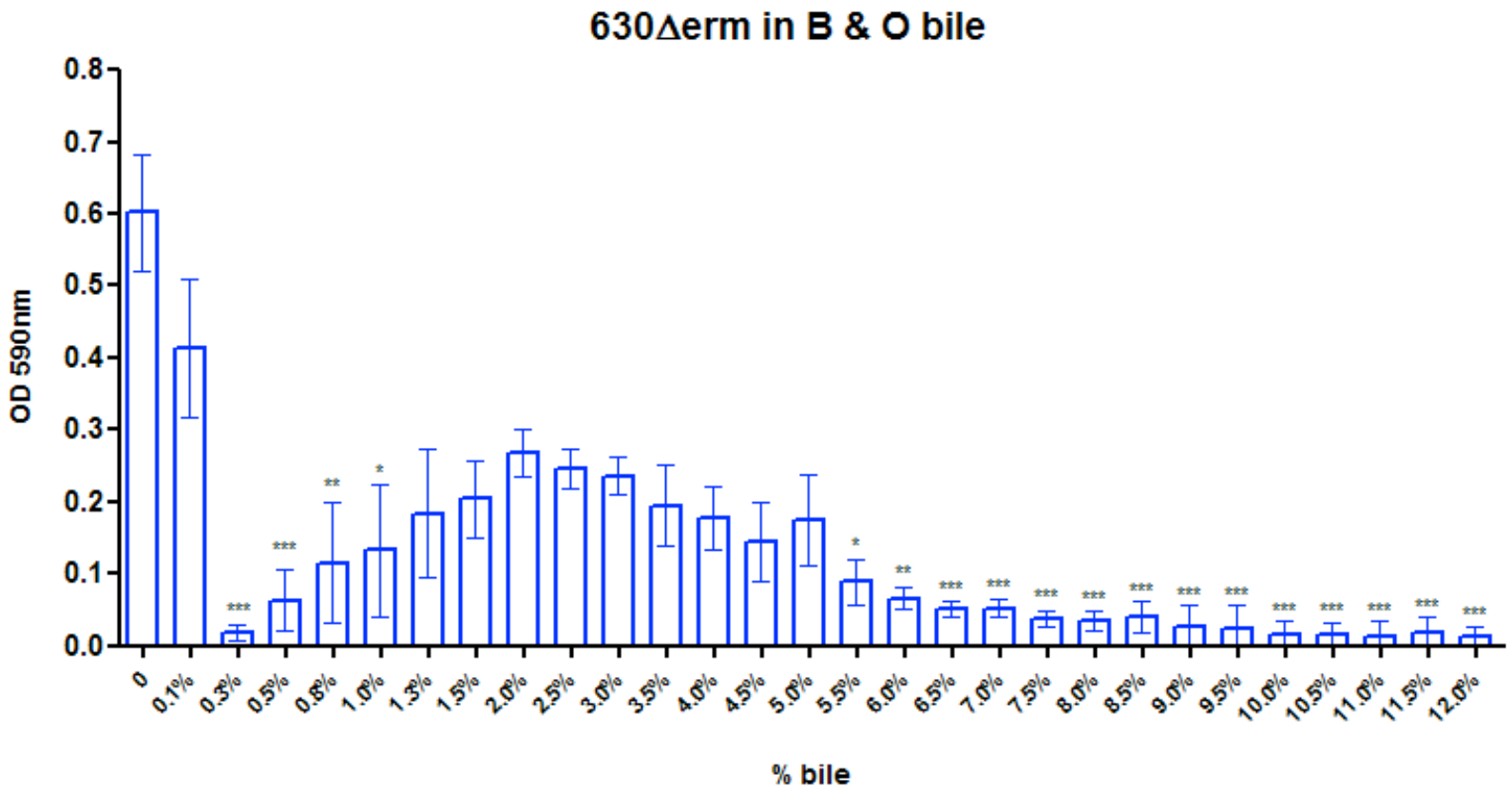


Figure 3.10 Extended range B & O bile broth assay for 630Δerm.

A refined range bile assay was conducted to assess the growth of 630Δerm in bile concentrations ranging from 0.1 to 12%, with particular emphasis at the lower range of 0.1 to 1.5%. OD590 nm was measured at 24 hours post inoculation; error bars represent data from three biological and three independent experimental replicates. Significance is denoted by asterisks (* = $P<0.05$, ** = $P<0.01$, *** = $P<0.001$).

3.5.2. **B & O bile agar assay**

To complement the broth assays, strains R20291, 630Δerm and *C. acetobutylicum* were spotted onto agar containing B & O bile at 0.1%, 0.5%, 1%, 3%, 5%, 10, and 20%. The results of this assay support those obtained from the broth assay, showing that R20291 was able to tolerate higher bile concentrations than 630Δerm (Table 3.3). In the broth assay the highest concentration of B & O bile tested was 16%, neither 630Δerm nor R20291 were able to grow in these broths, however, visible growth of R20291 was observed on agar containing 20% bile. Growth of *C. acetobutylicum* was not observed in any of the bile agar plates supporting the findings of the broth assay.

3.5.1. **Growth curve**

The growth of 630Δerm in the presence of various concentrations of B & O bile was assessed over time. Addition of B & O bile to the growth medium severely inhibited the growth of 630Δerm (Figure 3.11). After 9 hours the exposed samples had reached a third of the cell density seen in the unexposed control.

% B & O bile	R20291	630Δerm	*C. acetobutylicum*
0	+++	+++	+++
0.1	++	++	-
0.5	++	++	-
1	++	++	-
3	++	++	-
5	++	+/++	-
10	++	-	-
20	+	-	-

Key	
-	No growth
+	Poor growth
++	Good growth
+++	Full growth

Table 3.3 Growth of *C. difficile* and *C. acetobutylicum* on solid medium containing B & O bile

Colony formation of *C. difficile* and *C. acetobutylicum* on RCM agar containing B & O bile was assessed after 48 hours incubation. Growth was recorded according to the key above where poor growth = sporadic colonies observed in the inoculum area, good growth = colonies present in the majority of the inoculum area and full growth = dense colonies present over the entire inoculum area. Results represent the average result of three independent experiments.

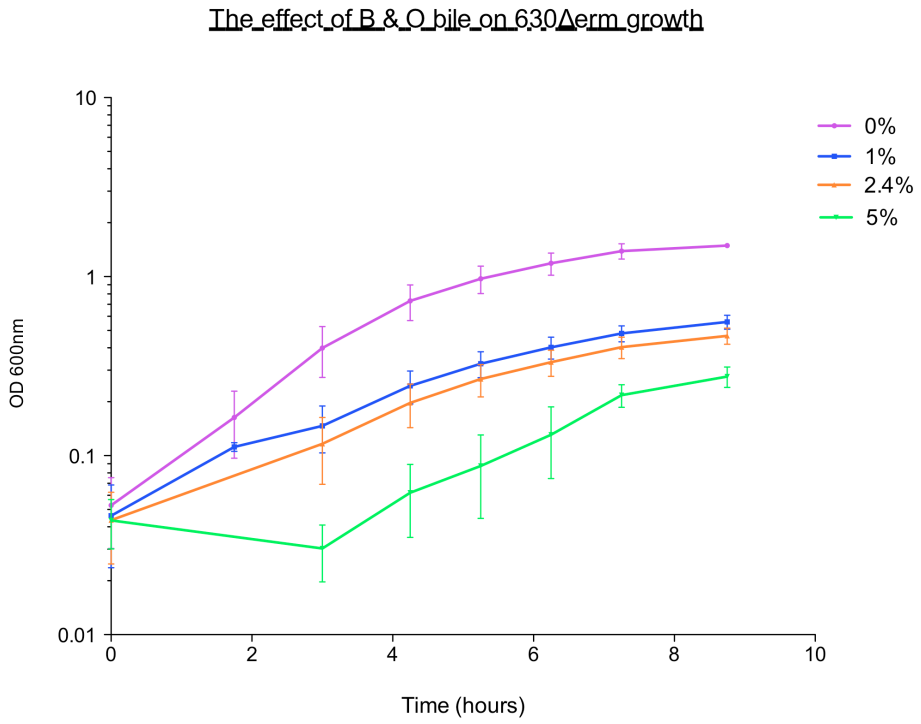


Figure 3.11 Effect of B & O bile on the growth of 630Δerm.

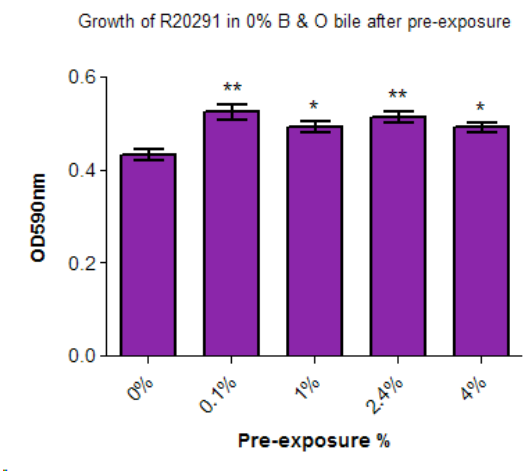
The growth of *C. difficile* strain 630Δerm was measured in 1%, 2.4% and 5% B & O bile and compared to growth in BHI medium without the addition of bile. Samples of the culture were taken at regular time intervals and the optical density measured at 600 nm. Cultures remained in an anaerobic environment during sample collection. Error bars represent the standard deviation resulting from three independent experiments.

3.5.2. **Bile adaptation assay**

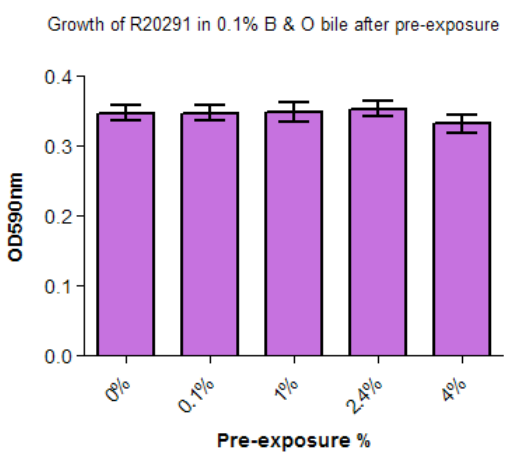
It has been reported that several bacterial species, such as *L. monocytogenes*, show increased tolerance to high levels of bile after exposure to sub-lethal levels [270]. To investigate whether such a response could be observed in *C. difficile*, cells were exposed to 0.1%, 1%, 2.4% or 4% B & O bile for 15 minutes before being used as inocula for a standard bile broth assay.

R20291 pre-exposed to 0.1 - 4% B & O bile and grown in rich media without bile, showed a significant increase in growth compared to non pre-exposed cultures (Mann Whitney test $P \leq 0.04$) (Figure 3.12, graph A). No significant difference was observed between pre-exposed and non pre-exposed cultures grown in 0.1%, 2.4% and 4% bile concentrations (Figure 3.12, graphs B, D & E). Pre-exposure to 0.1%, 1% and 2.4% led to significant increases in growth when cultured in broth containing 1% bile ($P<0.05$). Pre-exposure to 0.1% to 4% bile led to a significant increase in the growth of bacteria grown in 8% and 10% bile ($P \leq 0.004$), for those grown in 8% bile this equated to an average increase of approximately 1.5-fold (Figure 3.12, graph F & G). A small increase in growth was observed for bacteria pre-exposed to 2.4% bile grown in 12% bile broths ($P=0.006$) (Figure 3.12).

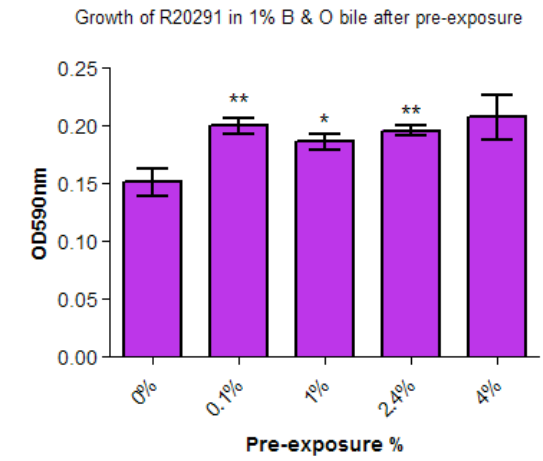
A


Growth of R20291 in 0% B & O bile after pre-exposure
OD590nm
0.0
0.2
0.4
0.6
**
*
**
*
0%
0.1%
1%
2.4%
4%
Pre-exposure %

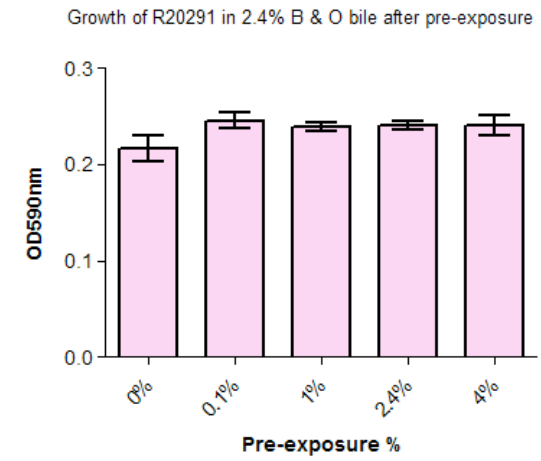
B


Growth of R20291 in 0.1% B & O bile after pre-exposure
OD590nm
0.0
0.1
0.2
0.3
0.4
0%
0.1%
1%
2.4%
4%
Pre-exposure %

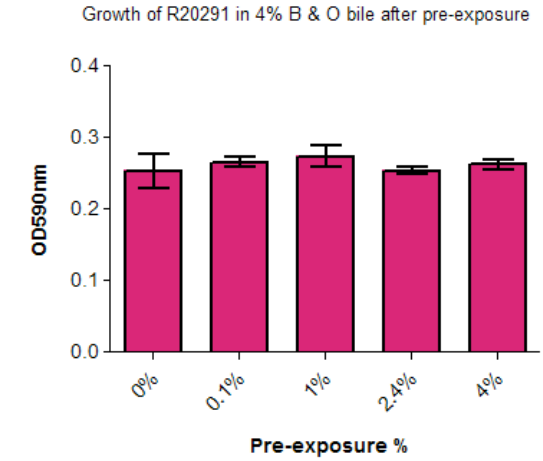
C


Growth of R20291 in 1% B & O bile after pre-exposure
OD590nm
0.00
0.05
0.10
0.15
0.20
0.25
**
*
**
0%
0.1%
1%
2.4%
4%
Pre-exposure %

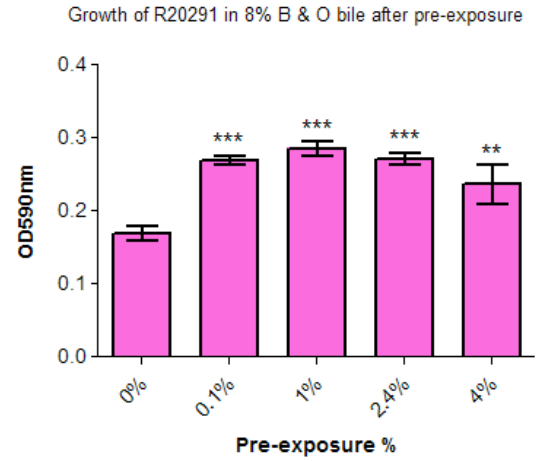
D


Growth of R20291 in 2.4% B & O bile after pre-exposure
OD590nm
0.0
0.1
0.2
0.3
0%
0.1%
1%
2.4%
4%
Pre-exposure %

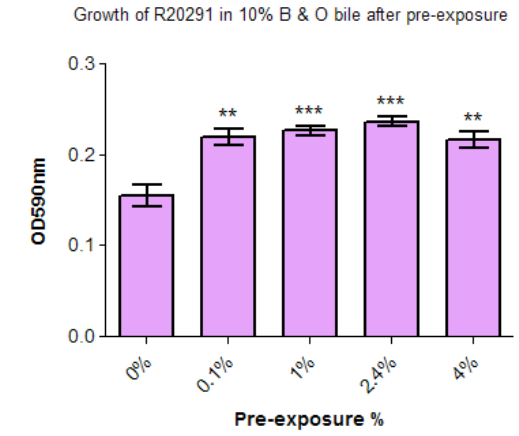
E


Growth of R20291 in 4% B & O bile after pre-exposure
OD590nm
0.0
0.1
0.2
0.3
0.4
0%
0.1%
1%
2.4%
4%
Pre-exposure %

F


Growth of R20291 in 8% B & O bile after pre-exposure
OD590nm
0.0
0.1
0.2
0.3
0.4
***
***
***
**
0%
0.1%
1%
2.4%
4%
Pre-exposure %

G



H

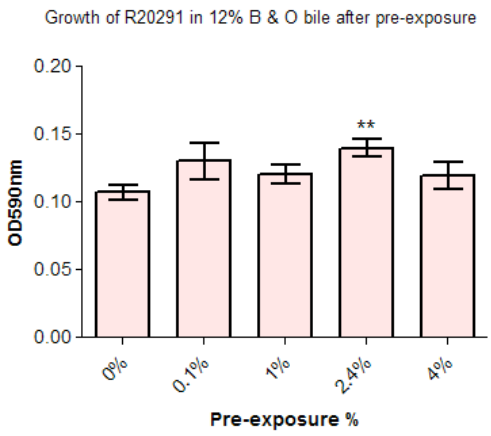


Figure 3.12 Comparison of the growth of R20291 in broth containing B & O bile after pre-exposure to B & O bile.

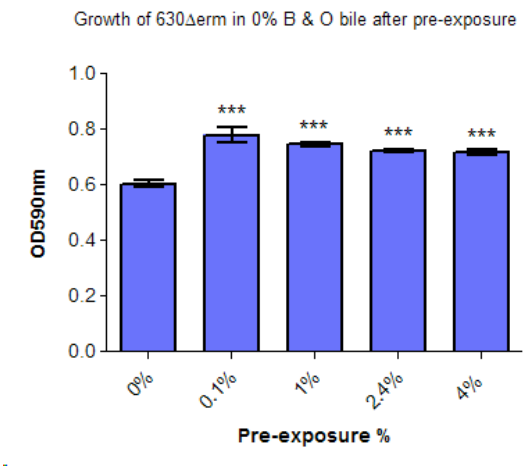
Graph A shows the growth of *C. difficile* strain R20291 in BHI broth after a 15 minute pre-exposure to 0.1%, 1%, 2.4% and 4% B & O bile, a non pre-exposed control is also included. Graphs B-H show the growth of R20291 in BHI broth containing 0.1 to 12% bile after a 15 minute pre-exposure to 0.1%, 1%, 2.4% and 4% bile. A non pre-exposed control is included for each condition. Growth was measured by optical density at 590 nm. Bidirectional error bars indicate the standard deviation resulting from three independent experiments. Significance is denoted by asterisks (* = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$).

630Δerm pre-exposed to 0.1% - 4% B & O bile in rich media without bile, showed a highly significant increase in growth compared to non pre-exposed cultures (Mann Whitney test $P<0.0001$). Highly significant differences were seen for growth in broth containing 0.1%, 1% and 8% bile for all pre-exposure conditions compared to non pre-exposed cultures ($P\leq0.004$) (Figure 3.13, graphs B, C & F). For growth in 1% and 8% there is an average increase in growth of approximately 2-fold. For cultures grown in 2.4% bile only those pre-exposed to 0.1% and 4% B & O bile showed a significant increase in growth compared to the non pre-exposed control ($P=0.0002$ and $P=0.002$, respectively) (Figure 3.13, graph D). For cultures grown in 4% bile only those pre-exposed to 0.1%, 2.4% and 4% B & O bile showed a significant increase in growth compared to the non pre-exposed control ($P=0.0009$, $P=0.024$ and $P=0.01$, respectively) (Figure 3.13, graph E). For cultures grown in 10% bile a significant increase in growth was observed for all of the pre-exposure condition compared to the non pre-exposed control ($P<0.012$) (Figure 3.13, graph G). Finally, a significant difference was observed for cultures pre-exposed to 4% bile grown in broth containing 12% bile.

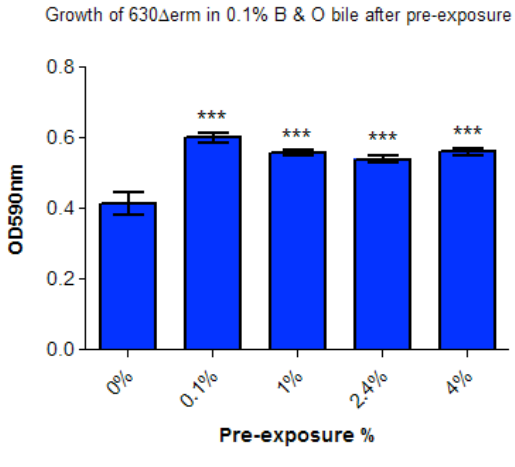
3.5.1. **Expression profiling in response to bile stress**

Bacteria often adapt to environmental stresses by altering expression of their genes to limit stress-induced damage [311-313]. To investigate the response of *C. difficile* strain 630Δerm to B & O bile stress, RNA was extracted from three replicate of mid-log phase cultures grown under control and test conditions for analysis. RNA was hybridised against a 630 genomic DNA control to allow multi-way comparisons between all RNA samples and to remove dye bias.

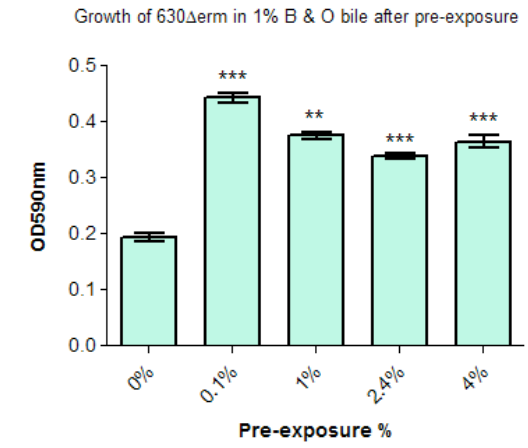
A


Growth of 630Δerm in 0% B & O bile after pre-exposure
OD590nm
1.0
0.8
0.6
0.4
0.2
0.0
***
***
***
***
0%
0.1%
1%
2.4%
4%
Pre-exposure %

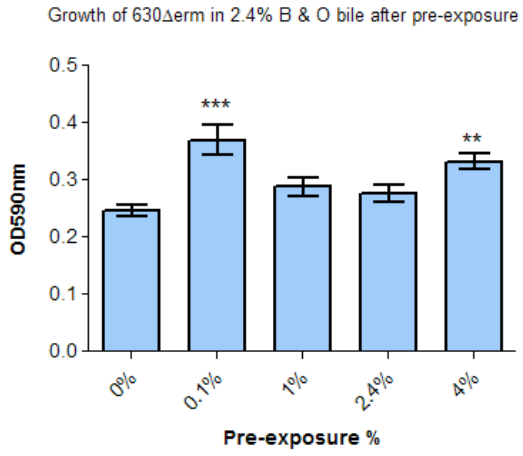
B


Growth of 630Δerm in 0.1% B & O bile after pre-exposure
OD590nm
0.8
0.6
0.4
0.2
0.0
***
***
***
***
0%
0.1%
1%
2.4%
4%
Pre-exposure %

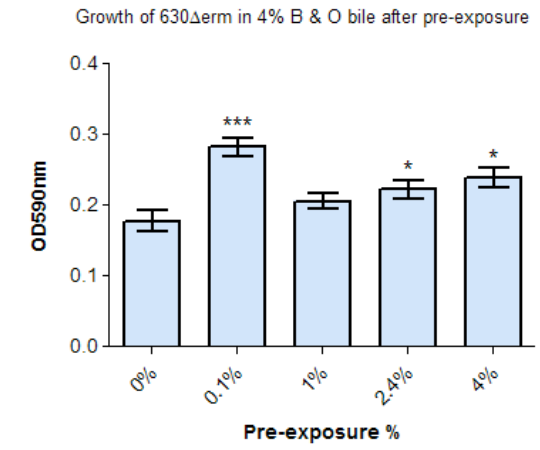
C


Growth of 630Δerm in 1% B & O bile after pre-exposure
OD590nm
0.5
0.4
0.3
0.2
0.1
0.0
***
**
***
***
0%
0.1%
1%
2.4%
4%
Pre-exposure %

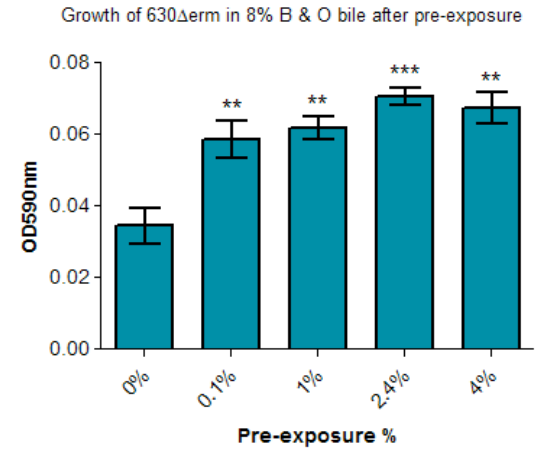
D


Growth of 630Δerm in 2.4% B & O bile after pre-exposure
OD590nm
0.5
0.4
0.3
0.2
0.1
0.0
***
**
0%
0.1%
1%
2.4%
4%
Pre-exposure %

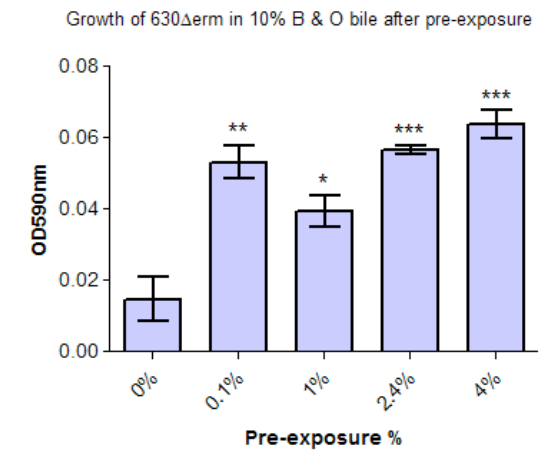
E


Growth of 630Δerm in 4% B & O bile after pre-exposure
OD590nm
0.4
0.3
0.2
0.1
0.0
***
*
*
0%
0.1%
1%
2.4%
4%
Pre-exposure %

F


Growth of 630Δerm in 8% B & O bile after pre-exposure
OD590nm
0.08
0.06
0.04
0.02
0.00
**
**
***
**
0%
0.1%
1%
2.4%
4%
Pre-exposure %

G



H

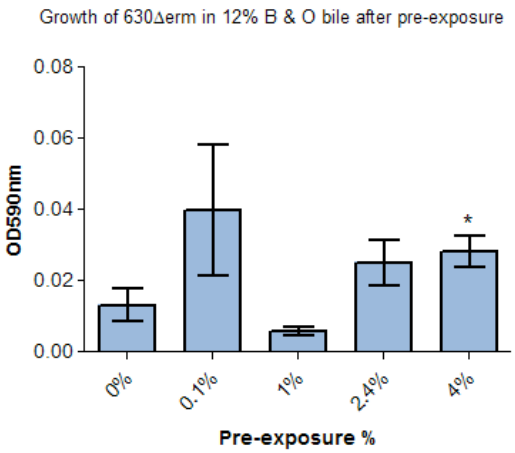


Figure 3.13 Comparison of the growth of 630Δerm in broth containing B & O bile after pre-exposure to B & O bile.

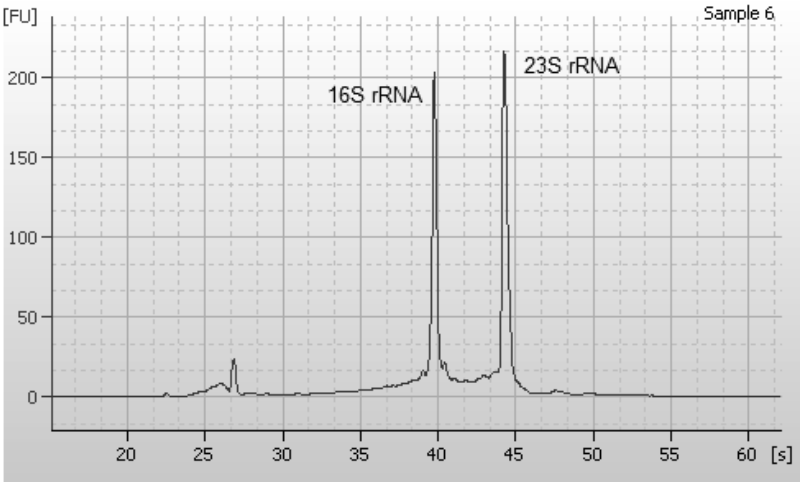
Graph A shows the growth of *C. difficile* strain 630Δerm in BHI broth after a 15 minute pre-exposure to 0.1%, 1%, 2.4% and 4% B & O bile, a non pre-exposed control is also included. Graphs B-H show the growth of 630Δerm in BHI broth containing 0.1 to 12% bile after a 15 minute pre-exposure to 0.1%, 1%, 2.4% and 4% bile. A non pre-exposed control is included for each condition. Growth was measured by optical density at 590 nm. Bidirectional error bars indicate the standard deviation resulting from three independent experiments. Significance is denoted by asterisks (* = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$).

Three experiments were performed, i) Long shock - bacteria were exposed to 0.1% or 2.4% bile for 15 minutes, ii) High concentration - bacteria were exposed to 2.4% or 4% bile for 15 minutes, iii) Short shock - bacteria were exposed to 0.1% or 2.4% bile for 1 minute. A non-exposed control was included to establish whether the presence of bile induced a change in gene expression.

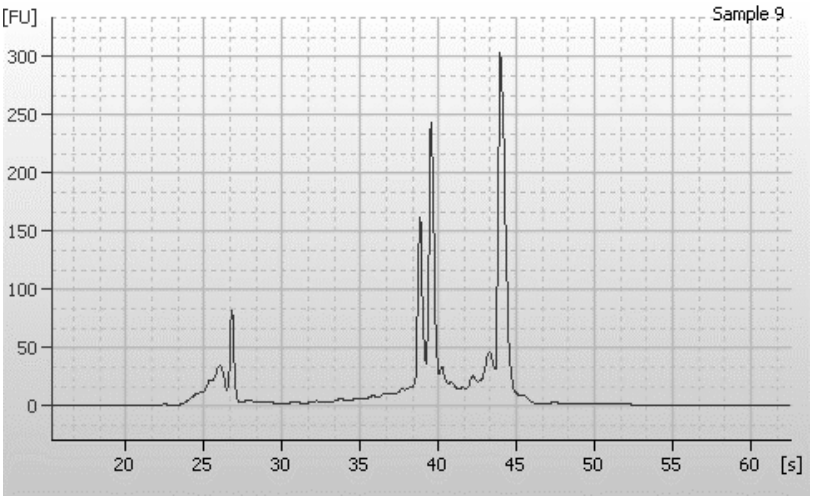
3.5.1.1. RNA quality assessment

After extraction and purification, RNA was assessed spectrophotometrically for purity and concentration with some samples examined for RNA integrity using a bioanalyzer. Figure 3.14 A shows the electropherogram for a control RNA sample (no bile exposure). The 16S and 23S rRNA peaks are sharp and well defined, suggesting good RNA integrity. The RNA from bacteria exposed to bile gave duel peaks for either the 16S rRNA (0.1% exposure samples only) (Figure 3.14 B) or both the 16S and 23S rRNA peaks (0.1% and 2.4% exposure samples) (Figure 3.14.C). This was an anomalous result suggesting two similar sized fragments for each rRNA, however the peaks are defined and there was little background. Twin peaks were not observed in any of the unexposed samples.

A


Sample 6
[FU]
16S rRNA
23S rRNA
200
150
100
50
0
20
25
30
35
40
45
50
55
60
[s]

B


Sample 9
[FU]
300
250
200
150
100
50
0
20
25
30
35
40
45
50
55
60
[s]

C

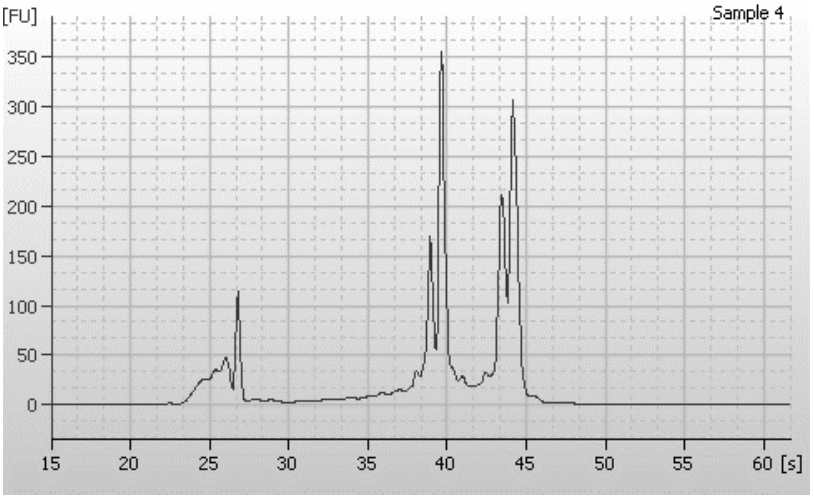

Sample 4
[FU]
350
300
250
200
150
100
50
0
15
20
25
30
35
40
45
50
55
60
[s]

Figure 3.14 Electropherograms of sample RNA.

To evaluate the integrity of the RNA extracted from *C. difficile* cultures grown with or without bile, samples of RNA treated with DNase were assessed using a bioanalyzer (Agilent Technologies). Three examples of the electropherograms are shown, representing each of the different growth conditions. Figure A represents an average profile generated by control RNA (unexposed). The rRNA ratio (23S vs. 16S) = 1.1 generating sharp defined peaks for each rRNA, 5S RNA was also visible (peak between 25 and 30 seconds). Figure B represents an average profile generated by RNA from test samples exposed to 0.1% bile for 15 minutes. The rRNA ratio (23S vs. 16S) = 1.7. A distinct double peak was seen for 16S rRNA and a single well defined peak is seen for 23S rRNA. A larger peak is seen for 5S RNA compared to the control sample. Figure C represents an average profile generated by RNA from test samples exposed to 2.4% bile for 15 minutes. The rRNA ratio (23S vs. 16S) = 0.8. A distinct double peak is seen for 16S rRNA and for 23S rRNA. A larger peak is again seen for 5S RNA compared to the control sample.

3.5.1.2. Experiment 1 – Long exposure

Analysis of the transcriptional response of *C. difficile* to bile stress has not been previously conducted; therefore the experimental design was based on preliminary data obtained in the bile assays and by assessment of similar experiments in other organisms [283, 314]. In the initial bile shock experiment, mid-log phase bacteria were exposed to two bile concentrations 0.1% and 2.4% for 15 minutes. 0.1% was chosen due to its relatively small inhibitory effect on growth, 2.4% was chosen as the bacteria were able to grow to approximately half the optical density of unexposed cultures (Figure 3.10). These levels are physiologically relevant as bile concentrations in the human intestine range from 0.2-2% [315]. Test and control samples from triplicate experiments each with three biological replicates were labelled and hybridised to a *C. difficile* 630 microarray and quantification of gene expression was determined *in silico*.

The data were normalised to the control RNA data from each experiment to control data of another, minimizing variability and allowing conclusions to be drawn from the different experiments. After normalisation, each gene from the unexposed control RNA should give a normalisation value of 1 (Figure 3.15). Figure 3.15 shows the majority of the genes to be tightly arranged around the central line suggesting reproducibility.

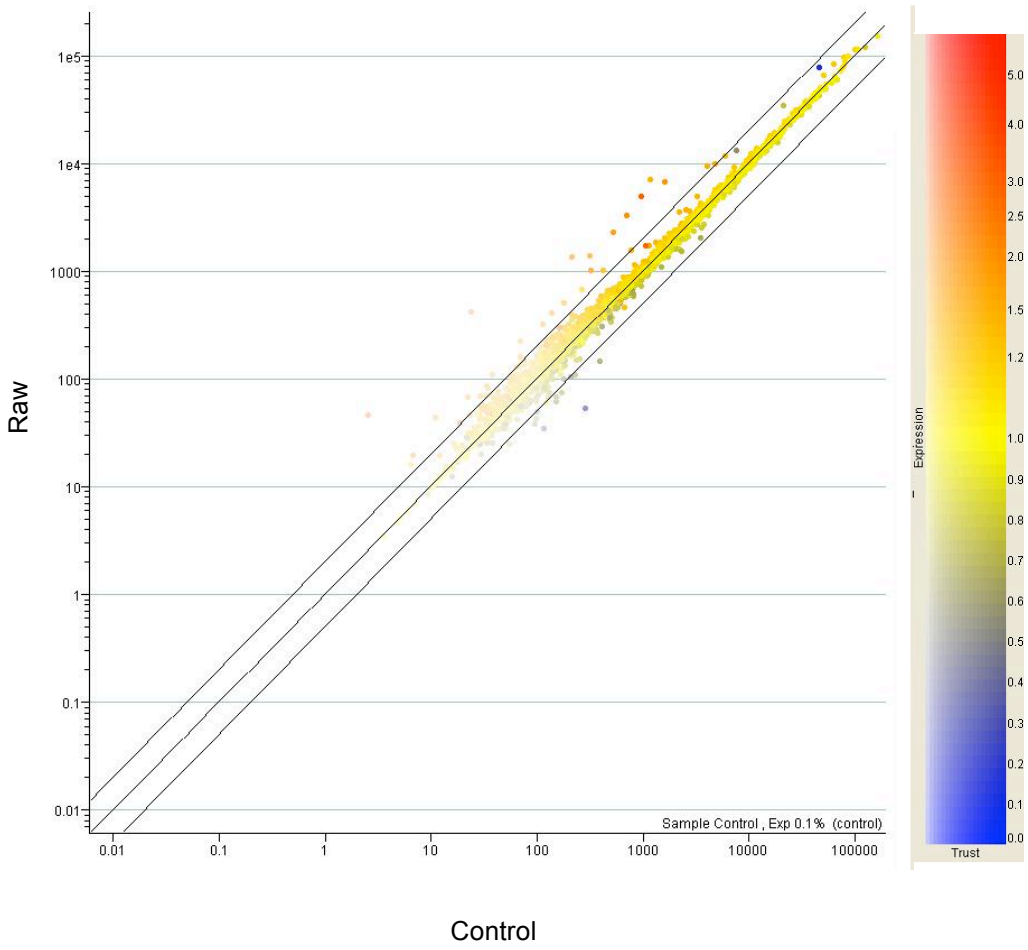


Figure 3.15 Comparison of fluorescence of control verses control RNA samples.

Normalised fluorescence generated by control RNA from experiment one is compared to that obtained in experiment two to assess reproducibility between the experiments. Each data point represents a spot on the microarray that is specific for a unique gene. The intensity of the colour of the data point gives an indication of the reliability of the data (based on fluorescence intensity), the points are colour coded according to their expression in comparison to the control data point with red representing increased expression and blue representing decreased expression. In this plot almost all of the data points are tightly arranged around the central line (x=y) suggesting expression of the gene represented by that data point was the same in each experiment.

Robust statistical assessment of the changes in gene expression in control verses test samples in each experiment failed to reveal any significant changes in the level of expression in bile exposed samples. Presented graphically, it is apparent that the majority of genes did not show differential expression under the test conditions compared to the control i.e. data points lie along the central midline (Figure 3.16). There are however, a number of genes whose expression appears to have changed, these genes are shown as data points positioned away from the midline indicating increased expression (red) or decreased expression (blue) compared to the control sample. The tramlines represent a 2-fold change in either direction. The intensity of the colour conveys the confidence with which one can conclude that the expression of this gene was different to that of the control. Several genes appear to show differential expression in the 0.1% test sample, however this was not supported when scrutinised statistically using the Benjamini-Hochberg false discovery rate multiple testing correction (Figure 3.16). A similar situation was true for the 2.4% test sample; although there was an increase in the number of genes apparently showing differential expression these were not statistically significant (Figure 3.17). When represented in a line diagram the differences between the control and test samples from the two experimental conditions can be easily seen (Figure 3.18). Few differences were observed in the expression of genes after exposure to 0.1% bile compared to control conditions (horizontal yellow lines). Exposure to 2.4% bile on the other hand revealed a number of genes that showed either an increase in expression (red lines) or a decrease in expression (blue line).

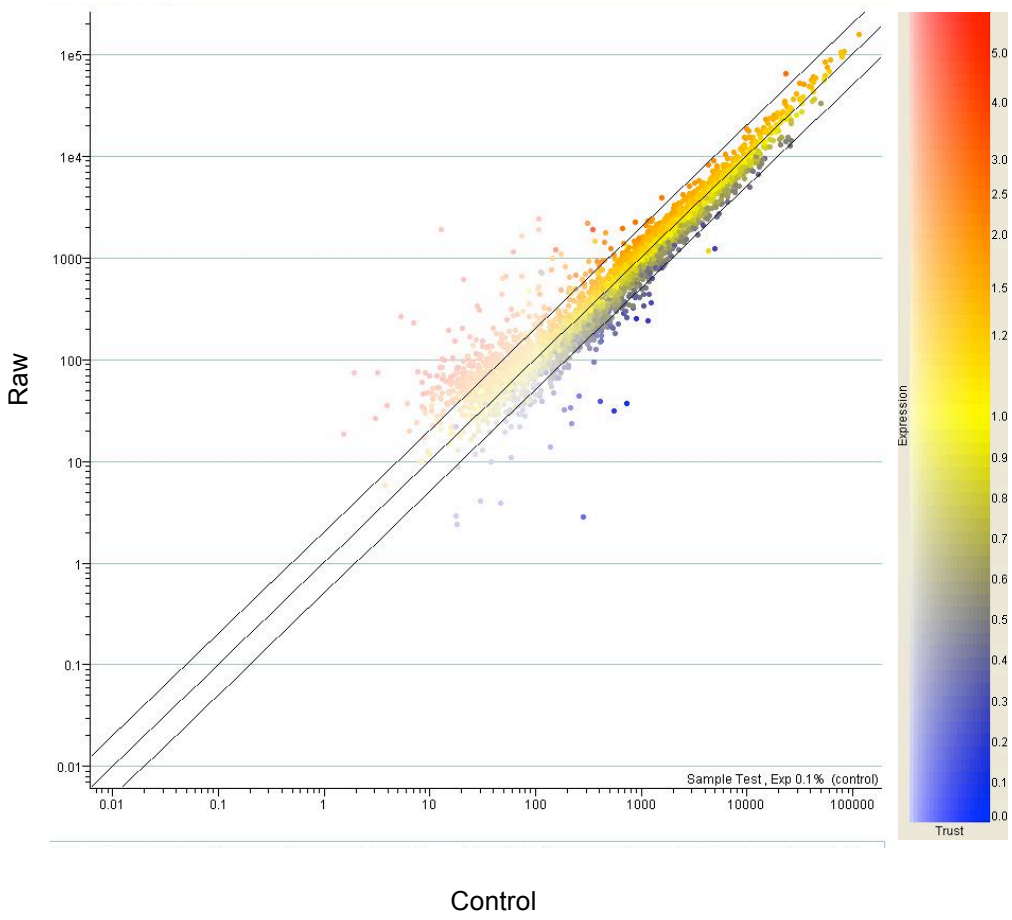


Figure 3.16 Scatter plot of the fluorescence profile of the 0.1% exposed test samples compared to unexposed controls.

The normalised fluorescence generated by hybridisation of the 0.1% test RNA was directly compared to the fluorescence of the unexposed control. Where fluorescence was identical a ratio of 1:1 was given and these samples lie along the x=y central line. Where differences in the fluorescence exist in the test sample spots are plotted off this line. The tramlines represent a 2-fold deviation from the x=y line thus, spots outside these lines represent genes whose expression was more than 2-fold increased (red spots) or decreased (blue spots). The intensity of the colour of each spot also reveals the predicted confidence for the result generated; high intensity suggests a reliable result. A small number of spots are found outside the 2-fold difference area, statistical analysis of these genes will confirm whether these differences are significant.

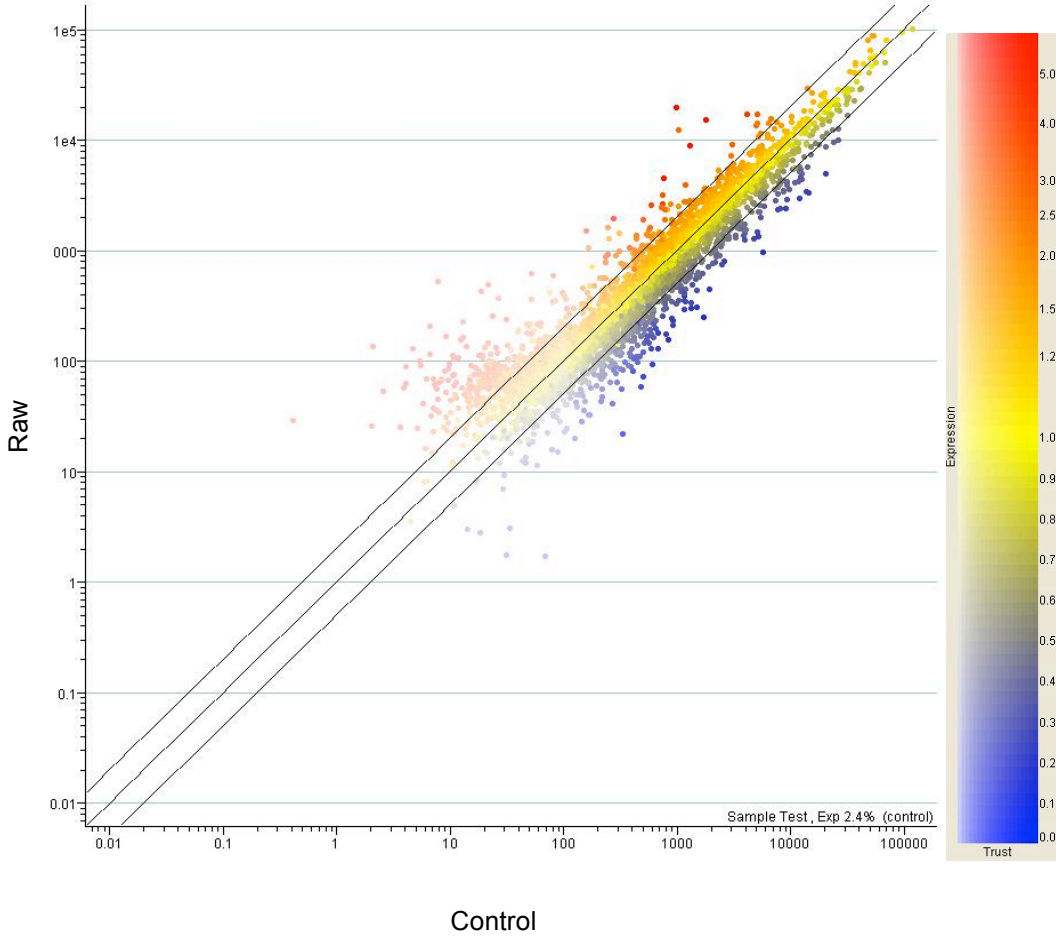


Figure 3.17 Scatter plot of the fluorescence profile of the 2.4% exposed test samples compared to unexposed controls.

The normalised fluorescence generated by hybridisation of the 2.4% test RNA was directly compared to the fluorescence of the unexposed control. Where fluorescence was identical a ratio of 1:1 was given, these samples lie along the x=y central line. These differences in fluorescence are represented by spots outside this area (see Figure 3.16 for further details). In comparison to the scatter plot generated for the 0.1% test sample, a significant increase in the number of genes whose expression is more than 2-fold different to that of the control was observed. Additionally, a proportion of these spots suggest a high level of reliability.

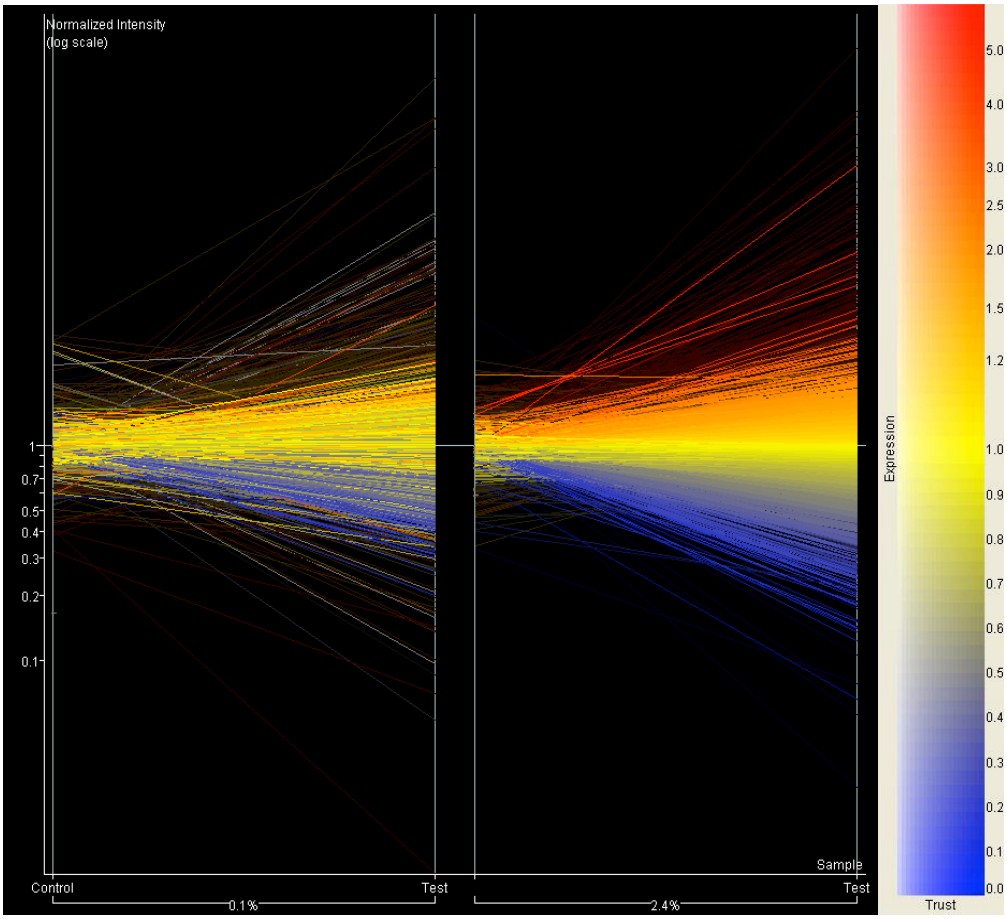


Figure 3.18 Comparison of gene expression levels in control and test samples.

The normalised fluorescence data from each of the 3723 genes was plotted as an individual line on a scale where all genes in the control sample should be 1 (normalised to themselves). After exposure to 0.1 or 2.4% bile the expression level is plotted and the points joined and colour coded. Lines are coloured to indicate the level of expression, yellow indicates no change (=1), red indicates an increase in expression (>1) and blue represents a decrease in expression (<1). In the 0.1% test condition the majority of genes seem to be represented by yellow lines indicating little or no change in expression under the two conditions. The results obtained for the 2.4% test condition show a larger number of changes with several red and blue lines indicating increased and decreased levels of expression respectively. Each of these lines was inspected to analyse the probability and fold-change in expression for each gene under the test and control conditions.

To extract useful information from the microarray data for future experiments the multiple testing correction was omitted from the ANOVA analysis. This revealed a number of genes whose expression was notably different in the test samples compared to the control samples, but which did not fulfil the required statistical stringency (ANOVA with Benjamini-Hochberg false discovery rate multiple testing correction) (see appendix section 9.2). This selection of genes was scrutinised to identify those which might show differential expression based on four parameters i) a relatively significant level of normalised fluorescence i.e. >2000 units (florescence is measured from 0 to 65535), ii) a DNA control normalisation value of between 0.9 and 1.1, iii) a t-test *P* value of less than 0.025, and iv) a change in expression of at least +/- 1.4-fold.

Analysis of samples exposed to 0.1% bile

Table 3.4 shows a selection of genes whose expression has changed in the 0.1% exposed test samples compared to the control samples. Although these genes did not pass the stringent statistical parameters required to support their differential expression, they provide useful information regarding the response of the bacteria to bile challenge. Samples with a positive fold change represent genes whose expression had increased, while negative values represent genes that showed a reduction in expression compared to the control, a summary of the results is shown in Table 3.5. Three genes CD1661 (transposase-like protein B), CD1305 (DNA polymerase III PolC-type) and CD2666 (PTS system, glucose-specific IIa component) showed the most significant increase in expression (1.7-1.8-fold). Thre most significant

decrese in expression was seen in gene CD3089 (PTS system, IIabc component) in the 0.1% test sample compared to the control. All the genes are grouped by their classification based on Cluster of Orthologous Groups (COG) analysis of the protein domains found in the protein sequence [239].

Functional category	Code	Gene	Fold change	*P* value	Product	Additional category/information
Energy production and conversion	C	CD0716	-1.4	0.016	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	*cooS*
Amino acid transport and metabolism	E	CD3670	1.4	0.004	putative selenocysteine lyase	
		CD0241	-1.4	0.023	phosphoserine phosphatase	
Nucleotide transport and metabolism	F	CD2335	1.5	0.002	inosine-5'-monophosphate dehydrogenase	*guaB*
		CD2438	-1.4	0.021	cytidine deaminase	*cdd*
		CD0186	-1.6	0.006	dihydroorotate dehydrogenase, catalytic subunit	*pyrD*
Carbohydrate transport and metabolism	G	**CD2666**	**1.7**	**0.022**	**PTS system, glucose-specific IIa component**	*crr*
		CD1191	1.5	0.004	putative fructose-1,6-bisphosphatase	*fbp*
		CD2322	1.5	0.013	transketolase	*tkt*
		CD3089	**-1.8**	**0.025**	**PTS system, IIabc component**	
Coenzyme transport and metabolism	H	CD1709	-1.4	0.001	putative molybdopterin-guanine biosynthesis protein	
		CD3421	-1.5	0.013	porphobilinogen deaminase	*hemC*
Transcription	K	CD0670	1.6	0.002	putative regulatory protein	
Replication, recombination and repair	L	**CD1661**	**1.8**	**0.012**	**transposase-like protein B**	*tlpB*

		CD1305	**1.7**	**0.024**	**DNA polymerase III PolC-type**	*dnaF*
Cell wall/membrane/envelope biogenesis	M	CD3488	-1.5	0.014	utp--glucose-1-phosphate uridylyltransferase (general stress protein 33)	*gtaB*
		CD2796b	-1.5	<0.001	putative cell surface protein	
Cell motility intracellular trafficking, secretion and vesicular transport	NU	CD0251	-1.5	0.015	flagellum-specific ATP synthase	*fliI*
		CD0247	-1.6	0.023	flagellar hook-basal body complex protein	*fliE*
Posttranslational modification, protein turnover, chaperones	O	CD3283	-1.6	0.02	putative pyruvate formate-lyase 3 activating enzyme	*pflE*
Inorganic ion transport and metabolism	P	CD3403	-1.4	0.009	putative formate/nitrite transporter	
General function prediction only	R	CD1176	1.4	0.006	conserved hypothetical protein	
		CD2173	-1.4	0.04	putative peptidase	
		CD2561	-1.4	0.012	putative phosphatase	
		CD2181	-1.6	0.017	putative aromatic compounds hydrolase	*nox*
Function unknown	S	CD1404	1.5	0.006	putative oligopeptide transporter	
		CD1645	-1.4	0.006	putative membrane protein	

No COG information	CD0939	1.4	0.017	hypothetical phage protein
	CD1768	1.6	0.017	putative membrane protein
	CD1219	-1.6	0.018	conserved hypothetical protein

Table 3.4 List of genes showing differential expression after exposure to 0.1% B & O bile.

Genes whose expression was altered in the test samples are listed in the table with their predicted function, the fold change in expression and *P* value. Selected genes were derived from a one-way ANOVA test without a multiple testing correction using GeneSpring and classified based on COG domains in the protein sequence using the JGI Integrated Microbial Genomics database. Where multiple COG domains were present this information is also included in the additional information. Genes in bold are referred to in the text.

Gene classification	No. of genes differentially expressed during bile shock:	
	Overexpressed	Underexpressed
Energy production and conversion	0	1
Amino acid transport and metabolism	1	1
Nucleotide transport and metabolism	1	2
Carbohydrate transport and metabolism	3	1
Coenzyme transport and metabolism	0	2
Transcription	1	0
Replication, recombination and repair	2	0
Cell wall/membrane/envelope biogenesis	0	2
Cell motility intracellular trafficking, secretion and vesicular transport	0	2
Posttranslational modification, protein turnover, chaperones	0	1
Inorganic ion transport and metabolism	0	1
General function prediction only	1	3
Function unknown	1	1
No COG information	2	1
Total	**12**	**18**

Table 3.5 Classes and frequency of genes differentially expressed during 15 minute exposure to 0.1% B & O bile.

Genes were classified based on COG domains found in the protein sequence using the JGI Integrated Microbial Genomics database. Where multiple COG domains were present the gene was categorised in the predominant class.

Analysis of samples exposed to 2.4% bile

Table 3.6 shows a selection of genes whose expression has increased or decreased in the 2.4% bile exposed test sample compared to the control, a summary of this information is shown in Table 3.7. The most significant increase was seen for gene CD1693 annotated as a putative dinitrogenase iron-molybdenum cofactor, expression appears to have increased by 3.9-fold, closely followed by CD2444, a putative sigma 54 modulation protein (3.2-fold increase) and CD1692, an ArsR-family transcriptional regulator (3-fold increase). The most significant decrease in expression was observed for genes CD2366 (putative exported protein) (4.2-fold), CD0231 (putative flagellar hook-associated protein) (4.1-fold) and CD0252 (flagellar protein) (4-fold). A greater number of genes showed potentially significant differences in expression with many showing a greater fold change in expression than those seen in the 0.1% samples (after removal of the multiple testing correction). These differences can be visualised on a line diagram (Figure 3.19). Here, only the genes selected after statistical analysis of the data obtained from the 2.4% bile exposed samples are shown allowing visual assessment of the expression of the comparative genes in the 0.1% test sample. Comparison of both gene lists from the two experiments revealed five genes whose expression was altered in both experiments (CD0241, CD2796, CD0247, CD2181 & CD1404). Gene CD1404, annotated as a putative oligopeptide transporter which may act by facilitating transport of small oligopeptides into the bacterium, was up-regulated in the 0.1% and 2.4% test samples by 1.5 and 1.4-fold respectively with highly significant *P* values. Several motility associated genes showed a change in expression in

the 2.4% test samples (18 in total) and in all cases this was a reduction in expression compared to the control (from -1.5 to 4-fold). Genes classed into the COG category 'posttranslational modification, protein turnover and chaperones' contained 8 genes whose expression was upregulated (between 1.4-2.7-fold), additionally several putative membrane associated protein coding genes were also differentially expressed. Finally, there was no evidence that genes annotated as being associated with bile resistance, specifically CD0065, CD3215 and CD3216, were differentially expressed under these experimental conditions [6].

Functional category	Code	Gene	Fold change	*P* value	Product	Additional category/information
Energy production and conversion	**C**	CD0117	2.2	0.001	putative subunit of oxidoreductase	
		CD1157	1.6	<0.001	putative anaerobic nitric oxide reductase flavorubredoxin	*norV*
		CD3467	1.5	0.007	ATP synthase epsilon chain (partial)	*atpC*
		CD0713	1.4	0.012	putative adenosylcobamide-dependent radical SAM superfamily protein	
		CD1655	-1.4	0.011	putative Na+/H+ antiporter	
		CD2171	-2.0	0.001	putative Sodium:dicarboxylate symporter	
		CD0394	-2.2	0.013	D-lactate dehydrogenase	*ldhA,* Coenzyme transport and metabolism, General function prediction only
		CD0395	-2.5	0.018	E-cinnamoyl-CoA:R-phenyllactate CoA transferase large subunit	*hadA*
Cell cycle control, cell division, chromosome partitioning	**D**	CD1143	-1.4	0.018	septum formation protein	*maf*
		CD1145	-1.9	0.021	rod shape-determining protein	*mreB2*
Amino acid transport and metabolism	**E**	CD0179	1.7	0.021	NAD-specific glutamate dehydrogenase	*gluD*
		CD2674	1.6	0.009	oligopeptide ABC transporter, permease protein	*appC,* Inorganic ion transport and metabolism
		CD1665	1.4	0.014	cysteine synthase A	*cysK*
		CD1279	-1.5	0.004	cysteine desulfurase	*iscS2*
		CD0853	-1.7	0.002	oligopeptide ABC transporter, permease protein	*oppB,* Inorganic ion transport and metabolism
		CD2172	-1.7	0.021	probable amino-acid ABC transporter, ATP-binding protein	*cbiO*

		CD0241	-2.5	<0.001	phosphoserine phosphatase	
		CD0855	-2.7	<0.001	oligopeptide ABC transporter, substrate-binding lipoprotein	*oppA*
		CD0888	-3.1	0.003	arginine decarboxylase	*speA*
Nucleotide transport and metabolism	**F**	CD1261	1.6	0.006	putative ribonucleotide-diphosphate reductase	
		CD2447	1.5	0.01	histidine triad nucleotide-binding protein	
		CD0220	-1.9	0.019	amidophosphoribosyltransferase	*purF,* Carbohydrate transport and metabolism, General function prediction only
		CD0224	-1.9	0.025	phosphoribosylamine--glycine ligase	*purD*
Carbohydrate transport and metabolism	**G**	CD0582	2	<0.001	putative PEP-utilising kinase	
		CD3091	-2.5	0.005	trehalose-6-phosphate hydrolase	*treA*
Coenzyme transport and metabolism	**H**	CD0243	-2.8	0.013	conserved hypothetical protein	
Translation, ribosomal structure and biogenesis	**J**	**CD2444**	**3.2**	**0.004**	**putative sigma 54 modulation protein**	
		CD0103	1.7	0.01	tRNA pseudouridine synthase 1	*truA1*
		CD1282	1.7	0.01	putative alanyl-tRNA synthetase	*alaS*
		CD2432	1.6	0.025	glycyl-tRNA synthetase beta chain	*glyS*
		CD1161	1.6	0.001	50S ribosomal protein L21	*rplU*

Transcription	**K**	**CD1692**	**3**	**<0.001**	**ArsR-family transcriptional regulator**	
Cell wall/membrane/envelope biogenesis	**M**	CD2789	-1.4	0.007	cell surface protein	*cwp66*
		CD0242	-1.9	0.003	conserved hypothetical protein	
		CD0226	-2.1	<0.001	putative transglycosylase	
		CD2796	-2.3	0.001	putative cell surface protein	
		CD2768	-3.3	<0.001	putative cell-wall hydrolase	
Cell motility	**N**	CD0270	-1.6	0.013	putative flagellar motor switch protein	*fliM*
		CD0256	-1.9	0.005	chemotaxis protein	*motA*
		CD0254	-2.0	0.019	putative basal-body rod modification protein	*flgD*
		CD0269	-2.1	0.001	putative flagellar basal-body rod protein	
		CD0249	-2.3	0.019	flagellar motor switch protein	*fliG*
		CD0268	-2.3	0.001	flagellar basal-body rod protein	*flgG*
		CD0232	-2.6	0.007	flagellar hook-associated protein	*flgL*
		CD0253	-2.7	0.015	putative flagellar hook-length control protein	*fliK*
		CD0245	-3.2	0.007	flagellar basal-body rod protein	*flgB*
		CD0264	-3.9	<0.001	signal recognition particle complex, GTP-binding subunit	*flhF*
		CD0231	**-4.1**	**<0.001**	**putative flagellar hook-associated protein**	*flgK*

Cell motility intracellular trafficking, secretion and vesicular transport	**NU**	CD0262	-1.5	0.009	flagellar export protein	*flhB*
		CD0258	-1.7	0.024	flagellar basal body-associated protein	*fliL*
		CD0250	-2.3	0.014	flagellar assembly protein	*fliH*
		CD0246	-2.3	0.009	flagellar basal-body rod protein	*flgC*
		CD0247	-2.6	<0.001	flagellar hook-basal body complex protein	*fliE*
		CD0263	-2.7	0.002	flagellar export protein	*flhA*
		CD0252	**-4.0**	**0.007**	**flagellar protein**	***fliJ***
Posttranslational modification, protein turnover, chaperones	**O**	**CD0881**	**2.7**	**0.003**	**putative membrane protein**	
		CD3284	2.3	0.004	probable protease	
		CD1828	1.9	0.007	cell division protein	*ftsH1*
		CD0194	1.9	<0.001	60 kDa chaperonin	*groEL*
		CD3157	1.8	0.005	SsrA-binding protein	*smpB*
		CD1690	1.6	0.011	thioredoxin	*trxA1*
		CD2461	1.4	0.022	chaperone protein	*dnaK*
		CD0703	1.4	0.008	protease	
Inorganic ion transport and metabolism	**P**	CD3377	1.8	0.023	magnesium transporting ATPase, P-type 1	*mgtA*
		CD1695	1.5	0.008	putative arsenical pump membrane protein	
		CD2365	-1.8	<0.001	putative sulfonate ABC transporter, solute-binding lipoprotein	

		CD0324	-2.1	0.002	putative cobalt transport protein	*cbiM*
		CD0327	-2.3	0.021	cobalt ABC transporter, ATP-binding protein	
		CD2361	-2.5	0.013	putative aliphatic sulfonate ABC transporter, ATP-binding protein	
Intracellular trafficking, secretion, and vesicular transport	**U**	CD0059	-1.5	0.004	preprotein translocase SecE subunit	*secE*
Defence Mechanisms	**V**	CD2149	1.7	0.008	putative exported protein	
		CD1469	-2.0	0.002	putative penicillin-binding protein	
General function prediction only	**R**	CD0730	2.3	0.001	putative iron-sulfur protein	
		CD2540	1.6	0.022	NADH oxidase	
		CD3041	1.4	0.009	proline iminopeptidase	
		CD2181	-1.7	0.009	putative aromatic compounds hydrolase	*nox*
		CD0240	-1.8	0.009	putative glycosyltransferase	*pepI*
		CD0878	-1.8	0.005	ABC transporter, permease protein	
		CD2367	-2.2	<0.001	putative permease	
		CD0893	-2.5	0.003	iron-dependent hydrogenase	
Function unknown	**S**	**CD1693**	**3.9**	**0.008**	**putative dinitrogenase iron-molybdenum cofactor**	
		CD1590	2	<0.001	putative membrane protein	
		CD1404	1.4	0.007	putative oligopeptide transporter	
		CD1215	-1.4	0.016	putative pyrophosphokinase	
		CD2531	-1.5	0.005	putative membrane protein	
		CD0227	-2.3	0.006	conserved hypothetical protein	
		CD3100	-2.7	0.002	putative C4-dicarboxylate anaerobic carrier	
No COG information		CD0910	1.7	0.009	putative phage protein	

		CD3327	1.6	0.013	conjugative transposon conserved hypothetical protein
		CD0735	1.5	0.01	hypothetical protein
		CD3674	1.5	0.009	methyltransferase (putative glucose inhibited division protein B) (pseudogene)
		CD2542	1.5	0.007	putative exported protein
		CD3668	-1.5	0.001	putative transcriptional regulator
		CD1720	-1.6	0.019	putative hydantoinase
		CD2063	-1.6	0.002	putative membrane protein
		CD2127	-1.8	0.01	putative membrane protein
		CD2062	-1.8	<0.001	conserved hypothetical protein
		CD2127	-1.8	0.01	putative membrane protein
		CD1147	-1.9	0.017	putative membrane protein
		CD2363	-2.0	0.02	hypothetical protein
		CD2362	-2.1	0.006	putative aliphatic sulfonates ABC transporter, permease protein (pseudogene)
		CD3145	-3.8	0.002	putative serine-aspartate-rich surface anchored fibrinogen-binding protein
		CD2366	**-4.2**	**<0.001**	**putative exported protein**

Table 3.6 List of genes showing differential expression after exposure to 2.4% B & O bile.

Genes whose expression was altered in the test samples are listed in the table with their predicted function, the fold change in expression and P value. Selected genes were derived from a one-way ANOVA test without a multiple testing correction using GeneSpring and classified based on COG domains in the protein sequence using the JGI Integrated Microbial Genomics database. Where multiple COG domains were present this information is also included in the additional information. Genes in bold are mentioned in the text.

Gene classification	No. of genes during bile shock:	
	Overexpressed	Underexpressed
Energy production and conversion	4	3
Cell cycle control, cell division, chromosome partitioning	0	2
Amino acid transport and metabolism	3	6
Nucleotide transport and metabolism	2	2
Carbohydrate transport and metabolism	1	1
Coenzyme transport and metabolism	0	1
Translation, ribosomal structure and biogenesis	5	0
Transcription	1	0
Cell wall/membrane/envelope biogenesis	0	5
Cell motility	0	11
Cell motility intracellular trafficking, secretion and vesicular transport	0	7
Posttranslational modification, protein turnover, chaperones	8	0
Inorganic ion transport and metabolism	2	4
Intracellular trafficking, secretion, and vesicular transport	0	1
Defence mechanisms	1	1
General function prediction only	3	5
Function unknown	3	4
No COG information	5	11
Total	**38**	**65**

Table 3.7 Classes and frequency of genes differentially expressed during 15 minute exposure to 2.4% B & O bile.

Genes were classified based on COG domains found in the protein sequence using the JGI Integrated Microbial Genomics database. Where multiple COG domains were present the gene was categorised in the primary allocated class.

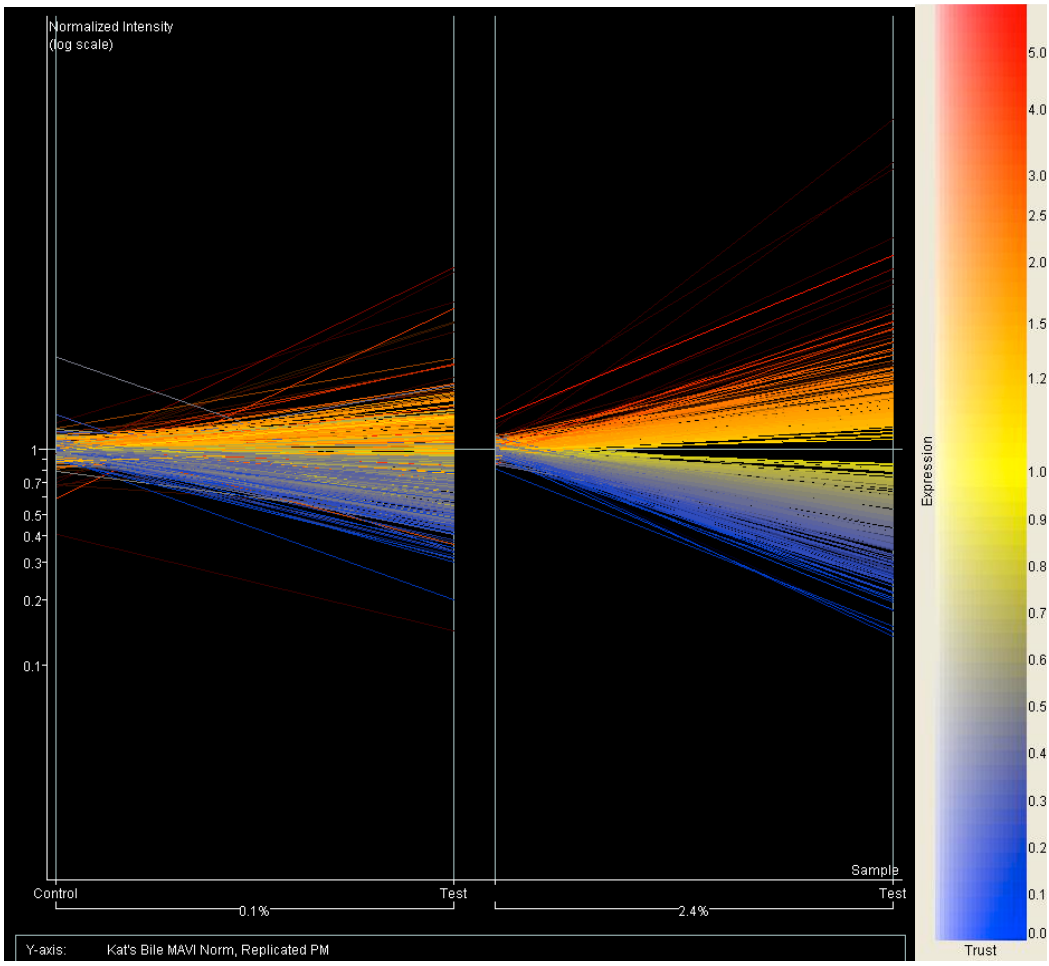


Figure 3.19 Comparison of 457 differentially expressed genes in 0.1% and 2.4% bile exposed samples.

Statistical analysis of the 2.4% bile exposed samples produced a list of 457 genes that were differentially expressed in the test sample with a *P* value of ≤0.05. Each gene is represented by a coloured line according to 2.4% bile test sample. The change in expression can be seen by a deviation of the line from horizontal at point 1. Genes expressed in the control should start at 1 (after normalisation) on the y axis. For comparison the same genes are shown for 0.1%, clear differences can be seen between the groups.

3.5.1.3. Experiment 2 – High concentration exposure

In the second experiment bacteria were exposed to 2.4% bile and to 4% bile. After hybridisation, the data were normalised and subjected to statistical analysis. No statistically significant differences were seen in the expression of genes in the unexposed cultures compared to the bile exposed cultures. Therefore, no further investigation of this data was undertaken.

3.5.1.4. Experiment 3 – Short exposure

In this final experiment, bacteria were subjected to a short one minute shock to establish whether an immediate stress response could be observed after addition of B & O bile. As with the previous experiments, comparison of the expression profiles revealed no statistically significant differences between the test and control samples. Removal of the multiple testing correction allowed non-significant differences to be scrutinised. Changes in expression of small number of genes were identified between the test and control samples. However, several of these genes showed poor levels of fluorescence in both samples. Furthermore, this experiment was performed only once therefore reliable conclusions cannot be drawn.

3.6. Discussion

Bile undoubtedly plays an important role in defending the GIT from invading pathogenic microorganisms. Its potent digestive activity can have dramatic effects on bacterial membranes, particularly those of Gram-positive bacteria, leading to cell lysis. To overcome these potentially catastrophic actions, many bacteria have evolved mechanisms to evade bile by adapting their physiology and gene expression, through for example, membrane modification and the utilisation of efflux pumps and porins [245].

For *C. difficile*, the role of bile in the pathogenicity of this organism has not yet been fully explored. As a coloniser of the GIT it is understood that *C. difficile* will have developed mechanisms to combat or tolerate the destructive effects of bile salts, particularly those found in the small intestine. The aetiological agent of CDI is the spore, this highly resistant cell state confers a high level of protection from the acidic environment of the stomach and from bile found in the upper GIT. It has been suggested that bile also plays a key role in sporulation and germination, acting as an environmental cue indicating the position of *C. difficile* in the GIT and by acting as a host-derived signal to initiate germination [316]. Therefore, there are two key questions to be addressed i) what is the tolerance of this bacterium to bile and ii) by what mechanism(s) is this tolerance conferred?

To address the first question, simple assays were developed to determine the tolerance of two pathogenic strains of *C. difficile* to B & O bile and porcine

bile in solid and liquid medium. Strains 630Δerm and R20291 were both able to grow in bile concentrations of 8% vol/vol of B & O bile in solid and liquid medium. This is a very high level compared to that found in the GIT, generally between 0.2-2%, although it is unlikely that B & O bile is representative of human bile [245, 315]. Assays investigating porcine bile, which better reflects the composition of human bile, could not be completed due to problems dissolving the commercially available preparation to generate a uniform solution [245].

The growth of both strains was significantly more inhibited in low bile concentration (0.3-1.5%). As bile concentration increased, growth also increased, leading to growth levels that were not significantly different to the growth observed in cultures grown without bile. This trend continued until a maximum concentration of bile was reached which had a significant effect on growth, for 630Δerm this occurred at 5% bile concentration and for R20291 at 7.5%. Further increases in bile concentration resulted in significant reduction in growth compared to cultures grown in the absence of bile. Overall, the data suggests that at low bile concentration, protective mechanism(s) that enable the bacteria to survive in this potent detergent are either not activated, or only partially activated. As the concentration of bile increases, defence mechanism(s) are likely to be activated, protecting the bacterium to facilitate growth. At a certain concentration, observed in this study as 5% and 7.5% for 630Δerm and R20291 respectively, the defence mechanism(s) employed by the bacterium become overwhelmed and unable to prevent the effects of bile on the bacterium. One could suggest that these mechanisms may

include efflux pumps, changes to membrane architecture or by other, as yet undefined, mechanism [175]. It may also be possible that increasing bile concentrations signal the bacterium to end vegetative growth and begin sporulation, thereby protecting it from the unfavourable environmental conditions. Bile is known to act as a signalling molecule in the body and it is therefore possible that it may also act as a signalling molecule to bacteria [317]. Levels of bile acids in the small intestine are very high and are thought to aid the suppression of colonisation by microorganisms [245, 318]. Thus, high bile concentrations could act as a signal to the bacterium that the environment is unfavourable and it should not undergo germination or that it should begin the process of sporulation.

It is also interesting to note the relatively higher bile tolerance shown by the epidemic strain R20291 compared to 630Δerm. R20291, which is a clinical isolate has, in our laboratory, undergone very minimal subculture, whereas 630Δerm was generated by multiple rounds of subculture [131]. It is possible that this difference in tolerance relates to the history of these isolates, for example due to a mutation, however it could also suggest an additional feature of R20291 which has aided its epidemicity [55]. This could be further explored by looking at additional clinical and laboratory strains, or by comparative sequence analysis of R20291 and 630Δerm, unfortunately the genome sequence is not available for the latter strain.

Sorg and Sonenshein (2008) have suggested that the loss of indigenous microflora in the large intestine may have been advantageous for *C. difficile*

colonisation [169]. These indigenous bacteria are responsible for conversion of primary bile salts (cholate derivatives) to secondary bile salts (deoxycholate) which are inhibitory to *C. difficile* vegetative cells [169]. Thus eradication of these bacteria by antibiotic treatment leads to a reduction in the concentration of secondary bile salts allowing *C. difficile* to colonise the colon. Therefore the reported ability of R20291 to withstand higher bile salt concentrations may allow colonisation of the colon where some indigenous gut flora and consequently secondary bile salts, still remain. However, it is difficult to fully support this hypothesis as the levels at which the most significant differences occurred are unlikely to be physiologically relevant.

The main limitation of the *in vitro* bile assays is that they cannot accurately represent the conditions found *in vivo*. The environment of the GIT is influenced by a number of factors including the type and number of indigenous and pathogenic flora, diet and a number of physiological parameters. Additionally, the composition of bile itself can vary between individuals due to genetic or dietary factors [252, 315]. Thus, as it is almost impossible to recreate such an environment *in vitro* and conclusions drawn from these assays should be considered as clues upon which further investigations can be based. Secondly, as the composition of the commercially available bile mixtures is unknown, it may not represent the mixture found *in vivo*. Ideally, the composition should represent the physiological mix of conjugated and deconjugated bile acids found in the colon. It seems likely that the dehydrated bile used in these assays is extracted directly from the animals gallbladder meaning that bacterially

mediated deconjugation will not have occurred. Oxgall and other bovine bile preparation have been used in a number of studies to examine the tolerance of Gram-positive organisms to bile, and was therefore chosen as an initial test medium [270, 283, 319].

The inocula used were late-log to stationary phase bacteria and therefore there may have been a mix of vegetative cells and spores. Vegetative cells may be lysed immediately by bile due to its potent detergent effects. Additionally, it has been postulated that the cell membrane of stationary phase bacteria shows higher resistance to bile associated damage, therefore it will be important to examine bacteria in different growth phases [245, 320]. Finally, bile is though to act as a germinant for *C. difficile* spores, if spores are present bile may trigger germination during which the bacterium can adapt to the presence of bile.

The growth of strain 630Δerm was monitored over time in bile containing media. The results showed that bile significantly reduced growth of the bacterium in a dose-dependent manner compared to the unexposed control. Bile has a number of potent activities affecting membrane integrity, protein confirmation, DNA and RNA integrity, and a number of, other as yet, undefined effects [263, 264]. The reduction in growth rate may be associated with the stress response, which was further investigated by the transcriptomics study. Damage to DNA and RNA may also affect central processes in the bacteria, leading to knock on effects to protein synthesis, metabolism and signalling.

Adaptation of *L. monocytogenes* to bile salts has been reported by Begley *et al*., (2002) where exposure to a sub lethal concentration of bile provided significant protection against challenge with lethal bile levels [270]. Similar observations have been made in *E. faecalis* [271, 272]. To investigate this trend in *C. difficile*, bacteria were exposed to a range of B & O bile concentrations before inoculation into bile containing growth medium. Both strains showed a significant increase in growth in broth without bile after pre-exposure. It is possible that this increase is a result of enhanced germination of spores present in the initial inocula. The inocula for these experiments was late-log phase cells which may have contained spores, further experiments using mid-log phase cells should provide clarification as to the effect of bile pre-exposure on growth.

For R20291 cultures the most significant effects in growth were seen in cultures grown in 8% and 10% bile. Exposure to all of the pre-exposure concentrations resulted in a significant increase in growth, compared non pre-exposed controls with high probabilities suggesting that this is not a chance result. For 630Δerm cultures pre-exposed to bile showed, in general, increased growth in bile containing broths. These largest differences were seen for pre-exposed cultures grown in 1% bile with an average 2-fold increase in growth compared to non pre-exposed cultures. Again this result was supported by small *P* values indicating that this result did not occur by chance. Large significant increases in growth were also observed for pre-exposed bacteria grown in 8% and 10% bile, however the relative growth of these cultures is small suggesting that this concentration of bile significantly

inhibits growth. The data also suggests that pre-exposure of 630Δerm is required for growth in low bile concentration such as 1%, however this does not appear to be true for R20291.

Taken together the data obtained for R20291 and 630Δerm suggest that pre-exposure has a positive effect on the ability of the bacteria to grow in bile, particularly in high concentrations. This corresponds with results reported for other bacterial species. Flahaut *et al*. (1996) observed cross-adaptation in *E. faecalis*, exposure to very low bile levels (0.08%) for as little as 5 seconds (known as flash adaptation) provided considerable protection against challenge with usually lethal bile levels; similar results have been seen in *L. monocytogenes* and *Bifidobacterium adolescentis* [270, 321, 322]. Both the studies by Flahaut *et al*. (1996) and Begley *et al*. (2002) on *E. faecalis* and *L. monocytogenes* respectively, showed that this adaptation was independent of protein synthesis. Begley *et al*. suggests that adaptation occurs at the membrane; as detergents form mixed micelles with lipids, rapid intercalation of bile with membrane lipids creates a mixed membrane that may prevent further damage by subsequent exposure to these detergents [245, 323]. Further experiments can be undertaken to assess whether a pre-exposure to bile can allow R20291 to grown in previously lethal levels of bile over 12%, as this was not investigated. The use of less complex bile mixture similar to those used in the Flahaut and Begley studies (1:1 mix of cholic acid and deoxycholic acid) and the addition of chloramphenicol may aid the investigation of this adaptation response in *C. difficile*. Finally, it is possible that other stresses may also have an effect on the tolerance of *C. difficile* to

bile by cross-adaptation [245]. Tolerance to bile may be mediated by general stress responses, exposure to stresses such as oxygen, low pH and temperature may lead to adaptations which increase *C. difficile's* tolerance to another stress such as bile. This type of cross-adaptation has been seen in *L. monocytogenes* and *S. typhimurium* [324-326].

The transcriptional response of *C. difficile* to bile challenge has not been investigated prior to this study. Consequently, the parameters for optimal assay conditions are not established. The two main parameters, length of exposure and bile concentration, were investigated and three experiments conducted i) a 15 minute exposure to a medium-high and a low bile concentration, ii) a 15 minute exposure to a high bile concentration, and iii) a short 1 minute exposure to a low and medium-high bile concentration. The exposure time of 15 minutes was chosen on the basis of similar studies with *Lactobacillus reuteri* [283]. The short exposure of 1 minute was used as this represented the minimum possible exposure time taking into account practical limitations. The bile shock was administered when bacteria had reached a mid-log phase growth with a view to repeat the assays with stationary phase cells, time permitting. The bile concentrations chosen were determined from the results of the bile assays and by physiological relevance. 0.1% was chosen for the low level shock as both R20291 and 630Δerm were able to grow to over 60% of their normal growth at this concentration. 2.4% bile was chosen as the medium-high shock concentration as at this value 630Δerm grew to approximately 50% of the optical density of the unexposed culture. Bile concentrations of 0.3 to 0.8%

were not chosen as these concentrations had significant inhibitory effects on growth that may have affected RNA yield. Finally, 4 % was chosen as the high shock concentration, as at this concentration 630Δerm was able to grow to approximately 40% compared to the unexposed culture thus allowing harvest of a sufficient number of bacteria for RNA extraction.

RNA extraction and purification was optimised to allow collection of a high yield of high quality RNA from mid log-phase cells. Assessment of some samples with the bioanalyzer revealed an interesting feature of the RNA from bile exposed cultures where twin peaks were observed for one or both of the rRNAs. One of the noted side effects of bile exposure is that it can cause RNA secondary structure formation, thus it was possible that the two peaks represent rRNA with slightly altered conformation [245]. Heat treatment is recommended to remove such anomalies, thus it is possible that the samples were not sufficiently heated prior to analysis [327]. However, in general the quality of the RNA was found to be acceptable.

Overall, the quality of the hybridisation was good however a number of technical issues did impact upon the results generated. Although these were minor issues, including poor placement of cover slips, scratching of the slide surface and the presence of dust, they may have contributed to a small number of spots failing to pass the rigorous quality control monitoring of the ImaGene software. This could have limited the number of genes that could be compared under the test and control conditions. Additionally, the split-pin method of spotting the PCR products can result in imperfect spotting leading

to rejection of spots during analysis. Such problems can be overcome by performing multiple replicates of samples; unfortunately this was not possible due to time constraints.

Scrutinised intensity data was imported to GeneSpring where data were normalised and interpreted using standard parameters for the BμG@S microarray. Initially, all the data obtained from each experiment were assessed for the presence of statistically significant differences between exposed and unexposed samples, and between the experiments. For each experiment no statistically significant differences were seen between exposed and unexposed samples. This would suggest that there were no differences in gene expression after bacteria were exposed to B & O bile. Further replicate experiments would need to be conducted to support this conclusion.

Removal of the multiple testing correction allowed a number of genes to be highlighted that potentially showed differential expression under bile stress. Although these genes were not considered to be differentially expressed, due to insufficient data, they provide insights into the possible response of *C. difficile* to bile challenge. An initial list of over 300 genes was generated for each bile exposed test sample with $P<0.05$ (after removal of the multiple testing correction). As the statistical parameters were less rigorous inspection of the data was required to identify differences that had a degree of reliability. Five genes were found to be differentially expressed after exposure to both bile concentrations compared to unexposed controls,

interestingly the difference in expression was almost identical in both bile concentrations. Quantitative RT-PCR could be performed to determine whether these genes are indeed differentially expressed. Two of these genes could be involved in the transport of substances into or out of the bacterium; the third is associated with motility. Therefore, targeted disruption of these genes using the ClosTron or by allele exchange could provide evidence of their role in bile tolerance [206].

The data were then assessed based on COG domains found in the protein sequence. By assigning genes to functional categories it may be possible to identify the general responses to bile. The induction of stress proteins is an important component of the bacterial general stress response to a range of stresses such as heat shock, salt stress, starvation and ethanol providing nonspecific protection [328]. In several Gram-positive bacteria, the alternative sigma factor σ^B is responsible for the regulation of a number of stress related processes [329]. *C. difficile*, unlike other sequenced clostridia, appears to have the *sigB* gene (CD0011). This is part of a 5-gene cluster that encodes two membrane proteins of unknown function, an anti-sigma-B factor antagonist (RsbV) and an anti-sigma-B factor (RsbW). One of the membrane proteins (CD0008) contains a Pfam motif (PF06103) that is found in general stress response proteins [6]. None of these genes were identified as being differentially expressed in this study. σ^B is involved in the regulation of *bilEA* and *bilEB* genes of *L. monocytogenes,* which are involved in bile resistance, homologues of these genes exist in *C. difficile* (CD3215 and CD3126 respectively). σ^B is also involved in the germination of *Bacillus cereus*, and

as bile salts have been implicated in germination of *C. difficile* it is not unreasonable to predict that σ^B may be differentially expressed in the presence of bile [6, 330]. Investigation of the microarray results for general stress response proteins failed to show changes in σ^B expression. Similarly a study by Emerson *et al*. (2008) which looked at the transcriptional response of *C. difficile* to a number of environmental stresses including heat, oxidative stress, low and high pH also failed to see upregulation of σ^B [331]. They conclude that this may be due to the duration of the shock (15 minutes) as previous studies of *B. subtilis* showed that σ^B was maximally induced after 3 minutes with significant reduction after 10 minutes [311]. As this study also used a 15 minute exposure time it is possible that upregulation of this gene was missed. Due to a lack of repeats the data generated after the 1 minute shock was insufficient for constructive analysis. In *B. subtilis* the alkali shock response is mediated by the RNA polymerase alternative sigma factor σ^w, however analysis of the *C. difficile* 630 genome sequence showed no close homologue of σ^w [331, 332]. Some similarities between the response of *C. difficile* and *B. subtilis* to alkali shock have been noted including the upregulation of the histidine synthesis pathway and tellurite and tellurium resistance proteins, none of which were altered in this experiment [313, 331, 332].

Stresses including heat, acidity and high osmolarity have been shown to induce expression of the heat-shock protein GroEL (Hsp60) [333]. A study by Hennequin *et al*. (2001) showed that expression of *groEL* mRNA was strongly induced yet not accompanied by high protein levels in *C. difficile* exposed to acidic conditions and high osmolarity [333]. Emerson *et al*. (2008) also showed an upregulation of *groEL* mRNA after heat (9-fold) and acid

shock (5-fold) but not in alkali shock [331]. Expression of *groEL* (CD0194) was upregulated in this study by 1.9-fold (P <0.001), suggesting a general stress response to the presence of 2.4% bile. The redox protein, thioredoxin, involved in the oxidative stress response, also showed a small increase in expression (1.6-fold) in the 2.4% bile exposed sample. Overall, expression of eight genes was upregulated in the category 'Posttranslational modification, protein turnover, chaperones', which would include genes encoding stress response proteins. The differences in expression are relatively low (1.4-2.7-fold) suggesting that 2.4% bile may not be sufficient to induce expression of the stress response in *C. difficile*.

Interestingly several genes associated with motility were found to be down regulated in this study, particularly for the 2.4% test condition. Two loci in the genome of *C. difficile* encode proteins which are flagella-associated, CD0226 to CD0240 and CD0245 to CD0271 [334]. Eighteen of the forty-two genes involved in motility, were shown to be down regulated by an average of 2.5-fold. This was a small fold reduction but as almost half of the genes from these flagella-associated loci show a reduction in expression compared to the unexposed control with an average of P=0.008. This suggests that reduced motility may be a true response of *C. difficile* after exposure to bile. Similar analysis of the response of in *B. cereus* to bile salts also showed a reduction in motility-associated gene expression, *Salmonella* spp. also show reduced expression of flagellar biosynthesis genes and reduced motility in the presence of bile [302, 305]. It is possible that the bile salts present in the B & O bile preparation were acting as location signals to *C. difficile*. It has been

postulated that the presence of deconjugated bile salts in the colon is an indicator of the bacteria's position in the GIT [169]. Biotransformation reactions carried out by the indigenous microflora in the colon deconjugate bile salts arriving from the small intestine; thus this environmental cue could signal that the bacterium is in the correct location reducing motility. *S. typhimurium* has been shown to use bile as an indicator of location in the GIT. Bacteria grown in the presence of bile show a reduction in the expression of invasion gene regulators rendering the bacteria less able to invade epithelial cells [305, 335, 336]. This sensory mechanism prevents the bacterium from initiating invasion when in the intestinal lumen, as the bacterium reaches the mucus layer bile concentrations decrease thus initiating invasion [245, 335].

Finally, differential expression of several putative membrane-associated genes was also identified from the 2.4% bile exposure analysis. Many of these are uncharacterised and it would be interesting to investigate these proteins further to identify whether bile exposure is leading to a change in membrane architecture. Targeted mutagenesis of these genes might lead to measurable changes in bile tolerance providing evidence for their role in the bile stress response. These preliminary studies help to form the foundation upon which other studies can be designed to maximise the generation of informative and reliable data, aiding our understanding of the transcriptional and physiological response of *C. difficile* to bile stress.

3.7. Conclusions & Future work

3.7.1. **Bile tolerance**

The bile assays using B & O bile gave a reproducible pattern of bile tolerance by both R20291 and 630Δerm, with *C. acetobutylicum* showing an inability to tolerate bile at all the concentrations tested on liquid and solid medium. The data suggest a difference in the tolerance of R20291 and 630Δerm, with the possibility that this significant increase in tolerance to bile acids being of benefit to this hyper epidemic strain. It is clear that these experiments touch on the true nature of the involvement of bile in pathogenicity and thus further investigation is required to fully understand the potential impact of these findings. A level of adaptation was also observed for both strains. Exposure to bile ranging from 0.1% to 4% led to significant increases in the ability of the bacteria to grow in higher levels of bile. Further investigation of these results and development of these assays are sure to provide a clearer picture about the adaptation of *C. difficile* cells to bile, and elucidate upon the increased tolerance observed to previously lethal levels. Clearly pre-adaptation may have clinical significance, many factors influence the composition and nature of bile in an individual, which in turn may have an effect on the ability of the bacterium to colonise and thrive in the colon. Thus further understanding of this process may aid design of treatment and management strategies.

3.7.2. **Transcriptomics**

Recent advances have been made in the field of Clostridial gene microarrays including the development of an R20291 microarray (Minton *et al.* University of Nottingham, unpublished). Technological advances have also improved efficiency and reduced cost to allow microarrays to be more accessible. This experiment was a preliminary assessment of whether measurable differences could be observed in *C. difficile* gene expression after exposure to bile stress. Although no significant differences were observed in the expression profile of strain 630Δerm it is possible that with modification of experimental parameters and increased replicates informative data can be generated. The differential tolerance shown by R20291 highlights the importance of comparative studies between this strain and 630Δerm.

Ideally, it was hoped that this would identify pathways and processes associated with the tolerance of bile in other organisms such as efflux pumps, changes to membrane architecture and general stress responses. Although the data provided evidence of downregulation of genes involved in motility and upregulation of genes involved in the stress response, the fold change was small. The results shown here do not provide solid conclusions about the response of *C. difficile* to bile but highlight the need for further investigation of the interaction of the bacterium with bile to further knowledge of its molecular pathogenicity.

Chapter 4

4. Development and investigation of a novel transposon delivery system

4.1. Abstract

This chapter describes the development and investigation of a novel tool designed to generate random insertions in the genome of *C. difficile*. A *C. difficile-E. coli* shuttle plasmid was constructed to deliver a derivative of Tn*10*, mini-Tn*10*, through filter mating to investigate its behaviour in *C. difficile*. Transposition of mini-Tn*10* is achieved through the actions of a specific transposase, which is located outside the element on an unstable plasmid leading to stable insertions. Two variations of the element were investigated containing either inside- (IE) or outside- (OE) ends of IS10 flanking the chloramphenicol resistance gene. Mating efficiencies of the *E. coli* donor with *C. difficile* strains 630, 630Δerm, CD37 and R20291 were found to be extremely low (~10^{-10} per donor). The presence of the delivery plasmid and the element was confirmed in all of the *C. difficile* transconjugants by PCR. Stability of the plasmid was investigated and revealed that the plasmid was unstable in *C. difficile* after removal of the selective antibiotic. Transposition of mini-Tn*10* was not observed in any of the isolated transconjugants. Addition of IPTG to promote expression of transposase via the P_{fac} promoter and serial passage failed to promote transposition. Transposase expression was demonstrated by real time RT-PCR however, no difference in the expression level was observed for bacteria grown with IPTG compared to those grown without IPTG. This study demonstrates the ability to deliver an 11 Kb plasmid via conjugation to both laboratory and clinically relevant strains of *C. difficile*. Optimisation of the filter mating protocol and refinement of the delivery plasmid may enable mini-Tn*10* to be used as a random mutagen of *C. difficile*.

4.2. Introduction

Clostridium difficile is notoriously difficult to manipulate genetically, the bacterium appears not to be naturally competent although there is evidence of the presence of homologues to the competence operon of *B. subtilis*. Additionally, its fragility has limited the use of common techniques such as electroporation. Currently only the toxins genes TcdA, TcdB and CDT are confirmed virulence factors, with many other putative factors under investigation. This limited knowledge of the molecular basis of pathogenicity has severely hindered the design, development and optimisation of therapeutics and control measures to reduce the economic and social impact of CDI. To this end both reverse (targeted mutagenesis) and forward (random mutagenesis) genetics approaches are being assessed. The following chapter describes the development and investigation of a potential random mutagenesis system for *C. difficile* using mini-Tn*10*.

4.2.1. **Introduction of DNA into *C. difficile***

The delivery of DNA into *C. difficile* is only possible via conjugation from donor bacteria such as *B. subtilis* and *E. coli*. This can occur via CTns or by plasmids containing broad host range origins of transfer that also contain origins of replication suitable for use in *C. difficile.* Prior to the mid 1990's, genetic investigation of *C. difficile* was limited due to the lack of methods to introduce heterologous DNA and was only possible via CTns [194, 337]. One early approach to clone genes into *C. difficile* utilised Tn*916* as a shuttle CTn

which contained vector DNA [193]. The plasmid vector could thus enter the Tn by recombination between homologous regions. This plasmid also possessed a region of homology to Tn*916*ΔE and was used as a vector for the introduction of heterologous DNA into *C. difficile* strain CD37 [190, 338]. A fragment of *tcdB* was cloned into pCI195 and the plasmid::transposon structure transferred into a *B. subtilis* strain carrying Tn*916*ΔE. pCI195 is unable to replicate in *B. subtilis* therefore colonies arising on chloramphenicol selective agar represented cells which had the plasmid inserted into the chromosome. These donor cells were then mated with *C. difficile* to allow transfer of the co-integrate transposon through filter mating [193].

In another example Haraldsen and Sonenshein (2003) modified Tn*916* by adding a copy of the *sigK* gene from *C. difficile* and transferred the element from a *B. subtilis* donor to CD196 (a toxinogenic strain) [12]. The transposon failed to insert at the preferred site and unexpectedly integrated via homologous recombination into the *sigK* gene.

This ability to introduce DNA using CTns facilitated investigation of gene silencing by antisense RNA. Minton *et al*. (2004) developed a co-integrative expression vehicle pMTL900 able to replicate in *E. coli* but not in Gram-positive hosts such as *B. subtilis* [205]. The plasmid would therefore only be maintained in *B. subtilis* by homologous recombination of plasmid borne Tn*916* and the chromosomal copy of Tn*916*. An experiment in which a small section of the 5' end of the *C. difficile virR* gene, a homologue of the *C. perfringens virR* virulence transcription factor, was cloned into the plasmid

next to the promoter allowing transcription of an antisense molecule. This plasmid construct was cloned into Tn*916*, integrated into a *B. subtilis* donor and the co-integrate transferred into *C. difficile* strain 630 by filter mating [205, 339]. Only one transconjugant was obtained in these experiments; however analysis revealed the expected integration of the element and production of antisense RNA. The transconjugant initially showed altered growth characteristics and impaired growth compared to the wild type, however this was shown to be unstable due to loss of its ability to produce antisense RNA.

In a second experiment Roberts *et al*. (2003) used Tn*916* as a vector to deliver antisense RNA to the adhesin encoding gene *cwp66* into the toxinogenic *C. difficile* strain 79-685 [203]. Mating with the *B. subtilis* donor generated several transconjugants that produced antisense RNA to *cwp66*. The level of Cwp66 in these transconjugants was unchanged in comparison to wild type and no evidence of a reduction in adhesion to Vero cells was observed. In this strain, Tn*916* inserted into different sites within the genome suggesting that it may be suitable for mutagenesis, however this precluded direct comparisons between the strains.

4.2.2. **Reverse genetics - Targeted disruption**

In the era of the DNA sequencing revolution in which it is possible to read 1 gigabase of DNA in half a day with a cost of as little as $0.001 per 1000 bases, the use of reverse genetics has grown phenomenally. The complete genomes of many model eukaryotic and prokaryotic organisms and the human genome has revolutionised the field of bioinformatics, generating a

wealth of data to aid our understanding of gene function and genome organisation [340, 341]. Comparative genomics allows genetic sequences to be assessed for similarity, identifying conserved genes, defining traits and highlighting potential common ancestry and/or function. Additionally by incorporating pathogen and host sequences, homology of bacterial proteins to eukaryotic factors can identify putative virulence determinants [342]. The growth of huge databases of genetic information, such as those managed by EMBL-EBI and GenBank, and the development of numerous prediction programs which use complex algorithms to assign function, have both seen considerable investment and enhancement over the last decade [340, 343, 344]. These online resources that are relatively simple to use, enable researchers from many disciplines to investigate sequence data. Consequently, from basic sequence data an enormous amount of valuable information can be gathered, helping to direct investigations of specific targets and make informed predictions of function and provenance. Genome annotation now represents one of the biggest challenges in genomics, a complex and lengthy process relying heavily on experimental data, which in many cases is lacking. Functional assignment by prediction software although extremely useful, is not definitive and laboratory confirmation is required for accurate assignment [340]. Therefore, the use of targeted disruption represents a useful avenue of research to explore the molecular pathogenicity of *C. difficile*.

Restriction/methylation systems (RM) may be responsible for a reduction in conjugation frequencies in numerous bacterial species and are therefore important systems to consider when designing vector systems. To successfully introduce plasmids this barrier must be circumvented by

methylation of the vector DNA or removal of the corresponding recognition sites [345, 346]. Experiments by Purdy *et al.* (2002) in two toxinogenic *C. difficile* strains, CD3 and CD6, highlighted the importance of RM systems and the necessity to consider these factors when developing conjugation based strategies in *C. difficile* [177]. Herbert *et al.* (2003) investigated gene transfer and RM systems in the toxinogenic strain 630 and showed that this strain does not show any restriction barrier to DNA transfer [347]. Additionally, no evidence of corresponding restriction endonuclease candidate genes was found in sequences in the proximity of the five identified methylase genes, suggesting that the methylase genes do not participate in RM and fulfil other functions [347].

Liyanage *et al.* (2001) describe the introduction of a suicide vector (pMTL31) carrying internal fragments of the gene *gldA* (encoding glycerol dehydrogenase) into CD37 by conjugative transfer from an *E. coli* donor CA434 [348]. This non-toxinogenic strain was shown not to have restriction activity and erythromycin resistant integrants were selected. However, only pin-prick colonies were observed suggesting the lethality of such a mutation.

Lastly O'Connor *et al.* (2006) describe the use of a replication-deficient plasmid to create reproducible targeted insertional inactivation of chromosomal genes by single crossovers [132]. The group observed that an *E. coli-C. perfringens* shuttle vector, pJIR1456, was unstable in *C. difficile*. The plasmid contains the pIP404 replicon, which is relatively unstable in *C. difficile*, producing a type of conditional lethal vector i.e. the plasmid is rapidly

lost when selection is removed. Researchers can take advantage of this to deliver gene constructs to the chromosome [132, 177]. Using the toxinogenic strain JIR8094 (strain 630 with a deletion of one *ermB* gene rendering it erythromycin sensitive) two genes were inactivated *rgaR* (CD3255) and *rgbR* (CD1089) through single crossover integration of the plasmid. Both genes are putative response regulators which encode proteins similar to the toxin gene regulator, VirR, part of the VirS/VirR two-component system which regulates several toxin and housekeeping genes in *C. perfringens* [9, 132, 349, 350]. The potential instability of this system was revealed by Southern hybridisation analysis, which showed that the plasmid could 'loop out 'of the chromosome at relatively low levels. However, their research represented an important advancement in the development of tools for the targeted mutagenesis of the *C. difficile* genome, and allowed the identification of genes that are putatively regulated by RgaR and RgbR proteins using microarray analysis.

4.2.2.1. Group-II introns and the TargeTron

Creation of mutants by recombination was an important step in the development of tools for *C. difficile* manipulation but was hindered by the low frequency of mutant isolation particularly due to the low frequency of transfer of elements to toxinogenic strains such as 630 (between 10^{-6} and 10^{-7} per donor cell) and the ability of the mutant to revert to wild-type. To improve this one could increase the frequency with which DNA is transferred by development of alternative suicide vectors, or allow integration of DNA by a process which is largely independent of the host recombination factors [148].

Group II introns are present in a wide range of bacterial genomes and in eukaryotic organelles, interrupting protein-coding or RNA genes. These catalytically active, RNA molecules self-splice out of the RNA transcripts of their host gene by a mechanism analogous to that of the excision of eukaryotic nuclear introns. Some group II introns are mobile and contain an open reading frame (ORF) containing a multidomain intron-encoded protein (IEP). A well-studied example of a mobile group II intron is the *L1.LtrB* intron from the conjugative plasmid associated *ltrB* gene of *Lactococcus lactis* [351]. The specificity of the process of integration is known as retro-homing and is largely determined by basepairing between sequences within the intron RNA and target site DNA. Investigation of this mobile intron revealed that by changing the intron sequence the target specificity can be changed allowing *L1.LtrB*-derived introns to be targeted to almost any gene using automated tools, these are generally referred to as 'TargeTrons' [352]. The IEP ORF can be removed and placed *in cis* or *in trans* orientation without affecting mobility, this property allows stable insertion of the element by preventing further splicing and mobility when the plasmid encoding the IEP is removed [148]. Furthermore the tools needed to allow integration are encoded by the element itself and are therefore independent of host factors. The only limitation of this system is that for re-targeting there is a requirement for a minimal recognition sequence consensus which is determined by the IEP [353].

The Lambowitz group used two important observations to further develop this tool; i) an intron can reside within another intron known as a twintron, and ii) the non-structural component of the IEP can be altered to carry 'cargo

sequence' without affecting the function of the intron. An artificial twintron was constructed where an antibiotic resistance marker interrupted by a group I intron was placed into *L1.LtrB* [354, 355]. The antibiotic resistance gene is placed in a reverse orientation within the *L1.LtrB* intron and is not expressed in the plasmid due to the group I intron (Figure 4.1). As the construct is transcribed the group I intron (in forward orientation) self-splices out of the RNA, generating a transcript of the intact antibiotic gene in reverse orientation. Retrotransposition of the mature group II intron RNA into the target gene confers antibiotic resistance to the bacterium allowing positive selection of the gene knock-out. This marker is known as the RAM element, which describes the restoration of the marker ORF made possible by passing through an RNA intermediate [130, 148, 354, 355].

A modified version of the TargeTron was developed by Chen *et al*. (2005) for use in *C. perfringens*. It contained a clostridial replicon, promoter and selectable marker. Their experiments proved that it was possible to target the TargeTron into a clostridial gene and produce stable mutants [356]. However, they did not make use of the RAM element instead they identified mutants using phenotypic screens and PCR, which is only useful for genes with a known phenotype. Heap *et al.* (2007) have developed a construct, named the ClosTron, which is specifically designed for use in clostridia with a functional RAM element for efficient identification of knock-out mutants.

Image unavailable due to copyright restrictions

Figure 4.1 Schematic of TargeTron/ClosTron mechanism of action, taken from Brüggemann & Gottschalk, 2009 [148]

A – An appropriate plasmid carrying the Group (Gp) II intron and RAM element (Gp I intron and antibiotic resistance gene) is transferred to the host organism, in *C. difficile* this is achieved through filter mating from an *E. coli* donor. The relative orientations of the Gp I and the RAM element are crucial for this system.

B – the plasmid borne RAM element does not confer antibiotic resistance due the interruption of the Gp I intron. The Gp I intron cannot self-splice from the gene transcript as it is in the reverse orientation.

C – The Gp II intron RNA is transcribed from a strong promoter on the plasmid. The Gp I intron is in the forward orientation in this transcript and self splices however, antibiotic resistance is not conferred at the antibiotic resistance gene is in the reverse orientation. The mature Gp II intron then undergoes retrotransposition into the target gene (in some bacteria in the population).

D – After retrotransposition into the target gene an intron containing a functional antibiotic resistance gene (lacking the Gp I intron) is present on the chromosome allowing selection

The *L1.LtrB* group II intron was modified by the addition of an *ermB* gene with the group I intron *td* from phage T4 inserted into a linker region between the ermB ORF and a *thlA* promoter, this replaced the natural and relatively inefficient *ermB* promoter [130].

Delivery of the ClosTron is achieved by conjugation of the plasmid pMTL007 from an *E. coli* donor and has been used successfully in *C. acetobutylicum*, *C. sporogenes*, *C. botulinum*, *C. beijerinckii*, *C. sordellii* and *C. difficile* [162, 206, 357]. One of the drawbacks of the system is that the insertion of the element could lead to polar effects within the mutants. Refinements such as addition of FRT sites allow the excision of the retargeted *ermB* gene through *in vivo* expression of the yeast FLP recombinase allowing further genes to be targeted within the mutant [148]. This novel tool will be invaluable for characterisation of virulence-associated genes in clostridia.

4.2.1. **Forward genetics - Random disruption**

To complement reverse genetic strategies forward genetics aims to identify the genetic basis of a phenotype without making assumptions about the genes involved, it is therefore possible to perform without sequence data. Although the genomes of several major pathogenic and commercially relevant clostridia have been sequenced the annotation is far from complete and while sequence comparisons can be employed to infer function this is limited to a select number of genes. The forward genetics approach can thus be used to characterise novel genes and regulatory processes to elucidate the molecular pathogenesis of pathogenic clostridia such as *C. difficile*.

4.2.1.1. Transposon mutagenesis

Transposon mutagenesis is a prime example of this technique; the ability to disrupt DNA randomly has been used extensively in bacteria and eukaryotic organisms to generate vast libraries of random insertion mutants [358-360]. These libraries can be built both quickly and economically and screened *in toto* for phenotypic changes in simple *in vitro* screening assays, or be subjected to *in vivo* investigation in model systems [179, 207]. The location of the transposon can then be identified to allow characterisation of the disrupted gene and its role in pathogenesis. This strategy has been used successfully in a number of organisms, and in conjunction with sequence data provides a powerful means of identifying virulence associated genes in bacterial pathogens.

In *Pseudomonas aeruginosa* an ordered transposon mutant library has been generated in the laboratory strain PAO1 and a clinical isolate, PA14. This comprehensive library contains individual mutants with almost saturation, deletions exist in 3985 of 4288 genes and this represents the first complete deletion mutant set for a bacterial species [342, 361, 362]. Other innovative approaches include the use of microarray technology to identify mutants from pooled populations which fail to survive *in vivo* model infection, allowing negative selection of mutants for further investigation [342, 363-365]. Transposons may also be used for signature-tagged mutagenesis which is particularly useful when investigating the virulence of bacterial pathogens, and for sequencing and protein fusion technologies [179, 366].

Tn10

The composite transposon Tn*10* encodes tetracycline resistance determinants with inverted repeats of IS*10* (Figure 4.2) and transposes by a nonreplicative mechanism akin to that of Tn*5* [367]. Tn*10* transposition involves the formation of a transpososome, a higher order protein-DNA complex thought to be an essential prerequisite for transposition. Like Tn*5* the right IS*10* element encodes a functional transposase protein whereas the left IS*10* element encodes a non-functional protein [368]. The ends of the IS*10* sequence are referred to as either outside ends (OE) or inside ends (IE) depending on their position in relation to the tetracycline resistance determinants (Figure 4.2). The OE and IE are 23 bp imperfect inverted repeat sequences which determine specific transposase binding and are all functional [369]. Each section of the IE and OE appear to have distinct function, for example positions 6-13 are believed to be required for binding of the transposase mediating pairing of the transposon ends allowing correct positioning before initiation of the chemical steps. The IE contains the sequence GATC, which is methylated by host *dam* methylase. The methylation state of this end controls transposition, the unmethylated and hemimethylated state is more active than the fully methylated form and this type of control prevents killing of the bacterium during non-replicative transposition [370, 371]. The inverted repeat on the OE differs from the IE inverted repeat at five positions resulting in the requirement of host proteins for transposition.

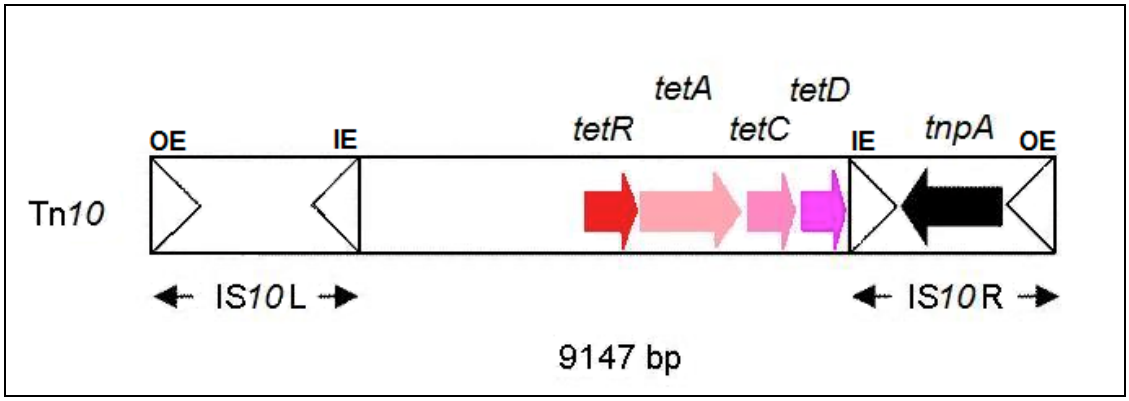


Figure 4.2 Schematic diagram of the genetic organisation of Tn*10* adapted from Hayes, 2003 [179].

Tn*10* is a composite transposon, additional genes are included encoding antibiotic resistance (shown by coloured arrows) flanked by two inversely orientated copies of IS*10* elements.

Host factors can play a significant role in the generation of the transpososome, integration host factor (IHF), the related histone-like protein (HU) and histone-like nucleoid structuring protein (H-NS) are involved in regulating assembly and/or function and target site selection [181]. IHF is a sequence-specific DNA binding protein that plays an important role in Tn*10* circle formation, binding specifically to the OE [372]. HU is an abundant, basic protein related to IHF but it lacks target DNA sequence specificity; it can also promote circle formation of Tn*10* but with reduced specificity. Reactions involving only the inside ends do not require host proteins whereas, for reactions involving only the outside ends, the availability of host proteins is critical [372].

Transposition of Tn*10* involves several steps beginning with binding of the transposase to the end sequences to form the transpososome. In conjunction with a divalent metal cation (Mg^{2+} or Mn^{2+}) three chemical steps then occur at each end of the transposon, first strand nicking, hairpin formation and hairpin resolution (Figure 4.3) [185, 373]. Donor DNA is nicked forming a double stranded break at the transposon-donor junction leading to hairpin formation, hairpin resolution then re-exposes the 3' OH groups to allow strand transfer. Transesterification leads to transfer of the 3' OH termini to phosphate groups on opposite strands of the target site separated by 9 base pairs, a process called strand transfer. After strand transfer is completed host repair functions fill in the 9 base pair gaps leading to duplication of the original short target sequence [181].

Image unavailable due to copyright restrictions

Figure 4.3. Schematic representation of the mechanism of Tn*10* transposition taken from Haniford 2006 [181].

Transposase (striped ovals) binds to the OE (black rectangles) of Tn*10* forming the transpososome. The transposase then catalyses the chemical steps in excision of the transposon from the donor DNA. After removal of the flanking DNA from the transposon ends the transpososome binds to the target DNA (grey rectangle) and catalyses the chemical steps in the strand transfer reaction. Host repair mechanisms fill in the 9 bp gaps generated during strand transfer leading to a small duplication of the original target site (checked rectangle).

Although the insertion of Tn*10* is essentially random it has been shown to have a preferred target site of a 6 base pair symmetrical consensus sequence (5' NGCTNAGCN 3') determined by the transposase [374]. As four of the six nucleotides on this consensus sequence are C/G, insertion into AT rich DNA is likely to be more randomly distributed [375]. The transposase is the only transposon encoded protein required for transposition leading to recognition of the transposon end sequence, generation of double stranded cuts at the donor ends and promotion of strand transfer into a target site [370].

Tn10 derivatives

When using transposons for mutagenesis, there are a number of considerations largely dependent upon the ultimate goal of the research and the availability of elements. The following describes the important considerations and potential adaptations to optimise the use of Tn*10* as a random mutagenesis tool.

Stability: To ensure that insertions are stable, the transposase would be located outside of the transposon itself, ideally on the transposon donor molecule, which is lost after subsequent transposition. These usually smaller derivatives are referred to as minitransposons. The use of minitransposons can counteract the problems associated with secondary transposition events including a) problems mapping the insertion due to multiple copies, b) promotion of rearrangement of adjacent DNA at high frequency, c) high frequency transposition of IS elements compared to the entire element, specifically in relation to composite transposons such as Tn*10*, d)

cotransduction problems of the transposon marker associated with using phage based delivery vehicles [375].

Specificity: An important characteristic of transposons is their ability to insert into numerous different locations. In reality, all transposons possess some degree of target site specificity and do not insert completely at random. Tn*10* shows a relatively high degree of target site specificity compared to other elements meaning some target genes may be "cold" spots whereas others may be "hot" spots. This can be overcome by screening larger numbers of insertions or by the development of derivatives with reduced target site specificity through mutations in the *transposase* gene [375]. For Tn*10* in addition to the 6 base pair symmetrical consensus sequence, the 8-10 bp located immediately adjacent to the target consensus sequence on either side can affect the efficiency of insertion into a particular site, possibly influencing DNA structure. Tn*10*-ATS (altered target specificity) contains a mutant transposase with changes at specific positions, this leads to a lower degree of insertion specificity in *E. coli* compared to wild type transposase [376]. Specifically, two mutations occur at two defined sites within the protein, patch I (atsI) and patch II, (atsII) which are involved in target site interactions. These mutations relax the specificity where certain base pairs are favoured over others in the insertion consensus sequence and do not significantly affect the frequency of transposition [376, 377].

Selection: Care should be taken when choosing a selectable marker considering the downstream applications of the system. The use of drug

resistance markers is particularly useful for the analysis of mutations, which produce phenotypes that are difficult to screen. To broaden the application of Tn*10*, new antibiotic resistance genes have been added instead of, or in addition to Tet^R, these include Kan^R, Cm^R and Ap^R. Therefore an appropriate selectable marker can be inserted for use in *C. difficile*.

Transposon size: In general smaller transposons are preferable, large elements are harder to handle, particularly on multicopy plasmids leading to the increased likelihood of instability. Furthermore, for Tn*10*- and Tn*5*-based transposons, the frequency of transposition is reduced by 40% for every kilobase of the transposon length [375].

Rearrangements/deletions: Almost all mobile elements can generate deletions adjacent to their site of insertion particularly Tn*10*. Tn*10* promotes deletions via the interaction of two intact IS*10* elements containing IEs and a target site adjacent and outside the element. Deletion/inversion events can also occur which lead to an inversion of the IS*10* element and a section of adjacent chromosomal DNA relative to its original location. Furthermore, all transposable elements can give rise to low frequency rearrangements in secondary transposition events. Small internal deletions prevent adjacent Tn*10* inversions/deletions.

Polarity: Insertions of large transposable elements such as Tn*10* can have polar effects on the expression of downstream genes. The extent of these effects largely depends on the strength of the transcription termination

signals in the transposon and the precise location of the insertion with respect to the target gene Rho-dependent termination sites [375, 378]. These effects may need to be considered when designing screening assays and during the characterisation of putative insertion mutants.

Way *et al.* (1984) described the development of derivatives of Tn*10* which can be more useful than the wild type element for genetic and physical analysis of bacteria [379]. These derivatives include one or more of the following properties i) new drug resistance markers ii) high frequency of transposition, iii) placement of the transposase gene outside the element, iv) internal deletions to prevent adjacent inversions/deletions by Tn*10*, v) addition of a *trp-lac* operon fusion segment to automatically generate a transcriptional fusion with the interrupted operon [379].

One of these derivatives, known as mini-Tn*10* has been used in a wide range of Gram-positive and Gram-negative bacteria to generate insertion libraries [197, 198, 380]. The elements are small, typically 400-3000 bp and can be engineered to carry a wide range of markers. Either the wild-type or ATS transposase is provided in *cis* to allow stable insertions. Each terminus contains short Tn*10* ends either OEs or IEs in inverted orientation flanking the selectable marker(s) [375, 379].

Petit *et al*. (1990) adapted a mini-Tn*10* element for use in the Gram-positive bacterium *B. subtilis* [197]. The features of this derivative include: replacement of the Tet^R gene with Cm^R, gene fusion of the ATS transposase to a

ribosome binding site functional in *B. subtilis*, and placement of a P_{tac} promoter upstream of the element allowing expression and transposition in a Gram-positive host. The Cm antibiotic resistance gene is flanked by 307 bp of IS*10* inverted repeats and short regions of the extremities of either the OE (42 bp) or the IE (23 bp). Two derivatives harbouring either OEs or IEs, found on plasmids pHV1248 and pHV1249 respectively, together with the transposase in *cis* orientation on a thermosensitive plasmid containing erythromycin and ampicillin selectable markers (Figure 4.4) [197]. The authors successfully demonstrated the delivery and transposition of both of the mini-Tn*10* derivatives producing random, high frequency stable insertions in *B. subtilis* [197]. Further studies using mini-Tn*10* have proved successful in other *Bacillus* species. Day *et al.* (2009) constructed a library of random insertion mutants covering 82% of *B. anthracis* genes using an ATS mini-Tn*10* derivative to investigate sporulation and germination [196]. Kamoun *et al.* (2009) generated a mini-Tn*10* library in *Bacillus thuringiensis,* an economically important strain used as a pesticide, and observed random single insertions allowing investigation of bacteriocin production [381]. Consequently mini-Tn*10* has proved to be an effective random mutagenesis tool in a number of *Bacillus* species.

4.2.2. **Transposon mutagenesis in *C. difficile***

The use of transposon mutagenesis has not been described in *C. difficile*, largely due to the problems associated with the introduction of DNA. The only method for successful introduction of DNA is via conjugation. Mullany *et al.* (1990) demonstrated that the conjugative transposon Tn*916* could be

introduced into *C. difficile* from *B. subtilis* and could integrate into the genome (described in detail in chapter 2) [337]. Tn*916* has been shown to integrate into a specific location in *C. difficile* strain CD37 it, however, in strain 630Δerm insertion occurs at multiple sites with preference for AT rich target sites [131, 194, 195]. This property has limited the usefulness of Tn*916* as a random mutagen in *C. difficile*.

4.3. Aims and objectives

This project aims to develop a new random mutagenesis system using a mini-Tn*10* transposon delivered via a conjugative plasmid vector. The major goals are three-fold, i) develop a method for efficient delivery of the mobile element into the genome of *C. difficile*, ii) characterise the movement of the mobile element within the genome and, iii) generate a library of insertion mutants which may be screened for mutations in genes and pathways associated with virulence.

Objectives:

1. Identify a suitable delivery plasmid for mini-Tn*10*.
2. Construct a delivery plasmid containing mini-Tn*10*.
3. Conjugate the plasmid into *C. difficile*.
4. Confirm presence of delivery plasmid/transposon in *C. difficile*.
5. Investigate the behaviour of the transposon in *C. difficile*.
6. Generate a library of insertion mutants.
7. Screen the library for mutations associated with virulence.

4.4. Methods

4.4.1. **Bacterial strains and culture conditions**

Clostridium difficile strain R20291 was cultured on BHI agar supplemented with 5% defibrinated horse blood and CDSS (BHIBS) or in pre-reduced BHI broth at 37^{o}C in an anaerobic chamber (Don Whitley Scientific Ltd, UK) with an atmosphere of 80% N_2, 10% H_2, 10% CO_2. *E. coli* strain CA434 (HB101 carrying the IncPβ conjugative plasmid, R702) was cultured on Luria-Bertani (LB) agar or broth at 37^{o}C aerobically [382]. All cultures were grown in 20 ml universals without shaking unless otherwise stated. Where appropriate, medium was supplemented with CDSS and antibiotics; ampicillin (Ap) (100 µg ml^{-1}), erythromycin (Em) (10 µg ml^{-1}), kanamycin (Kan) (50 µg ml^{-1}), chloramphenicol (Cm) (12.5/25 µg ml^{-1}), thiamphenicol (Tm) (100 µg ml^{-1}), lincomycin (Lm) (100 µg ml^{-1}) (all Sigma-Aldrich, UK). Table 4.1 describes the properties of the bacterial strains and plasmids.

Strain/plasmid	Relevant properties	Reference / Source
Bacterial strains		
Escherichia coli K12 DH5α MCR	General cloning vector	Bioline, UK
Escherichia coli K12 DH5α MCR (Ec)	Electrocompetent cloning vector	Bioline, UK
Escherichia coli CA434	HB101 carrying R702 (IncPβ) plasmid, Tra^+, Mob^+, Kan^R	Hedges and Jacob, 1974 [383]. Nigel Minton, University of Nottingham
Clostridium difficile R20291	Em^R, Tet^S	Nigel Minton, University of Nottingham/Jon Brazier, Anaerobe Reference Laboratory, Cardiff
Clostridium difficile 37	Rif^R	Hächler, *et al*., 1987 [230]. P. Mullany UCL Eastman Dental Institute, UK
Clostridium difficile 630	Tet^R, Em^R	Wust and Hardegger [384], P. Mullany UCL Eastman Dental Institute, UK
Clostridium difficile 630Δerm	Tet^R, Em^S	Hussain *et al.,* 2005 [131]; P. Mullany UCL Eastman Dental Institute, UK
Plasmids		
pHV1248	mini-Tn*10* with outside ends Ap^R, Tm/Cm^R, Em^R	Petit *et al.,* 1990 [197], N. Minton, University of Nottingham, UK
pHV1249	mini-Tn*10* with inside ends Ap^R, Tm/Cm^R, Em^R	Petit *et al.,* 1990 [197], N. Minton, Bacillus Genetic Stock Centre, USA
pMTL5401F	pCB102 replicon, P_{fac} promoter, Em^R, Ap^R	Heap *et al.*, 2007 [130], N. Minton, University of Nottingham, UK
pKM1048	*E. coli-Clostridium* shuttle, containing mini-Tn*10* with outside ends, pCB102, Ap^R, Tm/Cm^R, Em^R	This work
pKM1049	*E. coli-Clostridium* shuttle, containing mini-Tn*10* with inside ends, pCB102 Ap^R, Tm/Cm^R, Em^R	This work

Table 4.1 Bacterial strains and plasmids

Kan: kanamycin, Em: erythromycin, Ap: ampicillin, Cm/Tm: chloramphenicol/thiamphenicol, Rif: rifampicin; R = resistant, S = sensitive.

4.4.2. **Plasmid construction**

4.4.2.1. Generation of pKM1048

PCR

PCR primers containing *Bam*HI restriction sites were designed to amplify the target region (~3 Kb) from plasmid pHV1248 (Figure 4.4). PCR reactions consisted of 1X Buffer, 0.1 µM primers KM1 F and KM1 R (Table 4.2), 200 µM dNTPs, 5 U *Taq* DNA polymerase and 1-10 ng target DNA in a final volume of 100 µl (all PCR reagents New England Biolabs, UK unless otherwise stated). The standard PCR conditions were [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 50-58^{o}C 1.5 min, 72^{o}C 4-6 min] x 30-35 cycles [72^{o}C 10 min] x1 cycle.

After amplification PCR products were visualised by gel electrophoresis and purified using the PCR clean-up kit (Qiagen, UK). The PCR product and the vector plasmid (pMTL5401F) (Figure 4.5) were digested with *Bam*HI for 3 or 16 hours at 37^{o}C (all enzymes New England Biolabs, UK, unless otherwise stated). 10 U alkaline phosphotase (CIP) and 1X buffer was added to both samples and incubated for 1 hour at 37^{o}C, reactions were purified using the PCR clean-up kit. Digests were visualised by gel electrophoresis and DNA was extracted and purified using a gel extraction kit (Qiagen, UK). Digested pMTL5401F and the mini-Tn*10* PCR product were ligated for 16-18 hours at 22ºC. The reaction contained 0.3 µg pMTL5401F, 1 µg PCR product, 1X T4 ligase buffer with ATP, 1 weiss unit T4 ligase to a final volume of 50 µl.

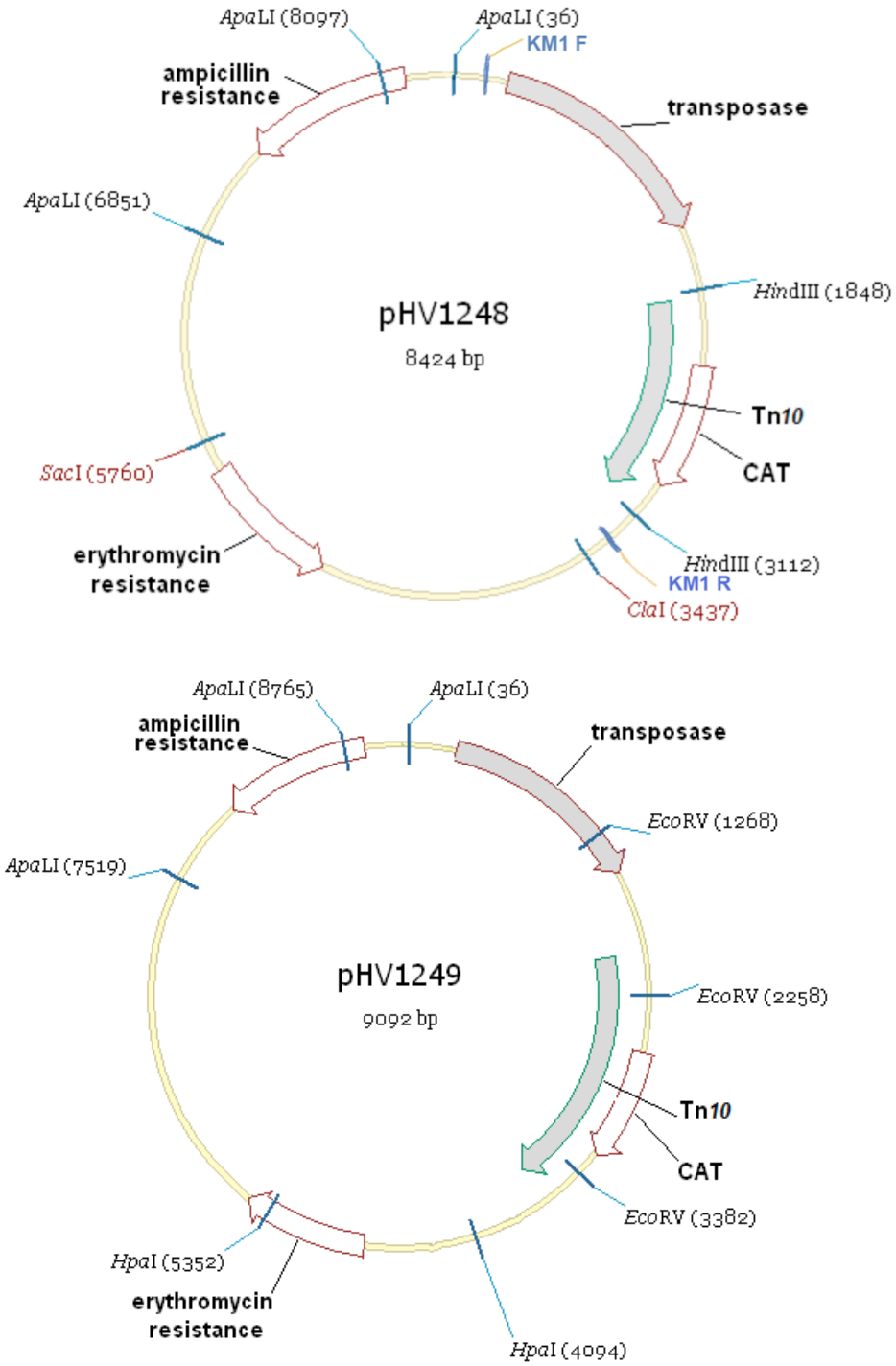


Figure 4.4 pHV1248 and pHV1249 [197]

The plasmids contain a mini-Tn*10* transposon, pHV1248 contains mini-Tn*10* with outside ends and pHV1249 contains mini-Tn*10* with inside ends. They each encode transposase, ampicillin, erythromycin and chloramphenicol resistance. Important restriction sites used for cloning are also labelled the enzymes in red cut at only one position. The size of each plasmid is shown in the centre of the diagram. The approximate positions of KM1 F and KM1 R primer pair are also shown on pHV1248.

Primer	Sequence (5' - 3')	Product size bp	Use/location	Additional information
KM1 F KM1 R	CT*GGATCC*GGCTCGTATAATGTG CT*GGATCC*ATGTATTGCGATATGC	2980	Amplification of *transposase* and mini-Tn*10* region	Both primers include a *Bam*H1 site (italics)
Tn10 F Tn10 R	CAGGTTAGTGACATTAG CAGACAAGTAAGCCTCC	762	Amplification of entire *cat* gene in mini-Tn*10*	
cat 1 cat 2	CTGGTTACAATAGCGACGGAGAG CCTGCTGTAATAATGGGTAGAAG	304	Amplification of a section of the *cat* gene	
amp 1 amp 2	CAGTGCTGCAATGATACCGC CATCGAACTGGATCTCAACAGC	610	Amplification of a section of the *bla* gene	
bla F bla R	CGCTCAGTGGAACGAAAACTC CCCTGATAAATGCTTC	1038	Amplification of entire *bla* gene	
ermB F ermB R	CTGACTTGCACCATATC CATAGTGACTGGCGATGCTGTC	1085	Amplification of entire *ermB* gene	
fac F fac R	CACAGCCATACCACAGCTTCC CCCAATACGCAAACCGCCTC	319	Amplification of entire P_{fac} region	
Tpase F Tpase R	CTTTCTAAGGCAGACCAACC CGACATCATAACGGTTCTGGC	1518	Amplification of entire *transposase* gene	
compTn10 F compTn10 R	CGGCAAGCTCTTTTAGG CATTGTCAGCAAATTGATCC	1394	Amplification of entire mini-Tn*10* region	
KM6 F KM6A R	CCGGGATCATATGACAAGATGTG CCCTATCCAAGAAGTGATGCC	234	Amplification of a region covering the *Hin*dIII site (position 4644 in pKM1048). Reverse primer located on plasmid backbone	
KM6B R	CTTGCGGGCGTTCATGCTC	652	Reverse primer located in *transposase* gene paired with KM6F	

RT-PCR Primer	Sequence	Product size bp	Use/location	Additional information
qCat F	ACCAACAAACGACTTTTAGTATAACCACAG	168	Amplification of a portion of *cat* in RT-PCR	Designed by SigmaGenosis
qCat R	CCTAACTCTCCGTCGCTATTGTAACC			
qTpase F	ACTTGTTGACTGGTCTGATATTCGTGAG	166	Amplification of a portion of *transposase* in RT-PCR	Designed by SigmaGenosis
qTpase R	GAATGCTCGCAAGGTCGGCTAG			
q16S F	GTAGCGGTGAAATGCGTAGATATTAG	80	Amplification of a portion of the *16S rRNA* gene from *C. difficile* in RT-PCR	Designed by SigmaGenosis commissioned by Dr. C Clayton
q16S R	CGTGCCTCAGCGTCAGTTAC			
qGyrA F	AACATCAGGAAGAGCACGTCCAGC	83	Amplification of a portion of the *GyrA* gene from *C. difficile* in RT-PCR	Designed by SigmaGenosis commissioned by Dr. C Clayton
qGyrA R	TCCCTATTGAAATAGCGGAAGAAATG			

Probe	Primer Pair	Product size bp	Use/location	
Pb-cat	cat 1 & cat 2	304	Southern blotting	
Pb-catH	KM6 F & KM6B R	652		
Pb-amp	amp 1 & amp 2	610		
Pb-Tn10	KM6 F & KM6A R			

Table 4.2 Primer sequences and probe information

The table lists the primers used for analysis of the construction of pKM1048/49, primers for RT-PCR, and primers used to create probes for Southern blotting. All primers synthesised by SigmaGenosys, UK.

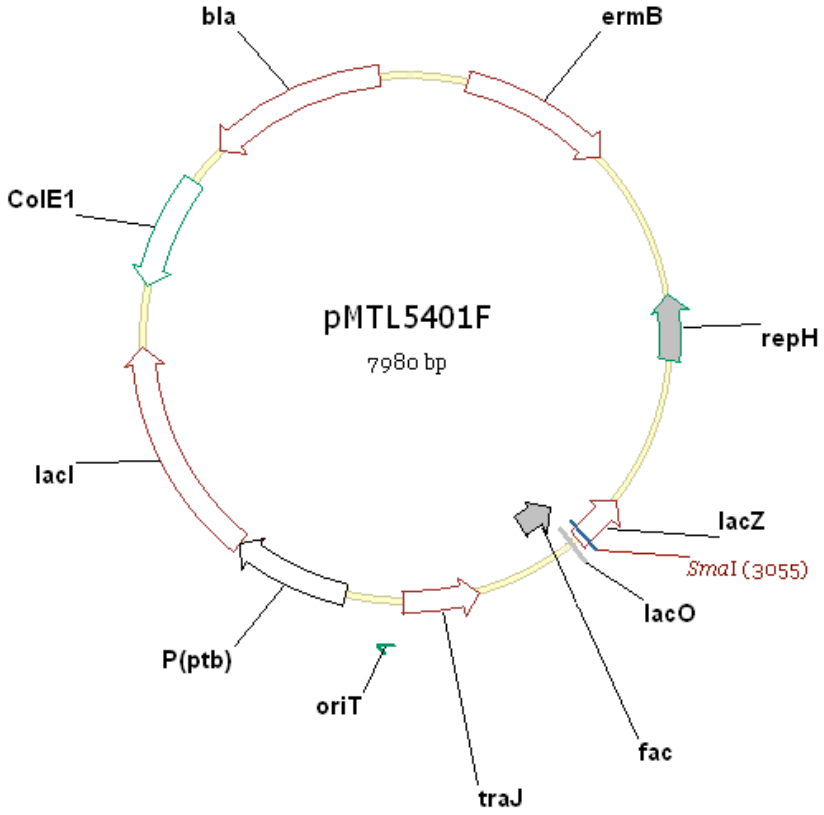


Figure 4.5 The *E. coli-Clostridium* shuttle plasmid pMTL5401F

Schematic representation of plasmid pMTL5401F. Important features are labelled including the antibiotic resistance genes *ermB* and *bla* providing resistance to erythromycin and ampicillin respectively. The plasmid contains the Gram-negative replicon ColE1, the Gram-positive replicon repH from pCB102, the transfer gene *traJ* and the promoter P_{fac} (fac).

Restriction digestion

Plasmid pHV1248 was digested with *Cla*I, *Sac*I and *Apa*LI simultaneously for 18 hours at 37°C. Digests were separated on a 0.8% agarose gel and the appropriate band (3.4 Kb) extracted using the gel extraction kit and eluted in 30 µl molH_2O. Fill-in of 3' recessed ends was achieved by adding 5 U of DNA polymerase I, Large (Klenow) fragment and 1X supplied buffer to the DNA fragment. The sample was incubated for 15 min at 25°C and the reaction stopped by addition of EDTA to a final concentration of 10 mM, and by heating to 75°C for 20 min. The reaction was purified using the PCR clean-up kit and eluted in 30 µl molH_2O. In parallel approximately 3 µg of pMTL5401F was digested with *Sma*I at 25°C for 18 hours. The digested plasmid was incubated with 10 U of CIP for 60 min at 37°C. The reaction was purified using the PCR clean-up kit and eluted in 30 µl molH_2O.

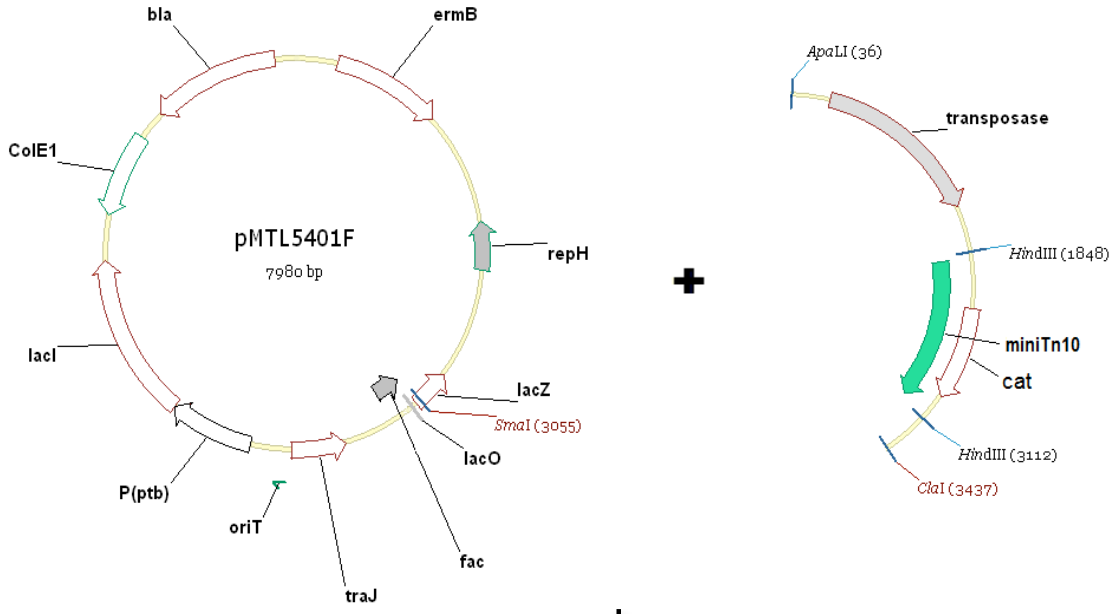
An approximation of concentration was determined from band intensities in the agarose gel and a ratio of vector (pMTL5401F) to insert (mini-Tn*10*) of approximately 1:5 was established. 5 U of T4 DNA ligase and 1X T4 DNA ligase buffer with ATP was added and the fragment mixture and incubated for 18 hours at approximately 20-22°C to generate pKM1048. pKM1048 was transformed into ElectroSHOX electrocompetent and/or DH5α chemically competent *E. coli* (Bioline, UK) (Figure 4.6 A) (as described in section 4.4.2.3 and 4.4.2.4).

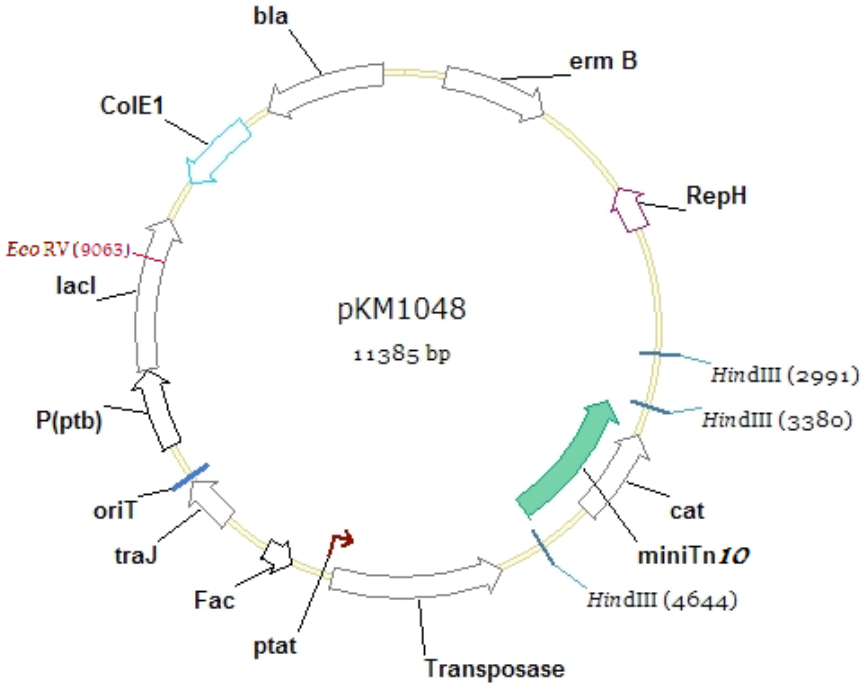
4.4.2.2. Generation of pKM1049

pHV1249 was digested with the restriction enzymes *Apa*L1 and *Hpa*I simultaneously at 37°C for 18 hours. Digests were separated on a 0.8%

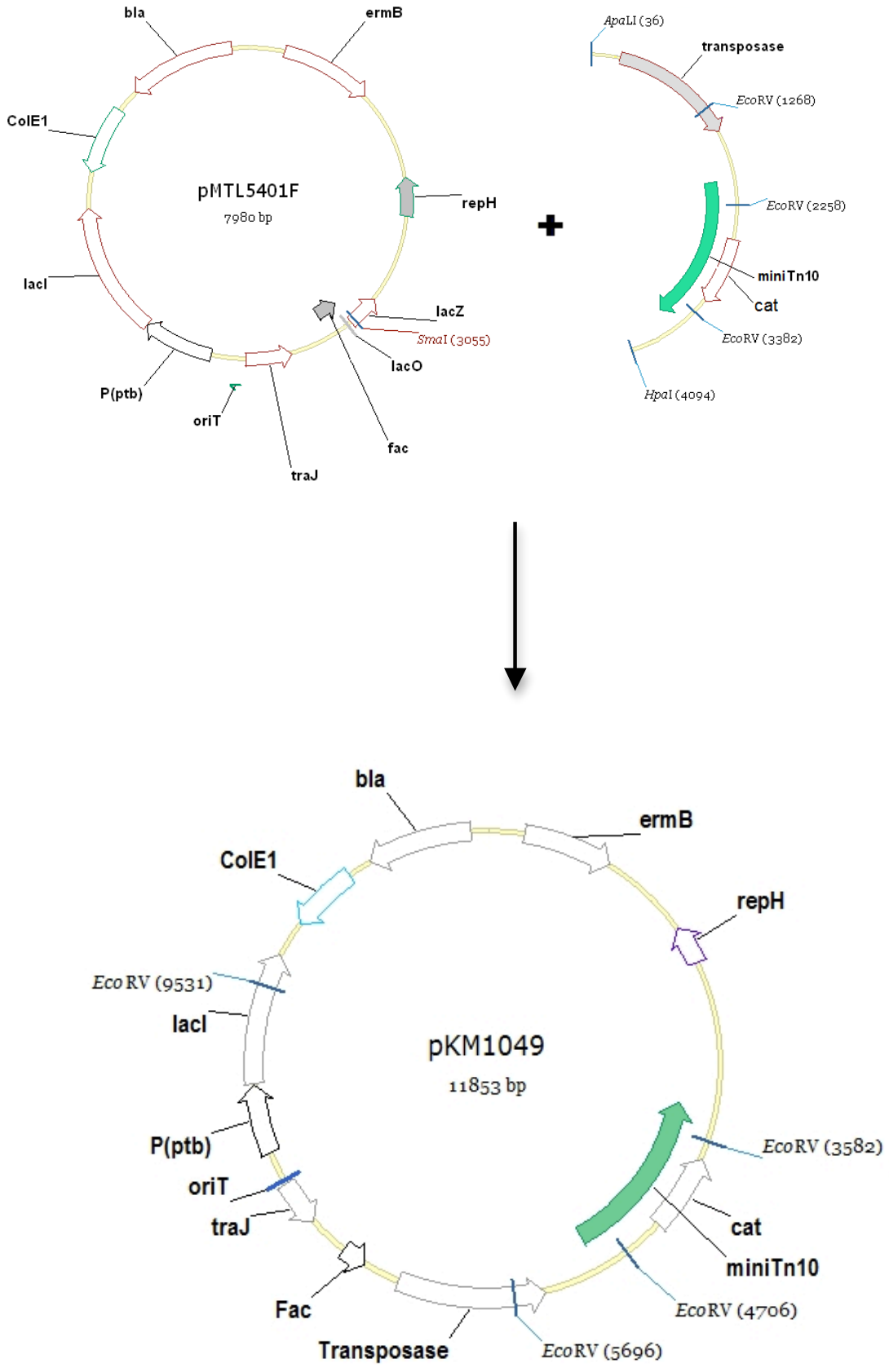
agarose gel and the appropriate band (3.9 Kb) extracted using the gel extraction kit and eluted in 30 µl molH_2O. Fill-in of 3' recessed ends was achieved by adding 5 U of DNA polymerase I, Large (Klenow) fragment and 1X supplied buffer to the DNA fragment. The sample was incubated for 15 minutes at 25^oC and reaction stopped by addition of EDTA to a final concentration of 10 mM, and by heating to 75^oC for 20 min. The reaction was purified using the PCR clean-up kit and eluted in 30 µl molH_2O. In parallel approximately 3 µg of pMTL5401F was digested with *Sma*I at 25^oC for 18 hours. The digested plasmid was incubated with 10 U of CIP for 60 min at 37^oC before purification using the PCR clean-up kit.

An approximation of concentration was made by assessment of products on agarose gels and a ratio of vector to insert of approximately 1:5 was established. 5 U of T4 DNA ligase and 1X T4 DNA ligase buffer with ATP was added and the fragment mixture and incubated for 18 hours at 20-22^oC to generate pKM1049. pKM1049 was transformed into α-select chemical competent *E. coli* (Figure 4.6 B) (se section 4.4.2.3).


bla
ermB
ColE1
pMTL5401F
7980 bp
repH
lacI
lacZ
SmaI (3055)
lacO
P(ptb)
oriT
traJ
fac
ApaLI (36)
transposase
HindIII (1848)
miniTn10
cat
HindIII (3112)
ClaI (3437)


bla
erm B
ColE1
RepH
EcoRV (9063)
lacI
pKM1048
11385 bp
HindIII (2991)
HindIII (3380)
P(ptb)
cat
oriT
miniTn10
traJ
HindIII (4644)
Fac
ptat
Transposase

A


bla
ermB
ColE1
pMTL5401F
7980 bp
repH
lacI
lacZ
SmaI (3055)
lacO
P(ptb)
oriT
traJ
fac
ApaLI (36)
transposase
EcoRV (1268)
EcoRV (2258)
miniTn10
cat
EcoRV (3382)
HpaI (4094)
bla
ermB
ColE1
repH
EcoRV (9531)
lacI
pKM1049
11853 bp
P(ptb)
EcoRV (3582)
oriT
cat
traJ
miniTn10
Fac
EcoRV (4706)
Transposase
EcoRV (5696)

B

Figure 4.6 Schematic cartoon of the generation of pKM1048 (A) & pKM1049 (B) delivery plasmids

Plasmid pHV1248 was digested with restriction enzymes *Apa*LI, *Cla*I and *Sac*I to allow isolation of the desired ~3.4 Kb fragment containing *transposase* and the mini-Tn*10* cassette. The fragment was ligated with pMTL5401F, digested with *Sma*I, to create pKM1048 (A). Plasmid pHV1249 was digested with restriction enzymes *Apa*LI and *Hpa*I to allow isolation of the desired ~3.9 Kb fragment containing *transposase* and the mini-Tn*10* cassette. The fragment was ligated with pMTL5401F, digested with *Sma*I, to create pKM1049 (B). Both plasmid constructs contain important features including: either inside (pKM1049) or outside (pKM1048) end based mini-Tn*10* cassette (green arrow) with associated *transposase* gene and chloramphenicol resistance (c*at*), erythromycin resistance (*ermB*), ampicillin resistance (*bla*), repH for replication in *C. difficile*, ColE1 for replication in *E. coli*, the promoter P_{fac} (fac) and the transfer gene *traJ*. Selected restriction sites are also highlighted.

4.4.2.3. Chemical transformation

DH5α and α-select (Bioline, UK) chemically competent *E. coli* were transformed with plasmid or ligation mixtures. Briefly: plasmid/ligation mixtures were added to 50 µl competent *E. coli* thawed on wet ice. Bacteria were incubated on ice for 45-60 min before being subjected to a heat shock at 42^{o}C for 45 seconds. Cells were returned to ice for 2 minutes before addition of 0.8 ml of pre-warmed SOC medium. Cultures were incubated at 37^{o}C with shaking (200 rpm) for 1 hour in 1.5 ml tubes. Bacteria were spread onto pre-warmed LB agar plates containing appropriate antibiotics before incubation at 37^{o}C for 16-18 hours.

4.4.2.4. Transformation by electroporation

Generation of electrocompetent CA434 bacteria: Bacteria were grown in 50-100 ml LB broth in 250-500 ml conical flasks at 37^{o}C with shaking to an OD of between 0.5-1.0. Cultures were chilled on ice for 30 minutes and centrifuged at 4000 x *g* for 15 min. Bacterial pellets were gently re-suspend in pre-chilled sterile water and centrifuged at 4000 x *g* for 15 min. Pellets were re-suspended in pre-chilled sterile 10% glycerol solution and centrifuged at 4000 x *g* for 15 min, pellets were then re-suspend in pre-chilled sterile 10% glycerol solution. Bacteria were divided into 50 µl aliquots and snap frozen in a dry ice and ethanol bath prior to storage at -80^{o}C.

Transformations using electrocompetent CA434 and ElectroSHOX *E. coli* (Bioline, UK) were carried out according to manufacturers guidelines (Bioline,

UK). Briefly 50 µl of *E. coli* was thawed on ice, 10-20 pmol of ligation mixture or plasmid was added and mixed gently. The transformation mixture was transferred to pre-cooled 0.1 cm gap electroporation cuvettes (Bio-Rad Laboratories, UK). Samples were placed in the electroporator and shocked under the following settings low range = 200 ohms, high range = 500 ohms, capacitance 25 µF, and 1.70 kV (Bio-Rad Laboratories, UK). 0.9 ml of pre-warmed SOC medium (Invitrogen, UK) was added to the bacteria followed by incubation at 37^{o}C with shaking (200 rpm) for 1-1.5 hours in 1.5 ml tubes. Bacteria were spread onto pre-warmed LB agar plates containing 50 µg ml^{-1} Kan, 100 µg ml^{-1} Ap, 12.5 µg ml^{-1} Cm, 25 µg ml^{-1} Cm or 50 µg ml^{-1} Cm as appropriate, with aerobic incubation at 37^{o}C for 18 hours. All transformations were accompanied by control transformation using pUC19 (supplied with the ElectroSHOX cells, Bioline, UK).

4.4.3. **Confirmation of constructs**

4.4.3.1. pKM1048

Twenty transformants were picked from the 12.5 or 25 µg ml^{-1} Cm selective LB plates and inoculated into 10 ml LB broth containing 12.5 µg ml^{-1} Cm. Cultures were incubated with shaking (200 rpm) for 16 hours at 37^{o}C. Plasmid was extracted from 3 ml of culture using a Mini-prep kit and eluted in molH_2O. Plasmids were assessed by restriction digest and PCR before the correct construct was selected and stored at -20 or -70^{o}C. Putative pKM1048 plasmids were digested with both *Hind*III and *Eco*RV in separate reactions. Approximately 400 to 800 ng of plasmid DNA was digested with

100 U (*Hind*III) or 80 U (*Eco*RV) of enzyme with appropriate 1X reaction buffer in a final volume of 25 µl for 16 hours at 37^{o}C. Digested DNA was separated on a 1% agarose gel with an appropriate marker; fragments were visualized on a transilluminator after staining with ethidium bromide.

Plasmid DNA was amplified using primer pair Tn10F/Tn10R (Table 4.2) in a standard PCR reaction containing 10-30 ng template, 2.5 U *Taq* DNA polymerase, 100 µM dNTPs, 1X standard *Taq* buffer, 0.1 µM of each primer (F & R) in a final volume of 50 µl. PCR conditions [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 50-58^{o}C 1.5 min, 72^{o}C 45 sec - 2 min] x 25-30 cycles [72^{o}C 10 min] x1 cycle, hold 4^{o}C. Samples were separated on a 0.8% agarose gel with an appropriate marker; bands were visualised on a transilluminator after staining with ethidium bromide. Desired constructs were stored at -20 or -70^{o}C.

4.4.3.2. pKM1049

Twelve transformants were picked from the 12.5 µg ml^{-1} Cm or 100 µg ml^{-1} Ap LB plates and inoculated into 10 ml LB broth containing 100 µg ml^{-1} Ap. Cultures were incubated with shaking (200 rpm) for 16 hours at 37^{o}C. Plasmid was extracted from 3 ml of culture using a Mini-prep kit and eluted in molH_2O. Plasmids were checked by digestion with *Eco*RV. Approximately 400 to 800 ng of plasmid DNA was digested with 80 U of *Eco*RV with 1X reaction buffer in a final volume of 25 µl at 37^{o}C for 16 hrs. Digested DNA was separated on a 1% agarose gel with an appropriate marker; bands were

visualized on a transilluminator after staining with ethidium bromide. Desired constructs were stored at -20 or -70°C.

4.4.4. **Plasmid sequencing**

Samples of each plasmid construct were sent for sequencing at the University of Cambridge where an Applied Biosystems 3730xl DNA Analyser was used to generate approximately 800 bp of sequence [385]. Electronic sequence results were analysed by eye using DNAMAN and by comparison to Vector NTI generated sequence files of the expected construct. In some cases the ClustalW2 alignment tool was used to compare sequencing data to expected sequence [386].

4.4.5. **Mating experiments**

Recipient *C. difficile* strains R20291, 630, 630Δerm and CD37 were grown in 10 ml BHI broth for 18-20 hours in anaerobic conditions at 37°C. These starter cultures were inoculated into 35-140 ml BHI broth (1% inoculum) and incubated without shaking in 50 ml universals (maximum 35 ml per tube) for 12-24 hours in anaerobic conditions at 37°C. Or starter cultures were inoculated into 35-50 ml BHI broth with 0.05% cysteine in 75 cm^3 vented tissue culture flasks and incubated with shaking (50 rpm) for 4-10 hours in anaerobic conditions at 37°C. If bacteria were grown with cysteine they were washed before mating: cells were centrifuged at 4000 x *g* for 10 min and re-suspended in 20 ml BHI broth, the process was repeated once.

Donor *E. coli* strains CA434-48 or CA434-49 were grown in 10 ml LB broth supplemented with 100 µg ml^{-1} Ap for 16-18 hours with shaking (200 rpm) at 37^{o}C. These starter cultures were used to inoculate (2%) 100 ml of LB containing 100 µg ml^{-1} Ap in a 500 ml conical flask. Cultures were incubated for 4-5 hours with shaking (200 rpm) at 37^{o}C.

E. coli cultures were centrifuged for 5 min at 5000 x *g* and washed twice with LB broth. *C. difficile* cultures were centrifuged at 3000 x *g* for 10-15 min in anaerobic conditions. The donor and recipient pellets were combined in anaerobic conditions and mixed gently but thoroughly with addition of 600 µl of BHI broth. The mating mixture was either a) spread over 3-15 cellulose nitrate filters (0.45 µm pore size, Sartorius stedim Biotech, UK) placed on pre-reduced BHIB agar; b) 15-25 µl of mating mixture was spotted directly onto the surface of pre-reduced BHIB agar. Plates were incubated for 20-24 hours in anaerobic conditions at 37^{o}C. Filters were transferred to sterile 20 ml universals and thoroughly washed with 0.6 ml of BHI broth, spot plates were washed with 1-1.2 ml of BHI broth. Cultures were then subjected to either a one-step or two-step selection protocol. One-step – culture was plated on 3-4 BHIBS selective plates containing 20, 50 or 100 µg ml^{-1} Tm per filter or spot mating plate. Plates were incubated for between 72 to 168 hours, and examined for transconjugants. Transconjugants were picked and re-streaked onto BHIBS agar plus 50 or 100 µg ml^{-1} Tm. Genomic DNA was extracted from the transconjugants and PCR was used to confirm the presence of *cat* and *amp* using the appropriate primers (Table 4.2). Two-step - culture was plated on BHIBS agar selective plates containing 100 µg

ml^{-1} Lm. Plates were incubated for between 72 to 168 hours, and examined for transconjugants at 24 hours intervals. Control plates were also included in initial experiments containing either *E. coli* or *C. difficile* alone and incubated alongside experimental plates. They were processed in the same manner as experimental plates post incubation. Transconjugants were replica plated onto BHIBS agar, incubated for 48 hours in anaerobic conditions before replica plating on BHIBS agar plus 100 µg ml^{-1} Tm and incubated for 48 hours in anaerobic conditions.

4.4.6. **PCR and Southern hybridisation**

4.4.6.1. PCR

PCR reactions were carried out using *Taq* DNA polymerase and standard *Taq* buffer as described previously under the following conditions: [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 50-58^{o}C 1.5 min, 72^{o}C 45 sec-3 min] x 25-35 cycles [72^{o}C 10 min] x1, hold 4^{o}C. Annealing temperatures and extension times were adjusted according to the primer set and expected product length (Table 4.2). Primers were designed with the aid of DNAMAN (Lynnon, USA), VectorNTI (Invitrogen, UK) and checked using SigmaAcustom oligonucleotide design/ordering online software (Sigma-Aldrich, UK) [387]. Figure 4.7 shows the location of primers used for probe generation.

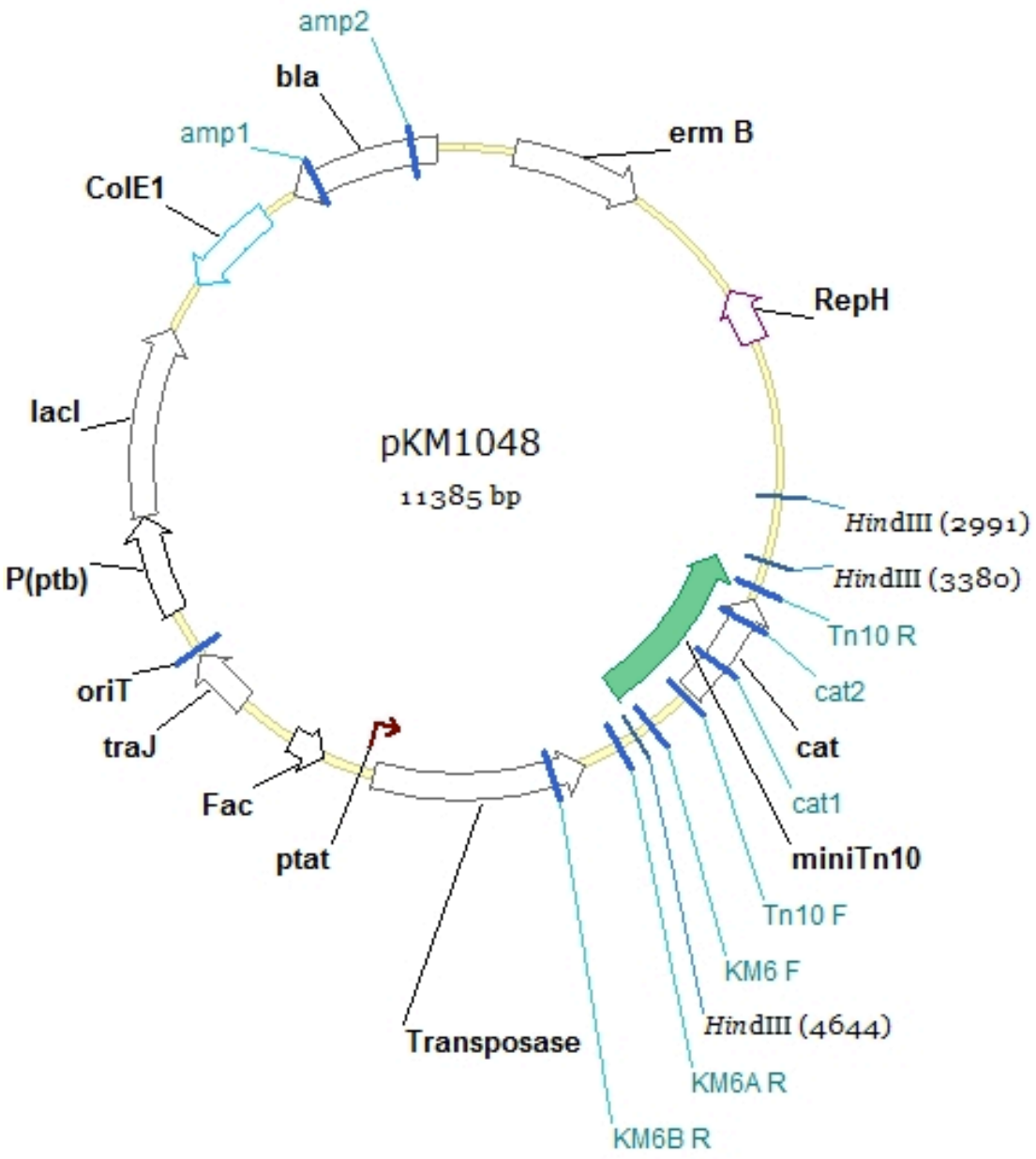


Figure 4.7 Location of primers on plasmid pKM1048.

Primer details can be found in Table 4.2.

4.4.6.2. Southern blotting and hybridisation

DNA was extracted from *C. difficile* transconjugants and digested with *Hind*III at 37^{o}C for 18 hours (see section 2.4.6.1). Approximately 300-500 ng of genomic DNA was digested with 200 U enzyme, approximately 250 ng of plasmid DNA was digested with 100 U of enzyme. Southern blotting was carried out according to the manufacturer's instructions in the ECL kit (GE Healthcare, UK) as described in section 2.4.6.3.

Four probes Pb-cat, Pb-catH, Pb-amp and Pb-Tn*10* were generated using the appropriate primers in a standard PCR reaction (Table 4.2). The PCR reaction was separated on a 1% agarose gel and the probe fragment were extracted using a gel extraction kit. The probes and ladder were labelled with the Amersham ECL direct nucleic acid labelling and detection system as described in section 2.4.6.5. Hybridisation and detection was carried out according to the methods described in section 2.4.6.6 (GE Healthcare, UK).

4.4.7. **Minimum inhibitory concentration (MIC) assay**

BHIB agar plates containing appropriate concentrations of Tm, Cm or Lm ranging from 0 to 1024 µg ml^{-1} were inoculated using a 36 pin automatic multipoint replicator (Mast Laboratories Ltd, Liverpool, UK) delivering approximately 1 x 10^{4} cfu spot^{-1} of 16-20 hour culture of *C. difficile* standardized to an OD600 nm of 0.8 prior to inoculation. Plates were incubated for 48 hours in anaerobic conditions at 37^{o}C. Biological replicates were performed in duplicate, experimental replicates were done in triplicate.

A control BHIB agar plate without antibiotics was inoculated with culture and incubated in aerobic conditions at 37^{o}C for 24 hours.

4.4.8. **Serial passage and replica plating**

C. difficile transconjugant cultures were grown for 24 hours in 10 ml BHI broth containing 100 µg ml^{-1} Tm at 37^{o}C in anaerobic conditions. 200 µl of this culture was used to inoculate 10 ml of fresh BHI broth containing 100 µg ml^{-1} Tm or BHI alone and grown for 24 hours as described above. This process was repeated 3 times. After the third passage a sample was taken and serially diluted 10-fold before onto BHIBS agar or BHIBS agar plus 100 µg ml^{-1} Tm. Plates were incubated for approximately 48 hours or until defined colonies were observed at 37^{o}C in anaerobic conditions. Plates were replica plated onto selective plates using sterile velvet squares. Three different replica plating experiments were performed; i) cultures grown without thiamphenicol selection and plated onto BHIBS plus 100 µg ml^{-1} Tm were replica plated onto BHIBS agar plus 100 µg ml^{-1} Tm and 100 µg ml^{-1} Lm; ii) cultures grown with thiamphenicol selection and plated onto BHIBS only agar were replicated onto BHIBS (control), BHIBS plus 100 µg ml^{-1} Tm and BHIBS plus 100 µg ml^{-1} Lm; iii) cultures grown with thiamphenicol selection and plated onto BHIBS plus 100 µg ml^{-1} Tm agar were replicated onto BHIBS plus 100 µg ml^{-1} Tm (control) and BHIBS plus 100 µg ml^{-1} Lm. All plates were incubated for 48 hours at 37^{o}C in anaerobic conditions before analysis.

4.4.9. **Promotion of transposition by induction of P_{fac} with IPTG**

10 ml BHI broth containing 100 µg ml^{-1} Tm and 100 mM IPTG were inoculated with a colony from four transconjugants (numbers 2, 7, 9 and 12), cultures were incubated for 24 hours in anaerobic conditions at 37°C. These cultures were then serially passaged (2% inocula) into BHI broth containing 100 µg ml^{-1} Tm and 100 mM IPTG 10 times. After passage 10 the cultures were serially diluted and plated on BHIBS agar containing 100 µg ml^{-1} Tm. 10 random colonies were picked and grown in BHI both containing 100 µg ml^{-1} Tm for 16-18 hours at 37°C. Genomic DNA was extracted (see section 2.4.6.1) and digested with *Hin*dIII, digests were separated on agarose gel using a suitable voltage. DNA was transferred to nitrocellulose membrane and probed with Pb-cat and Pb-amp, as described previously (section 2.4.6.3).

4.4.10. **PCR for determining site of integration**

PCRs were carried out using primers covering specific genes and regions of pKM1048. These included *bla*, *ermB*, *transposase*, and the P_{fac} region (Table 4.2). PCR conditions were as follows: 10-30 ng template, 2.5 U *Taq* DNA polymerase, 100 µM dNTPs, 1X standard *Taq* buffer, 0.1 µM of each primer (F & R) in a final volume of 50 µl. PCR conditions [94°C 4 min] x 1 cycle, [94°C 30 sec, 54°C 45 sec, 72°C 1 min] x 30 cycles, [72°C 10 min] x 1 cycle, hold 4°C. Samples were separated on 0.8-1% agarose gels with an appropriate marker; bands were visualised on a transilluminator after staining with ethidium bromide.

4.4.11. **Plasmid stability assay**

Three transconjugants were cultured independently in 10 ml BHI broth containing 100 µg ml^{-1} Tm for 16-18 hours in anaerobic conditions at 37°C. 10 ml BHI broth was inoculated with each culture and incubated for 16-18 hours at 37°C. Each culture was passaged, by 10-fold dilution (1 ml culture in 9 ml BHI broth) a further 4 times. Each culture from passage 1 to 5, including the original culture grown with antibiotic selection was serial diluted and plated onto BHIBS agar and BHIBS agar with 100 µg ml^{-1} Tm. Plates were incubated for 48 hours before enumeration of cfu ml^{-1}.

4.4.12. **Plasmid isolation from *C. difficile***

2 ml of 630Δerm::pKM1049 culture grown in 10 ml BHI broth with 100 µg ml^{-1} Tm for 16 hours at 37°C was centrifuged to harvest the bacteria. Bacteria were digested with 5 µl ml^{-1} lytic enzyme solution (Qiagen, UK) and 20 µg ml^{-1} lysozyme for 30 min at 37°C. Plasmid was extracted using a Mini-prep kit with plasmid eluted in 20 µl molH_2O. 10 µl of the elute was transformed into α-select chemical competent *E. coli* (Bioline, UK) (see section 4.4.2.3). Transformants were plated onto LB agar with 100 µg ml^{-1} Ap and incubated at 37°C for 16 hours. 5 to 10 colonies were selected and inoculated into 5 ml LB broth plus 100 µg ml^{-1} Ap; cultures were incubated at 37°C with shaking (200 rpm). Plasmid was extracted from 3 ml of culture using the Mini-prep kit with plasmid eluted in 30 µl molH_2O. 10 µl of this elute was digested with *Eco*RV for 18 hours at 37°C. The digest was separated on a 1% agarose gel with an appropriate marker.

4.4.13. **Measurement of transposase expression – RT-PCR**

4.4.13.1. Sample culture

R20291::pKM1048 was cultured in 10 ml BHI broth containing 100 μg ml^{-1} Tm for 18 hours without shaking at 37°C. This culture was used to inoculate 10 ml of BHI broth in vented 25 cm^3 tissue culture flasks containing 0.05% cysteine, 100 μg ml^{-1} Tm (control) or 10 ml BHI broth containing 0.05% cysteine, 100 μg ml^{-1} Tm and 100 μg ml^{-1} IPTG, to a final OD of 0.1 at 600 nm. Cultures were incubated with shaking (50 rpm) in aerobic conditions for 4-4.5 hours at 37°C.

4.4.13.2. RNA extraction

RNA was extracted as described previously using the Fast RNA pro blue extraction kit (see section 3.4.2.2). RNA was treated with DNA-free DNase as described previously (see section 3.4.2.3). Unless used immediately RNA was stored at -80°C.

4.4.13.3. Real-Time RT-PCR

RT-PCR (Reverse transcriptase-PCR) was carried out using the one-step Rotor-Gene SYBR Green PCR Kit (Qiagen, UK) according to the manufacturer's instructions. A master mix of 2X Rotor-Gene SYBR Green PCR Master Mix, 0.1X volume of Rotor-Gene RT Mix, 0.25-2.5 μM of forward and reverse primer was distributed to 0.1 ml tubes (Qiagen, UK). 1 to 100 ng of RNA was added and the volume adjusted to 10 μl using RNase-free water. RT-PCR for each sample was performed in triplicate and three independent

experiments were performed. Two primer pairs were designed against constitutively expressed genes *GyrA* and *16S* and two primer pairs against target genes *cat* and *transposase* by Sigma-Aldrich (UK) to generate products of approximately 100 bp (Table 4.2). The cycling conditions were as follows: Reverse transcription 10 min 55°C. PCR initial activation step 5 min 95°C, Two-step cycling-denaturation 5 sec at 95°C, combined annealing/extension 10 sec 58°C, Number of cycles 35–40 carried out using a Rotor-gene 3000 (Corbett Robotics, Australia). Control samples were included in each run these consisted of samples without template and samples without RT enzyme.

4.4.13.4. Reaction optimization

RT-PCR was carried out using 0.125 mM, 0.25 mM and 0.75 mM of GyrA and 16S primer pairs with varying quantities of template 1 ng, 10 ng and 100 ng. Results were analysed to determine the appropriate primer and template concentrations.

4.4.13.5. Calibration curve

700 ng of RNA was serially diluted 10-fold and added to standard optimized RT-PCR master mix with qGyrA, q16S or qTpase primers (at 0.25 mM) in triplicate. Rotor gene software was used to establish a threshold value for each reaction; the threshold value was then used to find the repCT value for the test RNA samples.

4.5. Results

4.5.1. **Vector construction and validation**

Plasmids pHV1248 and pHV1249, generated by Petit *et al.* (1990) were identified as suitable plasmids from which to obtain the mini-Tn*10* cassette and associated transposase [197]. Each plasmid contained a different version of the mini-Tn*10* containing either inside-end or outside-end IS*10* repeat sequences. Secondly the *E. coli-Clostridium* shuttle plasmid pMTL5401F (Figure 4.5) created by Heap *et al*. (2007) was chosen as the delivery plasmid as this is unstable in *C. difficile* [130, 388]. To create the desired construct, the mini-Tn*10* and associated transposase was cloned into the multiple cloning site (MCS) of pMTL5401F downstream of the P_{fac} promoter (a composite of the ferredoxin gene *fd* promoter fused to the operator of the lacZ operon). P_{fac} is reported to be inducible through the addition of IPTG however, this has not been determined in *C. difficile*.

4.5.1.1. pKM1048

Initially, the desired region was amplified by PCR using primers KM1F and KM1R designed to amplify the ~3 Kb region including the *transposase* gene with its promoter and ribosome binding site (RBS) and the complete mini-Tn*10* cassette (Figure 4.6). Amplification of the fragment did not generate a single product (data not shown). Attempts to optimise the amplification, including redesigning the primers and altering the reaction conditions, did not result in amplification of the expected product with multiple bands observed (data not shown).

In parallel, a wholly restriction digest-based construction of the plasmid was attempted. pHV1248 was digested with *Apa*LI and *Cla*I to isolate a 3.4 Kb region containing the *transposase* gene (including the promoter and RBS) and mini-Tn*10*. The fragment was cloned into the unique *Sma*I site of pMTL5401F to create pKM1048 (Figure 4.6). pKM1048 was transformed into competent *E. coil* and after growth, the plasmid was recovered and digested with *Eco*RV. Digestion with *Eco*RV enabled confirmation of the direction of the fragment insertion with respect to P_{fac} (Figure 4.6). The restriction profile of two plasmids revealed insertion of the fragment into the desired orientation (fragments of 7550 bp & 3835 bp) (Figure 4.8 lanes 1 & 4). PCR for the *cat* gene and sequencing of a section of the plasmid also confirmed the presence of mini-Tn*10* in these plasmids.

pKM1048 was transformed into chemical competent *E. coli* strain CA434 alongside two control transformations using pHV1248 and pUC19. Plasmid from fourteen colonies was digested with *Eco*RV with nine transformants showing an identical restriction fragment profile to the original plasmid (Figure 4.9 lanes 1, 6-12 &14). A *Hin*dIII digest was performed to confirm the presence of the mini-Tn*10* transposon in the construct; again all of the nine plasmids observed in the *Eco*RV digest were positive for mini-Tn*10* (Figure 4.10 lanes 1, 5-6, 8-12 & 14, note that due to an error plasmid in lane 5 and 7 were swapped in this gel). The expected fragment sizes were 9736, 1260 and 389 bp, however some of these fragments are very faint and are difficult to see. No plasmid is observed in lane three for either of the digests.

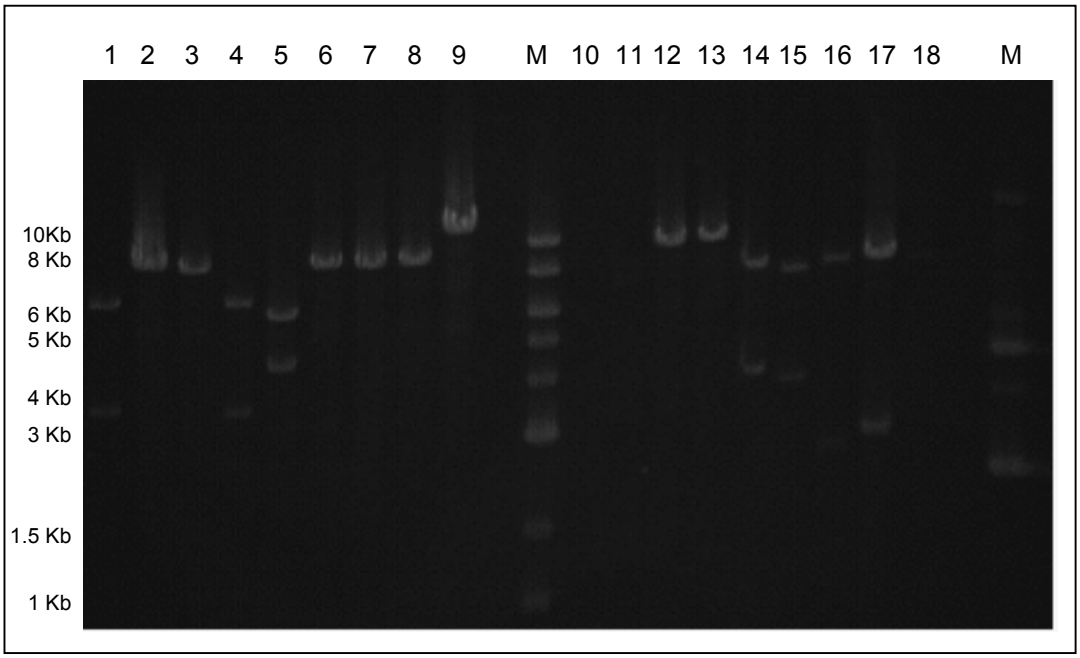


Figure 4.8 Digestion of putative pKM1048 plasmids.

Plasmid from 18 transformants was digested with *Eco*RV to identify plasmid constructs that contained the mini-Tn*10* insertion in the correct orientation. The digested plasmids in lanes 1 and 4 show the correct banding pattern with two fragments of 7550 and 3835 bp. M = 1 Kb ladder.

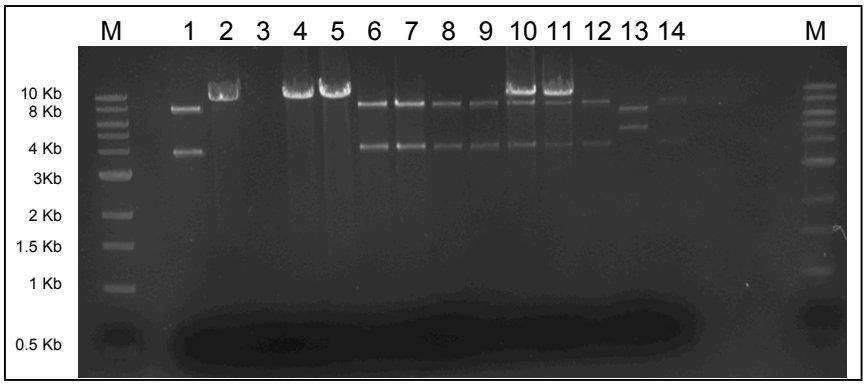


Figure 4.9 *Eco*RV digestion of plasmids from the transformation of CA434 with pKM1048.

Plasmids extracted from 14 transformants were digested with *Eco*RV to confirm the presence of pKM1048. 9 of the 14 plasmids have the expected restriction profile with two fragments of 7550 and 3835 bp in lanes 1, 6-12 and 14. No plasmid is seen in lane 3 and incomplete digestion of the plasmid is visible in lane 10 and 11. Plasmids in lane 2, 4 and 5 may represent incompletely digested plasmids or plasmid without the mini-Tn*10* insertion. The plasmid in lane 13 has an unexpected banding pattern. M = 1 Kb ladder.

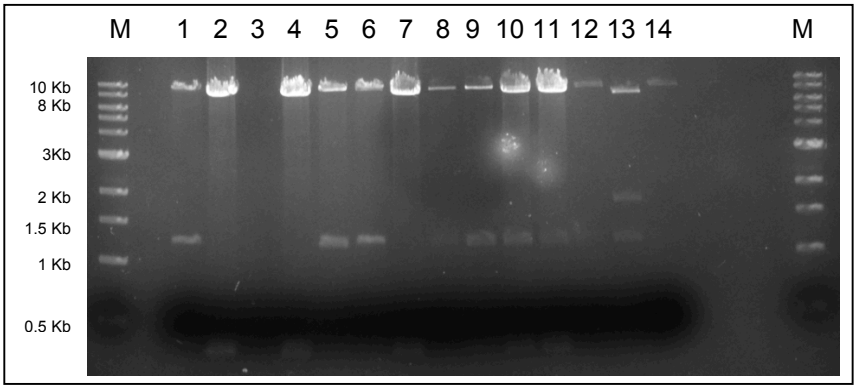


Figure 4.10 *Hin*dIII digestion of plasmids from the transformation of CA434 with pKM1048.

Plasmids extracted from 14 transformants were digested with *Eco*RV to confirm the presence of pKM1048. 9 of the 14 plasmids have the expected restriction profile with three fragments of 9736, 1260 and 389 bp in lanes 1, 5-6, 8-12 & 14. Note that due to a technical error plasmids in lanes 5 and 7 were swapped in this gel compared to the one for the *Eco*RV digest. As with the *Eco*RV digest no plasmid is seen in lane 3 and an unexpected banding pattern is seen for the plasmid in lane 13. The digest of plasmids in lane 2,4 and 7 reveals that the mini-Tn*10* fragment is not present in these plasmids. M = 1 Kb ladder.

4.5.1.2. pKM1049

pHV1249 was digested with *ApaL*I and *Hpa*I to isolate a fragment containing the *transposase* gene (including the RBS and promoter) and mini-Tn*10* (Figure 4.6). The fragment was cloned into pMTL5401F to generate pKM1049 and transformed into *E. coli* (Figure 4.6). Plasmid was extracted and digested with *Eco*RV to allow assessment of the orientation of the insertion (Figure 4.11 lane 9 & 12). Due to blunt ended ligation the fragment could ligate into the vector in two orientations with respect to P_{fac}. Four bands are expected in both cases, however the two high molecular weight bands were different depending on the orientation. One plasmid produced fragments of the expected size (6089 bp and 3835 bp) for insertion in the desired orientation (Figure 4.11 lane 9), and one plasmid produced fragments of the expected size (6609 bp and 3315 bp) for insertion in the opposite orientation (Figure 4.11 lane 12). Plasmids 9 (pKM1049-A) and 12 (pKM1049-B) were transformed into CA434 and confirmed by extraction and digestion with *Eco*RV (Figure 4.12).

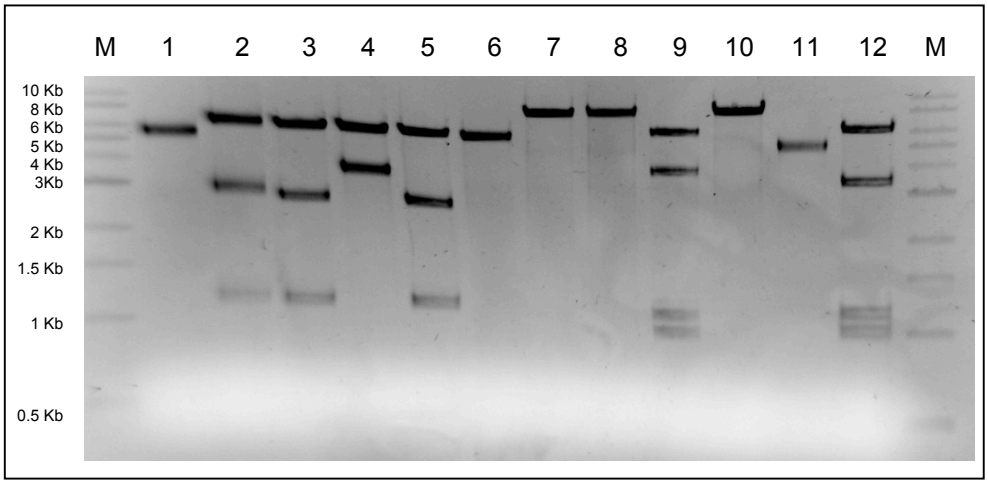


Figure 4.11 Agarose gel showing restriction digest fragments of putative pKM1049 transformants.

12 colonies were picked from Cm (lanes 1-6) or Ap (7-12) selective plates. Two colonies show the expected digest patterns indicating the insertion of the mini-Tn*10* fragment. Lane 9 contains plasmid with the insert ligated in the same direction of P_{fac} and the plasmid in lane 12 shows insertion in the opposite direction to P_{fac}. M = 1 Kb ladder.

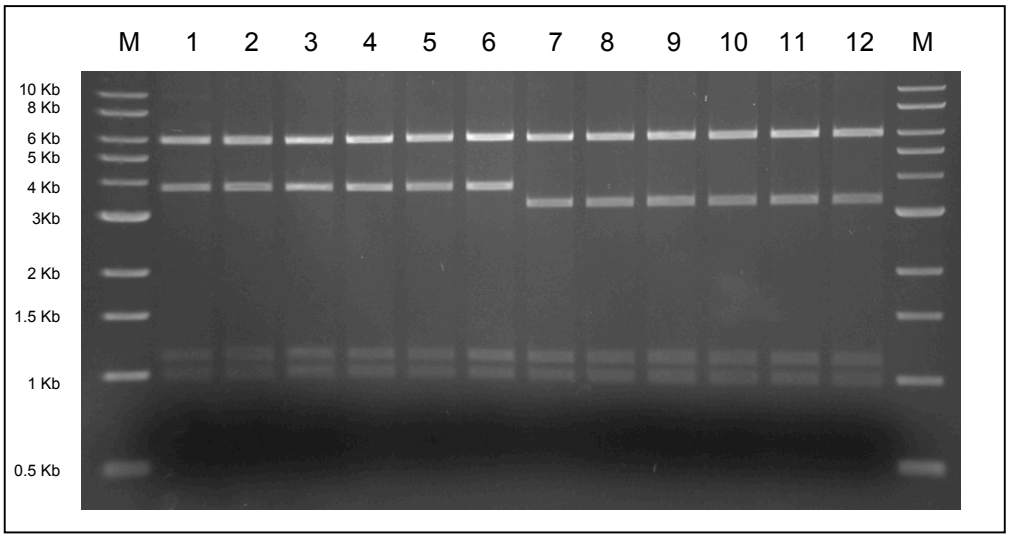


Figure 4.12 *Eco*RV digestion of plasmids pKM1049-A and pKM1049-B transformed into CA434

Six random colonies were picked from each transformation of pKM1049-A and pKM1049-B with electrocompetent CA434 *E. coli*. Plasmid DNA was extracted and digested with *Eco*RV before separation on 1% agarose gel with an appropriate marker. M = 1 Kb ladder, lane 1-6 = pKM1049-A transformants; 7-12 pKM1049-B transformants.

4.5.2. **Mating of CA434-48 and CA434-49 with *C. difficile***

4.5.2.1. CA434-48

CA434-48 (CA434 *E. coli* strain containing plasmid pKM1048) was mated with *C. difficile* strains 630, CD37, 630Δerm and R20291. Between two and seven independent mating experiments were performed depending on the strain. In one experiment strains were grown with the addition of cysteine and with shaking. Viable counts were conducted for each experiment, and the ratio of *C. difficile* to *E. coli* was calculated (ranging from 1:1 to 1:20) (Table 4.3).

In total nine independent mating experiments were performed with various strains, the appropriate level of thiamphenicol (Tm) for transconjugant selection was not known and therefore various levels were tested to ensure minimal background and maximum transconjugant retrieval. In the first experiment one putative transconjugant arose from the R20291 recipient and six arose from the 630 recipient. PCR initially confirmed the presence of the *cat* gene, however after subculture in nonselective and low level Tm selective broth (10 μg ml^{-1}) all of the samples were PCR-negative for *cat*. Freezer stocks were made for all of these putative transconjugants after the initial PCR positive result. In experiments two, four, five, six, seven and nine no transconjugants were obtained.

In the third mating experiment with 630Δerm, R20291 and CD37 transconjugants were initially selected with 20 or 50 μg ml^{-1} Tm. On the plates with 20 μg ml^{-1} Tm a large number of pin-prick colonies were observed

whereas larger defined colonies were observed on 50 μg ml^{-1} Tm selective plates. A total of fourteen putative transconjugants were picked and subcultured onto 20 and 50 μg ml^{-1} Tm selective plates. Of these, eight R20291 putative transconjugants only grew on 20 μg ml^{-1} Tm, three 630Δerm and three CD37 putative transconjugants grew on both 20 and 50 μg ml^{-1} Tm. All of these putative transconjugants were PCR positive for the presence of the *cat* and *bla* gene. The conjugation frequency per donor in these experiments was 5.1 x 10^{-10} for CD37 and 630Δerm, and 1.37 x 10^{-9} for R20291.

In experiment seven an alternate approach was used for selection of transconjugants, selection on Lm rather than Tm directly selected for the plasmid backbone rather than mini-Tn*10.* In experiment eight fourteen putative transconjugants were generated in strain R20291 from selection in 100 μg ml^{-1} Tm. The conjugation efficiency was 1.84 x 10^{-10} per CA434-48 donor and all of the transconjugants were PCR positive for the *cat* and *bla* genes.

Experiment	Recipient strain	No. of filters	Total *E. coli* cells per filter (approximately)	Total *C. difficile* cells per filter (approximately)	Approximate ratio *C. difficile* to *E. coli*	Level of selection μg ml^{-1}	No. of transcon jugants	Conjugation efficiency (per donor)	Supplementary info
1	R20291	5	2.10E+09	2.24E+08	1:10	Tm 20	1	9.52E-11	
	630	5	2.10E+09	4.20E+08	1:5	Tm 20	6	5.71E-10	
2	630Δerm	3	6.10E+07	6.93E+06	1:10	Tm 50	ND	ND	
	R20291	3	6.10E+07	8.13E+06	1:10	Tm 50	ND	ND	
	CD37	3	6.10E+07	5.87E+07	1:1	Tm 50	ND	ND	
3	630Δerm	3	1.95E+09	3.03E+08	1:7	Tm 20 + 50	3	5.13E-10	
	R20291	3	1.95E+09	4.47E+08	1:4	Tm 20 + 50	8	1.37E-09	
	CD37	3	1.95E+09	1.89E+08	1:10	Tm 20 + 50	3	5.13E-10	
4	630Δerm	10	6.15E+09	7.46E+08	1:10	Tm 100	ND	ND	
	R20291	10	6.15E+09	3.15E+08	1:20	Tm 100	ND	ND	
5	630Δerm	6	3.33E+09	1.63E+09	1:2	Tm 100	ND	ND	
	CD37	6	3.33E+09	7.29E+09	2:1	Tm 100	ND	ND	

Experiment	Recipient strain	No. of filters	Total *E. coli* cells per filter (approximately)	Total *C. difficile* cells per filter (approximately)	Approximate ratio *C. difficile* to *E. coli*	Level of selection μg ml^{-1}	No. of transcon jugants	Conjugation efficiency (per donor)	Supplementary info
6	R20291	5	2.28E+09	3.57E+08	1:6	Tm 100	ND	ND	
	CD37	5	2.28E+09	2.59E+09	1:1	Tm 100	ND	ND	
7	CD37	3	8.22E+09	3.09E+09	1:3	Lm 100	ND	ND	
	R20291	6	8.22E+09	2.18E+09	1:4	Lm 100	ND	ND	
8	R20291	15	5.07E+09	2.11E+09	2:5	Tm 100	14	1.84E-10	
9	R20291	10	4.1E+09	9.2E+08	1:4	Tm 100	ND	ND	Growth with additional cysteine
	630Δerm	5	4.1E+09	2.3E+09	1:2	Tm 100	ND	ND	

Table 4.3 Mating experiments between *C. difficile* and CA434-48.

In total 9 filter mating experiments were performed with strains CD37, 630, 630Δerm and R20291 and CA434-48. Transconjugants were generated in each of these strains. In initial experiments low levels of Tm selection or no continued selection led to loss of the plasmid. 14 transconjugants generated in experiment 8 in the R20291 recipient were maintained for further analysis. The presence of plasmid pKM1048 was confirmed by PCR for each transconjugant. For each experiment the number and ratio of cells mated is shown alongside the level of Tm selection and conjugation efficiency. ND = Not detected.

4.5.2.2. CA434-49

CA434-49 (*E. coli* strain CA434 containing plasmid construct pKM1049-A) was mated with *C. difficile* strains 630Δerm and R20291 in filter and spot mating. Four independent filter mating experiments were carried out for both *C. difficile* strains (Table 4.4). In each experiment the ratio of *C. difficile* to *E. coli* has been calculated, ranging from 1:10 to 4:1. All but one of the mating experiments failed to generate transconjugants. Fourteen transconjugants were generated in experiment three in the 630Δerm recipient strain. Spot mating was performed a total of five times with mating mixtures spotted directly onto agar rather than filters (Table 4.5). In experiments one to four no transconjugants were observed. In experiment five a 2-step selection procedure was attempted whereby the mating mixture was inoculated onto Lm rather than Tm agar. No transconjugants were observed.

Experiment	Recipient strain	No. of filters per strain	Total *E. coli* cells per filter	Total *C. difficile* cells per filter	Approximate ratio *C. difficile* to *E. coli*	No. of transconjugants	Conjugation efficiency (per donor)
1	R20291	10	1.50E+09	7.00E+08	1:2	ND	ND
	630Δerm			1.95E+09	1:1	ND	ND
2	R20291	15	5.65E+09	4.34E+08	1:10	ND	ND
	630Δerm			5.93E+09	1:1	ND	ND
3	R20291	5	1.05E+09	7.70E+08	7:10	ND	ND
	630Δerm			1.07E+09	1:1	14	1.33E-08
4	R20291	5	1.05E+09	3.05E+08	1:3	ND	ND
	630Δerm			9.45E+08	1:1	ND	ND

Table 4.4 Results of filter mating between CA434-49 and *C. difficile* strains R20291 and 630Δerm.

4 filter mating experiments were performed for each strain. Transconjugants were selected on 100 μg ml^{-1} Tm. For each experiment viable counts were taken, the approximate ratio of donor to recipient cells is shown. 14 transconjugants were generated in experiment 3 for the recipient 630Δerm. ND = Not detected

Experiment	Recipient strain	No. of filters per strain	Total *E. coli* cells per filter	Total *C. difficile* cells per filter	Approximate ratio *C. difficile* to *E. coli*	No. of putative transconjugants	Conjugation efficiency (per donor)
1	R20291	4	7.50E+08	1.38E+09	2:1	0	0
	630Δerm			2.75E+09	4:1	0	0
2	R20291	5	1.21E+09	1.16E+09	1:1	0	0
	630Δerm			9.50E+08	4:5	0	0
3	R20291	5	1.05E+09	7.70E+08	3:4	0	0
	630Δerm			1.07E+09	1:1	0	0
4	R20291	5	1.92E+09	5.39E+08	1:4	0	0
	630Δerm			5.81E+08	1:3	0	0
5	R20291	5	1.05E+09	3.05E+08	1:3	0	0
	630Δerm			9.45E+08	1:1	0	0

Table 4.5 Results of spot mating between CA434-49 and *C. difficile* strains R20291 and 630Δerm.

5 spot mating experiments were performed for each strain. In experiment 1-4 transconjugants were selected on 100 μg ml^{-1} Tm, for experiment 5 transconjugants were selected on 100 μg ml^{-1} Lm. In each experiment viable counts were taken and the approximate ratio of donor to recipient cells calculated. No transconjugants were generated from the spot mating experiments. ND = Not detected

4.5.3. **Analysis of transconjugants**

4.5.3.1. Minimum inhibitory concentrations (MIC) of *C. difficile* to various antibiotics

Both strain 630 and R20291 are resistant to Em, strain 630 has the Em resistance determinant *ermB*, but this is not present in R20291. Em belongs to the MLS_B class of antibiotics which includes Lm and clindamycin (Cli), it has been documented that some clinical isolates show high resistance to Em but not to other MLS_B compounds such as Cli [389]. Investigation of these strains revealed a nucleotide substitution within the 23S rDNA copy on the genome, this is likely to have resulted in resistance to Em. Therefore, it may be possible to use Lm or Cli to select for R20291 strains, which have acquired pKM1048/49 by conjugation. To investigate this, the MIC for these antibiotics was established for the recipient strains (Table 4.6). As anticipated, strain R20291 shows a comparable level of resistance to Em as strain 630 at >1024 μg ml^{-1} (the highest concentration tested). R20291 is sensitive to Lm at a significantly lower concentration than strain 630, 32 μg ml^{-1} and 512 μg ml^{-1}, respectively. Similarly, R20291 is sensitive to Cli at concentrations above 4 μg ml^{-1}, whereas strain 630 is sensitive to concentrations above 256 μg ml^{-1}. The profile of 630Δerm is similar to that of R20291 except that this strain was almost completely sensitive to Em (<2 μg ml^{-1}).

	Strain				
Antibiotic	CD37	630	630Δerm	R20291	R20291::pKM1048
Em	< 2 μg ml^{-1}	>1024 μg ml^{-1}	< 2 μg ml^{-1}	>1024 μg ml^{-1}	nt
Lm	32 μg ml^{-1}	512 μg ml^{-1}	32 μg ml^{-1}	32 μg ml^{-1}	100 μg ml^{-1}*
Cli	16 μg ml^{-1}	256 μg ml^{-1}	4 μg ml^{-1}	4 μg ml^{-1}	nt
Tc	< 2 μg ml^{-1}	128 μg ml^{-1}	128 μg ml^{-1}	< 2 μg ml-1	nt
Cm	8 μg ml^{-1}	4 μg ml^{-1}	8 μg ml^{-1}	< 2 μg ml^{-1}	>1024 μg ml^{-1}
Tm	4 μg ml^{-1}	4 μg ml^{-1}	4 μg ml^{-1}	4 μg ml^{-1}	>1024 μg ml^{-1}

Table 4.6 MICs for *C. difficile* strains and transconjugants to antibiotics.

The MIC of strains CD37, 630, 630Δerm and R20291, and 6 R20291::pKM1048 transconjugants to the MLS_B group antibiotics erythromycin (Em), lincomycin (Lm) and clindamycin (Cli); and tetracycline (Tc), chloramphenicol (Cm) and thiamphenicol (Tm) were established using the agar dilution method. The MICs given represent the average of three independent experiments. * = a full MIC was not completed for the transconjugants to Lm, only concentrations of 20, 50 and 100 μg ml-1 were examined. nt = not tested.

The resistance of all the strains to Tc, Cm and Tm was also tested. Strain CD37 and R20291 were sensitive to the lowest concentration of Tc tested (2 μg ml^{-1}), however, 630 and 630Δerm were both sensitive to concentrations above 128 μg ml^{-1}. All *C. difficile* strains were sensitive to Cm and Tm at concentrations above 8 μg ml^{-1}. Assessment of 13 R20291 transconjugants containing pKM1048 to Lm revealed that presence of the *ermB* gene on the plasmid conferred resistance up to 100 μg ml^{-1}. For this assay only concentrations of 20, 50 and 100 μg ml^{-1} were tested. The MIC of the transconjugants to Cm and Tm revealed that the presence of the *cat* gene conferred resistance to concentrations over 1024 μg ml^{-1}.

4.5.3.1. Analysis of R20291::pKM1048 transconjugants

PCR was carried out on each of the putative transconjugants to identify the *cat* in mini-Tn*10* and *bla* on the plasmid backbone (data not shown). Putative transconjugants from experiment 1, 3 and 8 were all PCR positive for both the *cat* and *bla* gene upon initial isolation. All PCR positive putative transconjugants were subjected to Southern hybridisation. *EcoR*V digests were blotted onto nitrocellulose membrane and probed using Pb-cat to identify the presence of mini-Tn*10* (Figure 4.13). The blot shows no hybridisation of Pb-cat to the wild type DNA and positive hybridisation to pKM1048 plasmid DNA. Positive hybridisation is seen at approximately 10 Kb in lane 1, 2 and 4.

In a second blot, PCR positive transconjugants from experiment three were assayed. DNA digested with *Eco*RV from nine transconjugants, three from each of the recipient strains CD37, R20291, and 630Δerm were probed with Pb-cat (Figure 4.14). A putative 630 transconjugant previously probed in first blot

(Figure 4.13, lane 4) was also included due to the poor quality of the original blot. Hybridisation of the probe is seen for this transconjugant at approximately 9 Kb. There is no hybridisation of Pb-cat to any of the remaining transconjugant samples, however positive hybridisation is seen to pKM1048.

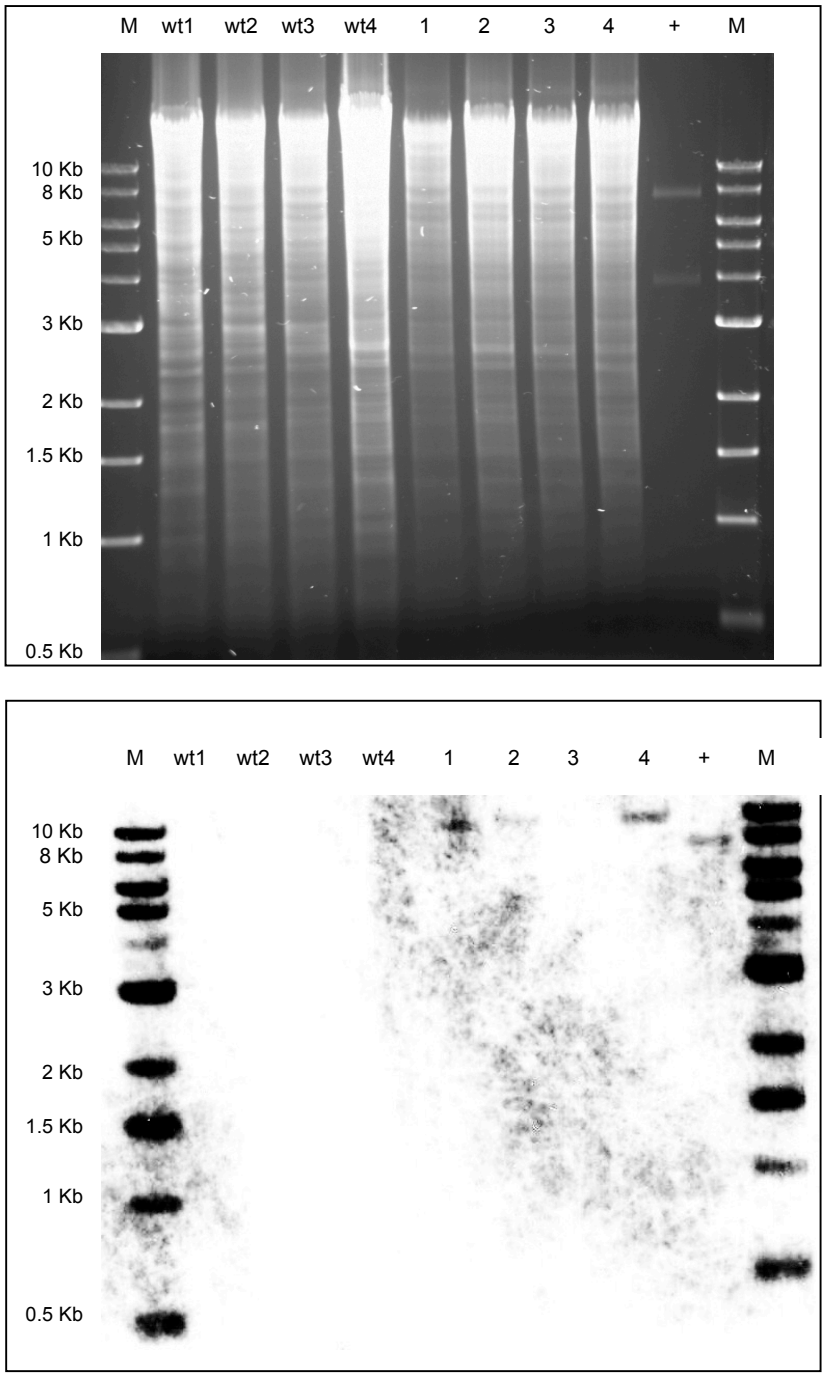


Figure 4.13 Digest of putative strain 630 transconjugants and Southern blot probed with Pb-cat

DNA from 4 putative 630::pKM1048 transconjugants was digested with *Eco*RV and transferred to a nitrocellulose membrane. The membrane was probed with labelled Pb-cat and exposed to photographic film for 15 hours. Faint bands of hybridisation can be seen at approximately 10 Kb for sample 1, 2 and 4. Faint hybridisation is also seen for the positive control (pKM1048) at approximately 8 Kb. No hybridisation of the probe is seen for any of the recipient strains. Lane M = 1 Kb ladder, + = pKM1048; wt1 = wild type CD37, wt2 = wild type R20291, wt3 = wild type 630, wt4 = wild type 630Δerm, lane 1-4 = 630 putative transconjugants.

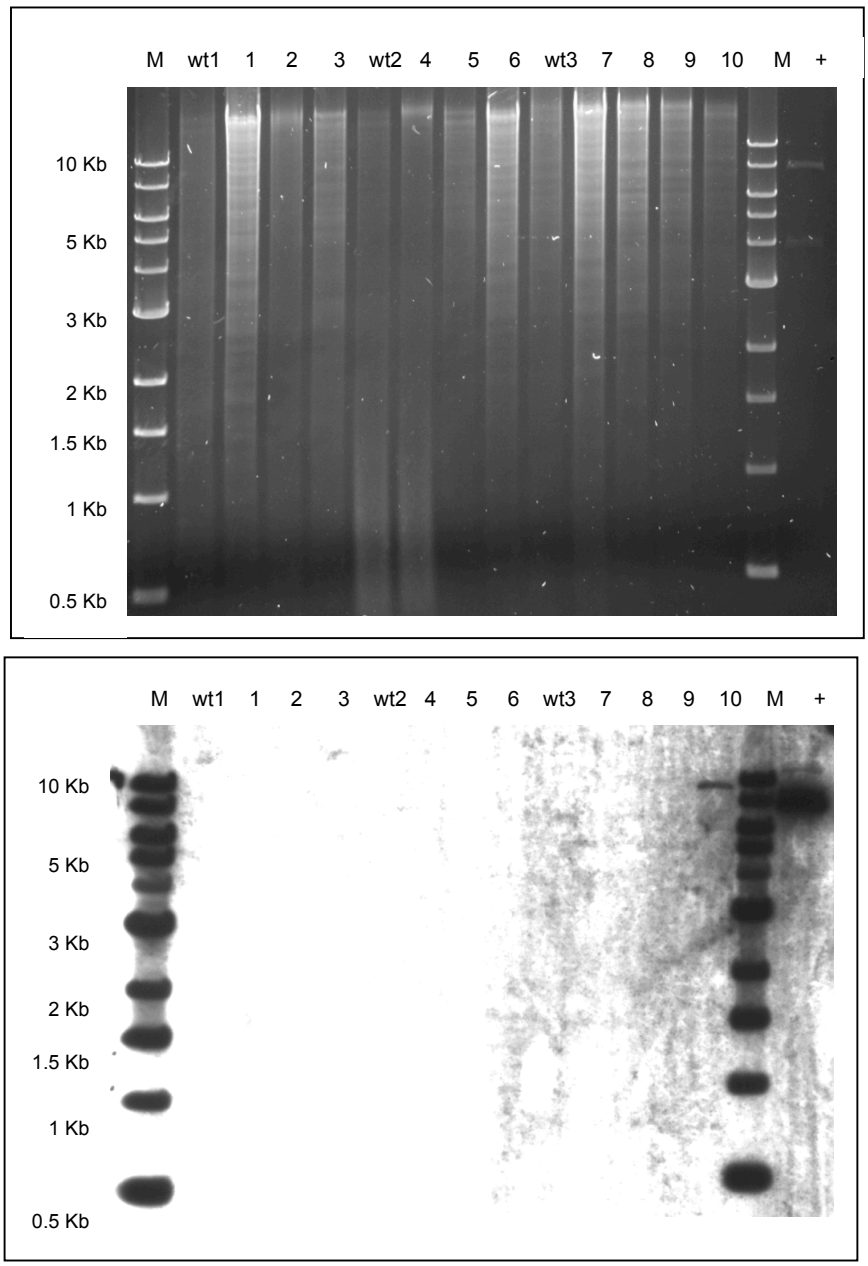
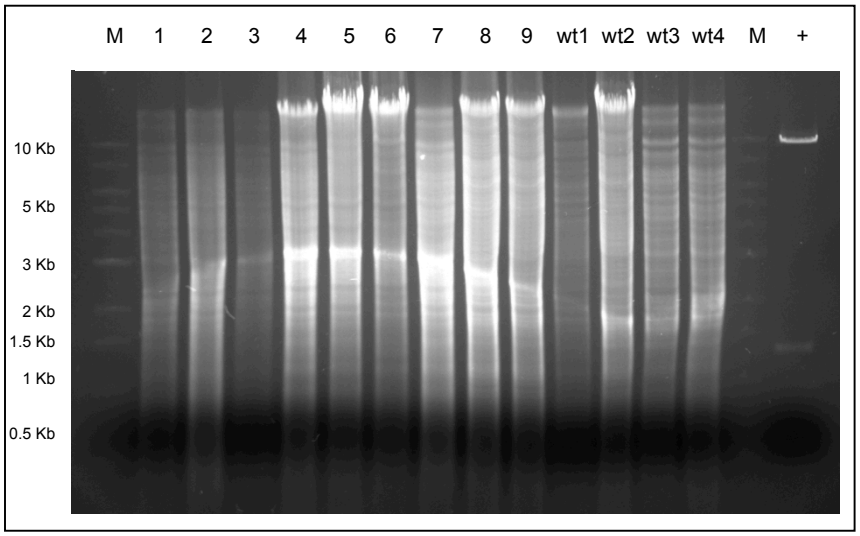


Figure 4.14 Digest and Southern hybridisation of putative transconjugants.

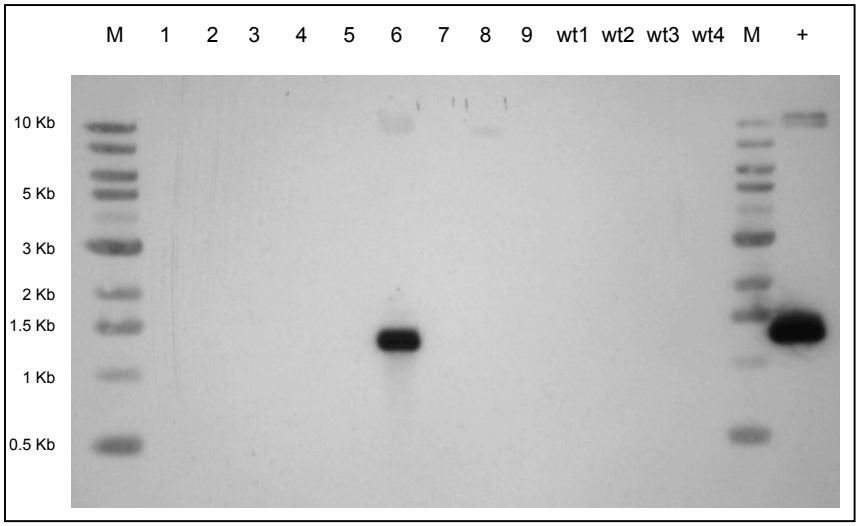
DNA digested with *Eco*RV was transferred to a nitrocellulose membrane and probed with Pb-cat. Positive hybridisation is seen for the plasmid control at approximately 8 Kb. No hybridisation of the probe is seen for any of the wild type strains. Positive hybridisation is seen in lane 9 at approximately 10 Kb. Lane M = 1 Kb ladder; + = pKM1048; wt1 = wild type CD37, wt2 = wild type R20291, wt3 = 630Δerm, lane 1-3 = CD37 putative transconjugants; lane 4-6 = R20291 putative transconjugants; 7-9 = 630Δerm putative transconjugants; 10 = 630 putative transconjugant that showed positive hybridisation in the first blot (Figure 4.13).

DNA from transconjugants generated in experiment one and three was digested with *Hin*dIII, including six R20291, one CD37 and two 630 recipient strains, and subject to Southern hybridisation (Figure 4.15A). The membrane was probed with two probes, Pb-cat for the *cat* gene on mini-Tn*10*, and Pb-amp for the *bla* gene on the plasmid backbone (Figure 4.15B & C). One of the putative transconjugants (Figure 4.15 B & C lane 6) showed hybridisation of both probes indicating the presence of *cat* and *amp*. The DNA is derived from a transconjugant in the CD37 recipient generated in experiment three. Hybridisation of Pb-cat to the positive control occurs on a fragment of approximately 1.2 Kb, this corresponds to the mini-Tn*10* fragment which is flanked by *Hin*dIII restriction sites. Similarly, hybridisation of Pb-amp to DNA from the positive control occurs on a fragment of approximately 10 Kb, this corresponds to the second fragment generated from *Hin*dIII digestion of pKM1048 (Figure 4.6A). An identical hybridisation pattern is seen in lane 6 for both probes. No hybridisation is observed to DNA from the recipient strains.

A


M 1 2 3 4 5 6 7 8 9 wt1 wt2 wt3 wt4 M +
10 Kb
5 Kb
3 Kb
2 Kb
1.5 Kb
1 Kb
0.5 Kb

B


M 1 2 3 4 5 6 7 8 9 wt1 wt2 wt3 wt4 M +
10 Kb
5 Kb
3 Kb
2 Kb
1.5 Kb
1 Kb
0.5 Kb

C

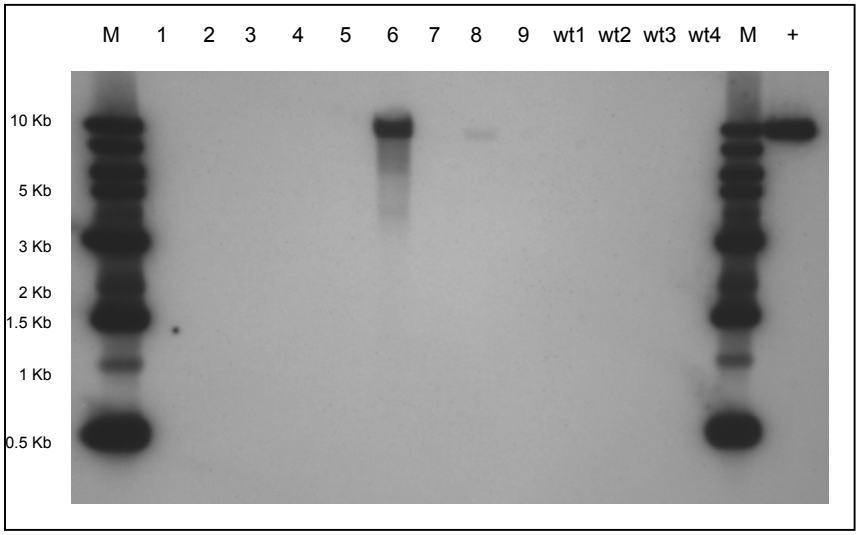

M 1 2 3 4 5 6 7 8 9 wt1 wt2 wt3 wt4 M +
10 Kb
5 Kb
3 Kb
2 Kb
1.5 Kb
1 Kb
0.5 Kb

Figure 4.15 Digest of putative transconjugant DNA from experiment 1 and 3 on agarose gel (A) and hybridisation of Pb-cat (B) and Pb-amp (C).

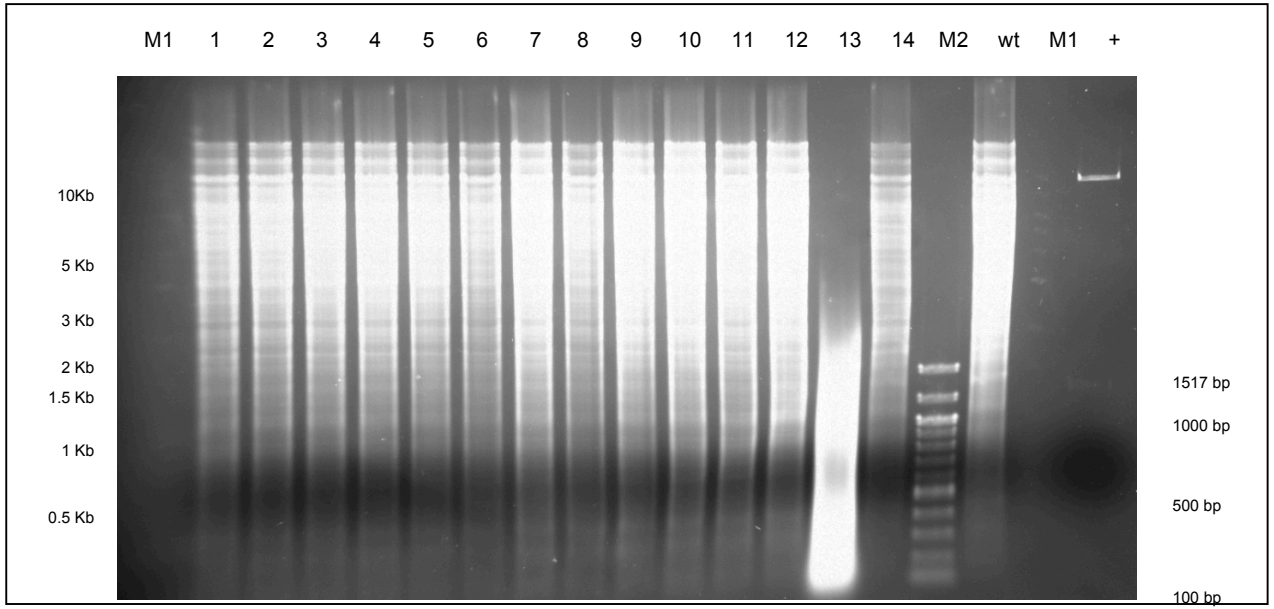
Genomic DNA was digested with *Hin*dIII and separated on a 1% agarose gel (figure A). DNA was transferred to a nitrocellulose membrane and probed with Pb-cat (figure B) or Pb-amp (figure C). Positive hybridisation for Pb-cat and Pb-amp is seen in lane 6 at approximately 1.2 Kb and 10 Kb, respectively. Positive hybridisation is also seen for the plasmid positive control in lane +. Lane M = 1 Kb ladder, + = pKM1048, lane 1-5 = R20291 putative transconjugants, lane 6 = CD37 putative transconjugant, lane 7= R20291 putative transconjugant (revived from experiment 1), lane 8-9 = 630 putative transconjugants (revived from experiment 1). wt1 = wild type CD37, wt2 = wild type 630, wt3 = R20291 wt4 = R20291 (alternative stock)

*Hin*dIII digested DNA from fourteen R20291 putative transconjugants from experiment eight was probed with Pb-cat and Pb-amp (Figure 4.16). The DNA extracted from one of the transconjugants (lane 13) is degraded and should be ignored. The remaining thirteen transconjugants were positive for both Pb-cat and Pb-amp hybridisation (Figure 4.16 B & C respectively). For Pb-cat, strong hybridisation can be seen at approximately 1.2 Kb identical to the plasmid control and for Pb-amp, hybridisation is seen at approximately 10 Kb, again identical to that seen for the plasmid.

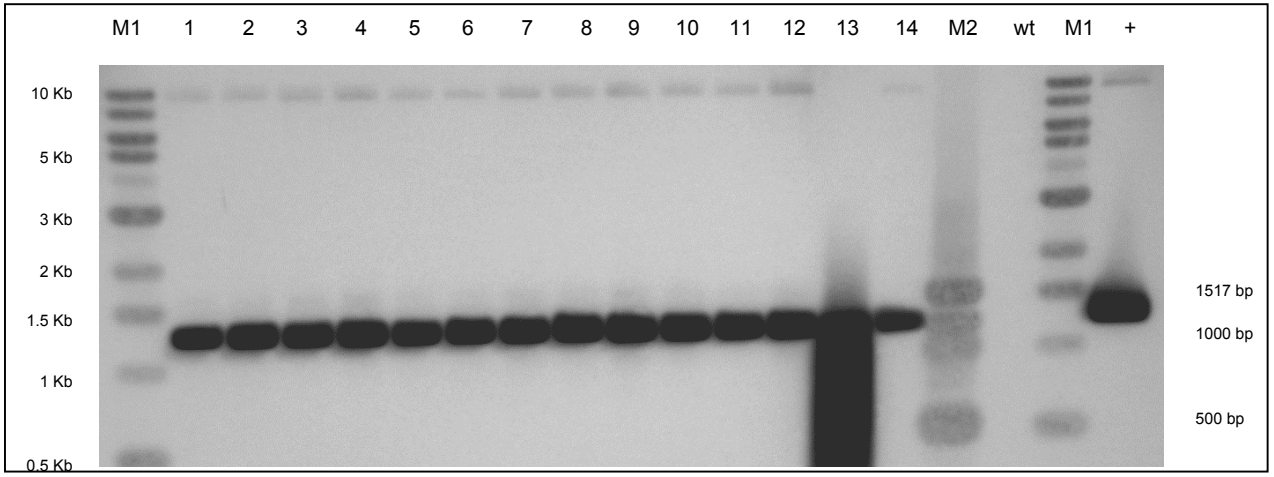
4.5.3.1. 630Δerm::pKM1049 transconjugants

The presence of mini-Tn*10* (*cat* gene) and the plasmid backbone (*bla* gene) was confirmed by PCR for all fourteen of the putative 630Δerm transconjugants generated in experiment three (Figure 4.17). Southern hybridisation was not performed using DNA from these transconjugants. The positive band seen in the wild type control lane at approximately 800 bp (lane wt) for amplification of the *bla* gene is a non-specific product and does not represent a positive result.

A


M1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 M2 wt M1 +
10Kb
5 Kb
3 Kb
2 Kb
1.5 Kb
1 Kb
0.5 Kb
1517 bp
1000 bp
500 bp
100 bp

B


M1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 M2 wt M1 +
10 Kb
5 Kb
3 Kb
2 Kb
1.5 Kb
1 Kb
0.5 Kb
1517 bp
1000 bp
500 bp

C

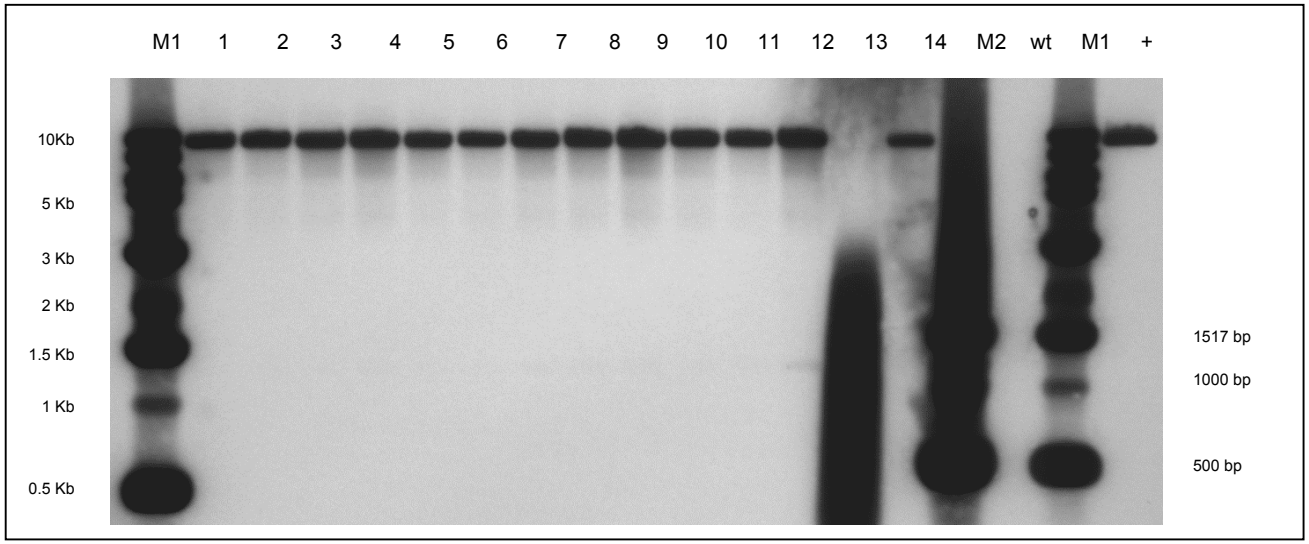
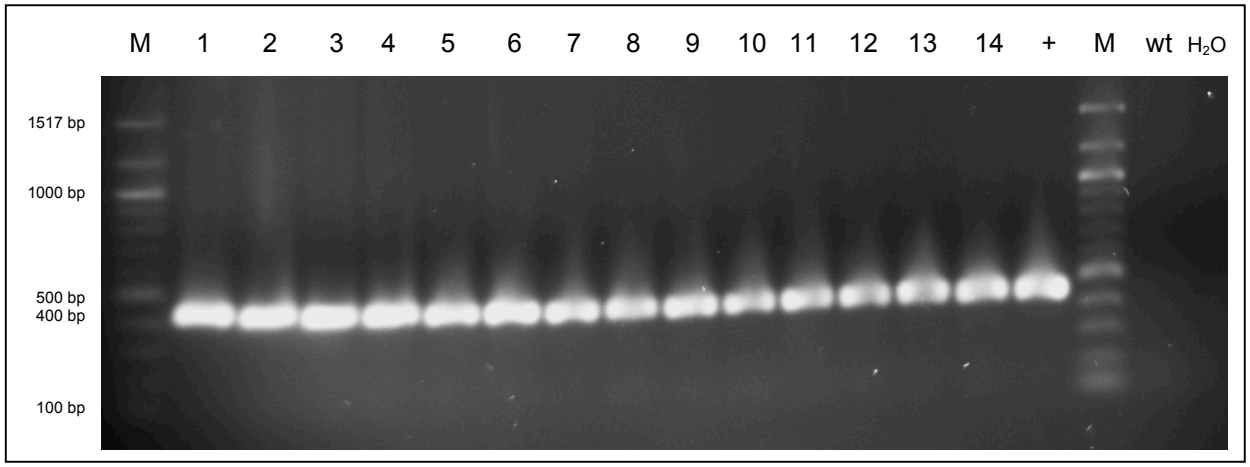


Figure 4.16 Restriction digest and Southern blot of putative R20291 transconjugants.

Figure A shows the *Hin*dIII digestion of DNA from 14 putative R20291 transconjugants containing pKM1048 separated on agarose gel. The DNA was transferred to a nitrocellulose membrane and probed with Pb-cat (figure B) and Pb-amp (figure C). Thirteen of fourteen putative transconjugants were positive for both *cat* (B) and *bla* (C). The DNA in lane 13 appears to have degraded producing a large smear on the gel and after hybridisation, this transconjugant was not subjected to any further analysis. Lane M1 = 1Kb ladder; M2 = 100 bp marker; 1-14 = R20291 putative transconjugants; + = pKM1048; wt = wild type R20291.

A



B

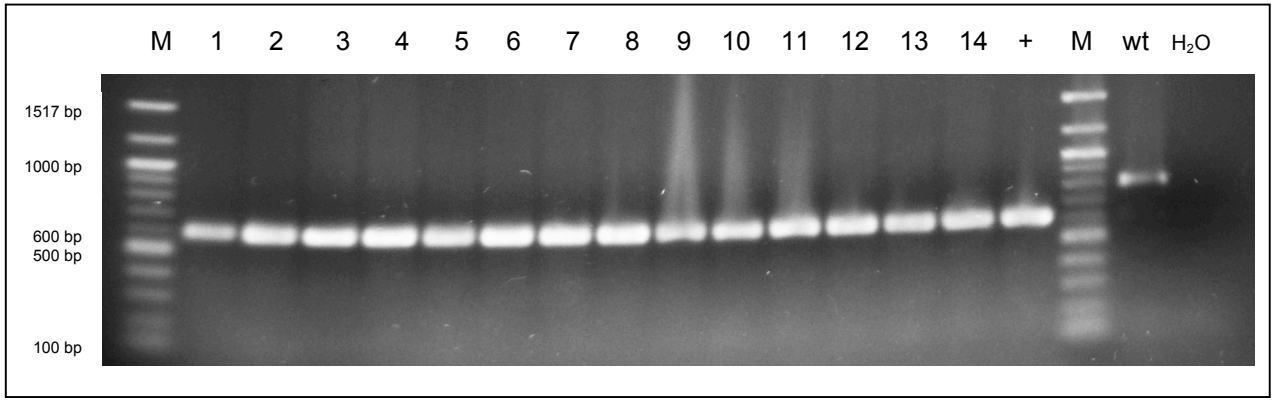


Figure 4.17 PCR amplification of 630Δerm putative transconjugants for presence of the *cat* gene(A) and *bla* gene (B).

Genomic DNA extracted from each transconjugant was used as template for a PCR reaction to amplify *cat* and *bla* to confirm the presence of both genes. PCR products were separated on 1% agarose gel. Lane M = 100 bp ladder; lane 1-14 = 630Δerm putative transconjugants; + = pKM1049; wt = wild type 630Δerm; H_2O = H_2O control.

4.5.4. **Investigation of plasmid fate in *C. difficile* transconjugants**

4.5.4.1. Sequence and PCR analysis

The hybridisation results provide evidence of the presence of *cat* in the *C. difficile* transconjugants, which was supported by resistance to Tm. However, *bla* is also present, indicating maintenance of the plasmid. Two possible explanations could explain this result, either the plasmid has been maintained in its original form or the plasmid has been integrated by homologous recombination into the genome of R20291. The plasmid sequence was analysed for regions of homology to the R20291 or 630 genome. Blastn analysis against the genome sequence of R20291 (Accession No. NC_013316) showed no regions of homology suggesting that recombination was unlikely to have occurred. Analysis of strain 630 (Accession No AM180355 (chromosome) & AM180356 (plasmid)) again revealed no sequence similarities except for the predicted identity to *ermB*. PCR was also carried out to assess whether regions of the plasmid had been broken to facilitate homologous recombination. PCRs covering the entire *bla*, *ermB*, *transposase* genes and P_{fac} revealed that all of these regions were intact (data not shown).

4.5.4.2. Southern hybridisation analysis

The second possible explanation may be that the plasmid is being maintained in the host cell with or without transposition of the transposon. To assess this hypothesis transconjugant and control DNA was digested with *Hin*dIII or left undigested and separated on agarose gel. The DNA was

transferred to nitrocellulose and probed with Pb-cat. Figure 4.18 shows the hybridisation pattern giving the same results as the previous blots for the digested DNA. The results are inconclusive for undigested DNA as it is impossible to distinguish the plasmid DNA from the genomic DNA in the hybridising band. In addition to the hybridisation, plasmid rescue was attempted to assess the fate of the plasmid. The standard Mini-prep procedure failed to generate visible plasmid from the transconjugants. Therefore, to amplify the plasmid the primary elute from a standard Mini-prep of a 630Δerm::pKM1049 transconjugant was transformed into competent *E. coli*. Plasmids isolated from transformants were then digested with *Eco*RV (Figure 4.19). The plasmid obtained from four transformant colonies produced a digest pattern identical to pKM1049.

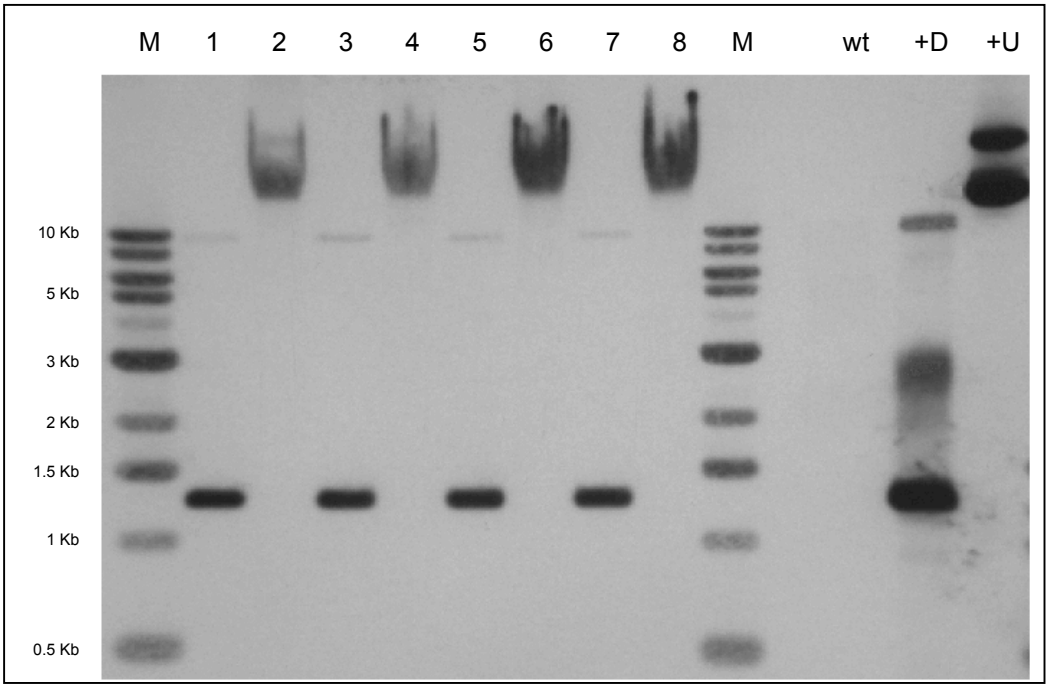


Figure 4.18 Southern hybridisation analysis of R20291 transconjugant DNA (pKM1048) undigested or digested with *Hin*dIII.

Transconjugant and plasmid DNA undigested or digested with *Hin*dIII was separated by gel electrophoresis and transferred to a nitrocellulose membrane. The membrane was then probed with Pb-cat. Dark bands represent hybridisation of the probe to DNA. The hybridisation of the digested DNA corresponds to previous blots, one band at approximately 1.2 Kb is seen. Hybridisation of the undigested DNA appears to be localised to a high molecular weight smear of a similar size to the undigested plasmid control (U) but show little discrimination between the predicted plasmid forms. Lanes: M = 1 Kb ladder; 1 = digested transconjugant DNA a.; 2 = undigested transconjugant DNA a.; 3 = digested transconjugant DNA b.; 4 = undigested transconjugant DNA b.; 5 = digested transconjugant DNA c.; 6 = undigested transconjugant DNA c.; 7 = digested transconjugant DNA d.; 8 = undigested transconjugant DNA d; wt = wild type *C. difficile* R20291, +D = digested pKM1048; +U = undigested pKM1048.

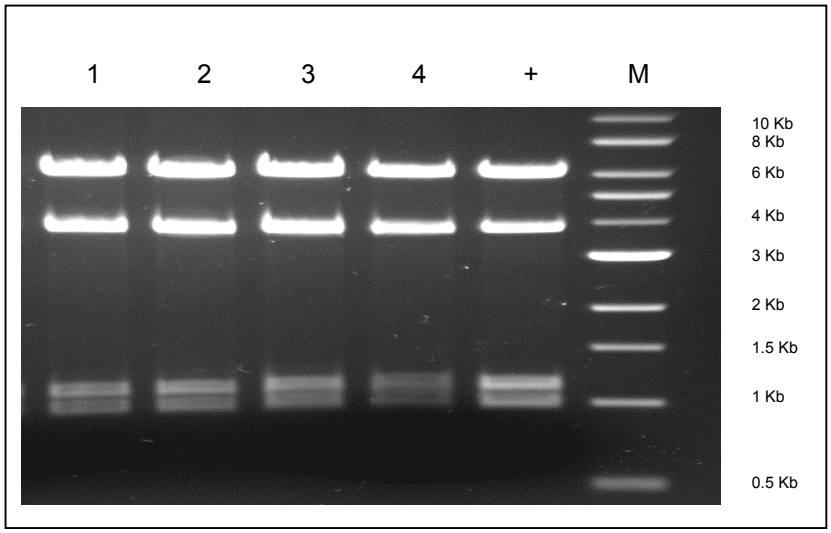


Figure 4.19 Digest of plasmid rescued from *C. difficile* 630Δerm transconjugants.

Plasmid DNA from four transconjugants was transformed into competent *E. coli* before extraction. Plasmids were digested with *Eco*RV and fragments separated on 1% agarose gel. Plasmid isolated from each transformant gave 4 fragments that are identical to the fragments generated from digestion of pKM1049. Lane1-4 = transformants; + = pKM1049; M = 1 Kb ladder.

4.5.4.3. Plasmid stability assay

pMTL5401F is described as an unstable plasmid within *C. difficile* due to the presence the repH replicon from pCB102, highly unstable in both strain 630 and in 027/NAP1 strains [177]. The plasmid may be maintained by the presence of antibiotics in the medium. To assess the stability of pMTL5401F and pKM1048 within R20291 a stability assay was conducted. Three independent transconjugants were cultured without antibiotic selection, and passaged four times. After each passage cultures were plated onto non-selective plates and selective plates. The number of colony forming units (cfu) was established for non-selective and selective plates and compared. After three passages in non-selective medium 99.9% of the transconjugant colonies did not contain the plasmid or mini-Tn*10* showing sensitivity to Tm (Table 4.7). An attempt was made to introduce pMTL5401F into R20291 by filter mating however, this failed to produce transconjugants and consequently no comparison can be made between the stability of pMTL5401F and pKM1048.

	passage 1	passage 2	passage 3
Transconjugant 1	77.9%	98.2%	99.9%
Transconjugant 2	74.8%	98.0%	99.9%
Transconjugant 3	77.9%	99.3%	99.9%

Table 4.7 Plasmid loss after serial passage in the absence of selection.

Three R20291 transconjugants were serially passaged in the absence of selective antibiotic and plated onto Tm selective agar. The number of colonies that were able to grow on selective media were enumerated and compared to the total number of cfu ml^{-1} in the inoculum to generate a percentage of cfu/ml which contained the plasmid. After three passages in non-selective medium 99.9% of the cfu had lost plasmid pKM1048 and associated Tm resistance.

4.5.5. **Promotion of transposition**

To assess whether transposition would occur after several generations, three R20291 transconjugants were inoculated into broth with and without thiamphenicol (a total of 30 cultures). Each broth was passaged three times in fresh media of the same constitution. Cultures grown without (1) or with (2) thiamphenicol were plated onto thiamphenicol selective agar plates, after growth of colonies each plate was replicated onto Lm selective agar. One set of cultures grown in Tm selection (3) were plated onto non-selective agar before replica plating onto Lm and Tm selective agar. The total number of colonies that grew on the Tm selective plates was enumerated and compared to the Lm replica plates (Table 4.8). All of the colonies were found to be Tm and Lm resistant.

The transposase and mini-Tn*10* fragment cloned from plasmids pHV1248 and pHV1249 included the promoter region of the transposase (P_{tat}), active in *B. subtilis*. It was not known whether this would be functional in *C. difficile* and thus the mini-Tn*10* delivery plasmid was designed so that P_{fac} could also be used to drive transcription of the *transposase* gene. Four R20291::pKM1048 transconjugants were cultured in broth with addition of IPTG and Tm. The cultures were passaged nine times in the same conditions. After ten passages, samples were plated and three random colonies chosen from each sample for further analysis. Digested DNA was probed with Pb-cat (Figure 4.20). The blot revealed that after ten passages in IPTG, transposition of mini-Tn*10* did not occur.

A

Transconjugant culture	Broth	Total Number of colonies		Difference
		Tm100 µg ml^{-1}	Lm 100 µg ml^{-1}	
1	1	47	47	0
	2	61	61	0
	3	64	64	0
	4	56	46	0
	5	46	46	0
2	1	32	32	0
	2	36	36	0
	3	37	37	0
	4	38	35	0
	5	35	35	0
3	1	45	45	0
	2	42	42	0
	3	48	48	0
	4	51	51	0
	5	47	47	0

B

Transconjugant culture	Broth	Total Number of colonies		Difference
		Tm100 µg ml^{-1}	Lm 100 µg ml^{-1}	
1	1	187	187	0
	2	173	173	0
	3	158	158	0
	4	177	177	0
	5	144	144	0
2	1	79	79	0
	2	87	87	0
	3	79	79	0
	4	86	86	0
	5	82	82	0
3	1	185	185	0
	2	188	188	0
	3	208	208	0
	4	176	176	0
	5	198	198	0

C

Transconjugant culture	Broth	Total Number of colonies		Difference
		Tm100 µg ml^{-1}	Lm 100 µg ml^{-1}	
1	1	60	60	0
	2	53	53	0
	3	42	42	0
	4	45	45	0
	5	55	55	0
2	1	40	40	0
	2	35	35	0
	3	29	29	0
	4	42	42	0
	5	33	33	0
3	1	43	43	0
	2	54	54	0
	3	65	65	0
	4	57	57	0
	5	51	51	0

Table 4.8 Results of transconjugant replica plating assay.

Three R20291 pKM1048 transconjugant cultures (A, B, & C) were grown in non-selective broth (1); Tm selective broth (2 & 3) and plated onto either Tm selective plates (1 & 2) or non-selective plates (3). Plates were incubated for 36 hours before plates were replica plated onto selective agar containing wither 100 µg ml^{-1} Tm or 100 µg ml^{-1} Lm and incubated for 48 hours. The total number of colonies on both plates was enumerated and plates were assessed for any difference in colony growth. No difference in the number of colonies that grew on the Tm plates compared to the Lm plates was observed for any of the experimental conditions.

To confirm that the intact plasmid had been maintained in the transconjugants after serial passage with IPTG, the membrane was re-probed with Pb-catH (Figure 4.21). Pb-catH covers a section of the plasmid backbone and part of mini-Tn*10* including one of the *Hin*dIII sites present in pKM1048. The blot shows hybridisation of the probe for all 12 of the transconjugants at two positions, approximately 10 Kb and 1.2 Kb. The hybridisation pattern for the transconjugants is identical to that for the positive (plasmid) control. The gel is slightly skewed however the lower bands in each lane are at the same position as the control band.

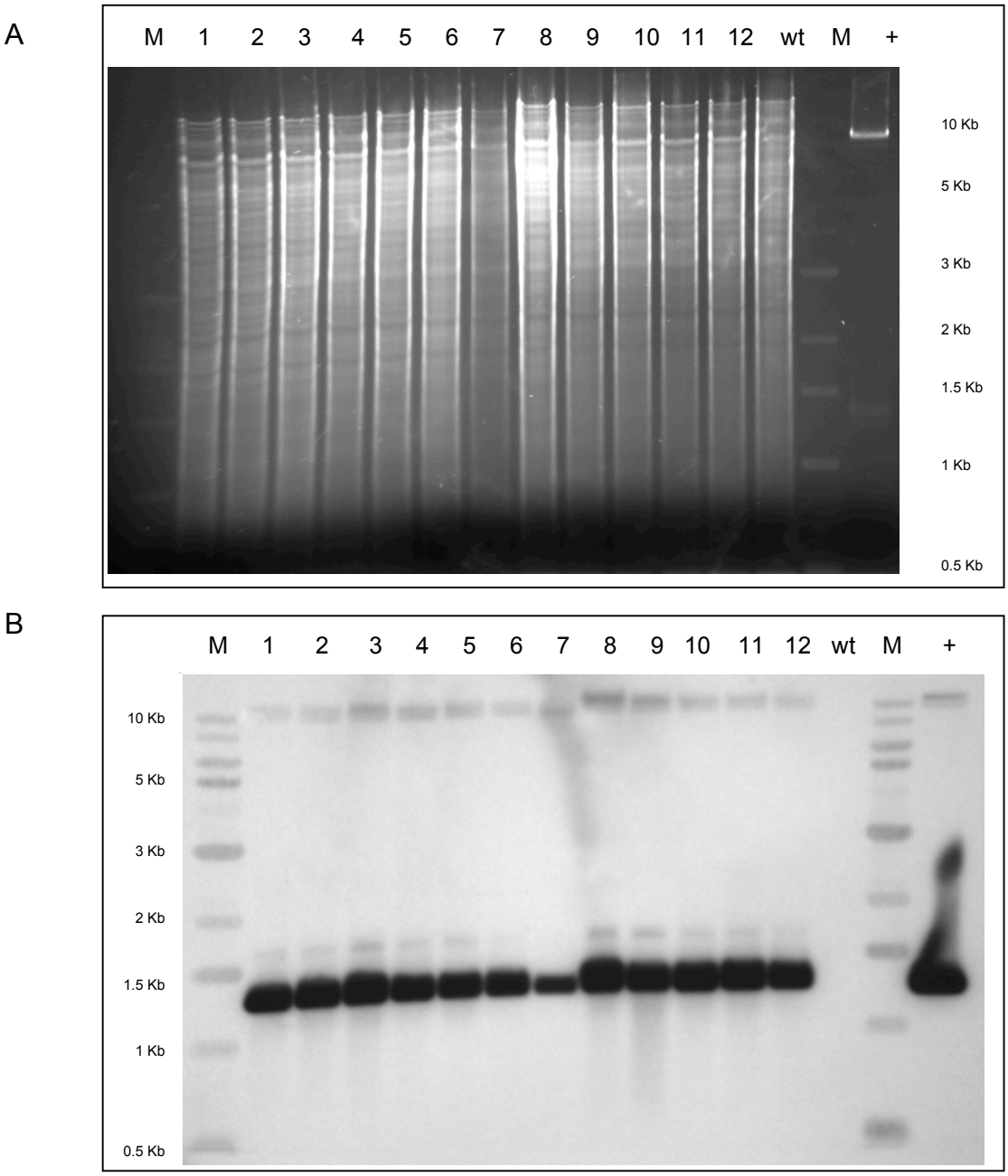


Figure 4.20 Agarose gel of IPTG passaged transconjugants digested with *Hin*dIII (A) and Southern hybridisation of digested DNA probed with Pb-cat (B).

Four transconjugants were grown in selective medium containing IPTG, passage was repeated a further 9 times. Cultures were plated on selective medium and 3 random colonies from each culture were selected and DNA extracted. The DNA was digested with *Hind*III and separated on a 0.8% agarose gel. DNA was transferred to nitrocellulose membrane and fixed by UV cross linking. The membrane was probed with labelled Pb-cat and appropriate marker. M = 1 Kb ladder; 1-3 = transconjugant 1; 4-6 = transconjugant 2; 7-9 = transconjugant 3.; 10-12 = transconjugant 4; wt = wild type R20291, + = pKM1048.

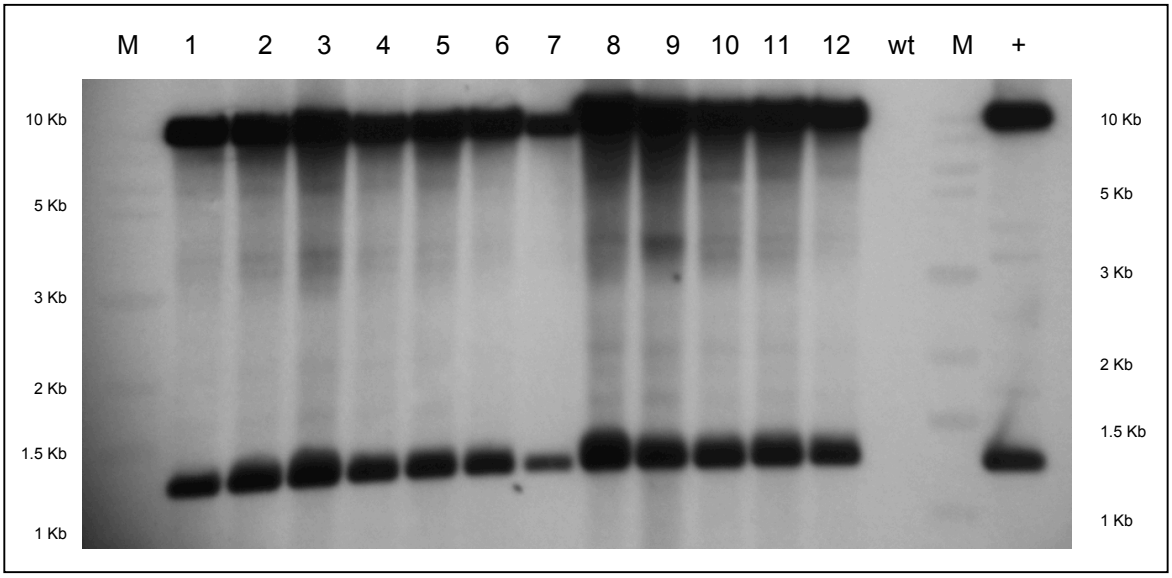


Figure 4.21 Southern hybridisation probed with Pb-catH.

Probe Pb-catH covers the *Hin*dIII restriction site downstream of the mini-Tn*10* element (see **Figure 4.7**). There is positive hybridisation of the probe for all 12 of the IPTG passaged R20291::pKM1048 transconjugants at two positions approximately 10 Kb and 1.2 Kb. Although the gel is slightly skewed and the first marker is faint the lower bands in each lane are at the same position. The hybridisation pattern for the transconjugants matches that seen for the positive control. As predicted two areas of hybridisation are observed as the probe binds to both the mini-Tn*10* and plasmid backbone on either side of a *Hin*dIII site meaning half of the probe will bind to one of the two main plasmid fragments and the other half will bind to the remaining fragment. The nonspecific ghost banding is likely to be due to non-specific hybridisation and/or insufficient washing before detection.

4.5.6. **RT-PCR**

P_{fac} is reported to be induced by IPTG, however inclusion of IPTG in the growth medium failed to promote the transposition of mini-Tn*10* in *C. difficile* strain R20291 to a detectable level. The expression of the transposase gene in the presence of IPTG was therefore investigated by RT-PCR. RNA was extracted from one R20291::pKM1048 transconjugant grown with and without IPTG. Two control primers pairs were designed to amplify the *16S* gene and *GyrA* gene as constitutively expressed standards. Two further primer pairs, qCat and qTpase, were designed to amplify the *cat* and *transposase* gene respectively. All of the primers were tested in control experiments to allow optimum design of the RT-PCR parameters, including establishing primer and RNA concentrations.

Positive amplification of each gene was observed in each RNA concentration tested, 100, 10 and 1 ng. A calibration curve was established for the *16S*, *GyrA* and transposase gene; the reaction efficiencies are within the acceptable range (from 86-100%) allowing real conclusions to be drawn form the test data (Table 4.9). The threshold values obtained from the calibration curve were used to establish the average C_T value (rep C_T) for expression of each gene in each experiment (Table 4.10). Three independent experiments were performed to assess the expression of the transposase gene with and without addition of IPTG. There was no significant difference in the expression. Assessment of the two constitutively expressed housekeeping genes showed almost identical expression in the two samples, supporting the results observed for the *transposase* gene.

Calibration curve values		
16S	E	0.857
	R^2	0.996
Tpase	E	0.913
	R^2	0.994
GyrA	E	1.020
	R^2	0.996

Table 4.9 Quantitation information for calibration curve.

The values for the efficiency of the reaction (E) and R^2 (coefficient of determination) are listed for each primer. Both measures are critical to evaluate the PCR efficiency.

	Gene		
Experiment 1	**Tpase**	**16S**	**GyrA**
Threshold Value	0.334	0.128	0.313
$repC_T$ value for samples with IPTG	12.52	5.56	20.77
$repC_T$ value for samples without IPTG	12.33	5.75	20.45
Experiment 2			
Threshold Value	0.334	0.128	0.313
$repC_T$ value for samples with IPTG	13.12	6.55	21.61
$repC_T$ value for samples without IPTG	13.39	6.70	21.45
Experiment 3			
Threshold Value	0.334	0.128	0.313
$repC_T$ value for samples with IPTG	12.80	6.45	21.90
$repC_T$ value for samples without IPTG	12.89	6.42	21.19
Average $repC_T$ for samples with IPTG	**12.81**	**6.19**	**21.43**
Average $repC_T$ for samples without IPTG	**12.87**	**6.29**	**21.03**

Table 4.10 Expression of *transposase* and with and without IPTG induction.

The $repC_T$ values (from triplicate samples) for the transposase, 16S and GyrA gene were calculated using the threshold values obtained from the calibration curve for each gene. The average $repC_T$ of three independent experiments is also shown. There is no significant difference between the expression in induced and non-induced samples. All threshold values were taken from the same RNA dilution with approximately 7 ng of template RNA.

4.6. Discussion

Reliable tools to generate random insertions in the genome of *C. difficile* were not available when this study began, however there was a clear need for their development to investigate the molecular pathogenicity of this important nosocomial pathogen. Transposon mutagenesis is a widely used technique allowing the generation of large libraries of insertion mutants that can be screened to identify genes involved in the specific functions. Petit *et al*. (1990) describe the adaptation and use of a Tn*10* derivative, mini-*Tn10*, for use in the Gram-positive bacterium *B. subtilis* [197]. Mini-Tn*10* was shown to generate random and stable insertions throughout the genome with one to two copies per clone. These experiments highlighted mini-Tn*10* as an attractive potential random mutagen of *C. difficile*.

Introduction of DNA into *C. difficile* is only possible through conjugative transfer from a suitable donor bacterium. Previous studies have shown that an *E. coli* strain CA434 could be used to deliver plasmids to *C. difficile* by direct filter mating [347]. The investigation of mini-Tn*10* insertion in *C. difficile* required the development of a suitable delivery plasmid. Several key attributes were required including a replicon for replication in *E. coli*, a replicon for replication in *C. difficile*, a selectable marker(s), transfer genes to work with the R702 plasmid, and a promoter active in *C. difficile* to drive expression of the Tn*10* transposase. Plasmid pMTL5401F is an *E. coil-Clostridium* shuttle plasmid and has been shown to be unstable in some

clostridia, there is also the potential to allow inducible expression of genes cloned downstream of P_{fac} [130].

Initially two methods, PCR and restriction digest were performed concurrently to create the construct pKM1048. PCR amplification of mini-*Tn10* failed to generate the desired fragment, it is possible that this was due to sub-optimal reaction conditions or primers. These possibilities were not investigated further to allow progression of the project. The restriction digest approach appeared to be relatively straightforward allowing direct insertion of a fragment containing the *transposase* gene and mini-Tn*10* into pMTL5401F. This method was thus repeated for the second mini-Tn*10* derivative (pKM1049).

Confirmation of the construct was achieved through PCR and restriction digest analysis. The plasmid construct was then transformed to the donor *E. coli* strain CA434. Problems were encountered when using the *cat* selection marker in *E. coli*, which significantly delayed the isolation of CA434-48. Selection with Cm at various concentrations was seen to be unreliable and failed to allow consistent growth of colonies thought to contain mini-Tn*10*. Further investigation of the use of *cat* in pervious studies revealed that weak expression of this gene in *E. coli* has been documented [390].

Filter mating experiments were performed with CA434-48/49 donor and *C. difficile* strains CD37, 630Δerm, 630 and R20291. Spot mating is usually performed for *E. coli*-*Clostridium* mating, whereas filter mating is used for

Bacillus-Clostridium mating as the filter aids the conjugation process [337]. Spot mating experiments between CA434-48 and strains R20291 and 630 failed to generate transconjugants, whereas filter mating did. The appropriate level of Tm selection for the transconjugants was initially unknown; therefore there was a degree of trial and error in determining the most appropriate parameters. Once transconjugants had been isolated an MIC was conducted to determine the level of resistance shown by *C. difficile* transconjugants containing the *cat* gene. High levels of resistance were observed for the transconjugants compared to the wild type strains allowing a standard concentration to be determined. Transconjugants were selected at various concentrations of Tm, although 100 μg ml^{-1} was used for the majority of experiments. This is significantly higher than the concentration recommended for selection for *C. difficile* containing *catP* (15 μg ml^{-1}) [130]. However, the MIC and selection level used in this study are based on the resistance conveyed by expression of the gene from the plasmid, it was not possible to determine the level of resistance shown by a chromosomally expressed *cat* gene. The resistance profiles of the transconjugants containing *cat* suggest that this is a useful marker in *C. difficile* however, the inconsistent results seen here and the previous demonstration of the low level of resistance conferred by this gene in *E. coil*, reveal that this is not a suitable maker for this study. Ideally other markers, including *catP* (derived from *C. perfringens* plasmids pIP401 & pJIR27) should be examined for their usefulness in the mini-Tn*10* element within both organisms.

The generation of transconjugants was extremely inefficient as no transconjugants were generated in the majority of mating experiments. Alterations to the protocol were investigated including initially selecting for the plasmid backbone through Lm selection before a secondary selection for mini-Tn*10* using Tm. No transconjugants were observed in this experiment. It is possible that the standard level of Tm used to select transconjugants was too high to identify those with the transposon on the chromosome, thus primary selection for the plasmid with Lm would evade this issue. A subsequent selection for Tm resistance at a lower level may allow identification of transconjugants with insertions of mini-Tn*10* on the chromosome.

Putative transconjugants were all assessed by PCR for the presence of the *cat* and *bla* gene, and all were initially positive. A number of transconjugants were not maintained in selective broth and were subsequently negative for *cat* and *bla*, confirmed by Southern hybridisation. This indicated that the plasmid and the element had been lost after removal of Tm selection. This result is consistent with the conclusion that the transposon had either not transposed or was not stably maintained in the chromosome of the recipient bacterium. If mini-Tn*10* had remained on the plasmid the reported instability of the repH replicon and the lack of selective pressure is likely to have led to the loss of the plasmid [130, 177]. Instability of the plasmid was confirmed after serial passage in non-selective media, proving pCB102 to be a suitable replicon for this system.

Hybridisation results for R20921::pKM1048 revealed that both mini-Tn*10* and the plasmid were present. The hybridisation profile was identical to the plasmid control suggesting that the complete plasmid had been maintained. It was also possible, but unlikely, that the plasmid had inserted into the genome of R20291 through homologous recombination. Investigation of the plasmid sequence revealed no significant regions of homology. Several PCRs were conducted to check for the presence of intact features of the plasmid including *ermB* and P_{fac}. A Southern blot of digested and undigested R20291::pKM1048 transconjugant DNA was probed with Pb-cat to assess whether the intact plasmid could be observed, unfortunately the result do not confirm the presence of intact plasmid. However, it was possible to obtain the original plasmid from the R20291::pKM1048 transconjugants confirming the presence of intact plasmid.

Plasmids containing two versions of mini-Tn*10* were generated by Petit *et al*. (1990), the element present on pHV1248 contained the *cat* gene flanked by two outside ends of IS*10* and the element from pHV1249 contained the *cat* gene flanked by the inside ends of IS*10* [197]. Host factors IHF and HU are known to be required for transposition of the element containing OEs but not for IEs in *E. coli*. Petit *et al*., showed that both elements were able to transpose in *B. subtilis*, therefore both elements were examined in *C. difficile*. From PCR and hybridisation analysis it appeared that there had been no transposition events, however it is possible that these events did occur but were not detected. Therefore, experiments aiming to promote transposition were performed. A replica platting assay was conducted with the aim of

identifying cells after serial passage that had retained Tm resistance yet lost Lm resistance, thus indicating the presence of mini-Tn*10* and loss of the plasmid. Unfortunately, no colonies were identified with this phenotype. Several hundred colonies were examined as transposition events might occur at very low frequencies. It is possible that an insufficient number of colonies were screened.

The promoter P_{fac}, a composite of the ferredoxin gene *fd* promoter fused to the operator of the lacZ operon, is present in pMTL5401F and was expected to allow IPTG inducible expression of downstream cloned genes [391]. This promoter was shown to be inducible in *C. sporogenes*; however it has not previously been investigated in *C. difficile* [130]. The cloned transposase and mini-Tn*10* element were placed downstream of this promoter potentially to allow inducible expression. The cloned fragments from pHV1248 and pHV1249 both contained the promoter P_{tac} and RBS region upstream of the transposase, which are active in *B. subtilis*. It was unknown whether these would be active in *C. difficile* and therefore the existence of the P_{fac} promoter would act as a backup to this to ensure the *transposase* was expressed allowing transposition of mini-Tn*10*. Transconjugants were therefore grown in the presence of IPTG to assess whether this would lead to transposition of the element. There was no evidence of transposition to the chromosome from the hybridisation analysis, blots were identical to those performed on transconjugants grown without IPTG. RT-PCR was then conducted to see whether the transposase was being expressed and whether the addition of IPTG had affected of the level of expression. The results revealed that the

transposase was being expressed and that addition of IPTG had no effect on the level of expression. As there was clearly expression of the transposase it is not clear why transposition of mini-Tn*10* did not occur. The element and the *transposase* gene are identical to those used in *B. subtilis* as the fragment was directly cloned into pMTL5401F, it is possible that it is not functional in *C. difficile*. Investigation of the codon usage in *C. difficile* revealed that there might be a shortage of codons for the cognate tRNA in *C. difficile* (Appendix) (personal communication Professor Mike Young, 2011). To overcome this issue the transposase gene could be re-synthesised for use specifically in *C. difficile*; unfortunately, this was not fully investigated due to time constraints.

During the design of this project both mini-Tn*10* and *mariner* were considered as potential elements for the random mutagenesis of *C. difficile*. Mini-Tn*10* was chosen as the preferred element to investigate due to the success seen in *B. subtilis*, with a view to investigate *mariner*-based transposon systems at a later date. Shortly after the end of this project Cartman and Minton (2010) published a paper describing the generation of random insertions of the *mariner*-transposable element *Himar1,* in the genome of *C. difficile* strain R20291 [388]. Cartman and Minton were able to show stable insertions of the element that were, in the majority, single insertions which occurred at a relatively high frequency (~4.5 x 10^{-4}). After generation of a library of undefined size they isolated two mutants, a sporulation/germination-defective clone with an insertion in *cspBA* and an auxotroph with an insertion in the

pyrimidine biosynthesis gene *pryB*. However, neither of these mutations has been complemented.

The pMTL-SC1 delivery vector for the *Himar1* element appears to be simpler and is significantly smaller (7.4 Kb) than the construct used in the present study. It contains *catP* as the selectable marker, which may be a more useful marker in *E. coli* and *C. difficile*. One of the potentially significant problems with the design of pKM1048/49 was that the plasmid vector was unmodified except for the addition of the *transposase* and mini-Tn*10*, largely due to unfeasible cloning options, thus P_{fac}, an unverified promoter in *C. difficile* was relied upon for expression of the transposase should its native promoter be non-functional.

By using the *C. difficile* promoter from TcdB, Cartman and Minton are able to express genes under its control in *C. difficile*. pMTL-SC1 contains pBP1 derived form *C. botulinum*, the authors assessed the efficiency of conjugation of the plasmid containing this replicon, and the segregational stability of the plasmid alongside the clostridial replicons pCB102 (*C. butyricum*) and pCD6 (*C. difficile*) in R20291. pBP1 was shown to have approximately the same segregational stability as pCB102, however, the element showed approximately 10-fold higher conjugation frequencies. These data suggest that the stability of the replicon used in the mini-Tn*10* delivery vector is comparable to the one used in this study although this is expressed per generation. The conjugation frequencies observed with the plasmids described are also significantly higher than the frequencies observed for

pKM1048/49, this may be due in part to the large size of these plasmids (approximately 11 Kb) compared to the plasmids assessed by Cartman and Minton which range from 3 Kb to 5.3 Kb and 7.4 Kb [392]

The development of a *mariner*-based random mutagenesis tool for *C. difficile* will hopefully enable a further understanding of the molecular pathogenesis of this organism. *Mariner* has been shown to enter the genome at AT sites, thus as *C. difficile* is AT rich this preference for AT insertion sites may not be a significant issue. It would also be useful to investigate the insertion sites in a larger sample to ensure the randomness of the system and confirm that predominantly single insertions occur. This initial study certainly suggested that this is an effective tool, however, further investigation of the use and efficiency of this tool will confirm whether it will be suitable for the generation of large stable insertion libraries, and whether it may be used in other *C. difficile* strains.

4.7. Conclusions and future perspective

A novel tool for the potential random mutagenesis of *C. difficile* has been developed based on an *E. coli-C. difficile* shuttle plasmid and mini-Tn*10*. The plasmid has been shown to transfer from the donor cell via conjugation to four *C. difficile* strains including the clinically relevant 027/NAP1 strain R20291. Low conjugation frequencies were observed however, once inside the recipient cell the plasmid conferred high levels of resistance to thiamphenicol. The plasmid was shown to be unstable in the absence of selective pressure; this facilitates stable insertions through loss of the transposase. It was not possible to demonstrate transposition of the element in the genome of *C. difficile.* The expression of the *transposase* gene was confirmed by RT-PCR with results indicating that there was no affect of IPTG on expression. It is not clear why the transposon failed to transpose in these bacteria, however it is possible that this is due to the very low limits of detection.

Although insertion mutants were not generated this work shows that it is possible to transfer relatively large plasmids into *C. difficile* strains. Clearly, further investigation is required to develop this potential tool and to understand why transposition failed to occur in these transconjugants. The development of a series of modular plasmids, and further characterisation of promoters and replicons suitable for use in *C. difficile* genetic systems, will surely aid the development and refinement of shuttle plasmids to allow efficient delivery of DNA.

The recent description of a *mariner*-based random mutagenesis system may diminish the importance of further developing the mini-Tn*10* tool. However, this initial report does not fully describe the usefulness of this tool and it is possible that other complementary tools, such as mini-transposon-based tools, will be equally important in the investigation of gene function and regulation in *C. difficile*.

Chapter 5

5. Future work - Development of a mini-Tn*916* transposon and delivery vector for *C. difficile* mutagenesis

5.1. Abstract

Transposon mutagenesis is a standard technique for the generation of random insertion libraries. The conjugative transposon Tn*916* has been used successfully to generate a library of transconjugants in the toxin producing strains 630Δerm and R20291. Development and testing of a mini-Tn*10* shuttle vector failed to generate transconjugants with stable genomic insertion of the transposon in R20291 or 630Δerm. As Tn*916* has been shown to enter the genome of CD37, 630Δerm and R20291, a 'mini' version of this transposon, containing the left and right ends of Tn*916* and a *cat* antibiotic resistance cassette (encoding resistance to chloramphenicol) was generated by SOEing PCR. A shuttle delivery plasmid was also designed and partially constructed to deliver this mini-transposon into *C. difficile* by filter mating with *E. coli* strain CA434.

5.2. Introduction

5.2.1. **Tn*916* and mini-transposons**

Mini-transposons represent a useful alternative to their bulky full-length ancestors. These compact elements are generally composed of functional ends of their parent required for transposition and a suitable marker to allow selection [375]. To ensure stability, the transposase or genes required for excision and insertion, are placed outside the element, finally minor modification to the transposase gene sequence can allow reduced target site specificity [375, 393]. Mini-Tn*10* has been used successfully in several bacteria including *E. coli* and *B. subtilis,* generating stable random insertions [197, 394].

The conjugative transposon Tn*916* has been used as an insertional mutagen in *C. difficile* strain 630Δerm and R20291 (see Chapter 2). Unfortunately, this is not an ideal transposon for random mutagenesis due to preferential insertion into AT rich sites and insertion into hot spots, for example in CD37; furthermore, due to it large size (18 Kb) there is a greater chance of causing downstream effects after insertion [193, 210].

5.2.2. **Shuttle vectors and filter mating**

As described earlier, one of the major issues that have limited the investigation of *C. difficile* has been the lack of efficient and reliable tools for genetic manipulation. Recently, two methods have been described for the introduction of DNA into *C. difficile* and the targeted disruption of genes,

namely the ClosTron and the shuttle plasmid/recombination vectors [130, 132]. This has revolutionised the study of *C. difficile* allowing targeted inactivation of specific genes to investigate their function, particularly those associated with virulence [117, 126, 129]. Both tools rely on conjugation of plasmids to introduce DNA using a filter mating method akin to that used for the transfer of Tn*916* from a *B. subtilis* donor [205].

Tn*916,* although shown to lead to insertions in the genome of *C. difficile* is by no means an optimal tool for random mutagenesis of this important pathogen. Preferential insertion sites, multiple insertion and low transfer frequencies of this large element are major drawbacks for its use in generation of large insertion libraries. The investigation of this element, also documented in this thesis, has further stressed these drawbacks; frequency of transfer to the hyper-epidemic strain R20291 is extremely low, making generation of large insertion library in this strain infeasible. Investigation of the use of mini-Tn*10* as a random mutagen in *C. difficile* (Chapter 4) showed that this element does not lead to random insertions in *C. difficile* strains 630Δerm and R20291 in its present form. Therefore, refinement of Tn*916* by creation of a mini derivative may represent an alternative tool for mutagenesis, which could significantly increase the utility of this transposon and further our understanding of the molecular pathogenicity of *C. difficile*. As with other mini-transposons, the derivative will be a fraction of the size of the parent and can be engineered with a more flexible resistance marker to facilitate its use in a variety of strains. The *xis* and *int* genes will be placed outside of the transposon ends to allow stable insertion and repeated rounds

of mutagenesis and to allow generation of a saturating insertion library if desired. Only the essential sequences required for transposition will be maintained, specifically the left and right end of Tn*916* with a length of flanking sequence. It is not clear how much sequence will be required for transposition. However, this can be adjusted accordingly after assessment of the construct in *C. difficile*. Likewise, a suitable resistance marker will be included in the construct to allow selection of transposon insertions. This may also need to be exchanged, so in this work, consideration will be given to the ease with which this can be achieved.

To ensure that minimal redundant sequence is included in the construct, the derivative will be made using splicing overlap extension PCR (SOEing PCR), a relatively simple and efficient method for generating gene constructs [395, 396]. The mini-transposon will be placed downstream of the *xis* and *int* genes essential for excision, mobilisation and insertion. The *tet*(M) gene with the *tet*(M) promoter (P_{tetM}) will be placed upstream of the *xis* and *int* genes. Plasmid pMTL5401F used for delivery of mini-Tn*10* has been shown to successfully conjugate with strains 630Δerm and R20291; this construct was also unstable upon removal of antibiotic selection (see Chapter 4). The plasmid carries appropriate genes for replication in the Gram-negative and Gram-positive host, two selectable markers and the *lacZ* reporter gene to allow blue/white selection during cloning. This chapter describes the steps conducted to generate a mini-Tn*916* derivative and a suitable delivery vector. Although this work is not complete it sets out a rationale for the generation of a potentially novel tool for random mutagenesis of *C. difficile*.

5.3. Aims and objectives

The aim of this study is to generate a mini version of Tn*916* and a suitable delivery vector for use as a potential mutagenesis tool in *C. difficile*.

Objectives:

1. Generate a mini-Tn*916* using splicing by overlap extension (SOEing) PCR that includes a suitable resistance marker for strains 630Δerm and R20291.
2. Use stepwise cloning to create a plasmid to deliver the mini-Tn*916* to *C. difficile.*
3. Transfer the plasmid to CA434 to facilitate transfer to *C. difficile* strains 630Δerm and R20291 using filter mating.
4. Test the construct in *C. difficile* and refine if necessary.
5. Generate a library of transconjugants and screen them for mutations affecting cytotoxicity, sporulation/germination, bile and *p*-cresol tolerance.

5.4. Materials and methods

5.4.1. **Generation of mini-Tn*916***

Three 1 Kb fragments of the left and right end of Tn*916* from pAM120 (Figure 5.1, Table 5.1) and the *cat* gene from mini-Tn*10* in pHV1248 (Figure 4.4) were amplified by PCR independently using chimeric primers (Table 5.2). Each fragment was confirmed by gel electrophoresis (0.8% agarose) and extracted and purified as using a gel extraction kit (Qiagen, UK). A two step SOEing PCR approach was used to generate the desired 3 Kb fragment – Initially fragment 2 (*cat*) and fragment 3 (Tn*916* left end) were fused using PCR with primers CAT-1 and 916L-2 using a standard PCR reaction with addition of 17.5 mM $MgCl_2$. The resulting 2 Kb fragment was visualised on agarose gel before extraction and purification. This 2 Kb fragment was joined by PCR to fragment 3 (Tn*916* right end) using primer 916R-1 and 916L-2 using standard PCR with addition of 17.5 mM $MgCl_2$. The resulting 3 Kb fragment contained the *cat* gene flaked by the right and left end of Tn*916* (Figure 5.2). The 3 Kb PCR fragment was visualised on agarose gel before extraction and purification. The fragment was cloned into pCR2.1-TOPO to generate pCR2.1-TOPO::mini-Tn*916*. Each 1 Kb fragment and the complete mini-Tn*916* were sequenced by the department of Biochemistry, University of Cambridge, UK using an Applied Biosystems 3730xl DNA Analyzer, (96 capillary array sequencer).

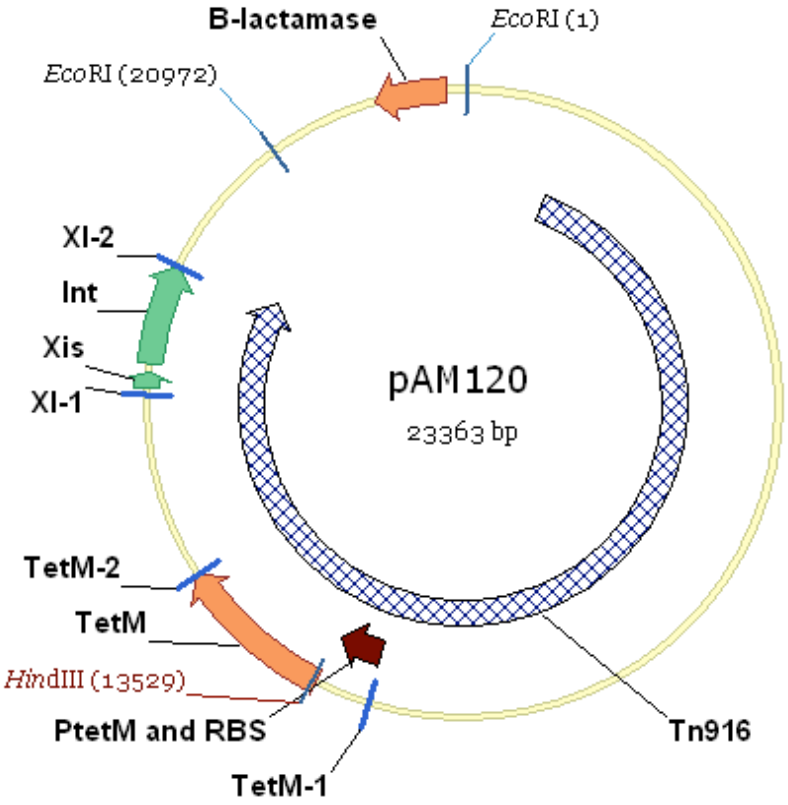


Figure 5.1 Schematic diagram of pAM120

pAM120 was used as template for the amplification of P_{tetM}/*tet*(M) and the left and right end of Tn*916*. The primer binding sites are indicated on the map.

Plasmid	Marker	Features	Reference
pAM120	*tet*(M)	Tn*916*	Gawron-Burke & Clewell; 1984 [232]
pHV1248	*bla, cat, ermB*	mini-Tn*10*	Petit *et al*., 1990 [197]
pMTL5401F	*bla, ermB*		Heap *et al*., 2007 [130]
pMTL85141	*catP*		Heap *et al*., 2009 [392]
pCR2.1-TOPO		Ap^R & Kan^R	Invitrogen
pCR2.1-TOPO::*tet(M)*	*bla, tet*(M)		This work
pCR2.1-TOPO::*xis/int*	*bla*		This work
pCR2.1-TOPO::mini-Tn*916*	*bla, cat*	mini-Tn*916*	This work
pKLM1.1	*bla, ermB, tet*(M)		This work
pKLM1.2	*bla, ermB, tet*(M)		This work
pKLM916	*bla, cat, ermB, tet*(M)	mini-Tn*916*	This work

Marker gene	Resistance	Abbreviation
bla	Ampicillin	Ap
cat	Chloramphenicol	Cm
catP	Chloramphenicol	Cm
ermB	Erythromycin/ Lincomycin	Em/Lm
tet(M)	Tetracycline	Tc

Table 5.1 Plasmids used in the study

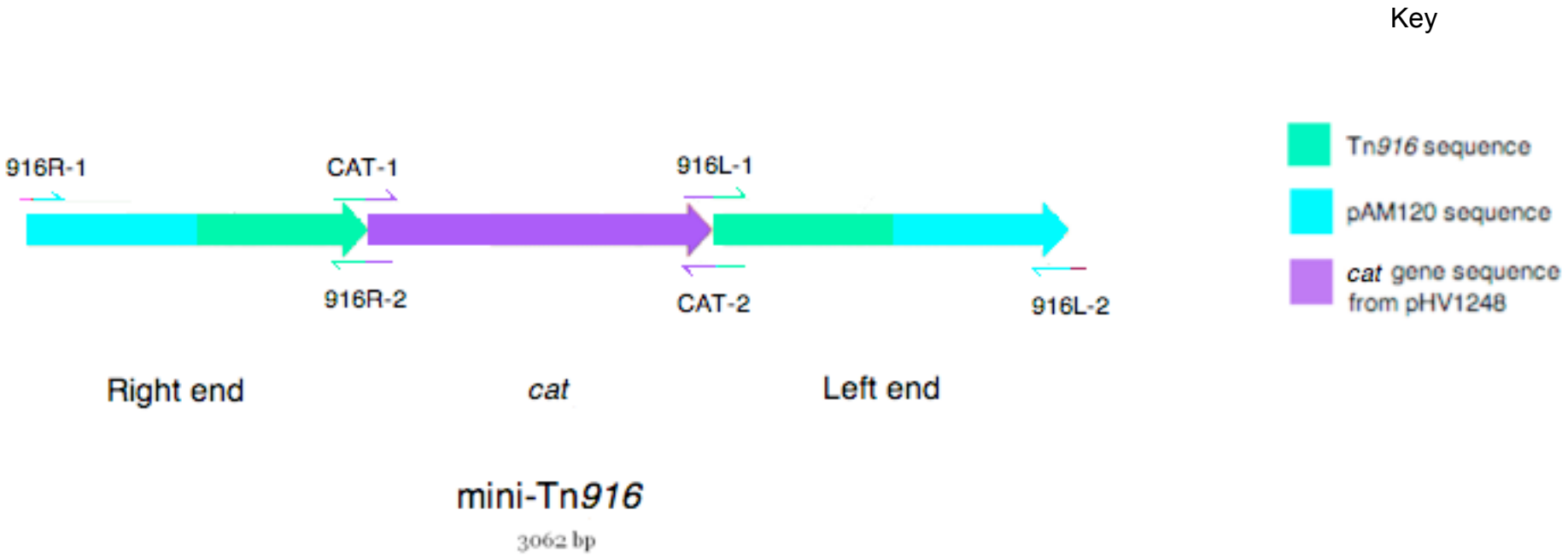


Figure 5.2 Schematic representation of mini-Tn*916*.

Mini-Tn*916* is composed of three sections: 500 bp of the right and left end of Tn*916* together with 500 bp of adjacent plasmid sequence from pAM120 flank the *cat* gene from pHV1248. The primer binding sites are shown for the SOEing PCR fragments

Primer	Sequence	Length	Feature
SOEing PCR- Mini-Tn*916*			
916R-1	GA**CCGCGG**GATAGAAGATGGACATAGC	27	*Sac*II site at 5' end
916R-2	**CATTAGGGGATTCATCAG**AAGCACCACCACAATATCATC	39	
CAT-1	GATATTGTGGTGGTGCTT**CTGATGAATCCCCTAATG**	36	Blue sequence = *cat* gene sequence
CAT -2	CTATATCCATCACTATATCCATC**CGACTGTAAAAAGTAC**	39	
916L-1	**GTACTTTTTACAGTCG**CGTGTTGAGATTGATGGATATAG	39	
916L-2	GA**CCCGGG**CCTTATAGTTAGAAATGTAATC	30	*Xma*I site at 5' end
tet(M) gene with RBS and P_{tetM}			
TetM-1	CT**CTCGAG**GAAATCAATGAGTTTTTAACGAC	31	*Xho*I at 5' end
TetM-2	CA**CTCGAG**GC**ACTAGT**CAACATAAAATACACTAAG	35	*Xho*I and *Spe*I at 5' end
xis/int genes			
XI-1	CG**ACTAGT**AAAGCTGTTCGATTGCTGG	27	*Spe*I site at 5' end
XI-2	CA**TCTAGA**TAT**CCGCGG**TCATTTGTACTACTAAGC	35	*Xba*I and *Sac*II at 5' end

Table 5.2 Primers for SOEing PCR and vector construction.

The table lists the primer sequences used in construction of min-Tn*916* and for cloning of the *tet*(M) and *xis*/*int* fragments.

Fragment 1 - Primers were designed to amplify approximately 500 bp of the right end of Tn*916* along with 500bp of plasmid pAM120 to generate a 1 Kb fragment. To facilitate cloning primer 916R-1 was designed with a *Sac*II restriction site at the 5'-end. The sequence of primer 916R-2 contains 18 bp of the *cat* gene and 21 bp of the right end of Tn*916*.

Fragment 2 – Primers were designed to amplify the 1 Kb *cat* gene from pHV1248. Primer CAT-1 contains 18 bp of the right end of Tn*916* and 18 bp of the start of the *cat* gene. Primer CAT-2 contains 22 bp of the left end of Tn*916* and 16 bp of the start of the *cat* gene.

Fragment 3 - Primers 916L-1 and 916L-2 were designed to amplify approximately 500 bp of the left end of Tn*916* and 500bp of the plasmid pAM120 to generate a 1 Kb fragment. The sequence of primer 916L-1 contains 16 bp of the *cat* gene and 23 bp of the left end of Tn*916*. To facilitate cloning a *Xma*I restriction site was included in primer 916L-2.

PCR reactions were carried out using *Taq* DNA polymerase and standard buffer according to manufacturer's instructions under the following conditions: [94^{o}C 4 min] x 1 cycle, [94^{o}C 30 sec, 50-60^{o}C 1.5 min, 72^{o}C 1-3 min] x 25-35 cycles [72^{o}C 10 min] x1, hold 4^{o}C. Annealing temperatures and extension times were adjusted according to the primer set and expected product length.

5.4.2. **Gene cloning**

5.4.2.1. TOPO cloning

After purification 3' A overhangs were added to the PCR fragments mini-Tn*916*, P_{tetM}/*tet*(M) and xis/int. Fragments were incubated with dATP and 0.5 units *Taq* DNA polymerase in standard *Taq* buffer (New England Biolabs, UK) for 12 minutes at 72°C. The fragments were ligated into the pCR2.1-TOPO® vector according to the manufacturers protocol for chemically competent *E. coli* transformation (Invitrogen, UK). This led to generation of plasmids pCR2.1-TOPO::mini-Tn*916*, pCR2.1-TOPO::*tet*(M) and pCR2.1-TOPO::*xis*/*int*.

5.4.2.2. Transformations

TOPO cloning: 2 µl of the TOPO cloning reaction was added to one vial of one-shot chemically competent TOP10 *E. coli* (Invitrogen, UK). Cells were incubated for 5-10 minutes on ice and heat-shocked for 30 seconds at 42°C. Cells were immediately placed on ice before addition of 250 µl SOC medium (Invitrogen, UK) and incubated at 37°C with shaking (200 rpm). Cells were spread on pre-warmed LB agar containing 50 µg ml^{-1} Ap and 40 mg ml^{-1} X-gal and incubated for 15 hours at 37°C (both Sigma-Aldrich, UK).

Chemical transformation: the ligation mixture was added to 50 µl of α-select chemically competent cells (Bioline, UK) thawed on ice. Cells were incubated on ice for 45-60 min before heat shock at 42°C for 45 seconds and returned to ice for 2 minutes. 0.8 ml of pre-warmed SOC medium was added

and cultures were incubated at 37^{o}C with shaking (200 rpm) for 1 hour. Cells were spread on LB agar plates containing appropriate antibiotics before incubation at 37^{o}C for 18 hours.

5.4.2.3. Analysis of transformants

After incubation 10-20 white colonies were selected for analysis. Colonies were inoculated into 10 ml of LB broth containing 50 μg ml^{-1} Ap and incubated for 15 hours with shaking at 37^{o}C. Plasmids were extracted from each of the cultures using a Mini-prep kit and digested with *Eco*RI in standard buffer for 4 hours at 37^{o}C. Digested plasmid and an appropriate molecular weight marker were separated on a 0.8% agarose gel and stained with ethidium bromide to visualise the inserted fragments. Plasmids that appeared to contain the correct insertion (based on size) were sent for sequencing (University of Cambridge).

5.4.3. **Generation of the *tet*(M) and *xis*/*int* fragments**

The *tet*(M) gene including the ribosome binding site and promoter region was PCR amplified from plasmid pAM120 using primers TetM-1 and TetM-2, the *xis*/*int* fragment was amplified from pAM120 using primers XI-1 and XI-2 using the standard PCR conditions. The fragments were separated on a 0.8% agarose gel before extraction and purification. 3' A overhangs were added by incubation of the fragment with dATP and 0.5 units *Taq* DNA polymerase in standard *Taq* buffer for 12 minutes at 72^{o}C. The fragments

were ligated into the TOPO vector and transformed into competent TOP10 *E. coli* as described above.

5.4.4. **Generation of plasmid pKLM1.1 (Figure 5.3)**

pMTL5401F and pCR2.1-TOPO::*tet(M)* (TOPO vector containing the *tet*(M) fragment) were digested with *Xho*I in appropriate buffer for 4 hours at 37^{o}C. The pCR2.1-TOPO::*tet*(M) plasmid digest was separated on a 1% agarose gel and the desired 2.6 Kb fragment extracted and purified. The digested plasmid and *tet*(M) fragment were ligated with T4 DNA ligase and ligase reaction buffer at 20-22^{o}C for 16 hours (New England Biolabs, UK). Ligated plasmids were transformed into competent α-select *E. coli* as described above. Plasmids were selected on LB agar containing 100 µg ml^{-1} Ap and 40 mg ml^{-1} X-gal, with incubation for 16 hours at 37^{o}C. 20 colonies were picked and inoculated into LB broth containing 100 µg ml^{-1} Ap and incubated with shaking at 37^{o}C for 6-8 hours. Plasmids were extracted using a Mini-prep kit and digested with *Apa*LI to confirm the ligation of the P_{tetM}/*tet*(M) fragment in the desired orientation.

5.4.5. **Generation of plasmid pKLM1.2 (Figure 5.4)**

pCR2.1-TOPO::*xis*/*int* (TOPO vector containing the *xis*/*int* fragment) and pKLM1.1 were digested with *Spe*I and *Xba*I for 4 hours at 37^{o}C in an appropriate buffer. Both digests were separated on an agarose gel and the desired fragments (*xis*/*int* ~1.5 Kb; pKLM1.1 ~10 Kb) were extracted and purified. The *xis*/*int* fragment was ligated to linearised pKLM1.1 using T4

DNA ligase and ligase reaction buffer at 20-22^{o}C for 16 hours. The ligation mixture was transformed into competent α-select *E. coli* (Bioline, UK) and transformants selected on LB agar containing 100 µg ml^{-1} Ap and 40 mg ml^{-1} X-gal. 20-100 colonies were inoculated in to LB broth containing 100 µg ml^{-1} Ap and incubated with or without shaking at 37^{o}C for 6-16 hours. Plasmids were extracted using a Mini-prep kit and plasmids digested with *Xho*I and *Sac*II to confirm the ligation.

5.4.6. **Generation of the plasmid pKLM916 (Figure 5.5)**

Plasmid pKLM1.2 was digested with *Sma*I at 25^{o}C for 4 hours and pCR2.1-TOPO::mini-Tn*916* was digested with *Bsr*BI for 4 hours at 37^{o}C. Digested plasmids were separated on agarose gel and the desired fragments extracted and purified (mini-Tn*916* ~ 3 Kb; pKLM1.2 ~ 12 Kb). The two fragments were ligated using T4 DNA ligase in T4 reaction buffer at 20-22^{o}C for 16 hours. The ligation mixture was transformed into competent α-select *E. coli* and transformants selected on LB agar containing 100 µg ml^{-1} Ap and 40 mg ml^{-1} X-gal.

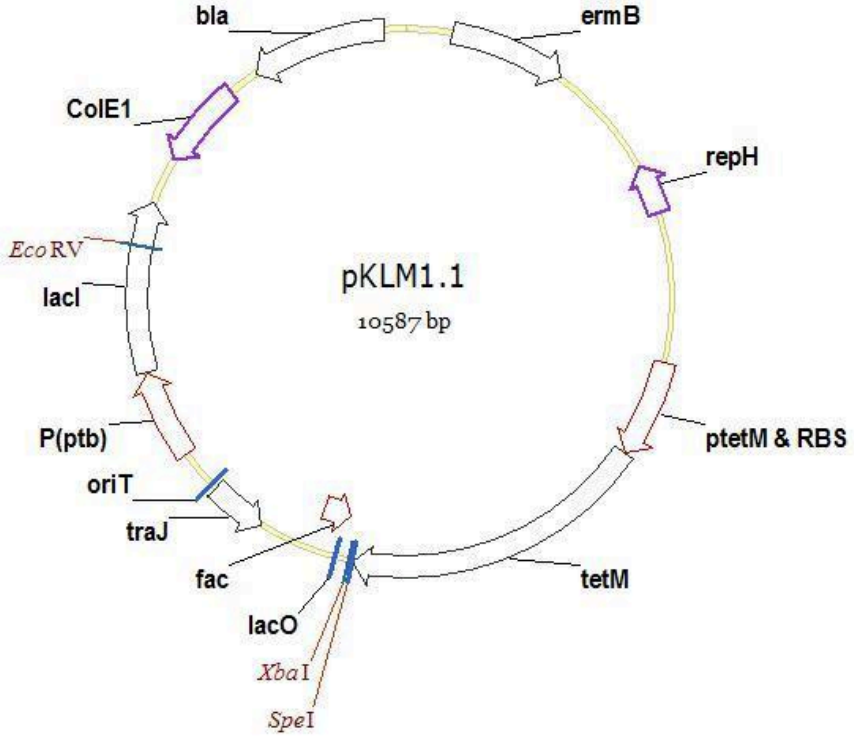


Figure 5.3 Schematic representation of pKLM1.1.

pKLM1.1 is composed of the pMTL5401F backbone with insertion of the P_{tetM} promoter and the *tet*(M) antibiotic resistance determinant

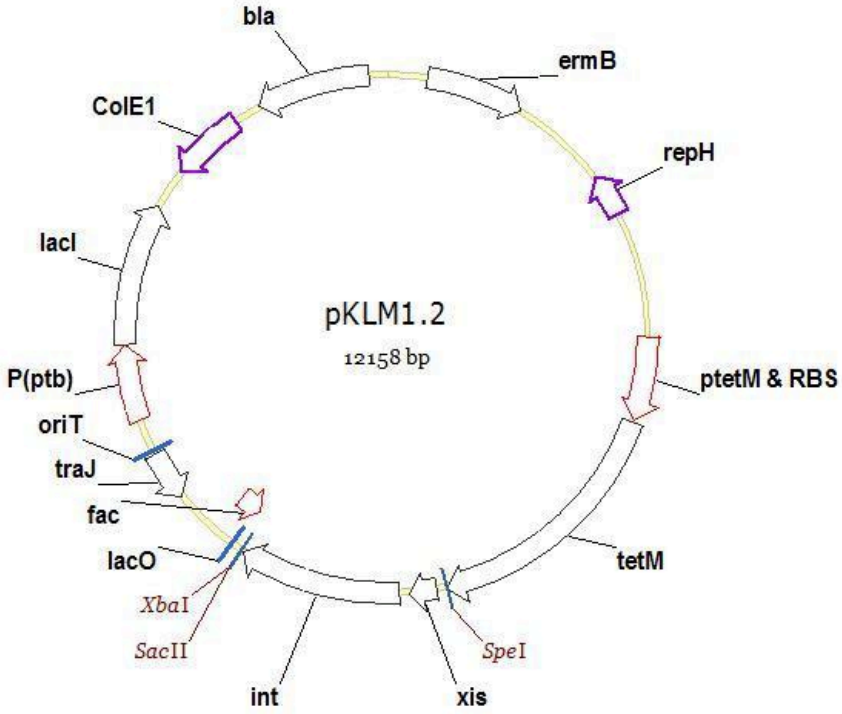


Figure 5.4 Schematic representation of pKLM1.2

pKLM1.2 is composed of the pKLM1.1 with the addition of the *xis* and *int* genes downstream of the P_{tetM} promoter and *tet*(M).

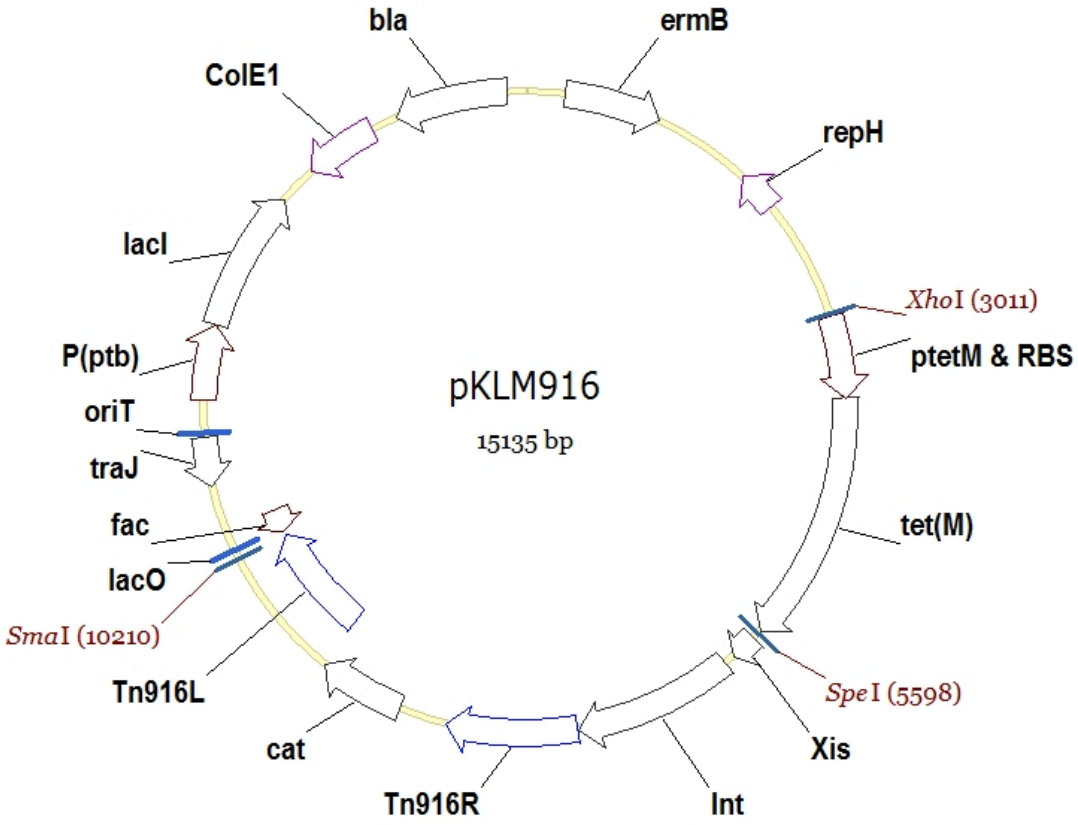


Figure 5.5 Schematic representation of plasmid pKLM916.

pKLM916 is composed of the pKLM1.2 with addition of mini-Tn*916* downstream of the *xis* and *int* genes. Confirmation of the structure of the construct was achieved through PCR and sequencing across the fragment joins.

5.4.7. **Generation of mini-Tn*916*M**

Two approximately 1.2 Kb sections of pAM120 containing either the left or right end of Tn*916* were amplified by primers 916L-3 & 916L-4 and 916R-3 & 916R-4 respectively (Table 5.3) using standard PCR conditions described above except that *Taq* DNA polymerase and *Taq* buffer were replaced by Vent polymerase and buffer (New England Biolabs, UK) to increase fidelity. These fragments were used as template to amplify each of the two mini-Tn*916*M sections – right and left end with primers 916M-1 and 916M-2 and 916M-3 and 916M-4 respectively. Primers 916M-1 and 916M-4 were used to SOE the two fragments together (template in equal ratio) to create a 2 Kb fragment (Figure 5.6). The fragment was cloned into PCR2.1-TOPO::Tn*916*M.

Plasmid pMTL82141 was transformed into α-select competent *E. coli* and transformants were selected on LB agar supplemented with 25 μg ml^{-1} Cm. Plasmid was extracted using a Mini-prep kit [392]. pMTL85141 and pCR2.1-TOPO::Tn*916*M were digested with *Fse*I and *Pme*I. The pMTL82141 digest was separated on agarose gel and the *catP* fragment extracted and purified. Linearised pCR2.1-TOPO::Tn*916*M and the *catP* fragment were ligated with T4 DNA ligase in ligase buffer at 20-22^{o}C for 15 hours. The ligation mixture was transformed into TOP10 competent *E. coli*, and transformants were selected on LB agar supplemented with 25 μg ml^{-1} Cm.

Primer	Sequence 5' - 3'	Length	Feature
916L-3	GGTTGAAGAAGCCTATCAAG	20	
916L-4	CAAAGCGTACCGTTCAAG	18	
916R-3	GTACCTGCATTTCCATGG	18	
916R-4	GACGTAGAGCAGAGAATTTC	20	
916M-1	GATCTAGAGATAGAAGATGGACATAGC	27	*Xba*I
916M-2	CTTTCGATTCGCTAACACTCGCGTTTAAACATATTGGCCGGCCGCCCTCTGCGGTAATACGTGG	64	*Pme*I & *Fse*I
916M-3	CCACGTATTACCGCAGAGGGCGGCCGGCCAATATGTTTAAACGCGAGTGTTAGCGAATCGAAAG	64	*Fse*I & *Pme*I
916M-4	GATCTAGAGCCTTATAGTTAGAAATGTAATC	31	*Xba*I

Table 5.3 Primers used for generation of mini-Tn*916*M

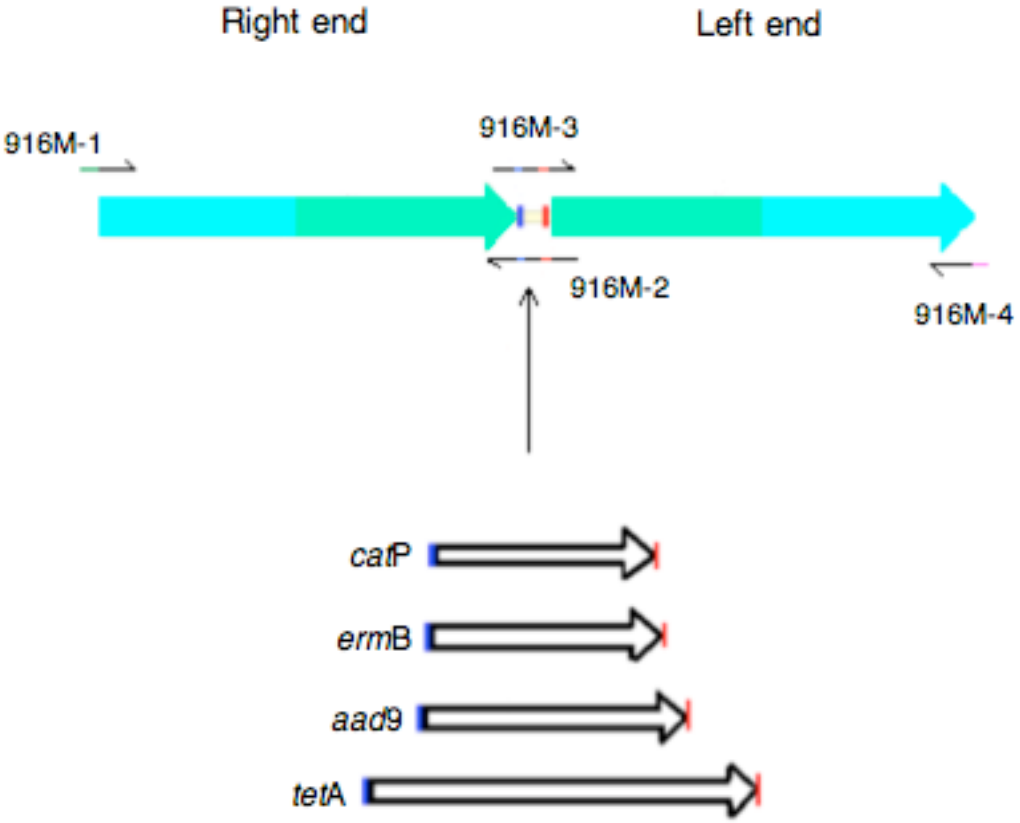


Figure 5.6 Schematic representation of mini-Tn*916*M

SOEing PCR was used to construct the mini-Tn*916*M transposon for use in pKLM1.2.

5.5. Results and discussion

There is a clear need for a random mutagenesis system in *C. difficile*. Attempts to develop a mini-Tn*10* based system failed to lead to demonstrable insertions in the genome of *C. difficile* strains 630Δerm and R20291. Thus, as Tn*916* has been shown to insert into the genome of *C. difficile* at multiple sites, attention turned to the modification of Tn*916* to generate a mini version termed mini-Tn*916*. Like other transposons, Tn*916* contains certain sequences that are essential to their function, while others perform more ancillary functions and can be considered non-essential. Therefore it is possible to make a functional mini-transposon which is significantly smaller than the parent yet retains the ability to excise and insert into DNA.

5.5.1. **Design and construction of mini-Tn*916***

A mini-Tn*916* was designed to include a section of each end of Tn*916,* which are required to work with the *xis* and *int* genes to allow insertion and excision of the transposon. Sequence from plasmid pAM120 flanking either side of Tn*916* was also included to aid mobilisation of the transposon. Initially 500 bp of flanking and transposon sequence was chosen as this was believed to be a suitable length to maintain the small size of the construct and provide sufficient sequence for the actions of Xis and Int to promote transposition. To allow selection of transconjugants containing mini-Tn*916* insertions, the *cat* gene from mini-Tn*10* was included in the initial construct. Each fragment was independently generated before being used as template for SOEing

PCR. Initially a one-step approach was used to generate the 3 Kb construct, this failed to generate the desired fragment although a faint product was observed in repeated reactions containing supplementary $MgCl_2$ (Figure 5.7 & Figure 5.8 lanes 1-4). A PCR using this fragment as template also failed to generate a sufficient quantity of the desired fragment. Modifications to optimise the PCR reaction such as changing the annealing temperature, changing the Mg^{2+} concentration and addition of DMSO failed to sufficiently improve the yield. A two-step approach was then attempted to first ligate two fragments, either one and two, or two and three before addition of the final section. Again various PCR reactions were performed to optimise the yield of the desired fragment. Amplification of fragments two and three led to a high yield of the required 2 Kb fragment, whereas amplification of fragments one and two did not lead to generation of the ligated 2 Kb fragment (Figure 5.8). In general, addition of $MgCl_2$ seems to have improved the stringency of the reactions leading to a greater predominance of the higher molecular weight products in the SOEing PCR reactions. Mg^{2+} ions form a complex with the dNTPs and are required as a co-factor for *Taq* DNA polymerase stabilising the interaction of the primer with the template. The fragment containing sections two and three was then ligated to the final section to generate mini-Tn*916*. SOEing PCR reactions are known to be difficult to optimize and for these reactions an increase in Mg^{2+} ions and a two-step approach enabled successful formation of the chimeric fragment.

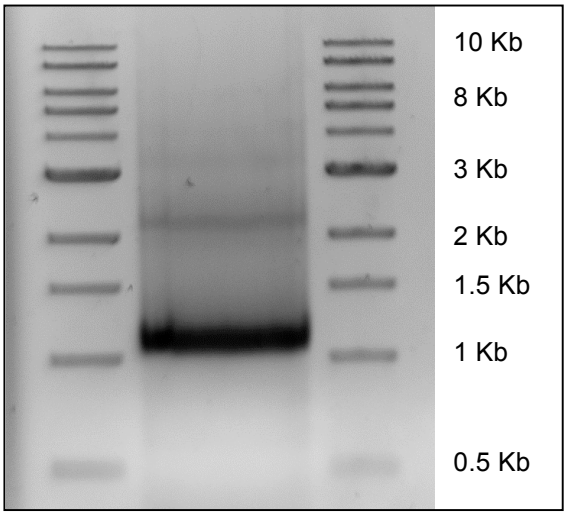


Figure 5.7 Generation of mini-Tn*916* using a 1-step SOEing PCR approach.

All three fragments were used a template to generate the 3 Kb mini-Tn*916* transposon. A 3 Kb was not amplified with a 1 Kb product predominating. A faint 2 Kb product can also be seen. PCR was carried out using standard conditions with a 3 minute extension time.

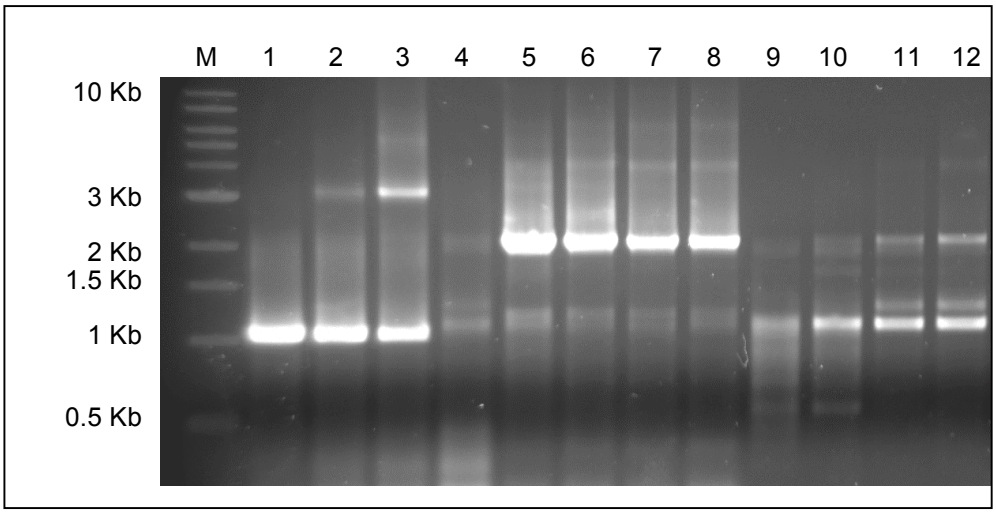


Figure 5.8 Generation of mini-Tn*916* using SOEing PCR in a 1-step and 2-step approach

SOEing PCR was performed with i) equal ratios of all three fragments in a 1-step reaction (lanes 1-4), ii) equal ratios of fragment 2 and 3 for step 1 in a 2-step approach (lanes 5-8) and iii) equal ratios of fragment 1 and 2 for step 1 in a 2-step approach (lanes 9-12). For each group 4 reactions were performed with varying MgCl concentration, either 1.5 mM (lane 1, 5 & 9), 2.5 mM (lane 2, 6 & 10), 5 mM (lane 3, 7 & 11) or 7.5 mM (lane 4, 8 & 12). M = 1 Kb ladder. PCR was performed using standard conditions with an extension time of 3 minutes for reaction i and 2 minutes for reactions ii and iii. The 1-step reaction failed to produce the desired 3 Kb product in reactions 1, 2 & 4, however a 3 Kb product is seen for reaction 3 (lane 3). In the 2-step approach a 2 Kb product is seen for all 4 reactions with fragment 2 and 3 as template, for reaction 3 the desired 2 Kb fragment is not specifically amplified using fragments 1 and 2 as template, although a faint band is seen in lanes 11 & 12 the 1 Kb fragments predominates.

5.5.2. **Design and generation of mini-Tn*916*M**

The inclusion of the *cat* gene in the initial construct was based on concurrent work with mini-Tn*10*. This gene was shown to be a useful marker in *C. difficile* however, difficulties were encountered during selection in *E. coli*. Additionally, a publication from Heap *et al*. (2009) described the generation of a series of plasmids based on a four unit module construction to allow generation of a range of shuttle plasmids for use in *C. difficile* [392]. This helped to form an alternative design for mini-Tn*916* where the antibiotic resistance cassette could be exchanged simply and quickly. Mini-Tn*916*M was made in a similar manner to mini-Tn*916* using SOEing PCR, however this construct was made with two sections, the left and right end of Tn*916*, but contained two unique restriction sites in the centre. The relatively rare restriction sites *Pme*I and *Fse*I are used in the pMTL80000 modular plasmid series to allow exchange of the antibiotic resistance genes (Figure 5.6). Thus, inclusion of these sites means that any of the resistance genes including *catP* (chloramphenicol/thiamphenicol), *aad9* (spectinomycin), *tetA* (tetracycline) and *ermB* (erythromycin) can be included in the mini-Tn*916*M through a simple digestion and ligation reaction. Samples of these plasmids were obtained and initially the *catP* gene was cloned. *catP* appears to be superior to *cat* and has been included in the modular plasmids and in the ClosTron plasmids [392]. Unfortunately, due to time constraints construction of mini-Tn*916*M was not completed.

5.5.3. **Design and construction of the delivery plasmid**

The initial design of mini-Tn*916* was based on the assumption that this would be delivered on a modified version of pMTL5401F, previously used as the vector for delivery of mini-Tn*10*. This appeared to be a suitable delivery plasmid as it contained the relevant features for replication in *E. coli* and *C. difficile*, a selectable marker, transfer genes to allow mobilisation of the plasmid from *E. coli* during mating and an unstable replicon, to facilitate loss of the delivery plasmid after transposition of the element.

In addition to the features present on pMTL5401F, a sufficiently strong promoter was required to allow expression of the *xis* and *int* genes to mobilise mini-Tn*916*. P_{tetM} has been shown to be functional and de-repressible in *C. difficile*, the *tet*(M) gene also offered a secondary means of selection and therefore both elements were included in the design. Finally, the *xis* and *int* genes were placed downstream of the P_{tetM} promoter, as *int* is essential for integration and *xis* enhances excision of mini-Tn*916*. Restriction enzyme cleavage sites were placed into the primer sequences to facilitate cloning of these sections to generate the plasmids pKLM1.1 and pKLM1.2 (Figure 5.4 & Figure 5.3). This was relatively straightforward although a large number of transformant colonies had to be screened to isolate pKLM1.2. Confirmation of the constructs was achieved through restriction digestion and sequencing.

Unfortunately, problems were encountered when attempts were made to clone the mini-Tn*916* section into pKLM1.2. Digestion of pKLM1.2 with the restriction enzyme *Sac*II, chosen as this was not found in pMTL5401F or the

P_{tetM}/*tet*(M) and *xis/int* fragment sequence, generated two fragments. The sequence of pMTL5401F was provided with the plasmid from the Minton Laboratory (University of Nottingham, UK). The sequence did not contain a *Sac*I site however this site was present in the plasmid, confirmed through restriction digestion. This meant that this enzyme could not be used to clone in the mini-Tn*916* fragment. Initially several attempts were made to blunt the *Sac*I-cut end of the mini-Tn*916* fragment from pCR2.1-TOPO::Tn*916* using mung bean nuclease. This would have allowed blunt-ended ligation of the fragment into a *Sma*I site present in pKLM1.2. This failed to generate the desired construct. Investigation of the sequence of pKLM1.2 and pCR2.1-TOPO::Tn*916* revealed that enzyme *Bsr*BI could be used to allow blunt-ended ligation into the pKLM1.2 vector. Unfortunately, due to time constraints generation of pKLM1.3 was not completed and verified, and therefore no experiments were performed in *C. difficile*.

When designing the primers for mini-Tn*916*M the exterior primers were generated with an *Xba*I site rather than a *Sac*I site. Although this would mean that the fragment could insert in either direction, the plasmid constructs could easily be screened for this insertion through restriction analysis. Additionally, the choice of suitable restriction sites was limited due to the sites present on the plasmid backbone and the mini-Tn*916* fragment. Unfortunately, due to time constraints this section of cloning was not completed, the final construct made was plasmid pKLM1.2 (Figure 5.3), this can be further developed to allow delivery of the element into *C. difficile*.

5.6. Conclusions

The initial design of the mini-Tn*916* element was based on information present at the time. Molecular tools for the investigation of *C. difficile* are continually developing and expanding and, as such, the design of both the element and the delivery vector evolved quite rapidly during the relatively slow cloning procedures.

The major development for this idea will be to adapt the restriction sites at the end of the mini-Tn*916* element and the P_{tetM}/*tet*(M) fragment for use with the pMTL80000 modular system [392]. This system was designed to facilitate easy cloning in Clostridia without the need to redesign plasmids for each cloning experiment. The modules allow the user to insert the appropriate sections for each specific project. In the case of this project, these would include a Gram-negative replicon (ColE1), a Gram-positive replicon (RepH from pCB102), a multiple cloning site (MCS), an *ermB* gene for selection and *tra* genes to mediate transfer of the plasmid from the donor to recipient cell. All of these features are available on the plasmids in this series and can be easily cloned to generate the desired vector. The MCS would then allow insertion of the three sections – P_{tetM}/*tet*(M), *xis/int* and mini-Tn*916,* potentially as a single fragment. After construction this could then be tested in a range of *C. difficile* strains to assess whether mini-Tn*916* is functional and to characterise its behaviour. Unfortunately, time contains prevented completion of this project. However, it represents a potentially useful mutagenesis system for *C. difficile*.

6. **General conclusions**

The investigation of the molecular pathogenicity of *C. difficile* has been severely hindered by the lack of available tools. Prior to the commencement of this thesis tools for the random mutagenesis of *C. difficile* were not available. Introduction of DNA is only possible through conjugative transfer, demonstrated by the use of conjugative transposons such as Tn*916* and using *E. coli-Clostridium* shuttle plasmids. Previous investigation of the behaviour of Tn*916* in strain CD37 and 630Δerm found that insertion occurs at a specific hot spot in the former strain, and apparently at random in the latter strain. This thesis describes the generation of a Tn*916* insertion library in the epidemic strain R20291 and investigation of the insertion sites of this conjugative transposon in this strain and in 630Δerm. Sequencing of the insertion sites revealed that Tn*916* inserts into an AT rich 15 bp sequence. This motif is found predominantly in intergenic regions of the genome of strain 630 leading to an insertion bias of this element into these regions. Insertion of Tn*916* into the genome of the epidemic strain R20291 was also found to occur predominantly in intergenic regions. This insertion bias limits the use of this transposon as a random mutagen in these *C. difficile* strains, however, it has the potential to be used to gain a further understanding of these poorly described regions of the genome.

The R20291::Tn*916* insertion library was screened using four assays to identify defects in the ability of the bacterium to cause cytotoxicity, tolerate

bile and *p*-cresol, and in its ability to sporulate/germinate due to insertion of the transposon. Although no non-wild type phenotypes were observed this library is available for further scrutiny with a variety as biological assays to elucidate upon the molecular pathogenesis of *C. difficile*.

The bile screening assays suggest that R20291 is inherently more tolerant to high concentrations of bile than 630Δerm. It was found that low concentrations of bile were more inhibitory than medium to high levels for both strains, which may be due to the inability of low bile levels to elicited the protective response(s) utilised by the bacterium. This was supported by data showing that pre-exposure of cultures to bile could enhance the growth of cultures in bile containing broth. However, bile is reported to act as a germinant for *C. difficile* spores and it is not known whether this can account for the increased growth observed after pre-exposure. Thus this study could be further developed to gain a better understanding of the tolerance of *C. difficile* to bile and the effects of pre-exposure on vegetative cells and spores independently.

To investigate the mechanisms involved in the tolerance of, and response to bile shock, the transcription profile of 630Δerm grown in several bile concentrations was examined. The results of these experiments did not show significant changes in gene expression in bile exposed cultures compared to non-exposed cultures. Reducing the statistical stringency for one of these experiments produced a list of genes whose expression was altered between these two conditions. Interestingly, genes associated with

stress responses failed to show significant changes in expression. As these changes would have been expected, it may suggest that the parameters of the experiment were not optimal; therefore optimisation of the exposure duration and bile concentration are required. Significant differences were observed between R20291 and 630Δerm's ability to tolerate high bile levels therefore, one could carry out similar expression analysis using an R20291 microarray. These data could provide potential insights into the stress response, specific bile tolerance mechanisms and the germination process of *C. difficile,* which can be further investigated by gene knockout studies.

The lack of tools available for the random mutagenesis of *C. difficile* prompted the development of novel tools to generate insertion libraries. A vector has been developed to allow the delivery of mini-Tn*10* to various *C. difficile* strains. It was possible, with low efficiency, to conjugate this vector into strain R20921 and 630Δerm, however transposition of mini-Tn*10* was not demonstrated. Several possible reasons for this outcome have been suggested including problems with expression of the transposase, a lack of activity of the transposase in *C. difficile*, sub-optimal selection of transconjugants or a low frequency transposition. Unfortunately, the precise reason for the apparent lack of transposition events has not been determined and requires further investigation.

A second transposon-based tool has also been partially constructed with a mini derivative of Tn*916*, mini-Tn*916*. The identification of an insertion bias of Tn*916* limits the usefulness of this transposon, however, a mini-version

may prove easier to manipulate and introduce into *C. difficile*. Importantly, preferential insertion into intergenic regions would facilitate investigation of these poorly described regions, potentially providing useful information about gene regulation and genome architecture.

The recent description of a *mariner*-based random mutagenesis tool for *C. difficile* should not detract from the need to develop further tools for the random mutagenesis of *C. difficile*. The availability of multiple tools will facilitate comprehensive investigation of the molecular genetics of this important pathogen.

7. **Acknowledgements**

I thank Professor Nigel Minton, University of Nottingham for supplying *C. difficile* strain R20291 and *E. coli* CA434, plasmids pMTL5401F and pHV1248. I also acknowledge Professor Minton, Dr. John Heap and Dr. Steve Cartman for their useful discussions during this work. Dr. Gemma Langridge, Dr. Tim Perkins and colleagues at the Wellcome Trust Sanger Institute for sequencing the Tn*916* insertion sites in the R20291::Tn*916* and 630Δerm::Tn*916*ΔE library, for determining the insertion motif and for investigating the frequency of this motif in the genome of strain 630. Dr Richard Stabler at LSHTM for running RNA samples on the bioanalyzer and for technical assistance with the transcriptomics work, particularly when assessing the microarray data. Dr. C Clayton for organising the design and supply of primer pairs q16S and qGyrA (RT-PCR) with SigmaGenosis Ltd. Finally, I thank Nazilla Jafari for providing information regarding the growth of *C. difficile* strains R20291 and 630Δerm in cysteine containing media with shaking.

8. **References**

1. Hall I, O'Toole E. Intestinal flora in newborn infants with a description of a new pathogenic anaerobe, *Bacillus difficilis.* American journal of diseases of children 1935;49:390-402.
2. Kelly CP, LaMont JT. *Clostridium difficile* infection. Annu Rev Med 1998;49:375-90.
3. Viscidi R, Willey S, Bartlett JG. Isolation rates and toxigenic potential of *Clostridium difficile* isolates from various patient populations. Gastroenterology 1981;81:5-9.
4. McFarland LV, Brandmarker SA, Guandalini S. Pediatric *Clostridium difficile*: a phantom menace or clinical reality? J Pediatr Gastroenterol Nutr 2000;31:220-31.
5. Bartlett JG, Moon N, Chang TW, Taylor N, Onderdonk AB. Role of *Clostridium difficile* in antibiotic-associated pseudomembranous colitis. Gastroenterology 1978;75:778-82.
6. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeno-Tarraga AM, Wang H, Holden MT, Wright A, Churcher C, Quail MA, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B, Parkhill J. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet 2006;38:779-86.
7. Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum.* J Bacteriol 2001;183:4823-38.
8. Sebaihia M, Peck MW, Minton NP, Thomson NR, Holden MT, Mitchell WJ, Carter AT, Bentley SD, Mason DR, Crossman L, Paul CJ, Ivens A, Wells-Bennik MH, Davis IJ, Cerdeno-Tarraga AM, Churcher C, Quail MA, Chillingworth T, Feltwell T, Fraser A, Goodhead I, Hance Z, Jagels K, Larke N, Maddison M, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, White B, Whithead S, Parkhill J. Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes. Genome Res 2007;17:1082-92.
9. Shimizu T, Ohtani K, Hirakawa H, Ohshima K, Yamashita A, Shiba T, Ogasawara N, Hattori M, Kuhara S, Hayashi H. Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. Proc Natl Acad Sci U S A 2002;99:996-1001.
10. Bruggemann H, Baumer S, Fricke WF, Wiezer A, Liesegang H, Decker I, Herzberg C, Martinez-Arias R, Merkl R, Henne A, Gottschalk G. The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. Proc Natl Acad Sci U S A 2003;100:1316-21.

11. Farrow KA, Lyras D, Rood JI. Genomic analysis of the erythromycin resistance element Tn*5398* from *Clostridium difficile.* Microbiology 2001;147:2717-28.
12. Haraldsen JD, Sonenshein AL. Efficient sporulation in *Clostridium difficile* requires disruption of the *sigmaK* gene. Mol Microbiol 2003;48:811-21.
13. Braun V, Mehlig M, Moos M, Rupnik M, Kalt B, Mahony DE, von Eichel-Streiber C. A chimeric ribozyme in *Clostridium difficile* combines features of group I introns and insertion elements. Mol Microbiol 2000;36:1447-59.
14. Aslam S, Hamill R, Musher D. Treatment of *Clostridium difficile*-associated disease: old therapies and new strategies. Lancet Infect Dis 2005;5:549-57.
15. Borriello SP. Pathogenesis of *Clostridium difficile* infection. J Antimicrob Chemother 1998;41 Suppl C:13-9.
16. Kelly CP, Pothoulakis C, LaMont JT. *Clostridium difficile* colitis. N Engl J Med 1994;330:257-62.
17. Rupnik M, Wilcox MH, Gerding DN. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol 2009;7:526-36.
18. Poxton IR, McCoubrey J, Blair G. The pathogenicity of *Clostridium difficile.* Clin Microbiol Infect 2001;7:421-7.
19. Bartlett J. *Clostridium difficile*-associated Enteric Disease. Curr Infect Dis Rep 2002;4:477-83.
20. McDonald L, Killgore G, Thompson A, Owens RJ, Kazakova S, Sambol S, Johnson S, Gerding D. An epidemic, toxin gene-variant strain of *Clostridium difficile*. N Engl J Med 2005;353:2433-41.
21. Mylonakis E, Ryan ET, Calderwood SB. *Clostridium difficile*--Associated diarrhea: A review. Arch Intern Med 2001;161:525-33.
22. Gan SI, Beck PL. A new look at toxic megacolon: an update and review of incidence, etiology, pathogenesis, and management. Am J Gastroenterol 2003;98:2363-71.
23. Barbut F, Petit J. Epidemiology of *Clostridium difficile*-associated infections. Clin Microbiol Infect 2001;7:405-10.
24. Arroyo L, Kruth S, Willey B, Staempfli H, Low D, Weese J. PCR ribotyping of *Clostridium difficile* isolates originating from human and animal sources. J Med Microbiol 2005;54:163-6.
25. Indra A, Lassnig H, Baliko N, Much P, Fiedler A, Huhulescu S, Allerberger F. *Clostridium difficile*: a new zoonotic agent? Wien Klin Wochenschr 2009;121:91-5.
26. Rupnik M. Is *Clostridium difficile*-associated infection a potentially zoonotic and foodborne disease? Clin Microbiol Infect 2007;13:457-9.
27. Gould L, Limbago B. *Clostridium difficile* in food and domestic animals: a new foodborne pathogen? Clin Infect Dis 2010;51:577-82.
28. Songer J, Trinh H, Killgore G, Thompson A, McDonald L, Limbago B. *Clostridium difficile* in retail meat products, USA, 2007. Emerg Infect Dis 2009;15:819-21.
29. Rodriguez-Palacios A, Reid-Smith R, Staempfli H, Daignault D, Janecko N, Avery B, Martin H, Thomspon A, McDonald L, Limbago B,

Weese J. Possible seasonality of *Clostridium difficile* in retail meat, Canada. Emerg Infect Dis 2009;15:802-5.
30. Weese J, Avery B, Rousseau J, Reid-Smith R. Detection and enumeration of *Clostridium difficile* spores in retail beef and pork. Appl Environ Microbiol 2009;75:5009-11.
31. Weese J, Reid-Smith R, Avery B, Rousseau J. Detection and characterization of *Clostridium difficile* in retail chicken. Lett Appl Microbiol 2010;50:362-5.
32. Bakri M, Brown D, Butcher J, Sutherland A. *Clostridium difficile* in ready-to-eat salads, Scotland. Emerg Infect Dis 2009;15:817-8.
33. Jöbstl M, Heuberger S, Indra A, Nepf R, Köfer J, Wagner M. *Clostridium difficile* in raw products of animal origin. Int J Food Microbiol 2010;138:172-5.
34. Kelly CP, LaMont JT. *Clostridium difficile*--more difficult than ever. N Engl J Med 2008;359:1932-40.
35. Freeman J, Bauer M, Baines S, Corver J, Fawley W, Goorhuis B, Kuijper E, Wilcox M. The changing epidemiology of *Clostridium difficile* infections. Clin Microbiol Rev 2010;23:529-49.
36. Plowman R, Graves N, Griffin M, Roberts JA, Swan AV, Cookson B, Taylor L. The Socio-economic Burden of Hospital Acquired Infection.; 1999.
37. HPA. Surveillance of Healthcare Associated Infections Report: 2008. http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1216193833496. Health Protection Agency Centre for Infections; 2008, pp. 1-53.
38. National *Clostridium difficile* Standards Group: Report to the Department of Health. J Hosp Infect 2004;56 Suppl 1:1-38.
39. Statistics OfN. http://www.statistics.gov.uk/CCI/nscl.asp?ID=5014. 2010.
40. Health-&-Safety-Executive. HSE investigation into outbreaks of *Clostridium difficile* at Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust. http://www.hse.gov.uk/healthservices/hospitalinfect/stokemandeville.htm; 2006.
41. Health-Protection-Agency. Health Protection Report. www.hpa.org.uk/hpr/archives/2010/hpr2410.pdf; 2010.
42. Razavi B, Apisarnthanarak A, Mundy L. *Clostridium difficile*: emergence of hypervirulence and fluoroquinolone resistance. Infection 2007;35:300-7.
43. Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, Frost E, McDonald L. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. Lancet 2005;366:1079-84.
44. National Office of Statistics. Deaths attributed to *C. difficile* and MRSA to 2009 www.statistics.gov.uk/cci/nugget.asp?id=1735 & www.statistics.gov.uk/cci/nugget.asp?id=1067.; 2010.
45. MacCannell D, Louie T, Gregson D, Laverdiere M, Labbe A, Laing F, Henwick S. Molecular analysis of *Clostridium difficile* PCR ribotype 027 isolates from Eastern and Western Canada. J Clin Microbiol 2006;44:2147-52.

46. Matamouros S, England P, Dupuy B. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Mol Microbiol 2007;64:1274-88.
47. Freeman J, Fawley W, Baines S, Wilcox M. Measurement of toxin production by *Clostridium difficile*. Lancet 2006;367:982-3; author reply 3-4.
48. Spigaglia P, Mastrantonio P. Molecular analysis of the pathogenicity locus and polymorphism in the putative negative regulator of toxin production (TcdC) among *Clostridium difficile* clinical isolates. J Clin Microbiol 2002;40:3470-5.
49. Cloud J, Kelly C. Update on *Clostridium difficile* associated disease. Curr Opin Gastroenterol 2007;23:4-9.
50. Geric B, Carman R, Rupnik M, Genheimer C, Sambol S, Lyerly D, Gerding D, Johnson S. Binary toxin-producing, large clostridial toxin-negative *Clostridium difficile* strains are enterotoxic but do not cause disease in hamsters. J Infect Dis 2006;193:1143-50.
51. Fawley WN, Underwood S, Freeman J, Baines SD, Saxton K, Stephenson K, Owens RC, Jr., Wilcox MH. Efficacy of hospital cleaning agents and germicides against epidemic *Clostridium difficile* strains. Infect Control Hosp Epidemiol 2007;28:920-5.
52. Akerlund T, Persson I, Unemo M, Noren T, Svenungsson B, Wullt M, Burman LG. Increased sporulation rate of epidemic *Clostridium difficile* Type 027/NAP1. J Clin Microbiol 2008;46:1530-3.
53. Birgand G, Blanckaert K, Carbonne A, Coignard B, Barbut F, Eckert C, Grandbastien B, Kadi Z, Astagneau P. Investigation of a large outbreak of *Clostridium difficile* PCR-ribotype 027 infections in northern France, 2006-2007 and associated clusters in 2008-2009. Euro Surveill 2010;15.
54. Kuijper EJ, Barbut F, Brazier JS, Kleinkauf N, Eckmanns T, Lambert ML, Drudy D, Fitzpatrick F, Wiuff C, Brown DJ, Coia JE, Pituch H, Reichert P, Even J, Mossong J, Widmer AF, Olsen KE, Allerberger F, Notermans DW, Delmee M, Coignard B, Wilcox M, Patel B, Frei R, Nagy E, Bouza E, Marin M, Akerlund T, Virolainen-Julkunen A, Lyytikainen O, Kotila S, Ingebretsen A, Smyth B, Rooney P, Poxton IR, Monnet DL. Update of *Clostridium difficile* infection due to PCR ribotype 027 in Europe, 2008. Euro Surveill 2008;13.
55. Clements AC, Magalhaes RJ, Tatem AJ, Paterson DL, Riley TV. *Clostridium difficile* PCR ribotype 027: assessing the risks of further worldwide spread. Lancet Infect Dis 2010;10:395-404.
56. Freeman J, Wilcox M. Antibiotics and *Clostridium difficile*. Microbes Infect 1999;1:377-84.
57. Thomas C, Stevenson M, Riley T. Antibiotics and hospital-acquired *Clostridium difficile*-associated diarrhoea: a systematic review. J Antimicrob Chemother 2003;51:1339-50.
58. Delmée M. Laboratory diagnosis of *Clostridium difficile* disease. Clin Microbiol Infect 2001;7:411-6.
59. McFarland L. Antibiotic-associated diarrhea: epidemiology, trends and treatment. Future Microbiol 2008;3:563-78.

60. Scottish Intercollegiate Guidelines Network. Antibiotic prophylaxis in surgery. A national clinical guideline. http://www.sign.ac.uk/pdf/sign104.pdf. 2008.
61. Wren S, Ahmed N, Jamal A, Safadi B. Preoperative oral antibiotics in colorectal surgery increase the rate of *Clostridium difficile* colitis. Arch Surg 2005;140:752-6.
62. Wiström J, Norrby S, Myhre E, Eriksson S, Granström G, Lagergren L, Englund G, Nord C, Svenungsson B. Frequency of antibiotic-associated diarrhoea in 2462 antibiotic-treated hospitalized patients: a prospective study. J Antimicrob Chemother 2001;47:43-50.
63. Starr J, Martin H, McCoubrey J, Gibson G, Poxton I. Risk factors for *Clostridium difficile* colonisation and toxin production. Age Ageing 2003;32:657-60.
64. Bartlett J. Clinical practice. Antibiotic-associated diarrhea. N Engl J Med 2002;346:334-9.
65. Bignardi G. Risk factors for *Clostridium difficile* infection. J Hosp Infect 1998;40:1-15.
66. Owens RJ, Donskey C, Gaynes R, Loo V, Muto C. Antimicrobial-associated risk factors for *Clostridium difficile* infection. Clin Infect Dis 2008;46 Suppl 1:S19-31.
67. Shah S, Lewis A, Leopold D, Dunstan F, Woodhouse K. Gastric acid suppression does not promote clostridial diarrhoea in the elderly. QJM 2000;93:175-81.
68. McFarland L, Clarridge J, Beneda H, Raugi G. Fluoroquinolone use and risk factors for *Clostridium difficile*-associated disease within a Veterans Administration health care system. Clin Infect Dis 2007;45:1141-51.
69. Linsky A, Gupta K, Lawler E, Fonda J, Hermos J. Proton pump inhibitors and risk for recurrent *Clostridium difficile* infection. Arch Intern Med 2010;170:772-8.
70. Jayatilaka S, Shakov R, Eddi R, Bakaj G, Baddoura W, DeBari V. *Clostridium difficile* infection in an urban medical center: five-year analysis of infection rates among adult admissions and association with the use of proton pump inhibitors. Ann Clin Lab Sci 2007;37:241-7.
71. Cunningham R, Dale B, Undy B, Gaunt N. Proton pump inhibitors as a risk factor for *Clostridium difficile* diarrhoea. J Hosp Infect 2003;54:243-5.
72. Mullany P. The dynamic bacterial genome. Cambridge: Cambridge University Press; 2005.
73. Casjens S. The diverse and dynamic structure of bacterial genomes. Annu Rev Genet 1998;32:339-77.
74. He M, Sebaihia M, Lawley T, Stabler R, Dawson L, Martin M, Holt K, Seth-Smith H, Quail M, Rance R, Brooks K, Churcher C, Harris D, Bentley S, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren B, Dougan G, Parkhill J. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proc Natl Acad Sci U S A 2010;107:7527-32.

75. Spigaglia P, Mastrantonio P. Comparative analysis of *Clostridium difficile* clinical isolates belonging to different genetic lineages and time periods. J Med Microbiol 2004;53:1129-36.
76. Musher D, Aslam S, Logan N, Nallacheru S, Bhaila I, Borchert F, Hamill R. Relatively poor outcome after treatment of *Clostridium difficile* colitis with metronidazole. Clin Infect Dis 2005;40:1586-90.
77. Pepin J, Alary M, Valiquette L, Raiche E, Ruel J, Fulop K, Godin D, Bourassa C. Increasing risk of relapse after treatment of *Clostridium difficile* colitis in Quebec, Canada. Clin Infect Dis 2005;40:1591-7.
78. Hinkson P, Dinardo C, DeCiero D, Klinger J, Barker RJ. Tolevamer, an anionic polymer, neutralizes toxins produced by the BI/027 strains of *Clostridium difficile*. Antimicrob Agents Chemother 2008;52:2190-5.
79. Johnson S. Recurrent *Clostridium difficile* infection: a review of risk factors, treatments, and outcomes. J Infect 2009;58:403-10.
80. Klein E, Boster D, Stapp J, Wells J, Qin X, Clausen C, Swerdlow D, Braden C, Tarr P. Diarrhea etiology in a Children's Hospital Emergency Department: a prospective cohort study. Clin Infect Dis 2006;43:807-13.
81. McFarland L. Update on the changing epidemiology of *Clostridium difficile*-associated disease. Nat Clin Pract Gastroenterol Hepatol 2008;5:40-8.
82. Kyne L, Warny M, Qamar A, Kelly C. Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. Lancet 2001;357:189-93.
83. Torres J, Lyerly D, Hill J, Monath T. Evaluation of formalin-inactivated *Clostridium difficile* vaccines administered by parenteral and mucosal routes of immunization in hamsters. Infect Immun 1995;63:4619-27.
84. Ryan E, Butterton J, Smith R, Carroll P, Crean T, Calderwood S. Protective immunity against *Clostridium difficile* toxin A induced by oral immunization with a live, attenuated *Vibrio cholerae* vector strain. Infect Immun 1997;65:2941-9.
85. Ward S, Douce G, Figueiredo D, Dougan G, Wren B. Immunogenicity of a *Salmonella typhimurium* aroA aroD vaccine expressing a nontoxic domain of *Clostridium difficile* toxin A. Infect Immun 1999;67:2145-52.
86. Pavliakova D, Moncrief J, Lyerly D, Schiffman G, Bryla D, Robbins J, Schneerson R. *Clostridium difficile* recombinant toxin A repeating units as a carrier protein for conjugate vaccines: studies of pneumococcal type 14, *Escherichia coli* K1, and *Shigella flexneri* type 2a polysaccharides in mice. Infect Immun 2000;68:2161-6.
87. Dendukuri N, Costa V, McGregor M, Brophy J. Probiotic therapy for the prevention and treatment of *Clostridium difficile*-associated diarrhea: a systematic review. CMAJ 2005;173:167-70.
88. Dendukuri N, Brophy J. Inappropriate use of meta-analysis to estimate efficacy of probiotics. Am J Gastroenterol 2007;102:201; author reply 2-4.
89. Lewis S. Response to the article: McFarland LV. Meta-analysis of probiotics for the prevention of antibiotic-associated diarrhea and the treatment of *Clostridium difficile* disease. Am J Gastroenterol 2006; 101:812-22. Am J Gastroenterol 2007;102:201-2.

90. Dubberke E, Wertheimer A. Review of current literature on the economic burden of *Clostridium difficile* infection. Infect Control Hosp Epidemiol 2009;30:57-66.
91. National Audit Office. Improving patient care by reducing the risk of hospital acquired infections: a progresss report. http://www.nao.org.uk/publications/0304/improving_patient_care.aspx . 2004.
92. Department of Health, Chief Medical Officer. Winning ways: working together to reduce healthcare associated infection in England. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4064682. 2003.
93. Gerding D, Muto C, Owens RJ. Measures to control and prevent *Clostridium difficile* infection. Clin Infect Dis 2008;46 Suppl 1:S43-9.
94. Salyers AA, Whitt DD. Bacterial pathogenesis : a molecular approach. Washington, D.C.: ASM Press; 1994.
95. Borriello S, Davies H, Kamiya S, Reed P, Seddon S. Virulence factors of *Clostridium difficile*. Rev Infect Dis 1990;12 Suppl 2:S185-91.
96. Johnson S, Gerding D. *Clostridium difficile*--associated diarrhea. Clin Infect Dis 1998;26:1027-34; quiz 35-6.
97. Voth DE, Ballard JD. *Clostridium difficile* toxins: mechanism of action and role in disease. Clin Microbiol Rev 2005;18:247-63.
98. Voth DE, Martinez OV, Ballard JD. Variations in lethal toxin and cholesterol-dependent cytolysin production correspond to differences in cytotoxicity among strains of *Clostridium sordellii.* FEMS Microbiol Lett 2006;259:295-302.
99. Busch C, Schömig K, Hofmann F, Aktories K. Characterization of the catalytic domain of *Clostridium novyi* alpha-toxin. Infect Immun 2000;68:6378-83.
100. Amimoto K, Noro T, Oishi E, Shimizu M. A novel toxin homologous to large clostridial cytotoxins found in culture supernatant of *Clostridium perfringens* type C. Microbiology 2007;153:1198-206.
101. Jank T, Aktories K. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. Trends Microbiol 2008;16:222-9.
102. Thelestam M, Chaves-Olarte E. Cytotoxic effects of the *Clostridium difficile* toxins. Curr Top Microbiol Immunol 2000;250:85-96.
103. Schirmer J, Aktories K. Large clostridial cytotoxins: cellular biology of Rho/Ras-glucosylating toxins. Biochim Biophys Acta 2004;1673:66-74.
104. Just I, Gerhard R. Large clostridial cytotoxins. Rev Physiol Biochem Pharmacol 2004;152:23-47.
105. Rupnik M, Pabst S, von Eichel-Streiber C, Urlaub H, Söling H. Characterization of the cleavage site and function of resulting cleavage fragments after limited proteolysis of *Clostridium difficile* toxin B (TcdB) by host cells. Microbiology 2005;151:199-208.
106. Carter G, Rood J, Lyras D. The role of toxin A and toxin B in *Clostridium difficile*-associated disease: Past and present perspectives. Gut Microbes 2010;1:58-64.
107. Pothoulakis C. Effects of *Clostridium difficile* toxins on epithelial cell barrier. Ann N Y Acad Sci 2000;915:347-56.

108. Collier R. Understanding the mode of action of diphtheria toxin: a perspective on progress during the 20th century. Toxicon 2001;39:1793-803.
109. Ho J, Greco A, Rupnik M, Ng K. Crystal structure of receptor-binding C-terminal repeats from *Clostridium difficile* toxin A. Proc Natl Acad Sci U S A 2005;102:18373-8.
110. Na X, Kim H, Moyer M, Pothoulakis C, LaMont J. gp96 is a human colonocyte plasma membrane binding protein for *Clostridium difficile* toxin A. Infect Immun 2008;76:2862-71.
111. Egerer M, Giesemann T, Jank T, Satchell K, Aktories K. Auto-catalytic cleavage of *Clostridium difficile* toxins A and B depends on cysteine protease activity. J Biol Chem 2007;282:25314-21.
112. Reineke J, Tenzer S, Rupnik M, Koschinski A, Hasselmayer O, Schrattenholz A, Schild H, von Eichel-Streiber C. Autocatalytic cleavage of *Clostridium difficile* toxin B. Nature 2007;446:415-9.
113. Braun V, Hundsberger T, Leukel P, Sauerborn M, von Eichel-Streiber C. Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. Gene 1996;181:29-38.
114. Cohen S, Tang Y, Silva JJ. Analysis of the pathogenicity locus in *Clostridium difficile* strains. J Infect Dis 2000;181:659-63.
115. Rupnik M, Dupuy B, Fairweather NF, Gerding DN, Johnson S, Just I, Lyerly DM, Popoff MR, Rood JI, Sonenshein AL, Thelestam M, Wren BW, Wilkins TD, von Eichel-Streiber C. Revised nomenclature of *Clostridium difficile* toxins and associated genes. J Med Microbiol 2005;54:113-7.
116. Mani N, Dupuy B. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proc Natl Acad Sci U S A 2001;98:5844-9.
117. Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL. Repression of *Clostridium difficile* toxin gene expression by CodY. Mol Microbiol 2007;66:206-19.
118. Tan K, Wee B, Song K. Evidence for holin function of *tcdE* gene in the pathogenicity of *Clostridium difficile*. J Med Microbiol 2001;50:613-9.
119. Rupnik M, Avesani V, Janc M, von Eichel-Streiber C, Delmée M. A novel toxinotyping scheme and correlation of toxinotypes with serogroups of *Clostridium difficile* isolates. J Clin Microbiol 1998;36:2240-7.
120. Rupnik M. Heterogeneity of large clostridial toxins: importance of Clostridium difficile toxinotypes. FEMS Microbiol Rev 2008;32:541-55.
121. Stabler RA, He M, Dawson L, Martin M, Valiente E, Corton C, Lawley TD, Sebaihia M, Quail MA, Rose G, Gerding DN, Gibert M, Popoff MR, Parkhill J, Dougan G, Wren BW. Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. Genome Biol 2009;10:R102.
122. Alfa M, Kabani A, Lyerly D, Moncrief S, Neville L, Al-Barrak A, Harding G, Dyck B, Olekson K, Embil J. Characterization of a toxin A-negative, toxin B-positive strain of *Clostridium difficile* responsible for a nosocomial outbreak of *Clostridium difficile*-associated diarrhea. J Clin Microbiol 2000;38:2706-14.

123. Drudy D, Fanning S, Kyne L. Toxin A-negative, toxin B-positive *Clostridium difficile*. Int J Infect Dis 2007;11:5-10.
124. Riegler M, Sedivy R, Pothoulakis C, Hamilton G, Zacherl J, Bischof G, Cosentini E, Feil W, Schiessel R, LaMont JT, Wenzl E. *Clostridium difficile* toxin B is more potent than toxin A in damaging human colonic epithelium in vitro. J Clin Invest 1995;95:2004-11.
125. Lyerly D, Saum K, MacDonald D, Wilkins T. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun 1985;47:349-52.
126. Lyras D, O'Connor JR, Howarth PM, Sambol SP, Carter GP, Phumoonna T, Poon R, Adams V, Vedantam G, Johnson S, Gerding DN, Rood JI. Toxin B is essential for virulence of *Clostridium difficile.* Nature 2009;458:1176-9.
127. Sambol S, Tang J, Merrigan M, Johnson S, Gerding D. Infection of hamsters with epidemiologically important strains of *Clostridium difficile*. J Infect Dis 2001;183:1760-6.
128. Douce G, Goulding D. Refinement of the hamster model of *Clostridium difficile* disease. Methods Mol Biol 2010;646:215-27.
129. Kuehne S, Cartman S, Heap J, Kelly M, Cockayne A, Minton N. The role of toxin A and toxin B in *Clostridium difficile* infection. Nature 2010;467:711-3.
130. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium.* J Microbiol Methods 2007;70:452-64.
131. Hussain HA, Roberts AP, Mullany P. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn*916*DeltaE enters the genome of this strain at multiple sites. J Med Microbiol 2005;54:137-41.
132. O'Connor JR, Lyras D, Farrow KA, Adams V, Powell DR, Hinds J, Cheung JK, Rood JI. Construction and analysis of chromosomal *Clostridium difficile* mutants. Mol Microbiol 2006;61:1335-51.
133. Karlsson S, Burman LG, Akerlund T. Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. Microbiology 1999;145 (Pt 7):1683-93.
134. Dupuy B, Sonenshein A. Regulated transcription of *Clostridium difficile* toxin genes. Mol Microbiol 1998;27:107-20.
135. Karlsson S, Dupuy B, Mukherjee K, Norin E, Burman L, Akerlund T. Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. Infect Immun 2003;71:1784-93.
136. Karlsson S, Lindberg A, Norin E, Burman LG, Akerlund T. Toxins, butyric acid, and other short-chain fatty acids are coordinately expressed and down-regulated by cysteine in *Clostridium difficile.* Infect Immun 2000;68:5881-8.
137. Gonçalves C, Decré D, Barbut F, Burghoffer B, Petit J. Prevalence and characterization of a binary toxin (actin-specific ADP-ribosyltransferase) from *Clostridium difficile*. J Clin Microbiol 2004;42:1933-9.

138. Rupnik M, Grabnar M, Geric B. Binary toxin producing *Clostridium difficile* strains. Anaerobe 2003;9:289-94.
139. Carter GP, Lyras D, Allen DL, Mackin KE, Howarth PM, O'Connor JR, Rood JI. Binary toxin production in *Clostridium difficile* is regulated by CdtR, a LytTR family response regulator. J Bacteriol 2007;189:7290-301.
140. Schwan C, Stecher B, Tzivelekidis T, van Ham M, Rohde M, Hardt W, Wehland J, Aktories K. *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. PLoS Pathog 2009;5:e1000626.
141. Stubbs S, Rupnik M, Gibert M, Brazier J, Duerden B, Popoff M. Production of actin-specific ADP-ribosyltransferase (binary toxin) by strains of *Clostridium difficile*. FEMS Microbiol Lett 2000;186:307-12.
142. Perelle S, Gibert M, Bourlioux P, Corthier G, Popoff M. Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. Infect Immun 1997;65:1402-7.
143. Jhung M, Thompson A, Killgore G, Zukowski W, Songer G, Warny M, Johnson S, Gerding D, McDonald L, Limbago B. Toxinotype V *Clostridium difficile* in humans and food animals. Emerg Infect Dis 2008;14:1039-45.
144. Hecht G. Innate mechanisms of epithelial host defense: spotlight on intestine. Am J Physiol 1999;277:C351-8.
145. Delmée M, Avesani V, Delferriere N, Burtonboy G. Characterization of flagella of *Clostridium difficile* and their role in serogrouping reactions. J Clin Microbiol 1990;28:2210-4.
146. Tasteyre A, Barc M, Collignon A, Boureau H, Karjalainen T. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. Infect Immun 2001;69:7937-40.
147. McCoubrey J, Starr J, Martin H, Poxton IR. *Clostridium difficile* in a geriatric unit: a prospective epidemiological study employing a novel S-layer typing method. J Med Microbiol 2003;52:573-8.
148. Bruggemann H, Gottschalk G. Clostridia : molecular biology in the post-genomic era. Wymondham: Caister Academic; 2009.
149. Janoir C, Péchiné S, Grosdidier C, Collignon A. Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. J Bacteriol 2007;189:7174-80.
150. Wright A, Wait R, Begum S, Crossett B, Nagy J, Brown K, Fairweather N. Proteomic analysis of cell surface proteins from *Clostridium difficile*. Proteomics 2005;5:2443-52.
151. Waligora AJ, Hennequin C, Mullany P, Bourlioux P, Collignon A, Karjalainen T. Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. Infect Immun 2001;69:2144-53.
152. Calabi E, Calabi F, Phillips A, Fairweather N. Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. Infect Immun 2002;70:5770-8.
153. Sára M, Sleytr U. S-Layer proteins. J Bacteriol 2000;182:859-68.
154. Borriello SP, Welch AR, Barclay FE, Davies HA. Mucosal association by *Clostridium difficile* in the hamster gastrointestinal tract. J Med Microbiol 1988;25:191-6.

155. Lawley TD, Croucher NJ, Yu L, Clare S, Sebaihia M, Goulding D, Pickard DJ, Parkhill J, Choudhary J, Dougan G. Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. J Bacteriol 2009;191:5377-86.
156. Vonberg RP, Kuijper EJ, Wilcox MH, Barbut F, Tull P, Gastmeier P, van den Broek PJ, Colville A, Coignard B, Daha T, Debast S, Duerden BI, van den Hof S, van der Kooi T, Maarleveld HJ, Nagy E, Notermans DW, O'Driscoll J, Patel B, Stone S, Wiuff C. Infection control measures to limit the spread of *Clostridium difficile.* Clin Microbiol Infect 2008;14 Suppl 5:2-20.
157. Setlow P. I will survive: DNA protection in bacterial spores. Trends Microbiol 2007;15:172-80.
158. Savage AM, Alford RH. Nosocomial spread of *Clostridium difficile.* Infect Control 1983;4:31-3.
159. Stragier P, Losick R. Molecular genetics of sporulation in *Bacillus subtilis.* Annu Rev Genet 1996;30:297-41.
160. Errington J. *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. Microbiol Rev 1993;57:1-33.
161. Hoch JA. Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis.* Annu Rev Microbiol 1993;47:441-65.
162. Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, Wilcox MH, Stephenson K. Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J Bacteriol 2009;191:7296-305.
163. Stephenson K, Lewis RJ. Molecular insights into the initiation of sporulation in Gram-positive bacteria: new technologies for an old phenomenon. FEMS Microbiol Rev 2005;29:281-301.
164. Worner K, Szurmant H, Chiang C, Hoch JA. Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum.* Mol Microbiol 2006;59:1000-12.
165. Akerlund T, Svenungsson B, Lagergren A, Burman L. 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. J Clin Microbiol 2006;44:353-8.
166. Kamiya S, Ogura H, Meng X, Nakamura S. Correlation between cytotoxin production and sporulation in *Clostridium difficile*. J Med Microbiol 1992;37:206-10.
167. Karlsson S, Burman L, Akerlund T. Induction of toxins in *Clostridium difficile* is associated with dramatic changes of its metabolism. Microbiology 2008;154:3430-6.
168. Wilson KH, Kennedy MJ, Fekety FR. Use of sodium taurocholate to enhance spore recovery on a medium selective for *Clostridium difficile.* J Clin Microbiol 1982;15:443-6.
169. Sorg JA, Sonenshein AL. Bile salts and glycine as cogerminants for *Clostridium difficile* spores. J Bacteriol 2008;190:2505-12.
170. Sorg JA, Sonenshein AL. Chenodeoxycholate is an inhibitor of *Clostridium difficile* spore germination. J Bacteriol 2009;191:1115-7.
171. Paredes-Sabja D, Bond C, Carman RJ, Setlow P, Sarker MR. Germination of spores of *Clostridium difficile* strains, including isolates

from a hospital outbreak of *Clostridium difficile*-associated disease (CDAD). Microbiology 2008;154:2241-50.
172. Wilcox MH, Fawley WN. Hospital disinfectants and spore formation by *Clostridium difficile.* Lancet 2000;356:1324.
173. Burns DA, Heap JT, Minton NP. SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. J Bacteriol 2010;192:657-64.
174. Cartman ST, Heap JT, Kuehne SA, Cockayne A, Minton NP. The emergence of 'hypervirulence' in *Clostridium difficile.* Int J Med Microbiol 2010;300:387-95.
175. Gunn JS. Mechanisms of bacterial resistance and response to bile. Microbes Infect 2000;2:907-13.
176. Wilson KH. Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination. J Clin Microbiol 1983;18:1017-9.
177. Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol 2002;46:439-52.
178. Stabler RA, Gerding DN, Songer JG, Drudy D, Brazier JS, Trinh HT, Witney AA, Hinds J, Wren BW. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J Bacteriol 2006;188:7297-305.
179. Hayes F. Transposon-based strategies for microbial functional genomics and proteomics. Annu Rev Genet 2003;37:3-29.
180. Comfort NC. From controlling elements to transposons: Barbara McClintock and the Nobel Prize. Trends Genet 2001;17:475-8.
181. Haniford DB. Transpososome dynamics and regulation in Tn*10* transposition. Crit Rev Biochem Mol Biol 2006;41:407-24.
182. Plasterk RH. RNA silencing: the genome's immune system. Science 2002;296:1263-5.
183. Reznikoff WS. Tn*5* as a model for understanding DNA transposition. Mol Microbiol 2003;47:1199-206.
184. Reznikoff WS. Transposon Tn*5.* Annu Rev Genet 2008;42:269-86.
185. Haren L, Ton-Hoang B, Chandler M. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol 1999;53:245-81.
186. Clewell DB, Flannagan SE, Ike Y, Jones JM, Gawron-Burke C. Sequence analysis of termini of conjugative transposon Tn*916.* J Bacteriol 1988;170:3046-52.
187. Scott JR, Churchward GG. Conjugative transposition. Annu Rev Microbiol 1995;49:367-97.
188. Salyers AA, Shoemaker NB, Stevens AM, Li LY. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol Rev 1995;59:579-90.
189. Clewell DB, Flannagan SE, Jaworski DD. Unconstrained bacterial promiscuity: the Tn*916*-Tn*1545* family of conjugative transposons. Trends Microbiol 1995;3:229-36.
190. Clewell DB, Gawron-Burke C. Conjugative transposons and the dissemination of antibiotic resistance in streptococci. Annu Rev Microbiol 1986;40:635-59.

191. Rice LB. Tn*916* family conjugative transposons and dissemination of antimicrobial resistance determinants. Antimicrob Agents Chemother 1998;42:1871-7.
192. Courvalin P, Carlier C. Transposable multiple antibiotic resistance in *Streptococcus pneumoniae.* Mol Gen Genet 1986;205:291-7.
193. Mullany P, Wilks M, Puckey L, Tabaqchali S. Gene cloning in *Clostridium difficile* using Tn*916* as a shuttle conjugative transposon. Plasmid 1994;31:320-3.
194. Mullany P, Wilks M, Tabaqchali S. Transfer of Tn*916* and Tn*916* delta E into *Clostridium* difficile: demonstration of a hot-spot for these elements in the *C. difficile* genome. FEMS Microbiol Lett 1991;63:191-4.
195. Wang H, Roberts AP, Mullany P. DNA sequence of the insertional hot spot of Tn*916* in the *Clostridium difficile* genome and discovery of a Tn*916*-like element in an environmental isolate integrated in the same hot spot. FEMS Microbiol Lett 2000;192:15-20.
196. Day WA, Jr., Rasmussen SL, Carpenter BM, Peterson SN, Friedlander AM. Microarray analysis of transposon insertion mutations in *Bacillus anthracis*: global identification of genes required for sporulation and germination. J Bacteriol 2007;189:3296-301.
197. Petit MA, Bruand C, Janniere L, Ehrlich SD. Tn*10*-derived transposons active in *Bacillus subtilis.* J Bacteriol 1990;172:6736-40.
198. Steinmetz M, Richter R. Easy cloning of mini-Tn*10* insertions from the *Bacillus subtilis* chromosome. J Bacteriol 1994;176:1761-3.
199. Hartl D, Lohe A, Lozovskaya E. Modern thoughts on an ancyent marinere: function, evolution, regulation. Annu Rev Genet 1997;31:337-58.
200. Rubin E, Akerley B, Novik V, Lampe D, Husson R, Mekalanos J. In vivo transposition of *mariner*-based elements in enteric bacteria and mycobacteria. Proc Natl Acad Sci U S A 1999;96:1645-50.
201. Hendrixson D, Akerley B, DiRita V. Transposon mutagenesis of *Campylobacter jejuni* identifies a bipartite energy taxis system required for motility. Mol Microbiol 2001;40:214-24.
202. Murray G, Morel V, Cerqueira G, Croda J, Srikram A, Henry R, Ko A, Dellagostin O, Bulach D, Sermswan R, Adler B, Picardeau M. Genome-wide transposon mutagenesis in pathogenic *Leptospira* species. Infect Immun 2009;77:810-6.
203. Roberts AP, Hennequin C, Elmore M, Collignon A, Karjalainen T, Minton N, Mullany P. Development of an integrative vector for the expression of antisense RNA in *Clostridium difficile.* J Microbiol Methods 2003;55:617-24.
204. Williams DR, Young DI, Young M. Conjugative plasmid transfer from Escherichia coli to *Clostridium acetobutylicum*. J Gen Microbiol 1990;136:819-26.
205. Minton N, Carter G, Herbert M, O'Keeffe T, Purdy D, Elmore M, Ostrowski A, Pennington O, Davis I. The development of *Clostridium difficile* genetic systems. Anaerobe 2004;10:75-84.
206. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods 2009.

207. Hensel M, Holden DW. Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi. Microbiology 1996;142 (Pt 5):1049-58.
208. Burrus V, Pavlovic G, Decaris B, Guedon G. Conjugative transposons: the tip of the iceberg. Mol Microbiol 2002;46:601-10.
209. Franke AE, Clewell DB. Evidence for a chromosome-borne resistance transposon (Tn*916*) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. J Bacteriol 1981;145:494-502.
210. Roberts AP, Mullany P. A modular master on the move: the Tn*916* family of mobile genetic elements. Trends Microbiol 2009;17:251-8.
211. Caparon MG, Scott JR. Excision and insertion of the conjugative transposon Tn*916* involves a novel recombination mechanism. Cell 1989;59:1027-34.
212. Scott JR, Bringel F, Marra D, Van Alstine G, Rudy CK. Conjugative transposition of Tn*916*: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate. Mol Microbiol 1994;11:1099-108.
213. Courvalin P, Carlier C. Tn*1545*: a conjugative shuttle transposon. Mol Gen Genet 1987;206:259-64.
214. Caillaud F, Carlier C, Courvalin P. Physical analysis of the conjugative shuttle transposon Tn*1545.* Plasmid 1987;17:58-60.
215. Caillaud F, Courvalin P. Nucleotide sequence of the ends of the conjugative shuttle transposon Tn*1545.* Mol Gen Genet 1987;209:110-5.
216. Jennert KC, Tardif C, Young DI, Young M. Gene transfer to *Clostridium cellulolyticum* ATCC 35319. Microbiology 2000;146 Pt 12:3071-80.
217. Chang TW, Lauermann M, Bartlett JG. Cytotoxicity assay in antibiotic-associated colitis. J Infect Dis 1979;140:765-70.
218. Brazier JS. The diagnosis of *Clostridium difficile*-associated disease. J Antimicrob Chemother 1998;41 Suppl C:29-40.
219. Knoop FC, Owens M, Crocker IC. *Clostridium difficile*: clinical disease and diagnosis. Clin Microbiol Rev 1993;6:251-65.
220. Jacobs J, Rudensky B, Dresner J, Berman A, Sonnenblick M, van Dijk Y, Yinnon AM. Comparison of four laboratory tests for diagnosis of *Clostridium difficile*-associated diarrhea. Eur J Clin Microbiol Infect Dis 1996;15:561-6.
221. Lyerly DM, Krivan HC, Wilkins TD. *Clostridium difficile*: its disease and toxins. Clin Microbiol Rev 1988;1:1-18.
222. Yucesoy M, McCoubrey J, Brown R, Poxton IR. Detection of toxin production in *Clostridium difficile* strains by three different methods. Clin Microbiol Infect 2002;8:413-8.
223. Musher DM, Manhas A, Jain P, Nuila F, Waqar A, Logan N, Marino B, Graviss EA. Detection of *Clostridium difficile* toxin: comparison of enzyme immunoassay results with results obtained by cytotoxicity assay. J Clin Microbiol 2007;45:2737-9.
224. Walters BA, Roberts R, Stafford R, Seneviratne E. Relapse of antibiotic associated colitis: endogenous persistence of *Clostridium difficile* during vancomycin therapy. Gut 1983;24:206-12.

225. Vanholder R, De Smet R, Lesaffer G. *p*-cresol: a toxin revealing many neglected but relevant aspects of uraemic toxicity. Nephrol Dial Transplant 1999;14:2813-5.
226. Scheline RR. Metabolism of phenolic acids by the rat intestinal microflora. Acta Pharmacol Toxicol (Copenh) 1968;26:189-205.
227. Hafiz S, Oakley CL. *Clostridium difficile*: isolation and characteristics. J Med Microbiol 1976;9:129-36.
228. Dawson LF, Stabler RA, Wren BW. Assessing the role of *p*-cresol tolerance in *Clostridium difficile*. J Med Microbiol 2008;57:745-9.
229. Selmer T, Andrei PI. *p*-Hydroxyphenylacetate decarboxylase from Clostridium difficile. A novel glycyl radical enzyme catalysing the formation of *p*-cresol. Eur J Biochem 2001;268:1363-72.
230. Hächler H, Kayser F, Berger-Bächi B. Homology of a transferable tetracycline resistance determinant of *Clostridium difficile* with *Streptococcus* (*Enterococcus*) *faecalis* transposon Tn*916*. Antimicrob Agents Chemother 1987;31:1033-8.
231. Weyer E, Rettger L. A comparative study of six different strains of the organism commonly concerned in large-scale production of butyl alcohol and acetone by the biological process. J Bacteriol 1927;14:399-424.
232. Gawron-Burke C, Clewell DB. Regeneration of insertionally inactivated streptococcal DNA fragments after excision of transposon Tn*916* in *Escherichia coli*: strategy for targeting and cloning of genes from gram-positive bacteria. J Bacteriol 1984;159:214-21.
233. Roberts AP, Cheah G, Ready D, Pratten J, Wilson M, Mullany P. Transfer of Tn*916*-like elements in microcosm dental plaques. Antimicrob Agents Chemother 2001;45:2943-6.
234. Langridge G, Phan M, Turner D, Perkins T, Parts L, Haase J, Charles I, Maskell D, Peters S, Dougan G, Wain J, Parkhill J, Turner A. Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. Genome Res 2009;19:2308-16.
235. Quail M, Kozarewa I, Smith F, Scally A, Stephens P, Durbin R, Swerdlow H, Turner D. A large genome center's improvements to the Illumina sequencing system. Nat Methods 2008;5:1005-10.
236. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 2008;18:1851-8.
237. Bailey T, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 1994;2:28-36.
238. Bailey T, Stafford-Noble W. The MEME Suite, Version 4.5.0 http://meme.nbcr.net/meme4_5_0/intro.html. 2010.
239. JGI IMG, database. http://imgweb.jgi-psf.org/cgi-bin/w/main.cgi JGI Intergrated Microbial Genomes database. University of California; Version 3.1 May 2010.
240. Dudley MN, McLaughlin JC, Carrington G, Frick J, Nightingale CH, Quintiliani R. Oral bacitracin vs vancomycin therapy for *Clostridium difficile*-induced diarrhea. A randomized double-blind trial. Arch Intern Med 1986;146:1101-4.

241. Young GP, Ward PB, Bayley N, Gordon D, Higgins G, Trapani JA, McDonald MI, Labrooy J, Hecker R. Antibiotic-associated colitis due to *Clostridium difficile*: double-blind comparison of vancomycin with bacitracin. Gastroenterology 1985;89:1038-45.
242. Rocha EP. The organization of the bacterial genome. Annu Rev Genet 2008;42:211-33.
243. Cheung JK, Awad MM, McGowan S, Rood JI. Functional analysis of the VirSR phosphorelay from *Clostridium perfringens.* PLoS One 2009;4:e5849.
244. Rodés J. Textbook of hepatology: from basic science to clinical practice. 3rd ed. editors, Rodés J, Benhamou J-P, Blei A, Reichen J, Rizzetto M. Oxford: Blackwell Pub.; 2007.
245. Begley M, Gahan CG, Hill C. The interaction between bacteria and bile. FEMS Microbiol Rev 2005;29:625-51.
246. Scotti E, Gilardi F, Godio C, Gers E, Krneta J, Mitro N, De Fabiani E, Caruso D, Crestani M. Bile acids and their signaling pathways: eclectic regulators of diverse cellular functions. Cell Mol Life Sci 2007;64:2477-91.
247. Martini F. Fundamentals of anatomy & physiology. 6th ed., International ed. San Francisco; Benjamin Cummings; 2004.
248. Floch MH. Bile salts, intestinal microflora and enterohepatic circulation. Dig Liver Dis 2002;34 Suppl 2:S54-7.
249. Sleisenger MH, Fordtran JS, Feldman M, Friedman LS, Brandt LJ. Sleisenger & Fordtran's gastrointestinal and liver disease: pathophysiology, diagnosis, management. 8th ed. edited by Feldman M, Friedman LS, Lawrence J. Brandt. ed. Philadelphia, Pa.: Saunders; 2006.
250. Ridlon JM, Kang DJ, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. J Lipid Res 2006;47:241-59.
251. Monte MJ, Marin JJ, Antelo A, Vazquez-Tato J. Bile acids: chemistry, physiology, and pathophysiology. World J Gastroenterol 2009;15:804-16.
252. Zakim D, Boyer TD. Hepatology : a textbook of liver disease. 4th ed. Philadelphia; London: Saunders; 2003.
253. Midtvedt T. Microbial bile acid transformation. Am J Clin Nutr 1974;27:1341-7.
254. Hofmann AF, Sjovall J, Kurz G, Radominska A, Schteingart CD, Tint GS, Vlahcevic ZR, Setchell KD. A proposed nomenclature for bile acids. J Lipid Res 1992;33:599-604.
255. McGarr SE, Ridlon JM, Hylemon PB. Diet, anaerobic bacterial metabolism, and colon cancer: a review of the literature. J Clin Gastroenterol 2005;39:98-109.
256. Duncan HE, Edberg SC. Host-microbe interaction in the gastrointestinal tract. Crit Rev Microbiol 1995;21:85-100.
257. Fujisawa T, Mori M. Influence of bile salts on beta-glucuronidase activity of intestinal bacteria. Lett Appl Microbiol 1996;22:271-4.
258. Noh DO, Gilliland SE. Influence of bile on cellular integrity and beta-galactosidase activity of *Lactobacillus acidophilus.* J Dairy Sci 1993;76:1253-9.

259. Lowe PJ, Coleman R. Membrane fluidity and bile salt damage. Biochim Biophys Acta 1981;640:55-65.
260. Coleman R, Lowe PJ, Billington D. Membrane lipid composition and susceptibility to bile salt damage. Biochim Biophys Acta 1980;599:294-300.
261. Nakayama F. Composition of gallstone and bile: species difference. J Lab Clin Med 1969;73:623-30.
262. King T, Ferenci T, Szabo EA. The effect of growth atmosphere on the ability of *Listeria monocytogenes* to survive exposure to acid, proteolytic enzymes and bile salts. Int J Food Microbiol 2003;84:133-43.
263. Merritt ME, Donaldson JR. Effect of bile salts on the DNA and membrane integrity of enteric bacteria. J Med Microbiol 2009;58:1533-41.
264. Prieto AI, Ramos-Morales F, Casadesus J. Bile-induced DNA damage in *Salmonella enterica.* Genetics 2004;168:1787-94.
265. Inagaki T, Moschetta A, Lee YK, Peng L, Zhao G, Downes M, Yu RT, Shelton JM, Richardson JA, Repa JJ, Mangelsdorf DJ, Kliewer SA. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. Proc Natl Acad Sci U S A 2006;103:3920-5.
266. Zakim D. Biochemical Pharmacology and Toxicology Vol. 1: Methodological aspects of drug metabolizing enzymes. New York ; Chichester: Wiley; 1985.
267. Wilson MA. Microbial inhabitants of humans : their ecology and role in health and disease. Cambridge University Press; 2005.
268. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A 1998;95:6578-83.
269. Madigan MT, Martinko JM, Parker J, Brock TDBom. Brock biology of microorganisms. 9th ed. Prentice Hall International; 2000.
270. Begley M, Gahan CG, Hill C. Bile stress response in *Listeria monocytogenes* LO28: adaptation, cross-protection, and identification of genetic loci involved in bile resistance. Appl Environ Microbiol 2002;68:6005-12.
271. Flahaut S, Frere J, Boutibonnes P, Auffray Y. Comparison of the bile salts and sodium dodecyl sulfate stress responses in *Enterococcus faecalis.* Appl Environ Microbiol 1996;62:2416-20.
272. Rince A, Le Breton Y, Verneuil N, Giard JC, Hartke A, Auffray Y. Physiological and molecular aspects of bile salt response in *Enterococcus faecalis.* Int J Food Microbiol 2003;88:207-13.
273. Finlay BB, Falkow S. Common themes in microbial pathogenicity revisited. Microbiol Mol Biol Rev 1997;61:136-69.
274. Prouty AM, Schwesinger WH, Gunn JS. Biofilm formation and interaction with the surfaces of gallstones by *Salmonella* spp. Infect Immun 2002;70:2640-9.
275. Ma D, Cook DN, Hearst JE, Nikaido H. Efflux pumps and drug resistance in gram-negative bacteria. Trends Microbiol 1994;2:489-93.
276. Li XZ, Nikaido H. Efflux-mediated drug resistance in bacteria: an update. Drugs 2009;69:1555-623.
277. Cabeen MT, Jacobs-Wagner C. Bacterial cell shape. Nat Rev Microbiol 2005;3:601-10.

278. Lin J, Huang S, Zhang Q. Outer membrane proteins: key players for bacterial adaptation in host niches. Microbes Infect 2002;4:325-31.
279. Thanassi DG, Cheng LW, Nikaido H. Active efflux of bile salts by *Escherichia coli.* J Bacteriol 1997;179:2512-8.
280. Provenzano D, Klose KE. Altered expression of the ToxR-regulated porins OmpU and OmpT diminishes *Vibrio cholerae* bile resistance, virulence factor expression, and intestinal colonization. Proc Natl Acad Sci U S A 2000;97:10220-4.
281. Paulsen IT, Park JH, Choi PS, Saier MH, Jr. A family of gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from gram-negative bacteria. FEMS Microbiol Lett 1997;156:1-8.
282. Bina JE, Mekalanos JJ. *Vibrio cholerae tolC* is required for bile resistance and colonization. Infect Immun 2001;69:4681-5.
283. Whitehead K, Versalovic J, Roos S, Britton RA. Genomic and genetic characterization of the bile stress response of probiotic *Lactobacillus reuteri* ATCC 55730. Appl Environ Microbiol 2008;74:1812-9.
284. Dunne C, O'Mahony L, Murphy L, Thornton G, Morrissey D, O'Halloran S, Feeney M, Flynn S, Fitzgerald G, Daly C, Kiely B, O'Sullivan GC, Shanahan F, Collins JK. In vitro selection criteria for probiotic bacteria of human origin: correlation with *in vivo* findings. Am J Clin Nutr 2001;73:386S-92S.
285. Fuller R. Probiotics in man and animals. J Appl Bacteriol 1989;66:365-78.
286. Gopal-Srivastava R, Hylemon PB. Purification and characterization of bile salt hydrolase from *Clostridium perfringens.* J Lipid Res 1988;29:1079-85.
287. Dussurget O, Pizarro-Cerda J, Cossart P. Molecular determinants of *Listeria monocytogenes* virulence. Annu Rev Microbiol 2004;58:587-610.
288. Pace JL, Chai TJ, Rossi HA, Jiang X. Effect of bile on *Vibrio parahaemolyticus.* Appl Environ Microbiol 1997;63:2372-7.
289. Osawa R, Arakawa E, Okitsu T, Yamai S, Watanabe H. Levels of thermostable direct hemolysin produced by *Vibrio parahaemolyticus* O3:K6 and other serovars grown anaerobically with the presence of a bile acid. Curr Microbiol 2002;44:302-5.
290. Krukonis ES, DiRita VJ. From motility to virulence: Sensing and responding to environmental signals in *Vibrio cholerae.* Curr Opin Microbiol 2003;6:186-90.
291. Olier M, Rousseaux S, Piveteau P, Lemaitre JP, Rousset A, Guzzo J. Screening of glutamate decarboxylase activity and bile salt resistance of human asymptomatic carriage, clinical, food, and environmental isolates of *Listeria monocytogenes.* Int J Food Microbiol 2004;93:87-99.
292. Sleator RD, Wemekamp-Kamphuis HH, Gahan CG, Abee T, Hill C. A PrfA-regulated bile exclusion system (BilE) is a novel virulence factor in *Listeria monocytogenes.* Mol Microbiol 2005;55:1183-95.
293. Gilmore MS. The enterococci : pathogenesis, molecular biology, and antibiotic resistance. Washington, D.C.: ASM Press; Oxford: Blackwell; 2002.

294. Flores C, Maguilnik I, Hadlich E, Goldani LZ. Microbiology of choledochal bile in patients with choledocholithiasis admitted to a tertiary hospital. J Gastroenterol Hepatol 2003;18:333-6.
295. Waar K, van der Mei HC, Harmsen HJ, Degener JE, Busscher HJ. Adhesion to bile drain materials and physicochemical surface properties of *Enterococcus faecalis* strains grown in the presence of bile. Appl Environ Microbiol 2002;68:3855-8.
296. Aktories K, Wilkins TD. *Clostridium difficile*. Berlin ; London: Springer; 2000.
297. Setlow P. Spore germination. Curr Opin Microbiol 2003;6:550-6.
298. Bliss DZ, Johnson S, Clabots CR, Savik K, Gerding DN. Comparison of cycloserine-cefoxitin-fructose agar (CCFA) and taurocholate-CCFA for recovery of *Clostridium difficile* during surveillance of hospitalized patients. Diagn Microbiol Infect Dis 1997;29:1-4.
299. Weese JS, Staempfli HR, Prescott JF. Isolation of environmental *Clostridium difficile* from a veterinary teaching hospital. J Vet Diagn Invest 2000;12:449-52.
300. Samuel P, Holtzman CM, Meilman E, Sekowski I. Effect of neomycin and other antibiotics on serum cholesterol levels and on 7alpha-dehydroxylation of bile acids by the fecal bacterial flora in man. Circ Res 1973;33:393-402.
301. Solheim M, Aakra A, Vebo H, Snipen L, Nes IF. Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. Appl Environ Microbiol 2007;73:5767-74.
302. Kristoffersen SM, Ravnum S, Tourasse NJ, Okstad OA, Kolsto AB, Davies W. Low concentrations of bile salts induce stress responses and reduce motility in *Bacillus cereus* ATCC 14579. J Bacteriol 2007;189:5302-13.
303. Cerda-Maira FA, Ringelberg CS, Taylor RK. The bile response repressor BreR regulates expression of the *Vibrio cholerae* breAB efflux system operon. J Bacteriol 2008;190:7441-52.
304. Pfeiler EA, Azcarate-Peril MA, Klaenhammer TR. Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus.* J Bacteriol 2007;189:4624-34.
305. Prouty AM, Brodsky IE, Manos J, Belas R, Falkow S, Gunn JS. Transcriptional regulation of *Salmonella enterica* serovar *Typhimurium* genes by bile. FEMS Immunol Med Microbiol 2004;41:177-85.
306. Malik-Kale P, Parker CT, Konkel ME. Culture of *Campylobacter jejuni* with sodium deoxycholate induces virulence gene expression. J Bacteriol 2008;190:2286-97.
307. Ye RW, Wang T, Bedzyk L, Croker KM. Applications of DNA microarrays in microbial systems. J Microbiol Methods 2001;47:257-72.
308. Bryant PA, Venter D, Robins-Browne R, Curtis N. Chips with everything: DNA microarrays in infectious diseases. Lancet Infect Dis 2004;4:100-11.
309. Warburg O, Christian W. Isolation and crystallization of enolase. Biochem Z 1942;310:384-421.

310. Bug@s-base. http://bugs.sgul.ac.uk/bugsbase/tabs/array.php?design_id=20&action=designs&l=a.
311. Helmann JD, Wu MF, Kobel PA, Gamo FJ, Wilson M, Morshedi MM, Navre M, Paddon C. Global transcriptional response of *Bacillus subtilis* to heat shock. J Bacteriol 2001;183:7318-28.
312. Cotter P, Hill C. Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiol Mol Biol Rev 2003;67:429-53, table of contents.
313. Petersohn A, Brigulla M, Haas S, Hoheisel JD, Volker U, Hecker M. Global analysis of the general stress response of *Bacillus subtilis.* J Bacteriol 2001;183:5617-31.
314. Prouty AM, Brodsky IE, Falkow S, Gunn JS. Bile-salt-mediated induction of antimicrobial and bile resistance in *Salmonella typhimurium.* Microbiology 2004;150:775-83.
315. Feldman M, Friedman LS, Sleisenger MH. Sleisenger & Fordtran's gastrointestinal and liver disease: pathophysiology, diagnosis, management. 7th ed. edited by Feldman M, Friedman LS, Sleisenger MH. ed. Philadelphia ; London: Saunders; 2002.
316. Sorg JA, Sonenshein AL. Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid. J Bacteriol 2010.
317. Hylemon PB, Zhou H, Pandak WM, Ren S, Gil G, Dent P. Bile acids as regulatory molecules. J Lipid Res 2009;50:1509-20.
318. Binder HJ, Filburn B, Floch M. Bile acid inhibition of intestinal anaerobic organisms. Am J Clin Nutr 1975;28:119-25.
319. Begley M, Sleator RD, Gahan CG, Hill C. Contribution of three bile-associated loci, bsh, pva, and btlB, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes.* Infect Immun 2005;73:894-904.
320. Leverrier P, Dimova D, Pichereau V, Auffray Y, Boyaval P, Jan G. Susceptibility and adaptive response to bile salts in *Propionibacterium freudenreichii*: physiological and proteomic analysis. Appl Environ Microbiol 2003;69:3809-18.
321. Flahaut S, Hartke A, Giard JC, Benachour A, Boutibonnes P, Auffray Y. Relationship between stress response toward bile salts, acid and heat treatment in *Enterococcus faecalis.* FEMS Microbiol Lett 1996;138:49-54.
322. Schmidt G, Zink R. Basic features of the stress response in three species of bifidobacteria: *B. longum*, *B. adolescentis*, and *B. breve.* Int J Food Microbiol 2000;55:41-5.
323. le Maire M, Champeil P, Moller JV. Interaction of membrane proteins and lipids with solubilizing detergents. Biochim Biophys Acta 2000;1508:86-111.
324. Begley M, Hill C, Gahan CG. Identification and disruption of *btlA*, a locus involved in bile tolerance and general stress resistance in *Listeria monocytogenes.* FEMS Microbiol Lett 2003;218:31-8.
325. Leyer GJ, Johnson EA. Acid adaptation induces cross-protection against environmental stresses in *Salmonella typhimurium.* Appl Environ Microbiol 1993;59:1842-7.

326. Lou Y, Yousef AE. Adaptation to sublethal environmental stresses protects *Listeria monocytogenes* against lethal preservation factors. Appl Environ Microbiol 1997;63:1252-5.
327. Krupp G. Stringent RNA quality control using the Agilent 2100 bioanalyzer - Application note. Agilent Technologies; 2005.
328. Hecker M, Schumann W, Volker U. Heat-shock and general stress response in *Bacillus subtilis.* Mol Microbiol 1996;19:417-28.
329. van Schaik W, Abee T. The role of *sigmaB* in the stress response of Gram-positive bacteria -- targets for food preservation and safety. Curr Opin Biotechnol 2005;16:218-24.
330. de Vries YP, Hornstra LM, Atmadja RD, Schaik W, de Vos WM, Abee T. Deletion of *sigB* in *Bacillus cereus* affects spore properties. FEMS Microbiol Lett 2005;252:169-73.
331. Emerson JE, Stabler RA, Wren BW, Fairweather NF. Microarray analysis of the transcriptional responses of *Clostridium difficile* to environmental and antibiotic stress. J Med Microbiol 2008;57:757-64.
332. Wiegert T, Homuth G, Versteeg S, Schumann W. Alkaline shock induces the *Bacillus subtilis sigma*(*W*) regulon. Mol Microbiol 2001;41:59-71.
333. Hennequin C, Collignon A, Karjalainen T. Analysis of expression of GroEL (Hsp60) of *Clostridium difficile* in response to stress. Microb Pathog 2001;31:255-60.
334. Janvilisri T, Scaria J, Thompson AD, Nicholson A, Limbago BM, Arroyo LG, Songer JG, Grohn YT, Chang YF. Microarray identification of *Clostridium difficile* core components and divergent regions associated with host origin. J Bacteriol 2009;191:3881-91.
335. Prouty AM, Gunn JS. *Salmonella enterica* serovar *typhimurium* invasion is repressed in the presence of bile. Infect Immun 2000;68:6763-9.
336. Wells CL, Jechorek RP, Erlandsen SL. Inhibitory effect of bile on bacterial invasion of enterocytes: possible mechanism for increased translocation associated with obstructive jaundice. Crit Care Med 1995;23:301-7.
337. Mullany P, Wilks M, Lamb I, Clayton C, Wren B, Tabaqchali S. Genetic analysis of a tetracycline resistance element from *Clostridium difficile* and its conjugal transfer to and from *Bacillus subtilis.* J Gen Microbiol 1990;136:1343-9.
338. Rubens CE, Heggen LM. Tn*916* delta E: a Tn*916* transposon derivative expressing erythromycin resistance. Plasmid 1988;20:137-42.
339. Shimizu T, Ba-Thein W, Tamaki M, Hayashi H. The virR gene, a member of a class of two-component response regulators, regulates the production of perfringolysin O, collagenase, and hemagglutinin in *Clostridium perfringens.* J Bacteriol 1994;176:1616-23.
340. Stein L. Genome annotation: from sequence to biology. Nat Rev Genet 2001;2:493-503.
341. Stein LD. Human genome: end of the beginning. Nature 2004;431:915-6.
342. Burrack LS, Higgins DE. Genomic approaches to understanding bacterial virulence. Curr Opin Microbiol 2007;10:4-9.

343. Brusic V. The growth of bioinformatics. Brief Bioinform 2007;8:69-70.
344. Ouzounis CA, Valencia A. Early bioinformatics: the birth of a discipline--a personal view. Bioinformatics 2003;19:2176-90.
345. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Biotechnology (N Y) 1992;10:190-5.
346. Davis TO, Henderson I, Brehm JK, Minton NP. Development of a transformation and gene reporter system for group II, non-proteolytic *Clostridium botulinum* type B strains. J Mol Microbiol Biotechnol 2000;2:59-69.
347. Herbert M, O'Keeffe TA, Purdy D, Elmore M, Minton NP. Gene transfer into *Clostridium difficile* CD630 and characterisation of its methylase genes. FEMS Microbiol Lett 2003;229:103-10.
348. Liyanage H, Kashket S, Young M, Kashket ER. *Clostridium beijerinckii* and *Clostridium difficile* detoxify methylglyoxal by a novel mechanism involving glycerol dehydrogenase. Appl Environ Microbiol 2001;67:2004-10.
349. Ba-Thein W, Lyristis M, Ohtani K, Nisbet IT, Hayashi H, Rood JI, Shimizu T. The *virR*/*virS* locus regulates the transcription of genes encoding extracellular toxin production in *Clostridium perfringens.* J Bacteriol 1996;178:2514-20.
350. Rood JI. Virulence genes of *Clostridium perfringens.* Annu Rev Microbiol 1998;52:333-60.
351. Matsuura M, Saldanha R, Ma H, Wank H, Yang J, Mohr G, Cavanagh S, Dunny GM, Belfort M, Lambowitz AM. A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Genes Dev 1997;11:2910-24.
352. Mohr G, Smith D, Belfort M, Lambowitz AM. Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. Genes Dev 2000;14:559-73.
353. Perutka J, Wang W, Goerlitz D, Lambowitz AM. Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. J Mol Biol 2004;336:421-39.
354. Karberg M, Guo H, Zhong J, Coon R, Perutka J, Lambowitz AM. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. Nat Biotechnol 2001;19:1162-7.
355. Zhong J, Karberg M, Lambowitz AM. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. Nucleic Acids Res 2003;31:1656-64.
356. Chen Y, McClane BA, Fisher DJ, Rood JI, Gupta P. Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. Appl Environ Microbiol 2005;71:7542-7.
357. Emerson JE, Reynolds CB, Fagan RP, Shaw HA, Goulding D, Fairweather NF. A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. Mol Microbiol 2009;74:541-56.

358. Salama NR, Manoil C. Seeking completeness in bacterial mutant hunts. Curr Opin Microbiol 2006;9:307-11.
359. Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC. Global transposon mutagenesis and a minimal Mycoplasma genome. Science 1999;286:2165-9.
360. Hamer L, DeZwaan TM, Montenegro-Chamorro MV, Frank SA, Hamer JE. Recent advances in large-scale transposon mutagenesis. Curr Opin Chem Biol 2001;5:67-73.
361. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson MV, Manoil C. Comprehensive transposon mutant library of *Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A 2003;100:14339-44.
362. Liberati NT, Urbach JM, Miyata S, Lee DG, Drenkard E, Wu G, Villanueva J, Wei T, Ausubel FM. An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. Proc Natl Acad Sci U S A 2006;103:2833-8.
363. Rengarajan J, Bloom BR, Rubin EJ. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. Proc Natl Acad Sci U S A 2005;102:8327-32.
364. Lamichhane G, Tyagi S, Bishai WR. Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. Infect Immun 2005;73:2533-40.
365. Chan K, Kim CC, Falkow S. Microarray-based detection of *Salmonella enterica* serovar *Typhimurium* transposon mutants that cannot survive in macrophages and mice. Infect Immun 2005;73:5438-49.
366. Mecsas J. Use of signature-tagged mutagenesis in pathogenesis studies. Curr Opin Microbiol 2002;5:33-7.
367. Bender J, Kleckner N. Genetic evidence that Tn*10* transposes by a nonreplicative mechanism. Cell 1986;45:801-15.
368. Foster TJ, Davis MA, Roberts DE, Takeshita K, Kleckner N. Genetic organization of transposon Tn*10.* Cell 1981;23:201-13.
369. Way JC, Kleckner N. Essential sites at transposon Tn*10* termini. Proc Natl Acad Sci U S A 1984;81:3452-6.
370. Mizuuchi K. Transpositional recombination: mechanistic insights from studies of *mu* and other elements. Annu Rev Biochem 1992;61:1011-51.
371. Roberts D, Hoopes BC, McClure WR, Kleckner N. IS*10* transposition is regulated by DNA adenine methylation. Cell 1985;43:117-30.
372. Morisato D, Kleckner N. Tn*10* transposition and circle formation in vitro. Cell 1987;51:101-11.
373. Haniford DB, Benjamin HW, Kleckner N. Kinetic and structural analysis of a cleaved donor intermediate and a strand transfer intermediate in Tn*10* transposition. Cell 1991;64:171-9.
374. Halling SM, Kleckner N. A symmetrical six-base-pair target site sequence determines Tn*10* insertion specificity. Cell 1982;28:155-63.
375. Kleckner N, Bender J, Gottesman S. Uses of transposons with emphasis on Tn*10.* Methods Enzymol 1991;204:139-80.

376. Bender J, Kleckner N. IS*10* transposase mutations that specifically alter target site recognition. EMBO J 1992;11:741-50.
377. Halling S, Simons R, Way J, Walsh R, Kleckner N. DNA sequence organization of IS*10*-right of Tn*10* and comparison with IS*10*-left. Proc Natl Acad Sci U S A 1982;79:2608-12.
378. Ciampi M. Rho-dependent terminators and transcription termination. Microbiology 2006;152:2515-28.
379. Way JC, Davis MA, Morisato D, Roberts DE, Kleckner N. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene 1984;32:369-79.
380. Dartois V, Djavakhishvili T, Hoch JA. Identification of a membrane protein involved in activation of the KinB pathway to sporulation in *Bacillus subtilis.* J Bacteriol 1996;178:1178-86.
381. Kamoun F, Fguira IB, Tounsi A, Abdelkefi-Mesrati L, Sanchis V, Lereclus D, Jaoua S. Generation of Mini-Tn*10* transposon insertion mutant library of *Bacillus thuringiensis* for the investigation of genes required for its bacteriocin production. FEMS Microbiol Lett 2009;294:141-9.
382. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning : a laboratory manual. 2nd ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory; 1989.
383. Hedges RW, Jacob AE. Transposition of ampicillin resistance from RP4 to other replicons. Mol Gen Genet 1974;132:31-40.
384. Wüst J, Hardegger U. Transferable resistance to clindamycin, erythromycin, and tetracycline in *Clostridium difficile*. Antimicrob Agents Chemother 1983;23:784-6.
385. The University of Cambridge. http://www.bio.cam.ac.uk/~pflgroup/DNA_Facility/About.html. 2006-2009.
386. EMBL-EBI. ClustalW2 - Multiple Sequences Alignment, (http://www.ebi.ac.uk/Tools/clustalw2/index.html). 2006-2009.
387. Sigma-Aldrich. http://www.sigmaaldrich.com/life-science/custom-oligos.html. 2006-2009.
388. Cartman ST, Minton NP. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile.* Appl Environ Microbiol 2010;76:1103-9.
389. Schmidt C, Löffler B, Ackermann G. Antimicrobial phenotypes and molecular basis in clinical strains of *Clostridium difficile*. Diagn Microbiol Infect Dis 2007;59:1-5.
390. Ballester S, Alonso JC, López P, Espinosa M. Comparative expression of the pC194 *cat* gene in *Streptococcus pneumoniae*, *Bacillus subtilis* and *Escherichia coli*. Gene 1990;86:71-9.
391. Fox ME, Lemmon MJ, Mauchline ML, Davis TO, Giaccia AJ, Minton NP, Brown JM. Anaerobic bacteria as a delivery system for cancer gene therapy: *in vitro* activation of 5-fluorocytosine by genetically engineered clostridia. Gene Ther 1996;3:173-8.
392. Heap JT, Pennington OJ, Cartman ST, Minton NP. A modular system for *Clostridium* shuttle plasmids. J Microbiol Methods 2009;78:79-85.
393. de Lorenzo V, Herrero M, Jakubzik U, Timmis K. Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and

chromosomal insertion of cloned DNA in gram-negative eubacteria. J Bacteriol 1990;172:6568-72.
394. Takiff H, Baker T, Copeland T, Chen S, Court D. Locating essential Escherichia coli genes by using mini-Tn*10* transposons: the *pdxJ* operon. J Bacteriol 1992;174:1544-53.
395. White BA. Methods in molecular biology. 67, PCR cloning protocols : from molecular cloning to genetic engineering. 1997.
396. Chen B-Y, Janes HW. PCR cloning protocols. 2nd ed. edited by Chen B-Y & James HW. ed. Totowa, N.J.: Humana Press; 2002.

9. **Appendix**

9.1. Statistical analysis of extended range bile assay

Kruskal-Wallis test	
P value	< 0.0001
Exact or approximate P value?	Gaussian Approximation
P value summary	****
Do the medians vary significantly (P < 0.05)	Yes
Number of groups	29
Kruskal-Wallis statistic	208.3

Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
0 vs 0.1%	51.82	No	ns
0 vs 0.3%	232.9	Yes	***
0 vs 0.5%	224.9	Yes	***
0 vs 0.8%	204.3	Yes	***
0 vs 1.0%	185.7	Yes	***
0 vs 1.3%	185.8	Yes	***
0 vs 1.5%	154.5	Yes	**
0 vs 2.0%	124.9	Yes	*
0 vs 2.5%	110.9	No	ns
0 vs 3.0%	94.71	No	ns
0 vs 3.5%	106.2	No	ns
0 vs 4.0%	83.04	No	ns
0 vs 4.5%	87.1	No	ns
0 vs 5.0%	75.21	No	ns
0 vs 5.5%	68.43	No	ns
0 vs 6.0%	73.04	No	ns
0 vs 6.5%	74.24	No	ns
0 vs 7.0%	81.6	No	ns
0 vs 7.5%	113.7	No	ns
0 vs 8.0%	165.3	Yes	***
0 vs 8.5%	150.3	Yes	**
0 vs 9.0%	122.9	Yes	*
0 vs 9.5%	146	Yes	**
0 vs 10.0%	180.8	Yes	***
0 vs 10.5%	180	Yes	***
0 vs 11.0%	223.7	Yes	***
0 vs 11.5%	203.1	Yes	***
0 vs 12.0%	240.9	Yes	***
0.1% vs 0.3%	181.1	Yes	**
0.1% vs 0.5%	173.1	Yes	**
0.1% vs 0.8%	152.5	Yes	*
0.1% vs 1.0%	133.9	No	ns

0.1% vs 1.3%	134	No	ns
0.1% vs 1.5%	102.7	No	ns
0.1% vs 2.0%	73.06	No	ns
0.1% vs 2.5%	59.06	No	ns
0.1% vs 3.0%	42.89	No	ns
0.1% vs 3.5%	54.39	No	ns
0.1% vs 4.0%	31.22	No	ns
0.1% vs 4.5%	35.28	No	ns
0.1% vs 5.0%	23.39	No	ns
0.1% vs 5.5%	16.61	No	ns
0.1% vs 6.0%	21.22	No	ns
0.1% vs 6.5%	22.42	No	ns
0.1% vs 7.0%	29.78	No	ns
0.1% vs 7.5%	61.83	No	ns
0.1% vs 8.0%	113.5	No	ns
0.1% vs 8.5%	98.44	No	ns
0.1% vs 9.0%	71.06	No	ns
0.1% vs 9.5%	94.22	No	ns
0.1% vs 10.0%	129	No	ns
0.1% vs 10.5%	128.2	No	ns
0.1% vs 11.0%	171.8	Yes	**
0.1% vs 11.5%	151.3	Yes	*
0.1% vs 12.0%	189.1	Yes	**
0.3% vs 0.5%	-8	No	ns
0.3% vs 0.8%	-28.61	No	ns
0.3% vs 1.0%	-47.22	No	ns
0.3% vs 1.3%	-47.11	No	ns
0.3% vs 1.5%	-78.39	No	ns
0.3% vs 2.0%	-108.1	No	ns
0.3% vs 2.5%	-122.1	No	ns
0.3% vs 3.0%	-138.2	No	ns
0.3% vs 3.5%	-126.7	No	ns
0.3% vs 4.0%	-149.9	No	ns
0.3% vs 4.5%	-145.8	No	ns
0.3% vs 5.0%	-157.7	Yes	*
0.3% vs 5.5%	-164.5	Yes	*
0.3% vs 6.0%	-159.9	Yes	*
0.3% vs 6.5%	-158.7	Yes	*
0.3% vs 7.0%	-151.3	Yes	*
0.3% vs 7.5%	-119.3	No	ns
0.3% vs 8.0%	-67.61	No	ns
0.3% vs 8.5%	-82.67	No	ns
0.3% vs 9.0%	-110.1	No	ns
0.3% vs 9.5%	-86.89	No	ns
0.3% vs 10.0%	-52.11	No	ns
0.3% vs 10.5%	-52.89	No	ns
0.3% vs 11.0%	-9.278	No	ns
0.3% vs 11.5%	-29.83	No	ns
0.3% vs 12.0%	8	No	ns
0.5% vs 0.8%	-20.61	No	ns
0.5% vs 1.0%	-39.22	No	ns
0.5% vs 1.3%	-39.11	No	ns
0.5% vs 1.5%	-70.39	No	ns
0.5% vs 2.0%	-100.1	No	ns
0.5% vs 2.5%	-114.1	No	ns

0.5% vs 3.0%	-130.2	No	ns
0.5% vs 3.5%	-118.7	No	ns
0.5% vs 4.0%	-141.9	No	ns
0.5% vs 4.5%	-137.8	No	ns
0.5% vs 5.0%	-149.7	No	ns
0.5% vs 5.5%	-156.5	Yes	*
0.5% vs 6.0%	-151.9	Yes	*
0.5% vs 6.5%	-150.7	No	ns
0.5% vs 7.0%	-143.3	No	ns
0.5% vs 7.5%	-111.3	No	ns
0.5% vs 8.0%	-59.61	No	ns
0.5% vs 8.5%	-74.67	No	ns
0.5% vs 9.0%	-102.1	No	ns
0.5% vs 9.5%	-78.89	No	ns
0.5% vs 10.0%	-44.11	No	ns
0.5% vs 10.5%	-44.89	No	ns
0.5% vs 11.0%	-1.278	No	ns
0.5% vs 11.5%	-21.83	No	ns
0.5% vs 12.0%	16	No	ns
0.8% vs 1.0%	-18.61	No	ns
0.8% vs 1.3%	-18.5	No	ns
0.8% vs 1.5%	-49.78	No	ns
0.8% vs 2.0%	-79.44	No	ns
0.8% vs 2.5%	-93.44	No	ns
0.8% vs 3.0%	-109.6	No	ns
0.8% vs 3.5%	-98.11	No	ns
0.8% vs 4.0%	-121.3	No	ns
0.8% vs 4.5%	-117.2	No	ns
0.8% vs 5.0%	-129.1	No	ns
0.8% vs 5.5%	-135.9	No	ns
0.8% vs 6.0%	-131.3	No	ns
0.8% vs 6.5%	-130.1	No	ns
0.8% vs 7.0%	-122.7	No	ns
0.8% vs 7.5%	-90.67	No	ns
0.8% vs 8.0%	-39	No	ns
0.8% vs 8.5%	-54.06	No	ns
0.8% vs 9.0%	-81.44	No	ns
0.8% vs 9.5%	-58.28	No	ns
0.8% vs 10.0%	-23.5	No	ns
0.8% vs 10.5%	-24.28	No	ns
0.8% vs 11.0%	19.33	No	ns
0.8% vs 11.5%	-1.222	No	ns
0.8% vs 12.0%	36.61	No	ns
1.0% vs 1.3%	0.1111	No	ns
1.0% vs 1.5%	-31.17	No	ns
1.0% vs 2.0%	-60.83	No	ns
1.0% vs 2.5%	-74.83	No	ns
1.0% vs 3.0%	-91	No	ns
1.0% vs 3.5%	-79.5	No	ns
1.0% vs 4.0%	-102.7	No	ns
1.0% vs 4.5%	-98.61	No	ns
1.0% vs 5.0%	-110.5	No	ns
1.0% vs 5.5%	-117.3	No	ns
1.0% vs 6.0%	-112.7	No	ns
1.0% vs 6.5%	-111.5	No	ns

1.0% vs 7.0%	-104.1	No	ns
1.0% vs 7.5%	-72.06	No	ns
1.0% vs 8.0%	-20.39	No	ns
1.0% vs 8.5%	-35.44	No	ns
1.0% vs 9.0%	-62.83	No	ns
1.0% vs 9.5%	-39.67	No	ns
1.0% vs 10.0%	-4.889	No	ns
1.0% vs 10.5%	-5.667	No	ns
1.0% vs 11.0%	37.94	No	ns
1.0% vs 11.5%	17.39	No	ns
1.0% vs 12.0%	55.22	No	ns
1.3% vs 1.5%	-31.28	No	ns
1.3% vs 2.0%	-60.94	No	ns
1.3% vs 2.5%	-74.94	No	ns
1.3% vs 3.0%	-91.11	No	ns
1.3% vs 3.5%	-79.61	No	ns
1.3% vs 4.0%	-102.8	No	ns
1.3% vs 4.5%	-98.72	No	ns
1.3% vs 5.0%	-110.6	No	ns
1.3% vs 5.5%	-117.4	No	ns
1.3% vs 6.0%	-112.8	No	ns
1.3% vs 6.5%	-111.6	No	ns
1.3% vs 7.0%	-104.2	No	ns
1.3% vs 7.5%	-72.17	No	ns
1.3% vs 8.0%	-20.5	No	ns
1.3% vs 8.5%	-35.56	No	ns
1.3% vs 9.0%	-62.94	No	ns
1.3% vs 9.5%	-39.78	No	ns
1.3% vs 10.0%	-5	No	ns
1.3% vs 10.5%	-5.778	No	ns
1.3% vs 11.0%	37.83	No	ns
1.3% vs 11.5%	17.28	No	ns
1.3% vs 12.0%	55.11	No	ns
1.5% vs 2.0%	-29.67	No	ns
1.5% vs 2.5%	-43.67	No	ns
1.5% vs 3.0%	-59.83	No	ns
1.5% vs 3.5%	-48.33	No	ns
1.5% vs 4.0%	-71.5	No	ns
1.5% vs 4.5%	-67.44	No	ns
1.5% vs 5.0%	-79.33	No	ns
1.5% vs 5.5%	-86.11	No	ns
1.5% vs 6.0%	-81.5	No	ns
1.5% vs 6.5%	-80.31	No	ns
1.5% vs 7.0%	-72.94	No	ns
1.5% vs 7.5%	-40.89	No	ns
1.5% vs 8.0%	10.78	No	ns
1.5% vs 8.5%	-4.278	No	ns
1.5% vs 9.0%	-31.67	No	ns
1.5% vs 9.5%	-8.5	No	ns
1.5% vs 10.0%	26.28	No	ns
1.5% vs 10.5%	25.5	No	ns
1.5% vs 11.0%	69.11	No	ns
1.5% vs 11.5%	48.56	No	ns
1.5% vs 12.0%	86.39	No	ns
2.0% vs 2.5%	-14	No	ns

2.0% vs 3.0%	-30.17	No	ns
2.0% vs 3.5%	-18.67	No	ns
2.0% vs 4.0%	-41.83	No	ns
2.0% vs 4.5%	-37.78	No	ns
2.0% vs 5.0%	-49.67	No	ns
2.0% vs 5.5%	-56.44	No	ns
2.0% vs 6.0%	-51.83	No	ns
2.0% vs 6.5%	-50.64	No	ns
2.0% vs 7.0%	-43.28	No	ns
2.0% vs 7.5%	-11.22	No	ns
2.0% vs 8.0%	40.44	No	ns
2.0% vs 8.5%	25.39	No	ns
2.0% vs 9.0%	-2	No	ns
2.0% vs 9.5%	21.17	No	ns
2.0% vs 10.0%	55.94	No	ns
2.0% vs 10.5%	55.17	No	ns
2.0% vs 11.0%	98.78	No	ns
2.0% vs 11.5%	78.22	No	ns
2.0% vs 12.0%	116.1	No	ns
2.5% vs 3.0%	-16.17	No	ns
2.5% vs 3.5%	-4.667	No	ns
2.5% vs 4.0%	-27.83	No	ns
2.5% vs 4.5%	-23.78	No	ns
2.5% vs 5.0%	-35.67	No	ns
2.5% vs 5.5%	-42.44	No	ns
2.5% vs 6.0%	-37.83	No	ns
2.5% vs 6.5%	-36.64	No	ns
2.5% vs 7.0%	-29.28	No	ns
2.5% vs 7.5%	2.778	No	ns
2.5% vs 8.0%	54.44	No	ns
2.5% vs 8.5%	39.39	No	ns
2.5% vs 9.0%	12	No	ns
2.5% vs 9.5%	35.17	No	ns
2.5% vs 10.0%	69.94	No	ns
2.5% vs 10.5%	69.17	No	ns
2.5% vs 11.0%	112.8	No	ns
2.5% vs 11.5%	92.22	No	ns
2.5% vs 12.0%	130.1	No	ns
3.0% vs 3.5%	11.5	No	ns
3.0% vs 4.0%	-11.67	No	ns
3.0% vs 4.5%	-7.611	No	ns
3.0% vs 5.0%	-19.5	No	ns
3.0% vs 5.5%	-26.28	No	ns
3.0% vs 6.0%	-21.67	No	ns
3.0% vs 6.5%	-20.47	No	ns
3.0% vs 7.0%	-13.11	No	ns
3.0% vs 7.5%	18.94	No	ns
3.0% vs 8.0%	70.61	No	ns
3.0% vs 8.5%	55.56	No	ns
3.0% vs 9.0%	28.17	No	ns
3.0% vs 9.5%	51.33	No	ns
3.0% vs 10.0%	86.11	No	ns
3.0% vs 10.5%	85.33	No	ns
3.0% vs 11.0%	128.9	No	ns
3.0% vs 11.5%	108.4	No	ns

3.0% vs 12.0%	146.2	No	ns
3.5% vs 4.0%	-23.17	No	ns
3.5% vs 4.5%	-19.11	No	ns
3.5% vs 5.0%	-31	No	ns
3.5% vs 5.5%	-37.78	No	ns
3.5% vs 6.0%	-33.17	No	ns
3.5% vs 6.5%	-31.97	No	ns
3.5% vs 7.0%	-24.61	No	ns
3.5% vs 7.5%	7.444	No	ns
3.5% vs 8.0%	59.11	No	ns
3.5% vs 8.5%	44.06	No	ns
3.5% vs 9.0%	16.67	No	ns
3.5% vs 9.5%	39.83	No	ns
3.5% vs 10.0%	74.61	No	ns
3.5% vs 10.5%	73.83	No	ns
3.5% vs 11.0%	117.4	No	ns
3.5% vs 11.5%	96.89	No	ns
3.5% vs 12.0%	134.7	No	ns
4.0% vs 4.5%	4.056	No	ns
4.0% vs 5.0%	-7.833	No	ns
4.0% vs 5.5%	-14.61	No	ns
4.0% vs 6.0%	-10	No	ns
4.0% vs 6.5%	-8.806	No	ns
4.0% vs 7.0%	-1.444	No	ns
4.0% vs 7.5%	30.61	No	ns
4.0% vs 8.0%	82.28	No	ns
4.0% vs 8.5%	67.22	No	ns
4.0% vs 9.0%	39.83	No	ns
4.0% vs 9.5%	63	No	ns
4.0% vs 10.0%	97.78	No	ns
4.0% vs 10.5%	97	No	ns
4.0% vs 11.0%	140.6	No	ns
4.0% vs 11.5%	120.1	No	ns
4.0% vs 12.0%	157.9	Yes	*
4.5% vs 5.0%	-11.89	No	ns
4.5% vs 5.5%	-18.67	No	ns
4.5% vs 6.0%	-14.06	No	ns
4.5% vs 6.5%	-12.86	No	ns
4.5% vs 7.0%	-5.5	No	ns
4.5% vs 7.5%	26.56	No	ns
4.5% vs 8.0%	78.22	No	ns
4.5% vs 8.5%	63.17	No	ns
4.5% vs 9.0%	35.78	No	ns
4.5% vs 9.5%	58.94	No	ns
4.5% vs 10.0%	93.72	No	ns
4.5% vs 10.5%	92.94	No	ns
4.5% vs 11.0%	136.6	No	ns
4.5% vs 11.5%	116	No	ns
4.5% vs 12.0%	153.8	Yes	*
5.0% vs 5.5%	-6.778	No	ns
5.0% vs 6.0%	-2.167	No	ns
5.0% vs 6.5%	-0.9722	No	ns
5.0% vs 7.0%	6.389	No	ns
5.0% vs 7.5%	38.44	No	ns
5.0% vs 8.0%	90.11	No	ns

5.0% vs 8.5%	75.06	No	ns
5.0% vs 9.0%	47.67	No	ns
5.0% vs 9.5%	70.83	No	ns
5.0% vs 10.0%	105.6	No	ns
5.0% vs 10.5%	104.8	No	ns
5.0% vs 11.0%	148.4	No	ns
5.0% vs 11.5%	127.9	No	ns
5.0% vs 12.0%	165.7	Yes	**
5.5% vs 6.0%	4.611	No	ns
5.5% vs 6.5%	5.806	No	ns
5.5% vs 7.0%	13.17	No	ns
5.5% vs 7.5%	45.22	No	ns
5.5% vs 8.0%	96.89	No	ns
5.5% vs 8.5%	81.83	No	ns
5.5% vs 9.0%	54.44	No	ns
5.5% vs 9.5%	77.61	No	ns
5.5% vs 10.0%	112.4	No	ns
5.5% vs 10.5%	111.6	No	ns
5.5% vs 11.0%	155.2	Yes	*
5.5% vs 11.5%	134.7	No	ns
5.5% vs 12.0%	172.5	Yes	**
6.0% vs 6.5%	1.194	No	ns
6.0% vs 7.0%	8.556	No	ns
6.0% vs 7.5%	40.61	No	ns
6.0% vs 8.0%	92.28	No	ns
6.0% vs 8.5%	77.22	No	ns
6.0% vs 9.0%	49.83	No	ns
6.0% vs 9.5%	73	No	ns
6.0% vs 10.0%	107.8	No	ns
6.0% vs 10.5%	107	No	ns
6.0% vs 11.0%	150.6	Yes	*
6.0% vs 11.5%	130.1	No	ns
6.0% vs 12.0%	167.9	Yes	**
6.5% vs 7.0%	7.361	No	ns
6.5% vs 7.5%	39.42	No	ns
6.5% vs 8.0%	91.08	No	ns
6.5% vs 8.5%	76.03	No	ns
6.5% vs 9.0%	48.64	No	ns
6.5% vs 9.5%	71.81	No	ns
6.5% vs 10.0%	106.6	No	ns
6.5% vs 10.5%	105.8	No	ns
6.5% vs 11.0%	149.4	No	ns
6.5% vs 11.5%	128.9	No	ns
6.5% vs 12.0%	166.7	Yes	*
7.0% vs 7.5%	32.06	No	ns
7.0% vs 8.0%	83.72	No	ns
7.0% vs 8.5%	68.67	No	ns
7.0% vs 9.0%	41.28	No	ns
7.0% vs 9.5%	64.44	No	ns
7.0% vs 10.0%	99.22	No	ns
7.0% vs 10.5%	98.44	No	ns
7.0% vs 11.0%	142.1	No	ns
7.0% vs 11.5%	121.5	No	ns
7.0% vs 12.0%	159.3	Yes	*
7.5% vs 8.0%	51.67	No	ns

7.5% vs 8.5%	36.61	No	ns
7.5% vs 9.0%	9.222	No	ns
7.5% vs 9.5%	32.39	No	ns
7.5% vs 10.0%	67.17	No	ns
7.5% vs 10.5%	66.39	No	ns
7.5% vs 11.0%	110	No	ns
7.5% vs 11.5%	89.44	No	ns
7.5% vs 12.0%	127.3	No	ns
8.0% vs 8.5%	-15.06	No	ns
8.0% vs 9.0%	-42.44	No	ns
8.0% vs 9.5%	-19.28	No	ns
8.0% vs 10.0%	15.5	No	ns
8.0% vs 10.5%	14.72	No	ns
8.0% vs 11.0%	58.33	No	ns
8.0% vs 11.5%	37.78	No	ns
8.0% vs 12.0%	75.61	No	ns
8.5% vs 9.0%	-27.39	No	ns
8.5% vs 9.5%	-4.222	No	ns
8.5% vs 10.0%	30.56	No	ns
8.5% vs 10.5%	29.78	No	ns
8.5% vs 11.0%	73.39	No	ns
8.5% vs 11.5%	52.83	No	ns
8.5% vs 12.0%	90.67	No	ns
9.0% vs 9.5%	23.17	No	ns
9.0% vs 10.0%	57.94	No	ns
9.0% vs 10.5%	57.17	No	ns
9.0% vs 11.0%	100.8	No	ns
9.0% vs 11.5%	80.22	No	ns
9.0% vs 12.0%	118.1	No	ns
9.5% vs 10.0%	34.78	No	ns
9.5% vs 10.5%	34	No	ns
9.5% vs 11.0%	77.61	No	ns
9.5% vs 11.5%	57.06	No	ns
9.5% vs 12.0%	94.89	No	ns
10.0% vs 10.5%	-0.7778	No	ns
10.0% vs 11.0%	42.83	No	ns
10.0% vs 11.5%	22.28	No	ns
10.0% vs 12.0%	60.11	No	ns
10.5% vs 11.0%	43.61	No	ns
10.5% vs 11.5%	23.06	No	ns
10.5% vs 12.0%	60.89	No	ns
11.0% vs 11.5%	-20.56	No	ns
11.0% vs 12.0%	17.28	No	ns
11.5% vs 12.0%	37.83	No	ns

Table Analyzed	Extended range 630deltaerm
Kruskal-Wallis test	
P value	< 0.0001
Exact or approximate P value?	Gaussian Approximation
P value summary	****
Do the medians vary significantly (P < 0.05)	Yes
Number of groups	29
Kruskal-Wallis statistic	242.7

Dunn's Multiple Comparison Test	Difference in rank sum	Significant? P < 0.05?	Summary
0 vs 0.1%	23.78	No	ns
0 vs 0.3%	208.9	Yes	***
0 vs 0.5%	155.8	Yes	***
0 vs 0.8%	135.3	Yes	**
0 vs 1.0%	124.7	Yes	*
0 vs 1.3%	81.54	No	ns
0 vs 1.5%	67.93	No	ns
0 vs 2.0%	38.6	No	ns
0 vs 2.5%	48.72	No	ns
0 vs 3.0%	53.93	No	ns
0 vs 3.5%	75.03	No	ns
0 vs 4.0%	84.43	No	ns
0 vs 4.5%	100.4	No	ns
0 vs 5.0%	83.99	No	ns
0 vs 5.5%	127.2	Yes	*
0 vs 6.0%	141	Yes	**
0 vs 6.5%	160.8	Yes	***
0 vs 7.0%	159.5	Yes	***
0 vs 7.5%	182	Yes	***
0 vs 8.0%	184.5	Yes	***
0 vs 8.5%	177.8	Yes	***
0 vs 9.0%	203.4	Yes	***
0 vs 9.5%	210.7	Yes	***
0 vs 10.0%	226	Yes	***
0 vs 10.5%	219.7	Yes	***
0 vs 11.0%	227.7	Yes	***
0 vs 11.5%	219.5	Yes	***
0 vs 12.0%	225.2	Yes	***
0.1% vs 0.3%	185.2	Yes	**
0.1% vs 0.5%	132	No	ns
0.1% vs 0.8%	111.5	No	ns
0.1% vs 1.0%	100.9	No	ns
0.1% vs 1.3%	57.76	No	ns
0.1% vs 1.5%	44.15	No	ns
0.1% vs 2.0%	14.82	No	ns

0.1% vs 2.5%	24.94	No	ns
0.1% vs 3.0%	30.15	No	ns
0.1% vs 3.5%	51.25	No	ns
0.1% vs 4.0%	60.65	No	ns
0.1% vs 4.5%	76.6	No	ns
0.1% vs 5.0%	60.21	No	ns
0.1% vs 5.5%	103.4	No	ns
0.1% vs 6.0%	117.3	No	ns
0.1% vs 6.5%	137	No	ns
0.1% vs 7.0%	135.7	No	ns
0.1% vs 7.5%	158.3	Yes	*
0.1% vs 8.0%	160.7	Yes	*
0.1% vs 8.5%	154	Yes	*
0.1% vs 9.0%	179.7	Yes	**
0.1% vs 9.5%	186.9	Yes	**
0.1% vs 10.0%	202.2	Yes	***
0.1% vs 10.5%	195.9	Yes	**
0.1% vs 11.0%	203.9	Yes	***
0.1% vs 11.5%	195.8	Yes	**
0.1% vs 12.0%	201.4	Yes	***
0.3% vs 0.5%	-53.17	No	ns
0.3% vs 0.8%	-73.61	No	ns
0.3% vs 1.0%	-84.28	No	ns
0.3% vs 1.3%	-127.4	No	ns
0.3% vs 1.5%	-141	No	ns
0.3% vs 2.0%	-170.3	Yes	**
0.3% vs 2.5%	-160.2	Yes	*
0.3% vs 3.0%	-155	Yes	*
0.3% vs 3.5%	-133.9	No	ns
0.3% vs 4.0%	-124.5	No	ns
0.3% vs 4.5%	-108.6	No	ns
0.3% vs 5.0%	-124.9	No	ns
0.3% vs 5.5%	-81.78	No	ns
0.3% vs 6.0%	-67.89	No	ns
0.3% vs 6.5%	-48.11	No	ns
0.3% vs 7.0%	-49.44	No	ns
0.3% vs 7.5%	-26.89	No	ns
0.3% vs 8.0%	-24.44	No	ns
0.3% vs 8.5%	-31.17	No	ns
0.3% vs 9.0%	-5.5	No	ns
0.3% vs 9.5%	1.722	No	ns
0.3% vs 10.0%	17.06	No	ns
0.3% vs 10.5%	10.78	No	ns
0.3% vs 11.0%	18.72	No	ns
0.3% vs 11.5%	10.61	No	ns
0.3% vs 12.0%	16.22	No	ns
0.5% vs 0.8%	-20.44	No	ns
0.5% vs 1.0%	-31.11	No	ns
0.5% vs 1.3%	-74.22	No	ns
0.5% vs 1.5%	-87.83	No	ns
0.5% vs 2.0%	-117.2	No	ns

0.5% vs 2.5%	-107	No	ns
0.5% vs 3.0%	-101.8	No	ns
0.5% vs 3.5%	-80.74	No	ns
0.5% vs 4.0%	-71.33	No	ns
0.5% vs 4.5%	-55.39	No	ns
0.5% vs 5.0%	-71.78	No	ns
0.5% vs 5.5%	-28.61	No	ns
0.5% vs 6.0%	-14.72	No	ns
0.5% vs 6.5%	5.056	No	ns
0.5% vs 7.0%	3.722	No	ns
0.5% vs 7.5%	26.28	No	ns
0.5% vs 8.0%	28.72	No	ns
0.5% vs 8.5%	22	No	ns
0.5% vs 9.0%	47.67	No	ns
0.5% vs 9.5%	54.89	No	ns
0.5% vs 10.0%	70.22	No	ns
0.5% vs 10.5%	63.94	No	ns
0.5% vs 11.0%	71.89	No	ns
0.5% vs 11.5%	63.78	No	ns
0.5% vs 12.0%	69.39	No	ns
0.8% vs 1.0%	-10.67	No	ns
0.8% vs 1.3%	-53.78	No	ns
0.8% vs 1.5%	-67.39	No	ns
0.8% vs 2.0%	-96.72	No	ns
0.8% vs 2.5%	-86.6	No	ns
0.8% vs 3.0%	-81.39	No	ns
0.8% vs 3.5%	-60.29	No	ns
0.8% vs 4.0%	-50.89	No	ns
0.8% vs 4.5%	-34.94	No	ns
0.8% vs 5.0%	-51.33	No	ns
0.8% vs 5.5%	-8.167	No	ns
0.8% vs 6.0%	5.722	No	ns
0.8% vs 6.5%	25.5	No	ns
0.8% vs 7.0%	24.17	No	ns
0.8% vs 7.5%	46.72	No	ns
0.8% vs 8.0%	49.17	No	ns
0.8% vs 8.5%	42.44	No	ns
0.8% vs 9.0%	68.11	No	ns
0.8% vs 9.5%	75.33	No	ns
0.8% vs 10.0%	90.67	No	ns
0.8% vs 10.5%	84.39	No	ns
0.8% vs 11.0%	92.33	No	ns
0.8% vs 11.5%	84.22	No	ns
0.8% vs 12.0%	89.83	No	ns
1.0% vs 1.3%	-43.11	No	ns
1.0% vs 1.5%	-56.72	No	ns
1.0% vs 2.0%	-86.06	No	ns
1.0% vs 2.5%	-75.94	No	ns
1.0% vs 3.0%	-70.72	No	ns
1.0% vs 3.5%	-49.63	No	ns
1.0% vs 4.0%	-40.22	No	ns

1.0% vs 4.5%	-24.28	No	ns
1.0% vs 5.0%	-40.67	No	ns
1.0% vs 5.5%	2.5	No	ns
1.0% vs 6.0%	16.39	No	ns
1.0% vs 6.5%	36.17	No	ns
1.0% vs 7.0%	34.83	No	ns
1.0% vs 7.5%	57.39	No	ns
1.0% vs 8.0%	59.83	No	ns
1.0% vs 8.5%	53.11	No	ns
1.0% vs 9.0%	78.78	No	ns
1.0% vs 9.5%	86	No	ns
1.0% vs 10.0%	101.3	No	ns
1.0% vs 10.5%	95.06	No	ns
1.0% vs 11.0%	103	No	ns
1.0% vs 11.5%	94.89	No	ns
1.0% vs 12.0%	100.5	No	ns
1.3% vs 1.5%	-13.61	No	ns
1.3% vs 2.0%	-42.94	No	ns
1.3% vs 2.5%	-32.83	No	ns
1.3% vs 3.0%	-27.61	No	ns
1.3% vs 3.5%	-6.514	No	ns
1.3% vs 4.0%	2.889	No	ns
1.3% vs 4.5%	18.83	No	ns
1.3% vs 5.0%	2.444	No	ns
1.3% vs 5.5%	45.61	No	ns
1.3% vs 6.0%	59.5	No	ns
1.3% vs 6.5%	79.28	No	ns
1.3% vs 7.0%	77.94	No	ns
1.3% vs 7.5%	100.5	No	ns
1.3% vs 8.0%	102.9	No	ns
1.3% vs 8.5%	96.22	No	ns
1.3% vs 9.0%	121.9	No	ns
1.3% vs 9.5%	129.1	No	ns
1.3% vs 10.0%	144.4	No	ns
1.3% vs 10.5%	138.2	No	ns
1.3% vs 11.0%	146.1	No	ns
1.3% vs 11.5%	138	No	ns
1.3% vs 12.0%	143.6	No	ns
1.5% vs 2.0%	-29.33	No	ns
1.5% vs 2.5%	-19.22	No	ns
1.5% vs 3.0%	-14	No	ns
1.5% vs 3.5%	7.097	No	ns
1.5% vs 4.0%	16.5	No	ns
1.5% vs 4.5%	32.44	No	ns
1.5% vs 5.0%	16.06	No	ns
1.5% vs 5.5%	59.22	No	ns
1.5% vs 6.0%	73.11	No	ns
1.5% vs 6.5%	92.89	No	ns
1.5% vs 7.0%	91.56	No	ns
1.5% vs 7.5%	114.1	No	ns
1.5% vs 8.0%	116.6	No	ns

1.5% vs 8.5%	109.8	No	ns
1.5% vs 9.0%	135.5	No	ns
1.5% vs 9.5%	142.7	No	ns
1.5% vs 10.0%	158.1	Yes	*
1.5% vs 10.5%	151.8	Yes	*
1.5% vs 11.0%	159.7	Yes	*
1.5% vs 11.5%	151.6	Yes	*
1.5% vs 12.0%	157.2	Yes	*
2.0% vs 2.5%	10.12	No	ns
2.0% vs 3.0%	15.33	No	ns
2.0% vs 3.5%	36.43	No	ns
2.0% vs 4.0%	45.83	No	ns
2.0% vs 4.5%	61.78	No	ns
2.0% vs 5.0%	45.39	No	ns
2.0% vs 5.5%	88.56	No	ns
2.0% vs 6.0%	102.4	No	ns
2.0% vs 6.5%	122.2	No	ns
2.0% vs 7.0%	120.9	No	ns
2.0% vs 7.5%	143.4	No	ns
2.0% vs 8.0%	145.9	No	ns
2.0% vs 8.5%	139.2	No	ns
2.0% vs 9.0%	164.8	Yes	**
2.0% vs 9.5%	172.1	Yes	**
2.0% vs 10.0%	187.4	Yes	**
2.0% vs 10.5%	181.1	Yes	**
2.0% vs 11.0%	189.1	Yes	**
2.0% vs 11.5%	180.9	Yes	**
2.0% vs 12.0%	186.6	Yes	**
2.5% vs 3.0%	5.215	No	ns
2.5% vs 3.5%	26.31	No	ns
2.5% vs 4.0%	35.72	No	ns
2.5% vs 4.5%	51.66	No	ns
2.5% vs 5.0%	35.27	No	ns
2.5% vs 5.5%	78.44	No	ns
2.5% vs 6.0%	92.33	No	ns
2.5% vs 6.5%	112.1	No	ns
2.5% vs 7.0%	110.8	No	ns
2.5% vs 7.5%	133.3	No	ns
2.5% vs 8.0%	135.8	No	ns
2.5% vs 8.5%	129	No	ns
2.5% vs 9.0%	154.7	Yes	*
2.5% vs 9.5%	161.9	Yes	*
2.5% vs 10.0%	177.3	Yes	**
2.5% vs 10.5%	171	Yes	**
2.5% vs 11.0%	178.9	Yes	**
2.5% vs 11.5%	170.8	Yes	**
2.5% vs 12.0%	176.4	Yes	**
3.0% vs 3.5%	21.1	No	ns
3.0% vs 4.0%	30.5	No	ns
3.0% vs 4.5%	46.44	No	ns
3.0% vs 5.0%	30.06	No	ns

3.0% vs 5.5%	73.22	No	ns
3.0% vs 6.0%	87.11	No	ns
3.0% vs 6.5%	106.9	No	ns
3.0% vs 7.0%	105.6	No	ns
3.0% vs 7.5%	128.1	No	ns
3.0% vs 8.0%	130.6	No	ns
3.0% vs 8.5%	123.8	No	ns
3.0% vs 9.0%	149.5	Yes	*
3.0% vs 9.5%	156.7	Yes	*
3.0% vs 10.0%	172.1	Yes	**
3.0% vs 10.5%	165.8	Yes	**
3.0% vs 11.0%	173.7	Yes	**
3.0% vs 11.5%	165.6	Yes	**
3.0% vs 12.0%	171.2	Yes	**
3.5% vs 4.0%	9.403	No	ns
3.5% vs 4.5%	25.35	No	ns
3.5% vs 5.0%	8.958	No	ns
3.5% vs 5.5%	52.13	No	ns
3.5% vs 6.0%	66.01	No	ns
3.5% vs 6.5%	85.79	No	ns
3.5% vs 7.0%	84.46	No	ns
3.5% vs 7.5%	107	No	ns
3.5% vs 8.0%	109.5	No	ns
3.5% vs 8.5%	102.7	No	ns
3.5% vs 9.0%	128.4	No	ns
3.5% vs 9.5%	135.6	No	ns
3.5% vs 10.0%	151	No	ns
3.5% vs 10.5%	144.7	No	ns
3.5% vs 11.0%	152.6	No	ns
3.5% vs 11.5%	144.5	No	ns
3.5% vs 12.0%	150.1	No	ns
4.0% vs 4.5%	15.94	No	ns
4.0% vs 5.0%	-0.4444	No	ns
4.0% vs 5.5%	42.72	No	ns
4.0% vs 6.0%	56.61	No	ns
4.0% vs 6.5%	76.39	No	ns
4.0% vs 7.0%	75.06	No	ns
4.0% vs 7.5%	97.61	No	ns
4.0% vs 8.0%	100.1	No	ns
4.0% vs 8.5%	93.33	No	ns
4.0% vs 9.0%	119	No	ns
4.0% vs 9.5%	126.2	No	ns
4.0% vs 10.0%	141.6	No	ns
4.0% vs 10.5%	135.3	No	ns
4.0% vs 11.0%	143.2	No	ns
4.0% vs 11.5%	135.1	No	ns
4.0% vs 12.0%	140.7	No	ns
4.5% vs 5.0%	-16.39	No	ns
4.5% vs 5.5%	26.78	No	ns
4.5% vs 6.0%	40.67	No	ns
4.5% vs 6.5%	60.44	No	ns

4.5% vs 7.0%	59.11	No	ns
4.5% vs 7.5%	81.67	No	ns
4.5% vs 8.0%	84.11	No	ns
4.5% vs 8.5%	77.39	No	ns
4.5% vs 9.0%	103.1	No	ns
4.5% vs 9.5%	110.3	No	ns
4.5% vs 10.0%	125.6	No	ns
4.5% vs 10.5%	119.3	No	ns
4.5% vs 11.0%	127.3	No	ns
4.5% vs 11.5%	119.2	No	ns
4.5% vs 12.0%	124.8	No	ns
5.0% vs 5.5%	43.17	No	ns
5.0% vs 6.0%	57.06	No	ns
5.0% vs 6.5%	76.83	No	ns
5.0% vs 7.0%	75.5	No	ns
5.0% vs 7.5%	98.06	No	ns
5.0% vs 8.0%	100.5	No	ns
5.0% vs 8.5%	93.78	No	ns
5.0% vs 9.0%	119.4	No	ns
5.0% vs 9.5%	126.7	No	ns
5.0% vs 10.0%	142	No	ns
5.0% vs 10.5%	135.7	No	ns
5.0% vs 11.0%	143.7	No	ns
5.0% vs 11.5%	135.6	No	ns
5.0% vs 12.0%	141.2	No	ns
5.5% vs 6.0%	13.89	No	ns
5.5% vs 6.5%	33.67	No	ns
5.5% vs 7.0%	32.33	No	ns
5.5% vs 7.5%	54.89	No	ns
5.5% vs 8.0%	57.33	No	ns
5.5% vs 8.5%	50.61	No	ns
5.5% vs 9.0%	76.28	No	ns
5.5% vs 9.5%	83.5	No	ns
5.5% vs 10.0%	98.83	No	ns
5.5% vs 10.5%	92.56	No	ns
5.5% vs 11.0%	100.5	No	ns
5.5% vs 11.5%	92.39	No	ns
5.5% vs 12.0%	98	No	ns
6.0% vs 6.5%	19.78	No	ns
6.0% vs 7.0%	18.44	No	ns
6.0% vs 7.5%	41	No	ns
6.0% vs 8.0%	43.44	No	ns
6.0% vs 8.5%	36.72	No	ns
6.0% vs 9.0%	62.39	No	ns
6.0% vs 9.5%	69.61	No	ns
6.0% vs 10.0%	84.94	No	ns
6.0% vs 10.5%	78.67	No	ns
6.0% vs 11.0%	86.61	No	ns
6.0% vs 11.5%	78.5	No	ns
6.0% vs 12.0%	84.11	No	ns
6.5% vs 7.0%	-1.333	No	ns

6.5% vs 7.5%	21.22	No	ns
6.5% vs 8.0%	23.67	No	ns
6.5% vs 8.5%	16.94	No	ns
6.5% vs 9.0%	42.61	No	ns
6.5% vs 9.5%	49.83	No	ns
6.5% vs 10.0%	65.17	No	ns
6.5% vs 10.5%	58.89	No	ns
6.5% vs 11.0%	66.83	No	ns
6.5% vs 11.5%	58.72	No	ns
6.5% vs 12.0%	64.33	No	ns
7.0% vs 7.5%	22.56	No	ns
7.0% vs 8.0%	25	No	ns
7.0% vs 8.5%	18.28	No	ns
7.0% vs 9.0%	43.94	No	ns
7.0% vs 9.5%	51.17	No	ns
7.0% vs 10.0%	66.5	No	ns
7.0% vs 10.5%	60.22	No	ns
7.0% vs 11.0%	68.17	No	ns
7.0% vs 11.5%	60.06	No	ns
7.0% vs 12.0%	65.67	No	ns
7.5% vs 8.0%	2.444	No	ns
7.5% vs 8.5%	-4.278	No	ns
7.5% vs 9.0%	21.39	No	ns
7.5% vs 9.5%	28.61	No	ns
7.5% vs 10.0%	43.94	No	ns
7.5% vs 10.5%	37.67	No	ns
7.5% vs 11.0%	45.61	No	ns
7.5% vs 11.5%	37.5	No	ns
7.5% vs 12.0%	43.11	No	ns
8.0% vs 8.5%	-6.722	No	ns
8.0% vs 9.0%	18.94	No	ns
8.0% vs 9.5%	26.17	No	ns
8.0% vs 10.0%	41.5	No	ns
8.0% vs 10.5%	35.22	No	ns
8.0% vs 11.0%	43.17	No	ns
8.0% vs 11.5%	35.06	No	ns
8.0% vs 12.0%	40.67	No	ns
8.5% vs 9.0%	25.67	No	ns
8.5% vs 9.5%	32.89	No	ns
8.5% vs 10.0%	48.22	No	ns
8.5% vs 10.5%	41.94	No	ns
8.5% vs 11.0%	49.89	No	ns
8.5% vs 11.5%	41.78	No	ns
8.5% vs 12.0%	47.39	No	ns
9.0% vs 9.5%	7.222	No	ns
9.0% vs 10.0%	22.56	No	ns
9.0% vs 10.5%	16.28	No	ns
9.0% vs 11.0%	24.22	No	ns
9.0% vs 11.5%	16.11	No	ns
9.0% vs 12.0%	21.72	No	ns
9.5% vs 10.0%	15.33	No	ns

9.5% vs 10.5%	9.056	No	ns
9.5% vs 11.0%	17	No	ns
9.5% vs 11.5%	8.889	No	ns
9.5% vs 12.0%	14.5	No	ns
10.0% vs 10.5%	-6.278	No	ns
10.0% vs 11.0%	1.667	No	ns
10.0% vs 11.5%	-6.444	No	ns
10.0% vs 12.0%	-0.8333	No	ns
10.5% vs 11.0%	7.944	No	ns
10.5% vs 11.5%	-0.1667	No	ns
10.5% vs 12.0%	5.444	No	ns
11.0% vs 11.5%	-8.111	No	ns
11.0% vs 12.0%	-2.5	No	ns
11.5% vs 12.0%	5.611	No	ns

9.2. Raw transcriptomics data

Test 0.1% p=0.05
Kat's Bile MAVI
Norm

	Control 0.10%				Test 0.10%				
Systematic ID	**Normalized**	**t-test P-value**	**Raw**	**Control**	**Fold change**	**t-test *P* value**	**Test**	**Control**	**Product**
Cd630-0067 (1K9)	1.1	0.468	5445	4820	1.9	0.004	5143	2848	DNA-directed RNA polymerase beta' chain
Cd630-0070 (1N9)	1.0	0.667	1716	1724	1.5	0.011	1129	757	translation elongation factor G
Cd630-0087 (1O11)	1.1	0.557	7738	7123	2.0	0.027	8265	4252	50S ribosomal protein L18
Cd630-0092 (1L12)	1.1	0.592	3800	3323	2.4	0.000	3901	1567	methionine aminopeptidase
Cd630-0096 (1P12)	1.1	0.562	96705	77653	1.5	0.048	89662	60882	30S ribosomal protein S11
Cd630-0126 (2F16)	1.0	0.991	791	810	0.7	0.011	538	729	stage III sporulation protein D
Cd630-0138 (2B18)	1.5	0.398	82	34	3.2	0.020	58	18	conserved hypothetical protein
Cd630-0140 (2D18)	0.8	0.572	372	397	0.6	0.029	128	207	putative peptidase
Cd630-0144 (2H18)	1.1	0.619	1659	1573	1.4	0.036	1257	927	peptide chain release factor 2
Cd630-0166 (2F21)	1.2	0.539	126	100	2.3	0.012	105	47	putative amidohydrolase/peptidase
Cd630-0180 (2D23)	0.9	0.722	104	99	0.6	0.038	57	94	hypothetical protein
Cd630-0181 (2E23)	0.9	0.575	825	848	0.3	0.019	259	727	putative membrane-associated metalloprotease
Cd630-0184 (2H23)	1.0	0.991	546	513	0.4	0.042	189	351	aspartate carbamoyltransferase catalytic chain
Cd630-0186 (2B24)	1.0	0.890	7181	7131	0.6	0.006	3987	6558	dihydroorotate dehydrogenase, catalytic subunit
Cd630-0188 (2D24)	0.9	0.403	1413	1506	0.5	0.012	656	1142	putative ketopantoate reductase
Cd630-0201 (4A2)	1.0	0.885	1121	1118	0.7	0.040	450	620	putative chromate transporter
Cd630-0202 (4B2)	1.0	0.675	886	855	0.8	0.049	335	439	conserved hypothetical protein
Cd630-0215 (4G3)	0.9	0.728	650	664	0.4	0.020	175	375	transposase
Cd630-0217 (4A4)	1.1	0.754	981	850	0.5	0.044	319	581	putative nitroreductase
Cd630-0219 (4C4)	0.8	0.463	916	1012	0.5	0.027	325	652	phosphoribosylaminoimidazole-succinocarboxamide synthase
Cd630-0230 (4F5)	1.0	0.638	4033	3950	0.6	0.050	1928	2841	putative flagellar biosynthesis protein
Cd630-0235 (4C6)	1.0	0.925	12046	11954	0.8	0.009	7540	9354	flagellar protein FliS
Cd630-0241 (4A7)	1.0	0.933	12746	12360	0.7	0.023	6565	9128	phosphoserine phosphatase
Cd630-0247 (4G7)	1.0	0.909	8943	9004	0.5	0.023	4709	7647	flagellar hook-basal body complex protein
Cd630-0251 (4C8)	1.0	0.713	49203	51295	0.6	0.015	33109	50395	flagellum-specific ATP synthase
Cd630-0253 (4E8)	1.0	0.937	15899	15870	0.5	0.049	7828	13067	putative flagellar hook-length control protein
Cd630-0254 (4F8)	1.0	0.661	22850	23297	0.6	0.048	13141	20850	putative basal-body rod modification protein
Cd630-0256 (4H8)	1.0	0.697	28011	28251	0.6	0.043	15376	24126	chemotaxis protein
Cd630-0295 (2G1)	1.0	0.874	66	64	3.8	0.025	76	21	putative iron-sulfur-binding protein

Cd630-0332 (2D6)	0.9	0.432	1523	1663	0.6	0.026	802	1308	putative exosporium glycoprotein
Cd630-0334 (2F6)	1.1	0.611	2007	1973	1.5	0.008	1934	1283	aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase; acetaldehyde dehydrogenase [acetylating]; pyruvate-formate-lyase deactivase
Cd630-0335 (2G6)	0.5	0.663	25	28	0.4	0.009	19	51	hypothetical protein
Cd630-0356 (2D9)	1.0	0.574	1156	1105	1.3	0.009	1026	779	excisionase
Cd630-0382 (2F12)	1.1	0.724	43	38	3.3	0.014	73	22	hypothetical protein
Cd630-0397 (2M2)	1.1	0.612	24574	21243	1.2	0.011	18251	14933	R-phenyllactate dehydratase medium subunit
Cd630-0406 (2N3)	1.0	0.775	1995	2152	1.3	0.041	1345	1046	putative tRNA (Uracil-5-)-methyltransferase
Cd630-0407 (2O3)	1.5	0.368	92	59	4.4	0.038	86	25	hypothetical protein
Cd630-0409 (2I4)	1.0	0.921	479	475	0.6	0.043	194	306	putative replication initiation protein
Cd630-0447 (2P8)	0.9	0.659	571	578	0.5	0.018	258	472	methylaspartate mutase
Cd630-0460 (2M10)	0.8	0.473	577	676	0.6	0.033	215	319	ABC transporter, permease protein
Cd630-0467 (2L11)	0.9	0.672	228	247	0.7	0.022	83	124	putative hydrolase
Cd630-0482 (1C1)	1.2	0.644	54	43	3.3	0.015	92	25	putative lantibiotic resistance two-component sensor kinase
Cd630-0491 (1D2)	1.1	0.742	44	41	4.8	0.046	154	31	PTS system, IIa component
Cd630-0494 (1G2)	0.9	0.843	46	41	2.5	0.048	41	16	PTS system, IId component
Cd630-0515 (1E5)	0.7	0.385	1542	2278	0.4	0.001	444	1056	D-alanyl-D-alanine carboxypeptidase
Cd630-0516 (1F5)	1.7	0.525	70	28	7.0	0.015	70	10	putative transcription antiterminator
Cd630-0524 (1F6)	1.0	0.868	2079	2046	0.7	0.037	1051	1401	putative peptidase
Cd630-0525 (1G6)	1.0	0.701	3785	3824	0.7	0.029	2490	3434	putative aminobenzoyl-glutamate transporter (pseudogene)
Cd630-0533 (1G7)	0.9	0.477	617	655	0.3	0.032	215	512	chemotaxis protein
Cd630-0535 (1A8)	0.9	0.684	892	939	0.2	0.034	251	886	chemotaxis protein
Cd630-0539 (1E8)	1.0	0.808	3599	3638	0.4	0.045	1702	3419	chemotaxis protein
Cd630-0541 (1G8)	1.0	0.858	1134	1098	0.3	0.028	289	681	chemotaxis protein methyltransferase
Cd630-0542 (1H8)	1.0	0.610	2109	2157	0.5	0.022	1072	1857	putative chemotaxis-related protein-glutamate methylesterase
Cd630-0588 (2N14)	0.6	0.433	307	403	0.4	0.042	144	310	hypothetical protein
Cd630-0592 (2J15)	1.3	0.514	102	63	3.6	0.004	137	38	putative membrane protein
Cd630-0638 (2N20)	1.0	0.954	45	41	3.9	0.040	67	20	hypothetical protein
Cd630-0644 (2L21)	0.9	0.839	158	169	0.7	0.018	77	116	two-component sensor histidine kinase
Cd630-0663g (3B1)	1.0	0.430	13200	12889	1.2	0.027	8309	7127	toxin A
Cd630-0666 (3F1)	1.0	0.655	8029	7853	1.2	0.029	5129	4436	putative lantibiotic ABC transporter, peremase protein
Cd630-0670 (3B2)	1.0	0.735	4531	4597	1.6	0.002	3285	2050	putative regulatory protein
Cd630-0707 (3G6)	1.0	0.990	2724	2794	0.7	0.019	1240	1701	putative signaling protein
Cd630-0716 (3H7)	0.9	0.434	8394	8927	0.7	0.016	5628	7676	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase
Cd630-0777 (1E15)	1.0	0.807	69	73	0.5	0.045	31	59	putative membrane protein
Cd630-0779 (1G15)	1.0	0.944	740	719	0.6	0.004	251	456	putative amidohydrolase
Cd630-0791 (1C17)	1.4	0.464	69	54	6.7	0.018	98	14	conserved hypothetical protein
Cd630-0799 (1C18)	1.1	0.784	40	35	8.1	0.015	194	15	putative CoA-transferase

Cd630-0802 (1F18)	1.0	0.872	622	591	0.6	0.012	156	283	putative CoA-transferase
Cd630-0809 (1E19)	1.0	0.528	1109	1087	0.7	0.034	367	524	conserved hypothetical protein
Cd630-0815 (1C20)	1.2	0.722	50	42	5.7	0.019	85	18	probable transporter
Cd630-0820 (1H20)	1.0	0.868	806	776	1.4	0.043	751	512	two-component response regulator
Cd630-0839 (1C23)	1.0	0.901	233	258	0.8	0.002	263	314	putative membrane-associated CAAX amino terminal protease
Cd630-0846 (1B24)	1.1	0.439	222	189	1.4	0.036	170	122	hypothetical protein
Cd630-0868 (3P2)	1.1	0.427	3828	3424	1.4	0.042	2268	1652	conserved hypothetical protein
Cd630-0875 (3O3)	1.0	0.389	15567	16470	0.9	0.008	9035	10436	ABC transporter, permease protein
Cd630-0891 (3O5)	0.9	0.576	4594	4973	0.8	0.045	2354	3087	putative agmatinase
Cd630-0907 (3O7)	1.0	0.977	110	107	1.5	0.045	97	65	putative phage regulatory protein
Cd630-0917 (3I9)	1.1	0.633	20	20	3.5	0.048	41	11	putative phage recombination protein
Cd630-0930 (3N10)	0.9	0.404	397	423	0.6	0.025	143	229	phage protein
Cd630-0939 (3O11)	1.1	0.416	3676	3453	1.4	0.017	3121	2260	hypothetical phage protein
Cd630-0963 (4O14)	1.1	0.667	261	192	1.6	0.030	185	107	phage protein
Cd630-0964 (4P14)	0.9	0.499	1002	1057	0.7	0.041	602	818	phage protein
Cd630-0967 (4K15)	0.9	0.551	813	875	0.6	0.012	436	666	hypothetical phage protein
Cd630-0974 (4J16)	2.2	0.429	147	74	4.4	0.020	203	39	hypothetical phage protein
Cd630-0991 (4K18)	1.0	0.669	550	568	0.7	0.049	212	304	3-isopropylmalate dehydratase small subunit
Cd630-1009 (4M20)	0.9	0.838	1426	1492	1.6	0.008	1140	728	GntR-family transcriptional regulator
Cd630-1017 (4M21)	1.3	0.539	249	139	1.8	0.047	132	76	putative ABC transporter, permease/ATP-binding protein
Cd630-1019 (4O21)	1.1	0.517	41	39	4.9	0.045	69	16	conserved hypothetical protein
Cd630-1036b (4J24)	1.1	0.533	379	314	0.6	0.047	177	317	putative cell surface N-acetylmuramoyl-L-alanine amidase
Cd630-1039 (4M24)	0.3	0.641	53	279	0.1	0.027	31	554	conserved hypothetical protein
Cd630-1054 (4D14)	0.8	0.434	376	505	0.4	0.037	235	505	butyryl-CoA dehydrogenase
Cd630-1064 (4F15)	0.9	0.673	518	579	0.8	0.040	416	523	LacI-family transcriptional regulator (catabolite control protein)
Cd630-1094 (4D19)	0.8	0.457	58	70	2.1	0.017	122	54	conjugative transposon protein
Cd630-1107 (4A21)	1.1	0.906	16	6	9.2	0.004	35	4	putative membrane protein
Cd630-1110 (4D21)	2.8	0.548	423	23	2.5	0.027	73	25	putative conjugal transfer protein
Cd630-1138 (4H24)	1.0	0.854	21421	20221	0.8	0.031	14316	17773	electron transport complex protein
Cd630-1147 (3A14)	1.0	0.980	9599	9565	0.8	0.017	4786	5944	putative membrane protein
Cd630-1168 (3F16)	0.9	0.507	552	586	0.8	0.002	365	482	putative membrane protein
Cd630-1175 (3E17)	1.0	0.418	25067	25787	1.2	0.016	15614	13459	acetate kinase
Cd630-1176 (3F17)	1.1	0.619	18563	17833	1.4	0.006	13079	9516	conserved hypothetical protein
Cd630-1185 (3G18)	1.1	0.572	2044	1836	1.6	0.045	1596	1026	putative signaling protein
Cd630-1191 (3E19)	1.0	0.762	6370	6278	1.5	0.004	3779	2523	putative fructose-1,6-bisphosphatase
Cd630-1211 (3A22)	0.9	0.488	2725	2849	0.8	0.004	1362	1747	putative acetyltransferase
Cd630-1219 (3A23)	0.9	0.530	6867	7126	0.6	0.018	3189	5049	conserved hypothetical protein

Cd630-1222 (3D23)	1.0	0.576	5129	5348	0.5	0.043	2639	4417	putative tyrosine recombinase
Cd630-1249 (3P14)	1.0	0.916	571	528	0.7	0.032	234	344	radical SAM-superfamily protein
Cd630-1253 (3L15)	0.8	0.399	1620	2028	0.5	0.042	783	1450	30S ribosomal protein S16
Cd630-1257 (3P15)	1.0	0.684	3417	3457	0.7	0.028	1662	2486	50S ribosomal protein L19
Cd630-1258 (3I16)	1.0	0.624	1628	1656	0.6	0.040	789	1229	putative GTPase
Cd630-1261 (3L16)	1.0	0.842	7150	6786	0.8	0.031	4455	5477	putative ribonucleotide-diphosphate reductase
Cd630-1271 (3N17)	0.9	0.523	5807	6353	1.4	0.034	3832	2720	conserved hypothetical protein
Cd630-1305 (3P21)	1.1	0.672	16633	14732	1.7	0.024	15287	9464	DNA polymerase III PolC-type
Cd630-1318 (3M23)	1.0	0.715	11338	10793	0.8	0.006	6604	8190	polyribonucleotide nucleotidyltransferase
Cd630-1319 (3N23)	1.0	0.794	338	329	0.5	0.035	176	301	putative polysaccharide deacetylase
Cd630-1322 (3I24)	1.0	0.963	456	389	0.7	0.039	247	370	aspartokinase
Cd630-1370 (4I6)	1.5	0.578	57	30	5.0	0.036	67	16	phage protein
Cd630-1376 (4O6)	0.8	0.639	119	122	0.7	0.039	66	93	putative membrane protein
Cd630-1390 (4M8)	1.2	0.760	32	22	1.4	0.033	33	24	putative membrane protein
Cd630-1404 (4K10)	1.0	0.563	16830	15969	1.5	0.006	15587	10338	putative oligopeptide transporter
Cd630-1412 (4K11)	0.9	0.622	492	519	0.5	0.001	181	350	sigma-54 dependent regulatory protein
Cd630-1417 (4P11)	0.9	0.518	1334	1521	0.4	0.026	300	703	putative ATP-binding protein
Cd630-1438 (5E2)	1.4	0.459	152	131	1.3	0.039	115	84	putative multiprotein complex assembly protein
Cd630-1449 (5H3)	1.1	0.862	871	893	0.6	0.017	373	596	putative GTP cyclohydrolase I
Cd630-1497 (5A10)	1.0	0.886	3536	3652	1.2	0.011	2453	2008	hypothetical protein
Cd630-1543 (5O3)	1.5	0.576	77	42	2.4	0.016	63	27	putative acyl carrier protein phosphodiesterase
Cd630-1550 (5N4)	1.0	0.969	3072	3178	0.6	0.028	1260	1831	imidazoleglycerol-phosphate dehydratase
Cd630-1576 (5P7)	1.1	0.550	3248	2765	0.5	0.011	1436	2524	putative arylesterase
Cd630-1578 (5J8)	0.9	0.410	1077	1146	0.7	0.034	620	877	putative membrane-associated ribonuclease
Cd630-1582 (5N8)	0.8	0.525	77	85	0.4	0.006	59	135	histidinol dehydrogenase
Cd630-1592 (5P9)	0.8	0.495	73	79	3.0	0.022	98	38	potassium-transporting ATPase B chain
Cd630-1607 (5O11)	0.8	0.482	58	62	3.1	0.019	85	29	ABC transporter, ATP-binding protein
Cd630-1610 (5J12)	1.1	0.814	220	121	4.2	0.027	136	35	putative membrane protein
Cd630-1611 (5K12)	1.0	0.885	101	98	2.1	0.026	76	34	putative exported protein
Cd630-1645 (5E16)	1.0	0.741	6091	5616	0.7	0.006	3515	4941	putative membrane protein
Cd630-1646 (5F16)	1.0	0.746	4347	4027	0.7	0.045	2113	2943	putative dipeptidase
Cd630-1657 (5A18)	1.0	0.701	1906	1877	1.3	0.049	1025	746	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein
Cd630-1661 (5E18)	1.0	0.804	7646	7512	1.8	0.012	5456	3156	transposase-like protein B
Cd630-1663 (5G18)	1.1	0.462	1327	1208	1.5	0.017	846	556	guanine deaminase
Cd630-1673 (5A20)	1.0	0.952	1935	1815	0.7	0.003	637	908	putative membrane protein
Cd630-1675 (5C20)	1.0	0.863	1897	1971	1.6	0.033	923	614	putative transcriptional regulator
Cd630-1677 (5E20)	0.7	0.512	1874	2063	0.4	0.006	438	1177	putative membrane protein
Cd630-1696 (5H22)	1.1	0.549	283	271	1.5	0.014	263	185	putative sodium:amino acid symporter

Cd630-1709 (5E24)	1.0	0.867	3891	3845	0.7	0.001	2210	3084	putative molybdopterin-guanine biosynthesis protein
Cd630-1710 (5F24)	1.0	0.471	2085	2052	0.7	0.023	1215	1637	hypothetical protein
Cd630-1760 (5I19)	1.0	0.911	1229	1220	0.5	0.004	236	438	putative iron-sulfur protein
Cd630-1767 (5P19)	1.0	0.887	52	48	3.5	0.034	88	29	glyceraldehyde-3-phosphate dehydrogenase 1
Cd630-1768 (5I20)	1.0	0.863	40183	42545	1.6	0.017	34810	22585	putative membrane protein
Cd630-1774 (5O20)	1.3	0.651	140	124	2.4	0.024	108	47	ABC transporter, substrate-binding protein
Cd630-1779 (5L21)	1.1	0.561	2600	2543	0.7	0.049	1499	1954	conserved hypothetical protein
Cd630-1802 (5K24)	1.0	0.769	461	484	0.7	0.039	288	403	putative hydrolase
Cd630-1803a (5L24)	1.3	0.493	92	55	3.6	0.017	86	28	putative cell surface protein
Cd630-1812 (1N13)	1.2	0.496	375	350	2.3	0.046	1421	447	putative two-somponent system response regulator (partial)
Cd630-1838 (1P16)	1.2	0.417	1158	988	0.8	0.041	726	864	shikimate kinase
Cd630-1846 (1P17)	0.7	0.648	108	96	2.5	0.011	134	50	conserved hypothetical protein
Cd630-1850 (1L18)	1.1	0.551	125	94	3.2	0.014	149	47	AraC-family transcriptional regulator
Cd630-1854 (1P18)	1.7	0.456	109	37	3.9	0.045	96	24	putative conjugative transposon membrane exported protein
Cd630-1880 (1J22)	1.1	0.755	93	83	2.3	0.018	97	42	hypothetical protein
Cd630-1885 (1O22)	1.0	0.948	42	42	3.1	0.014	63	20	putative acetyl xylan esterase
Cd630-1895 (1I24)	1.0	0.863	3871	3531	0.7	0.002	1954	2890	putative membrane protein
Cd630-1927 (6A4)	1.1	0.915	675	264	0.3	0.010	28	79	ABC transporter, ATP-binding protein (pseudogene)
Cd630-1931 (6E4)	0.9	0.698	5033	5823	1.4	0.014	2388	1689	SOS regulatory protein
Cd630-1936 (6B5)	1.0	0.878	3155	3086	1.3	0.036	2456	1816	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
Cd630-1937 (6C5)	0.9	0.604	11864	12489	0.8	0.020	7491	10047	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
Cd630-1943 (6A6)	1.0	0.532	1700	1623	1.6	0.010	1571	990	conserved hypothetical protein
Cd630-1946 (6D6)	1.0	0.937	555	510	0.5	0.026	232	400	AraC-family transcriptional regulator
Cd630-2003 (6N1)	1.0	0.981	1664	1656	1.2	0.030	1151	942	putative efflux pump
Cd630-2004 (6O1)	1.2	0.473	341	293	1.9	0.016	303	165	hypothetical protein
Cd630-2043 (6I7)	0.9	0.370	781	841	0.6	0.021	369	565	putative hydrolase
Cd630-2051 (6I8)	1.0	0.756	1792	1791	0.8	0.044	1003	1228	putative membrane protein
Cd630-2052 (6J8)	1.1	0.381	4894	4541	0.8	0.011	2959	3733	putative lipoprotein
Cd630-2065 (6O9)	1.1	0.524	2277	2164	1.4	0.013	1809	1227	putative regulatory protein
Cd630-2075 (6I11)	0.9	0.692	562	584	0.7	0.049	207	307	xanthine permease
Cd630-2079 (6M11)	1.1	0.680	333	268	0.5	0.036	95	169	xanthine dehydrogenase, molybdenum binding subunit
Cd630-2081 (6O11)	0.8	0.513	33	39	2.1	0.027	49	25	xanthine dehydrogenase iron-sulfur binding subunit
Cd630-2108 (6B15)	1.1	0.452	4160	3466	0.9	0.001	2319	2747	putative drug/sodium antiporter
Cd630-2120 (6F16)	1.1	0.547	13070	11286	0.8	0.044	7460	9364	putative 2-nitropropane dioxygenase
Cd630-2122 (6H16)	1.2	0.541	116	94	3.8	0.022	164	47	putative cations transporter
Cd630-2123 (6A17)	1.0	0.792	2329	2291	1.1	0.015	1521	1367	putative multiprotein-complex assembly protein
Cd630-2129 (6G17)	1.4	0.502	56	48	4.8	0.039	97	21	putative membrane-associated protease
Cd630-2144 (6F19)	1.0	0.870	295	280	0.5	0.047	70	154	putative membrane protein

Cd630-2172 (6B23)	1.0	0.568	5203	5300	0.8	0.011	2775	3561	probable amino-acid ABC transporter, ATP-binding protein
Cd630-2173 (6C23)	1.0	0.745	4239	3945	0.7	0.040	2232	3069	putative peptidase
Cd630-2174 (6D23)	1.1	0.812	1549	1430	0.7	0.042	891	1221	probable amino-acid ABC transporter, substrate-binding protein
Cd630-2179 (6A24)	1.0	0.948	2524	2436	0.5	0.006	985	1751	putative oxidoreductase, molybdopterin-binding subunit
Cd630-2181 (6C24)	1.0	0.911	7696	7272	0.6	0.017	3516	5743	putative aromatic compounds hydrolase
Cd630-2208 (8O3)	0.9	0.722	239	261	0.4	0.037	75	171	MarR-family transcriptional regulator
Cd630-2235 (8J7)	0.9	0.838	505	454	0.6	0.012	194	304	RpiR-family transcriptional regulator
Cd630-2252 (8K9)	1.0	0.756	1676	1630	1.6	0.039	1188	716	L-lysine 2,3-aminomutase
Cd630-2322 (6I18)	1.0	0.984	9611	9337	1.5	0.013	7433	5137	transketolase
Cd630-2335 (6N19)	1.1	0.693	4926	4569	1.5	0.002	4274	2842	inosine-5'-monophosphate dehydrogenase
Cd630-2363 (6J23)	1.0	0.854	10042	9471	0.7	0.041	4514	6239	hypothetical protein
Cd630-2365 (6L23)	1.1	0.553	23546	21263	0.7	0.007	12372	16524	putative sulfonate ABC transporter, solute-binding lipoprotein
Cd630-2371 (6J24)	1.0	0.927	2576	2542	0.8	0.041	1663	2287	L-aspartate oxidase
Cd630-2374 (6M24)	1.1	0.576	611	536	0.4	0.009	207	423	hypothetical protein
Cd630-2388 (7C2)	0.9	0.750	691	806	0.6	0.037	225	375	putative peptidoglycan- binding/hydrolysing protein
Cd630-2435 (7B8)	1.0	0.793	5792	5873	0.7	0.016	3070	4144	DNA repair protein
Cd630-2438 (7E8)	1.0	0.926	4413	4369	0.7	0.021	2366	3216	cytidine deaminase
Cd630-2442 (7A9)	0.9	0.497	955	1072	0.6	0.015	341	561	putative stage IV sporulation protein
Cd630-2444 (7C9)	1.0	0.867	9047	9387	1.6	0.047	9584	6463	putative sigma 54 modulation protein
Cd630-2456 (7G10)	1.0	0.545	1508	1428	1.6	0.035	1282	752	ABC transporter, ATP-binding protein
Cd630-2459 (7B11)	1.1	0.483	2455	2184	1.5	0.007	1696	1123	glucokinase
Cd630-2468 (7C12)	1.1	0.701	64	58	2.1	0.039	63	31	putative exported protein
Cd630-2473 (7H12)	1.0	0.754	9150	8941	0.7	0.030	7693	10723	30S ribosomal protein S20
Cd630-2493 (7D15)	1.3	0.609	161	139	2.1	0.047	165	80	selenocysteine-specific elongation factor
Cd630-2496 (7G15)	0.9	0.412	2624	2774	1.4	0.043	1602	1183	selenide,water dikinase selenide,water dikinase
Cd630-2500 (7C16)	1.0	0.959	1710	1627	0.6	0.022	838	1197	argininosuccinate lyase
Cd630-2517 (7D18)	0.9	0.557	637	659	1.6	0.043	570	352	putative regulatory protein
Cd630-2521 (7A19)	1.1	0.656	369	355	2.4	0.001	255	106	leucyl-tRNA synthetase
Cd630-2526 (7F19)	0.9	0.548	405	456	0.4	0.044	158	352	glycogen branching enzyme
Cd630-2541 (7E21)	1.1	0.527	21384	19986	1.3	0.007	12065	9388	sodium:dicarboxylate symporter family protein
Cd630-2544 (7H21)	0.9	0.588	912	908	0.6	0.007	269	470	putative membrane protein
Cd630-2545 (7A22)	1.0	0.480	2324	2384	0.6	0.020	947	1412	cellulose synthase catalytic subunit [UDP-forming]
Cd630-2549 (7E22)	1.0	0.540	13036	13357	0.8	0.014	6707	8390	putative sugar transporter, permease protein
Cd630-2553 (7A23)	1.0	0.875	3380	3129	0.7	0.041	1579	2289	phosphotransferase system, C component
Cd630-2556 (7D23)	1.0	0.946	1379	1358	0.6	0.016	724	1048	putative bifunctional protein [includes B component of phosphotransferase system and transcriptional antiterminator]
Cd630-2561 (7A24)	0.9	0.512	13261	14148	0.7	0.012	8592	12160	putative phosphatase
Cd630-2570 (7J13)	1.3	0.789	463	655	2.7	0.020	266	87	two-component sensor histidine kinase
Cd630-2591 (7O15)	1.0	0.623	3217	3126	0.9	0.046	1755	2009	low-specificity L-threonine aldolase

Cd630-2604f0 (7L17)	1.0	0.924	130	144	0.5	0.041	50	111	binary ADP-ribosyltransferase Cdt toxin (pseudogene)
Cd630-2611 (7N18)	1.1	0.607	2449	2368	1.9	0.022	2081	1107	5'-methylthioadenosine/S- adenosylhomocysteine nucleosidase
Cd630-2612 (7O18)	1.0	0.997	4638	4560	1.3	0.008	2706	2107	putative amino acid permease
Cd630-2630 (7I21)	1.0	0.945	6582	6405	1.3	0.036	3509	2759	glycerol-3-phosphate dehydrogenase [NAD(P)+]
Cd630-2640 (7K22)	0.9	0.686	905	1006	0.3	0.012	206	532	putative ATP-binding protein
Cd630-2644 (7O22)	0.5	0.440	397	480	0.3	0.028	125	371	sporulation sigma-E factor processing peptidase
Cd630-2650 (7M23)	2.0	0.442	3328	690	0.5	0.015	223	403	putative cell division protein
Cd630-2655 (7J24)	1.0	0.988	3763	3698	0.7	0.030	1823	2410	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase
Cd630-2656 (7K24)	1.0	0.854	1567	1558	0.8	0.030	855	1217	stage V sporulation protein D (sporulation specific penicillin-binding protein)
Cd630-2664 (7K1)	1.0	0.513	3770	3529	1.5	0.020	2640	1738	putative UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase
Cd630-2665 (7L1)	1.0	0.999	3319	3215	1.4	0.035	2467	1714	AraC-family transcriptional regulator
Cd630-2666 (7M1)	1.0	0.511	6497	6270	1.7	0.022	5746	3476	PTS system, glucose-specific IIa component
Cd630-2667 (7N1)	0.9	0.673	9688	10733	1.3	0.001	7871	5992	PTS system, glucose-specific IIbc component
Cd630-2670 (7I2)	1.0	0.755	3107	3038	1.4	0.014	2329	1674	oligopeptide ABC transporter, ATP-binding protein
Cd630-2701 (7P5)	0.8	0.490	2336	2283	1.5	0.007	2173	1406	putative lipoprotein
Cd630-2710 (7I7)	1.1	0.709	594	435	0.7	0.027	203	284	conserved hypothetical protein
Cd630-2734 (7J10)	1.2	0.633	177	106	2.4	0.043	82	39	putative malate-2H(+)/Na(+)-lactate antiporter
Cd630-2738 (7O10)	0.9	0.499	977	1040	0.5	0.001	446	855	putative cytosine permease
Cd630-2740 (7I11)	0.9	0.508	6394	7152	1.3	0.029	4221	3356	putative histidyl-tRNA synthetase
Cd630-2757 (8B1)	1.0	0.871	4141	4241	1.3	0.001	2791	2225	putative phospholipase
Cd630-2776 (8F3)	1.0	0.670	4001	3933	1.2	0.047	2542	2118	putative glycosyl transferase
Cd630-2778 (8H3)	1.1	0.610	6191	5860	1.2	0.048	3895	3339	putative polysaccharide biosynthesis protein
Cd630-2796b (8C7)	0.9	0.447	7627	8252	0.7	0.000	4465	6716	putative cell surface protein
Cd630-2805 (8F8)	1.0	0.575	2321	2180	0.8	0.009	1383	1793	holliday junction DNA helicase
Cd630-2806 (8G8)	1.0	0.876	1722	1698	0.6	0.024	832	1235	holliday junction DNA helicase
Cd630-2817 (8B10)	1.0	0.654	1149	1126	1.3	0.023	674	521	ABC transporter, ATP-binding/permease protein
Cd630-2820 (8E10)	0.9	0.794	427	429	0.6	0.000	161	254	putative membrane protein
Cd630-2821 (8F10)	0.9	0.401	542	592	0.8	0.036	288	358	putative membrane protein
Cd630-2841 (10B13)	1.1	0.569	277	224	1.8	0.028	190	111	putative amidohydrolase
Cd630-2863 (10H15)	1.5	0.389	47	23	4.3	0.012	61	16	putative sigma-54-dependent transcriptional regulator
Cd630-2866 (10C16)	1.0	0.787	112	106	1.5	0.009	109	72	putative regulatory protease
Cd630-2871 (10H16)	1.1	0.908	113	56	4.2	0.004	133	34	altronate hydrolase (N-terminus)
Cd630-2874 (10C17)	1.1	0.515	53	41	2.0	0.040	62	27	putative drug/sodium antiporter

Cd630-2879 (10H17)	1.1	0.514	596	492	1.7	0.036	411	256	hypothetical protein
Cd630-2883 (10D18)	1.1	0.789	127	88	3.0	0.003	166	51	PTS system, lichenan-specific IIc component
Cd630-2888 (10A19)	1.3	0.629	64	43	1.5	0.029	38	25	ABC transporter, permease protein
Cd630-2899 (10D20)	1.7	0.608	6840	1587	0.5	0.019	169	315	phage protein
Cd630-2948 (8D14)	1.3	0.505	32	24	4.1	0.040	77	18	putative phage anti-repressor
Cd630-2962 (8B16)	1.0	0.709	690	664	0.7	0.032	433	583	hypothetical protein
Cd630-2979 (8C18)	1.0	0.794	296	298	0.6	0.027	121	196	conserved hypothetical protein
Cd630-3003 (8C21)	0.9	0.649	15	16	5.7	0.005	46	8	probable carbohydrate hydrolase (N-terminus)
Cd630-3010 (8B22)	1.1	0.678	63	52	2.5	0.026	76	31	putative cytochrome C biogenesis protein
Cd630-3018 (8B23)	1.0	0.752	2087	2004	1.4	0.044	2390	1721	conserved hypothetical protein
Cd630-3030 (8F24)	0.8	0.710	110	114	1.7	0.002	95	54	PTS system, maltose and glucose-specific IIbc component
Cd630-3031 (8G24)	1.0	0.706	927	963	1.6	0.049	1000	640	transcription antiterminator
Cd630-3040 (8P13)	1.2	0.460	1342	1251	1.7	0.015	1126	659	conserved hypothetical protein
Cd630-3055 (8O15)	1.0	0.559	1591	1635	0.7	0.025	792	1027	putative ABC transporter, permease protein
Cd630-3060 (8L16)	1.0	0.858	905	900	0.7	0.005	657	909	putative maltose-6'-phosphate glucosidase
Cd630-3069 (8M17)	1.1	0.476	3343	3207	1.3	0.014	1943	1508	PTS system, IId component
Cd630-3089 (8I20)	0.9	0.417	34185	36982	0.5	0.025	14314	26188	PTS system, IIabc component
Cd630-3100 (8L21)	1.0	0.975	12614	12561	0.7	0.035	9360	12705	putative C4-dicarboxylate anaerobic carrier
Cd630-3102 (8N21)	1.1	0.472	814	693	1.4	0.047	549	383	putative dipeptidase
Cd630-3107 (8K22)	1.0	0.571	1270	1188	0.7	0.039	690	890	putative transcription antiterminator
Cd630-3110 (8N22)	0.9	0.666	459	443	0.8	0.012	227	284	putative phosphofructokinase
Cd630-3112 (8P22)	1.2	0.446	758	594	0.7	0.048	425	535	putative amidohydrolase
Cd630-3116 (8L23)	1.1	0.447	1738	1543	0.7	0.009	1120	1602	PTS system, beta-glucoside-specific IIabc component
Cd630-3119 (8O23)	1.0	0.960	582	535	0.7	0.044	295	406	putative glycosyltransferase
Cd630-3143 (9G2)	1.0	0.527	800	818	0.6	0.009	324	527	conserved hypothetical protein
Cd630-3159 (9B5)	1.2	0.485	203	138	2.5	0.020	158	62	putative acetyltransferase
Cd630-3161 (9D5)	0.8	0.428	632	729	0.6	0.011	232	383	putative ABC transporter, permease permease
Cd630-3181 (9H7)	0.9	0.861	538	515	0.4	0.023	129	326	putative amidohydrolase
Cd630-3200 (9D10)	1.4	0.442	129	65	2.5	0.040	90	33	putative ABC transporter, permease protein
Cd630-3204 (9H10)	1.2	0.461	148	109	2.3	0.017	155	69	putative membrane protein
Cd630-3207 (9C11)	0.8	0.526	697	783	0.3	0.027	154	419	putative Na+ driven multidrug efflux pump
Cd630-3211 (9G11)	1.0	0.965	608	609	0.5	0.017	186	330	conserved hypothetical protein
Cd630-3222 (9J1)	1.0	0.865	99	99	1.6	0.025	78	53	L-serine dehydratase
Cd630-3223 (9K1)	0.9	0.535	241	248	1.3	0.026	190	140	dihydrodipicolinate synthase

Cd630-3224 (9L1)	1.0	no replicates	64	64	3.9	0.008	133	33	aspartate-semialdehyde dehydrogenase
Cd630-3241 (9M3)	0.9	0.819	55	58	2.2	0.015	85	38	proline reductase
Cd630-3244 (9P3)	1.2	0.396	94	82	3.3	0.011	195	64	proline reductase subunit proprotein
Cd630-3257 (9M5)	1.0	0.693	1148	1178	0.6	0.030	330	514	putative polysaccharide deacetylase
Cd630-3258 (9N5)	2.8	0.460	1386	1248	0.6	0.003	714	1148	iron-only hydrogenase
Cd630-3259 (9O5)	0.9	0.487	1340	1401	0.7	0.024	676	923	putative membrane protein
Cd630-3277 (9I8)	1.0	0.880	684	643	0.5	0.047	263	525	PTS system, IIc component
Cd630-3283 (9O8)	1.0	0.869	3925	3934	0.6	0.020	2017	3255	putative pyruvate formate-lyase 3 activating enzyme
Cd630-3285 (9I9)	0.8	0.591	432	433	0.4	0.024	180	556	glucose-6-phosphate isomerase
Cd630-3290 (9N9)	1.1	0.591	6214	6015	1.2	0.022	3718	3055	putative protein translocase
Cd630-3297 (9M10)	0.9	0.405	656	690	0.6	0.023	209	335	putative membrane protein
Cd630-3301 (9I11)	1.2	0.466	433	377	1.8	0.006	237	133	ATP-dependent protease La
Cd630-3322 (9F13)	1.1	0.581	1484	1447	0.6	0.016	791	1359	hypothetical protein
Cd630-3328 (9D14)	0.9	0.658	794	824	0.6	0.042	380	633	congative transposon conserved hypothetical protein
Cd630-3329 (9E14)	0.9	0.494	1029	1142	0.5	0.032	454	858	conjugative transposon conserved hypothetical protein
Cd630-3335 (9C15)	0.9	0.581	2045	2109	0.6	0.014	1166	1836	conjugative tyransposon conserved hypothetical protein
Cd630-3339 (9G15)	1.1	0.826	389	303	0.5	0.023	154	271	conjugative transposon membrane protein
Cd630-3342 (9B16)	1.0	no replicates	55	55	6.5	0.025	105	16	putative phage membrane protein
Cd630-3380 (9H20)	0.7	0.394	588	797	0.3	0.009	225	602	conjugative transposon protein
Cd630-3384 (9D21)	1.0	0.952	376	358	0.6	0.017	141	233	conjugative transposon membrane protein
Cd630-3386 (9F21)	1.5	0.518	224	69	2.6	0.014	115	42	conjugative transposon protein
Cd630-3392 (9D22)	0.7	0.389	30	37	2.1	0.046	57	27	putative collagen-binding surface protein
Cd630-3403 (9G23)	1.0	0.792	5560	5534	0.7	0.009	2894	4154	putative formate/nitrite transporter
Cd630-3413 (9I13)	1.6	0.506	40	33	5.0	0.013	48	9	putative nucleotide pyrophosphatase
Cd630-3421 (9I14)	1.0	0.687	4424	4647	0.7	0.013	2159	3149	porphobilinogen deaminase
Cd630-3456 (9L18)	0.9	0.583	721	747	0.8	0.036	379	491	putative 5-formyltetrahydrofolate cyclo-ligase
Cd630-3457 (9M18)	1.0	0.773	773	722	0.6	0.022	279	478	putative exported protein
Cd630-3472 (9L20)	1.1	0.621	16182	14866	1.5	0.025	10961	7522	ATP synthase B chain
Cd630-3488 (9L22)	1.0	0.872	6215	6195	0.7	0.014	3239	4877	utp--glucose-1-phosphate uridylyltransferase (general stress protein 33)
Cd630-3490 (9N22)	1.0	0.914	1247	1239	0.8	0.008	604	824	stage II sporulation protein e
Cd630-3491 (9O22)	0.9	0.536	2152	2208	0.6	0.017	857	1480	putative exopolyphosphatase
Cd630-3498 (9N23)	1.0	0.748	9579	9056	0.7	0.006	5289	7178	stage V sporulation protein B
Cd630-3499 (9O23)	1.0	0.813	1723	1735	0.7	0.032	923	1266	stage V sporulation protein T
Cd630-3512 (10D1)	1.0	0.905	422	415	1.4	0.001	389	271	type IV pilus assembly protein

Cd630-3525 (10A3)	1.1	0.520	2393	2284	1.4	0.036	1582	1051	putative iron ABC transporter, solute-binding protein
Cd630-3549 (10A6)	0.9	0.415	1340	1559	0.6	0.037	679	1092	putative DNA polymerase III, delta' subunit
Cd630-3561 (10E7)	0.9	0.545	775	859	0.7	0.046	491	657	conserved hypothetical protein
Cd630-3565 (10A8)	1.0	0.760	4851	4651	0.8	0.007	3167	4153	GntR-family transcriptional regulator
Cd630-3574 (10B9)	0.7	0.518	97	109	0.5	0.002	31	54	putative membrane protein
Cd630-3601 (10E12)	0.9	0.796	771	790	0.5	0.013	272	449	D-alanyl-D-alanine carboxypeptidase
Cd630-3610 (10N1)	0.9	0.461	1314	1511	0.7	0.042	599	856	putative glyoxalase
Cd630-3623 (10K3)	1.0	0.572	924	892	0.8	0.048	447	528	putative ABC transporter, ATP-binding protein
Cd630-3627 (10O3)	1.1	0.607	550	462	0.6	0.008	155	262	putative dihydrodipicolinate synthase
Cd630-3632 (10L4)	0.9	0.599	368	394	0.5	0.042	98	191	putative isochorismatase
Cd630-3650 (10N6)	1.0	0.774	2120	2147	0.7	0.010	1195	1705	putative signaling protein
Cd630-3670 (10J9)	1.0	0.928	3799	4021	1.4	0.004	3017	2129	putative selenocysteine lyase

Bile 2.4% p=0.05
Kat's Bile MAVI Norm

	Control 2.40%				**Test 2.40%**				
Systematic ID	**Normalized**	**t-test P-value**	**Raw**	**Control**	**Fold change**	**t-test *P* value**	**Test**	**Control**	**Product**
Cd630-3238 (9J3)	1.2	0.510	62	59	3.4	0.050	90	30	conserved hypothetical protein
Cd630-2460 (7C11)	1.0	0.758	7952	7768	1.6	0.050	7666	5345	chaperone protein
Cd630-1318 (3M23)	1.0	0.811	6650	6971	0.7	0.050	4113	5591	polyribonucleotide nucleotidyltransferase
Cd630-3138 (9B2)	1.0	0.616	8907	9256	0.8	0.049	5026	6248	transcription antiterminator
Cd630-1038 (4L24)	1.0	0.791	422	382	0.5	0.049	112	198	putative helicase
Cd630-0044 (1L6)	1.1	0.906	57	37	2.3	0.048	43	20	putative membrane protein
Cd630-0592 (2J15)	1.0	0.936	457	445	0.4	0.048	119	320	putative membrane protein
Cd630-2759 (8D1)	1.0	0.991	764	765	0.6	0.048	221	380	conserved hypothetical protein
Cd630-1230 (3D24)	0.9	0.625	426	497	0.5	0.048	110	220	sigma K
Cd630-1616 (5P12)	1.1	0.714	2119	1931	0.5	0.048	442	895	putative signaling protein
Cd630-2654 (7I24)	1.0	0.945	4701	4854	0.6	0.048	2053	3487	phospho-N-acetylmuramoyl-pentapeptide-transferase
Cd630-1438 (5E2)	0.7	0.656	90	90	2.2	0.048	222	96	putative multiprotein complex assembly protein
Cd630-3595 (10G11)	1.0	0.955	125	125	1.7	0.048	122	74	aminopeptidase

Cd630-0295 (2G1)	1.0	0.874	60	64	6.0	0.048	113	19	putative iron-sulfur-binding protein
Cd630-0517 (1G5)	1.0	0.860	1553	1575	1.7	0.048	1837	1064	putative membrane protein
Cd630-0290 (2B1)	1.4	0.537	55	30	7.1	0.048	92	15	hypothetical protein
Cd630-2459 (7B11)	1.0	0.966	2315	2337	1.3	0.047	1563	1267	glucokinase
Cd630-0618 (2J18)	1.0	0.864	425	414	0.3	0.047	79	214	putative transcriptional regulator
Cd630-0875 (3O3)	1.0	0.648	20752	21377	0.6	0.047	8740	14913	ABC transporter, permease protein
Cd630-0255 (4G8)	1.1	0.670	6902	6833	0.5	0.047	1818	3476	flagellar hook protein
Cd630-1979 (6E10)	1.0	0.855	433	433	1.4	0.046	350	228	ABC transporter, substrate-binding protein
Cd630-0818 (1F20)	1.0	0.980	343	333	1.7	0.046	172	99	putative 6-phospho-beta-glucosidase
Cd630-1309 (3L22)	1.1	0.608	1542	1393	1.5	0.046	1209	843	translation initiation factor IF-2
Cd630-1873 (1K21)	0.9	0.843	45	61	11.3	0.046	1529	158	ABC transporter, permease protein
Cd630-0604 (2L16)	1.0	0.804	1069	1080	1.5	0.046	845	582	conserved hypothetical protein
Cd630-1091 (4A19)	1.0	0.913	180	176	2.3	0.046	209	114	integrase
Cd630-1681 (5A21)	1.0	0.770	1846	1880	1.2	0.046	1502	1278	TetR-family transcriptional regulator
Cd630-0536 (1B8)	1.0	0.919	943	917	0.3	0.045	168	550	chemotaxis protein
Cd630-2605f4 (7J18)	0.9	0.801	382	378	0.5	0.045	95	193	ADP-ribosyltransferase (pseudogene)
Cd630-1148 (3B14)	1.0	0.940	11758	11729	0.6	0.045	5174	8544	putative penicillin-binding protein
Cd630-2527 (7G19)	1.0	0.942	4226	4267	1.3	0.045	3996	3025	TetR-family transcriptional regulator
Cd630-1780 (5M21)	1.1	0.514	192	162	1.9	0.045	93	48	putative methylase
Cd630-2681 (7L3)	1.3	0.587	1388	1137	2.4	0.045	1546	562	putative Calcium-chelating exported protein
Cd630-2943 (8G13)	1.2	0.863	41	41	2.3	0.045	40	15	putative phage replication protein
Cd630-0388 (2L1)	1.0	0.573	3043	2972	1.3	0.045	2261	1743	PTS system, beta-glucoside-specific IIabc component
Cd630-1339 (4J2)	1.1	0.701	7561	6978	1.5	0.044	5802	4244	aspartate aminotransferase
Cd630-0588 (2N14)	0.9	0.670	314	300	0.6	0.044	131	224	hypothetical protein
Cd630-3099 (8K21)	1.0	0.757	17078	17545	0.7	0.044	8225	11204	putative amidohydrolase
Cd630-1421 (4L12)	0.9	0.659	2133	2239	0.5	0.044	814	1399	putative signaling protein
Cd630-2220 (8K5)	1.0	0.878	473	480	1.5	0.044	396	243	putative cAMP-binding regulatory protein
Cd630-0375 (2G11)	1.6	0.659	16	7	24.6	0.044	54	3	conjugative transposon protein
Cd630-2830 (8G11)	1.0	0.961	17200	17363	1.1	0.044	13564	11957	putative exported protein
Cd630-1977 (6C10)	1.0	0.808	491	523	0.4	0.044	94	278	DNA mismatch repair protein
Cd630-1645 (5E16)	1.0	0.991	5536	5583	1.6	0.043	5527	3212	putative membrane protein
Cd630-3423 (9K14)	1.0	0.934	548	535	0.3	0.043	131	307	cobalt-precorrin-6a reductase precorrin-6x reductase

Cd630-2564 (7D24)	0.9	0.765	487	529	1.3	0.043	426	327	MerR-family transcriptional regulator
Cd630-0381 (2E12)	1.2	0.603	57	48	2.5	0.043	55	26	putative conjugative transposon replication initiation factor
Cd630-3044 (8L14)	1.0	0.824	5243	5329	1.2	0.043	3758	3015	putative phosphosugar-binding transcriptional regulator
Cd630-0744 (3D11)	1.0	0.781	5173	5006	1.5	0.042	4250	2776	putative chemotaxis protein
Cd630-1168 (3F16)	0.9	0.512	475	474	1.2	0.042	411	348	putative membrane protein
Cd630-3536 (10D4)	1.0	0.972	135	133	0.7	0.042	67	88	putative phosphonate metabolism protein
Cd630-2437 (7D8)	1.0	0.487	14254	14947	0.5	0.042	4345	8446	GTP-binding protein
Cd630-0236 (4D6)	1.0	0.940	17735	17724	0.6	0.042	6549	10402	flagellar protein
Cd630-0422 (2N5)	1.1	0.474	907	855	2.6	0.042	1272	437	conserved hypothetical protein
Cd630-3488 (9L22)	1.0	0.940	5231	5284	0.7	0.042	2215	3252	utp--glucose-1-phosphate uridylyltransferase (general stress protein 33)
Cd630-0987 (4O17)	1.0	0.775	638	648	0.6	0.042	243	396	putative phosphoesterase
Cd630-0410 (2J4)	0.9	0.782	278	265	1.9	0.042	321	141	putative DNA replication protein
Cd630-1350 (4M3)	1.1	0.560	214	202	0.5	0.041	67	135	putative lantibiotic ABC transporter, permease protein
Cd630-2876 (10E17)	1.0	0.834	34	35	4.1	0.041	75	18	putative ferrichrome ABC transporter, permease protein
Cd630-2086 (6L12)	0.9	0.408	1240	1365	0.5	0.041	499	931	putative iron-sulfur flavoprotein
Cd630-0403 (2K3)	1.1	0.547	4230	4015	1.8	0.041	4716	2421	putative fructose-bisphosphate aldolase
Cd630-1471 (5G6)	1.0	0.451	2535	2435	1.9	0.041	3255	1737	putative membrane protein (putative phage infection protein)
Cd630-1254 (3M15)	1.0	0.933	7763	7779	1.2	0.041	5712	4821	conserved hypothetical protein
Cd630-3581 (10A10)	0.8	0.430	1977	2437	0.5	0.040	550	1225	conserved hypothetical protein
Cd630-1592 (5P9)	1.2	0.453	35	27	6.4	0.040	112	18	potassium-transporting ATPase B chain
Cd630-1458 (5A5)	0.9	0.481	625	641	0.7	0.039	292	408	putative flavoprotein
Cd630-0941 (3I12)	1.0	0.911	3741	3749	1.4	0.039	3208	2178	phage protein
Cd630-2406 (7E4)	1.0	0.976	1146	1084	1.5	0.039	794	544	putative lipoprotein
Cd630-2922 (10B23)	0.8	0.423	570	670	0.8	0.039	318	430	phage terminase large subunit
Cd630-3217 (9E12)	1.4	0.440	221	114	2.3	0.039	137	62	putative nogalamycin resistance protein
Cd630-0239 (4G6)	1.1	0.632	65526	55031	1.6	0.039	88983	52720	flagellin subunit
Cd630-3650 (10N6)	1.0	0.946	2499	2496	0.7	0.039	1093	1418	putative signaling protein
Cd630-0820 (1H20)	1.1	0.628	945	867	1.6	0.038	699	441	two-component response regulator
Cd630-2980 (8D18)	0.9	0.485	751	788	0.4	0.038	154	415	conserved hypothetical protein
Cd630-3048 (8P14)	0.9	0.747	668	750	0.7	0.038	277	437	PTS system, IIc component
Cd630-0441 (2J8)	0.9	0.429	503	617	0.5	0.038	197	432	putative arginine utilization sigma-54 dependent regulatory protein
Cd630-2604f2	1.0	0.859	215	197	0.2	0.038	20	124	binary ADP-ribosyltransferase Cdt toxin (pseudogene)

(7M17)									
Cd630-3088 (8P19)	1.0	0.777	12983	13040	1.8	0.038	12786	7139	putative cellobiose-phosphate degrading protein
Cd630-0257 (4A9)	1.1	0.698	24713	22854	0.5	0.037	9270	16317	chemotaxis protein
Cd630-0332 (2D6)	0.9	0.504	1390	1610	1.3	0.037	1426	1137	putative exosporium glycoprotein
Cd630-0396 (2L2)	0.9	0.597	40487	43202	0.6	0.037	16893	31439	R-phenyllactate dehydratase activator
Cd630-2161 (6G21)	0.8	0.398	1122	1443	0.4	0.037	360	795	putative membrane protein
Cd630-3168 (9C6)	1.0	0.868	470	429	1.7	0.037	361	235	conserved hypothetical protein
Cd630-1903 (6A1)	0.6	0.531	34	39	3.8	0.036	49	13	putative lipoprotein signal peptidase
Cd630-2682 (7M3)	1.0	0.767	70554	72927	1.2	0.036	80886	69065	pyruvate-flavodoxin oxidoreductase
Cd630-3432 (9L15)	1.0	0.946	2144	2125	0.3	0.036	214	911	pthreonine-phosphate decarboxylase
Cd630-1032 (4L23)	1.0	0.998	62	65	4.0	0.036	57	15	putative exported protein
Cd630-1167 (3E16)	1.1	0.718	830	753	1.8	0.036	912	481	putative phage-related integrase/recombinase
Cd630-2827 (8D11)	1.0	0.990	1299	1244	0.3	0.035	256	745	putative MarR-family regulatory protein
Cd630-0001 (1I1)	1.0	0.959	352	334	0.5	0.035	110	221	chromosomal replication initiator protein
Cd630-3420 (9P13)	1.0	0.958	3233	3351	0.7	0.035	1308	1705	porphyrin biosynthesis protein includes: uroporphyrinogen-III methyltransferase and uroporphyrinogen-III synthase]
Cd630-1688 (5H21)	1.0	0.802	5212	5259	0.9	0.035	2714	3083	two-component response regulator
Cd630-0663a (2L24)	1.0	0.929	2457	2492	0.6	0.035	1022	1763	toxin A
Cd630-3101 (8M21)	0.9	0.773	2152	2485	0.6	0.035	738	1296	PTS system, IIbc component
Cd630-0663h (3C1)	1.0	0.630	2582	2535	1.2	0.035	1701	1439	toxin A
Cd630-1698 (5B23)	1.0	0.796	1318	1316	1.1	0.035	931	838	riboflavin biosynthesis protein
Cd630-3310 (9J12)	1.0	0.970	2287	2294	1.4	0.035	1868	1327	putative D-isomer specific 2-hydroxyacid dehydrogenase
Cd630-0163 (2C21)	0.9	0.816	155	158	0.2	0.034	32	165	hypothetical protein
Cd630-0530 (1D7)	1.0	0.867	110	101	0.6	0.034	33	56	putative membrane protein
Cd630-0157 (2E20)	1.0	0.732	305	284	0.6	0.034	95	155	putative membrane protein (pseudogene)
Cd630-0051 (1K7)	1.1	0.403	3754	3565	1.5	0.034	3230	1950	glutamyl-tRNA synthetase
Cd630-1955 (6E7)	1.0	0.757	638	630	0.7	0.034	280	385	ABC transporter, ATP-binding protein
Cd630-2807 (8H8)	1.0	0.592	1404	1397	0.8	0.033	645	807	crossover junction endodeoxyribonuclease RuvC
Cd630-2105 (6G14)	1.4	0.417	41	23	3.1	0.033	43	14	ABC transporter, ATP-binding protein
Cd630-1108 (4B21)	1.5	0.517	48	28	2.4	0.033	60	26	putative DNA-repair protein
Cd630-1754 (5K18)	1.0	0.991	20	20	0.1	0.033	3	18	putative ABC transporter, permease protein
Cd630-2269 (8L11)	0.8	0.474	136	143	2.2	0.033	197	92	PTS system, IIbc component
Cd630-2101 (6C14)	1.0	0.908	982	981	0.5	0.033	254	602	putative FAD-binding subunit of oxidoreductase

Cd630-2678 (7I3)	1.1	0.905	42	37	2.4	0.033	47	19	butyrate--acetoacetate CoA-transferase subunit B
Cd630-1694 (5F22)	0.9	0.582	1295	1446	1.9	0.033	1457	716	hypothetical protein
Cd630-2653 (7P23)	1.0	0.979	5707	5640	0.5	0.033	1534	4012	UDP-N-acetylmuramoylalanine--D-glutamate ligase
Cd630-0904 (3L7)	1.0	0.899	867	831	0.5	0.033	326	566	phage integrase
Cd630-0079 (1O10)	0.9	0.608	37802	40851	0.8	0.032	16366	22192	30S ribosomal protein S3
Cd630-1445 (5D3)	1.0	0.790	4142	4134	0.3	0.032	755	2440	para-aminobenzoate synthase glutamine amidotransferase component II
Cd630-0308 (2D3)	1.0	0.455	1346	1413	1.3	0.032	1034	782	putative membrane protein
Cd630-1745 (5I17)	1.0	0.918	665	640	0.6	0.032	230	392	hypothetical protein
Cd630-0228 (4D5)	1.1	0.440	2129	1915	0.6	0.032	754	1323	flagellar motor switch protein
Cd630-2849 (10B14)	1.1	0.503	305	258	0.5	0.032	70	140	putative bifunctional protein [includes phosphonoacetaldehyde phosphonohydrolase; 2-aminoethylphosphonate:pyruvate aminotransferase]
Cd630-1866 (1L20)	1.0	0.842	440	417	0.5	0.032	105	241	putative conjugative transposon conserved hypothetical protein
Cd630-1117 (4C22)	1.3	0.724	45	25	4.0	0.032	58	16	conserved hypothetical protein
Cd630-0646 (2N21)	1.3	0.535	93	49	2.1	0.031	62	30	putative lantibiotic ABC transporter, permease protein
Cd630-2624 (7K20)	1.1	0.794	3418	3541	0.5	0.031	998	2032	putative exported protein
Cd630-0116 (2D15)	1.1	0.409	3337	3164	2.2	0.031	3720	1756	putative oxidoreductase, thiamine diP-binding subunit
Cd630-2393 (7H2)	1.0	0.930	1182	1196	0.6	0.031	409	721	putative transglycosylase
Cd630-1195 (3A20)	0.8	0.516	357	405	0.4	0.031	97	218	stage III sporulation protein AD
Cd630-1413 (4L11)	0.9	0.455	872	932	0.6	0.031	391	611	putative membrane protein
Cd630-0686 (3B4)	1.0	0.862	6290	6343	0.5	0.031	2741	4677	50S ribosomal protein L35
Cd630-2963 (8C16)	1.1	0.473	2921	2787	0.4	0.031	638	1518	putative peptidoglycan-binding exported protein
Cd630-2030 (6L5)	1.0	0.667	1158	1197	1.3	0.031	953	710	ornithine carbamoyltransferase
Cd630-2786b (8B5)	1.0	0.629	902	946	2.0	0.031	977	488	putative cell surface protein
Cd630-1303 (3N21)	1.1	0.471	1652	1496	2.1	0.030	1758	891	putative NAD-dependent deacetylase (Sir2-family regulatory)
Cd630-2259 (8J10)	1.0	0.915	1483	1454	0.4	0.030	485	1075	conserved hypothetical protein
Cd630-1009 (4M20)	1.0	0.739	1823	1984	1.6	0.030	1449	989	GntR-family transcriptional regulator
Cd630-2962 (8B16)	1.0	0.919	767	824	0.5	0.030	218	441	hypothetical protein
Cd630-1625 (5A14)	1.1	0.546	1771	1620	1.8	0.030	1605	948	two-component sensor histidine kinase
Cd630-2927 (10G23)	1.0	0.803	2694	2632	1.9	0.030	2763	1677	putative phage DNA-binding protein
Cd630-1802 (5K24)	1.0	0.990	790	784	0.5	0.030	274	521	putative hydrolase
Cd630-1575 (5O7)	1.0	0.985	327	337	0.3	0.029	69	213	putative exported protein
Cd630-1746 (5J17)	1.0	0.953	55	50	3.4	0.029	80	27	sodium/glutamate symporter

Cd630-0508 (1E4)	1.0	0.627	44252	45147	0.8	0.029	24928	30726	conjugative transposon tetracycline resistance protein
Cd630-2476 (7C13)	1.0	0.729	3168	3434	1.6	0.029	3434	1980	conserved hypothetical protein
Cd630-0894 (3J6)	1.0	0.783	6715	6675	0.6	0.029	2744	4564	iron-dependent hydrogenase
Cd630-3078 (8N18)	1.4	0.412	310	172	2.1	0.029	214	107	putative membrane protein
Cd630-2866 (10C16)	1.0	0.876	173	181	1.4	0.029	110	76	putative regulatory protease
Cd630-2828 (8E11)	1.0	0.731	2571	2494	1.3	0.029	1778	1363	putative aromatic amino acid aminotransferase
Cd630-2346 (6I21)	0.8	0.399	835	987	0.5	0.029	464	1069	putative membrane protein
Cd630-3330 (9F14)	1.0	0.885	1245	1262	1.4	0.029	1196	871	putative transposon-related DNA-binding protein
Cd630-1545 (5I4)	1.1	0.530	2467	2372	1.7	0.029	2140	1167	conserved hypothetical protein
Cd630-0351 (2G8)	1.0	0.421	3112	3321	1.3	0.029	2757	2120	hypothetical protein
Cd630-1474 (5B7)	1.1	0.769	2379	2195	1.8	0.029	2581	1279	putative ruberythrin
Cd630-0193 (4A1)	1.0	0.839	3204	3258	1.3	0.028	2751	2062	10 kDa chaperonin
Cd630-3027 (8C24)	1.0	0.864	4511	4496	1.4	0.028	3333	2491	PTS system, glucose-specific IIa component
Cd630-3188 (9G8)	1.0	0.780	758	786	0.7	0.028	353	533	hypothetical protein
Cd630-0202 (4B2)	1.0	0.993	643	645	0.5	0.028	210	388	conserved hypothetical protein
Cd630-0349 (2E8)	1.1	0.586	1589	1494	1.6	0.028	1448	856	conserved hypothetical protein (pseudogene)
Cd630-3033 (8I13)	1.1	0.709	153	147	1.6	0.028	130	81	thioredoxin
Cd630-2648 (7K23)	1.2	0.438	1799	1402	0.7	0.028	616	943	putative exported protein
Cd630-2276 (8K12)	1.0	0.903	1417	1448	2.1	0.028	1848	1050	putative sodium:solute symporter
Cd630-1301 (3L21)	1.0	0.928	1732	1710	1.1	0.028	1162	1001	putative membrane protein
Cd630-2553 (7A23)	1.0	0.829	3243	3227	1.2	0.027	2268	1881	phosphotransferase system, C component
Cd630-3392 (9D22)	1.1	0.824	32	29	3.3	0.027	62	22	putative collagen-binding surface protein
Cd630-3067 (8K17)	1.0	0.786	1863	1869	1.4	0.027	1543	1051	PTS system, IIa component
Cd630-1256 (3O15)	0.9	0.493	801	891	0.4	0.027	242	566	tRNA (Guanine-n(1)-)-methyltransferase
Cd630-1007 (4K20)	1.0	0.923	260	269	2.3	0.027	561	206	conserved hypothetical protein
Cd630-0562 (1D11)	1.0	0.950	642	635	0.5	0.027	163	350	putative endonuclease
Cd630-3373 (9A20)	1.4	0.409	1540	948	2.2	0.027	1113	488	magnesium transporting ATPase, P-type
Cd630-2463 (7F11)	1.0	0.863	9219	9344	1.7	0.027	8907	5534	heat-inducible transcription repressor
Cd630-3523 (10G2)	0.9	0.593	590	594	0.5	0.027	164	354	dimethyladenosine transferase
Cd630-1848 (1J18)	1.3	0.793	31	13	3.6	0.027	31	9	putative conjugative transposon protein
Cd630-3419 (9O13)	0.9	0.637	3313	3638	1.6	0.026	3865	2283	delta-aminolevulinic acid dehydratase
Cd630-0899 (3O6)	0.9	0.469	1451	1586	1.4	0.026	1497	1066	DNA polymerase IV
Cd630-3628 (10P3)	0.9	0.452	637	730	0.4	0.026	182	392	putative sugar-phosphate isomerase

Cd630-1177 (3G17)	1.0	0.932	13890	13888	0.8	0.026	6827	9086	DeoR-family transcriptional regulator (fatty acid and phospholipid biosynthesis regulator)
Cd630-2128 (6F17)	1.5	0.582	77	57	2.6	0.026	67	28	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
Cd630-1253 (3L15)	1.0	0.832	1900	1880	0.5	0.026	629	1244	30S ribosomal protein S16
Cd630-2500 (7C16)	1.0	0.778	1625	1600	0.5	0.026	489	958	argininosuccinate lyase
Cd630-3139 (9C2)	1.0	0.926	188	181	1.6	0.026	176	109	putative transcriptional regulator
Cd630-0056 (1P7)	1.0	0.592	4341	4306	1.3	0.026	3177	2390	conserved hypothetical protein
Cd630-0222 (4F4)	1.0	0.866	4456	4526	0.6	0.026	1948	3088	phosphoribosylglycinamide formyltransferase
Cd630-0889 (3M5)	1.1	0.584	15507	12930	0.7	0.026	6437	8961	S-adenosylmethionine decarboxylase proenzyme
Cd630-3149 (9H3)	1.0	0.994	1855	1853	0.4	0.026	355	1031	putative DNA helicase
Cd630-0271 (4G10)	1.0	0.805	3625	3713	0.6	0.026	1364	2239	putative flagellar motor switch protein
Cd630-2784b (8G4)	1.0	0.794	5993	5944	0.5	0.026	2268	4533	putative cell surface protein
Cd630-2386 (7A2)	1.3	0.591	50	34	4.3	0.026	59	15	conserved hypothetical protein
Cd630-2462 (7E11)	1.0	0.820	15476	14813	1.2	0.025	9454	7910	heat shock protein
Cd630-0817 (1E20)	1.0	0.975	6991	6964	1.3	0.025	5848	4252	putative transposase-like protein B
Cd630-3393 (9E22)	1.0	no replicates	13	13	3.1	0.025	50	14	putative RNA methyltransferase
Cd630-2432 (7G7)	1.0	0.999	5014	4901	1.6	0.025	5120	3329	glycyl-tRNA synthetase beta chain
Cd630-0077 (1M10)	0.9	0.468	17053	19724	0.8	0.025	11064	14611	30S ribosomal protein S19
Cd630-1579 (5K8)	0.9	0.543	2290	2646	0.7	0.025	963	1354	putative two-component histidine kinase
Cd630-1909 (6G1)	1.0	0.954	26	28	4.8	0.025	75	17	putative ethanolamine/propanediol utilization protein
Cd630-1391 (4N8)	0.8	0.660	74	80	0.4	0.025	28	92	putative esterase
Cd630-2701 (7P5)	1.0	0.998	3307	3282	1.3	0.025	2221	1805	putative lipoprotein
Cd630-2640 (7K22)	1.1	0.496	917	902	0.3	0.025	194	576	putative ATP-binding protein
Cd630-0224 (4H4)	1.0	0.965	6799	6775	0.5	0.025	2329	4436	phosphoribosylamine--glycine ligase
Cd630-3239 (9K3)	1.0	0.523	950	961	4.3	0.024	2573	583	conserved hypothetical protein
Cd630-1964 (6F8)	0.9	0.576	338	380	1.4	0.024	298	206	conserved hypothetical protein
Cd630-0326 (2F5)	1.0	0.915	486	542	0.3	0.024	121	500	cobalt transport protein
Cd630-3651 (10O6)	1.0	0.939	3065	3049	0.9	0.024	1753	2052	putative metallo-beta-lactamase superfamily protein
Cd630-0215 (4G3)	0.9	0.637	556	638	0.4	0.024	161	364	transposase
Cd630-1934 (6H4)	0.9	0.551	1479	1624	0.6	0.024	530	980	conserved hypothetical protein
Cd630-0258 (4B9)	1.0	0.926	7814	7799	0.6	0.024	3265	5469	flagellar basal body-associated protein
Cd630-0924 (3P9)	1.5	0.471	39	20	8.5	0.023	68	9	hypothetical phage protein
Cd630-1220 (3B23)	1.1	0.487	3897	3673	0.4	0.023	1276	2835	ADP-ribose pyrophosphatase

Cd630-3377 (9E20)	1.0	0.834	5519	5277	1.8	0.023	4994	2748	magnesium transporting ATPase, P-type 1
Cd630-1047 (4E13)	1.1	0.595	2235	2051	0.7	0.023	1089	1544	putative cell surface protein
Cd630-2004 (6O1)	1.0	0.872	424	424	3.3	0.023	885	234	hypothetical protein
Cd630-1651 (5C17)	1.0	0.946	526	538	0.5	0.023	178	319	putative signaling protein
Cd630-0729 (3E9)	1.0	0.614	2048	2097	0.3	0.023	489	1341	putative glycine cleavage system H protein
Cd630-3633 (10M4)	0.9	0.680	782	770	0.4	0.023	212	521	putative membrane protein
Cd630-1837 (1O16)	0.9	0.685	1567	1801	0.4	0.023	531	1181	shikimate dehydrogenase
Cd630-2285 (6L13)	0.8	0.471	845	901	0.2	0.023	99	446	GntR-family transcriptional regulator
Cd630-2452 (7C10)	1.0	0.983	1189	1277	0.3	0.023	198	646	hypothetical protein
Cd630-2018 (6P3)	1.0	0.623	1374	1359	0.7	0.022	538	776	AraC-family transcriptional regulator
Cd630-2540 (7D21)	1.0	0.555	15184	15407	1.6	0.022	16547	10110	NADH oxidase
Cd630-0171 (2C22)	1.0	0.600	1697	1658	1.8	0.022	1941	1062	putative DNA-binding protein
Cd630-3335 (9C15)	1.0	0.749	1933	1960	1.3	0.022	1678	1312	conjugative tyransposon conserved hypothetical protein
Cd630-3620 (10P2)	1.0	0.917	1792	1776	0.8	0.022	864	1042	hypothetical protein
Cd630-2450 (7A10)	1.0	0.465	1653	1682	0.7	0.022	930	1213	putative ribosomal protein L11 methyltransferase
Cd630-2545 (7A22)	1.0	0.933	1972	1968	0.8	0.022	988	1287	cellulose synthase catalytic subunit [UDP-forming]
Cd630-2933 (10E24)	1.1	0.387	276	219	2.0	0.022	274	146	hypothetical phage protein
Cd630-1975 (6A10)	1.0	0.555	1466	1533	0.7	0.022	727	927	tRNA delta(2)-isopentenylpyrophosphate transferase
Cd630-2461 (7D11)	0.9	0.375	15557	16844	1.4	0.022	12127	8839	chaperone protein
Cd630-2605f1 (7O17)	0.9	0.471	525	548	0.5	0.022	148	327	ADP-ribosyltransferase (pseudogene)
Cd630-1011 (4O20)	1.2	0.495	595	510	1.9	0.022	439	218	glucosamine-6-phosphate deaminase
Cd630-3425 (9M14)	1.0	0.719	2951	3184	0.5	0.022	1155	2455	putative cobalamin biosynthesis protein
Cd630-0179 (2C23)	1.0	0.506	5942	5805	1.7	0.021	8185	5031	NAD-specific glutamate dehydrogenase
Cd630-0942 (3J12)	1.0	0.735	1855	1786	1.8	0.021	2444	1315	hypothetical phage protein
Cd630-0902 (3J7)	1.0	0.815	1126	1147	1.7	0.021	1199	703	putative cation efflux rotein
Cd630-2616 (7K19)	1.0	0.771	1514	1511	1.3	0.021	1328	1023	hypothetical protein
Cd630-1145 (3G13)	1.0	0.975	34555	34637	0.5	0.021	12485	23295	rod shape-determining protein
Cd630-2172 (6B23)	1.1	0.430	5946	5580	0.6	0.021	2454	4076	probable amino-acid ABC transporter, ATP-binding protein
Cd630-1550 (5N4)	1.0	0.701	3007	3040	0.5	0.021	1013	2168	imidazoleglycerol-phosphate dehydratase
Cd630-0353 (2A9)	1.1	0.673	2207	2121	1.3	0.021	1357	1038	hypothetical protein
Cd630-3356 (9H17)	1.0	0.631	2336	2382	1.3	0.021	1630	1250	putative transcriptional regulator
Cd630-0809 (1E19)	1.0	0.919	937	959	0.7	0.021	440	581	conserved hypothetical protein

Cd630-0327 (2G5)	1.1	0.412	9809	8852	0.4	0.021	2544	5753	cobalt ABC transporter, ATP-binding protein
Cd630-1589 (5M9)	1.1	0.571	646	557	1.4	0.021	438	321	putative ribose ABC transporter, substrate-binding lipoprotein
Cd630-2363 (6J23)	1.0	0.648	9275	9178	0.5	0.020	3756	7576	hypothetical protein
Cd630-0183 (2G23)	1.0	0.798	701	745	0.3	0.020	157	533	putative exported protein
Cd630-0452 (2M9)	0.9	0.612	3257	3409	1.3	0.020	2324	1716	putative DNA repair protein
Cd630-2823 (8H10)	1.0	0.839	923	934	0.4	0.020	223	486	conserved hypothetical protein
Cd630-1119 (4E22)	0.7	0.774	20	18	8.8	0.020	81	9	putative lipoprotein
Cd630-0272 (4H10)	1.0	0.665	4486	4760	0.4	0.020	1712	3682	conserved hypothetical protein
Cd630-0204 (4D2)	1.0	0.950	2017	1960	0.5	0.020	647	1141	putative signaling protein
Cd630-1751b (5P17)	1.0	0.991	83	88	3.1	0.020	110	36	putative cell surface-associated cysteine protease
Cd630-1551 (5O4)	1.0	0.897	54	51	2.0	0.020	42	21	putative imidazole glycerol phosphate synthase subunit
Cd630-1503 (5G10)	1.0	0.815	751	771	0.5	0.020	258	483	putative transcriptional regulator
Cd630-1088 (4F18)	1.2	0.378	81	67	2.3	0.019	83	32	putative membrane protein
Cd630-0249 (4A8)	1.0	0.713	27605	28534	0.4	0.019	8723	20400	flagellar motor switch protein
Cd630-0413 (2M4)	0.7	0.581	44	54	3.4	0.019	57	17	putative single-strand binding protein
Cd630-3578 (10F9)	0.9	0.485	1159	1261	1.4	0.019	1010	726	putative resolvase
Cd630-3621 (10I3)	1.0	0.983	1394	1354	0.4	0.019	296	759	conserved hypothetical protein
Cd630-1761 (5J19)	1.0	0.945	1587	1604	1.3	0.019	1407	1051	conserved hypothetical protein
Cd630-0254 (4F8)	1.0	0.833	18882	21094	0.5	0.019	7344	14992	putative basal-body rod modification protein
Cd630-3375 (9C20)	1.0	0.996	931	936	2.3	0.019	1035	472	Mabnesium transporting ATPase protein C
Cd630-2536 (7H20)	0.9	0.452	1284	1325	0.5	0.019	344	741	two-component response regulator
Cd630-3111 (8O22)	1.0	0.973	868	853	0.5	0.019	256	537	conserved hypothetical protein
Cd630-1099 (4A20)	1.0	0.857	57	58	3.6	0.019	67	22	two-component response regulator
Cd630-2345 (6P20)	1.0	0.762	2850	2908	0.6	0.019	1143	1862	LysR-family regulatory protein
Cd630-1783 (5P21)	1.0	0.722	847	905	0.3	0.019	167	474	two component sensor histidine kinase
Cd630-1210 (3H21)	1.0	0.415	1678	1650	0.6	0.019	589	1036	putative exported protein
Cd630-1720 (5P13)	1.0	0.795	11371	11311	0.6	0.019	4840	7593	putative hydantoinase
Cd630-0220 (4D4)	1.0	0.964	11615	11514	0.5	0.019	4168	7850	amidophosphoribosyltransferase
Cd630-3337 (9E15)	1.5	0.554	37	19	4.1	0.019	49	13	conjugative transposon membrane protein
Cd630-2413 (7D5)	1.5	0.422	113	78	2.2	0.019	107	51	conserved hypothetical protein
Cd630-2169 (6G22)	0.9	0.573	224	259	1.7	0.019	301	167	iron-sulfur binding protein
Cd630-0379 (2C12)	0.9	0.768	216	202	1.6	0.019	177	112	conjugative transposon protein
Cd630-3608 (10L1)	0.9	0.665	418	419	0.3	0.018	89	286	ABC transporter, ATP-binding protein

Cd630-1143 (3E13)	1.0	0.782	14702	14879	0.7	0.018	6418	8944	septum formation protein
Cd630-3271 (9K7)	1.0	0.829	2493	2565	0.4	0.018	917	2168	hypothetical protein
Cd630-1236 (3K13)	1.0	0.942	6091	6335	0.8	0.018	3141	3959	putative exported protein
Cd630-3622 (10J3)	1.0	0.654	937	890	0.6	0.018	361	625	putative phosphonoacetate hydrolase
Cd630-3490 (9N22)	0.9	0.626	1062	1112	0.8	0.018	618	784	stage II sporulation protein e
Cd630-1071 (4E16)	0.9	0.779	425	520	0.5	0.018	152	296	conserved hypothetical protein
Cd630-3524 (10H2)	0.9	0.592	825	830	0.4	0.018	199	465	putative DNA-binding protein
Cd630-0550 (1H9)	1.0	0.880	586	579	1.3	0.018	453	347	putative membrane protein
Cd630-0395 (2K2)	1.0	0.814	43426	44207	0.4	0.018	9570	24011	E-cinnamoyl-CoA:R-phenyllactate CoA transferase large subunit
Cd630-2958 (8F15)	1.0	0.933	895	860	0.6	0.018	407	666	V-type sodium ATP synthase subunit E
Cd630-3280 (9L8)	0.9	0.432	1216	1356	0.8	0.018	522	675	putative transcription antiterminator
Cd630-3584 (10D10)	1.0	0.995	162	124	3.0	0.018	129	43	putative ABC transporter, permease protein
Cd630-1599 (5O10)	0.9	0.593	264	269	0.3	0.018	60	194	putative phosphomethylpyrimidine kinase
Cd630-3247 (9K4)	1.1	0.575	148	130	3.4	0.017	386	102	putative electron transfer protein
Cd630-0325 (2E5)	1.0	0.895	6212	6156	0.4	0.017	1580	3463	cobalt transport protein
Cd630-1147 (3A14)	1.0	0.971	7977	8008	0.5	0.017	3189	5922	putative membrane protein
Cd630-2445 (7D9)	1.2	0.551	1648	1491	2.8	0.017	2349	780	probale sensory protein
Cd630-3544 (10D5)	1.0	0.889	902	920	1.8	0.017	863	500	AsnC-family transcriptional regulator
Cd630-0908 (3P7)	0.9	0.619	55	55	3.9	0.017	87	24	phage protein
Cd630-2439 (7F8)	1.0	0.795	343	355	0.4	0.017	73	223	putative membrane-associated kinase and phosphatase
Cd630-1384 (4O7)	1.0	0.984	888	881	0.3	0.016	217	616	conserved hypothetical protein
Cd630-3531 (10G3)	1.0	0.727	113	111	2.5	0.016	105	42	putative phophoesterase
Cd630-0177 (2A23)	1.0	0.838	928	919	2.1	0.016	1286	596	cyclopropane-fatty-acyl-phospholipid synthase
Cd630-3185 (9D8)	1.1	0.435	881	823	0.5	0.016	252	463	conserved hypothetical protein (pseudogene)
Cd630-2772 (8B3)	0.7	0.397	256	353	0.3	0.016	85	313	putative teichuronic acid biosynthesis glycosyl transferase
Cd630-1166 (3D16)	1.0	0.897	4222	4476	1.3	0.016	3772	2701	homoserine dehydrogenase
Cd630-0758 (1B13)	0.9	0.632	492	527	1.8	0.016	459	261	pyruvate formate-lyase activating enzyme
Cd630-1607 (5O11)	1.3	0.708	41	32	2.6	0.016	64	21	ABC transporter, ATP-binding protein
Cd630-1215 (3E22)	1.0	0.984	5426	5470	0.7	0.016	2195	3114	putative pyrophosphokinase
Cd630-1935 (6A5)	1.1	0.557	2030	1883	1.9	0.016	1788	970	stage V sporulation protein S
Cd630-0197 (4E1)	1.0	0.408	1556	1643	1.4	0.016	1329	961	conserved hypothetical protein (pseudogene)
Cd630-2173 (6C23)	1.0	0.987	5682	5633	0.4	0.016	1443	3948	putative peptidase
Cd630-3347 (9G16)	1.0	0.825	1666	1655	0.7	0.016	775	1085	putative membrane protein

Cd630-0821 (1A21)	1.0	0.989	676	680	1.7	0.016	753	428	two-component sensor histidine kinase
Cd630-2013 (6K3)	1.0	0.573	1514	1428	0.6	0.016	554	842	TetR-family transcriptional regulator
Cd630-0218 (4B4)	1.0	0.994	3733	3726	0.5	0.015	1293	2763	phosphoribosylaminoimidazole carboxylase catalytic subunit
Cd630-3061 (8M16)	1.1	0.437	609	492	1.7	0.015	463	265	PTS system, IIbc component
Cd630-0230 (4F5)	1.0	0.970	4966	5076	0.5	0.015	1856	3494	putative flagellar biosynthesis protein
Cd630-2509 (7D17)	1.0	0.860	3568	3526	0.4	0.015	891	2180	probable 6-phospho-alpha-glucosidase
Cd630-0739 (3G10)	1.0	0.980	4958	5095	0.8	0.015	2467	3156	putative exported protein
Cd630-2067 (6I10)	1.0	0.552	2294	2243	0.7	0.015	1009	1396	putative exported protein
Cd630-0192 (2H24)	1.1	0.458	2574	2443	1.3	0.015	1868	1378	cardiolipin synthetase
Cd630-2454 (7E10)	0.9	0.451	1295	1373	0.4	0.015	335	775	hypothetical protein
Cd630-0975 (4K16)	1.0	0.938	5192	5443	0.5	0.015	1623	3005	hypothetical phage protein
Cd630-1841 (1K17)	1.0	0.785	881	882	0.3	0.015	111	468	putative signaling protein
Cd630-1294 (3M20)	1.1	0.659	3579	3282	1.6	0.015	2888	1894	putative membrane-associated peptidase
Cd630-1096 (4F19)	1.0	0.839	194	197	2.1	0.015	206	102	putative lantibiotic ABC transporter, permease protein
Cd630-0253 (4E8)	0.9	0.436	13153	15136	0.4	0.015	4611	12281	putative flagellar hook-length control protein
Cd630-3354 (9F17)	1.0	0.651	976	1002	0.7	0.014	401	613	conserved hypothetical protein
Cd630-1765 (5N19)	1.0	0.887	299	288	1.4	0.014	266	199	Nudix-family hydrolase
Cd630-0307 (2C3)	1.0	0.997	1848	1914	1.5	0.014	1706	1117	putative 'four-disulfide core' protein
Cd630-1665 (5A19)	1.0	0.932	4634	4486	1.4	0.014	3630	2652	cysteine synthase A
Cd630-0250 (4B8)	0.9	0.612	14677	16779	0.4	0.014	5187	11932	flagellar assembly protein
Cd630-0334 (2F6)	1.2	0.436	2301	2033	2.0	0.013	2827	1456	aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase; acetaldehyde dehydrogenase [acetylating]; pyruvate-formate-lyase deactivase
Cd630-2228 (8K6)	1.2	0.426	1287	1162	1.6	0.013	1101	703	putative membrane protein
Cd630-0270 (4F10)	1.0	0.942	7106	7692	0.6	0.013	3519	5501	putative flagellar motor switch protein
Cd630-2656 (7K24)	1.0	0.846	1391	1381	1.4	0.013	1092	798	stage V sporulation protein D (sporulation specific penicillin-binding protein)
Cd630-0003 (1K1)	1.1	0.442	506	442	0.5	0.013	152	294	putative RNA-binding mediating protein
Cd630-3670 (10J9)	1.0	0.834	4413	4348	1.2	0.013	3151	2581	putative selenocysteine lyase
Cd630-2639 (7J22)	0.9	0.604	1354	1390	0.3	0.013	301	785	conserved hypothetical protein
Cd630-0266 (4B10)	1.0	0.851	6735	6819	0.3	0.013	1289	4680	RNA polymerase sigma factor for flagellar operon
Cd630-2361 (6P22)	1.0	0.753	12332	12418	0.4	0.013	3450	8576	putative aliphatic sulfonate ABC transporter, ATP-binding protein
Cd630-1556 (5L5)	1.0	0.785	1108	1106	0.5	0.013	298	620	putative polysaccharide deacetylase
Cd630-0394 (2J2)	1.0	0.710	15020	15492	0.5	0.013	5508	12124	D-lactate dehydrogenase

Cd630-3117 (8M23)	1.0	0.977	1777	1743	0.7	0.013	617	870	putative beta-glucoside bgl operon transcription antiterminator
Cd630-1863 (1I20)	1.0	0.956	66	60	2.2	0.013	72	32	putative conjugative transposon conserved hypothetical protein
Cd630-2705 (7L6)	1.0	0.803	2463	2508	0.3	0.013	439	1467	putative amidohydrolas
Cd630-0740 (3H10)	1.0	0.857	3338	3394	0.5	0.013	963	1985	putative aminotransferase
Cd630-1239 (3N13)	1.1	0.454	2691	2464	0.8	0.013	1435	1720	putative regulatory protein
Cd630-3327 (9C14)	1.0	0.849	8824	8864	1.6	0.013	8453	5493	conjugative transposon conserved hypothetical protein
Cd630-0243 (4C7)	1.0	0.817	11133	11514	0.3	0.013	4437	12516	conserved hypothetical protein
Cd630-2001 (6L1)	0.9	0.513	743	819	1.5	0.012	798	517	conserved hypothetical protein
Cd630-3208 (9D11)	0.9	0.472	762	785	0.5	0.012	226	435	MarR-family transcriptional regulator
Cd630-2990 (8F19)	0.9	0.555	962	1080	0.4	0.012	235	640	ABC transporter, ATP-binding protein
Cd630-0499 (1D3)	1.0	0.991	1748	1694	1.5	0.012	1488	995	putative conjugative transposon replication initiation factor
Cd630-3332 (9H14)	1.0	0.897	6635	6635	0.8	0.012	3435	4530	TetR-family transcriptional regulator
Cd630-1171 (3A17)	1.0	0.876	885	940	1.5	0.012	849	562	electron transfer flavoprotein beta-subunit
Cd630-0158 (2F20)	1.0	0.825	900	887	0.3	0.012	199	645	putative membrane protein (pseudogene)
Cd630-1526 (5N1)	1.0	0.689	5730	5766	1.3	0.012	4778	3567	dihydroorotase
Cd630-2773 (8C3)	0.9	0.461	352	431	0.2	0.012	111	437	putative beta-glycosyltransferase
Cd630-0713 (3E7)	1.0	0.786	3894	3894	1.4	0.012	3653	2544	putative adenosylcobamide-dependent radical SAM superfamily protein
Cd630-3075 (8K18)	1.5	0.516	4074	1866	6.2	0.012	8901	1276	PTS system, IIabc component
Cd630-1307 (3J22)	1.0	0.979	1178	1162	1.3	0.012	976	782	transcription elongation protein
Cd630-1548 (5L4)	1.0	0.958	819	822	0.4	0.012	218	509	putative ATP phosphoribosyltransferase
Cd630-2534 (7F20)	1.1	0.535	978	857	0.5	0.011	293	563	ABC transporter ATP-binding protein
Cd630-1690 (5B22)	1.0	0.937	6184	6260	1.6	0.011	6062	3689	thioredoxin
Cd630-0687 (3C4)	1.1	0.494	2816	2652	0.6	0.011	1016	1705	50S ribosomal protein L20
Cd630-1185 (3G18)	1.1	0.378	2363	2198	1.7	0.011	2129	1164	putative signaling protein
Cd630-0434 (2J7)	1.1	0.621	442	387	1.7	0.011	333	203	putative drug/sodium antiporter
Cd630-1655 (5G17)	0.9	0.471	6799	7445	0.7	0.011	3756	5197	putative Na+/H+ antiporter
Cd630-2389 (7D2)	1.0	0.826	663	676	0.4	0.011	175	508	putative beta-lactamase inducer
Cd630-1212 (3B22)	1.0	0.513	1873	1890	0.5	0.011	727	1367	spermine/spermidine acetyltransferase
Cd630-1216 (3F22)	1.0	0.902	3451	3441	0.5	0.011	986	2091	putative exported protein
Cd630-0539 (1E8)	1.1	0.409	3802	3368	0.3	0.011	788	2130	chemotaxis protein
Cd630-0195 (4C1)	1.0	0.978	2056	2094	1.6	0.011	1853	1186	putative membrane protein
Cd630-2992 (8H19)	1.0	0.806	600	556	0.6	0.011	254	444	hypothetical protein
Cd630-1538 (5J3)	1.0	0.855	1683	1682	0.8	0.010	863	1017	putative signaling protein

Cd630-1483 (5C8)	1.2	0.574	192	140	0.7	0.010	44	67	putative aliphatic sulfonates ABC transporter, ATP-binding protein
Cd630-0153 (2A20)	0.9	0.732	53	56	2.7	0.010	86	30	4-hydroxyphenylacetate decarboxylase, catalytic subunit
Cd630-1024 (4L22)	1.0	0.444	839	787	2.1	0.010	950	425	spermidine/putrescine ABC transporter, ATP-binding protein
Cd630-2447 (7F9)	1.0	0.796	10176	10055	1.5	0.010	10331	6748	histidine triad nucleotide-binding protein
Cd630-0522 (1D6)	1.1	0.543	736	647	0.3	0.010	78	285	putative signaling protein
Cd630-3671 (10K9)	1.1	0.426	3599	3460	0.8	0.010	1622	2038	stage 0 sporulation protein J
Cd630-2127 (6E17)	1.0	0.803	21891	22040	0.5	0.010	7032	12999	putative membrane protein
Cd630-1282 (3I19)	1.0	0.980	25039	25453	1.7	0.010	26768	15173	putative alanyl-tRNA synthetase
Cd630-0735 (3C10)	1.1	0.393	7334	7040	1.5	0.010	5867	3943	hypothetical protein
Cd630-0473 (2J12)	1.0	0.735	1354	1340	0.2	0.010	232	919	MarR-family transcriptional regulator
Cd630-0917 (3I9)	1.1	0.709	35	30	3.3	0.010	76	21	putative phage recombination protein
Cd630-3461 (9I19)	0.9	0.407	4235	4667	1.2	0.010	3300	2874	putative regulator of cell growth
Cd630-3236 (9P2)	1.2	0.641	105	91	6.3	0.010	236	31	putative membrane protein
Cd630-0103 (2G13)	1.0	0.767	3528	3662	1.7	0.010	3703	2124	tRNA pseudouridine synthase 1
Cd630-2716 (7P7)	0.9	0.762	2152	2349	0.4	0.009	611	1495	hypothetical protein
Cd630-2674 (7M2)	1.0	0.699	4331	4006	1.6	0.009	3946	2623	oligopeptide ABC transporter, permease protein
Cd630-3283 (9O8)	1.0	0.916	3477	3654	0.7	0.009	1622	2296	putative pyruvate formate-lyase 3 activating enzyme
Cd630-0240 (4H6)	1.1	0.755	7222	7122	0.5	0.009	3344	5938	putative glycosyltransferase
Cd630-0732 (3H9)	1.0	0.500	1959	1864	1.6	0.009	1545	951	putative radical SAM superfamily protein
Cd630-3439 (9K16)	1.0	0.757	4489	4933	0.5	0.009	1345	2829	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase
Cd630-0028 (1L4)	1.0	0.864	4233	4152	0.5	0.009	1302	2674	putative DNA-binding protein
Cd630-3674 (10N9)	1.0	0.592	4690	4580	1.5	0.009	4206	2814	methyltransferase (putative glucose inhibited division protein B) (pseudogene)
Cd630-2649 (7L23)	1.0	0.684	2347	2471	0.5	0.009	671	1299	putative exported protein
Cd630-3097 (8I21)	1.0	0.883	2782	2786	1.2	0.009	1370	1130	PTS system, IIabc component
Cd630-1933 (6G4)	1.1	0.518	1805	1752	0.3	0.009	349	1170	putative membrane protein
Cd630-3299 (9O10)	1.0	0.759	1737	1757	1.2	0.009	1403	1173	probable transporter
Cd630-0262 (4F9)	1.0	0.988	12992	12994	0.7	0.009	5272	7703	flagellar export protein
Cd630-3416 (9L13)	1.1	0.794	1375	1401	2.0	0.009	1403	699	ABC transporter, permease protein
Cd630-0910 (3J8)	1.0	0.530	3070	3104	1.7	0.009	3520	2113	putative phage protein
Cd630-2181 (6C24)	0.9	0.537	6769	7523	0.5	0.009	3088	5375	putative aromatic compounds hydrolase
Cd630-0246 (4F7)	1.0	0.757	9969	9981	0.4	0.009	3249	7584	flagellar basal-body rod protein
Cd630-3258 (9N5)	0.9	0.425	1061	1225	0.8	0.009	566	727	iron-only hydrogenase
Cd630-3041 (8I14)	0.9	0.530	7521	8094	1.4	0.009	7237	5001	proline iminopeptidase

Cd630-3613 (10I2)	1.1	0.475	1361	1236	1.4	0.008	965	694	conserved hypothetical protein
Cd630-2149 (6C20)	1.0	0.724	5698	5971	1.7	0.008	7596	4301	putative exported protein
Cd630-3453 (9I18)	1.0	0.850	2175	2157	0.5	0.008	578	1277	N-acetylglucosamine-6-phosphate deacetylase
Cd630-2009 (6N2)	1.0	0.873	111	102	2.4	0.008	156	60	putative topoisomerase (partial)
Cd630-1693 (5E22)	1.0	0.519	8323	7952	3.9	0.008	17264	4117	putative dinitrogenase iron-molybdenum cofactor
Cd630-0472 (2I12)	0.9	0.528	2513	2630	0.4	0.008	586	1541	putative iron-sulfur flavoprotein
Cd630-1695 (5G22)	1.0	0.699	3963	4084	1.5	0.008	4399	2804	putative arsenical pump membrane protein
Cd630-2983 (8G18)	0.7	0.645	342	321	0.3	0.008	64	255	conserved hypothetical protein
Cd630-1587 (5K9)	1.0	0.853	111	101	2.2	0.008	107	50	putative ribose ABC transporter, ATP-binding protein
Cd630-2533 (7E20)	1.5	0.439	64	35	20.0	0.008	206	7	ABC transporter, permease protein
Cd630-0703 (3C6)	1.0	0.740	5420	5470	1.4	0.008	7633	5202	protease
Cd630-3087 (8O19)	1.1	0.665	1614	1525	1.7	0.008	1551	952	phosphosugar-binding transcriptional regulator
Cd630-3374 (9B20)	1.8	0.394	336	95	3.6	0.008	176	51	putative membrane protein
Cd630-2360 (6O22)	1.0	0.879	5896	5815	0.3	0.007	1475	5045	hypothetical protein
Cd630-2542 (7F21)	1.1	0.430	4347	4122	1.5	0.007	4517	2969	putative exported protein
Cd630-1404 (4K10)	1.0	0.774	17621	17915	1.4	0.007	14477	10199	putative oligopeptide transporter
Cd630-1284 (3K19)	1.0	0.754	4723	4638	0.8	0.007	2323	2831	putative membrane protein
Cd630-0252 (4D8)	0.9	0.591	20547	23724	0.2	0.007	3439	13614	flagellar protein
Cd630-0232 (4H5)	1.0	0.650	33064	34086	0.4	0.007	10159	25969	flagellar hook-associated protein
Cd630-1489 (5A9)	1.0	0.888	1541	1496	0.5	0.007	489	900	D-methionine ABC transporter, ATP-binding protein
Cd630-1485 (5E8)	1.0	0.879	805	783	3.6	0.007	1912	501	conserved hypothetical protein
Cd630-2024 (6N4)	1.0	0.810	1912	1914	1.4	0.007	1529	1025	ABC transporter, ATP-binding protein
Cd630-2416 (7G5)	1.2	0.558	13	9	34.4	0.007	526	8	putative transcription antiterminator
Cd630-3547 (10G5)	0.9	0.542	301	292	0.4	0.007	53	143	conserved hypothetical protein
Cd630-1172 (3B17)	1.0	0.995	1103	1099	1.5	0.007	1092	724	electron transfer flavoprotein alpha-subunit
Cd630-0245 (4E7)	1.0	0.824	11530	12364	0.3	0.007	2367	7627	flagellar basal-body rod protein
Cd630-1844 (1N17)	1.0	no replicates	86	86	1.4	0.007	103	73	putative cell surface protein (pseudogene)
Cd630-2789b (8G5)	1.0	0.694	9279	9138	0.7	0.007	4194	5821	cell surface protein
Cd630-0211 (4C3)	1.0	0.825	2082	2080	0.2	0.007	301	1333	CTP:phosphocholine cytidylyltransferase
Cd630-0438 (2N7)	0.9	0.442	1155	1291	0.4	0.007	431	1071	putative exported protein
Cd630-1151 (3E14)	1.0	0.802	4677	4637	1.3	0.007	4087	3016	cell division topological specificity factor
Cd630-3076 (8L18)	1.5	0.527	3325	1462	5.6	0.007	4547	764	putative tagatose (1)-phosphate-kinase
Cd630-1389 (4L8)	0.9	0.730	491	512	0.4	0.007	136	327	putative chloromuconate cycloisomerase

Cd630-1828 (1N15)	1.0	0.842	4470	4634	1.9	0.007	5506	2749	cell division protein
Cd630-1275 (3J18)	1.0	0.901	2751	2733	1.4	0.007	2307	1654	GTP-sensing transcriptional pleiotropic repressor
Cd630-3467 (9O19)	1.0	0.992	4735	4731	1.5	0.007	4544	3130	ATP synthase epsilon chain (partial)
Cd630-3098 (8J21)	0.9	0.518	1420	1686	0.7	0.006	746	1137	transcription antiterminator
Cd630-2471 (7F12)	1.1	0.469	894	796	2.6	0.006	1010	410	conserved hypothetical protein
Cd630-1417 (4P11)	1.0	0.941	1510	1498	0.4	0.006	475	1009	putative ATP-binding protein
Cd630-2362 (6I23)	1.0	0.862	9160	9403	0.5	0.006	4187	8909	putative aliphatic sulfonates ABC transporter, permease protein (pseudogene)
Cd630-1999 (6J1)	1.0	0.915	736	732	2.1	0.006	996	462	flavodoxin
Cd630-0227 (4C5)	1.0	0.628	6072	5921	0.4	0.006	2254	5241	conserved hypothetical protein
Cd630-2737 (7N10)	1.1	0.594	980	811	0.6	0.006	309	530	putative carbon-nitrogen hydrolase
Cd630-1261 (3L16)	1.1	0.407	9985	9546	1.6	0.006	8662	5229	putative ribonucleotide-diphosphate reductase
Cd630-3376 (9D20)	1.0	0.982	2713	2499	2.4	0.006	2798	1165	putative transcriptional regulator
Cd630-0184 (2H23)	1.2	0.446	690	522	0.5	0.006	152	334	aspartate carbamoyltransferase catalytic chain
Cd630-0985 (4M17)	1.0	0.818	2551	2563	0.7	0.006	1110	1553	putative drug/sodium antiporter
Cd630-3010 (8B22)	0.9	0.607	63	66	2.5	0.006	65	25	putative cytochrome C biogenesis protein
Cd630-1304 (3O21)	1.1	0.526	3430	3236	1.8	0.006	3052	1663	putative mannosyl-glycoprotein endo-beta-N-acetylglucosamidase
Cd630-0607 (2O16)	1.0	0.910	2227	2229	1.6	0.006	2021	1288	putative transposase-like protein B
Cd630-0358 (2F9)	1.2	0.382	373	308	1.5	0.005	271	188	conjugative transposon conserved hypothetical protein
Cd630-0878 (3J4)	1.0	0.976	23143	23199	0.6	0.005	11645	20522	ABC transporter, permease protein
Cd630-1630 (5F14)	0.9	0.569	483	512	0.3	0.005	147	507	putative membrane-associated metal-dependent hydrolase
Cd630-3157 (9H4)	1.0	0.877	3697	3861	1.8	0.005	4122	2285	SsrA-binding protein
Cd630-2358 (6M22)	1.0	0.813	538	496	0.6	0.005	172	283	putative oligoendopeptidase
Cd630-3091 (8K20)	1.0	0.749	17477	18059	0.4	0.005	5694	14087	trehalose-6-phosphate hydrolase
Cd630-1785 (5J22)	1.0	0.664	3560	3451	0.7	0.005	1789	2636	argininosuccinate synthase
Cd630-1943 (6A6)	1.0	0.606	1785	1832	1.4	0.005	2464	1723	conserved hypothetical protein
Cd630-2374 (6M24)	0.8	0.411	640	645	0.5	0.005	216	392	hypothetical protein
Cd630-3428 (9P14)	1.1	0.421	2563	2397	0.2	0.005	306	1471	probable cobalt-precorrin-6y C(5)-methyltransferase
Cd630-3103 (8O21)	1.0	0.495	1388	1333	1.4	0.005	1129	780	putative PTS syste, IIa component
Cd630-3430 (9J15)	1.0	0.795	1104	1139	0.2	0.005	243	894	cobalt-precorrin-8x methylmutase
Cd630-3426 (9N14)	1.1	0.680	1344	1239	0.2	0.005	174	793	precorrin-4 C(11)-methyltransferase
Cd630-1397 (4L9)	1.0	0.896	513	516	0.4	0.005	125	289	conserved hypothetical protein
Cd630-2543 (7G21)	1.0	0.946	1897	1973	0.6	0.005	917	1475	putative polysaccharide O-acetylase
Cd630-2492 (7C15)	0.9	0.573	526	592	0.2	0.005	91	383	putative sensor histidine kinase

Cd630-2666 (7M1)	1.0	0.807	9675	9865	1.2	0.005	7535	6181	PTS system, glucose-specific IIa component
Cd630-1170 (3H16)	0.9	0.582	535	577	2.4	0.005	773	322	conserved hypothetical protein
Cd630-2531 (7C20)	1.1	0.374	5066	4651	0.6	0.005	2241	3410	putative membrane protein
Cd630-3438 (9J16)	1.0	0.790	1647	1797	0.2	0.005	296	1151	bifunctional adenosylcobalamin biosynthesis protein [includes: adenosylcobinamide kinase and adenosylcobinamide-phosphate guanylyltransferase]
Cd630-0256 (4H8)	1.1	0.627	30258	29247	0.5	0.005	9096	17545	chemotaxis protein
Cd630-2234 (8I7)	1.0	0.931	177	178	0.4	0.005	40	106	cAMP-binding regulatory protein
Cd630-0840 (1D23)	1.0	0.918	474	483	0.4	0.005	138	330	putative hydrolase
Cd630-1594 (5J10)	1.0	0.964	914	897	1.5	0.004	742	483	putative O-acetylserine sulfhydrylase
Cd630-2262 (8M10)	0.9	0.883	750	711	1.6	0.004	653	446	probable proton-dependent oligopeptide transporter
Cd630-0606 (2N16)	0.9	0.465	2754	3269	1.6	0.004	3045	1931	chloride ion channel protein
Cd630-0393 (2I2)	1.1	0.437	3579	3337	0.6	0.004	1406	2459	putative membrane protein
Cd630-2012 (6J3)	1.0	0.908	2287	2251	0.7	0.004	1019	1465	putative hydrolase
Cd630-2138 (6H18)	0.9	0.557	1078	1164	0.4	0.004	350	880	uridylate kinase
Cd630-2750 (7K12)	1.1	0.412	1021	912	0.5	0.004	361	760	accessory gene regulator
Cd630-2371 (6J24)	1.0	0.674	1868	1860	1.3	0.004	1273	979	L-aspartate oxidase
Cd630-2444 (7C9)	1.0	0.859	8400	8787	3.2	0.004	17033	5041	putative sigma 54 modulation protein
Cd630-2334 (6M19)	1.1	0.469	1507	1373	1.4	0.004	1165	843	PTS system, mannitol-specific IIbc component
Cd630-2055 (6M8)	1.0	0.898	690	669	1.8	0.004	893	477	hypothetical protein
Cd630-2179 (6A24)	1.0	0.944	1937	2003	0.4	0.004	439	1106	putative oxidoreductase, molybdopterin-binding subunit
Cd630-1279 (3N18)	1.0	0.834	7072	6952	0.7	0.004	2991	4356	cysteine desulfurase
Cd630-3284 (9P8)	1.0	0.905	5723	5991	2.3	0.004	7278	2978	probable protease
Cd630-0654 (2N22)	0.9	0.425	343	368	0.4	0.004	116	293	putative ABC transporter, permease protein
Cd630-1627 (5C14)	1.0	0.809	335	328	0.4	0.004	93	194	putative D-alanyl-D-alanine carboxypeptidase
Cd630-0059 (1K8)	1.0	0.854	7358	7811	0.7	0.004	3762	5684	preprotein translocase SecE subunit
Cd630-1963 (6E8)	1.1	0.439	517	469	2.2	0.004	746	335	conserved hypothetical protein
Cd630-2819 (8D10)	1.0	0.906	3058	3034	1.9	0.004	3416	1694	putative amino acid racemase
Cd630-0541 (1G8)	1.0	0.810	1103	1145	0.2	0.004	131	677	chemotaxis protein methyltransferase
Cd630-0540 (1F8)	1.0	0.960	555	538	0.2	0.004	84	330	chemotaxis protein
Cd630-2757 (8B1)	1.0	0.512	4020	4159	0.8	0.004	2147	2708	putative phospholipase
Cd630-2641 (7L22)	1.0	0.889	963	938	0.3	0.004	178	580	conserved hypothetical protein
Cd630-2815 (8H9)	1.0	0.907	1246	1217	1.7	0.004	1200	674	hypothetical protein
Cd630-1778 (5K21)	1.0	0.538	1772	1739	0.6	0.004	622	1073	conserved hypothetical protein

Cd630-2455 (7F10)	0.9	0.596	838	901	0.3	0.004	105	415	conserved hypothetical protein
Cd630-1614 (5N12)	1.0	0.997	2554	2544	0.4	0.003	636	1330	putative sodium extrusion ABC transporter, ATP-binding protein
Cd630-2039 (6M6)	0.9	0.629	306	307	1.6	0.003	288	183	conserved hypothetical protein
Cd630-0688 (3D4)	1.0	0.786	2616	2512	0.2	0.003	445	1903	conserved hypothetical protein
Cd630-0769 (1E14)	1.0	0.852	545	546	0.5	0.003	204	429	putative oxidoreductase
Cd630-3254 (9J5)	1.0	0.621	3052	3208	0.4	0.003	922	2233	putative serine protease
Cd630-3561 (10E7)	0.9	0.775	538	528	0.5	0.003	227	507	conserved hypothetical protein
Cd630-0356 (2D9)	1.0	0.903	1422	1447	1.5	0.003	1404	964	excisionase
Cd630-0242 (4B7)	1.0	0.832	10757	11346	0.5	0.003	5487	10365	conserved hypothetical protein
Cd630-0881 (3M4)	1.0	0.981	9734	9544	2.7	0.003	13586	4949	putative membrane protein
Cd630-0866 (3N2)	1.0	0.632	1890	1927	1.4	0.003	1582	1132	putative exported protein
Cd630-2354 (6I22)	1.1	0.817	615	586	0.6	0.003	284	468	glycine reductase complex component B alpha and beta subunits
Cd630-1747 (5K17)	1.0	0.901	2861	2904	0.7	0.003	1421	1881	adenosine deaminase
Cd630-1686 (5F21)	1.0	0.881	724	758	0.3	0.003	112	383	putative DNA-binding protein
Cd630-0533 (1G7)	1.0	0.869	874	891	0.1	0.003	72	495	chemotaxis protein
Cd630-3074 (8J18)	1.4	0.457	4966	2816	8.0	0.003	15206	1799	tagatose-1,6-bisphosphate aldolase
Cd630-0893 (3I6)	1.0	0.869	10153	10419	0.4	0.003	3499	8663	iron-dependent hydrogenase
Cd630-0888 (3L5)	1.1	0.408	14127	12709	0.3	0.003	2909	9132	arginine decarboxylase
Cd630-0534 (1H7)	0.9	0.699	1072	1155	0.3	0.003	166	608	chemotaxis protein CheC
Cd630-0229 (4E5)	1.0	0.654	5536	5238	0.5	0.003	1814	3361	negative regulator of flagellin synthesis (anti-sigma-d factor)
Cd630-2745 (7N11)	0.8	0.421	558	644	0.6	0.003	251	434	adenine phosphoribosyltransferase
Cd630-3422 (9J14)	0.9	0.512	1189	1288	0.2	0.003	159	826	sirohydrochlorin cobaltochelatase
Cd630-1222 (3D23)	1.0	0.833	5455	5658	0.5	0.003	1626	3310	putative tyrosine recombinase
Cd630-2388 (7C2)	1.0	0.776	633	633	0.3	0.003	117	380	putative peptidoglycan- binding/hydrolysing protein
Cd630-2309 (6L16)	1.1	0.754	342	327	2.3	0.002	373	169	hypothetical protein
Cd630-1469b (5E6)	1.0	0.374	8845	8389	0.5	0.002	3002	5863	putative penicillin-binding protein
Cd630-2063 (6M9)	1.0	0.665	5108	5187	0.6	0.002	2090	3322	putative membrane protein
Cd630-0223 (4G4)	1.0	0.814	5855	6102	0.5	0.002	1835	3457	bifunctional purine biosynthesis protein [includes: phosphoribosylaminoimidazolecarboxamide formyltransferase and IMP cyclohydrolase]
Cd630-1190 (3D19)	1.0	0.726	3641	3460	0.5	0.002	994	2110	putative acetyltransferase
Cd630-0545 (1C9)	1.0	0.623	3958	3840	0.5	0.002	1356	2590	putative lipoprotein
Cd630-3609 (10M1)	1.1	0.467	566	505	0.4	0.002	152	385	hypothetical protein
Cd630-0339 (2C7)	1.0	0.916	515	507	0.3	0.002	111	390	two-component response regulator

Cd630-3619 (10O2)	1.0	0.533	1976	1912	0.6	0.002	801	1360	putative acetyltransferase
Cd630-2861 (10F15)	1.0	0.939	74	68	4.6	0.002	100	22	putative membrane protein
Cd630-0678 (3B3)	1.0	0.873	8635	8832	1.3	0.002	7022	5128	putative transcriptional regulator
Cd630-2786a (8A5)	1.1	0.518	436	410	1.9	0.002	427	227	putative cell surface protein
Cd630-0263 (4G9)	1.0	0.691	14384	14419	0.4	0.002	4269	11689	flagellar export protein
Cd630-2014 (6L3)	1.0	0.972	1677	1574	0.6	0.002	554	959	dihydroxy-acid dehydratase
Cd630-3353 (9E17)	1.1	0.562	734	614	0.5	0.002	187	372	GntR-family transcriptional regulator
Cd630-0853 (3I1)	1.0	0.867	11760	11947	0.6	0.002	4653	7756	oligopeptide ABC transporter, permease protein
Cd630-1027 (4O22)	1.1	0.538	1062	919	2.4	0.002	1394	541	spermidine/putrescine ABC transporter, substrate-binding lipoprotein
Cd630-0558 (1H10)	1.0	0.800	4946	4867	0.5	0.002	1719	3095	probable glutaminase
Cd630-2524 (7D19)	1.0	0.663	4078	3958	0.4	0.002	774	2067	nicotinate-nucleotide adenylyltransferase
Cd630-1793 (5J23)	1.1	0.503	1066	900	2.9	0.002	1634	578	putative membrane protein
Cd630-3614 (10J2)	1.1	0.503	2344	2218	1.7	0.002	2456	1402	conserved hypothetical protein
Cd630-3513 (10E1)	1.1	0.405	327	286	2.3	0.002	331	144	pilin
Cd630-0324 (2D5)	1.0	0.873	29696	30198	0.5	0.002	12590	26074	putative cobalt transport protein
Cd630-3427 (9O14)	1.1	0.553	3601	3251	0.4	0.002	912	2343	probable cobalt-precorrin-6y C(15)-methyltransferase
Cd630-3100 (8L21)	1.0	0.703	17939	17919	0.4	0.002	4120	10919	putative C4-dicarboxylate anaerobic carrier
Cd630-2330 (6I19)	1.0	0.687	5338	5496	0.3	0.002	1255	3942	xanthine phosphoribosyltransferase
Cd630-0027 (1K4)	1.0	0.918	3694	3789	0.7	0.002	1589	2297	DNA repair protein
Cd630-3145 (9A3)	1.0	0.968	15569	15226	0.3	0.002	2394	9030	putative serine-aspartate-rich surface anchored fibrinogen-binding protein
Cd630-0591 (2I15)	1.0	0.823	2019	1957	0.3	0.002	353	1139	putative cation transporting ATPase
Cd630-2760 (8E1)	0.9	0.442	2285	2370	0.5	0.002	864	1588	putative signaling protein
Cd630-1221 (3C23)	1.0	0.755	2046	2085	0.3	0.001	332	967	putative membrane protein
Cd630-1478 (5F7)	0.9	0.524	1080	1345	3.7	0.001	2629	730	putative ferrous iron transport protein A
Cd630-1237 (3L13)	1.0	0.731	3956	4051	0.5	0.001	1268	2513	hypothetical protein (fragment)
Cd630-1446 (5E3)	1.0	0.709	2162	2341	0.3	0.001	341	1305	para-aminobenzoate synthase component I
Cd630-2761 (8F1)	0.9	0.467	3887	4239	0.5	0.001	1508	2813	putative N-acetylmuramoyl-L-alanine amidase
Cd630-2510 (7E17)	1.0	0.807	3203	3228	0.6	0.001	1074	1838	PTS system, IIbc component
Cd630-2065 (6O9)	1.0	0.943	2264	2181	2.2	0.001	2857	1300	putative regulatory protein
Cd630-1161 (3G15)	1.1	0.630	5202	4985	1.6	0.001	7379	4365	50S ribosomal protein L21
Cd630-1779 (5L21)	1.0	0.753	3384	3277	0.5	0.001	1178	2204	conserved hypothetical protein
Cd630-2565 (7E24)	1.1	0.636	785	700	1.5	0.001	780	509	putative transcriptional antiterminator

Cd630-3143 (9G2)	1.0	0.991	691	690	0.7	0.001	287	412	conserved hypothetical protein
Cd630-2796b (8C7)	0.9	0.454	7224	7918	0.4	0.001	2264	5178	putative cell surface protein
Cd630-0117 (2E15)	1.1	0.429	8069	7542	2.2	0.001	11560	5416	putative subunit of oxidoreductase
Cd630-1800 (5I24)	0.9	0.513	2968	3389	0.4	0.001	1092	2422	tellurium resistance protein
Cd630-2229 (8L6)	0.9	0.436	125	144	4.5	0.001	349	79	putative membrane protein
Cd630-1576 (5P7)	0.9	0.401	3485	3629	0.5	0.001	957	1931	putative arylesterase
Cd630-3668 (10P8)	0.9	0.639	7546	8348	0.7	0.001	3277	4932	putative transcriptional regulator
Cd630-0268 (4D10)	1.0	0.788	16216	15812	0.4	0.001	3980	9326	flagellar basal-body rod protein FlgG
Cd630-2647 (7J23)	1.0	0.625	1823	1949	0.4	0.001	619	1655	putative membrane protein
Cd630-0887 (3K5)	1.0	0.510	4971	5187	0.5	0.001	1818	3606	GntR-family transcriptional regulator
Cd630-3341 (9A16)	1.0	0.870	1089	1020	1.5	0.001	898	606	conjugative transposon conserved hypothetical protein (pseudogene)
Cd630-3642 (10N5)	1.0	0.902	2092	2111	0.6	0.001	763	1319	putative selenocysteine synthase
Cd630-2171 (6A23)	1.0	0.901	8050	7970	0.5	0.001	2583	5235	putative Sodium:dicarboxylate symporter
Cd630-2767b (8E2)	1.0	0.635	6104	5799	0.4	0.001	1756	3911	putative cell surface protein
Cd630-0707 (3G6)	1.0	0.734	2727	2698	0.6	0.001	1005	1721	putative signaling protein
Cd630-0730 (3F9)	1.0	0.999	10558	10597	2.3	0.001	15588	6686	putative iron-sulfur protein
Cd630-0728 (3D9)	1.0	0.581	4676	4457	0.6	0.001	1507	2572	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit
Cd630-3077 (8M18)	1.3	0.498	1073	717	2.9	0.001	1076	379	phosphosugar-binding transcriptional regulator
Cd630-0269 (4E10)	1.0	0.583	13488	13837	0.4	0.001	4255	9108	putative flagellar basal-body rod protein
Cd630-2270 (8M11)	1.2	0.476	234	184	3.7	0.001	363	102	putative 1-phosphofructokinase
Cd630-3116 (8L23)	1.0	0.806	3635	3745	0.7	0.001	1397	2114	PTS system, beta-glucoside-specific IIabc component
Cd630-0226 (4B5)	1.0	0.956	15207	15050	0.5	0.000	5738	11783	putative transglycosylase
Cd630-0282 (4B12)	1.0	0.854	2121	2073	2.0	0.000	2134	1041	putative phosphoesterase
Cd630-3673 (10M9)	1.0	0.726	5057	4824	0.5	0.000	1824	3591	putative stage 0 sporulation protein
Cd630-3241 (9M3)	1.3	0.722	78	30	3.9	0.000	78	20	proline reductase
Cd630-2365 (6L23)	1.0	0.974	25340	25640	0.6	0.000	10829	18992	putative sulfonate ABC transporter, solute-binding lipoprotein
Cd630-1997 (6H12)	1.1	0.430	1722	1665	2.2	0.000	1901	895	putative proline iminopeptidase
Cd630-3625 (10M3)	0.9	0.498	2994	3102	0.4	0.000	716	1874	putative signaling protein
Cd630-0267 (4C10)	1.0	0.945	4523	4643	0.4	0.000	1388	3244	putative exported protein
Cd630-1692 (5D22)	1.0	0.963	4832	4865	3.0	0.000	9147	3047	ArsR-family transcriptional regulator
Cd630-3414 (9J13)	1.4	0.445	445	302	1.7	0.000	363	210	ABC transporter, solute-binding protein
Cd630-0341 (2E7)	1.0	0.992	2096	2073	0.4	0.000	552	1444	conserved hypothetical protein
Cd630-1157 (3C15)	1.0	0.628	10506	10187	1.6	0.000	13531	8135	putative anaerobic nitric oxide reductase flavorubredoxin

Cd630-0261 (4E9)	1.0	0.882	3881	3868	0.5	0.000	1595	2945	flagellar export protein
Cd630-0194 (4B1)	1.1	0.624	12837	12269	1.9	0.000	14495	7570	60 kDa chaperonin
Cd630-0241 (4A7)	0.9	0.719	11324	12564	0.4	0.000	4971	12370	phosphoserine phosphatase
Cd630-1590 (5N9)	1.0	0.979	4060	4004	2.0	0.000	4516	2296	putative membrane protein
Cd630-0231 (4G5)	1.0	0.940	24468	24284	0.2	0.000	4941	20098	putative flagellar hook-associated protein
Cd630-1953 (6C7)	1.0	0.995	7079	7050	0.2	0.000	977	5607	conserved hypothetical protein
Cd630-0796 (1H17)	1.0	0.400	2787	2923	1.2	0.000	1974	1686	LacI-family transcriptional regulator
Cd630-3272 (9L7)	1.0	0.995	4647	4724	0.6	0.000	1946	3159	putative membrane protein
Cd630-0247 (4G7)	1.0	0.980	10454	10897	0.4	0.000	2853	7380	flagellar hook-basal body complex protein
Cd630-2062 (6L9)	1.0	0.685	6790	6524	0.6	0.000	2643	4682	conserved hypothetical protein
Cd630-2329 (6P18)	1.1	0.775	6248	6117	0.4	0.000	1564	3819	putative transaldolase
Cd630-0582 (2P13)	1.0	0.849	5333	5031	2.0	0.000	6173	3108	putative PEP-utilising kinase
Cd630-0559 (1A11)	0.9	0.474	5118	5407	0.4	0.000	1385	3842	conserved hypothetical protein
Cd630-0549 (1G9)	1.0	0.780	2505	2565	2.0	0.000	2861	1423	putative exported protein
Cd630-1025 (4M22)	0.9	0.512	1230	1320	1.8	0.000	1653	923	spermidine/putrescine ABC transporter, permease protein
Cd630-0617 (2I18)	1.0	0.958	881	882	0.3	0.000	145	560	putative membrane protein
Cd630-0538 (1D8)	0.9	0.436	2849	3169	0.1	0.000	251	1686	putative methyl accpeting chemotaxis protein
Cd630-2366 (6M23)	1.0	0.828	18874	19989	0.2	0.000	3379	14184	putative exported protein
Cd630-2367 (6N23)	1.0	0.784	10077	10343	0.5	0.000	4273	9224	putative permease
Cd630-0265 (4A10)	1.0	0.941	8875	8593	0.3	0.000	1334	5133	flagellar number regulator
Cd630-0855 (3K1)	1.0	0.983	13417	13256	0.4	0.000	3598	9622	oligopeptide ABC transporter, substrate-binding lipoprotein
Cd630-1205 (3C21)	1.0	0.508	5206	5421	0.5	0.000	1811	3416	geranyltranstransferase
Cd630-1479 (5G7)	0.8	0.423	3730	4606	0.4	0.000	1323	3256	ferrous iron transport protein B
Cd630-1477 (5E7)	1.1	0.671	514	444	6.9	0.000	1959	273	putative ferrous iron transport protein
Cd630-2806 (8G8)	1.0	0.799	1594	1587	0.6	0.000	675	1044	holliday junction DNA helicase
Cd630-3515 (10G1)	0.9	0.474	1655	1745	0.4	0.000	454	1164	bifunctional protein [includes: UDP-N- acetylglucosamine pyrophosphorylase and glucosamine-1-phosphate N-acetyltransferase]
Cd630-0854 (3J1)	1.0	0.652	8743	8567	0.4	0.000	1985	5427	oligopeptide ABC transporter, permease protein
Cd630-0535 (1A8)	0.8	0.618	839	782	0.2	0.000	93	601	chemotaxis protein
Cd630-0264 (4H9)	1.0	0.640	18430	18550	0.2	0.000	3007	11798	signal recognition particle complex, GTP- binding subunit
Cd630-2003 (6N1)	1.0	0.582	1851	1811	2.4	0.000	2785	1152	putative efflux pump
Cd630-2768 (8F2)	1.0	0.743	9711	9858	0.3	0.000	2421	8100	putative cell-wall hydrolase

9.3. Codon usage in C. difficile

The codon usage of *C. difficile* was compared to that of the *transposase* gene from pHV1248 using the Graphical Codon Usage Analyzer version 2.0 (http://gcua.schoedl.de/seqoverall_v2.html) based on information from www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=1496&aa=1&style=N and the plasmid sequence of pHV1248. The results are shown graphically with the percentage of codons representing each amino acid from the *transposase* gene shown in red and the *C. difficile* reference in black. Where the usage in the *transposase* gene is significantly higher this highlights a possible shortage of the cognate tRNAs in *C. difficile*.