Differential Expression of *Clostridium difficile* and phage genes during a one-step growth curve

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Thekra sideeq Theanon Al-Tayawi

Department of Infection, Immunity and Inflammation

College of Medicine, Biological Sciences and Psychology

University of Leicester

University Road, Leicester, LE1 7RH

March 2019

Differential Expression of *Clostridium difficile* and CDHS1 phage genes during a one-step growth curve

Thekra sideeq Theanon Al-Tayawi

Abstract

Clostridium difficile is one of the most dangerous pathogens that causes significant human morbidity and mortality in hospital settings. Bacteriophages are viruses of bacteria that are ubiquitous in nature. Most strains of *C. difficile* carry one or more prophages that can support the virulence of the bacteria. However, during the lytic cycle, temperate phages can affect virulence genes expression. In order to investigate the impact of phage on the infection potential of *C. difficile*, high-throughput RNA sequencing reactions (RNA-seq) were conducted on a genome-wide transcriptomic analysis of the epidemic *C. difficile* strain R20291. The finding from RNA-seq analysis shows that at different growth time points the expression of a majority of strain R20291 genes changed for R20291 infected with phage CDHS1 compared with uninfected R20291. Those R20291 genes involved in core DNA binding and transcription were among the genes most highly expressed. Genes encoding metabolic processes and permease were among the very low expression levels group (i.e. were down-regulated). Both RNA-seq and qPCR techniques showed significant down regulation in genes coding toxin A, B and the binary toxin in *C. difficile* infected with phage CDHS1. *Galleria mellonella* larvae were used as an infection model to analyse phage effects on *C. difficile* virulence. This showed that larvae infected with lysogenic *C. difficile* containing φCDHS1 had reduced survival whilst less virulent bacteria which were resistant to φCDHS1 and wild type *C. difficile,* were less infective. In addition, the more pathogenic strains (lysogenic) produced greater immune responses in *G. mellonella* than the relatively less-pathogenic forms (resistant to φCDHS1 and wild type). This study suggests that temperate phage can mimic the virulence enhancing aspect for lytic phage.

Acknowledgements & & dedications

The first to deserve all praise and thank giving is Allah, most merciful and most benevolent.

First of all, I would like to express my sincere gratitude to my supervisor Dr. Shaun Heaphy and Professor Martha Clokie for the constant support of my PhD study, for his patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis.

My special thanks goes to Dr. Janet Nale for her valuable supervision at the start of this work and for her kind donation of strains from the bank anaerobic reference laboratory strain bank in the lab. I would also like to express my sincere thanks to Dr. Primrose Freestone for his help with Scientific and linguistic proofreading. I am also grateful to each member of the Genetics Department, Leicester University. In addition, I am grateful for everyone in Leicester University for supporting me during my PhD journey.

I would also like to take the opportunity to thank all my colleagues in Lab 230 and all members of the Department of Infection, Immunity and Inflammation Department, Leicester University for their help during my PhD project. Also, I do not want to forget to thank the university of Leicester for giving me an opportunity to complete the doctoral study.

My sincere thanks also goes to my sponsor The Iraqi Ministry of Higher Education and Scientific Research (MOHESR) for funding my scholarship. I also acknowledge the role of the Iraqi cultural attache and University of Mosul for generous support.

Last but not least, I could not have completed my PhD journey without love, patience, and understanding from my amazing family, my brothers and my sisters. And the last person I dedicate this thesis to is the light of my eyes, the pure spirit of my father and mother.

Thekra Altayawi

Contents

List of tables

List of figures

List of Abbreviations

General Abbreviation	
φ	Phage symbol
ANOVA	Analysis of variance
BA	Bacteriological agar
BHI	Brain heart infusion
BLAST	Basic local alignment search tool
$CaCl_2$	Calcium Chloride
CCEY	Cycloserine, cefoxitin and egg yolk agar (Brazier's selective medium)s
CDAD	*Clostridium difficile* associated diarrhoea
CDI	*Clostridium difficile* infection
cDNA	Complementary Deoxyribonucleic acid
CDs	Coding DNA sequences
CFU/ml	Colony-forming unit per millilitre
CO_2	Carbon dioxide
DHB	Defribinated horse blood

DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylene diaminetetra acetic acid
FAB	Fastidious anaerobic broth
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GC	Guanine-cytosine
GTP	Guanosine-5'-triphosphate
H_2	Hydrogen
mRNA	Messenger Ribonucleic acid
$MgCl_2$	Magnesium chloride
MOI	Multiplicity of infection (the ratio of adsorbed phage particles to bacteria)
MW	Molecular weight
N_2	Nitrogen
NCBI	National centre for biotechnology information
NGS	Next generation sequencing
NTC	No template control
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFU/ml	Plaque-forming unit per millilitre
PMC	Pseudomembranous colitis
RAST	Rapid Annotation using Subsystem Technology server
RNA	Ribonucleic acid
RPM	Revolutions per minute

RT	Ribotype
SDS	Sodium dodecyl sulphate
TAE	Tris-acetate-EDTA
TBE	Tris-Borate-EDTA
TE	Tris-EDTA
ToxA	Toxin A
Tox B	Toxin B
TE	Tris-EDTA

List of units	
Bp	Basepairs
Cm	Centimeter
Kbp	Kilo basepairs
M	Molar
Mg	Milligram
Ml	Millilitre
Mm	Milli molar
Ml	Microliter
Mg	Microgram
Nm	Nanometer
U	Units
°C	Degree centigrade

Chapter 1 Introduction

Chapter 1 General Introduction

1.1 *Clostridium difficile* (*Peptoclostridium difficile*)

Clostridium difficile (*C. difficile*) is a Gram positive, obligate anaerobic spore-forming bacilli, which is motile, which can exist as both toxigenic and non-toxigenic forms (Lamont, 2002; Collins *et al.*, 1994). *C. difficile* was first isolated from human stool by Hall and O`Toole at Colombia University in 1935 as *Bacillus difficilis*, a component of the normal intestinal flora of newborn infants (Hall and O'Toole, 1935; Bartlett, 2008). The *difficile* name of this species of bacteria was used to reflect the difficulty with its culture and isolation (Kelly and LaMont, 2008). The microorganism is present in approximately 4–6 % of non-hospitalised healthy adults as part of their normal microbiota (Miyajima *et al.*, 2011; Pepin *et al.*, 2004; d'Herelle, 2007). While the rate of asymptomatic carriage in hospitalised adults and for those in long-term care facilities is estimated to be 20–50 %, carriage rates have been reported in healthy newborns of up to 70 % (Riggs *et al.*, 2007; McFarland *et al.*, 1989; Jangi and Lamont, 2010). Even though *C. difficile* is found as a part of the intestinal flora of healthy infants, the microorganism can be found in environments such as soil, water, and even meat products (Simango, 2006; Hensgens *et al.*, 2012).

The normal gut flora and host immune responses work together to prevent *C. difficile* spore germination, spore proliferation, and toxin production. But exposure to antibiotics can result in development of active *C. difficile* disease due to disruption of the gut microbiome and host immune responses(Speert, 2000; Taslim and Taslim, 2009). *C. difficile* was considered as a direct pathogen only in the past three decades but has since become one of the most common causes of nosocomial infections associated with antibiotic therapy (Bartlett *et al.*, 1978; Viswanathan *et al.*, 2010; George *et al.*, 1978). *C. difficile* causes a range of illness from mild diarrhea, to *C. difficile*- associated diarrhea, CDAD) and life-threatening pseudomembranous colitis (PMC) and/or toxic megacolon with perforations of the colon (Janka and O'Grady, 2009; Kim *et al.*, 2008). Although *C. difficile* infection (CDI) can cause a serious disease, the majority of infected persons present with no symptoms (Donskey *et al.*, 2015). In general, typical features of CDI include watery diarrhoea, abdominal pain and cramps, lower-quadrant tenderness, fever, leukocytosis, and hypoalbuminaemia (McGowan *et al.*, 2011).

Recent reports have showed that CDI usually occurs due to the disruption of the normal gut microbiota in the bowel, as a result of the use of broad spectrum antibiotics such as ampicillin, amoxicillin, cephalosporins, fluoroquinolones and clindamycin, or other therapeutics such as cytotoxic anticancer drugs (Leffler and Lamont, 2015; Smits *et al.*, 2016; Kwok *et al.*, 2012; Klingler *et al.*, 2000). In addition to CDI occurring in humans, several studies have confirmed CDI in non-human species including pigs, horses, primates, rabbits, rats, dogs and cats (Norman *et al.*, 2009; Hopman *et al.*, 2011; Clooten *et al.*, 2008; Theriot *et al.*, 2014).

Although all *C. difficile* strains are considered potentially pathogenic, epidemiological and genetic studies has shown that the infection is more frequently associated with hyper-virulent strains such as *C. difficile* PCR-ribotype 027 (Stabler *et al.*, 2009a). In addition to the severe effects on human life, *C. difficile* can cause significant financial losses in hospitals because of the need for additional prescription medication and extra intensive care (Wiegand *et al.*, 2012; Eggertson, 2004; Kyne *et al.*, 2002). For this reason, novel and more effective strategies such the use of bacteriophages have been investigated for treatment and prevention CDI. In order to advance such applications, it is necessary to identify the essential genetic information of *C. difficile* and its related phages.

1.1.1 *C. difficile* pathogenesis

Healthy adults host are protected from *C. difficile* colonisation and infection by a mechanism known as colonisation resistance that means the presence of the resident colonic microbiota works synergistically with the host (Gorbach *et al.*, 1988; Sorbara and Pamer, 2018; Lawley and Walker, 2013; Buffie and Pamer, 2013). This mechanism of resident microbiota protection include a number of factors such as the production of hydrogen peroxide, lactic acid, short chain fatty acids and bacitracin in addition to competition for mucosal binding sites and nutrients (Naaber *et al.*, 2004; Knaus *et al.*, 2017). Furthermore, Pham *et al*. (2014) confirmed that epithelial IL-22RA1-mediated fucosylation supports intestinal colonization resistance by pathogens and enhances host-microbiota to reduce overcolonization by opportunistic pathogens (Pham *et al.*, 2014). However, under antimicrobial therapy conditions, *C. difficile* infection accelerates the disruption of these protective mechanisms especially those by the resident colonic microbiota.

C. difficile pathogenesis can be divided into two steps after contamination by endogenous or exogenous *C. difficile* spores. Firstly, spores germination in the small intestine upon exposure to bile acids to form the vegetative active cells (Sorg and Sonenshein, 2008; Francis *et al.*, 2013). The resulting vegetative cells can adhere to the mucus layer and then penetration the enterocytes with the aid of flagella and proteases (Deneve *et al.*, 2009).

Secondly, after this first step of infection process (germination and colonisation), major virulence toxins, TcdA and TcdB, and for some strains an additional toxin, the binary toxin, immediately are produced and causing clinical CDI symptoms (Figure 1.1) (Janoir, 2016; Deneve *et al.*, 2009). Briefly, the toxins bind to and enter intestinal epithelial cells, and cause damage as demonstrated by actin disaggregation and cytoskeletal rearrangement. Finally, the end result of toxin activity due to actin polymerization in the intestine is opening of tight junctions, fluid secretion, mucosal damage and ultimately cell death and inflammation (Bartlett, 1994; Fekety, 1997).

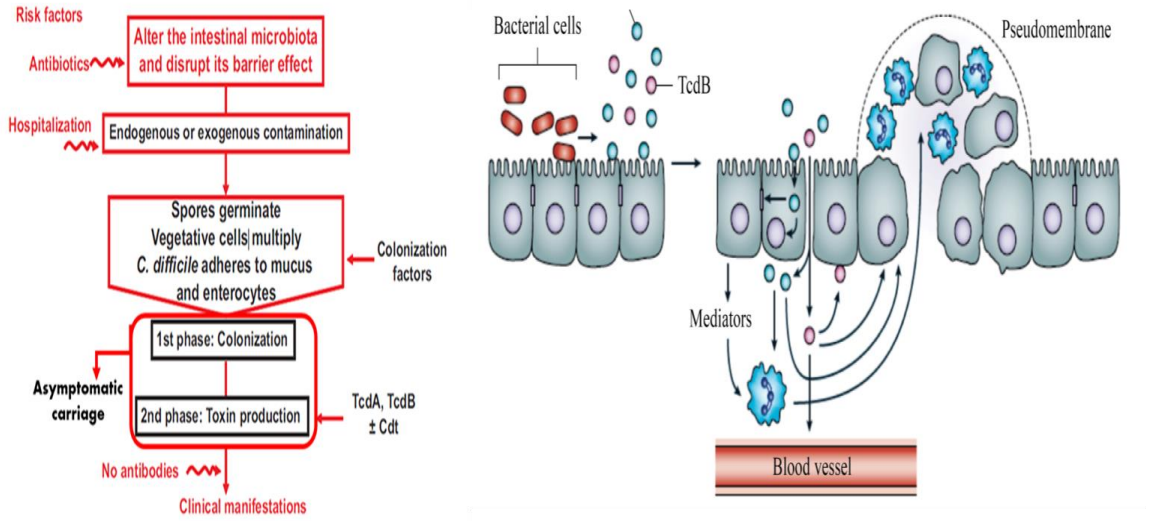


Figure 1.1. *Clostridium difficile* pathogenesis. The figure illustrates how *C. difficile* colonises the intestine, disrupts the host gut microbiota and adheres (via spores or vegetative cells) to the gut mucosa to secrete toxins in toxigenic strains. By the catalytic (enzymatic) protein activity of each toxin, toxins can insert into the cytoplasm through the membrane. The toxins can cause disruption of the tight junctions and loosening of the epithelial barrier, consequently cell death or release inflammatory mediators that can attract neutrophils. TcdB shown it has the ability to enter bloodstream in animal (Deneve *et al.*, 2009; Rupnik *et al.*, 2009).

1.1.2 Virulence factors of *C. difficile*

The virulence is a measure of the ability of an organism to invade the tissues of the host and produce disease. The description of the severity of this invasion can be determined

by expression levels of virulence factors. Virulence factors of bacteria include varieties of molecules that help bacteria to invade the host, cause host cell damage, and evade host defences (Baron and Peterson, 1996). These molecules can contribute to *C. difficile* pathogenicity and enable the bacteria to achieve colonization of a niche in the host, immune-evasion, immune-suppression and obtain nutrition from the host. Most of these virulence factors of bacteria are proteins and are encoded by genes located in chromosomal DNA, bacteriophage DNA or plasmids (Keen, 2012).

Over the last decade, the height of the epidemic of *C. difficile* infections especially RT 027 strains is attributed to the possession of several virulence factors, which include toxic and non-toxic virulence factors. These factors enable the pathogen to persist and disseminate through spore formation, the production of toxins in their vegetative form, and finally host colonization factors (Janoir, 2016).

1.1.2.1 Toxic virulence factors & regulation of their production

a) The Large Glycosylating Toxins, Toxin A and Toxin B

The severe symptoms of CDI are mainly caused by two exotoxins: toxin A (TcdA) (an enterotoxin and cytotoxin) and toxin B (TcdB) (a cytotoxin); most pathogenic strains of *C. difficile* produce both toxins (Bartlett *et al.*, 1978; Drudy *et al.*, 2007; Leffler and Lamont, 2015; Bartlett *et al.*, 1978). Together, detection of toxin A and toxin B is essential for *C. difficile* to be discriminated from pathogenic and non-pathogenic *C. difficile* strains (Geric *et al.*, 2004; Rupnik *et al.*, 2001; Kuehne *et al.*, 2010). Both of these toxins are proteins, pro-inflammatory, and are closely related proteins belonging to the Rho-glucosylating toxin or Large Clostridial Toxin (LCT) family which target small guanosine triphosphatases (GTPases) in the patient gut (Carman *et al.*, 2011; Jank *et al.*, 2015). This toxin group have the ability to inactivate GTPases leading to disaggregation of the host cell cytoskeleton, loss of the tight junctions between the epithelial cells, and ultimately cell death (Bella *et al.*, 2016a; Elliott *et al.*, 2017; Thelestam and Chaves-Olarte, 2000; Bella *et al.*, 2016b).

Recent studies pointed to the important role of TcdA and TcdB toxins in the pathogenesis of *C. difficile* associated diarrhoea (Pothoulakis and Lamont, 2001; Warny *et al.*, 2005). They present also as a main target for enzyme based or molecular diagnostic tests for detecting pathogenic *C. difficile* (Carroll, 2011; Crobach *et al.*, 2009). TcdA (308-kDa) and TcdB (270-kDa) toxins are encoded by the *tcdA* and *tcdB* genes, respectively, which

are located on a 19.6-kb of the pathogenicity locus named (*PaLoc*) of chromosomally integrated DNA sequence (shown in Figure 1.2) (Rupnik *et al.*, 2009; Hammond and Johnson, 1995). Voth and Ballard confirmed that *tcdA* and *tcdB* genes are expressed optimally in the late log and stationary phases in response to the nutrient limitation or accumulation of growth inhibiting substances occurring during the growth cycle (Voth and Ballard, 2005; Dupuy *et al.*, 2008). Although both toxins are regarded as major virulence factors of *C. difficile,* Lyras et al. demonstrated that TcdB is the major virulence factor by using tcdA and tcdB mutants in the *C. difficile* 630 strain (Lyras *et al.*, 2009). Furthermore, toxin B producing *C. difficile* strains which induce severe symptoms in infected patients are also able to cause diarrhea in the absence of toxin TcdA or CdtB (Janezic *et al.*, 2015).

It has been shown that the expression of TcdA and TcdB depend on environmental conditions and different regulators, including the availability of specific nutrients, changes in temperature, and fluctuation in the redox potential (Onderdonk *et al.*, 1976; Yamakawa *et al.*, 1996; Karlsson *et al.*, 2003; Bouillaut *et al.*, 2015; Haslam *et al.*, 1986). Generally, the transcription of the toxin genes occurs during the stationary phase of the bacterial growth cycle in which nutrients are exhausted and accumulation of growth inhibiting substances occurs (Hundsberger *et al.*, 1997; Dupuy and Sonenshein, 1998; Darkoh *et al.*, 2015). In addition, the presence of sugars and specific amino acids also inhibits toxin gene expression (Dupuy and Sonenshein, 1998; Karasawa *et al.*, 1997; Karlsson *et al.*, 1999)

b) Regulation of Expression *TcdA* and *TcdB* and Secretion of A&B Toxins

Since *C. difficile* possess the ability to produce toxins, therefore they need additional genes to organize the expression and excretion of these toxins. So, in addition to the *tcdA* and *tcdB* genes, the *PaLoc* also contains three other accessory genes that are thought to be involved in regulation of toxin production or release from the cell: *tcdR, tcdC* and *tcdE* (Figure 1.2)(Mani and Dupuy, 2001; Bouillaut *et al.*, 2015; Monot *et al.*, 2015; Smits *et al.*, 2016).

The *tcdR* (formerly known as *tcdD*) gene is one of putative regulatory genes, upstream of *tcdB* which encoding an alternative RNA polymerase sigma factor that is responsible for *tcdA* and *tcdB* expression as a positive regulator gene of toxin A and B production (Figure 2) (Mani and Dupuy, 2001; Moncrief *et al.*, 1997). TcdR protein is similar to the extra-

cytoplasmic function (ECF) sigma factor family (group 5 σ^{70}) such as *C. botulinum* (BotR), *C. tetani* (TetR) and *C. perfringens* (UviA), which is required for the initiation of toxin gene transcription as a σ factor of the RNA polymerase (RNAP) in the major pathogenic clostridia (Dupuy and Matamouros, 2006). Moreover, TcdR also activates its own expression and is also induced by one of several environmental signals (Mani *et al.*, 2002; Raffestin *et al.*, 2005). Hundsberger et al. demonstrated that the expression of *TcdR* is significantly increased in stationary phase and decreased in the presence of glucose in their growth conditions (Hundsberger *et al.*, 1997). In addition, expression of *tcdR* is significantly sensitive to environmental factors such as temperature, the cellular growth phase, and the ingredients of the medium. For example, *C. difficile* toxins proteins are induced at 37°C (Mani and Dupuy, 2001; Karlsson *et al.*, 2003).

Evidence for a major role of *tcdR* in activation of the expression of the toxins encoded by *tcdA* and *tcdB* has been confirmed when excessive release of toxins as a result of the up-regulation of toxin synthesis genes were observed (Moncrief *et al.*, 1997). Another evidence is that the *tcdD* gene positive regulator for producing toxins when it was shown in a heterologous *Clostridium perfringens* system that *tcdD* encodes an alternative sigma factor involved in transcription of toxin genes (Mani and Dupuy, 2001).

Another putative toxin non coding gene is *tcdC*, an anti-sigma factor situated downstream of *tcdA*, which is transcribed in opposite direction from the two toxin genes at highest expression levels during early exponential phase of growth (when there is a low toxin expression); this suggests that TcdC can be considered a negative regulator of the toxin genes (Figure 1.3) (Matamouros *et al.*, 2007; Hughes and Mathee, 1998). At the same time, its expression has declined as long as cells reach a stationary phase, the expression of *tcdA* and *tcdB* consistently increase (high toxin expression), indicating to its a negative regulatory role in both toxin genes expression (Dupuy *et al.*, 2008; Hundsberger *et al.*, 1997).

The *tcdE* gene encoding TcdE is a small open reading frame (ORF), a putative holin, of molecular weight 19-kDa which is located between *tcdB* and *tcdA* (122 bp downstream of *tcdB* and 727 bp upstream of *tcdA*) in the pathogenic locus (*PaLoc*) (Figure 1.3) (Tan *et al.*, 2001; Monot *et al.*, 2011). Dove and coworkers sequenced and analyzed TcdE (Dove *et al.*, 1990) as a protein with activity patterns similar to holins which have cytolytic action and is encoded by certain bacteriophages (Tan *et al.*, 2001). In addition,

the TcdE protein shares homology with the bacteriophage holin proteins, is involved in the release of progeny phages from the host bacterium, which shows its necessity for the extracellular release of both toxins through permeabilization process of the *C. difficile* cell wall (Govind *et al.*, 2015; Govind and Dupuy, 2012). Therefore, *C. difficile* toxin gene expression depends on the decline in *tcdC*, *tcdD*-increased expression, and *tcdE*-mediated release from the cell (Voth and Ballard, 2005). Moreover, the increase in *tcdE* expression shows its importance in the release of *C. difficile* toxins and suggests it might synergistically contribute to the hypervirulence of strains such as RT027 (Vohra and Poxton, 2011).

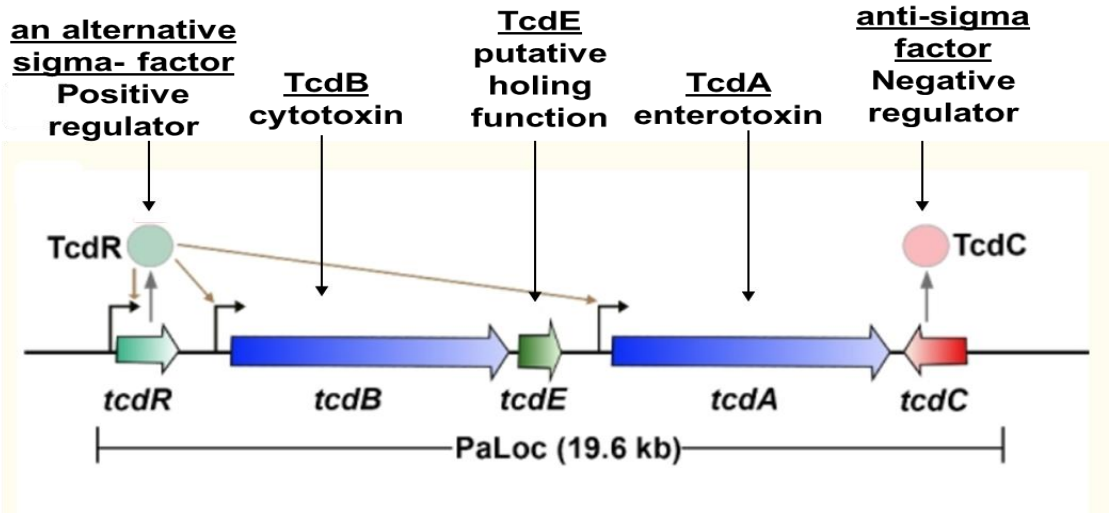


Figure 1.3. The Pathogenic locus (PaLoc) of *Clostridium difficile.*
The image shows the schematic representation of the 19.6 kb *C. difficile* PaLoc showing the genetic organization of *tcdA* and *tcdB* genes encoding the two major toxins (TcdA and TcdB respectively), and their negative and positive regulatory genes (*tcdC* and *tcdR* respectively) and the holin-like gene (*tcdE*) (Chandrasekaran and Lacy, 2017).

c) Other regulator genes that control toxin production

Expression of the LCTs and other virulence factors is regulated by a system known an accessory gene regulator (Agr) quorum-signalling locus. The Agr system is a Gram-positive specific quorum signalling system that enables the pathogenic bacteria to rapidly adapt the stress in their environment and modulate the expression of virulence-associated factors (Martin *et al.*, 2013; Millan *et al.*, 2016; Janzon *et al.*, 1989; Huelsenbeck *et al.*, 2009; Abdelnour *et al.*, 1993; Rutherford and Bassler, 2012). For example, the Agr quorum-sensing locus which was described firstly in *Staphylococcus aureus*, comprises the *agrACDB* genes and contributes to virulence by controlling the expression and

production of cell-surface colonization factors and other virulence exoproteins (Arya and Princy, 2013; Novick and Geisinger, 2008).

Currently, analysis of the genome of Clostridial species has revealed that homologues of the accessory genes regulator (*agr*) locus is present in a number of bacteria besides staphylococci. For instance, they have been found in various clostridial species such as *C. difficile*, *C. perfringens C. botulinum*, *C. sporogenes* and *C. acetobutylicum* (Patel, 2011). In addition to these clostridial species, the *agr* locus genes have also been found in other bacteria, including *Lactobacillus plantarum* (Diep *et al.*, 1994), *Enterococcus faecalis* (Nakayama *et al.*, 2001) and *Listeria monocytogenes* (Autret *et al.*, 2003). Further, the a *agr* locus is not restricted to the *C. difficile* RT027 strains but also exists in other pathogenic *C. difficile* such as PCR ribotypes 001 and 017 (Martin *et al.*, 2013).

Generally, The accessory gene regulator (*agr*) quorum sensing system of *C. difficile* R20291 regulates toxin production and includes a four gene operon consisting of *agrA*, *agrC*, *agr*D and *agr*B, respectively (Stabler *et al.*, 2009a; Darkoh *et al.*, 2015). The AgrD is encoded by *agrD* gene which makes an autoinducing prepeptide (AIP), which is processed by the transmembrane protein AgrB leading to release of an autoinducer peptide (AIP). Consequently, the AgrC histidine kinase protein binds the extracellular AIP, which then activates its ATPase activity leading to phosphorylation of AgrA which in turn regulates transcription of its target genes (Darkoh and DuPont, 2017).

Sequence analysis shows that many of the *Clostridia* have two physically separated sets of *agr* loci: *agr1* and *agr2*. The *agr1* locus is incomplete and consists only of the AIP generation genes (*agrB1* and *agrD1*), whereas the *agr2* locus is complete operon and contains both AIP generation and response genes *agrB2D2* and *agrC2A2*, respectively (Stabler *et al.*, 2009a; Montfort-Gardeazabal, 2017). The genomic analysis of *Clostridium spp.* showed that all *C. difficile* strains such as the non-hypervirulent strains 630, *C. perfringens* and *C. botulinum* encode only *agr1*, whereas the hypervirulent strains, such as R20291/ RT027 encode both *agr1* and *agr2* loci (Darkoh *et al.*, 2016; Patel, 2011). Interestingly, the two *agr* loci play a central role in toxin production and virulence factors (Darkoh *et al.*, 2016). Only when the *Agr1* (*agrB1D1*) locus is deleted, transcription of toxin genes is repressed in both hypervirulent strain R20291 and non-hypervirulent strain 630. Furthermore, mRNA transcripts of the *tcdA* and *tcdB* which are genes involved in the production of the toxin peptides, is also reported to be decreased in the 630 and

R20291 *agr1* mutants (Darkoh and DuPont, 2017). In addition, Cooksley and co-workers (2010) demonstrated that inactivation of *agrl* and *agr2*, respectively, affected sporulation in *C. botulinum* mutant strains (Cooksley *et al.*, 2010). However, an *agrB2D2* deletion mutant in the wild-type that has *agrB1D1* genes is able to produce toxin due its possibility to generate the TI signal (Agr-like quorum sensing peptide) (Darkoh *et al.*, 2016).

The role of *agr* in the regulation of *C. difficile* toxin expression has also been reported. The work of Othani et al. revealed that a mutant in *C. perfringens agrBD* strain 13, showed a clear reduction in toxin gene expression for both *tcdA* and *tcdB* genes (Ohtani *et al.*, 2009). Martin et al also reported that *C. difficile* R20291 produced lower levels of toxin A (TcdA) when there was an insertion in the *agrA2* gene (Darkoh *et al.*, 2016). Both sporulation and neurotoxin production were also drastically reduced in *agrD* mutant strains of *C. botulinum* (Cooksley *et al.*, 2010). In contrast, the work of Patel (2011) demonstrated that *agrB* has no significant role in the regulation of virulence factors in *C. difficile* 630, although *agrB2* and *agrA* mutants showed significantly reduced toxicity of both A and B toxins towards vero cells and HT29 cells (Patel, 2011).

d) Binary toxin (CDT)

Recently, there has been demonstrated a new hypervirulent *C. difficile* strain identified as the epidemic BI/NAP1/027 produces an additional toxin which has been termed the *C. difficile* transferase (CDT; or binary toxin) which has been associated with more severe disease (Elliott *et al.*, 2017; Perelle *et al.*, 1997). CDT is produced by 20% to 32% of *C. difficile* strains according to a recent report (Eckert *et al.*, 2015). Binary toxin is an actin-specific ADP ribosyltransferase encoded by two genes, *cdtA* and *cdtB*, which are located on the *C. difficile* chromosome at a locus 6.2-kb separate from the *Paloc* called *cdt* locus (cdtLoc), with a positive transcriptional regulator named *cdtR* (Figure 1.3) (Hamm *et al.*, 2006; Carter *et al.*, 2007; Perelle *et al.*, 1997). A later study confirmed that CdtR regulates the production of toxins A and B in some *C. difficile* strains (Lyon *et al.*, 2016).

The *C. difficile* CDT was first described in 1988 (Popoff *et al.*, 1988) and belongs to the family of clostridial binary toxins that also includes *C. perfringens* iota toxin, *C. spiroforme* CST toxin, *C. botulinum* C2 toxin, as well as *Bacillus cereus* vegetative insecticidal protein (or VIP) (Carter *et al.*, 2007; Stiles *et al.*, 2016; Barth *et al.*, 2004). At least 20% of *C. difficile* strains produce an extra toxin, binary toxin (CDT), the presence of which have been associated with more severe disease (Barbut *et al.*, 2005;

Bacci *et al.*, 2011), although this is not considered to be evidence as a major cause of the clinical CDI signs (Eckert *et al.*, 2015; Gerding *et al.*, 2014). In addition, the incidence of *C. difficile* strains which produce only CDT is very low and such strains tend to yield a moderate symptom of enteritis in lab animal infection experiments (Knapp *et al.*, 2016; Carman *et al.*, 2011). However, a recent study revealed for the first time that the LC693 *C. difficile* strain, which does not belong to the hypervirulente RT027 and RT078 groups, produces binary toxin and was found to cause severe diarrhea in a patient in Xiangya Hospital in China (Li *et al.*, 2018).

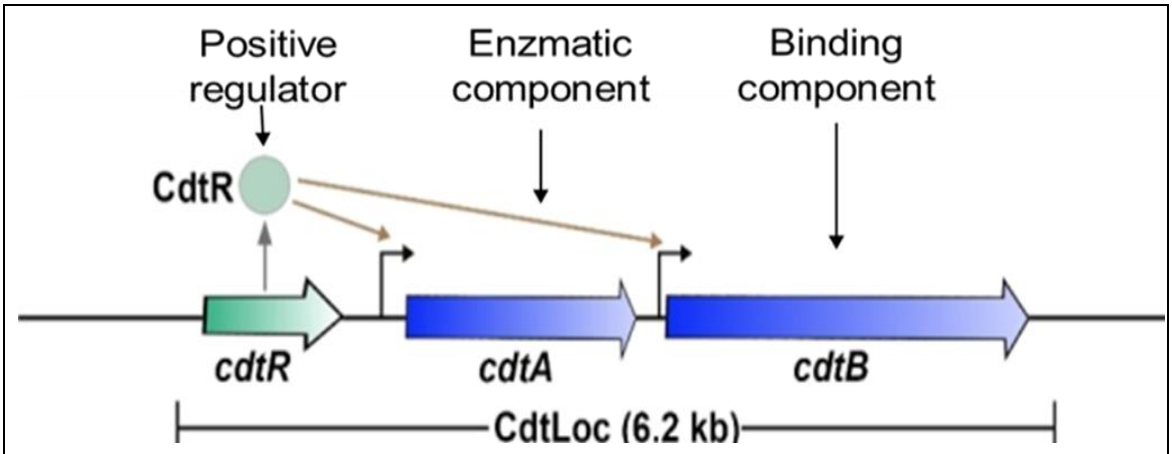


Figure 1.3. Schematic representation of the binary toxin locus (CdtLoc) region.
The image shows the binary toxin locus (CdtLoc) region which is containing the following genes: CDT-encoding genes, *cdt*A and *cdt*B, are shown in blue. The regulatory gene *cdt*R is shown in light green which positively regulates the synthesis of CdtA and CdtB.

1.1.2.2 Other *C. difficile* virulence factors (Non- toxin virulence factors)

The new epidemic strains of *C. difficile* have non-toxin virulence factors which play role in *C. difficile* colonization, proliferation and maintenance in the host gut (Viswanathan *et al.*, 2010).

1.1.2.2.1 Adhesion factors

Adhesion factors play important role in the ability of bacteria to adhere the gut surface in their host, which is urgently required for the full demonstration of virulence (Janoir, 2016). In 1979, it was shown for the first time to describe that attached *C. difficile* cells were present to the surface in washed biopsy sample from patients with PMC under the

microscope (Borriello, 1979). There are various factors that can impact the ability of *C. difficile* to adhere to the gut, considered in the next sections:

a) Fimbriae

A fimbria (plural fimbriae), is a cell-surface multisubunit protein polymer appendage that can be found on many Gram-negative, some Gram-positive bacteria and archaea (Mandlik *et al.*, 2008; Proft and Baker, 2009). It is regarded as one of the main mechanisms of virulence for *Staphylococcus, Streptococcus, E. coli, Haemophilus influenzae* and *Bordetella pertussis* by supporting the bacteria's ability to attach to the host and cause disease (Wayoff and Jankowski, 1989; Connell *et al.*, 1996; Spurbeck *et al.*, 2011). Many strains of *C. difficile* have fimbriae which are polar or peritrichous, 4–9 nm in diameter and 6 μm in length. These factors were detected in 1988, and are used for specific attachment in the gut so that they cannot be removed easily (Borriello *et al.*, 1990; Borriello *et al.*, 1988).

b) Flagella

A flagellum (plural flagella) is another bacterial adhesin which can play important role in bacterial pathogenicity. *C. difficile* is known to have peritrichious flagella which consist of two structural components, the flagellin FliC and the flagellar cap protein FliD, and both of these have been shown to participate in the process of attachment to the intestinal mucus s(Tasteyre *et al.*, 2001; Stevenson *et al.*, 2015). Flagellum-mediated motility is an essential virulence factor which is required to colonise the stomach by most gastrointestinal pathogens such as *Helicobacter pylori* (Kirov *et al.*, 2004), *Listeria monocytogenes* (O'Neil and Marquis, 2006), *Vibrio anguillarum* (Ormonde and Rstedt, 2000), *Campylobacter jejuni* (Grant *et al.*, 1993). Furthermore, the role for flagellum components as adhesins for attachment to host cell surfaces is not restricted to *C. difficile*, but also enteropathogenic *Escherichia coli* flagellum enhance the adherence to the surface of epithelial cells (Giron *et al.*, 2002). Moreover, the FliC and FliD proteins of *Pseudomonas aeruginosa* mediate the adherence of the organisms to host cells and colonization of the respiratory tract (Arora *et al.*, 1998). Recently, flagella have been shown to play a role in the formation and development of complex communities known as biofilms during pathogen survived *in vivo* in *C. difficile* (Dapa *et al.*, 2013) and in a number of pathogens (Blair *et al.*, 2008; Kirov *et al.*, 2004). Baban's group has also demonstrated that flagella in strain R20291, do play important role in colonisation and adherence of intestinal epithelium during infection (Baban *et al.*, 2013).

c) Surface layer proteins (SLP)

One of the most common surface structures on bacteria that has been detected in all *Clostridium difficile* strains is the S- layer proteins (SLP). The S-layer of the vegetative *C. difficile* is considered as another virulence factor that is composed of two distinct surface layer proteins, the HMW (high molecular weight) and LMW (low molecular weight) proteins, which are derived by proteolytic cleavage of the precursor SlpA (Cerquetti *et al.*, 2000; Waligora *et al.*, 2001). S-layers are also found on both other Gram-positive and Gram-negative bacteria and are highly prevalent in the archae (Sára and Sleytr, 2000; Albers *et al.*, 2011). Recent studies have shown that *C. difficile* SLP are involved in growth and survival, and contribute to host cell adhesion and induction of cytokine production (Calabi *et al.*, 2001; Cerquetti *et al.*, 2000; Bianco *et al.*, 2011). The importance of *C. difficile* SLP as a major contributor to bacterial adherence was also underlined by the study of Merrigan and colleagues, which demonstrated that the bacterial attachment to host cells significantly reduced after pre-treatment of host cells with crude or purified SlpA subunits, or incubation of vegetative bacteria with anti-SlpA antisera (Merrigan *et al.*, 2013).

d) Physicochemical properties

The physicochemical properties of a bacterium can enhance its capability of adhering to surfaces. *C. difficile* cells are hydrophilic and carry a net positive charge which is evenly dispersed over the cell wall (Krishna *et al.*, 1996). Charged interactions with the negatively charged human gut can therefore lead to colonisation (Waligora *et al.*, 1999).

1.1.2.2.2 Hydrolytic enzymes

Other non-toxogenic factors may also contribute to bacterial virulence as species of anaerobic bacteria produce a comprehensive array of hydrolytic enzymes (Steffen and Hentges, 1981; Hafiz and Oakley, 1976). Seddon *et al.* (1990) demonstrated that all strains of *C. difficile* have the ability to produce hydrolytic enzymes such as hyaluronidase, chondroitin-4-sulphatase, collagenase and heparinase, although the heparinase was usually less in levels (Seddon *et al.*, 1990). Further research also revealed that these hydrolytic enzymes result in breakdown of connective tissue and epithelial cell inflammation thereby promoting host cell invasion (Seddon *et al.*, 1990; Borriello, 1998).

1.1.2.2.3 *C. difficile* spores

For many bacteria the possession of spores is an important virulence factor. Similar to other members of the *Clostridium* genus, *C. difficile* produces endospores that spread the infections since vegetative cells, being anaerobic, cannot survive outside of the host (Akerlund *et al.*, 2008). More importantly, those spores which are heat-resistant also play an important role in the dissemination, perseverance and pathogenesis of *C. difficile* through enable it to survive harsh conditions (Lyerly *et al.*, 1988; McDonald *et al.*, 2007; Jump *et al.*, 2007). Interestingly, only one third of all patients infected by *C. difficile* developed diarrhea, while the remaining two thirds were asymptomatic carriers due to the spore-forming bacterium *C. difficile*, in turn these asymptomatic patients serve as a silent reservoir of *C. difficile* infection (Kelly and LaMont, 1998; Vaishnavi, 2010). In addition, *C. difficile* spores in turn enable the bacteria to persist and spread CDI by the faecal-oral route or through patients touching contaminated surfaces, including the hands of healthcare staff (Claro *et al.*, 2014; Joshi *et al.*, 2017).

1.1.2.2.4 *C. difficile* capsule

It has been recognised since the early years of this century that bacterial capsules enhance pathogen's virulence potential. Although it was unclear whether *C. difficile* has a capsule as other members of the genus *Clostridium*, research has demonstrated that *C.difficile* strains can possess a capsule which may be involved in opsonization against phagocytosis; this would indicate the presence of an anti-phagocytic factor on the bacteria (Davies and Borriello, 1990; Borriello *et al.*, 1990; Dailey *et al.*, 1987; Strelau *et al.*, 1989). Other studies suggested that the polysaccharide capsule of Gram positive bacteria such as *Bacteroides fragilis* can be chemically characterized to become a virulence factor by its interfering with the phagocytosis and killing by polymorphonuclear leukocytes (PMNs) (Kasper, 1976; Klempner, 1984; Onderdonk *et al.*, 1977).

1.1.2.2.5 Bacteriophage carriage

Another important factor that can be contributed to the pathogenicity of many bacteria through transduction of virulence genes are bacteriophages (Boyd, 2012; O'Sullivan *et al.*, 2019). Bacteriophages are one of the extra virulence factor group which are termed mobile genetic elements (MGEs) that can influence virulence production such as plasmids, transposons or bacteriophages (Wagner and Waldor, 2002; Henn *et al.*, 2010; Rankin *et al.*, 2011). The study of the relationship between bacteriophages and the

transmission of pathogenicity to bacterial hosts was first carried out by Frobisher and Brown as early as 1927, where they discovered that after exposure of nontoxigenic streptococci to the filtered supernatants of toxigenic streptococcal cultures acquired these bacteria the ability to produce toxin (Frobisher and Brown, 1927). Furthermore the bacteriophages contribute to virulence of bacteria is also supported by Freeman (1951) when found that avirulent strains of *Corynebacterium diphtheriae* can cause diphtheria after infection with a bacteriophage transformed it into a virulent strain (Freeman, 1951). Several studies have shown that bacteriophage can encode virulence factors for pathogens, such as Shiga toxin (*stx*) in *Escherichia coli* (O'Brien *et al.*, 1984), botulinum neurotoxins (BoNTXs) in *Clostridium botulinum* (Sakaguchi *et al.*, 2005) and cholera toxin in *Vibrio cholera* (Waldor and Mekalanos, 1996). Similarly, other studies have shown that phage infection of toxigenic *C. difficile* strains can cause an increase in toxin production (Goh *et al.*, 2007; Govind *et al.*, 2009; Sekulovic *et al.*, 2011; Hargreaves and Clokie, 2014; Fortier, 2018). Additionally, Certain prophage genes can also encode putative multidrug (antibiotic resistance) genes, proteases, and multiple factors that could interfere with host immune defences (Garneau *et al.*, 2018).

1.1.3 The *C. difficile* Genome

In the last 10 years or so a new group of highly virulent *C. difficile* strains has emerged to cause outbreaks of increased disease severity in North America and Europe (Kuijper *et al.*, 2006; O'Connor *et al.*, 2009). Most of the strains associated severe diarrhoea with high morbidity and mortality were of epidemic 027 origin (Goorhuis *et al.*, 2007; Hubert *et al.*, 2007). Prior to 2006, there is no *C. difficile* strain completely genomic sequenced until Sebaihia and his colleagues sequenced and annotated the first extreme genome of the non-epidemic *C. difficile* 630 (RT012) strain using the Sanger Centre, UK by whole-genome microarray analysis (Sebaihia *et al.*, 2006). *C. difficile* 630 strain, a fully virulent and multidrug-resistant strain, was isolated in 1982 from a patient with severe PMC (Stabler *et al.*, 2009a; Knight *et al.*, 2015).

Although all genomes of *C. difficile* strains are highly conserved, the new 027 strains that have emerged may contribute to the distinct phenotypic differences relating to motility, antibiotic resistance and toxicity compared to the parent 630 strain (Stabler *et al.*, 2009a). Since this time, several other *C. difficile* genomes ranging in size from 4.1 to 4.3 Mbp have been fully sequenced and annotated in different countries: G46 (RT027; isolated in the United Kingdom in 2006), M120 (RT078; isolated in the United Kingdom

in 2007) (Gaulton *et al.*, 2015), 196 (RT027; an historic non-epidemic; isolated in France in 1985), BI1 (RT027; isolated in the United States in 1988) and R20291 (RT027; isolated in the United Kingdom in 2006) (He *et al.*, 2010; Stabler *et al.*, 2009a; Brouwer *et al.*, 2012). Generally, over last 8 years ago, genetic typing and sequence analysis have shown that there are only five lineages of *C. difficile* which include the common significant PCR ribotypes: RT012, RT017, RT023, RT027, RT078 (Valiente *et al.*, 2016; Dingle *et al.*, 2011; He *et al.*, 2010). In 2009, Stabler and colleagues described the fully sequenced and annotated genome of a recent epidemic *C. difficile* (R20291, RT 027). This hypervirulent strain was isolated following an outbreak in UK and it is very closely related to the hypervirulent BI strains in North American (Stabler *et al.*, 2009a). Recent comparative genomic studies concerning *C. difficile* revealed that there is significant genetic differences between PCR-ribotype (027) strains and previously published PCR-ribotype 012 strain 630 which in turn produce a phenotypic effects including motility, antibiotic resistance and toxicity (Stabler *et al.*, 2009a). The whole genomes sequencing of *C. difficile* (R20291) strain is available at GenBank (NCBI, FN_545816.1). Using RAST server with *C. difficile* (R20291) strain revealed a large circular chromosome of 4,191,339 bp (4.3 Mb) in length and G+C content of 28.8% (Aziz *et al.*, 2008).

1.1.4 Treatment and prevention of CDI

Treatment of CDI has been employed mainly on the basis of the disease condition and its severity. A mild CDI can potentially be controlled by discontinuing the antibiotics that cause the infection (Dumyati *et al.*, 2012). But in moderate cases, treatment with other antibiotics like vancomycin and metronidazole can usually eliminate *C. difficile* (Zar *et al.*, 2007). In other cases, surgical intervention to remove the damaged section of the colon may be the only option for patient with serious symptoms and life threatening prognosis (Bingley and Harding, 1987; Leffler and Lamont, 2015). Recently a novel therapy has emerged called faecal microbiota transplanting (FMT) (Borody and Campbell, 2012). Probiotics such as *Lactobacillus* spp. and *Saccharomyces boulardii* can also be used as probiotics in the treatment and prevention of recurrent *C. difficile* infection through restoring the natural balance of microbiota in the gut (Tuohy *et al.*, 2003; Edwards-Ingram *et al.*, 2007). The appropriate prescription and the optimum use of antibiotics are the most important means to reduce the risk of CDI (Surawicz *et al.*, 2013). Prevention of *C. difficile* is not hard. Limiting use of antibiotics can keep the bacteria from over-colonizing the gut. Prescribing appropriate antibiotics can also prevent the

bacteria from every bothering a patient. Almost 50% of antibiotics prescribed in hospital settings are not appropriate or entirely safe which enables opportunistic pathogens such as *Clostridium difficile* to take advantage of the host changed environment. Wearing gloves and gown by hospital staff can help to limit or event prevent the spread of the disease (Landelle *et al.*, 2014; Dubberke, 2010).

1.1.5 Polymorphism in *C. difficile*

Polymorphism means two or more distinctly different phenotypes present in the same population of species. In other words, it is a description of variations in morph and form that distinguish normal organisms within a species from each other (Ford, 1965). (Dobzhansky, 1950) defined polymorphism as the term that expresses the biodiversity, genetic variation and adaptation in nature by which organism can usually retain variety of form in a population living in a different environment.

In *C. difficile* a genetic instability generates a high degree of polymorphism consisting of six main phylogenetic clades designated 1, 2, 3, 4, 5, and C-I (Stabler *et al.*, 2012; Dingle *et al.*, 2011). In addition, the report of (Nale *et al.*, 2012) indicated that indicated that high diversity of *C. difficile* 027 strains was associated with existence of mobile genetic elements including temperate bacteriophages. Although these strains belong to the same ribotype but they did differed in their pathogenicity not a proper sentence.

In another study, the complete genome sequence analysis revealed that a large proportion (11%) of the genome *C. difficile* strain 630 has mobile genetic elements, mainly in the form of conjugative transposons genetic. These mobile elements may be responsible in the phenotypic differences of this strain from others (Sebaihia *et al.*, 2006). Variation within *C. difficile* strains can also occur due to plasmids. (Clabots *et al.*, 1988) confirmed that plasmids were exist in about 30 (37%) of 82 clinical isolates in patients infected with *C.difficile.*

1.2 Bacteriophages

1.2.1 Introduction

Over the last decade bacteriophages (Bacterial virus or phage) have been shown to be the most widespread and abundant organisms and it estimated there are more than 10^{31} phage particles on the planet (Brussow and Hendrix, 2002; Squires, 2018; Hendrix *et al.*, 2000; LaFee and Buschman, 2017). Phages are not only the most abundant biological entities but probably also the most diverse (Hatfull *et al.*, 2008). The fight for survival between

bacteria and their viral predators (bacteriophages) is ever constant, therefore, bacteria attempt to create new defence mechanisms to face this attack. In response, the phages develop strategies in order either to overcome these host protection mechanisms or evade them.

1.2.2 History and taxonomy

Bacteriophages were discovered for the first time by a British pathologist Frederick Twort in 1915, who described glassy transformation of micrococci colonies without stated them as bacterial virus (Twort, 1915). In the next two years, Felix d'Herelle isolated an anti-microbe of *Shigella dysenteriae* and called these isolates the name of bacterial eater in 1917 (d'Herelle, 2007). It is known that there is considerable diversity among phages, and they are classified according to their genomic and morphological characteristics (Ackermann, 2009; Murphy *et al.*, 2012). After the invention of the electron microscope in 1940s, there was a significant development in the field of virus taxonomy in which genetics specialist were enabled to classify phages based upon their morphology (Orlova, 2012). Subsequently, virus were later classified by the Lwoff, Horne and Tournier based on their type of nucleic acid (DNA or RNA), capsid form and enveloped nature (Lwoff *et al.*, 1962). Gnomically, the genetic material of phages consists of double-stranded (ds) or single-stranded (ss) DNA or RNA, whereas their genome sizes ranged from very simple ~3.5 kb ssRNA in *Escherichia coli* MS2 phage to very complex ~500 kb dsDNA genome in *Bacillus* phage G virus (Villarroel *et al.*, 2018; Salmond and Fineran, 2015). Also, phage can include modified nucleotides as protection system against restriction enzymes (Salmond and Fineran, 2015). Morphologically, the phenotype of the phages differs greatly, therefore some have a tail or tailless, filamentous or pleomorphic, polyhedral, with lipid or lipoprotein envelopes(Ackermann, 2009; King *et al.*, 2011).

1.2.3 Structure

In general, bacteriophage has two main molecular components: nucleic acid and protein. Protein has been represented in the head (or capsid), an outer shell, which can vary in size and shape (Guttman *et al.*, 2005). The head, acts as a protective coating and encloses the phage genetic material. The bacteriophage can carry a single or double strand of RNA or DNA (ssRNA, dsRNA, ssDNA or dsDNA) in either a circular or linear arrangement (Lodish *et al.*, 2000). Sometimes, bacteriophage have a contractile sheath surrounding its tail as in the T4 phage (Figure 1.4) (Aksyuk *et al.*, 2009). Also at the end of the tail, some phages have tail fibres and a base plate. Both structures play a role in the attachment of

the phage to the bacterial cell thereby facilitates high viral infection efficiency (Hyman and van Raaij, 2018; Hanlon, 2007). However, approximately 95%, of all known bacteriophages has a tail and dsDNA (for example *E. coli* phage T4, whereas only 5% is tailless with either ssDNA, ssRNA or dsRNA genome (Ackermann, 1998) .In principle, The main effective differences between phages based essentially on their style of entry, genome incorporation and release (Gorski *et al.*, 2007).

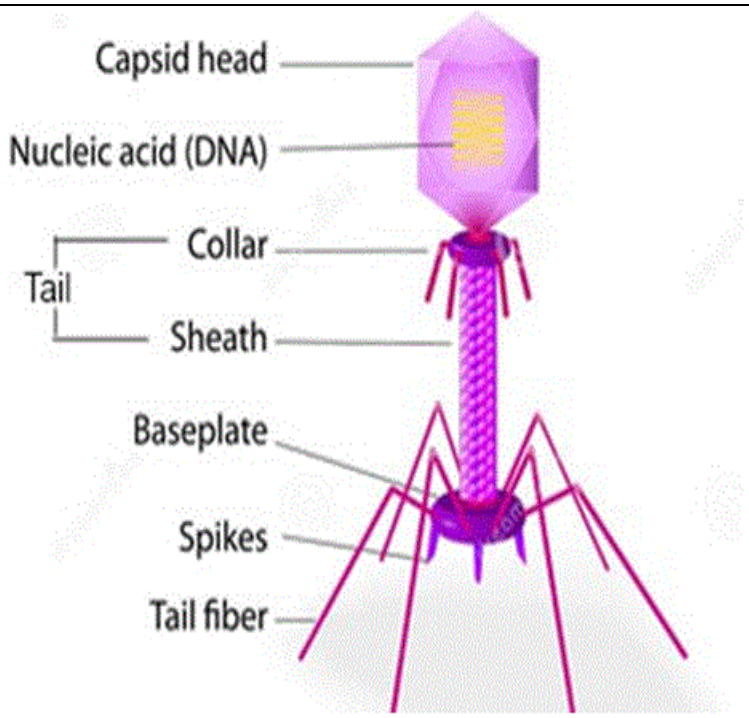


Figure 1.4. Illustration of the basic structural features of bacteriophages (typical bacteriophage structure) (Mansour, 2017).

1.2.4 Replication and life cycle of bacteriophage

All bacteriophages are obligate intracellular parasites of bacteria, and can be divided into main categories based according to the type of infection they cause: strictly lytic phages (or virulent) following a lytic cycle and lysogenic (or temperate phages) following a lysogenic cycle (Salmond and Fineran, 2015). In addition, Guttman *et al* state that some types of phages can adopt a pseudotemperate lifestyle in which they do not integrate the chromosome of their host but remain free in the cytoplasm in a linear or circular plasmid state (Guttman *et al.*, 2005).

Bacteriophage are known to possess distinctive receptor binding units which specifically bind to a host receptor. These specific receptors may be present on the cell envelope,

capsule, flagella or even conjugative pili, by their tail fibres or by some other structure in those phages that lack tail fibres (Young *et al.*, 2000). Therefore Phages which target Gram-positive bacteria seem can adhere to specific teichoic acids in peptidoglycan layer whereas phages for Gram-negative bacteria can bind to the different membrane structures such as proteins, lipopolysaccharides or sugars (Chatterjee, 1969; Heller and Braun, 1979; Brussow *et al.*, 2004; Spinelli *et al.*, 2006).

Phages have either a lytic or a lysogenic life cycles (Figure 1.5) (Dy *et al.*, 2014). With the lytic life cycle, phages (DNA/RNA) infect a host cell, resulting in quick lysis and death of the infected host cell in a very short period after injecting their genetic material inside the infected bacteria. Whereas, temperate bacteriophages (DNA) can insert their genome into the host in their lysogenic life cycle (Sulakvelidze *et al.*, 2001).
To infect a host bacterium in both lytic and lysogenic life cycles, phage will first interact with specific receptors on the host cell, adsorb and then inject its genome into the bacterial cell.

In the lytic life cycle, after penetration, the genome of virulent or lytic phages circularises in the cytoplasm of the bacteria (Figure 1.5). Then immediate early viral genes are transcribed by the host cell RNA polymerase, which can cause restructuring of bacterial metabolism, inactivation of host defences and protection of own DNA/RNA. Subsequently, middle phage genes are transcribed and redirecting the metabolic processes of the bacterium to produce new phages genome occurs. In the late stage of the reproductive cycle, late virus genes are expressed which encode for the remaining virion particles (capsid proteins) which self-assemble to release the progeny phages (Young *et al.*, 2000; McGrath and van Sinderen, 2007).

To degrade the cell wall of the infected bacteria, tailed phages have commonly two proteins which regulate the lysis process: an endolysin (often called lysin) and a holin (Ackermann, 1998). Once the holin, a small protein that forms oligomers, creates pores within the cytoplasmic membrane, endolysin enable phage particles to gain access to the peptidoglycan layer that it hydrolyses (Wang *et al.*, 2000). This results in cell rupture (lysis) and the release of a few to a few thousands of new progeny phage into the surrounding environment that will be able to re-infect susceptible bacteria nearby and the cycle will start over again (Young, 1992; McGrath and van Sinderen, 2007). However,

not all virulent phages cause death for host cell; filamentous phages can replicate and exit from the bacteria cell without killing the organism (Russel, 1995) .

In the lysogenic cycle, temperate bacteriophages have the same lytic capacity as virulent phages but in addition, they can select a nonlytic lifestyle (Guttman *et al.*, 2005) (Figure 1.5). Once the phage DNA insert their genome within minutes into the host cell cytoplasm, the phage can choose to initiate a lytic cycle or to integrate its DNA into the bacterial chromosome of its host to form a prophage. These integrate phage genomes (prophages) can replicate with the host DNA, either in a free, plasmid-like state (for example, phage P1) or integrated into the bacterial chromosome (for example, phage λ). In addition, under stress condition, prophages can turnover from the lysogenic state and switch back to lytic life cycle by exiting from bacterial genetic material to produce more virions that are finally released from the bacterium (Orlova, 2012; Lamont *et al.*, 1993; Feiner *et al.*, 2015).

A bacterial cell which carrird one or more prophages is called a lysogen and generally, it will be immunized to infect by lytic or lysogenic infection by other phages of the same group (Guttman *et al.*, 2005).

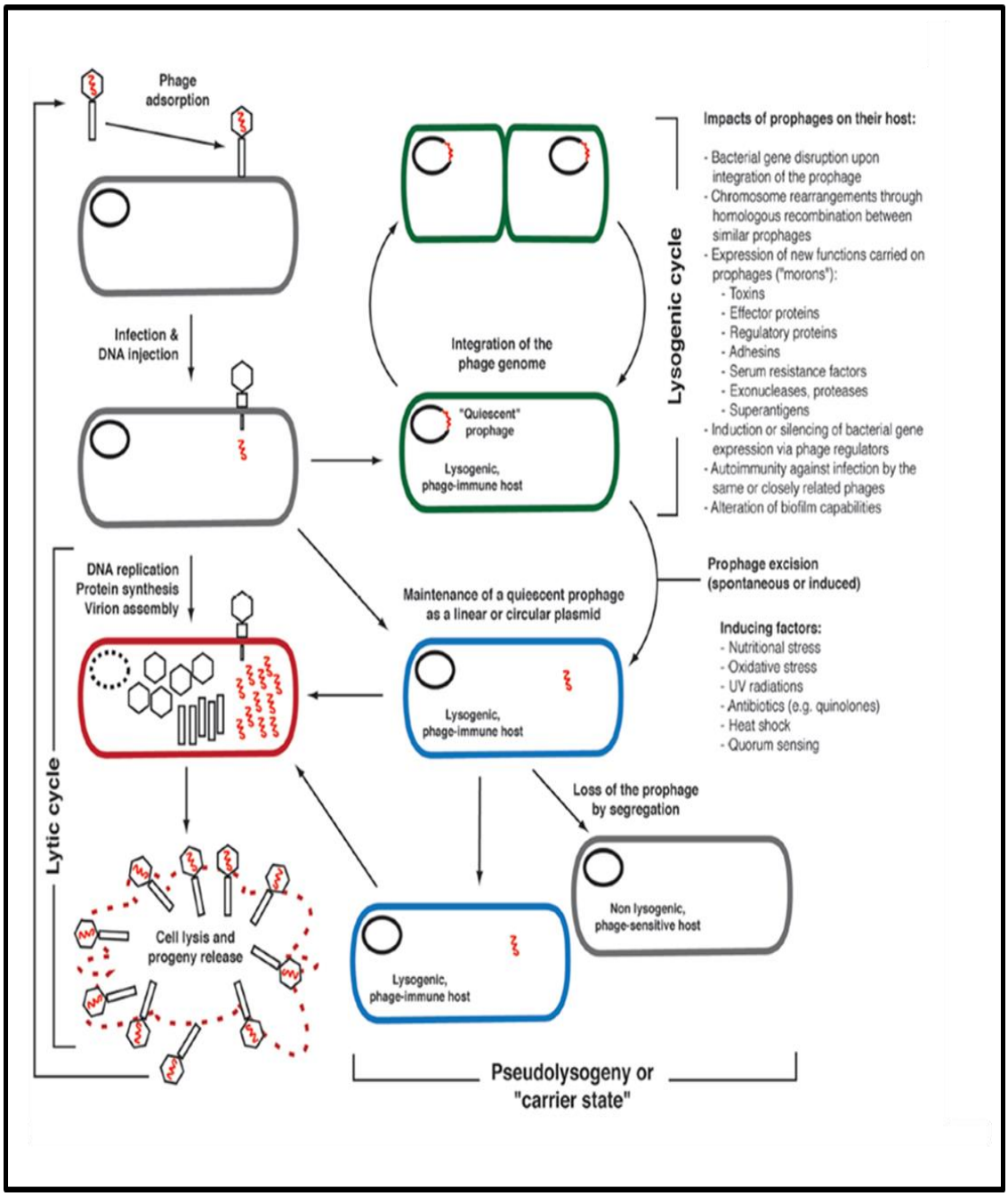


Figure 1.5. Different lifestyles adopted by bacteriophage. Virulent phages only follow the lytic cycle for their replication and will lead to the lysis of the host cell at the end of the cycle. Temperate phages have the choice to replicate through the lytic cycle like virulent phages, or they can select for the lysogenic cycle. Some phages also adopt a pseudotemperate lifestyle, i.e., they generally do not integrate the chromosome of their host and replicate as linear or circular plasmids within the cytoplasm (Fortier and Sekulovic, 2013).

1.2.5 One-Step Growth Curve

Bacteriophages, also termed more simply phages, are viruses that attack bacteria. Phages, like other viruses, cannot exist without a suit host. To understand the interaction between viruses and their hosts, Ellis and Delbruck proposed a technique by which can quantitate and monitor the growth of phage in a specific host (Ellis and Delbrück, 1939). In this technique all cells in a culture are simultaneously infected with phage so that the stages of viral replication, attachment, entry, uncoating, genome replication, capsid assembly and release occur as close to synchronously as possible. Simply, to perform this technique, the bacteria are mixed with phage and incubated for a short period of time. The mixture is then diluted to reduce the number of bacteria available for phage adsorption. Samples are removed at specified intervals and plated to quantitate the phage present in the culture. In addition to determination of three phage growth phaeses (Latent, Log and Stationary), the latent period (time required for the phage to cause cell lysis) and burst size (phages produced per infected cell) for a subset of phage– host combinations can be estimated by a one-step growth curve (Figure 1.6) (Adams, 1959). The phage latent period can be defined as the timing of phage induced host cell lysis, which typically is controlled by the phage protein complex known as a holin (Young *et al.*, 2000).

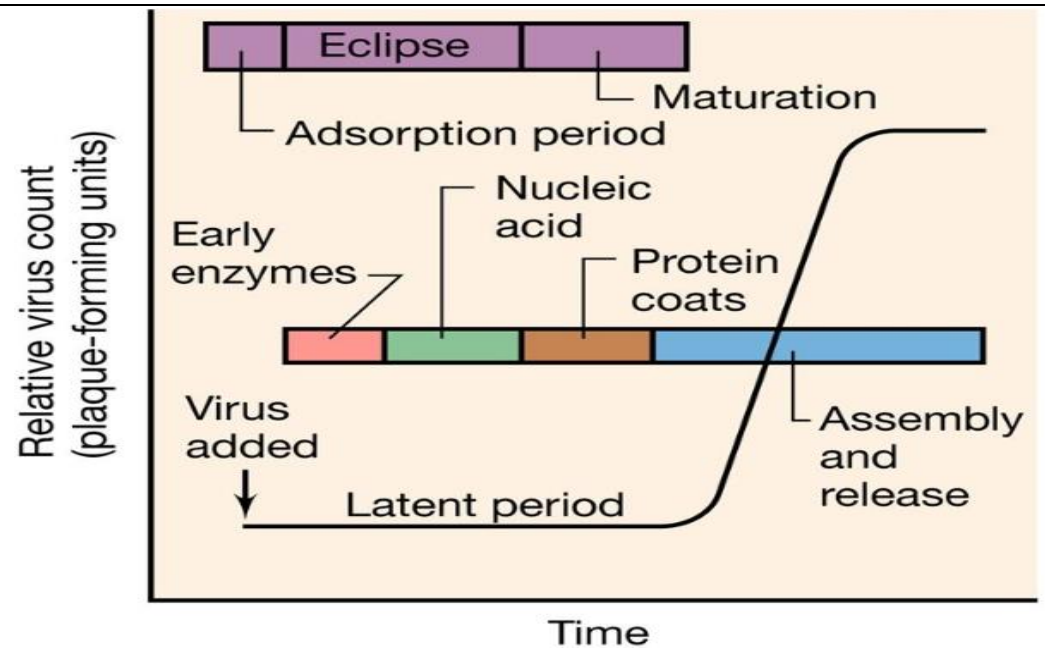


Figure 1.6. Illustration of growth stages of bacteriophage during one-step growth curve.

1.2.6 Phage Therapy

Several research studies have shown that strictly lytic phages are more suitable and preferable than temperate phages for phage therapy purposes (Pires *et al.*, 2015; Lin *et al.*, 2017; Fortier and Sekulovic, 2013). Although temperate phages are highly abundant and are more readily available than lytic phages, there are some limitations that can be used in the treatment (Weld *et al.*, 2004). One of the chief limitations is due to possess a lower killing potential and can transduce resistance and toxin genes from one bacterium to another which contribute to the spread of antibiotic resistance between bacteria, thus enhance bacterial virulence (Skurnik and Strauch, 2006; Mazaheri Nezhad Fard *et al.*, 2011; Kropinski, 2006). In addition to the ability of temperate phages to integrate into the genome of bacteria, it poses a risk when used in that they potentially can encode uncharacterized toxins (Nilsson, 2014). Also less desirable for therapeutics are those phages that such low virulent can be due to poor adsorption properties, low potential to evade bacterial defences, or poor replication characteristics (Abedon and Thomas-Abedon, 2010).

However, the emerging widespread appearance of resistance to conventional antibiotics over the last decades has driven a search for alternatives to control of resistant bacteria (Viertel *et al.*, 2014). As phages have the ability to target, degrade and penetrate into bacterial biofilm and form plaques (Nale *et al.*, 2016; Parasion *et al.*, 2014). Because of these properties, bacteriophage has possibility become to use as novel therapeutics for antibiotic resistant bacterial infections through phage therapy. Moreover, the specific nature of the phage–bacterial interaction make phages to be an optimal option with treatment of CDI by avoiding disruption of gut micro- biota (Peterfreund *et al.*, 2012). Bacteria-specific viruses have been used first for treatment against pathogens such as *Shigella dysenteriae* as early as 1919 (Chanishvili, 2012; Lin *et al.*, 2017).

1.2.7 C. difficile phages identified and sequenced to date

The majority of known *C. difficile* phages that identified to date are temperate, double strand DNA tailed viruses which belong to the Myoviridae or Siphoviridae families and have genomes of medium size (32– 57 kb)(Goh *et al.*, 2005a; Mahony *et al.*, 1985; Ramirez-Vargas *et al.*, 2018). In the 1980s, the number of bacteriophages have been used in the typing of *C. difficile* strains was very limited (Hargreaves and Clokie, 2014; Sell *et al.*, 1983). Recently, numbers of whole phages isolations and purified phage components, with a number of temperate *C. difficile* phages have been successfully

isolated and observed by the use of electron microscopy (EM) (Phothichaisri *et al.*, 2018; Hargreaves, 2012; Nale, 2013). Although high numbers of lytic and temperate *C. difficile* phages have been successfully isolated and observed by the use of electron microscopy (EM) (Figure 1.7), only several bacteriophages have been fully sequenced, annotated and become available in public databases in NCBI (Goh *et al.*, 2007; Mayer *et al.*, 2008; Govind *et al.*, 2006; Sebaihia *et al.*, 2006; Stabler *et al.*, 2009b; Horgan *et al.*, 2010; Sekulovic *et al.*, 2011; Hargreaves, 2012; Meessen-Pinard *et al.*, 2012; Alyousef, 2013; Nale *et al.*, 2016).

1.2.8 Temperate phage and virulence

The relationship between bacteriophages and emergence of virulent *C. difficile* strains has been investigated. Researchers have studied the effect of bacteriophages on bacterial evolution and suggested that the temperate phages are regarded as the main reason in prevalence and diversity of *C. difficile strains* (Brussow and Hendrix, 2002; Nale *et al.*, 2012). Brussow et al., 2004 showed that the presence of a prophage can induce increased expression of toxin genes in certain bacteria, for instance, production of the cholera toxin in *Vibrio cholerae* and the Shiga toxins in *E. coli* O157:H7 (Merabishvili *et al.*, 2009)

Previous reports have shown that temperate bacteriophages play a distinct role in the evolution and virulence of bacterial pathogenicity through involving their genetic material and integrated it into the chromosome of the host cell (Brussow *et al.*, 2004; Brown, 2006). Moreover, during the lysogenic cycle, this incorporation eventually may lead to the possibility of modifying some of genes, which subsequently will affect the phenotype of the bacteria (Brussow *et al.*, 2004).

phiCDHS1 is one of the most studied *C. difficile* phages (Hargreaves, 2012). It belongs to Caudovirales order and Siphoviridae family which have larger capsid and long tail length. The Siphovirus (SV) CDHS1 has capsid with size ~60 nm (Figure 1.7) (Hargreaves, 2012). The size of dsDNA genomes of phiCDHS1 is 41,474 bp (Figure 1.8). Previously, several studies shown that phiCDHS1 temperate phage can effectively kill the widespread epidemic strain of *C. difficile* R027 in bacterial liquid culture (Shan *et al.*, 2018; Thanki, 2016).

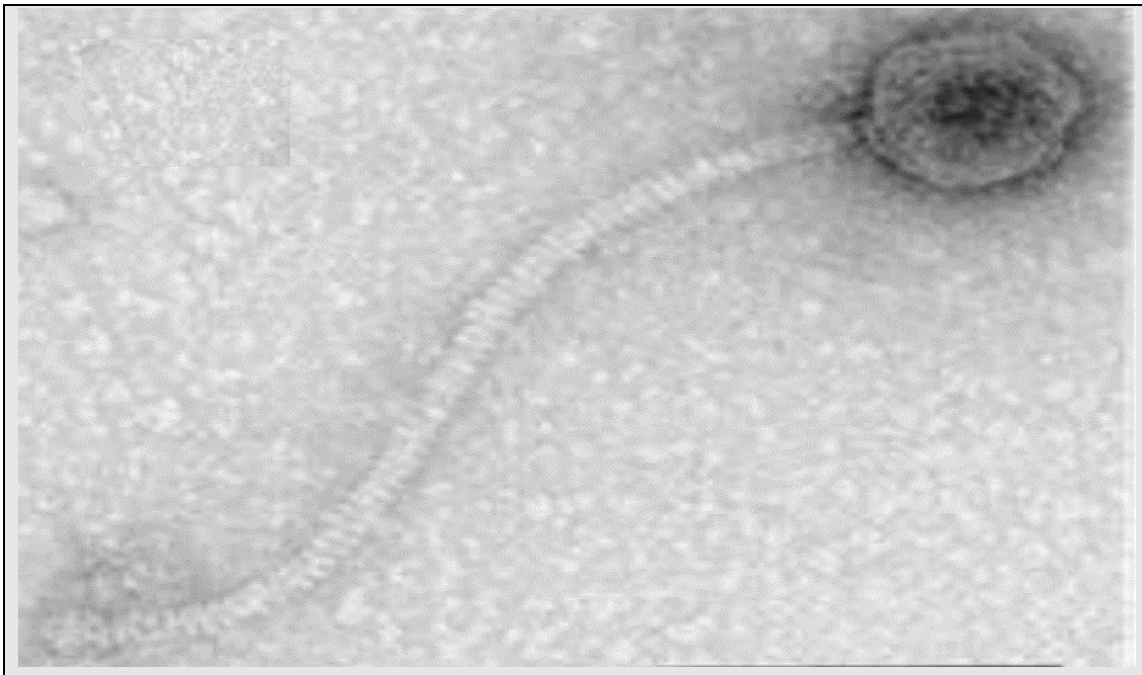


Figure 1.7. Morphology of purified CDHS1 phage particles with lytic activity against R027 *C. difficile* host strains (Hargreaves, 2012).

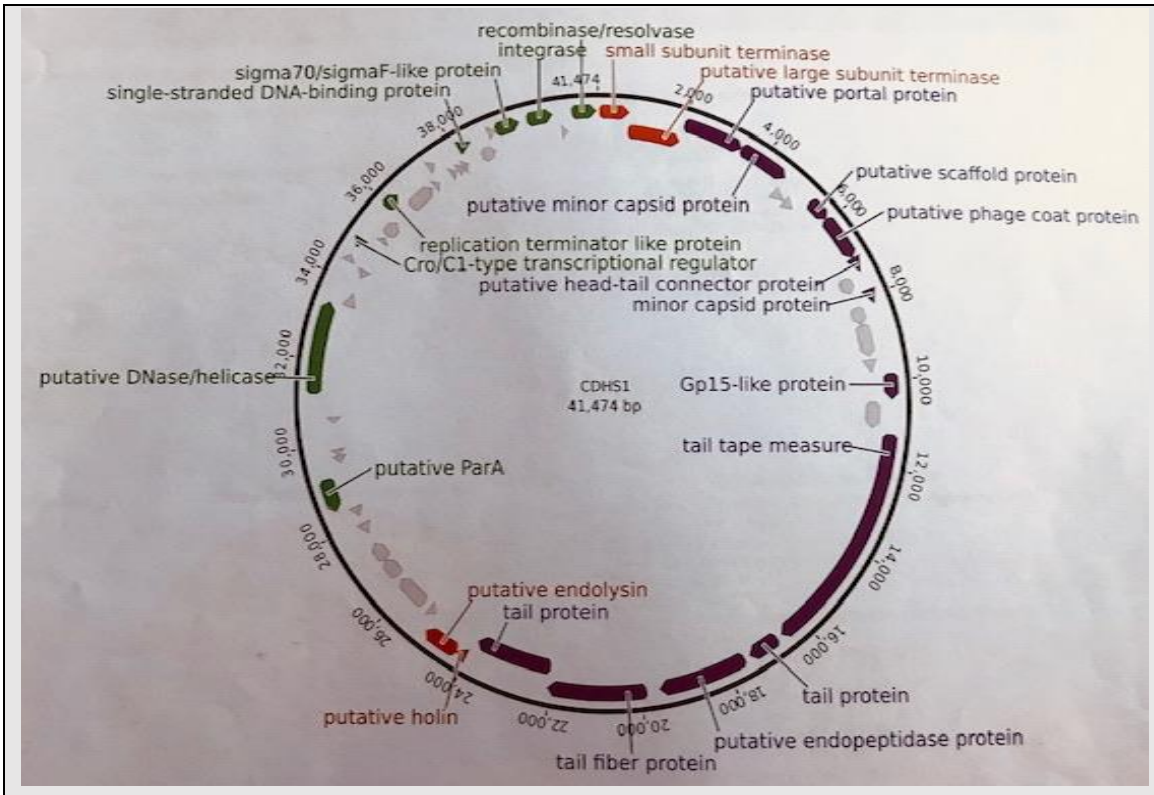


Figure 1.8. Circularized genomic map of bacteriophage PhiCDHS1.
The 41,474 bp genome assembled as a circle, shown here oriented to start at the terminase small subunit gene. Only the genes with predicted functions are indicated and suggest that approximately 55% encode hypothetical proteins.

1.3 Gene expression regulation in bacteria

1.3.1 Introduction

Gene expression is the process by which a protein can be synthesised based on the genetic information. Specifically, mRNA is transcribed using genetic codes on the DNA molecule as a template and translate these genetic codes to produce a protein (Figure 1.9). Transcription in bacteria is done by a single type of RNA polymerase, which needs to two important parts for initiation. First, the Pribnow box, a specific sequence (*TATAAT* of six nucleotides) is an essential part of a promoter site on DNA (Schaller *et al.*, 1975; Pribnow, 1975) by which RNA polymerase can recognise and bind to start transcription. Second, a sigma factor (σ factor) protein which is needed only for initiation of RNA synthesis by enabling specific binding of RNA polymerase to gene promoters (Gruber and Gross, 2003).

Translation is the process by which a polypeptide chain of protein is synthesised by translating the genetic code on the messenger RNA molecule as its guide (Slonczewski *et al.*, 2013) .

There are three ways (levels) by which the production of genes expressed is regulated (Figure 1.10). Firstly, transcriptional regulation in which genetic regulatory proteins can either affect the ability of RNA polymerase to bind to the promoter, terminate transcription prematurely by process of attenuation, or by binding of a metabolite to a riboswitch in mRNA (Slonczewski *et al.*, 2013).

Gene expression is controlled by system called gene regulation. It takes a lot of energy to make RNA and protein. Therefore some genes active all the time because their products are in constant demand. Other genes are turned off most of the time and are only switched on when their products are needed. Bacterial cells are extremely selective about what genes are expressed, in what amounts, and when. To achieve this selectivity, bacteria need to specific system of regulation to control gene expressed.

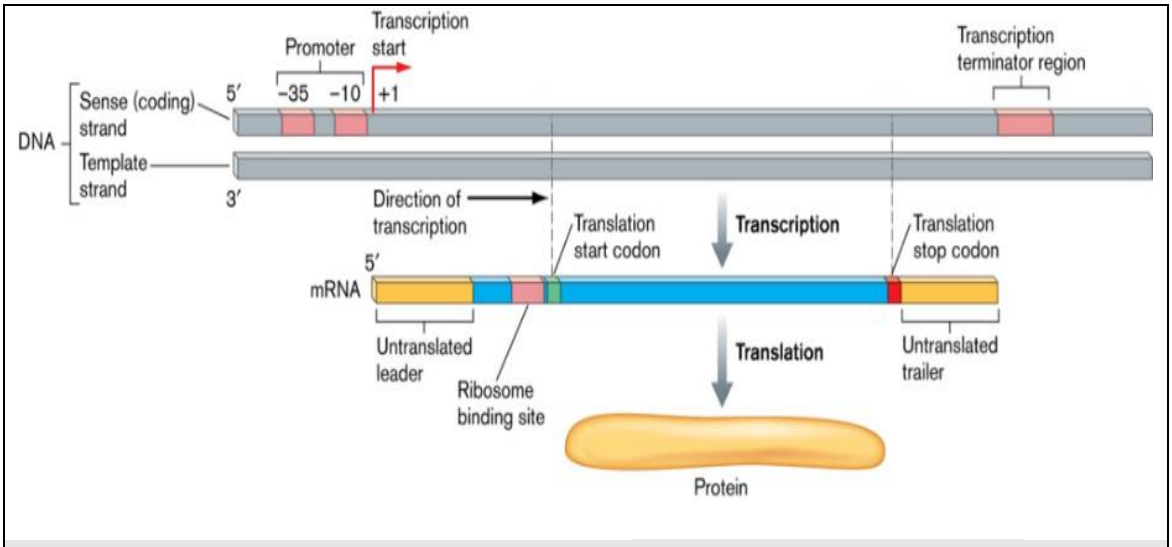


Figure 1.9. Gene structure and expression in prokaryotes (Pribnow, 1975).

Secondly, the gene encodes a protein can be controlled at the translational level which is called translational regulation function (Willey, 2014). In this regulation, various translational repressor proteins can bind to the mRNA and affect the rate of translation and lifetime of an mRNA transcript. Furthermore, translation can be controlled by binding of antisense RNA to the mRNA molecule and decide whether or not translation begins. Finally, post-transcriptional or post-translational regulation in which gene products (RNA and proteins) can be regulated after they are completely synthesized by degradation. In term of regulation of activity of the protein product, there are small molecules that can bind covalently and noncovalently to a protein and change its function such as phosphate and carbohydrate residues (post-translational modification) (Willey, 2014).

In addition to levels of control of gene expression, some bacteria can randomly or programme alter their DNA sequence to activate or repress a particular gene. For example, reversible overturning of DNA segment of cell-surface proteins to elude the host immune system by switching off or switching on of gene expression. Also, there is another level of regulation coordinating by controlling on mRNA stability. In this circumstance certain mRNA molecules can be degraded by an RNase activity (Slonczewski and Foster, 2013).

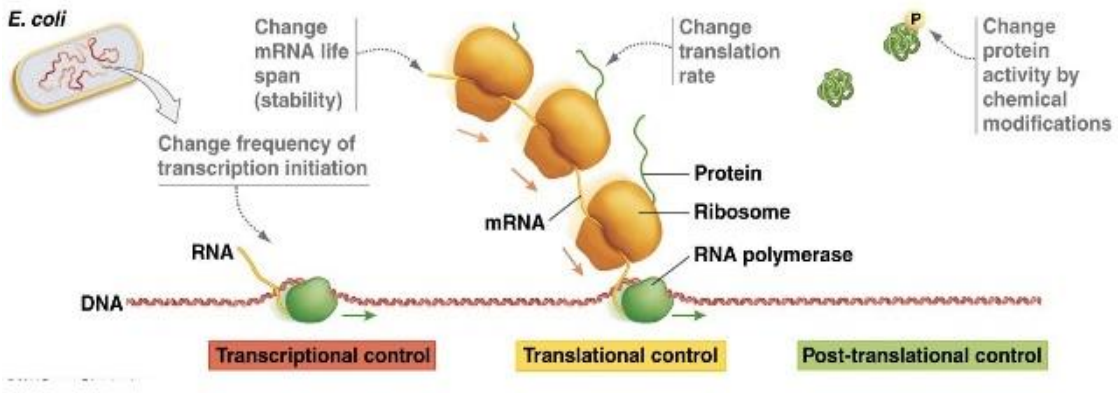


Figure 1.10. The levels of regulation in bacterial gene expression.
The image shows three different potential modes of regulation of gene expression in bacteria which include the following levels: Regulating the rate of transcription (transcriptional control), regulating the processing of RNA molecules and the stability of mRNA molecules (translation control) and regulating the rate of translation (post-translational control).

1.3.2 Gene expression analysis

Gene expression is the process by which genetic information (DNA) is used to synthesize gene products through the messenger RNA (mRNA). Most of these products are usually proteins, which go on to perform essential functions as enzymes, hormones and receptors. Whereas non-protein products which are coded by non-protein coding genes will be produced as functional RNA products such as ribosomal RNA (rRNA) or transfer RNA (tRNA). Gene expression analysis typically involves the isolation or capture of transcribed RNA within a sample, followed by its amplification and subsequent detection and quantitation of initial RNA isolated (Metzker, 2010; Kroghsbo et al., 2013). In the last decade, the study of gene expression was restricted to quantitative PCR (qPCR) analyses of candidate genes or relied on cross-species hybridization on microarrays. Recently, after the revolution in biological and diagnostic fields, gene expression analysis can be achieved through next generation sequencing technolog (Metzker, 2010; Kroghsbo *et al.,* 2013). With next generation sequencing technologies such as high-throughput RNA sequencing (RNA-seq), an increasing number of complete and draft bacterial genome sequences are now becoming available in public databases. RNA-Seq (RNA-sequencing), is also called whole transcriptome shotgun sequencing or high-throughput RNA sequencing (Morin *et al.,* 2008). RNA-Seq technology has

revolutionised the approach to transcriptome profiling that used to measure the precise value of the presence and quantity of RNA in biological sample at a specific time moment (Chu and Corey, 2012; Wang *et al.*, 2009).

1.3.3 Transcriptomics

The term "transcriptome" was first used in the 1990s (Pietu *et al.*, 1999; Velculescu *et al.*, 1997). Previously, various technologies have been developed to deduce and quantify the transcriptome. The first attempt to study on whole transcriptomes was on human brain in which involved capturing a partial transcriptome (about 609 mRNA sequences) and was first published in 1991 (Adams *et al.*, 1991). In the mid-1990s and 2000s, the transcriptomics technological field witnessed a significant development which made more it available and so become a widespread discipline in the biological sciences (Wang *et al.*, 2009; Nelson, 2001). Particularly in 2008, there was published another two human transcriptomes, composed of millions of transcript-derived sequences covering 16,000 genes, (Pan *et al.*, 2008; Sultan *et al.*, 2008) and by 2015 new one transcriptome per hundred individuals was being published (Mele *et al.*, 2015; Lappalainen *et al.*, 2013).

1.3.4 Transcriptomic contemporary techniques

Quantitative PCR (or qPCR), Microarrays and RNA-Seq are the most dominant contemporary technique assays for in-depth gene expression analysis in recent years (McGettigan, 2013; Nelson, 2001). Quantitative PCR (qPCR), or real-time PCR, and reverse transcription PCR (RT-PCR) are similar to the conventional PCR and are based on amplification of DNA by the polymerase chain reaction (PCR). The technique involves determination of exact amounts (relative or absolute) of amplified DNA in samples in real time which is lost in PCR (an end-point detection). PCR is considered the best choice to analyse the expression of a limited number of genes (Bustin *et al.*, 2009; Deepak *et al.*, 2007).

Several comparative analysis studies in the transcriptomic field revealed that RNA-Seq has a clear advantages in relative to hybridization-based approaches (Microarray) (Zhao *et al.*, 2014). Indeed, RNA-Seq is more sensitive in detecting very low expressed transcripts and more accurate in detecting expression of extremely abundant genes in comparison to microarray, which is based on expression quantification(Illumina, 2011). Moreover, RNA-Seq has the ability to quantify differential expressions at a large dynamic range of expression levels even with organisms lacking a reference genome by perform *de*

novo transcriptome assembly and differential expression analysis, and also to examine DNA variations (SNPs (Single nucleotide polymorphisms), insertions, deletions). In addition, RNA-Seq experiments can investigate both known transcripts and explore novel ones while by microarray technology only detects transcripts that correspond to the existing genomic sequencing information (Zhao *et al.*, 2014).

Taking all of these technologies into account, RNA-Seq is the first sequencing-based method that confers a survey of the entire transcriptome in a very high-throughput and quantitative manner (Mutz *et al.*, 2013). However, despite the benefits of the RNA-Seq, it is still more expensive than microarrays (Rao *et al.*, 2018; Zhao *et al.*, 2014).

1.3.5 The principle of RNA-Seq

RNA-Seq, also called whole-transcriptome shotgun sequencing, refers to the use of high-throughput sequencing technologies for characterizing the RNA content and composition of a given sample (Wang *et al.*, 2009). Briefly, In RNA-Seq technology, as transcripts cannot be used as a whole so it needs into randomly decomposed into short reads of up to several hundred base pairs before being converted into cDNA by reverse-transcribed. Thereafter adapters are ligated to each end of the cDNA. Sequencing can be done either unidirectional (single-end sequencing) or bidirectional (paired-end sequencing) and then directly aligned onto a reference genome database or assembled to obtain *de novo* transcripts in the case where transcript or genome information unavailable, proving a genome-wide expression profile (Wang *et al.*, 2009).

1.4 Transcriptomic studies

1.4.1 Studies conducted on bacterial transcriptome

Over the last two decades, numerous studies have focused on the impact of bacteriophages on mRNA of their bacterial host during infection (Zegans *et al.*, 2009; Zhao *et al.*, 2016; Grunberg-Manago, 1999). These studies have been conducted through genome-wide transcriptomic analysis which gives an opportunity to understand the information about expression of specific gene within whole host genome as well as their expression at specific time (Mojardin and Salas, 2016).

In principle, bacteria can modulate their gene expression in response to phage infection. One such method is the overexpression of some genes such as *ddl* and *dltABCD* as was conducted in *Lactococcus lactis* (Fallico *et al.*, 2011). It was also reported that certain genes encoded by *Escherichia coli* are expressed during bacteriophage PRD1 infection

(Poranen *et al.*, 2006). Genes such as *phoH, phoBR, phoA, pstSCAB,* and phoU are involved in the uptake and biosynthesis of arginine and methionine and the transcriptions of the genes are up-regulated in response to phosphate starvation (Poranen *et al.*, 2006). Similarly, the expression of several *nar* (nitrate reductase) genes encoded by *Pseudomonas aeruginosa* are up-regulated in response to environmental changes has been reported to be higher during phage infection (Ravantti *et al.*, 2008). On the hand, genes encoded by bacteria may be down-regulated in response to during viral infection. For instance, Multiple genes in *Lactococcus lactis*, responsible for nucleotide biosynthesis, amino acid metabolism and respiration proteins, was shown to be down-regulated during the Tuc2009 phage infection (Stuart Ainsworth *et al.*, 2013). Although it has been demonstrated that transcription of the (ccoO) gene encoding cytochrome oxidase was up-regulated in uninfected cells, *P. aeruginosa* is negatively regulated by Phage PRR1 Infection for this gene (Ravantti *et al.*, 2008). In another study, (Schachtele *et al.*, 1972)), showed that certain prophages can express alternative sigma factors to stimulate biofilm formation and inhibit sporulation, thereby enabling *Bacillus anthracis* to survive and grow in various harsh conditions.

Previous research has shown that some bacteriophages like phiCD38-2 regulate the expression of cell wall protein such as CwpV, which is highly conserved across *C. difficile* strains, and causes up-regulation of this gene to approximate 20-fold in the ribotype 027 lysogene (Sekulovic and Fortier, 2015). Furthermore, a number of the regulatory genes of *C. difficile* that encode transcriptional regulators and phosphotransferase system (PTS), which are involved in the uptake and metabolism of glucose, fructose, and glucitol /sorbitol, were also up-regulated the expression in the presence of the phiCD38-2 prophage (Sekulovic and Fortier, 2015). Interestingly, the expression of the number of *C. difficile* genes, in particular, which encode carbon metabolism and sugar transport, were down-regulated in infection (Sekulovic and Fortier, 2015).

1.4.2 Research conducted on the *C. difficile* Transcriptome

Despite the difficulties accompany working with *C. difficile*, recent advances in the field of microarray technology have enabled the ability to monitor host cell responses to viral infection at the level of mRNA synthesis in the infected and uninfected bacterial cell. More recently, RNA-seq has become the method of choice for measuring gene transcription in *C. difficile*. Several of the previous studies that employed microarray

analysis in host gene expression provided only a general picture and limited information about changes induced by viral infection. Currently, microarray analyses can detect the impact of specific gene transcripts of a phage within an infected host and at the same time has the ability to test any change at the whole host genome level (Duplessis *et al.*, 2005; Frye *et al.*, 2005). Chen et al 2005 reported the changes that occur in the whole host transcripts during lysogenic stage by using RNA- seq technology (Chen *et al.*, 2005).

In terms of the use of advanced sequencing technique to track changes during infection, there are only a few studies covering the entire viral replication cycle and whole-host gene expression of *C. difficile* at a single gene level and during one step growth. Previous genomic studies of *C. difficile* 027 strains showed that they encode mobile genetic elements including prophages and pathogenicity locus and SPI2 (Lyerly *et al.*, 1988; Voth and Ballard, 2005). A recent study has determined the global transcriptional changes of *C. difficile* carrying prophage by using high-throughput RNA sequencing during stationary phase (Sekulovic and Fortier, 2015).

1.5 Hypothesis of this study

Temperate phage can induce similar gene expression changes in *C. difficile* as virulent phage via inhibiting the expression of virulence genes during the lytic cycle.

1.6 Aim and objectives

The main aim of this project is to assess the hypothesis that infection with bacteriophage can affect the transcriptome of *C. difficile*. The objectives were:

- To detect differentially expressed (DE) genes between *C. difficile* infected with φCDHS1 bacteriophage and an uninfected control at each stage post-infection.

- To determine the effects of φCDHS1 bacteriophage on transcriptional responses of virulence genes of *C. difficile* at each stage post-infection in vitro.

- To evaluate the effects of temperate phage on the gene expression of *C. difficile* using RNA-seq and qPCR technique.

- To determine the pathogenicity of lysogenic, resistant to phiCDHS1 and wild type *C. difficile in vivo*.

Chapter 2
Materials and methods

Chapter 2. General Materials and methods

This chapter describes the general protocols and reagents used in this study.

2.1 Media, buffers and solutions preparation

All media, buffers and solutions used in this study are described in (Appendix 1 and Appendix 2) For all media and buffers preparations, Prestige Medical Autoclave, England was used to autoclave them at 121ºC, 15psi (1 bar) for 15 minutes.

2.2 Bacterial inoculum preparation and growth conditions

2.2.1 C. difficile source and Routine culture

Clostridium difficile strains R20291 and AIU (ribotype 027) were kindly provided by Prof Martha Clokie and Dr. Janet Nale.

C. difficile strains were from confirmed preserved cryogenic stock cultures (described below). These were routinely streaked and grown on blood agar plates and incubated at 37°C for 48 h under anaerobic condition (5 % CO2, 10 % H2 and 85 % N2) (Don Whitley Scientific, West Yorkshire, UK).A single colony from blood agar plate culture was sub-cultured into 5 ml of pre-reduce broth inside a bijou tube (Sterilin, UK) and incubated in an anaerobic chamber overnight (8-16 hours) at 37ºC. Pre-reduction was performed to eliminate the oxygen from the media by incubation of media anaerobically overnight. This Fastidious anaerobe broth (FAB) culture was used for spot test (CDHS1 phage titration), for plaque assays (CDHS1 phage propagation), freezer cultures preparation and the one step growth curve.

2.2.2 Freezer cultures of C. difficile (cryogenic stock)

Freezer cultures of all strains were prepared as follows: 1 ml of an overnight (FAB) culture was placed in an Eppendorf tube and centrifuged at room temperature for 5 minutes at 13,000xg (Eppendorf, Germany). The pellet was re-suspended in 200µl of glycerol taken from Protect Bacterial Preservers vial (Abtec, 39 UK). The suspension was transferred back into the vial and gently mix and immediately stored as cryopreservation stocks at -80˚C.

2.2.3 *C. difficile* confirmation

After storage overnight at -80°C, a vial of frozen cells was tested to be sure that the frozen cells are still viable, and to ensure the cryo-stock was free from contamination. The purity of the stock was checked by streaking one loop full of cryostock onto blood plates and

incubate for 48 h at 37°C in anaerobic conditions. The Gray/white, opaque, flat colonies colonies were examined under long wave UV light (365nm) to confirm their characteristic green-yellow fluorescence colour. Further confirmation of colonies has been done by using the *C. difficile* Test Kit (Oxoid, UK) based on latex agglutination according to the manufacturer's guidelines (Sharma and Vogel, 2014; Brazier, 1998a; Brazier, 1998b).

2.3 Propagation and Titration of bacteriophage

2.3.1 Plaque assay

To propagate the φCDHS1 on the double layer method (solid method), one colony of starter broth culture of *C. difficile* R20291/ribotype 027 was inoculated in pre-reduced FAB (Tubes of broth stored in anaerobic cabinet at 37°C previously to pre-warm the media and eliminate oxygen from it) and incubated at 37°C for 18-24 h. From the overnight FAB culture of *C. difficile*, 400 µl of the culture were mixed with 150 µl of phage stock titre into 10 ml of pre-warmed soft agar and salt buffer (ratio 1:1). The mixture was swirled and poured onto agar plates. The next day, the soft (top) layer containing phage was scraped off with a sterile L shaped spreader into 50 ml falcon tubes and left at least 5 hours at 4°C but preferably overnight to allow phage to dissociate from the soft agar. This was then centrifuged at 15, 000 x g for 15 minutes. The supernatant was filter-sterilized through a 0.22µm filter (Millipore, USA) into a sterile tube and stored at 4°C before the phage enumerations was performed by spot tests as described in section 2.3.2.2.

2.3.2 Bacteriophage titration by spot test

2.3.2.1 Preparation of media

The double agar layer method was used which consists of a semi-solid medium of soft agar and salts (Mahony *et al.*, 1985; Goh *et al.*, 2005b). Double concentrated soft agar was prepared with 74 g/l BHI and 1 % w/v of bacteriological agar (Mahony *et al.*, 1985). Double concentrated salt buffer containing 0.8 M $MgCl_2$ (Fisher Scientific, UK) and 0.2 M $CaCl_2$ (Argos Organics, UK) was prepared in distilled water. The soft agar and salt buffer were autoclaved separately and stored at 55ºC and combined immediately before use in equal volumes to give a final concentration of 37g/l of BHI, 0.5g agar, 0.4 M $MgCl_2$ and 0.1 M $CaCl_2$.

2.3.2.2 Spot test

The bacteriophage CDHS1 was detected by the spot test to determine their titre levels. BHI 1% agar plates (using 90 mm size plates) were prepared and marked into sections on the bottom lid. Approximately, 3 ml of soft agar and salt buffer which salt buffer (ratio 1:1) was added to a bijou vial containing 250 µl of overnight *C. difficile* culture in FAB. After inverting, the mixture was poured onto the BHI agar plate and left to set for 5 minutes. A 1/10 dilution of the phage lysate was prepared in BHI broth. Afterwards 10 µL samples of the enrichment supernatants of propagated phage previously were spotted on the surface of the prepared lawns, making sure to test every supernatant with all *C. difficile strains*. Drops were allowed to dry for around 10 minutes and then the plates were incubated aerobically at 37ºC overnight, after which they were all checked for complete lysis and plaques (Clokie and Kropinski, 2009).

2.4 One-step growth curve of CDHS1 phage

A one-step growth experiment was conducted according to technique reported previously (Ellis and Delbrück, 1939; Hyman and Abedon, 2009). This method was used to determine the early-log, mid-log and stationary phases of the phage replication (Luo *et al.*, 2016). Briefly, 1 ml of overnight *C. difficile* FAB culture (Appendix 1) was transferred to 100 ml pre-reduced BHI broth and incubated anaerobically until an OD_{550} (0.2 ± 0.02) (early logarithmic phase). The culture was then infected with phage at a Multiplicity of Infection (MOI) of 10. To calculate the multiplicity of infection (MOI), the following equation was used:

$$\text{MOI (for equal volumes)} = \frac{\text{Number of phages}}{\text{Number of bacteria}}$$

After inoculation phages were left to adsorb to *C. difficile* cells for 15 minutes in order to promote phage-host adsorption and during this time the phage infected culture was gently mixed at regular intervals. The mixture of the host and phage was pelleted by centrifugation at 3000 xg for 5 min to remove any free (non-absorbed) phage particles, only washed once. The re-suspended pellets in fresh pre-reduced BHI broth were incubated at 37°C. Aliquots of 10 ml of the host/phage suspension was taken at 10 min interval at time points over 100 min and the phage titre was determined by the spot method. Then, 1 ml was immediately used for OD_{550} reads and to calculate CFU/ml count. At the same time, 1 ml PFU/ml count was immediately conducted on a lawn of *C. difficile* on double agar method which is described in (2.3.2.2) (Miles *et al.*, 1938). For the PFU/ml

count the suspension was centrifuged at 15,000 xg for 2 min and filtered through a 0.22µm Millipore filter prior to spotting on a lawn of *C. difficile*. Phage enumeration was conducted in duplicate and the growth curve was repeated three times to obtain three technical repeats. The average of plaques number was taken at each time point to draw phage growth curve.

2.5 The growth curve of *C. difficile*

To determine the exponential growth phase starting and stationary phase, starter culture was prepared by inoculating a pure colony of *C. difficile* strain from BHI blood plate (48 h) to 5 ml of pre-reduced FA broth. The culture was incubated anaerobically overnight at 37 °C. After incubation, 0.5 ml of bacterial culture was inoculated into 50 ml pre-reduced BHI. Growth was determined by measuring OD_{550} at 0, 1, 2, 3, 4, 5, 6, 7, 8, 12 and 24 h. Also the aliquots were evaluated for bacterial levels by counting colony forming units per ml (CFU/ml). All growth curves were performed in triplicate.

2.6 Gene expression assay

To detect and quantify total RNA as an indicator of the gene expression at the transcriptomic level, RNA extractions were carried out.

2.6.1 RNA Extraction

2.6.1.1 Bacterial and phage growth conditions

C. difficile R20291, ribotype 027 (wild type) was grown in BHI broth and incubated until the OD600nm reached 0.18- 0.2. Then approximately 10 ml of CDHS1 phage (10^9 PFU/ml) (MOI=10) was added to the culture (100 ml) which was incubated for 20 min to allow phage-bacteria adsorption. The culture was centrifuged at 3000 xg, for 5 min and the pellet was re-suspended with 100 ml of pre-reduced BHI broth (washed once) and incubated for 100 min at 37 °C. A control sample of *C. difficile* R20291 was treated in exactly the same way but pre-reduced BHI was used rather than phage. Six aliquots with 10 ml volume were removed at each 10 min intervals from *C. difficile* R20291 and *C. difficile* R20291 infected with phage for RNA extraction step.

2.6.1.2 RNA extraction steps

RNA extraction was performed using Trizol reagent (Stewart *et al.*, 2002) in which 8 ml of *C. difficile* R20291 (wild type) and *C. difficile* R20291 infected by CDHS1 bacteriophage (aliquots from above step) were immediately centrifuged at 4000 rpm for 10 min at 4°C. The major part of the supernatant was discarded, and the bacterial pellet

was re suspended with 500 µl of ice-cold TRIzol reagent (Invitrogen), 100 µl of chloroform was added and vortexed for 15 sec. The bacterial suspension was then transferred into lysing matrix B tubes containing 0.5 g of acid-washed glass beads (106 micron; Sigma) and sonicated at a power setting of 6.5 in a PowerLyzer^{TM} 24 homogeniser (MO BIO, USA) for 45 seconds. The treatment was repeated after a 15 second cooling pause. After incubation for 1-2 min at room temperature, the samples were centrifuged at 12,000 xg for 15 min at 4°C. To precipitate the RNA, the upper aqueous phase containing the RNA was transferred into a 1.5 µl fresh Eppendorf tube, 250 µl isopropanol was added, and the suspension was vortexed for 15 sec and then left at room temperature for 15 min. The tubes were centrifuged at 12,000 xg for 10 min at 4°C and the supernatant removed. The pellet was washed with 500 µl 75% ethanol and re-centrifuged for 10 min at 12,000 xg at 4°C. The supernatant was discarded and then the pellet was air-dried for 15 min, and was solubilized in RNase-free water. Total RNA was stored at -80°C for further use. After each extraction, a small fraction was taken to check the quantity and integrity of the total RNA using a NanoDrop 1000 (Thermo Scientific, UK) at 230 nm, and by 1% (w/v) agarose gel electrophoresis.

2.6.2 Analysing RNA using Agarose gel electrophoresis

2.6.2.1 Gel Electrophoresis of RNA samples

Agarose gel electrophoresis was prepared using1% w/v dry agarose (Bioline) which was suspended in TBE buffer solution, boiled in a microwave until the solution became clear. The solution was allowed to cool and 0.2µg/mL of ethidium bromide was added, and then poured it into a casting tray. All RNA samples were mixed with 6x gel loading dye and heated at 65°C for 10 min before loaded into the wells. Electrophoreses was carried in 1x TBE buffer at 120 V for approximately 1 h. A long wave UV transilluminator was used to visualize RNA bands. The size of the RNA bands was determined by using the DNA ladder (1kb, New England Biolab).

2.6.2.2 DNase Treatment

Residual chromosomal DNA were removed from extracted nucleic acid using the Ambion® DNA-free™ DNase Treatment kit (Ambion, UK). This kit is designed to remove trace to moderate amounts of contaminating DNA (up to 50 µg DNA/mL RNA) from purified RNA to a level that is mathematically insignificant by RT-PC. Briefly the following steps were carried out:

1- **DNA Digestion:** A 0.1 volume of 10X DNase I Buffer and 1 µL rDNase I were added to the RNA, mix gently, and incubated at 37°C for 20–30 min.

2- DNase Inactivation: DNase Inactivation Reagent (typically 0.1 volume) was added and mix well and incubate 2min at room temperature, mixing occasionally.

3- Centrifugation: Spin it at 10,000 × g for 1.5 min and transfer the RNA to a fresh tube and stored at -80°C until further use.

2.7 Preparing qPCR reactions

2.7.1 Construction of complementary DNA (cDNA)

Synthesis of cDNA from purified total RNA was performed into three reactions according to the manufacturer's instructions using First Stand cDNA synthesis kit (Invitrogen, UK) protocol. Firstly, 1µl (0.5) µg/ml of total DNase treated RNA mixed was mixed with 1µl of random hexamer primers (Invitrogen) and 10 µl DNase-RNase free water (plus water to final concentration of 12 µl). The mixture was heated to 65°C for 5 min. Then 4µl of 5X First-Strand Reaction Buffer, 1µl of RiboLock RNase Inhibitor (20U/1µl), 2µl of 10mM dNTP and 1µl of Revert Aid M-MuL V RT (200U/µl) were added to the mixture. Secondly, a no-template control reaction was set up using all reagents in the kit without RNA sample to assess DNA contamination of reagents.

Thirdly, A positive and negative control RNA template (GAPDH) that are supplied with the kit were also set using all reagents in the kit with and without RNA template (Control GAPDH RNA) in positive and negative reaction respectively.

Finally, all three reactions were set to final volume of 20 µl by addition of H_2O and the mixtures incubated for 5 min at 65°C, followed by 60 min at 42°C. For termination, the reactions were heated at 70 for 5min.

2.7.2 Guide to Designing qPCR Primers

According to (Thornton and Basu, 2011), many parameters were conducted before designing primers for qPCR to ensure that the method will be performed correctly as following:

- Primer Length: min length: 18, max length 24 (best: 20 nt).
- Product size: for efficient amplification in real-time RT-PCR, amplified region must be no bigger than 200 bp; usually 80-150 bp.
- Melting temperature: the ideal melting temperature is 60°C (57°C - 63°C).

- GC content: Optimal primers also should have a GC percent of around 50-60% to ensure maximum product stability.

2.7.3 Virulence Genes of *C. difficile* R20291 and primer results

The expression levels of a number of genes were investigated using qPCR. Ten virulence genes were chosen from different papers (Darkoh *et al.*, 2016).Two candidate reference genes including 16s ribosomal RNA and Recombinase A. (16s rRNA and *recA*) were selected from commonly described housekeeping genes in other bacteria (Rocha *et al.*, 2015; Metcalf *et al.*, 2010). The selection of primer sequences was determined using the Primer3Plus (http://primer3.ut.ee/) and NCBI primer design programmes (https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). Specificity of the amplification product for each primer pair was verified by the presence of a single amplicon at the expected size in a 1% agarose gel electrophoresis. The list of virulence genes including primers, their sequence, expected amplicons is available in Table 2.1.

Virulence genes					
No.	Gene	locus_tag	Product	Primers (forward& reverse)	Amplicons size
1	tcdA	CDR20291_0584	Toxin A	F AGCTTTCGCTTTAGGCAGTG R TGGCTGGGTTAAGGTGTTGG	128
2	tcdB	CDR20291_0582	Toxin B	F GGAAGGTGGCTCAGGTCATA R GTCCATCCTGTTTCCCAAGC	201
3	Binary toxin(cdt)	CDR20291_2491	actin-specific ADP ribosyltransferase	F ATGGGAAGGACAAGCACTGT R CCTGCATAACCTGGAATAGCTG	156
4	tcdE	CDR20291_0583	putative cell wall hydrolase protein	F AGGAGGCGTTATGAATATGACAA R TGCTACTTTTCTGATTCCTCCATCT	166
5	dtxA(tcdC)	CDR20291_0585	putative exported protein	F AGGAGGCGTTATGAATATGACAA R TGCTACTTTTCTGATTCCTCCATCT	105
6	tcdD(tcdR)	CDR20291_0581	putative transcriptional regulator	F GCAAGAAATAACTCAGTAG R CTGTTTCTCCCTCTTCATAATG	135
7	agrB	CDR20291_2640	Accessory gene regulator B	F TTCCAACCCAACTAACACTTGC R CAGGAGGGTATCATGCAGACAAT	147
8	agr	CDR20291_3187	accessory gene regulator	F TCTCACTTTCACTAAGAGGTTTGT R CAGTTCACAGGAGGTTATCATGC	208
9	spmA	CDR20291_3377	spore maturation protein A	F ACCATTTCCAGCACCCAACA R TGGCACTCTGGATGGGGATA	176
10	fliA	CDR20291_0270	RNA polymerase sigma factor for flagellar	F AAGTGATAGAGAAGAGGAAGCTC R TGAAACACCTAGCACTTTCCC	174
Reference(Housekeeping) genes					
1	16S rRNA	1087908..1089531	16S ribosomal RNA	F GATGGACCCGCGTCTGATTA R CGTAGGAGTTTGGACCGTGT	120
2	RecA	CDR20291_1169	protein (recombinase A)	F ACTGGAGGACGTGCACTAAA R GCTTGCTTAAATGGTGGTGCT	146

Table 2.1. List of PCR primers of virulence genes used in this study

2.8 Quantitative Real-Time PCR (qRT-PCR)

2.8.1 qPCR Master Mix for mRNA

qPCR was performed using 7500 Fast Real Time PCR system with Fast SYBR Green Master mix (Thermo Fisher Scientific, USA).

- The cDNA samples were diluted 1:100 with RNA-free H_2O.
- The qPCR Master Mix was prepared as follows:

  * SYBR Green 2x master mix: 10µl

  * Forward primer (5mM): 2.5µl

  * Reverse primer (5mM): 2.5µl

- 5µl of each cDNA sample was mixed with 15µl of Master Mix plus wate to a final volume of 25µl.

- A 7500 Fast Real-Time PCR System thermocycler device was used for quantitative real-time PCR.

2.8.2 Normalisation of gene expression data

The qPCR values for targeted genes were normalized to the housekeeping genes *16s rRNA* and *recA*. These reference genes were selected because it produced better reproducible results compared to *gyrA* (DNA gyrase A). Relative gene expression fold changes were measured using the following equation:

$$\text{Fold Change of target gene} = \frac{2\text{^ (Test ReferenceCt} - \text{Test TargetCt)}}{2\text{^ (Control ReferenceCt} - \text{ Control TargetCt)}}$$

(Livak and Schmittgen, 2001)

2.9 RNA-seq analyses

2.9.1 Transcript isolation and sequencing.

To investigate gene expression changes caused by φCDHS1 infection of, R20291 *C. difficile* gene expression during the one step growth curve (early, middle and late growth phases) was analysed using RNA-Seq technology; all RNA samples isolated were sent BGI Company for sequencing. Briefly, the obtained cleaned sequencing reads (FASTQ datasets) (0, 10, 20, 30, 40, and 50 minutes) for each sample which were generated from Illumina MiSeq were imported into Kallisto software for aligning to the genomes of both phage and host. Then, the resulting files were used as a new annotation file to perform

further gene analysis and biological interpretation by Partek Genomics Suite software (version 6.6 Copyright©; 2014 Partek Inc., St. Louis, MO, USA) (Wang, 2016) . Then after, the read counts that align to annotated gene features could be compared between samples according to different time points to statistically test for differences in expression.

2.9.2 Statistical and other analysis used in this study

A recent epidemic and hypervirulent 027 (R2029) *C. difficile* were sequenced and annotated and revealed a large circular chromosome of 4,191,339 bp (4.3 Mb), GC 28.8 using RAST server (Stabler *et al.*, 2009a; Aziz *et al.*, 2008). The whole genomes of *C. difficile* (R20291) strain is available in GenBank FN545816.1. Predictions of prophage sequences within the R20291 genomes were investigated using PHASTER, a new implementation of the PHAST (Arndt *et al.*, 2016). Primer design and validation were performed by the Primer3 Input platform, qPCR data were subjected to statistical analysis using GraphPad Prism 7. Relevant analysis tests were performed to compute the statistical significance at $P < 0.05$ as statistically significant.

Chapter 3
Results & Discussion

Chapter 3. Results and discussion

3.1 Growth curve of *C. difficile* phiCDHS1 analysis

One of the objectives of my work was to look at the differential expression of bacterial and phage genes during a one-step growth curve experiment. One-step growth curves were done via measuring the characteristics of one synchronised infection cycle of the phage. In these experiments cells are infected at a high MOI to ensure all bacterial cells were infected. Normally virus particles with MOI of 10 are added to cells at 37°C and allowed to adsorb for 20 minutes. Cells are then washed to remove free virus particles, returned to anaerobic incubation and replication cycle initiated. Because the MOI in this experiment is high almost all cells would have been infected simultaneously and the phage release shows good evidence of synchronicity (Kasman *et al.*, 2002).

In general, the growth cycle starts with an eclipse period which is the time from injecting nuclear material into the cell to the appearance of the first intracellular phages. No infectious phage particles can be found either inside or outside the bacterial cell. The eclipse or latent period ends with the appearance of intracellular phage particles. It is a period of intense biochemical activity including attachment, entry, replication, transcription, translation and assembly of progeny phages. The eclipse or latent period ends when new phage particles exit the cell.

The latent period and the burst size were determined for this virus using the single-step growth curve as described by Adams & Yordpratum et al. (Adams, 1959; Yordpratum *et al.*, 2011). The burst size is the number of phage particles released by a cell after one infection cycle and can be calculated as the ratio of the final count of liberated phage particles (maximum phage yield in the first burst) to the initial count of infected bacterial cells during the latent period. The phage burst per infected cell was calculated as the division of the phage burst size by initial count of infected bacterial cells CFU (colony-forming unit) / ml (Figure 3.1). Both phage burst size and phage released per infected cell were calculated using the below equations:

$$\textbf{Phage burst size/ml} = \frac{\textbf{Final count of phage paticles PFU/ml}}{\textbf{Initial count of infected bacterial cells CFU/ml}}$$

$$\textbf{Phages released per infected cell} = \frac{\textbf{phage burst size}}{\textbf{Initial count of infected bacterial cells}}$$

The overall time of the phiCDHS1 growth curve was applied every 10 minutes up to 100 minutes (Figure 3.1). Based on one-step growth, as shown in two figures (Figure 1.1. a, b), a triphasic curve, including the latent phase, log phase, and plateau phase, was obtained. Using these data, the latent period was determined to be about 20 minutes. It is in agreement with the Goh et al & Sekulovic et al who suggest that although only a few one-step growth curve experiments have been published, latency periods was highly varied (from ~30 min to 2 h) (Goh *et al.*, 2005b; Sekulovic *et al.*, 2011).

At the end of the latent period, it is followed by an increase in the number of virus particles prepared to release via host-cell lysis. The burst size of CDHS1 phage was ~37 virions per infected cell which is in the range reported for other *C. difficile* phages (Sekulovic *et al.*, 2011; Goh *et al.*, 2005b). Notably, for phages with short latent periods, their burst sizes will be smaller than phages which had longer latent periods, these corresponded to that previously observed (Goh *et al.*, 2005b; Thanki, 2016). However, burst size of temperate phage CDHS1 can be smaller than for temperate phages of other bacterial species (Hendry and Fitz, 1974; Ashelford *et al.*, 1999).

C. difficle R20291 was also used in a separate experiment as control. The lytic properties of CDHS1 temperate phage is clear between infected culture and un-infected culture (Figure 3.2). The observations shows expected results, the optical density of the bacterial solution infected with phage CDHS1 dropped rapidly from the start of the infection until it finally tailed off close to zero after 80 min in contrast to those bacteria grown in the absence of CDHS1. These result agrees relatively well with that from Cao et al (Cao *et al.*, 2015). This is strong evidence of the lytic ability of phage CDHS1 after incubation with *C. difficle* R20291 at 10 multiplicity of infection (MOI).

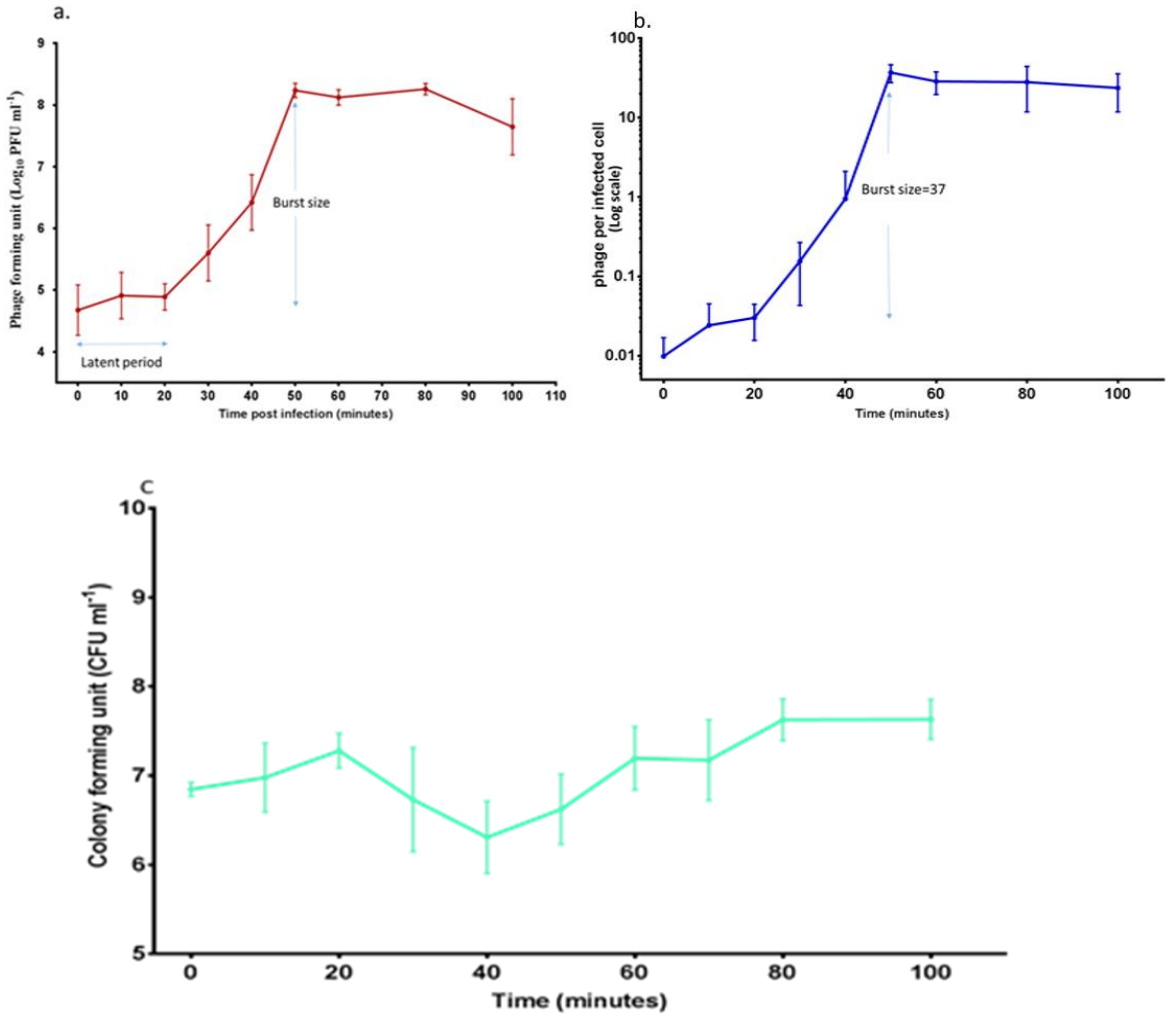


Figure 3.1. One-step growth of CDHS1 phage on *C. difficile* R20291 at a MOI of 10. A one step growth curve was carried out for 100 minutes and at regular 10 minutes intervals PFU/ml were determined. (a) Bacteriophages showed a latent period started at zero time and ended at 20 min of incubation and a log phase followed by a plateau phase at 50 minutes. (b) Based on the curve, the burst size of CDHS1 phage was ~37 phages per infected cell. (c) Mean of colony forming units (CFU) during infection was performed. For each data point mean of PFU/ml and CFU/ml were calculated from three biological replicates, with three technical repeats. The error bars indicate standard deviation.

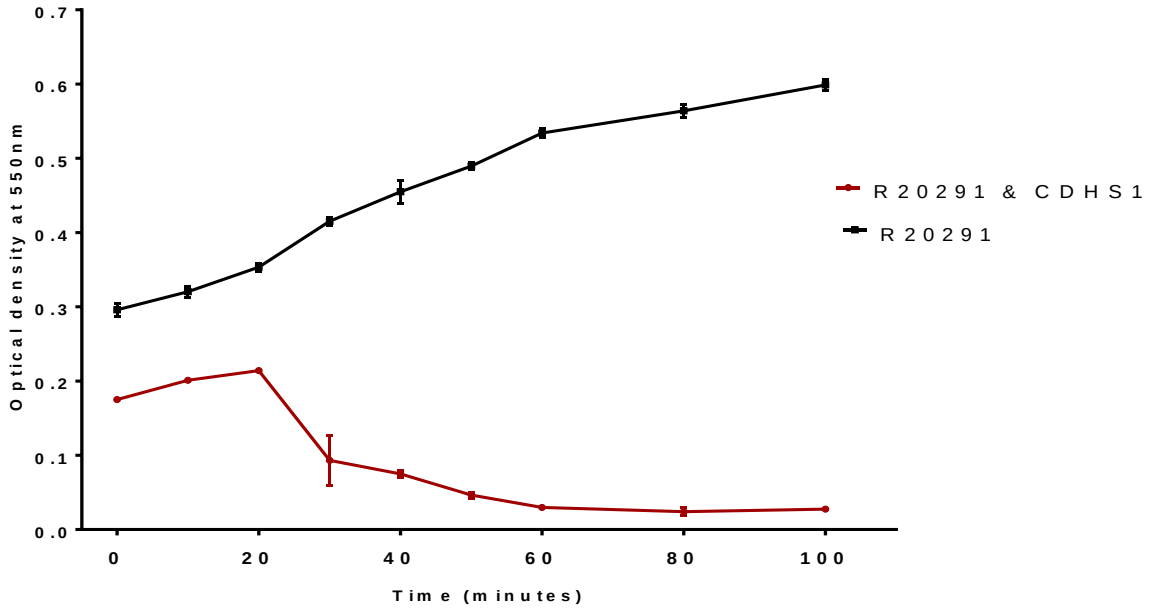


Figure 3.2. The Growth Curve of *C. difficile* at optical density (OD) 550 in the absence and presence of temperate CDHS1 Bacteriophage during one step growth curve. There is a clear drop in growth of infected bacteria compared to uninfected bacteria. For each data point mean of PFU/ml and CFU/ml were calculated from three biological replicates, with three technical repeats. The error bars indicate standard deviation.

3.2 Validation of RNA integrity

RNA integrity was determined using Nano-Drop OD260nm measurement and agarose gel electrophoresis.

3.2.1 Agarose gel electrophoresis

Electrophoresis of RNA extracted from infected *C. difficile* R20291 and uninfected with phi CDHS1 in a one step growth curve were done as described previously (2.6.2.1).

To obtain intact RNA without degradation, procedures to extract total RNA were performed more than once according to the method of Tedin & Blasi (Tedin and Blasi, 1996). Figure 3.3 A and B shows that the RNA bands of *C. difficile* R20291 (A) and *C. difficile* infected with phiCDHS1 (B) were inconsistent in terms of pattern. A possible explanation for these results may be the lack of a sonication step or glass beads which may be lead to lost RNA in the samples. Another possible explanation for this result may refer to the problems associated with RNA extraction using first protocol which was suggested for Gram negative bacteria such as *Salmonella*. These possibilities were investigated and the findings are findings in Figure 3-4 which show RNA extraction by using a second protocol which yields clearly comparable fragments. This result may be

explained by the fact that adding a sonication step yields more RNA. The present results in (Figure 3-4 A) which represent RNA from *C. difficile* R20291 are more significant than those with infected samples (Figure 3-4 B). It can explain this result, that it may be related to low concentration of RNA due to bacterial cell lysis after infected with bacteriophage. However, the consistent bands of RNA using the second protocol are important and encouraged to send these samples to RNA sequencing.

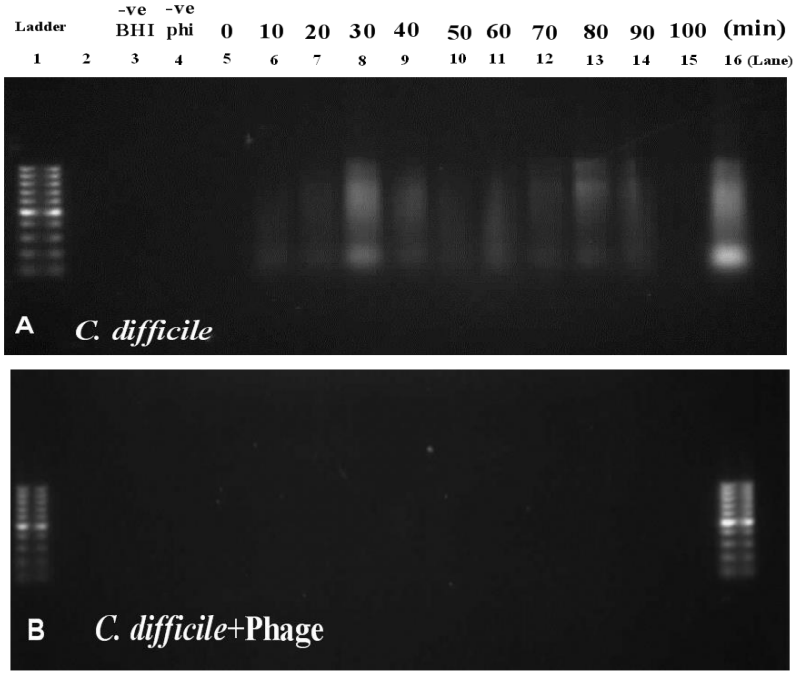


Figure 3.3. (A) and (B) Agarose gel electrophoresis of total RNA extraction procedure described in (2.6.2.1) from: *C. difficile* R20291 (A) and *C. difficile* infected with phi CDHS1 (B). For each gel, 10 microliters of 1kb ladder was loaded into lane one. Ten microliters (7 µl of sample +3 µl of loading dye) from *C. difficile* alone and *C. difficile* infected with phi CDHS1 (0, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 minutes) were loaded into lanes 5, 6, 7, 8, 9, 10, 11, 12,13, 14 and 15 respectively.

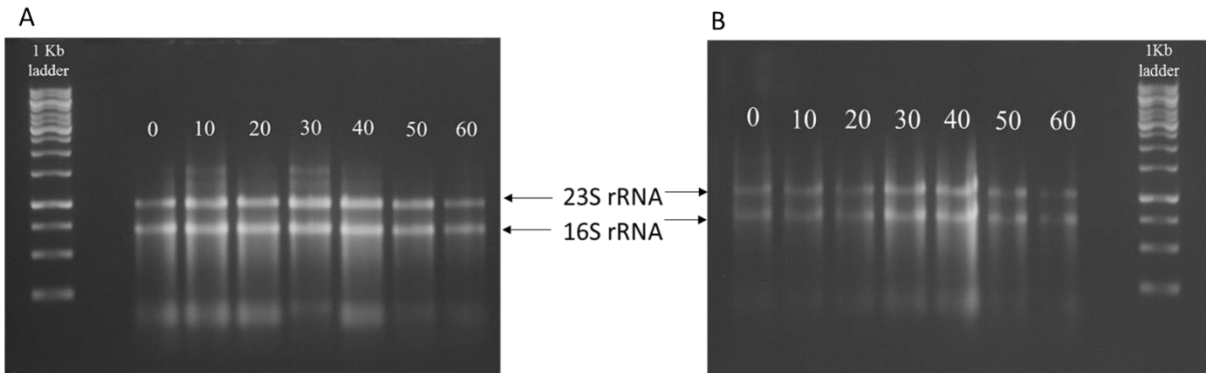


Figure 3.4. Agarose gel electrophoresis of total RNA extraction from *C. difficile* R20291. (A) and *C. difficile* infected with phi CDHS1 (B). For each gel electrophoresis, 10 microliters of 1kb ladder was loaded into lane one. Ten microliters (7.5 µl of sample +2.5 µl of loading dye) from *C. difficile* alone and *C. difficile* infected with phage CDHS1 (0, 10, 20, 30, 40, 50 and 60 minutes) were loaded in lanes of gels A and B. The arrows on both panels shows the clear two bands of 23S rRNA and 16S rRNA from *C. difficile.*

3.2.2 Measuring RNA integrity number (RIN)

To identify the integrity and concentration of extracted RNA, Nanodrop and Agilent RNA 6000 Nano-kit (Kidlington, UK) was used by the BGI Company. The purity of RNA was calculated by measuring the absorbance values 260/280 and 260/230. The concentration of RNA was measured by spectrophotometric quantification at 260 nm. All RNA test results are described in detail in Appendice 3 and 4.

Chapter 4
Transcriptional responses of *C.difficile* R20291 virulence genes induced by infection with CDHS1 bacteriophage *in vitro* using Real-time RT-PCR (qPCR)

Chapter 4. Transcriptional responses of *Clostridium difficile* R20291 virulence genes induced by infection with CDHS1 bacteriophage in vitro using Real-time RT-PCR (qPCR)

4.1 Introduction

Real-time RT-PCR is becoming a common tool for monitoring gene expression in biological research. The use of RT-PCR has regarded as a valuable technique to quantify mRNA during the process of infection and the abundance of phage or host gene transcripts (Clokie, 2009; Vandecasteele *et al.*, 2001). It has recently been used to detect expression as an alternative detection of bacteriophages to the plaque assay (Klopot *et al.*, 2017; Refardt, 2012). Additional applications have included detection of the medically important *Ehrlichia* species by amplification of a genus-specific disulfide bond formation protein gene (dsb) (Doyle *et al.*, 2005); monitoring the expression of the bacterial host during infection with bacteriophages (Poranen *et al.*, 2006; Ravantti *et al.*, 2008; Fallico *et al.*, 2011; Sekulovic and Fortier, 2015); and detection of the global host response *C.difficile* of in the presence of ΦCD38-2 prophage (Sekulovic and Fortier, 2015).

In a previous studies (Hacker and Ott, 1989; Wassenaar and Gaastra, 2001; Wagner and Waldor, 2002; Hernandez-Doria and Sperandio, 2018), they were stated that bacterial virulence properties can be altered by bacteriophages infection, and induced regulation of phage-encoded virulence factors. Herein, RT-PCR has been employed to quantify the expression of the virulence genes in *C.difficile* during infection with the CDHS1 bacteriophage.

4.2 Material and methods

The qPCR reactions were performed as described previously in section 2.8.

4.2.1 Validation of virulence genes expression data by RT-qPCR

In order to investigate the overall interactions between phage and host during all the stages of infection, Real-Time qPCR were performed to validate expression levels in 10 virulence genes that were thought might be responsive to phage infection at different time points in the growth cycle (early, middle and late phase) (Table 2.1). RT-qPCR can incidentally be used to validate the RNA-seq. Differential expression determined by RNA-seq (Eriksson *et al.*, 2003) was validated by real-time reverse transcriptase qPCR. The 10 virulence genes differently expressed were subject to the RNA-seq analysis. These genes were selected based on the hypothesis that some bacteriophages encode regulatory

factors that increase expression of virulence genes not encoded by the phage (Spanier and Cleary, 1980). Briefly, the significance of differential expression of virulence genes, relative to the control, was assessed by using SYBR Green (Sharkey *et al.*, 2004). In this validation, three independent RNA samples of *C. difficile* R20291 infected with CDHS1 Phage, plus a fourth independent sample of only *C. difficile* R20291, were subject to cDNA synthesis as described previously in section (2.7.1). The qPCRs were performed using 7500 Fast Real Time PCR system with Fast SYBR Green Master mix (Thermo Fisher Scientific, USA). Output data was analyzed using GraphPad Prism 6.

4.2.2 Validation of cDNA by agarose gel electrophoresis

Once the RNA quality has been checked by a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies,Inc.) and cDNA has been synthesized, PCR and Agarose gel electrophoresis was used to verify of cDNA by using *tcdA* primers that can later be amplified by quantitative RT-PCR (Figure 4.1).

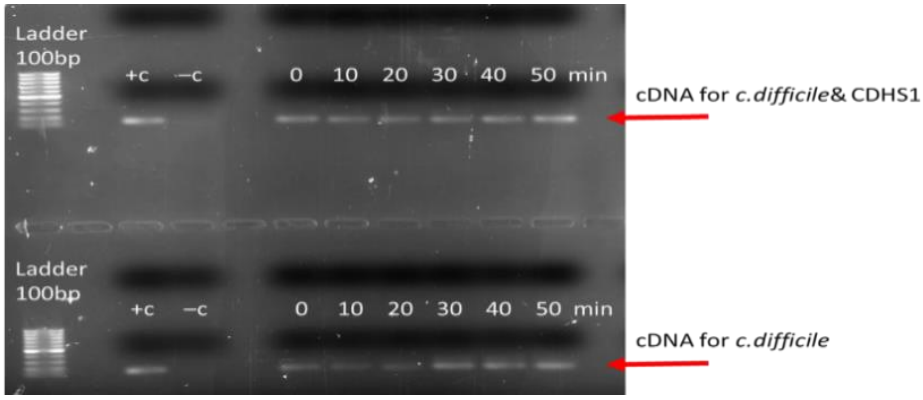


Figure 4.1. Agarose gel electrophoresis of PCR products of *tcdA* gene amplified from cDNA using total RNA from *C. difficile* R20291 infected with phage CDHS1and the uninfected *C. difficile* R20291 strain. The ladder used in this experiment was the 100 bp DNA Ladder. Controls: + c control positive genome DNA, -c control negative cDNA. (0-50) minutes fragments of cDNA (amplicon size 145 bp).

4.2.3 Preparing qPCR reactions

To quantify the expression of the virulence genes (Table 2.1) in *C. difficile* R20291 during infection with phage CDHS1, quantitative PCR was conducted. Amplification of specific PCR product (virulence genes cDNA) using the Fast SYBR® Green Master Mix protocol was carried out in triplicate for each sample, and three biological repeats for bacteria infected with phage and uninfected bacteria. For each reaction 1/100 dilution of

cDNA (25ng) was added to 250nM of the forward and reverse primers and 10µl Fast SYBR® Green Master Mix buffer (2X) to get the final volume of 20µl. The experiment was done with *recA*, and 16s *rRNA* housekeeping genes as reference controls. Within each experiment, no cDNA was used as negative control and a negative RT control was used to verify any contamination of the gDNA. The qPCR conditions that were used are shown in Table 4.1.

Table 4.1. Thermocycler conditions for qPCR

qPCR Step	Temperature(°C)	Time (sec)	Repeat cycle
Activation	95	20	Hold
Denaturation	95	3	40
Annealing/extension	60	30	

4.3 Quantification of virulence genes expression by qPCR

In this project, data obtained from qPCR were analysed to quantify gene expression levels for virulence genes by comparing the cycle threshold (Merabishvili *et al.*, 2009) values of the target genes in *C. difficile* R20291 infected with bacteriophage CDHS1 and non-infected *C. difficile* R20291 at different time points. In general, the double delta Ct (ΔΔCt) method is the one of the most widely used methods that use for calculating relative gene expression from qPCR data (Livak and Schmittgen, 2001).

These values were normalized to the reference genes as follows:

$$Relative\ expression\ of\ \text{of sample} * (\text{normalized}) =$$

$$2\text{^ (Reference } CT - \text{ Target } CT)$$

Sample* means Test or control

$$Ratio\ of\ gene\ expression\ of\ target\ gene(Fold\ change) = \frac{2\text{^ (Test Reference } CT - \text{ Test Target } CT)}{2\text{^ (Control Reference } CT - \text{ Control Target } CT)}$$

4.4 Results

4.4.1 *Relative quantification of expression of* Clostridium difficile virulence genes

The expression levels of the 10 virulence genes was determined using qPCR to quantify the transcriptome in both *C. difficile* R20291 infected with phage CDHS1 and the uninfected *C.difficile* R20291 strain as a control during lytic cycle.

4.4.2 The expression of all virulence genes within each time points

Based on the analysis of the gene expression in all virulence genes of interest genes, as expected, a similar relative expression was observed in most virulence genes that were down-regulated when *C.difficile* was infected with the phage. However, the level of the down- regulation differed between different time points. Therefore, to determine the impact of phage on gene expression levels of all virulence genes of interest at periods of 0 min, 10 min, 20 min, 30 min, 40 min, and 50 min were examined (Figure 4.2).

In terms of gene expression of all virulence genes at each time points, the *spmA* gene was more expressed in infected R20291 with CDHS1 at 30, 40 and 50 min into the growth cycle relative to uninfected R20291. However, down-expression of the *spmA* gene was found at the first three time points. The *FliA* gene was expressed less than other genes such as *tcdB* and *agrB* genes with exception at timepoints 0 and 10 min. Furthermore, the *FliA* gene was overexpressed at 0 and 10 minutes in infected *C. difficile* R20291, but the fold up-regulations were not particularly high in comparison to higher expression level of *spmA* gene at 50 min. The *tcdR* genen was also highly down-regulated in expression in R20291 during infection at 20 minutes. Overall, most genes were considerably reduced in expression in phage infected R20291, and only an increase in *spmA* gene expression levels at 30,40 and 50 minutes was observed.

4.4.3 The expression of each virulence genes at different time points

Additionally, the expression levels of each virulence gene at all time points were tested (Figure 4.3). The RT-PCR results showed that the toxin genes *tcdA,* Binary toxin (*cdta*), and *tcdE* were down-regulated in the infected R20291 at all time points after infection. In addition, for gene *tcdD,* expression was also decreased at all time points during infection. However, expression of gene *tcdC* was significantly up-regulated at most time points with exception of 0 min. Although qPCR results showed a relatively lower fold change in *agrB* and *agr* mRNA expression in infected R20291 than uninfected R20291, at 10 min both genes showed higher expression in R20291 during infection with CDHS1 bacteriophage

relative to uninfected R20291. Interestingly, lower expression of spore maturation protein A (*spmA*) gene was found in infected R20291at the first three time points, except for samples taken at 30, 40 and 50 min post infection. The expression of flagellar gene *FliA* was significantly reduced at 20 min and over time during the infection.

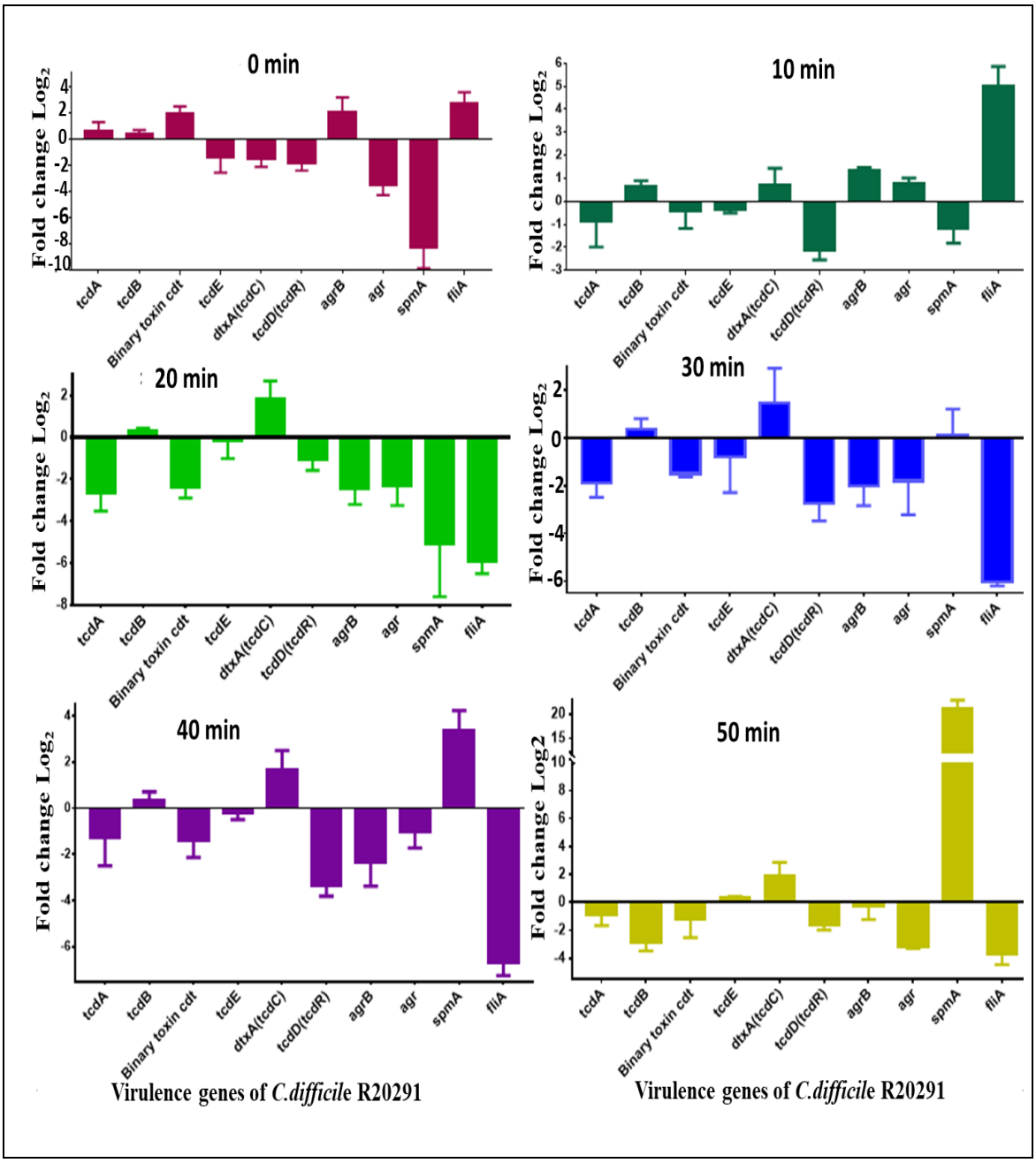


Figure 4.2. Comparison of all selected virulence genes at each time points during infection. Experiments were repeated using 3 replicate of 5 independent biological samples, error bar indicate SEM., $p<0.001$. Significantly values ranged between (-1> and 1<).

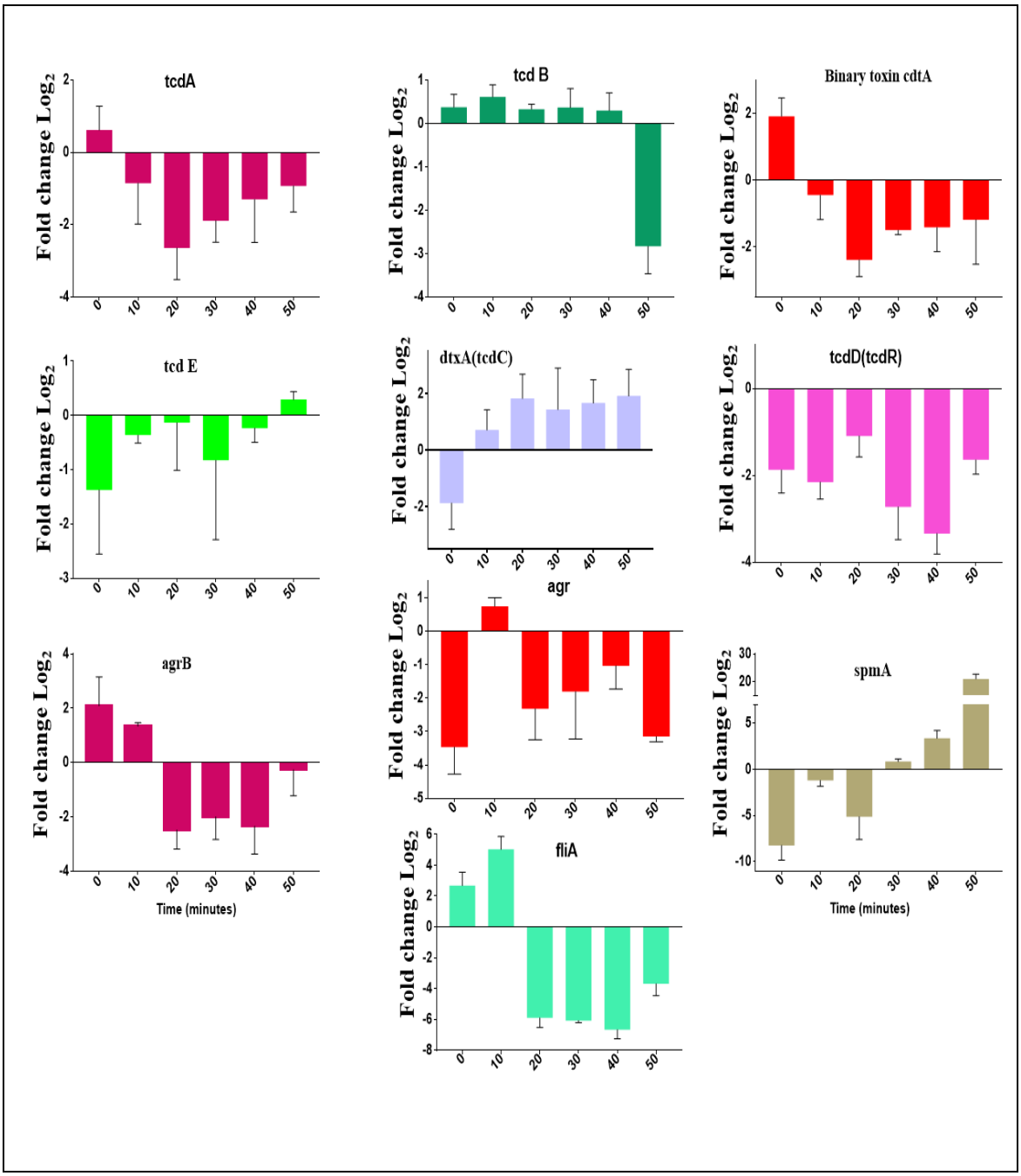


Figure 4.3. Comparison of all selected virulence genes at all time points during infection. Experiments were repeated using 3 replicate of 5 independent biological samples, error bar indicate SEM., $p<0.0001$. Significantly values ranged between (-1> and 1<).

4.5 Discussion

The aim of RT-qPCR experiment were to determine the relative transcript level quantification of several *C. difficile* virulence genes in infected strin R20291 with temperate phage CDHS1 and to study the transcript level of these virulence genes during

early, mid and late log phase and intact bacteria as control. Real time (RT-qPCR) analysis of mRNA was normalised by using at least two housekeeping reference genes to reduce any possibility of errors in expression analysis (Nicot *et al.*, 2005).

The results showed that the *tcdA*, *cdta* (Binary toxin) and *tcdE* transcript level decreased in log phase (20min, 30min 40min and 50min). Although this result was somewhat surprising since previous studies had identified that temperate phages can cause induction of toxin production by *C. difficile* (Bishai, 1988; Goh *et al.*, 2005a). However, it is possible that these phages carry genes for virulence factors such as exotoxins (Wagner and Waldor, 2002; Bishai, 1988). But there is a body of evidence to suggest that although of significant reduction of endotoxin (Et) release and cytokine production as a result of infection, lysis-deficient (LyD) bacteriophages can still kill the host bacterial cell (Matsuda *et al.*, 2005; Paul *et al.*, 2011). Furthermore Govind et al. confirmed that *C. difficile* carrying a prophage ΦCD119 can cause reduction in transcription of *tcdA* and *tcdB* (Govind *et al.*, 2009). Another observation indicates that phage treatment precluded toxin production by *C. difficile* (Meader *et al.*, 2013; Meader *et al.*, 2010). All these led to the suggestion that toxin gene transcription is reduced during infection with bacteriophage. However, previous studies have emphasized that phage can enhance the virulence of the bacterial host by encoding toxin genes or increased expression of host virulence genes which are not encoded by the phage (Spanier and Cleary, 1980; Wagner and Waldor, 2002; Sekulovic *et al.*, 2011).

Although the level of *tcdB* transcription as determined by real-time RT-PCR was not significantly changed at all time points, there was a clear reduction of toxin B production at the end of log phase (50min). Despite the phage infection of toxigenic strains increased toxin B production in lysogens (Goh *et al.*, 2005a); bacteriophage (ΦCD27) demonstrated a significant reduction in *C. difficile* toxin production (Meader *et al.*, 2013). This suggests that there are differences in toxin levels after infection from phage to another and this effect is specific for some phages more than others, but not in all.

Besides toxins A (TcdA) and B (TcdB), there is another toxin that adds more virulence to some *C.difficile* strains known as CDT binary toxin (Perelle *et al.*, 1997; Sundriyal *et al.*, 2009). In this study the expression of *cdtA* gene was repressed during log phase in infected cells than in non-infected cells as it happened with *tcdA* gene. This result suggests that there may be some correlation between toxin A and binary toxin. Despite the wide

studies in bacterial-phage interactions, currently, there is no study that has been done to determine the effects of bacteriophage on binary toxin production during bacterial infection.

Interestingly, low gene expression of *tcdE* at all times points during the growth cycle as a gene responsible for the release of toxin enzymes is fully consistent with the low level of toxin A and B production (Govind and Dupuy, 2012; Govind *et al.*, 2015). This is the most likely scenario as when the level of *C. difficile* toxins was decreased, consequently the result will be a decrease in the production of the enzyme which responsible for its release.

It is already known that there are two accessory genes have the opposite function of one another in *C. difficile*. The first is *tcdC*, a putative negative regulator of toxin A and B production and *tcdD* (*tcdR*), a putative positive regulator of toxin A&B production. It has been shown that the *tcdC* transcript is up-regulated in the R20291infected with phage CDHS1 at all time points compared to the uninfected strain R20291. In contrast, the down-expression of *tcdR* genes was suppressed in infected *C. difficile* during log phase. These findings are in agreement with those described by Govind et al., (2009) which showed a decrease in the expression of *C. difficile tcdA*, *tcdB*, *tcdR*, *tcdE*, and *tcdC* genes in CD119 lysogens. This suggests that phage infection has a negative effect on tcdR gene expression which corresponding to the lower A and B toxins production as it represented a positive regulator for these toxins. In addition, the up-regulation of tcdC (negative regulator) transcription has confirmed the effects of this gene on *tcdA* and *tcdB* genes levels through the negative action.

Accessory gene regulator quorum signaling system control gene expression of virulence factors by regulation of toxins A and B production (Darkoh *et al.*, 2016). The transcripts of *agr* and *agrB* were found to be down-regulated in the infected bacteria at log phase (20, 30, 40 and 50) minutes compared with the wild type R20291 strain. It has been previously demonstrated that the expression of *cdtA* and *cdtB* genes were repressed during log phase in infected cells than in non-infected cells. These results suggest that the *agr* and *agrB* play a central role in *C. difficile* toxin A and B production.

SpmA is one of important virulence factors encoded by *C. difficile* (Edwards *et al.*, 2016). The results shows that decrease in host gene expression in the infected bacteria which was not detected in uninfected *C. difficile* in the subject experiments. This suggested that

bacteriophage CDHS1 might modulate the spore production. Another virulence gene was detected for gene expression like *fliA* gene that participated the regulation for flagella motility *C. difficile* strains. By qPCR detection, the expression of the *fliA* gene of phage infected cells during log phase was down-regulated significantly compared with uninfected *C. difficile* R20291. In agreement, Cao et al. (2014) demonstrated that *fliA* and *fliD* gene expression levels were down-regulated significantly compared with MG1655 deltaflhDC and suggested that these genes could affect the motility of the lysogenic strain by phage in *Escherichia coli* (Cao *et al.*, 2014). The results shown here provides the theoretical basis for further research on the role of bacteriophage infection on expression of bacterial motility genes.

Chapter 5
The Transcriptional Changes in *C. difficile* R20291 during infection with φCDHS1 bacteriophage studied by RNA-Seq

Chapter 5

The Transcriptional Changes in *C. difficile* R20291 during infection with φCDHS1 bacteriophage studied by RNA-Seq

5.1 Introduction

5.1.1 Studying bacterial transcriptomes using RNA-seq

Over the past 20 years, microbial gene expression and functional genomics has been made easier by shifting from microarray technology to the next generation sequencing technologies. Mainly, RNA sequencing (RNA-seq) based on deep sequencing of cDNAs has greatly facilitated transcript mapping with single nucleotide resolution. High-throughput RNA sequencing (RNA-seq) techniques have become one of the most popular choices to study differential gene expression in prokaryotes and eukaryotes (Osmundson *et al.*, 2013). RNA-seq is used for the quantification and characterization of transcriptomes (Conesa *et al.*, 2016; Costa *et al.*, 2010). RNA-seq is considered more powerful and sensitive in the detection and monitor of any change in gene expression happening temporally following phage and host interaction (Ozsolak and Milos, 2011). For instance, several transcriptomic works have shown that phage genes are expressed at different times (e.g., early, middle and late, or early and late) according to their location in the genome (Legendre *et al.*, 2010). Further, Scaria et al study demonstrated that RNA-Seq was able to detect a number of transcriptionally active regions (novel genes) that were not part of the primary genome annotation (Scaria *et al.*, 2013).

Bacteriophage–host interactions have been intensely studied and generally, bacteriophages are thought to play important roles in adjusting the abundance, diversity, and synthesis of bacterial populations cycles through the process of host cell lysis (Suttle, 2007), and so significantly influencing bacterial populations and their function (Sullivan *et al.*, 2017; Rosenwasser *et al.*, 2016). Moreover, through these phage-host interactions, the bacteriophage can serve as a positive factor in bacterial physiology through gene transfer or by contributing directly to bacterial pathogenesis at the time of infection (Wagner and Waldor, 2002). Consequently, many studies have aimed at understanding how bacteriophages can induce or reduce the bacterial virulence (Laanto *et al.*, 2012; Fortier, 2018; Hsu *et al.*, 2019; Batinovic *et al.*, 2019). For instance, Wagner and Waldor emphasized that the alteration of bacterial virulence by bacteriophages is either due to phage-encoded virulence factors (Reidl and Mekalanos, 1995; Boyd, 2012; Castillo *et al.*,

2018), or alteration in bacterial virulence properties (Bishai, 1988), or regulation of phage-encoded virulence factors (Coleman *et al.*, 1989; Hernandez-Doria and Sperandio, 2018).

This study will discuss the use of RNA-seq transcriptome sequencing for studying gene expression in *C. difficile* R20291, and to use this technique to analyse bacterial gene expression during bacteriophage infection.

5.1.2 Experimental Design

To study the transcriptome of different infection stages of bacteriophage CDHS1 to host *C.difficile* R20291, RNA-SEQ analysis was conducted on datasets produced at six different time points of the phage infection and replication cycle within one step growth curve. Briefly, 24 RNA samples were sent BGI Company for sequencing as described in below:

Experimental Design
Triplicate experiments of *C.difficile* R20291 infected with φCDHS1 at time Point 0, 10, 20, 30, 40, 50 minutes 3x
Triplicate control (*C. difficile* R20291) at time Point 0, 10, 20, 30, 40, 50 minutes 3x

5.1.3 Next generation sequencing Analysis

A diagram summarising the different steps in the Illumina is depicted in Figure 5.1. Basically, Total RNA was prepared and sent to BGI BEING for sequencing. Briefly, in the BGI Company, rRNA was depleted using rRNA removal kits. Fragmentation buffer was added to make short fragments from the full length mRNA. Taking these short fragments as templates, random hexamer-primer were used to synthesize the first-strand (complementary) cDNA. Second-strand cDNA was synthesized using buffer, dATPs, dGTPs, dCTPs, dUTPs, RNase H and DNA polymerase I respectively after removing dNTPs. Short fragments were purified using a QiaQuick PCR extraction kit and resolved with EB buffer for end repair and to add a polyA tail. After that, the short fragments were

connected with sequencing adapters (A: 5'-AAGATACGGCGACCACCGAGATCT ACAC-3' and B: 5' - CAAGCAGAAGACGGCATACGAGAT - 3') (Buermans and Den Dunnen, 2014). Then, the uracil DNA glycosylase (UNG) enzyme was used to degrade the second-strand cDNA, and the product was purified by using a MiniElute PCR Purification Kit before PCR amplification. The library was then sequenced using a HiSeq™ 4000 sequencer and the activity pipeline of the experiment was described in Figure 5.1.

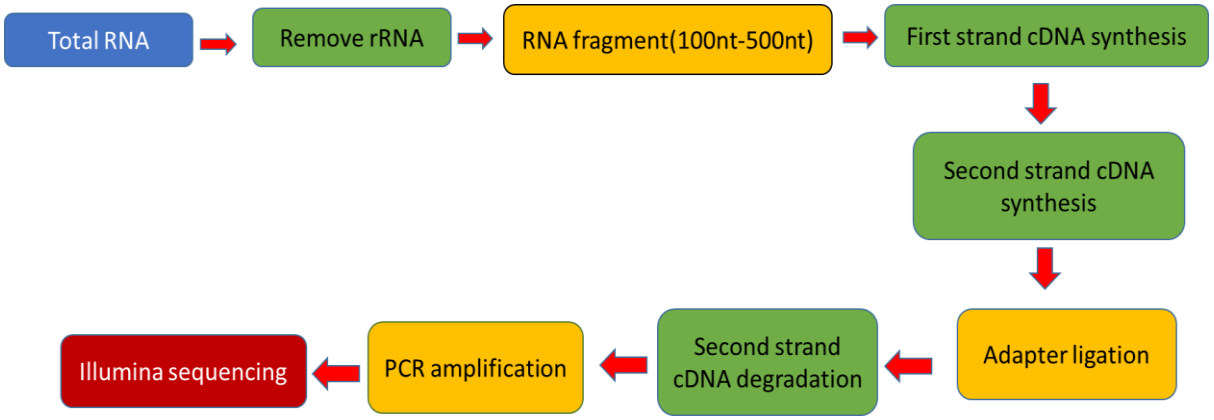


Figure 5.1. Experiment pipeline of Prokaryotic Strand-Specific Transcriptome

5.1.4 Next Generation sequencing Transcriptome Transcript Count

All reads data are sequenced by BGI, consequently, FASTQ format were gained ready to align to the transcriptome or genome of both *C. difficile* and CDHS1. Prior to alignment, all data detected through Quality Control to ensure the purity of the data and avoid any unsuitable alignment (Figure 5.2, and Appendix 5).

Briefly, 24 datasets in FASTQ format were provided for analysis. The technical and biological replicates were subjected to phage infection. The control contained the host only and no phage. Abundances of transcripts were calculated using reference genomes, targeted sequences and Kallisto software. Analysis has been performed using the read count estimates provided that were generated using Kallisto (Bray *et al.*, 2016), a recently published abundance estimation package that generates its output via fast pseudoalignment (Bray *et al.*, 2016).

A single matrix of read counts for each gene across all samples was generated from the individual Kallisto files. Differential expression analysis was performed using edgeR v3 (Robinson *et al.*, 2010) through R Studio v1.0.143, Rv3.4.0.

Genes with counts-per-million levels were filtered prior to analysis. Counts were subject to trimmed mean of M-values normalization (TMM) to permit comparison across samples. Differential expression was calculated using a generalized linear method (GLM) likelihood ratio test.

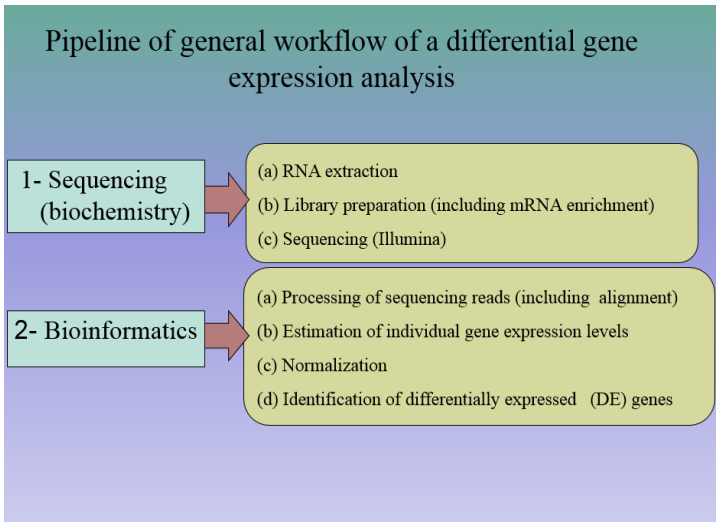


Figure 5.2. Steps for differential gene expression

5.2 Results and Discussion

RNA was sequenced from all replicates samples of *C. difficile* in each of two well-studied experimental conditions; Infected with bacteriophage and wild-type (WT) as control.

5.2.1 Bacterial gene expression during bacteriophage infection

In order to get insights into the bacteriophage-host interactions, RNA-seq gene expression has been previously applied by several researchers to study bacterial gene expression (Leskinen *et al.*, 2016; Sekulovic and Fortier, 2015). RNA-seq was performed with Illumina HiSeq^{TM}4000 on three independent replicate cultures of the *C. difficile* R20291 infected with phage φCDHS1 and compared with wild type R20291.

In this study, two methods were used to analyse RNA-seq data, Partek software data analysis and bioinformatic (interface R studio & package edge R). Depending on data analyses using Partek software, the expression mode of majority of R20291 genes (3514) were varied between no, low, moderate, or high for both R20291 infected with CDHS1 and uninfected R20291 (as control) at different time points. For all genes significantly

expressed at *p*- value 0.01, tables A, B, C, D, E and F (appendix 6) described top 20 up-regulated and down-regulated genes with their products / function for both R20291 and φCDHS1 during infection at 0, 10, 20, 30, 40 and 50 min, respectively. Briefly, although most R20291 genes involved in core DNA binding protein and transcription were among those showing higher expression (up-regulated), but still at low level comparing to expression levels of genes from bacteria infected with φCDHS1 (Table 5.1 A). Since bacterial genes would be expected to be moderately or more lowly expressed at different developmental stages, those genes encoding metabolic processes and phage releasing permease were especially at 0 and 10 min among the very low expression levels group (i.e. were down-regulated). Likewise, genes encoding putative membrane and hypothetical proteins, were also expressed at very low levels in phage infected R20291 (Table 5.1 A). As previously confirmed that the uninfected cells when reach at the late logarithmic growth phase (at 50 min post infection), the expression of genes encoding metabolic pathways were down-regulated (Ravantti *et al.*, 2008). However, on the opposite view point, Sekulovic and Fortier (2015) stated that ΦCD38-2 prophage increases the expression of metabolic genes of *C. difficile* R20291 in the lysogenic cycle. Study findings suggest that bacteriophage was able to supress the expression of genes related to metabolic processes at early stage of growth. Moreover, the results confirm that a temperate φCDHS1 has ability to inhibit *C. difficile* gene expression during one step growth curve, while a previous study suggested that bacteriophage can disrupt host DNA replication and that a virulent coliphage T4 can completely cease host gene expression within 20 min post infection (Ueno and Yonesaki, 2004; Bryan *et al.*, 2016). This indicates that temperate phage can terminate the expression of host genes during lytic cycle.

5.2.2 Bacteriophage CDHS1 gene expression during lytic growth

Temperate phages have the same lytic capacity as virulent phages but in addition, they can integrate to bacterial DNA and adopt lysogenic lifestyle (Guttman *et al.*, 2005). These two pathways are controlled, depending on the metabolic state of the bacterial cell and phage anti-repressor (Cro in phage λ) induces the cleavage of the phage CI repressor which in turn inhibits the transcription of most of the phage genes to start lysogenic cycle (Hernandez-Doria and Sperandio, 2018; Takeda *et al.*, 1977). In terms of changes involved phage gene expression during one step growth curve, it was observed across the whole phage genome and seemed to be greatly higher than for most of the bacterial genes

(Table 5.1 B, Appendix 6). Early genes such as those encoding Cro / C1-type transcriptional regulators (Transcription Repressor) (CDHS1_0037), single stranded DNA (CDHS1_0047) and helicase (CDHS1_0035) were strongly expressed to higher levels at 0 and 10 min during one step growth curve. Previous work has shown that the *Cro* gene which is required for lytic development, is the first gene which expressed increasingly from 5 to 10 min, and then less at 20 min (Liu *et al.*, 2013b).

In the current project, similar patterns of up-regulation were recorded at 20, 30, 40 and 50 min after φCDHS1 infection, the most up-regulated pathways (intermediate and late genes) were as expected related to the lysis enzyme system including holin (CDHS1_023), endolysin (CDHS1_025), and to structural proteins such as tail fiber protein (CDHS1_022). At the last phase of the phage reproduction cycle, bacteriophages encode holin and endolysin proteins to lyse the host cell and subsequently releasing the progeny virions (Rydman and Bamford, 2003; Ziedaite *et al.*, 2005). However, no apparent down-regulation pattern for φCDHS1 genes were observed in infected cells at all time points during infection. The general trend of φCDHS1 genes was that such pathways were further up-regulated.

Table 5.1. Summary of modulation of bacterial and φCDHS1 gene expression during one step growth curve.

A		
Clostridium difficile R20291	**Time**	**Up-regulated**
	0 min	Nucleoted and ATP binding protein
	10 min	DNA binding and ATP-binding / permease protein
	20 min	DNA binding pr., transcription and protein folding
	30 min	DNA binding pr., transcription and protein folding
	40 min	DNA binding pr., transcription
	50 min	Membrane and integral component of membrane
	Time	**Downregulated**
	0 min	Metabolic process, membrane and hypotheticl protein
	10 min	Metabolic process, integral component of membrane
	20 min	Putative membrane and hypotheticl protein
	30 min	Putative membrane protein and conserved hypothetical pr.
	40 min	Permease protein,putative membrane protein
	50 min	Permease protein,putative membrane protein
B		
Bacteriophage CDHS1	**Time**	**Up-regulated**
	0 min	Cro/C1-type transcriptional regulator, ssDNA-binding pr. & helicase
	10 min	Cro/C1-type transcriptional regulator, ssDNA-binding pr.
	20 min	Holin, Endolysin and tail fiber protein
	30 min	Holin, Endolysin and tail fiber protein
	40 min	Endolysin and tail fiber protein
	50 min	Tail fiber protein, Endolysin and minor capsid protein
	Time	**Downregulated**
	0 min	Not expressed
	10 min	Not expressed
	20 min	Not expressed
	30 min	Not expressed
	40 min	Not expressed
	50 min	Not expressed

5.2.3 Number of differentially expressed (DE) genes

In this study, two methods were used to analyse RNA-seq data. The number of significantly differentially expressed genes obtained between *C. difficile* infected with φCDHS1 bacteriophage at each stage post-infection is shown in Table 5.2 A, B which show the number of DE genes *C. difficile* infected with φCDHS1 depending on control by using Partek software data analysis and bioinformatic (interface R studio & package edge R) respectively. At the same time, Table 5.3 A &B show the number of DE genes *C. difficile* infected with φCDHS1 depending on 0 time baseline by using Partek software data analysis and bioinformatic (interface R studio & package edge R) respectively.

Taking all of this into account, it was decided to examine the transcriptional changes caused by phage CDHS1 in *C. difficile* R20291 after 0, 10, 20, 30, 40 and 50 min of infection, since in these early times there is a significant lysis of the cells that could interfere with the results. RNA-Seq technology was used to simultaneously analyze both host and viral transcriptomes. Moreover, differential expression analysis and Characterizations results by Partek software are compared to Bioinformatics software (edge R & R studio). The results of differential gene expression with analysis of RNA-seq data were similar to those using Partek software. At all-time points after infection, transcriptional regulator, DNAse / helicase proteins and hypothetical proteins were found to be up-regulated of φCDHS1 genes.

Table 5.2 Summary of genes differentially expressed (DE) between *C. difficile* R20291infected with φCDHS1 versus control at each time points using (**A)** Partek software, and (**B)** Bioinformatics /edge R & R studio.

A	**Infected R20291 vs. control**					
Condition	**0 min**	**10 min**	**20 min**	**30 min**	**40 min**	**50 min**
Signifcantly expressed	455	404	481	492	524	520
upregulated genes	27	45	48	50	68	70
Down regulated genes	428	359	433	442	456	450
bacterial Signifcantly expressed	442	379	453	462	498	496
upregulated bacterial genes ⇧	14	20	20	20	42	46
downregulated bacterial genes ⇩	428	359	433	442	456	450
Phage Signifcantly expressed	13	25	28	30	26	24
upregulated phage genes ⇧	13	25	28	30	26	24
downregulated phage genes ⇩	0	0	0	0	0	0
B	**Infected R20291 vs. control**					
Condition	**0 min**	**10 min**	**20 min**	**30 min**	**40 min**	**50 min**
Signifcantly expressed	367	82	85	93	123	123
upregulated genes	238	48	45	57	59	72
Down regulated genes	129	34	40	36	64	51
bacterial Signifcantly expressed	317	66	73	72	100	99
upregulated bacterial genes ⇧	188	32	33	36	36	48
downregulated bacterial genes ⇩	129	34	40	36	64	51
Phage Signifcantly expressed	50	16	12	21	23	24
upregulated phage genes ⇧	50	16	12	21	23	24
downregulated phage genes ⇩	0	0	0	0	0	0

Table 5.3 Summary of genes differentially expressed (DE) between *C. difficile* R20291infected with φCDHS1 compared to the time 0 baseline at each time points using (**A)** Partek software, and (**B)** Bioinformatics / edge R & R studio.

A	**Infected R20291 vs. 0 min baseline**				
Condition	**10 min**	**20 min**	**30 min**	**40 min**	**50 min**
Signifcantly expressed	**336**	**359**	**512**	**589**	**598**
upregulated genes	**58**	**95**	**160**	**190**	**197**
Down regulated genes	**278**	**264**	**352**	**399**	**401**
bacterial Signifcantly expressed	**309**	**324**	**478**	**556**	**561**
upregulated bacterial genes	**35**	**67**	**131**	**162**	**170**
downregulated bacterial genes	**274**	**257**	**347**	**394**	**391**
Phage Signifcantly expressed	**27**	**35**	**34**	**33**	**37**
upregulated phage genes	**23**	**28**	**29**	**28**	**27**
downregulated phage genes	**4**	**7**	**5**	**5**	**10**

B	**Infected R20291 vs. 0 min baseline**				
Condition	**10 min**	**20 min**	**30 min**	**40 min**	**50 min**
Signifcantly expressed	**142**	**180**	**256**	**303**	**328**
upregulated genes	**63**	**73**	**108**	**130**	**140**
Down regulated genes	**79**	**107**	**148**	**173**	**188**
bacterial Signifcantly expressed	**119**	**153**	**225**	**274**	**292**
upregulated bacterial genes	**40**	**47**	**78**	**104**	**113**
downregulated bacterial genes	**79**	**106**	**147**	**170**	**179**
Phage Signifcantly expressed	**23**	**27**	**31**	**29**	**36**
upregulated phage genes	**23**	**26**	**30**	**26**	**27**
downregulated phage genes	**0**	**1**	**1**	**3**	**9**

5.2.4 Virulence gene expression changes in infected *C. difficile* with phage CDHS1 using RNA-seq technology versus qPCR

RNA-seq has become to be a powerful tool for bacterial and bacteriophage transcriptome assessment. Therefore, data obtained from RNA-seq were analysed to quantify gene expression levels for virulence genes. As used before with qPCR, 10 virulence genes were selected from samples (*C. difficile* R20291 infected with phage CDHS1) encounter its

control (uninfected *C. difficile*) at different time points (Figure 5.3). Based on the results, a strong correlation was found between RNA-seq and qPCR with exception of binary toxin (*cdt*), *spmA* and *fliA* gene expression. The results revealed that *cdt* was 4-fold induced in RNA-seq versus 2-fold induction in qPCR at 0 min, while showed induction in RNA-seq versus repression in qPCR at other time points.

Out of 10 targeted genes, most were differentially expressed in relative control of which *spmA* and *fliA* were down-regulated in RNA-seq at all time point whereas, *spmA* were up-regulated in qPCR at 40 and 50 min. Moreover, qPCR result revealed that *fliA* were only up-regulated at 0 and 10 min. In *tcdC* gene, RNA-seq and qPCR displayed a ratio of nearly 1.5-fold repression at 30, 50 min in RNA-seq with slightly induction found at same time in qPCR. There was no significant difference in expression of *tcda*, *tcdB*, *tcdE*, *agrB* and *agr* in the data from qPCR and RNA-seq.

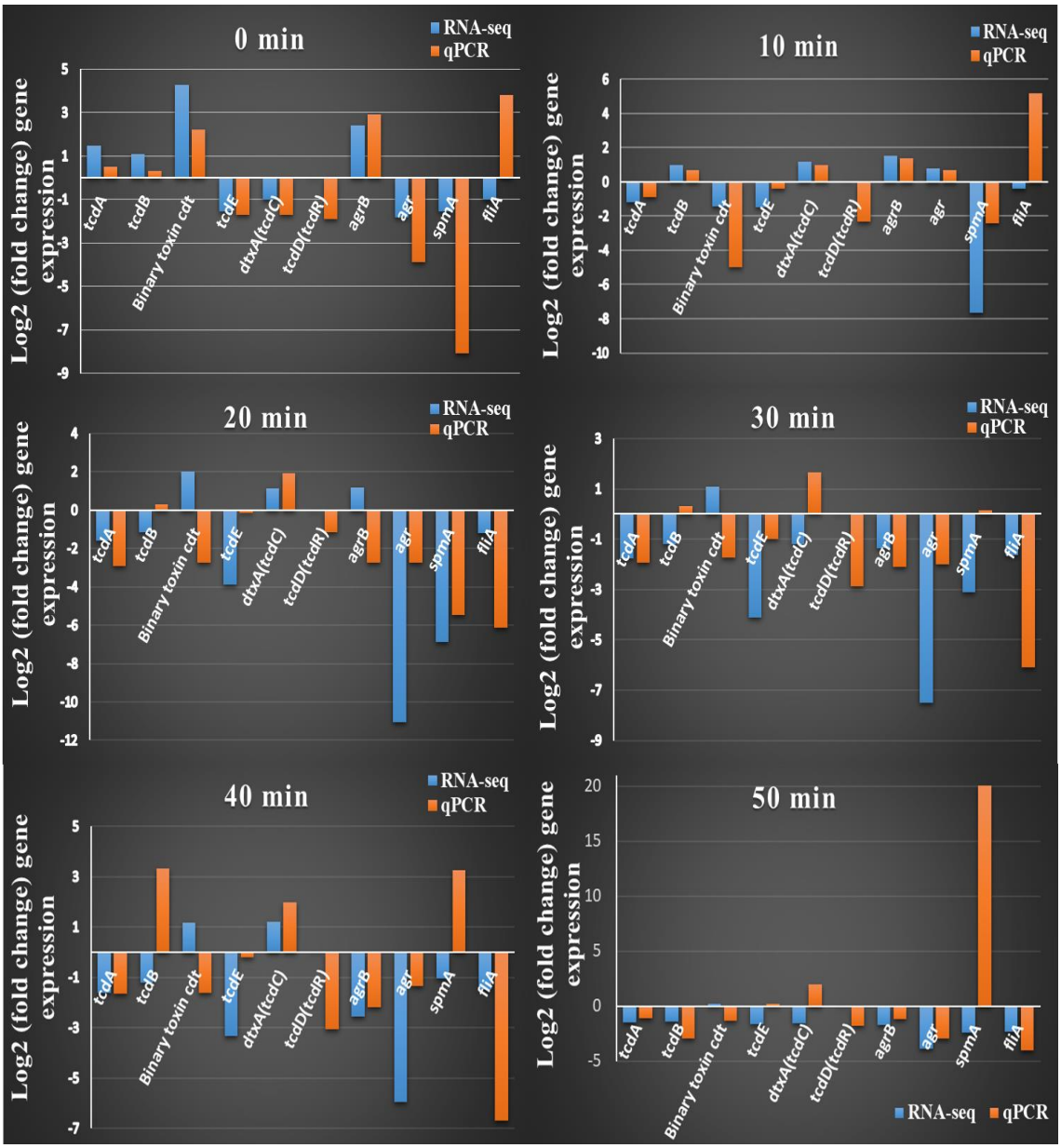


Figure 5.3. Comparison of selected *C. difficile* virulence genes at each time points during infection using RNA-seq and qPCR results. Experiments were repeated using triplicate of 5 independent biological samples, error bar indicate SEM., P > 0.05. Expression level changes were considered significant if they differed at least twofold (Log2 fold change) ranged between (-1> and 1<).

5.2.5 Phage gene expression during lytic growth according to RPKM data

As phage lack the machinery to express their own genes and replicate their own genomes, they must infect bacteria and exploit its host cell machinery. During infection,

expression of phage genes, and synthesis of new phage particles will start to form the lytic cycle. In starvation conditions, the phage genome integrates with the host chromosome to become a prophage which remains within their host for many generations before returning to the lytic cycle. The studies of phage-gene expression during infection is limited. In this section, the chapter will focus on the main phage genes that are significantly expressed during the one step growth curve according to their Reads per Kilobase per Million mapped reads (RPKM). From the previous experiments, phage gene can be classified into immediate/ early, intermediate and late maximum gene expression. Here, the subject focused on more significantly expressed among those genes.

5.2.5.1 Terminase small and large subunit

Terminase small (CDHS1_1) and large (CDHS1_2) Sub units genes are expressed together as late/ intermediate genes. Terminase is DNA packaging enzyme of phage λ and PaP3 which has been believed to plays an active role in the translocation of DNA into the head shell (Duffy and Feiss, 2002; Shen *et al.*, 2012). It is also involved in DNA cleavage once the capsid is full (Zhao *et al.*, 2013). Figure 5.4 showed that both genes which encode whole enzyme terminase (small and large subunit) is overexpressed at early late / late phase post infection. Thus, these results demonstrated that both proteins are required at the late stages of growth to backage phage DNA in the capsid without necessity to be present early (Wang *et al.*, 2005).

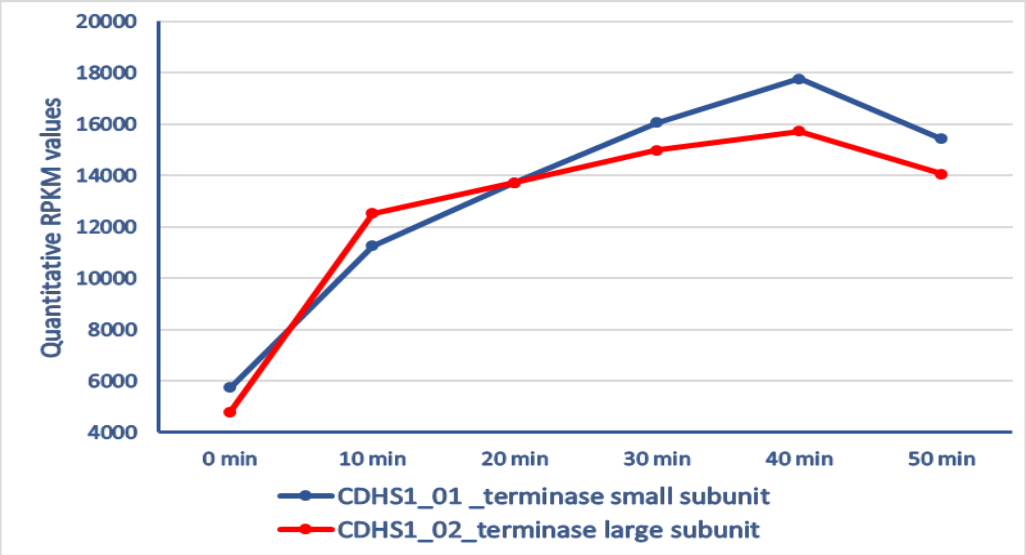


Figure 5.4. The expression characteristics of intermediate/ late genes: terminase small and large subunit transcripts in *C. difficile* R20291 infected with phage CDHS1. Quantitative RPKM value is used to represent the expression level at each time points.

5.2.5.2 Coat, head-tail connector and minor capsid proteins

These structural proteins have been previously observed as intermediate genes and continuously expressed in late stages as would be expected that structural proteins should be produced at early or intermediately onwards (Figure 5.5) (Liu *et al.*, 2013b). Clearly, head tail connector protein plays important role in virion assembly by joining the head and the tail at the last step of morphogenesis (Cardarelli *et al.*, 2010). Minor capsid protein is essential for stable capsid assembly of complete particles and when assembles with major capsid protein, icosahedral capsid and copy of the T7 genome will form to release in the medium (Ionel *et al.*, 2011).

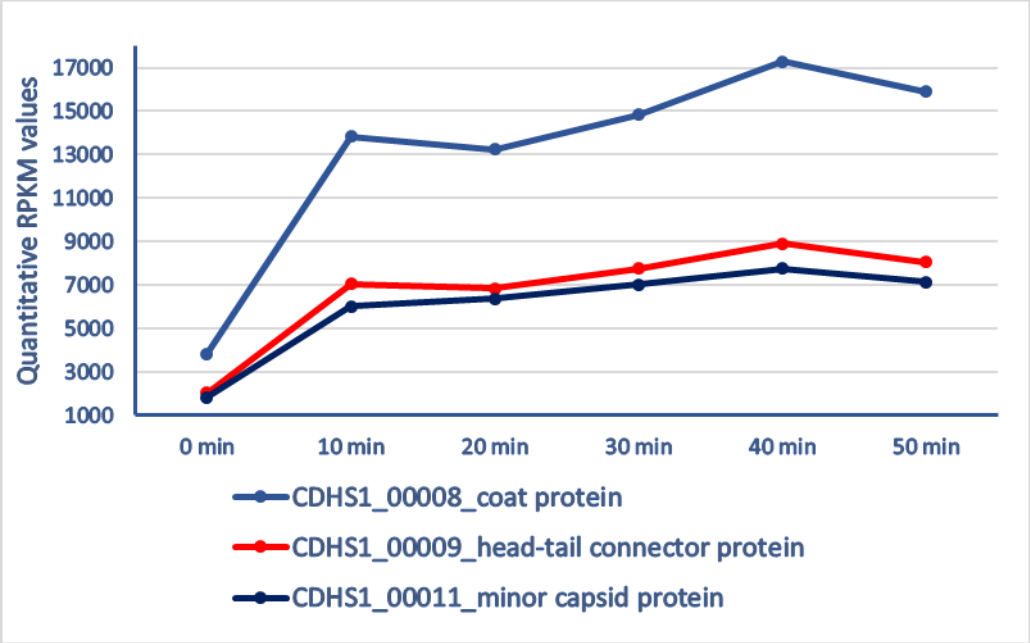


Figure 5.5. Expression of genes encoding intermediate/ late genes Coat, head-tail connector and minor capsid in *C. difficile* R20291 infected with phage CDHS1. Quantitative reads per kb per million (RPKM) values for Coat, head-tail connector and minor capsid were plotted over the infection time points.

5.2.5.3 Tail fiber and tail tape measure proteins

Tail proteins are also structural proteins which are produced as expected at intermediate / late stage. Tail fiber protein is structural component of the short non-contractile tail and help bacteriophage to attach to host lipopolysaccharides to mediate primary attachment to the host cell (Doval and Raaij, 2012). Tail tape measure protein serves as a base component for tail tube protein polymerization and acts as a template for tail length

determination (Belcaid *et al.*, 2011).The pattern for tail fiber and tail tape measure proteins presented in (Figure 5.6) was very similar to each other.

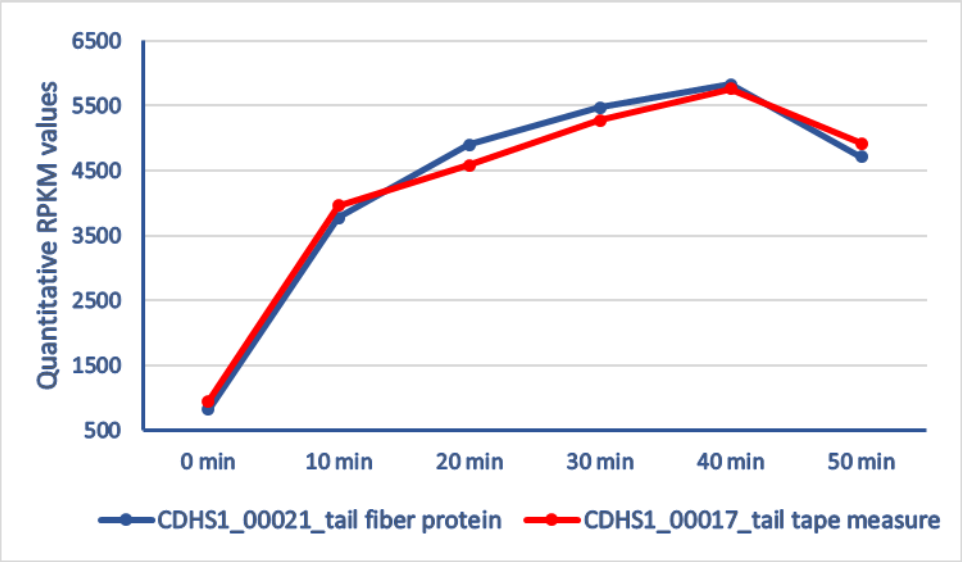


Figure 5.6. Changes of gene expression levels of Tail fiber and tail tape measure transcripts in *C. difficile* R20291 infected with phage CDHS1. Quantitative RPKM value is used to represent the expression level at each time points.

5.2.5.4 Holin, Scaffolding protein, integrase and sigma70/ sigma F-like protein

Holin is small membrane protein and essential for host lysis by bacteriophage. Although there is not much known about the holin mechanism, phage holins have been demonstrated to depolarize the bacterial cytoplasmic membrane to form holes allowing the passage of endolysin and disruption of the host cell wall (Wang *et al.*, 2000; Catalao *et al.*, 2013). Holins are hydrophobic proteins that start to oligomerize in cytoplasmic membrane at mid or late stages of phage replication (Grundling *et al.*, 2001; Young, 2013). The results of the current study showed that the holin gene was expressed at a higher level in early, mid and late phase (20, 30 and 40 min) (Figure 5.7.A) which corresponds well to that of other research (Delisle *et al.*, 2006). Expression of the gene for Scaffolding protein was observed as late in occurrence as this gene as this gene product is involved in DNA condensation in the capsid which should occur in the later/end stages of viral replication. It is required for successful condensation of DNA within the capsid. The interior of the prohead is filled with the gp8 protein. The scaffolding protein is lost from the structure during packaging (Dewey *et al.*, 2010; White

et al., 2011). The result in Figure 5.7.B could be considered as new evidence for that of late expression of the scaffolding gene.

Integrase is a viral enzyme as it is necessary for integration of the phage into the host genome by site-specific recombination (Fogg *et al.*, 2014). As shown for the holin and scaffolding proteins which discussed above, the integrase encoding gene is expressed at mid and late stages during the infection cycle (Figure 5.7.C). However, the stage in which enzyme releasing is unclear as the phage integrase is produced under conditions of stress, when bacteria enter stationary phase (Feiner *et al.*, 2015).

Sigma factors are initiation factors that promote the attachment of RNA polymerase to specific initiation sites and are then released. Sigma70 /sigma F-like protein gene is involved in the promotion of specific gene expression at the late stages of spore development in *Streptomyces* (Potuckova *et al.*, 1995). Based on the results in (Figure 5.7. D), sigma70 / sigma F-like protein is produced at early intermediate/ late stage during the cycle.

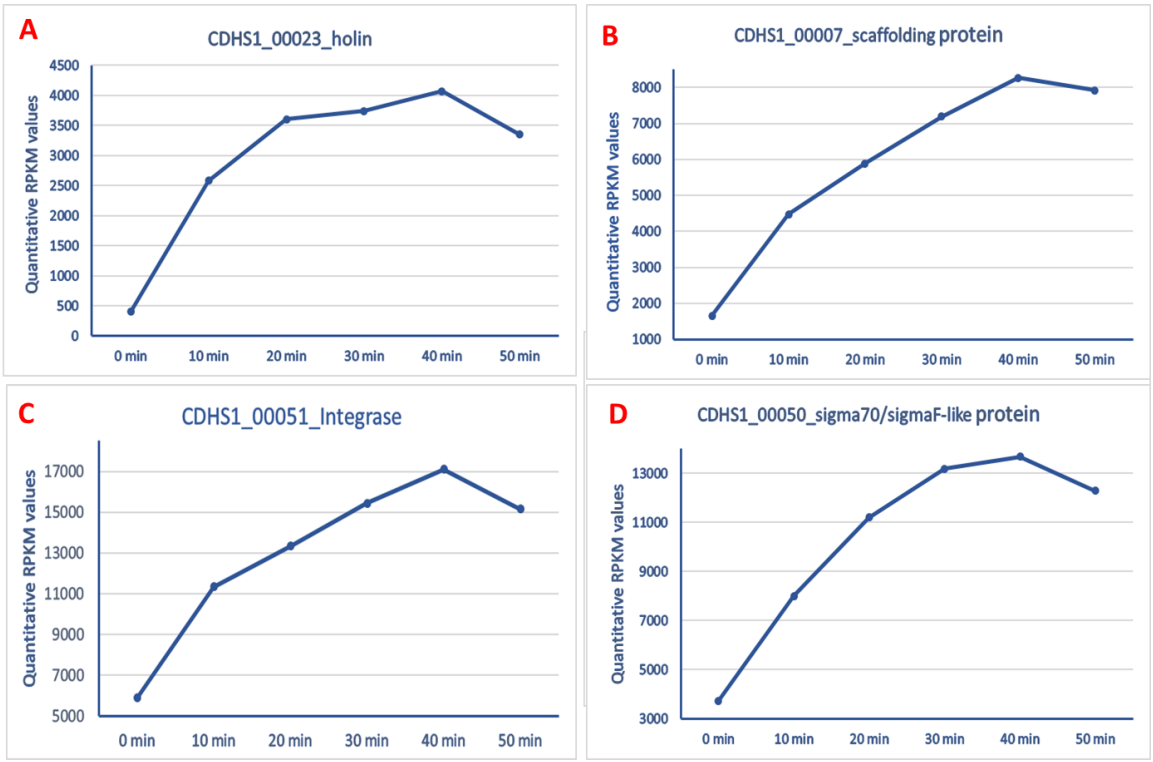


Figure 5.7. The expression characteristics of late phage replication cycle genes: (A) Holin (B) Scaffolding protein (c) integrase and (D) sigma70/ sigma F-like protein transcripts in *C. difficile* R20291 infected with phage CDHS1. Quantitative RPKM value is used to represent the expression level at each time points. Shown are the graduate expression at early phase values represent the average data calculated for the 0 and 10 min time points, and up- regulation at the intermediate/ late phase values, the average for the 20, 30, 40 and 50 min time points.

5.2.5.5 DNAse / Helicase, single-stranded DNA-binding protein and Cro /C1-type transcriptional regulator

As bacteriophage genomes are short, a large number of bacteriophage genes can be classified to the early, middle and late, based on the temporal pattern of their expression during infection (Hinton, 2010). Although, most CDHS1 phage genes that mentioned above were expressed at middle and late stage, the CDHS1 phage also has early genes such as DNAse / Helicase and single-stranded DNA-binding protein genes whichare expressed early only; this is as expected as the viral genome should be replicating at early stages during the infection (Figure 5.8 A & B). It can be seen that both genes are repressed at intermediate and late stages. DNA helicase is required to stimulate viral DNA replication and recombination (Gauss *et al.*, 1994). Single-stranded DNA-binding protein also expressed early as its role is to protect replicating DNA from nucleases. Single-

stranded DNA binding protein required for the elongation during viral DNA replication by strand displacement (Soengas *et al.*, 1995). The viral DNA strands displacement are transiently coated with the ssDNA-binding protein and therefore protected against nucleases (Martin *et al.*, 1989).

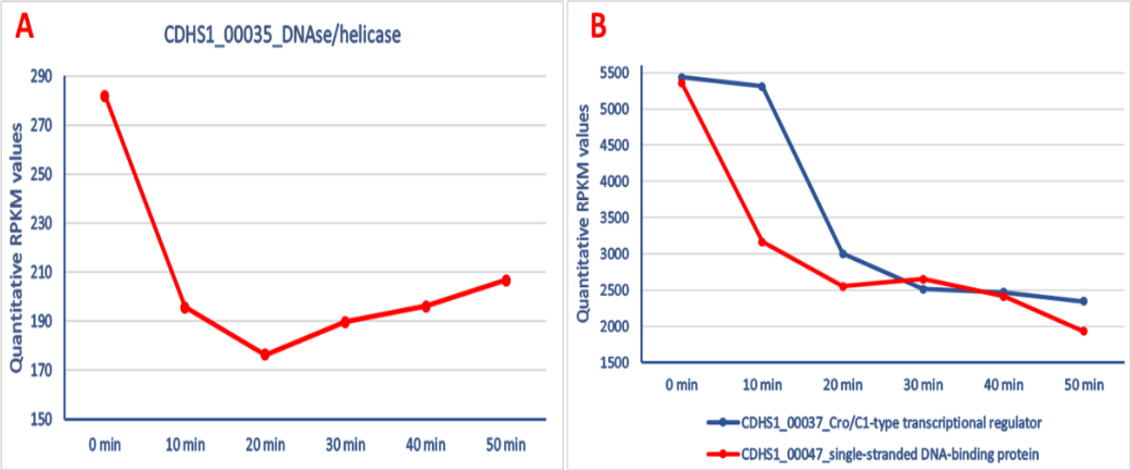


Figure 5.8. The expression characteristics of early genes: (A) DNAse / Helicase (B) single-stranded DNA-binding protein and Cro / C1-type transcriptional regulator transcripts in *C. difficile* R20291 infected with phage CDHS1. Quantitative RPKM value is used to represent the expression level at each time points. Shown are the over expression at early phase values represent the average data calculated for the 0 and 10 min time points, and down- regulation at the intermediate/ late phase values, the average for the 20, 30, 40 and 50 min time points.

Chapter 6
Evaluation of *C. difficile* R20291 virulence using *Galleria mellonella* larvae as an *in vivo* model

Chapter 6. Evaluation of *Clostridium difficile* R20291 types virulence using *Galleria mellonella* larvae as an *in vivo* model

6.1 Introduction

The *Galleria mellonella* larvae (also called wax worms or honeycomb moth) have become a widely adopted insect infection model to study several the infectivity of human pathogens. The main reason to use *G. mellonella* larvae in pathogen infection experiments is the larvae can be incubated at optimal physiological temperature for human pathogens 37°C and inexpensively obtained in large numbers at low cost (Rejasse *et al.*, 2012; Fuchs *et al.*, 2010). They are also simple to use relative to another laboratory animals such as mice and are easy to maintain without needing to special laboratory housing equipment. (Ramarao *et al.*, 2012). In addition, use of *G. mellonella* does not need ethical approval plus their short life span encourage the researchers to select them as alternative invertebrate models and a suitable host for high-throughput studies and pathogen mutant libraries (Tsai *et al.*, 2016; Desbois and Coote, 2012). Furthermore, the whole genome of *G. mellonella* has recently been sequenced and their transcriptome become available to use (Lange *et al.*, 2018; Vogel *et al.*, 2011). However, besides the advantages of using these insects as a model organism in laboratory experiments, there are also drawbacks to their use. For example, it is difficulty to perform genetic manipulations to get mutant strains defective in aspects of immunity. They are also not suitable for some bacterial species, and use of the larvae also needs to be pre-planned before purchasing them from the companies involved to ensure they arrive in time (Trevijano and Zaragoza, 2019).

In this study, the *G. mellonella* larvae was evaluated as an *in vivo* model for studying *C.difficile* R20291 virulence., as well as that of the lysogenic, φCDHS1 Resistant and wild type of *C. difficile* R20291 isolates.

6.2 Material and methods

6.2.1 Bacterial strains and culture conditions

To investigate the pathogenesis and gene expression of phage resistant and lysogenic *C. difficile in vivo, G. mellonella* was used as a an infection model for studying *C. difficile* pathogenesis (Ramarao *et al.*, 2012; Pereira *et al.*, 2018).

6.2.1.1 Isolation and purification of phage resistant, lysogenic *C. difficile*

Phage resistant and lysogenic *C. difficile* R20291 isolation was performed as described previously (Sorensen *et al.*, 2011) with some modifications. Briefly, *C. difficile* R20291 was streaked from cryopreserved on blood agar and incubated for 48h in 37˚C in an anaerobic chamber. Individual colonies of *C. difficile* R20291 were used to inoculate pre-reduced (FA) broth and incubated 18h in 37˚C in an anaerobic chamber. After incubation, each liquid culture was transferred to 10 ml of BHI broth, then they were incubated in 37˚C in an anaerobic chamber until OD_{550} = 0.2 ± 0. 01 which is about 10^7 CFU/ml. Following that, each obtained *C. difficile* R20291 culture in BHI above was diluted to 10^5 to obtain about 10^2 CFU/ml in the final volume of 1 ml. Then after, cultures were incubated overnight in 37˚C in an anaerobic chamber. After incubation, 100 µl from each culture was transferred to 10 ml of fresh BHI broth and further incubated in 37˚C in an anaerobic chamber until an OD_{550}= 0.2 ± 0.01 (λ=550) was reached. Cultures obtained after incubation were serially diluted from 10^0 to 10^{-7} and 10 µl of each bacterial dilution was applied on on 3-ml BHI semisolid agar overlays on Brucella agar (Oxoid, United Kingdom) supplemented with 5% blood, 5 mg/L hemin, and 1 mg/L vitamin K (Kim *et al.*, 2014). Similarly, the diluted cultures were also applied to semisolid agar overlays containing 250 µl of φCDHS1 at a titre of 10^9 / PFU. This procedure was repeated for each individual dilution. After an additional incubation of 24 h at 37°C inside the anaerobic chamber, single colonies were observed at higher dilutions. After which, the whole procedure was repeated five times. Then, colonies recovered from each phage-containing plate has been purified by re-streaking on fresh BHI blood agar and incubated in 37˚C for 24h. Re-streaking was repeated at least 3 times to be sure that all residues of phages from soft agar layer on initial supplemented Brucella Agar plates were eliminated. The last step single colony from the phage treatments was streaked onto blood agar plates and was identified to be *C. difficile* using 16s rRNA PCR technique. The sensitivity of resistant, lysogenic, and wild-type bacteria to phages phiCDHS1 was determined using a spot test with serially diluted phage lysate.

6.2.1.2 Validation of *Clostridium difficile* types

All *C. difficile* colonies were isolated from previous step (section 6.2.1.1) were collected and further characterized for the presence or absence of the phages using primers targeting Holin, CDHsipho with PCR to verify the identity of lysogenic *C. difficile* (Table 6.1). Additional individual PCRs were carried out to detect the presence of Medium tailed

myos gene specific for conserved *C. difficile* phage as several of isolates were negative for phage genes (Table 6.1). CDHS1 phage genes for control was performed by using phiCDHS1 DNA extract (section 6.2.1.2.4).

Table 6.1. Oligonucleotides used to detect differential gene expression of phage CDHS1 during infection by qPCR.

Name	Forward and reverse primers	Target CDHS1 genes	Expected product size (bp)
Holin	F 5'- GGACTAGGAGCAGTAGGGATAT-3'	Holin specific for CDHS1	181
	R 5'- CTTCTACGTCCGTTTTCATTGC-3'		
CDHsipho	F 5'- TTATGCGCTTTGCTRTTYAA-3'	CDHsipho specific	150
	R 5'- MGTTTTCATTGCTCCCATTT-3'		
Medium tailed myos gene	F 5'-AA CTTCGGGGATTTGTATGC-3'	Medium tailed myos gene	814
	R 5'- CAACAAATTGTATTGCATCAGC-3'		
Minor capsid protein	F 5'- AAGGGCTGTGCTTACTGGTG-3'	Minor capsid	147
	R 5'- ACCATTTGCCCTGCCATTTC-3'		
Head-tail connector protein	F 5'- AGGTCAGCAATCTCACAAGTGGCT-3'	Head-tail connector	163
	R 5'- TGGGGGAGAAATCCCACAGGAC-3'		
Tail fiber prorein	F 5'- GCACACAGTTAGGCGTTAGG - 3'	Tail fiber	152
	R 5'- GCTCCGTACCATCCCAACTA - 3'		

6.2.1.2.1 Polymerase Chain Reaction

PCR analyses were done using a thermal cycler (Biometra, Germany). The PCR was carried out with the primers in a total volume of 25 µl containing 2 µl of template DNA (~50 ng/µl), 2 µl forward and reverse primers (4 µM), 2 µl of 50 mM MgCl2, 1.25 µl of dNTPs (2.5 mM), 2.5 µl of 10 × PCR reaction buffer, 0.1 µl of 0.2 unit of BioTaq DNA polymerase (Bioline, London, UK), and 15.2 µl of DNase-RNase free water. DNA amplification conditions for all primers were: initial denaturation at 95°C for 2 min followed by 30 cycles of 92°C for 1 min, annealing at 55°C for 1 min, extension at 72°C for 90 sec, final extension for 5 min at 72°C, and hold at 8°C.

6.2.1.2.2 Detection of *C. difficile* lysogeny

To detect lysogeny of particular *C. difficile* phages, a single colony from a culture on a BHI blood plate was suspended in Chelex 100 (Sigma-Aldrich, USA) solution and incubated at 99.5˚C for 15 min to extract DNA. Detection of phage genes was done using PCR targeting φCDHS1 genes after extraction of phage DNA as described below (Section 6.2.1.2.3). PCR products were separated using a 1% w/v agarose gel containing 0.2 ml/L ethidium bromide. If a PCR a positive band was observed on the agarose gel then glycerol stocks of the lysogens were created and stored at -80°C until needed.

6.2.1.2.3 Detection of resistance *C. difficile* to φCDHS1

The resistance of cells from all colonies without detected lysogeny was confirmed using a method described before (Fischer *et al.*, 2013). Briefly each colony were transferred to about 8 ml of FA broth and incubated for 18h in 37˚C in an anaerobic chamber. After incubation, 500 µl of all cultures obtained this way were add to 3ml of soft agar, mixed and poured onto a BHI agar plate. When plates had set, serial dilution from 10^0 to 10^7 of the phage that the initial colony was supposed to be resistant to were spotted on the plate. Simultaneously, all phages serially diluted were spotted onto BHI agar plates covered with 3ml of soft agar containing 500 µl 18h culture of an ancestral culture or a phage susceptible *C. difficile* R20291 strain in FA broth. All plates were incubated for 24h in 37˚C in an anaerobic chamber. To determine whether any colony was resistant or susceptible to a particular phage PFU/ml counted from plate containing cells from this colony was compared to PFU/ml counted from plate containing an ancestral *C. difficile* R20291. Strains were considered as resistant if there was no any visible plaque spotted from the highest dilution which caused confluent lysis on plate containing ancestral *C. difficile* R20291.

6.2.1.2.4 Phage DNA extraction

To detect φCDHS1genes phage DNA was extracted by the phenol/chloroform/ isoamyl alcohol method from a high titter phage lysate 10^9 PFU/mL as described previously (Nale *et al.*, 2016). Then, 1ml of the lysate was transported into an eppendorf tube, 12.5µl of 1M magnesium chloride was added, after which the tube was gently mixed. Then the mixture was treated with 10µl of 30mg/ml DNaseI and 2µl of 100mg/ml RNase H (New England Biolabs, U.K). The tube was then briefly vortexed and incubated overnight at 37ºC. The following day, the mixture was filtered with a 0.22µm filter into a new tube.

Universal bacterial 16s rRNA PCR was performed to detect any contamination with the bacterial genome (Appendix 7). The following reagents were added in this exact sequence: 40µl of 0.5M Ethylene Di-amine Tetra-acetic Acid (EDTA), 5µl of 10mg/mL proteinase K (Fisher Scientific, UK), and 50µl of 10% Sodium dodecyl sulphate (SDS). The mixture was vigorously vortexed and incubated at 55^0C for 60 min and within this time the tube was vortexed twice, every 20 minutes during the incubation. An equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) (v/v/v) was added and the tube was inverted several times and centrifuged at 15000 xg for 5 minutes at 4^0C. This step was repeated 2-3 times until the white interphase disappeared. Then the aqueous layer on the top which containing the DNA was transported into a new tube and added to 1 mL of 95% ethanol and 50 µL of 3M sodium acetate solution (Fisher Scientific, UK) to precipitate the DNA. The tube was incubated in ice for 30 minutes and centrifuged at 21000xg for 20 minutes at 4^0C.

The entire liquid was removed carefully using a pipette and 1ml of 70% ethanol was added to wash the pellet. Following this, the isopropanol was carefully removed, and the DNA pellet was briefly air-dried and dissolved in 50µl of 5mM Tris-acetate pH 8-8.5. DNA concentration was determined using a Nano-Drop 1000 spectrophotometer (Thermo-scientific, USA), and stored at -20^0C.

6.2.2 *In vivo* virulence gene expression studies using the *Galleria mellonella* model

G. mellonella is an insect from the order Lepidoptera, the family *Pyralidae*, and is the only member of the genus (Kristensen *et al.*, 2007; Park *et al.*, 2017). Larvae of *G. mellonella* were obtained from Live Food UK Ltd. (Rooks Bridge, UK). On arrival, the larvae were stored immediately at 4°C and used within one week. Initially the larvae were weighed and individually selected with approximate weight of 0.25–0.30g for *in vivo* analysis (Abbasifar *et al.*, 2014). Before infection, the larvae were surface-sterilized by swabs dipped in 70% ethanol (Figure 6.1).

6.2.2.1 Bacterial Inoculum preparation, and *Galleria* infections

For the *G. mellonella* infection experiments, cultures of *C. difficile* R2029, a lysogenic *C. difficile* or phage resistant culture grown overnight in BHI were diluted 1/100 (to an optical density at 550) of 0.2. Then after, the culture was centrifuged at 15000 xg for 5 min and the resident pellet was re-suspended in cold BHI to a final concentration of 1×10^7 CFU/mL. Prior to inoculation with *C. difficile* strains into *G. mellonella* were washed

with phosphate-buffered saline (PBS) (Oxoid, England) pH 7.5. Bacterial inoculums with 10^5 CFU was administered by forced feeding (oral route) using a 10- µl Hamilton syringe pump in larvae (Figure 6-1). After infection, four randomly selected larvae were examined in each treatment at 37 °C for 0, 2, 24, 36 and 72 hours in plastic Petri dishes separately. The insects were kept without feeding throughout the experiment (Ramarao *et al.*, 2012). Two control groups of larvaes were treated in exactly the same time with pre-reduced BHI and untreated. After incubation, insects were considered dead when they lacking the ability to move and larval colour has changed to dark brown/black. Experiments were repeated three times and survival curves were plotted using the Kaplan-Meier method in GraphPad. Prism 6.

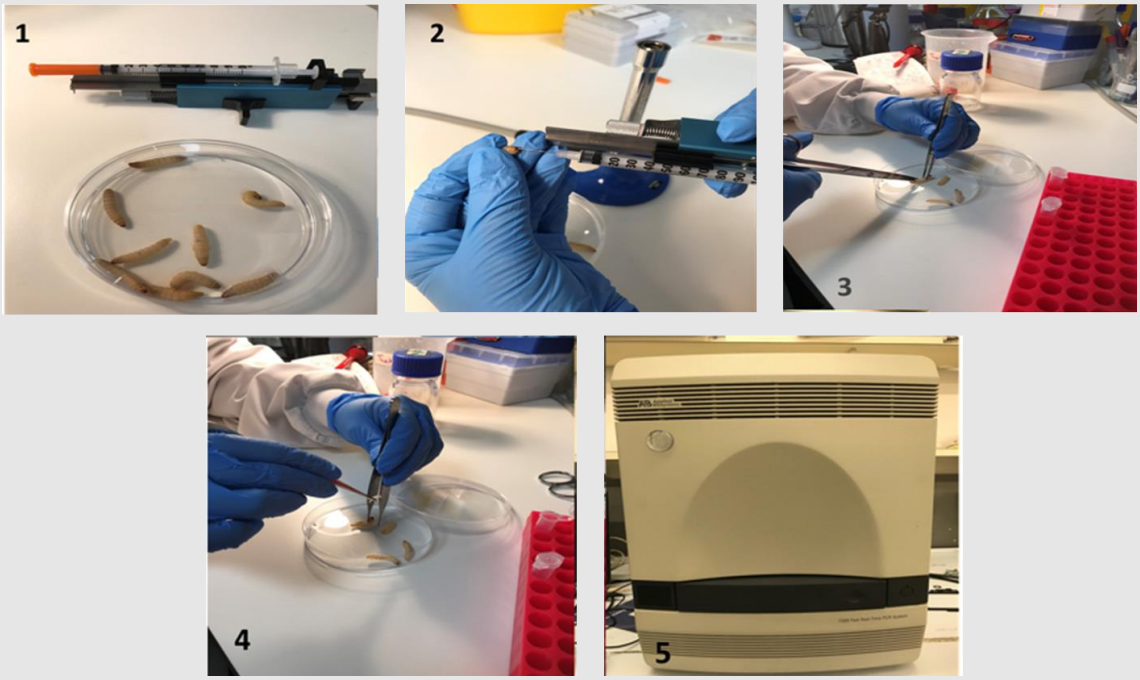


Figure 6.1. Images of forced feeding, dissection and removal of GI tract from *G. mellonella*, (1) and (2) show larvae in plate and forced feeding by three types of *C. difficile*. (3) Shows cutting between the last two prolegs up to the base of the head. (4) Shows prying open of the *Galleria* exposing the GI tract and the GI tract as a thin tubular structure. (5) Shows the 7500 Fast Real-Time PCR System device.

6.2.3 *C. difficile* strain recovery from the larvae

At the end of the experiment, all larvae were incubated had the number of dead counted and frozen at -20°C for 24 hours to ensure killing of the larvae. To determine CFU counts,

larvae from three treatments group and controls were thawed, dissected and had their gastrointestinal tract removed and suspended in 1ml PBS and vortexed for 30 seconds. 1 in 10 serial dilutions of each sample were conducted up to 10^{-7}. To detect CFU, all dilutions were spotted in 10µl aliquots on to Cefoxitin, Cycloserine and egg yolk (CCEY) agar plates (6.2.3.1) and incubated anaerobically at 37°C for 48h, with distinguishable colonies being counted.

6.2.3.1 Brazier s selective medium (CCEY)

To prepare Brazier's selective medium (CCEY) agar, 48g of (CCEY) powder medium (Oxoid, England) were suspended in 1l of distilled water and shaked well to dissolve thoroughly the ingredients. The medium was autoclaved and allowed to cool to 47ºC and then mixed with 50 ml of S2085 sterile egg yolk tellurite emulsion and 10 ml each of cycloserine and cefoxitin; this was mixed well before dispensing into plates, which were stored at 4ºC.

6.2.4 *G. mellonella* RNA extraction

Three larvae per treatment for each time point were subject to RNA extraction by using Stewart protocol with some modifications (Stewart *et al.*, 2002). Briefly, whole animals were homogenized in lysing matrix B tubes containing glass beads with 900 µl of Trizol reagent, 200 µl of cold chloroform and 200 µl of RNA stabilization solution according to the modified protocol. After sonication, all tubes were left at room temperature for 2 min and centrifuged for 15 min at 12,000 x g and 4˚C. To precipitate RNA, a clear top liquid was moved to a new clean tube and mixed with 500 µl cold isopropanol and incubated at R.T. for 15 min. The RNA precipitate was pelleted by spinning 10 min centrifugation at 12,000 x g and 4˚C and discarded the supernatant. Then, the RNA pellet was washed twice using 1ml of 75% ethanol and spun at 7,500 x g for 5 minutes at 4˚C. The RNA pellet was re-suspended with 30-50 µl of RNase, DNase free water and stored at -80˚C. A small amount was used to identify the integrity of total RNA by nonodrop (quantification) and by gel staining electrophoreses (qualification).

6.2.4.1 cDNA synthesis and qPCR

Complementary DNA (cDNA) construction and Real-time qPCR was done as described in section 2.7.1 and section 2.8 respectively. The *G. mellonella* transcriptome was screened with marker genes representing stress responses including those encoding

growth factors, and regulators of reproduction. All 16 gene-specific primers including the housekeeping gene 18S rRNA and actin are shown in appendix 8.

6.3 Results

6.3.1 Efficacy of lysogenic, resistant to φCDHS1 and wild type of *C.difficile* R20291 strains on *G. mellonella* infection

6.3.1.1 Mortality in *C. difficile* infected *Galleria* larvae depends on the pathogen type

The survivability of larvae from all three types of *C. difficile* R20291 groups are displayed in Figure 6-2. The susceptibility of *G. mellonella* to lysogenic, resistant to φCDHS1 and wild type of *C. difficile* R20291 isolates was examined. All larvae were forced feed at a *C. difficile* dose of 10^5 CFU and incubated at 37°C for up to 3 days. The results shown in Figure 6-2 show that all *C. difficile* infected larvaae had significantly lower survival rates when compared to BHI and uninfected larvae (as controls). These results show that virulence toward *G. mellonella* is species and strain-specific, so the lysogenic *C. difficile* R20291 exhibits significantly greater virulence than of wild type of *C. difficile* R20291 and φCDHS1 Resistant bacteria in *G. mellonella* model ($p < 0.0001$). The data indicate that those larvae infected with lysogenic *C. difficile* R20291 were the most susceptible to infection and showed a 43.1 % survival at 48 h compared with 54.17 % and 75 % survival for those larvae that infected with wild type *C. difficile* and *C. difficile* resistant to φCDHS1 respectively (Table 6.2).

Larvae that had been infected with lysogenic and WT *C. difficile* strains showed that survival decreased after 24 h from infection, while those infected with Resistant *C. difficile* to φCDHS1 strain decreased after 2 days. Notably, there was little difference in the virulence of lysogenic and wild type *C. difficile* compared with the larvae treated with BHI only. No significant mortality was observed in an uninfected group or in the infected group with BHI as all larvae still survived after 72 h of incubation at 37°C.

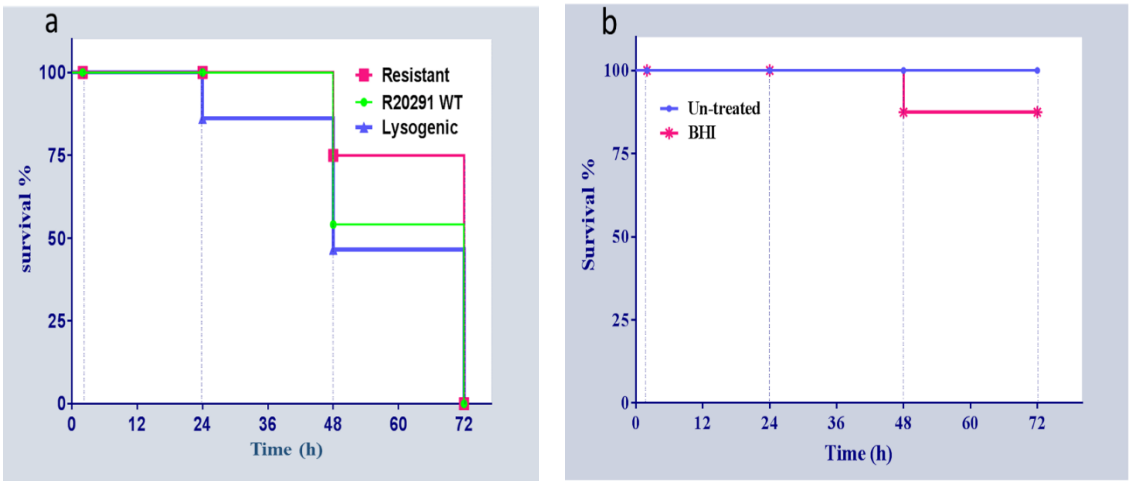


Figure 6.2. Survival curve showing (a) the impact of three groups of lysogenic, φCDHS1 resistant and wild type of *C. difficile* R20291 on the larvae life. (b) BHI infected larvae and untreated larvae used as a control. A Statistical significant difference in time of death was observed at (p value < 0.0001) using four larvae per group and experiment was repeated three times for each type of Clostridium difficile R20291. Survival data were plotted using GraphPad Prism 7.

Table 6.2. Survival proportion (%) of *G. mellonella* infected with lysogenic, Resistant to φCDHS1 and wild type of *Clostridium difficile* R20291. Three groups of 20 larvae were infected with 10^5 CFU of the indicated strain and kept at 37 C, and survival scored after (0, 2, 24, 48 and 72) h. BHI only infected larvae and untreated larvae used as a control.

Time (h)	Lysogenic	Resistant	R20291 WT	BHI	Un-ifected
0	100	100	100	100	100
2	100	100	100	100	100
24	86.11	100	100	100	100
48	43.07	75	54.17	100	100
72	0	6.25	0	87.5	100

6.3.1.2 Change in populations of *C. difficile* types in inoculated *G. mellonella* larvae

In order to monitor the fate of the injected *C.difficile* R20291 types, viable bacteria has been estimated by measuring all (CFU) counts in the gastrointestinal tract of larvae, 0, 2, 24, 48 and 72 hours post inoculation (Figure 6-3). We observed that there is a clear decrease in the CFU count at 0, 2 and 24 h post infection by all *C. difficile* types as a result of highly effective of immune defences of *Galleria.* Figure 6.3 shows at 2 days post infection, successive increases in bacterial CFU across all types of *C. difficile* were recorded. Specifically, this effect was noted significantly with wild type of *C.difficile* R20291 CFU counts at 4×10^6 in comparison to lysogenic and Resistant to φCDHS1 (6.4×10^5 and 2.8×10^5), respectively. After 72 h infection, the number of colonies rate in larvae treated with wild type of *C.difficile* R20291 group was higher than of those inoculated with lysogenic and resistant *C.difficile* R20291. However, after 3 days post infection, a greatest reduction in CFU across all the three treatment groups was found.

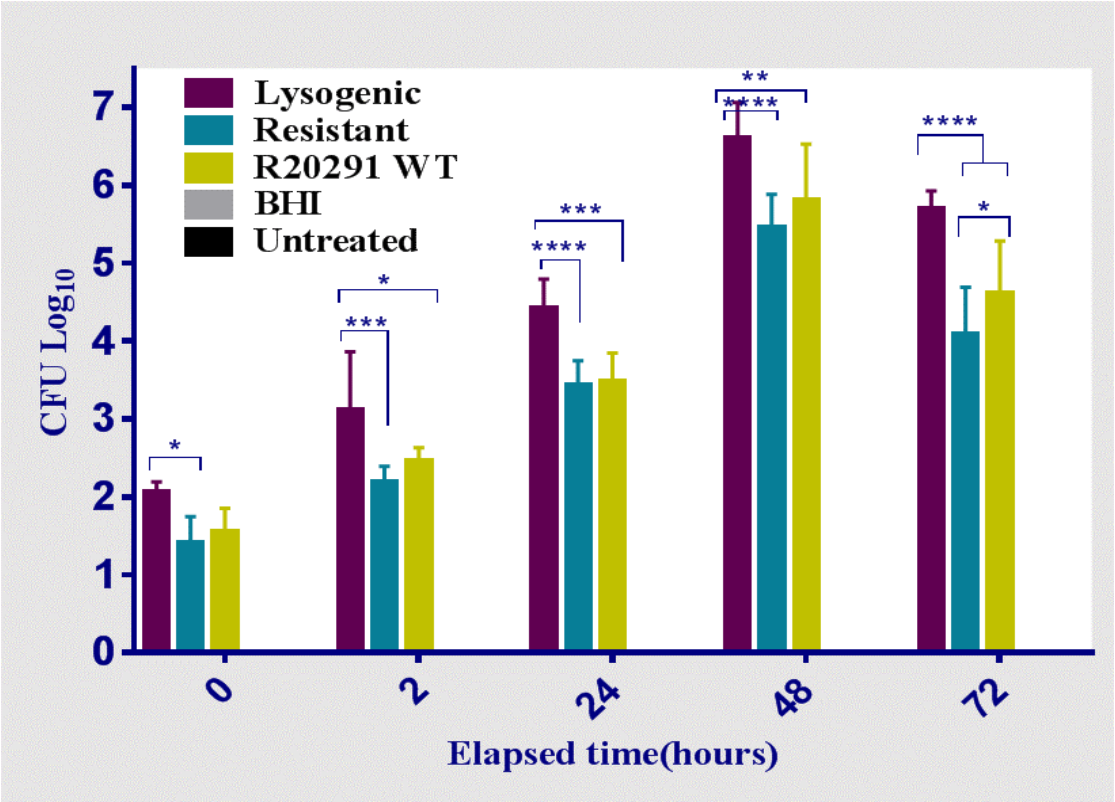


Figure 6.3. Survival curve showing the impact of three groups of lysogenic, bacteriophage φCDHS1 resistant and wild type *C. difficile* R20291 on the larvae life. BHI infected larvae and untreated larvae were used as a control. Larvae death in time was observed statistically Tukey's multiple comparisons test, two way ANOVA to ($p < 0.0001$) using four larvae per group and experiment was repeated three times. Error bars are SD of all the replicates. Data was analysed using Graph Pad Prism 7.

6.3.2 QRT-PCR Analysis of Insect Immunity-related Gene Expression

6.3.2.1 Lysogenic and φCDHS1 resistant *Clostridium difficile* R20291 strains induce antimicrobial peptide production in the host.

As *Galleria* can produce a broad spectrum of antimicrobial peptides in response to the microbial infection (Mukherjee *et al.*, 2010), this immune protection was used to determine whether lysogenic and φCDHS1 resistant bacteria were involved in determining he virulence of *C. difficile* R20291. Consequently, RT-PCR was done to assess expression levels of the genes for the immune effectors Moricin and Gloverin in the infected larvae. Larvae were infected with the three types of *C. difficile* R20291, and RNA was extracted at 0, 2, 24, 48 and 72 h post-infection. Transcriptional activation of immune-related genes is described as the fold change of expression in infected *Galleria*

relative to the un-infected control larvae and normalized using the housekeeping actin and 18S RNA gene (Figure 6.4).

RT-PCR on extracted mRNA showed that infection with all types of *C. difficile* R20291 resulted in an up-regulation of expression of the genes for the immune-related peptide compared to the levels in a BHI-injected control, with a significantly increased expression of gloverin in Lysogenic infected larvae than in the φCDHS1 resistant bacteria and wild type infected larvae especially after 48 and 72 h. The amounts of gloverin mRNAs were found to be induced about ~23-fold and ~33-fold in lysogenic *C. difficile* infected larvae at 48 h and 72 h post-infection. The moricin gene was also transiently expressed in larvae that were inoculated with the lysogenic and φCDHS1 Resistant *C. difficile* R20291 and wild type. In general, relative transcript levels in the larvae as examined by RT-qPCR showed up-regulation of all *C. difficile* R20291 in comparison to un-infected larvae at all time points.

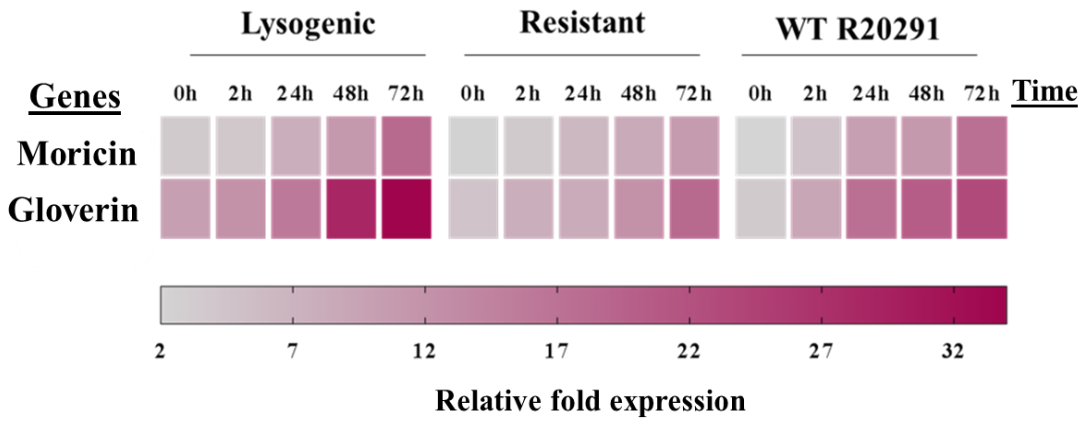


Figure 6.4. Expression of antimicrobial peptide genes in the infected larvae. The expression of Moricin and Gloverin genes was assayed in lysogenic, φCDHS1 resistant and wild type of *C. difficile* R20291 larvae by quantitative real-time RT-PCR. Basal expression in the infected larvae was calculated as a fold change relative to un-infected control larvae and normalized to the Actin and 18S rRNA housekeeping gene. The color gradient indicates fold changes in gene expression, with increasing red intensity representing up-regulation and decreasing red intensity representing own regulation. The figure is a representation of the mean fold change values of three independent experiments.

6.3.2.2 Infection with *C. difficile* R20291 types induces *G. mellonella* genes encoding growth hormones

It has been established, previously, that more pathogenic strains produced greater responses than non-pathogenic strains in *G. mellonella* larvae (Walters and Ratcliffe, 1983). In this experiment, to determine whether three types of *C. difficile* R20291 pathogenesis interferes with the endocrine regulation in *G. mellonella*, the expression of genes related to juvenile hormone group and ecdysteroides was monitored. Figure 6.5 shows the relative quantification of growth hormones in *G. mellonella* infected with lysogenic, phage resistant and wild type *C. difficile* R20291 by real-time quantitative RT-PCR. The expression of growth hormones genes was measured as transcript levels normalized against that of the 18S RNA and the actin housekeeping genes. The expression pattern of Juvenile hormone group in all RNA extracts of *G. mellonella* changed according to the time following infection with *C. difficile*. Indeed, analysis showed that the levels of Juvenile hormone-inducible (JHI), Juvenile hormone binding protein 3 (JHBP3) and Juvenile hormone binding protein 1 (JHBP1) were higher in larvae infected with the lysogenic *C. difficile* R20291 than in those infected with φCDHS1 resistant and wild type of *C. difficile* R20291 relative to un-infected larvae after one to three days post-infection (Figure 6.5 a). Importantly, infected larvae with lysogenic *C. difficile* R20291 after 48 and 72 h showed fold changes of 22.4, 18.2 and 12.8, 18.5 for Juvenile hormone-inducible (JHI) and Juvenile hormone binding protein 3 (JHBP3), respectively, which mean that the genes are increased in expression compared to the φCDHS1 resistant and wild type of *C. difficile* R20291. Moreover, with the lysogenic infected larvae, the highest level of genes expression was attributed to (JHI) and (JHBP3) at 24 hr with a 5.5 and 6 fold change respectively greater than the untreated sample. There was a transient impact of bacterial infection in larvae for ecdysteriod gene expression (Figure 6.5 b). Decreased levels of most genes expression were recorded throughout the period of 0 and 2 hours of three types of *C. difficile* R20291 infection. Furthermore, no significant difference was found in transcript levels of juvenile hormone and ecdysone between φCDHS1 resistant and wild type of *C. difficile* R2029.

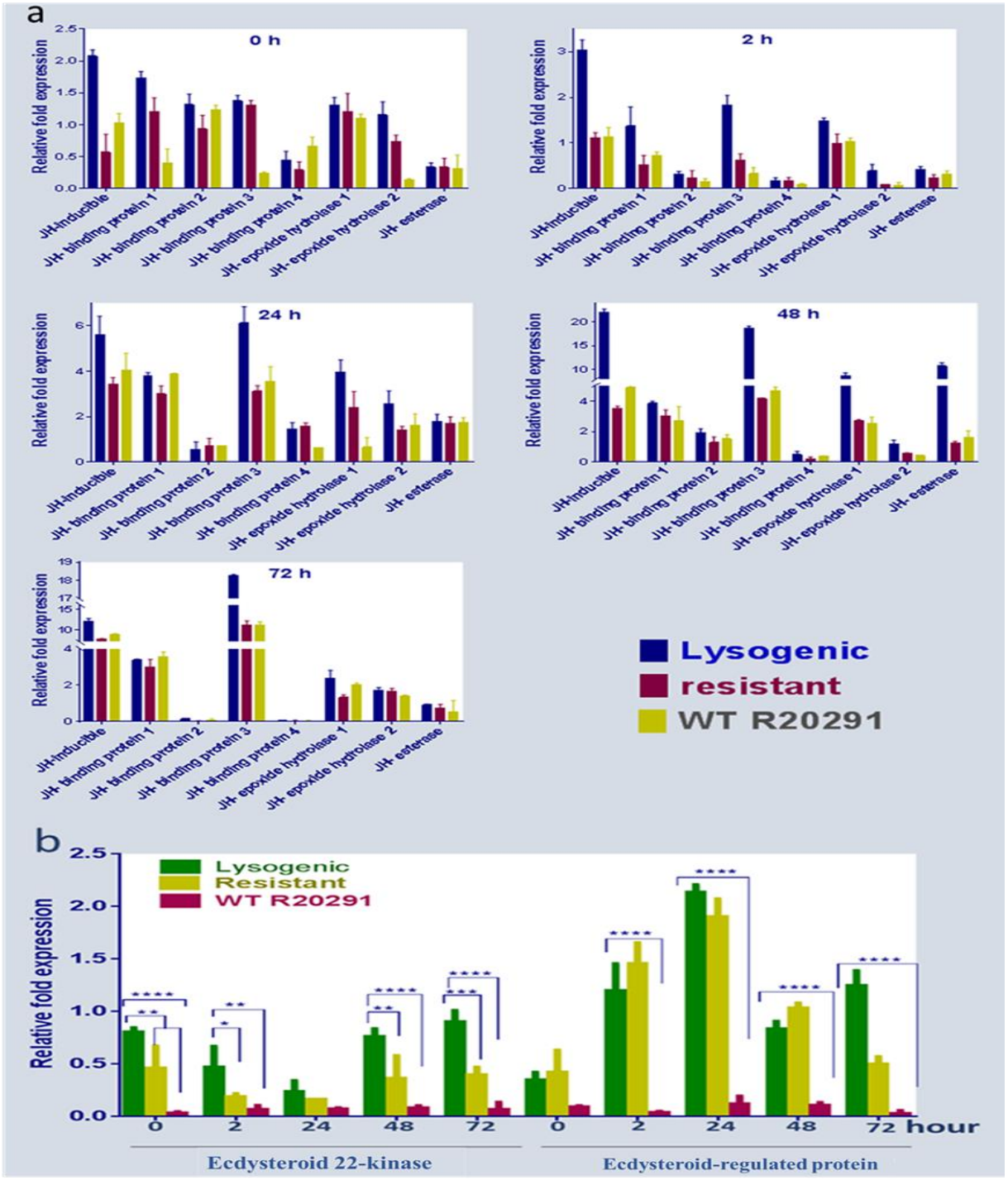


Figure 6.5. Relative gene expression ratios of genes encoding growth hormone associated proteins in *G. mellonella* following *C. difficile* R20291 infection determined by q-RT-PCR .The expression levels of genes encoding (a) JHI, JHBP 1, 2, 3, and 4, JHEH 1 and 2, and JHE and (b) Ecdysteroid 22-kinase and ecdysteroid-regulated protein, in *G. mellonella* larvae following forced feed with all *C. difficile* R20291 types were monitored at different time points after infection. Significantly induced expression levels are calculated relative to control uninfected larvae. Expression of each gene was normalized to the expression of the housekeeping genes 18S rRNA and actin. Results represent means of three independent determinations with standard deviations ±SD.

6.3.2.3 Lysogenic *C. difficile* R20291 induces growth factors gene expression in *G. mellonella*

To further investigate the pathogenicity and virulence among *C. difficile* R20291 types on *G. mellonella* development, the expression levels post- infection with three types of *C. difficile* for two growth factors -related genes, GME-string_Contig_704.0 and GME-string_Contig_233.0, were analyzed by RT-PCR. These growth factors represent as signal transduction in *G. mellonella.* Significantly induced expression levels are calculated relative to control larvae un-infected. Values were normalized against the expression levels of the housekeeping genes 18S rRNA and actin (Figure 6.6). Expression of the genes coding for putative stress-management factors (Contig_704.0 and Contig_233.0) showed clear increases in those larvae treated with lysogenic *C. difficile* at 72 hr with a change 13.79 -fold and 8.9, respectively greater than the untreated larvae.

Although of the induction in the gene expression of the host Contig_704.0 and Contig_233.0 also found at 72 h following φCDHS1 resistant and wild type of *C. difficile* R20291 infection, this induction was still lower than of which infected with lysogenic bacteria.

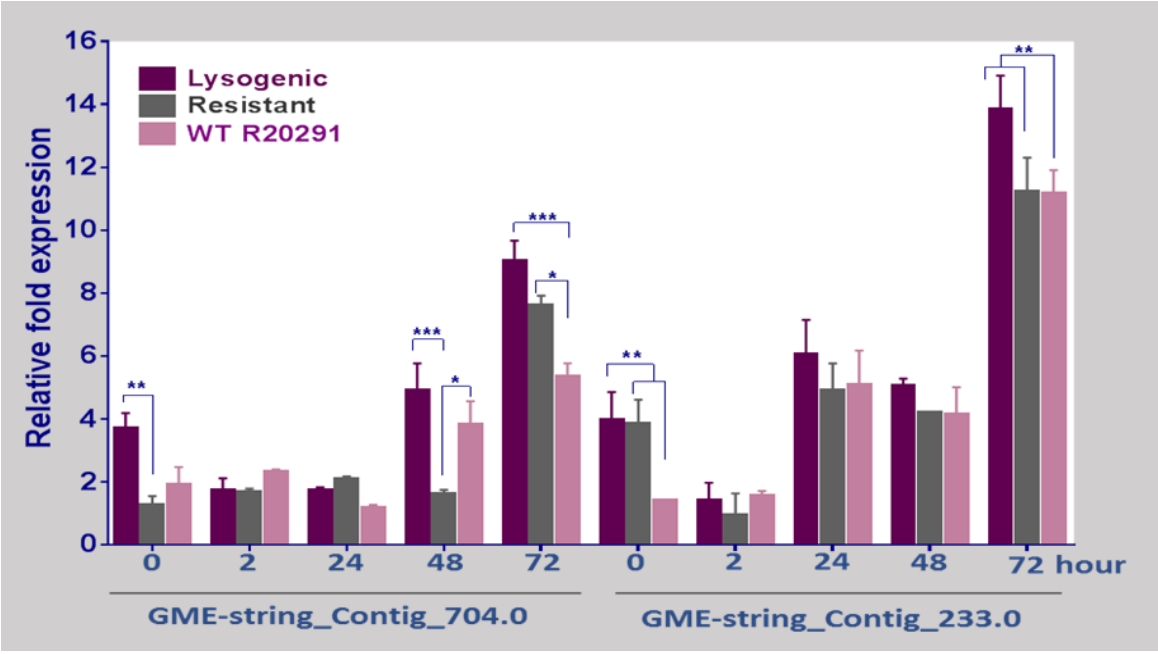


Figure 6.6. Quantitative RT-PCR analysis of genes encoding growth factors transcripts in total RNA isolated from *Galleria mellonella* larvae at different time points after infected with different *C. difficile* R20291 types. The expression levels of GME-string_Contig_704.0 and GME-string_Contig_233.0 was normalized to the expression of the housekeeping genes, actin and 18S rRNA. The error bars indicate standard deviation from three independent experiments.

6.4 Discussion

6.4.1 *G. mellonella* larvae can be used as an alternative model to assess *C. difficile* pathogenicity

G. mellonella larvae are being increasingly used as a simple *in vivo* model to study the virulence of wide range of bacterial pathogens in different studies, including gram-positive and gram-negative such as *Staphylococcus aureus* (Ferro *et al.*, 2016), *Streptococcus pneumoniae* (Evans and Rozen, 2012), *Streptococcus pyogenes* (Olsen *et al.*, 2011; Loh *et al.*, 2013), *Listeria monocytogenes* (Mukherjee *et al.*, 2010; Martinez *et al.*, 2017), *Bacillus cereus* (Salamitou *et al.*, 2000), *Mycobacterium abscessus* (Meir *et al.*, 2018), *Enterococus faecalis* (Aperis *et al.*, 2007), *P. aeruginosa* (Koch *et al.*, 2014), *Shigella* spp. (Barnoy *et al.*, 2017), *Klebsiella pneumonia* (Wand *et al.*, 2013; Diago-Navarro *et al.*, 2014), *Legionella pneumophila* (Harding *et al.*, 2013), *Yersinia pseudotuberculosis* (Champion *et al.*, 2009), *Yersinia pestis* (Desbois and Coote, 2012; Erickson *et al.*, 2011), *Francisella tularensis* (Aperis *et al.*, 2007), *Burkholderia pseudomallei* (Thomas *et al.*, 2013) and Uropathogenic *Escherichia coli* Alghoribi *et al.* (2014). Moreover *G. mellonella* is now employed to assess microbial virulence and pathogenicity of fungi such as *Aspergillus fumigatus* (Slater *et al.*, 2011), *Candida* spp. (Chen *et al.*, 2011), *Cryptococcus neoformans* (Mylonakis *et al.*, 2005; Vu and Gelli, 2010) and yeasts (Cotter *et al.*, 2000). In this current study, the larvae of the greater wax moth (*Galleria mellonella*) was established as a model host for the investigation of lysogenic, φCDHS1 resistant and wild type *C. difficile* R20291 pathogenesis. The virulence of each *C. difficile* types in the wax moth larvae host was estimated by time-to-death assays by monitoring infected larvae over 3 days. Control larvae that were inoculated with BHI remained alive except one after 72 h with no signs of illness. This study reports for this first time that *G. mellonella* is susceptible to infection with lysogenic *C. difficile* R20291 and φCDHS1 resistant bacteria. Here, it was demonstrated that innoculation of larvae with lysogenic *C. difficile* after 2 days resulted in statistically significant decreases in survival. Similarly, it is observed that similar numbers of larvae died over a three-day period when infected with lysogenic and wild type *C. difficile* R20291 (Figure 6.2 & Table 6.2). Indeed, reports suggest that temperate prophage carriage is associated with increased bacterial virulence (Davies *et al.*, 2016) and the expression of some phage-encoded genes like Shiga toxin (stx1, stx2) and cytotoxins can boost the pathogenicity (virulence) of *E. coli* (Wagner *et al.*, 2001). The one notable result was bacteria contain prophage (lysogenic) (Figure 6.), which demonstrated

unexpectedly high virulence in *G. mellonella* larvae than others and consequently it is impossible to use temperate phage in phage therapy.

6.4.2 *C.difficile* can induce immunological responses in *G. mellonella*

The innate immune response of *G. mellonella* were evaluated after infection by analyzing transcript levels of two AMPs, moricin and gloverin protein genes using RT-qPCR. Both peptides are regarded as a key components of insect innate immunity, which combat invading microbes in the hemolymph (Vilcinskas, 2011; Jiang *et al.*, 2010). Although all of the *C. difficile* types used in this study resulted in the induction of moricin and gloverin production, there were slight differences in the rate of the response (Figure 6.4). Lysogenic bacteria appeared to result in the induction of higher levels of gloverin and moricin, while infection with the phage resistant and wild type led to less levels of both peptides production. These findings are consistent with other studies showing that peptides with strong antibacterial activity, such as lysozyme, cecropin, gloverin, galiomycin, and moricin, which were strongly induced in the virulence pathogens a uropathogenic *E. coli* (UPEC) larvae (Heitmueller et.al 2017; Reigstad, 2007; Zitzmann, 2017(Zitzmann *et al.*, 2017)). Furthermore, previous study monitored that virulent *L. monocytogenes* enhanced expression of moricin and gloverin genes in *G. mellonella* (Mukherjee *et al.*, 2013).

As had been expected, some differences were present in the expression genes encoding growth hormone (juvenile hormone and ecdysone) profiles. Results showed that the lysogenic type of *C. difficile* R20291 was able to postpone the development of the infected insect host whereas resistant and wild type bacteria stimulate premature formation of pupae. Thus the induction of expression of genes mediating processing or binding of either juvenile hormone or ecdysone reveal that the lysogenic type can interfere with the endocrine system of the infected host resulting in postponement of its development. This is in agreement with the study of Mukherjee et al. (2010), in which virulent *L.monocytogenes* stimulated expression of genes in *G. mellonella* which are related to binding and metabolism of juvenile hormone. However, the expression of juvenile hormone was demonstrated to be maintained at a constant level when *G. mellonella* larvae were challenged with b-glucan (Mowlds *et al.*, 2010). In the current study the genes whose expression were induced in response to the different types of *C. difficile* R20291 infection including those encoding proteins related to stress responses and growth factors examined. There were significant increases in the expression levels of GME-

string_Contig_704.0 and GME-string_Contig_233.0. in larvae infected with lysogenic bacteria after 3 days. This is in accord with previous finding that these three genes are strongly expressed following exposure to this pathogen and that RNAi knock-down results in faster. Overall, *in vivo*-testing using *G. mellonella* larvae infected with different types of *C. difficile* R20291 showed that lysogene infected larvae had reduced survival after 48 h whilst less virulente resistant and wild type *C. difficile* shows less mortality to larvae. In addition the more pathogenic strains (lysogenic) produced greater immune responses than the relatively less-pathogenic forms (resistant and wild type).

Chapter 7
General Discussion, Conclusion and Future work

Chapter 7 General Discussion, Conclusion and Future work

7.1 Introduction and Discussion

Clostridium difficile is a Gram-positive, spore-forming obligate anaerobe bacterium that has emerged as the leading cause of antibiotic-associated nosocomial diarrhea in developed countries. *C. difficile* infections usually occur as a consequence of normal bowel flora distortion and reduction of the intestinal microbiota caused by antimicrobial drug therapy, with symptoms ranging from no effects to high incidences of morbidity and mortality. Since 2003, highly virulent strains, such as NAP1/027, have become the major cause of severe CDI outbreaks not only in Europe, but also in North and Central America, Asia, and Australia (Clements *et al.*, 2010). Interestingly, the high virulence of *C. difficile* is attributed to the ability of some strains to produce in addition to the main virulence factors toxins A and B (encoded by *tcdA and tcdB* genes) and their regulator genes (*tcdR*, *tcdC* and *tcdE*), another toxin (CDTa & CDTb) which are encoded by the *cdtA* and *cdtB* (McDonald *et al.*, 2005; Carter *et al.*, 2007). In addition to the severity of toxins A and B which can cause disruption and damage of the actin cytoskeleton in human intestinal epithelial cells, hypervirulent strains (027 ribotype) have A and B toxins as well as h binary toxin CDT. Collectively, compared with non-outbreak strains of *C. difficile* this trio of toxins can reduce susceptibility to the usual fluoroquinolone treatment antibiotics(McDonald *et al.*, 2005).

The main reason of *C. difficile* to persist and resist antibiotics or host immune system is due to its presence in spore form which ensures its ability to survive adverse conditions (Paredes *et al.*, 2012). Furthermore, the mobile genome of *C. difficile* such as transposons, insertion sequences and (pro) phages can also contribute to development antimicrobial resistance (Sebaihia *et al.*, 2006). Because of *C. difficile* resistance to antibiotics, many researchers have resorted to searching for potential alternative other than antibiotics to eliminate these bacteria. Therfore, the Eliava Institute has extensively used different lytic bacteriophage in clinical treatment of bacterial pathogens such as *S. aureus, E. coli, Streptococcus* spp., *S. dysenteriae*, *Salmonella* spp., *Proteus* spp., *Enterococcus* spp. and *P. aeruginosa* (Kutateladze and Adamia, 2008). Moreover, it becomes of interest that bactericine and viruses (bacteriophages) appear to be able to target *C. difficile* (Hargreaves and Clokie, 2014).

Although there are many obstacles to the clinical use of lytic bacteriophages as an antimicrobial therapy in humans, a recent study confirmed that bacteriophages can be used to treat bacterial infections because of its ability to selectively infect and kill bacteria without any negative effect on human or animal cells (Calap and Martínez, 2018). Recently, encouraging results have been obtained by Monteiro et al. indicate and enhance to use of temperate phages and their lytic cycle in therapy (Monteiro *et al.*, 2019). In regard to this point, studying phage-bacterial interactions is essential to understand phage behaviour which in turn can potentially help to exploit these phages in therapy. Hence, RNA sequencing (RNA-Seq) technology has been used to monitor these host-virus interactions at a high level of accuracy through analyse transcriptional changes of in both organisms.

This thesis consists of three sections; in the first section, an RNA-seq analysis was adopted, in an attempt to monitor transcriptomic changes in the *C. difficile* R20291 infected with φCDHS1 during one step growth curve. Therefore, total RNA were directly extracted from both *C. difficile* R20291 infected with φCDHS1 and uninfected as control at different time points within lytic cycle. The second section focused on transcriptional responses of *C. difficile* R20291 virulence genes which can be induced by infection with φCDHS1 in vitro using Real-time RT-PCR (qPCR). Subsequently, the result of virulence genes expression which obtained in qPCR were compared to those in RNA-seq. The last section of this study aimed to detect the severity of pathogenicity of lysogenic, phage CDHS1 resistant and wild type *C. difficile* R20291 using silkworm *Galleria mellonella.* To achieve this test, in addition to the survival test, the quantification of gene expression of several genes of *G. mellonella* associated to immune defence using qPCR technique were also performed.

In terms of the one step growth curve experiment, different phases (latent, exponential and stationary) of φCDHS1 growth curve were observed during infection with *C. difficile* R20291 strain / 027 ribotype (chapter 3). One step growth curve experiment was also conducted to calculate the burst size of φCDHS1 and burst sizes and number of phages which released per infected cell. Although there are relatively experiments published which have used the one-step growth curve so far *C. difficile* phages for this species as well also for *E. coli* phage, most show considerable variation in latency periods and burst sizes during infection (Bolger-Munro *et al.*, 2013; Mahony *et al.*, 1985; Abedon *et al.*, 2001). Moreover, previous literatures showed that the length of the latent period varies

according to the phage species, the incubating temperature and the condition of the host cell culture medium (Kutter and Sulakvelidze, 2004); (Weinbauer, 2004). One step growth of φCDHS1 in this study revealed that the burst size (Figure 3.1) was ~ 37 phages released per infected cell and latent period of 20 min which agreed with the findings of previous study (Thanki, 2016). While this was still significantly higher than other temperate phages of *C. difficile* such as ΦC2, ΦC5, ΦC6 and ΦC8, which showed small burst size ranged from 5-36 phages released per infected cell in previous study (Goh *et al.*, 2005b). Unlike φCDHS1, the temperate phage of *C. difficile* 56 released about 122 during one step growth (Mahony *et al.*, 1985). These are a strong evidence to the variation in burst size for phages even within same bacterial species.

Chapter 4 described the effects of infection φCDHS1 on transcriptional responses of virulence genes in *C. difficile* R20291 using Real-time RT-PCR (qPCR). Depending on the phage replication strategy, their impact on bacterial pathogenicity can be drastically different as it can range between adding extra toxin genes through prophage (lysogenic) or affect virulence gene expression during the lytic cycle. Real-time RT-PCR results of impact of temperate φCDHS1 on gene expression of 10 virulence genes of *C. difficile* during one step growth revealed that most *C. difficile* virulence genes were down-regulated vs. uninfected control at different time points within growth curve with some exceptions (Figure 4.2). For example, the expression of the *tcdA, cdta* (Binary toxin) and *tcdE* genes was at decreased level in early, intermediate and late log phase (20 to 50 min). Likewise, the level of *tcdB* transcription showed also a clear reduction at all time points. Although these results are inconsistent with the results of previous studies which confirmed that infection with bacteriophage can stimulate and enhance toxin production (Wagner and Waldor, 2002; Goh *et al.*, 2005a), it is possible that these phages carry genes encoded virulence factors such as exotoxins (Bishai, 1988; Wagner *et al.*, 2001). On the contrary, which was completely consistent with the research results, it was demonstrated that there was a clear suppression of toxin production when using a temperate φCD27 (Meader *et al.*, 2013) or during prophage stage of φCD119 (lysogen) in *C. difficile* (Govind *et al.*, 2009). These result suggest that there is a potential to use temperate phage therapeutically, as the work in the study suggests that temperate phage can inhibit toxin production in *C. difficile*. Also, for genes encoding toxins, the results showed a remarkable decrease in the level of gene expression of the regulatory genes which responsible for supporting (positive regulator, *tcdR*), suppressing (negative regulator,

tcdC) and releasing (*tcdE*) toxins. In contrast, the prophage φCD38-2 in a ribotype 027 strain results in the stimulation of toxin production (Sekulovic *et al.*, 2011). While the *tcdE* gene encodes a protein which is require to release of TcdA and TcdB proteins from *C. difficile* (Govind and Dupuy, 2012). A previous study by Meader et al. (2013) on CD27 phage has not only indicated to the drastic absence of vegetative *C. difficile* cells in cultures that treated with phage, but also to a significant increase in the number of spores compared to the untreated control bacteria (Meader *et al.*, 2013).

Flagella related *fliA* gene expression was also studied as it is considered a *C. difficile* virulence factor because of its important role in the motility of bacteria. The expression of *fliA* gene of infected cells during log phase (30, 40 and 50 min) was down-regulated significantly compared with uninfected *C. difficile* R20291. This result was consistent with another experiment which demonstrated that *fliA* and *fliD* genes levels was down-regulated significantly in lysogenic strain of *Escherichia coli* (Cao *et al.*, 2014). This suggests that flagella protein which encoded by *fliA* gene has an extracellular role during the phage lytic cycle, and the level of this protein will decrease as a result of a significant decrease in bacteria population due to phage infection.

Chapter 5 in this study describes intracellular bacteriophage-host interaction changes between *C. difficile* R20291 and φCDHS1 at a level of transcriptome using RNA-seq during the one step growth experiment. Over recent years, the Investigation of the transcriptional responses of the host to the phage infection has become a robust area of transcriptional research from RNA-seq (Leskinen *et al.*, 2016). Hence, RNA-seq has been performed and RNA-seq data in turn has been analysed using the read count estimates provided that were generated using Kallisto software (Bray *et al.*, 2016). The results of differential gene expression with analysis of RNA- seq data were similar to those of Partek. In this study, the preliminary results revealed that the expression mode of majority of R20291 genes (3514) were varied between no, low, moderate, or high for both R20291 infected with φCDHS1 and uninfected R20291 (as control) at different time points. Briefly, most of phage-induced differentially expressed genes of *C. difficile* which were analysed via RNA-seq analysis, were among those showing very low expression levels especially for genes encoding the metabolic processes, membrane and permease proteins especially at beginning of growth curve (0 and 10 min) (Table 5.1 A). As a previous study indicated to the importance genes encoding metabolic pathways in uninfected cells at early logarithmic growth phase which reach to down levels at the late logarithmic growth

phase (at 50 min post infection) (Ravantti *et al.*, 2008). This strongly suggests that infection with a temperate φCDHS1 inhibits the *C. difficile* metabolic processes in cell during growth cycle. The expression of phage gene during one step growth curve has been also determined. Phage-induced differentially expressed genes was observed across the whole phage genome and seemed to be greatly higher than for most of the bacterial genes (Table 5.1 B, Appendix 6). At 0 and 10 min after infection, phage early genes such those encoding Cro / C1-type transcriptional regulator, single stranded DNA and helicase were expressed significantly shortly after viral DNA entered the cell relative to control (Table 5.1, Figure 5.8). A previous study showed that the viral *Cro* gene is required at the early stage for lytic development and expressed increasingly from 5 to 10 min (Liu *et al.*, 2013a). Genes in the same pathway, related to lysis enzyme system, holin, endolysin and structural protein (tail fiber protein) often exhibit similar expression patterns at 20, 30, 40 and 50 min (intermediate and late genes) after φCDHS1 infection. According to holin and endolysin function, these proteins are required to lyse the host cell and subsequently releasing the progeny virions (Rydman and Bamford, 2003; Ziedaite *et al.*, 2005). In general, all results indicated that most φCDHS1 genes showed a high level of expression, while on the contrary, bacterial genes showed a low expression at all infection timecourse stages. This could be due to that bacteriophage genes interact and inhibit bacterial gene expression in order to synthesis and produce new virions. Chapter 5 also describes a comparison of changes in virulence gene expression in *C. difficile* infected with φCDHS1 using RNA-seq versus qPCR (Figure 5.3). A strong correlation was found between RNA-seq and qPCR with some exceptions. The expression of genes encoding terminase small, large Sub units, Coat, head-tail connector and minor capsid proteins were dramatically increased after infection to maximum level to early stationary phase, and they are considered to be intermediate and late phage genes (Figure 5.4 and 5.5). Likewise, the expression of late genes encoding tail fiber, tail tape measure, holin, scaffolding protein, integrase and sigma70/ sigma F-like proteins also showed a gradual increase in levels after middle stage within growth curve (Figure 5.6 and 5.7). As it is known, intermediate and late phage gene products are essential and required for process such as package of phage DNA in the capsid, virion assembly, attach to host cells or releasing of virions (Shen *et al.*, 2012; Cardarelli *et al.*, 2010; Doval and Raaij, 2012; Delisle *et al.*, 2006). The results also revealed that the expression level of phage early genes encoding DNAse/Helicase and single-stranded DNA-binding protein and Cro/C1-type transcriptional regulator were high initially and then decreased in the later stages (Figure

5.8). These genes are required to express at early stage during growth curve for stimulation of phage DNA replication, protection of replicating DNA from nucleases and turning to lytic cycle (Gauss *et al.*, 1994; Soengas *et al.*, 1995; Liu *et al.*, 2013b).

The final part of this study, chapter 6, aimed to evaluate whether a temperate phage CDHS1 can add virulence genes to the *C. difficile* R20291 during the lysogenic cycle (prophage), and compared virulence of wild type *C. difficile* and strains resistant to φCDHS1 and wildtype *C. difficile* R20291 using *G. mellonella* as an infection model. To estimate the virulence of each *C. difficile* types and monitored over 3 days by time-to-death assays. The result showed that numbers of survived larvae inoculated with lysogenic *C. difficile* significantly were decreased in survival after 2 days from infection while their numbers of dead larvae were similar to those with wild type *C. difficile* R20291 over a three-day period (Figure 6.2 & Table 6.2). Previously, Wagner and his group have stated that phage-encoded genes for Shiga toxin and cytotoxins can add extra virulence to pathogens such as *E.coli* O157:H7 (Wagner *et al.*, 2001). Moreover, the presence of prophage in non-toxigenic *Streptococci* strains can convert it to a toxigenic strain (Frobisher and Brown, 1927). Based on the current findings, *C.difficile*-carriage CDHS1 (a lysogen) became a more pathogenic bacteria than wild type, and this may be due to the addition of the virulence gene through the CDH1 genome to the wild type *C.difficile*.

Overall, *in vivo*-testing using *G. mellonella* larvae infected with different types of *C. difficile* R20291 showed that lysogen infected larvae had reduced survival after 48 h whilst less virulent, resist to φCDHS1 and wild type *C. difficile*, shows less mortality to larvae. In addition the more pathogenic strains (lysogenic) produced greater immune responses than the relatively less-pathogenic forms (resist to φCDHS1 and wild type).

7.2 Conclusion

It was found that the expression levels of most virulence genes especially genes related 3 main toxins (TcdA, TcdB and CDT) in *C. difficile* R20291 infected with temperate phage CDHS1 were down-regulated control at early, middle and late stages using Real-time RT-PCR. Likewise, flagella related *fliA* gene expression has also been showed down-regulation at late stages of growth curve. High expression of spore related *SpmA* gene at late stages during infection as result of stress due to phage infection were also

determined. All these findings boost the idea that temperate phage CDHS1 has the ability to inhibit toxins production in *C. difficile* R20291 during lytic cycle. In addition, depending on bacteriophage-host interaction changes level of transcriptome using RNA-seq, this work confirmed that the infection of *C. difficile* R2029 with temperate φCDHS1 revealed very low expression levels especially for genes encoding the metabolic processes, membrane and permease proteins especially at beginning of growth curve (0 and 10 min). This strongly suggests that a temperate φCDHS1 inhibits the metabolic processes in cell during growth cycle which in turn lead to terminate the growth process in host cell. This approach also add another research which confirm relative similarity between RNA-seq and Real-time RT-PCR (qPCR) analysis.

It was also found that lysogenic *C. difficile* R20291 containing phage CDHS1 was able to postpone the development and accelerate mortality in *G. mellonella* in relative to wild type and resist to φCDHS1 *C. difficile* R20291. Finally, the current data provides new insights into virulence of temperate phage and suggest that although temperate phages are highly abundant and are more readily available than lytic phages in phage therapy, it can be used under specific condition (highly MOI). At this point, there are no indications that *C. difficile* interaction with temperate phages would make it any less 'aggressive' than virulent phages.

7.3 Suggestions for future work

This section contains some recommendations for future work based on results obtained in the PhD thesis.

1. Examine the effects of bacteriophage CDHS1 on gene expression in *C. difficile* during lysogenic cycle and compare it with results that obtained from RNA-seq and Real-time RT-PCR (qPCR) in this research.
2. Apply more bioinformatics resources to visualize the whole genome for both *C. difficile* and φCDHS1 according to their gene ontology (genes with their products and biological functions).
3. Detection of gene expression on the level of protein product for *C. difficile* toxins, A, B, binary toxins and their gene regulation products during infection with φCDHS1 in lytic cycle to investigate if the toxin products decreased and corresponded to their own down- regulation gene expression which obtained in this research or not!

4. Detection of gene expression on the level of protein product for *C. difficile* toxins, A, B, binary toxins and their gene regulation products during infection with φCDHS1 in lytic cycle to investigate if the toxin products decreased and corresponded to their own down- regulation gene expression which obtained in this research or not!. In addition, and depending on these result, the therapeutic potential of φCDHS1 will be explored.
5. Future work will also aim to clarify the impact of φCDHS1 prophage carriage on the physiology of the hosts, and on how they relate to each other.
6. Further studies should be undertaken to investigate of other genes in *Galleria mellonella* that possible affected due to infection with *C. difficile* alone, and *C. difficile* infected with φCDHS. But unfortunately, it is difficult to specify these genes as no references indicated for that.

Chapter 8 Appendices

Chapter 8 Appendices

Appendix 1

Table. A list of all media used for growth *C. difficile* throughout this project.

All media were made up to 1000 ml using distilled water except for the semi-solid Brain Heart Infusion(up to 250 ml), and Prestige Medical Autoclave, England was used to autoclave under 121ºC, 15 psi (1bar) for 15minutes.

Media	Composition of media in 100 ml	Manufacturer	The purpose
Brain Heart Infusion broth	37 g BHI broth agar	Oxoid, UK	Growing *C. difficile*
Brain heart infusion agar, BHI 3.7% and agar 1%	37 g BHI broth 10 g Bacteriological agar	Oxoid, UK Oxoid, UK	- For Plaque assay and spot test of *C.difficile.* - For propagation of *C. difficile* phages.
BHI 1% (w/v) Agar 7% (v/v) defibrinated horse blood (DHB) plates (blood agar plates)	37 g BHI broth 10 g bacteriological agar 70 ml DHB (added after media has cooled)	Oxoid, UK Oxoid, UK Sigma-Aldrich, UK	Culturing *C. difficile*
Fastidious Anaerobic broth (FA), 2.97%	29.7 g FA	BioConnections, UK	Culturing *C. difficile*
BHI soft agar (BHI 3.7% and agar 0.4 %)	18.5 g BHI broth 2 g bacteriological agar 250 ml Distilled water	Oxoid, UK Oxoid, UK	For plaque assay and spot test of *C. difficile* phages
Brazier's Cefoxitin, Cycloserine and egg yolk (CCEY)	48 g CCEY 50 ml egg yolk emulsion 10 ml Cycloserine 250mg/l and Cefoxitin 8mg/l	BioConnections, UK Laboratories Ltd, UK Bio Connections, UK	Selection of lysogenic and phage resistant *C. difficile*

Appendix 2

Buffers and solutions

The table contains a list of buffers and solutions that were used throughout this study. All buffers, solutions and enzymes used during this study are autoclaved 121°C for 15 minutes and listed in this table.

Solutions & Buffers	Composition	Manufacturer	Purpose
Salt solution (0.4 M MgCl2, 0.1 M CaCl2)	40.6g MgCl2 0.75g CaCl2 250 ml Distilled water	Argos Organics, UK Fisher Scientific, UK	-For plaque assay and spot test of *C. difficile* phages. - For propagation of *C. difficile* phages.
0.5 M Ethylene-diamine-tetra acetic acid (EDTA)	EDTA 93.5g Dissolved in 450ml distilled H2O and autoclaved.	Sigma-Aldrich, USA	Making up buffer solutions
1 X TAE buffer	Tris-HCl (40mM) Acetic acid (20 mM) EDTA (2M) pH 8	Santa Cruz, USA Fisher Scientific, UK Sigma-Aldrich, USA	Used for agarose gel electrophoresis.
5 x TBE buffer	54g Tris-Base 27.5g boric acid 20ml 0.5M EDTA pH8 Up to 1L with H2O.	Fisher Scientific, UK Fisher Scientific, UK Sigma-Aldrich, USA	For pulsed-field gel-electrophoresis
Sodium dodecyl sulphate	Vary	Fisher Scientific, Belgium	Making up solutions.
% Chelex 100	0.25 g Chelex 5 ml Ultra-pure (UP) water		For Bacterial DNA extraction
Trizol reagent		Invitrogen,UK	For total RNA extraction

Appendix 3

Table showing RNA integrity, yields and DNA contamination for each RNA sample.

No.	Sample Name	Concentration (ng/µL)	Volume (µL)	Total Mass (µg)	OD 260/280	OD 260/230	RIN	23S/ 16S	Test Result
1	0 min 1st.tech.B&phage	430	43	18.49	1.65	0.74	9	1.4	Level A
2	10 min 1st.tech. &Phage	278	33	9.17	1.83	0.67	9.4	1.1	Level A
3	20 min 1st.tech. &Phage	325	41	13.33	1.55	0.61	9	1.2	Level A
4	30 min 1st.tech. &Phage	564	38	21.43	2.21	1.13	9.1	1.1	Level A
5	40 min 1st.tech. &Phage	294	45	13.23	1.92	0.57	8.6	0.8	Level A
6	50 min- 1st.tech.&Phage	378	41	15.5	1.72	0.68	9.1	1.4	Level A
7	0 min 2nd.tech. &Phage	438	41	17.96	2.1	1.5	9.3	1.2	Level A
8	10 min 2nd.tech. &Phage	428	40	17.12	1.71	0.67	9.3	1.7	Level A
9	20 min 2nd.tech. &Phage	220	33	7.26	1.74	0.6	9.3	1.5	Level A
10	30 min 2nd.tech. &Phage	744	39	29.02	1.86	0.91	9.4	1.3	Level A
11	40 min 2nd.tech. &Phage	184	29	5.34	1.86	0.6	9.1	1.2	Level A
12	50 min 2nd.tech. &Phage	290	43	12.47	1.56	0.63	9.1	1.4	Level A
13	0 min Bio B&Phage	110	34	3.74	1.96	1.18	9.2	1.2	Level A
14	10 min Bio B&Phage	105	42	4.41	1.88	1.01	9.8	1.2	Level A
15	20 min Bio B&Phage	134	40	5.36	1.95	1.09	9.9	1.3	Level A
16	30 min Bio B&Phage	218	38	8.28	1.97	1.2	9.5	1.5	Level A
17	40 min Bio B&Phage	64	43	2.75	1.91	0.98	9.7	1.6	Level A
18	50 min Bio B&Phage	224	35	7.84	1.95	1.25	9.4	9.4	Level A
19	0 min Bio B.	93	41	3.81	1.85	0.92	9.5	1	Level A
20	10 min Bio B.	234	41	9.59	2	1.66	9.9	15	Level A
21	20 min Bio B.	182	44	8.01	2.05	1.81	9.4	2.1	Level A
22	30 min Bio B.	154	46	7.08	2.03	1.7	9.4	1.9	Level A
23	40 min Bio B.	146	40	5.84	1.97	1.17	9.3	2.2	Level A
24	50 min Bio B.	127	46	5.84	2.05	1.57	9.3	2.2	Level A

Note

Tech. = Technical replicate. Bio. =Biological replicate. B. = *C.difficile*

Library Type: HiSeq Transcriptome.

1. The 260/280 and 260/230 absorbance ratio <1.8: It may lead to library construction failure, library production too low to sequence or insufficient sequencing data quantity; and it may affect randomicity and cause bias. The closer to the standard the smaller risk, conversely the bigger.

2. The test result based on the (RNA sequencing sample quality standards) explains whether the testing sample meets the requirement of library construction.

a) Level A means the sample is qualified, and the amount of sample satisfies two times library construction or more.

b) Level B means the sample is qualified, but the amount of sample only satisfies one time library construction.

c) Level C means the sample does not totally meet the requirements of library construction and sequencing. BGI can try to construct library but the sequencing quality is not fully guaranteed (23S/16S＜1.0.).

d) Level D means the sample does not meet the requirements of library construction and sequencing. BGI does not suggest using the sample.

Sample Test Methods

Method of concentration determination: Qubit Fluorometer, Agilent 2100, Nano Drop, Microplate Reader.

Method of OD260/280 & OD260/230 test: Nano Drop.

Method of 28S/18S & 23S/16S test: Agilent 2100.

Method of RIN test: Agilent 2100.

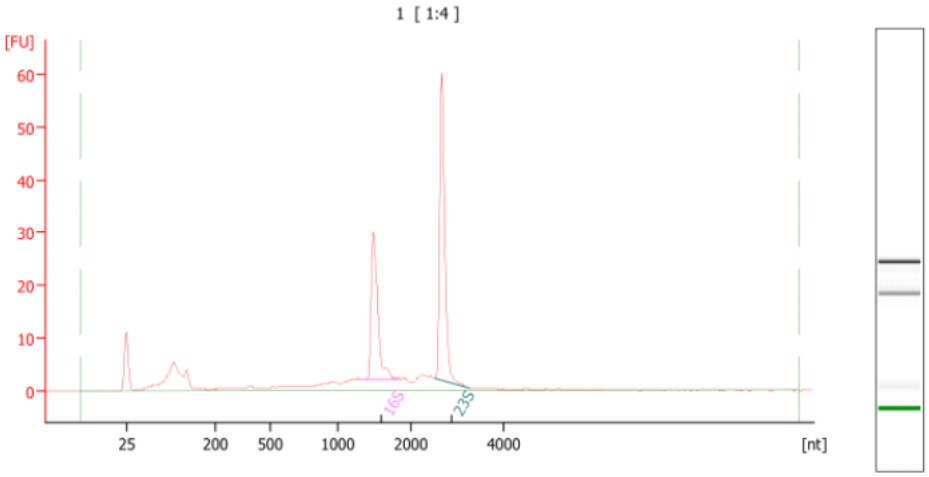
Appendix 4

Test results of Agilent 2100

The Agilent 2100 Bioanalyzer system provides sizing, quantitation, and purity assessments for DNA, RNA, and protein samples.

For RNA samples in this subject, Pre-treatment has been done. After the RNA sample melted, it was centrifuged and mixed and then processed as 3.2.2 section.

(1) Sample name: 1



Overall Results for sample 1 : 1

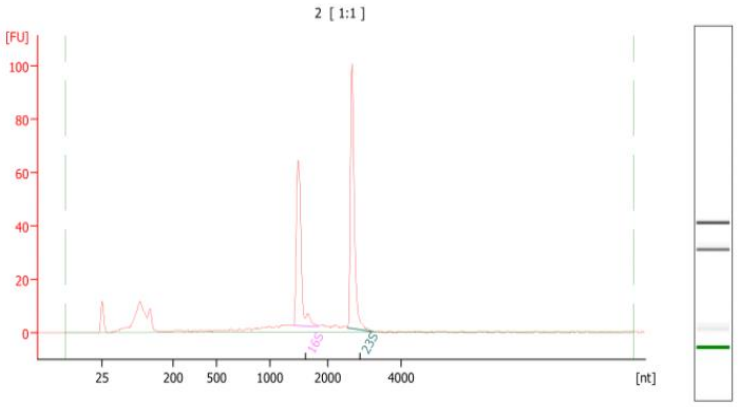
RNA Area:	190.4	rRNA Ratio [23s / 16s]:	1.4
RNA Concentration:	86 ng/µl	RNA Integrity Number (RIN):	9 (B.02.08)

Fragment table for sample 1 : 1

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,335	1,856	38.8	20.4
23S	2,491	3,273	55.4	29.1

Quality assessment of RNA isolation from samples have automatically been detected, and peaks characteristics such as RNA concentration, rRNA Ratio (23s/ 16s), RNA integrity number (RIN) and start size/end size for both 16s/23s were all calculated.

(2) Sample name: 2



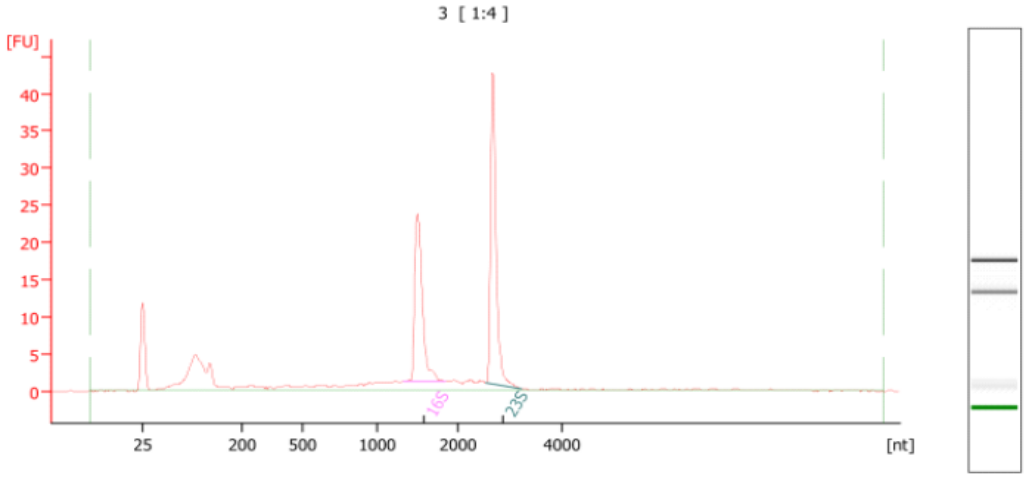
Overall Results for sample 2 : 2

RNA Area:	307.7	rRNA Ratio [23s / 16s]:	1.1
RNA Concentration:	139 ng/µl	RNA Integrity Number (RIN):	9.4 (B.02.08)

Fragment table for sample 2 : 2

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,406	1,828	78.7	25.6
23S	2,523	3,241	90.4	29.4

(3) Sample name: 3



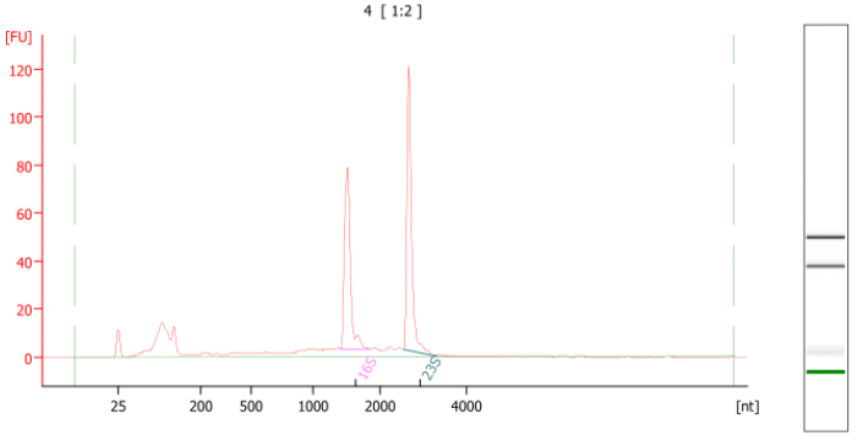
Overall Results for sample 3 : 3

RNA Area:	144.1	rRNA Ratio [23s / 16s]:	1.2
RNA Concentration:	65 ng/µl	RNA Integrity Number (RIN):	9 (B.02.08)

Fragment table for sample 3 : 3

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,330	1,840	32.0	22.2
23S	2,524	3,225	37.9	26.3

(4) Sample name: 4



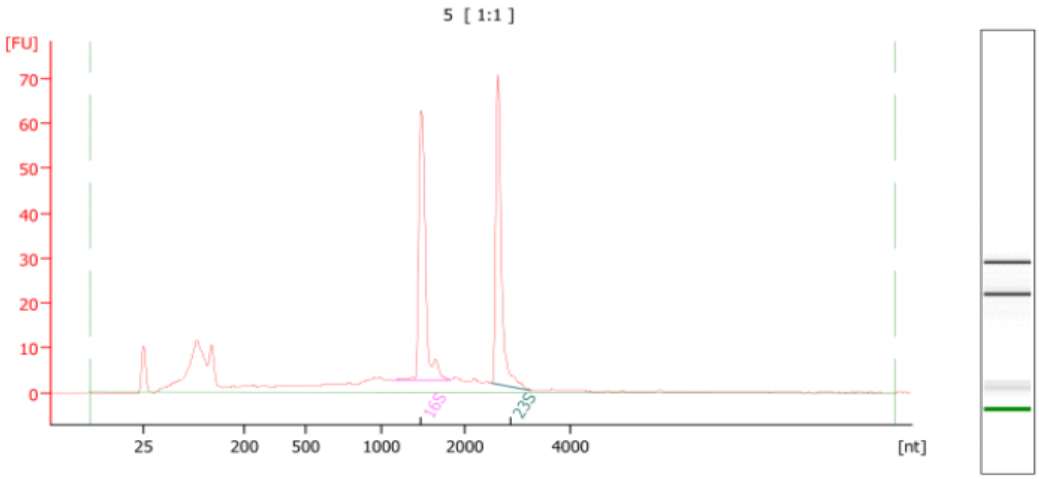
Overall Results for sample 4 : 4

RNA Area:	416.6	rRNA Ratio [23s / 16s]:	1.1
RNA Concentration:	188 ng/µl	RNA Integrity Number (RIN):	9.1 (B.02.08)

Fragment table for sample 4 : 4

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,409	1,859	100.1	24.0
23S	2,539	3,303	110.2	26.5

(5) Sample name: 5



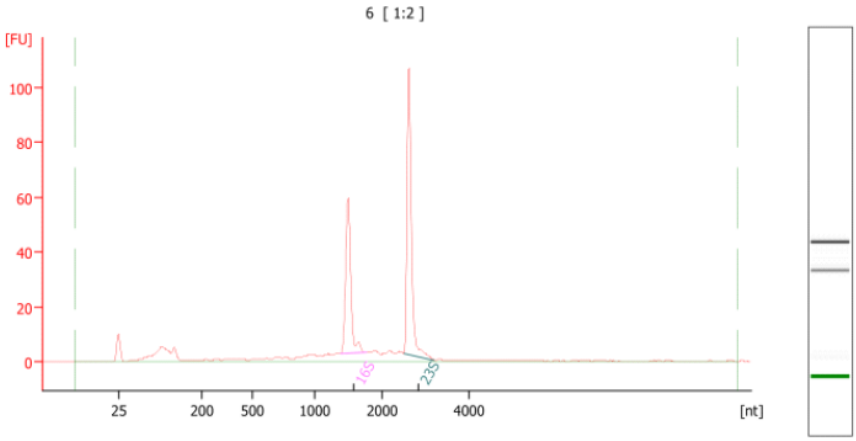
Overall Results for sample 5 : 5

RNA Area:	324.0	rRNA Ratio [23s / 16s]:	0.8
RNA Concentration:	147 ng/µl	RNA Integrity Number (RIN):	8.6 (B.02.08)

Fragment table for sample 5 : 5

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,128	1,822	75.2	23.2
23S	2,509	3,255	63.4	19.6

(6) Sample name: 6



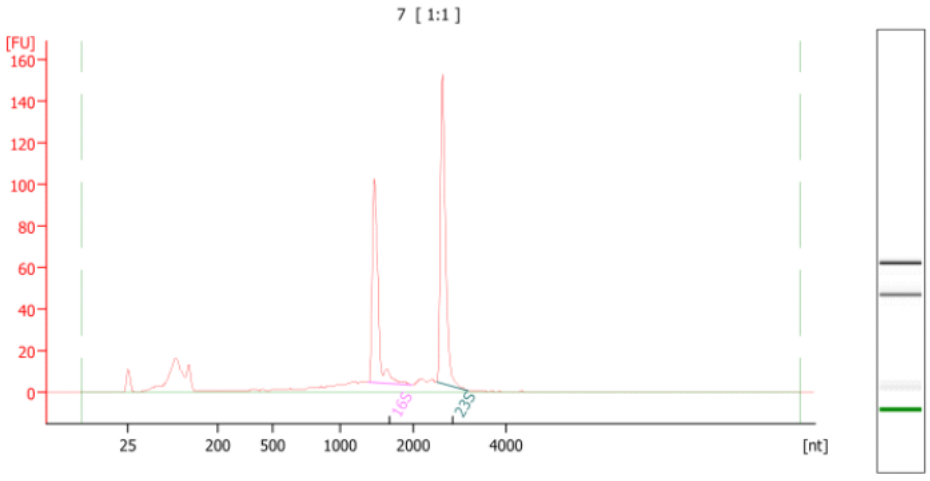
Overall Results for sample 6 : 6

RNA Area:	278.3	rRNA Ratio [23s / 16s]:	1.4
RNA Concentration:	126 ng/µl	RNA Integrity Number (RIN):	9.1 (B.02.08)

Fragment table for sample 6 : 6

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,389	1,771	64.9	23.3
23S	2,493	3,209	93.1	33.4

(7) Sample name: 7



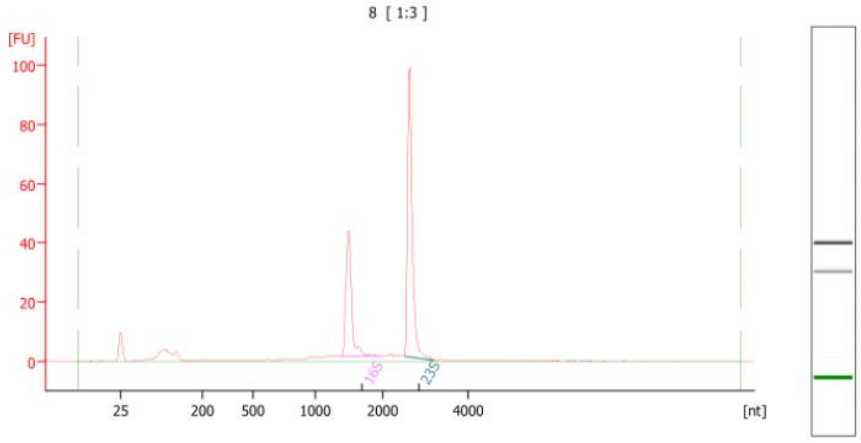
Overall Results for sample 7 : 7

RNA Area:	484.3	rRNA Ratio [23s / 16s]:	1.2
RNA Concentration:	219 ng/µl	RNA Integrity Number (RIN):	9.3 (B.02.08)

Fragment table for sample 7 : 7

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,386	1,975	118.7	24.5
23S	2,492	3,179	139.7	28.8

(8) Sample name: 8



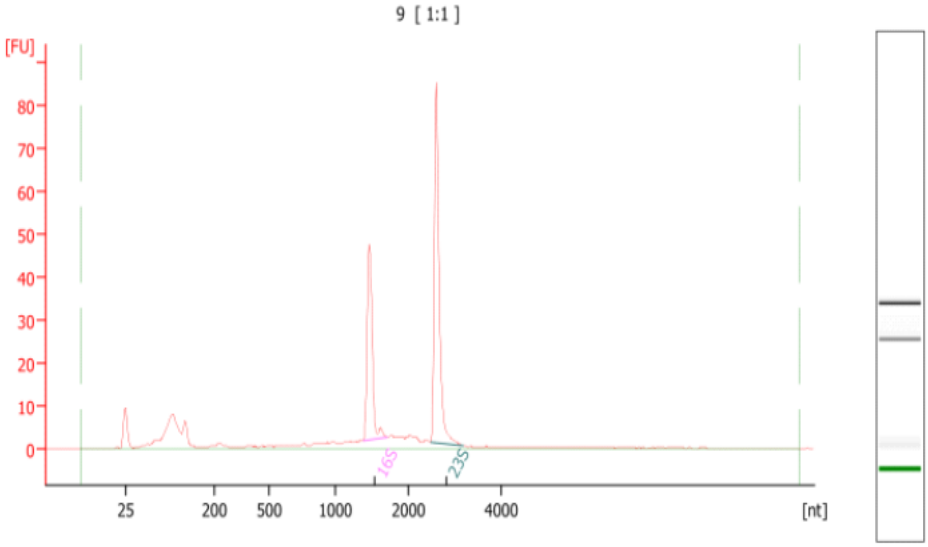
Overall Results for sample 8 : 8

RNA Area:	235.6	rRNA Ratio [23s / 16s]:	1.7
RNA Concentration:	107 ng/µl	RNA Integrity Number (RIN):	9.3 (B.02.08)

Fragment table for sample 8 : 8

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,374	1,994	54.9	23.3
23S	2,491	3,210	93.5	39.7

(9) Sample name: 9



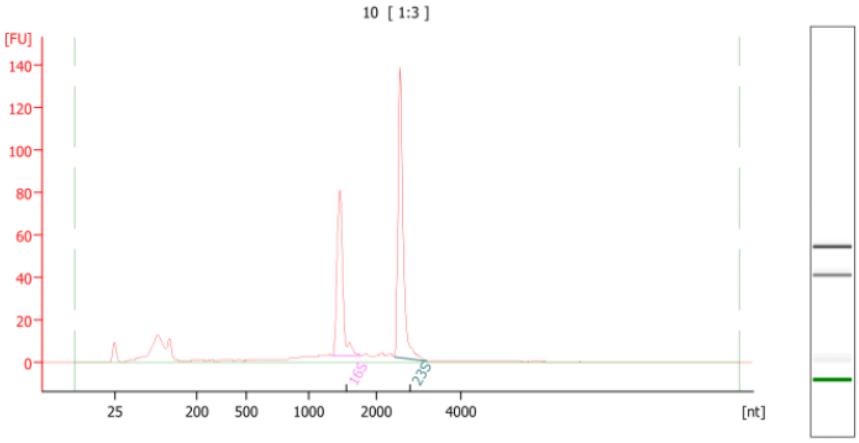
Overall Results for sample 9 : 9

RNA Area:	243.0	rRNA Ratio [23s / 16s]:	1.5
RNA Concentration:	110 ng/µl	RNA Integrity Number (RIN):	9.3 (B.02.08)

Fragment table for sample 9 : 9

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,378	1,704	51.0	21.0
23S	2,457	3,164	76.4	31.4

(10) Sample name: 10



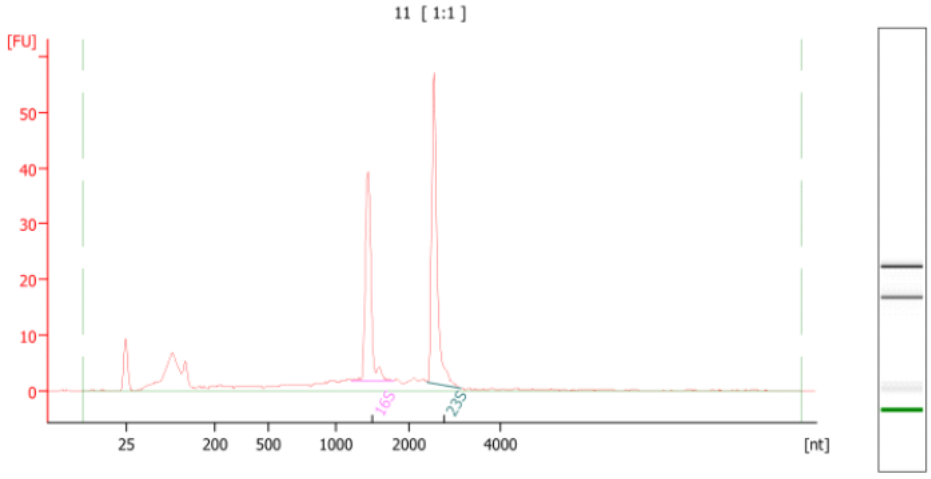
Overall Results for sample 10 : 10

RNA Area:	411.0	rRNA Ratio [23s / 16s]:	1.3
RNA Concentration:	186 ng/µl	RNA Integrity Number (RIN):	9.4 (B.02.08)

Fragment table for sample 10 : 10

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,363	1,790	100.9	24.5
23S	2,439	3,213	135.0	32.9

(11) Sample name: 11



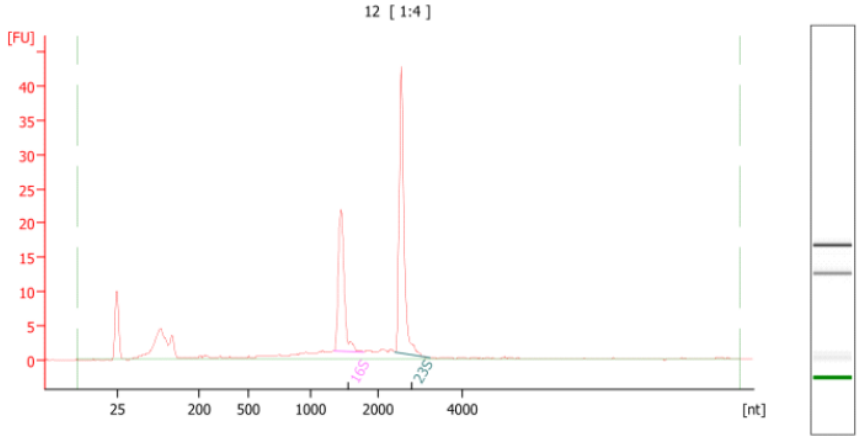
Overall Results for sample 11 : 11

RNA Area:	202.7	rRNA Ratio [23s / 16s]:	1.2
RNA Concentration:	92 ng/µl	RNA Integrity Number (RIN):	9.1 (B.02.08)

Fragment table for sample 11 : 11

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,224	1,800	44.2	21.8
23S	2,424	3,150	53.1	26.2

(12) Sample name: 12



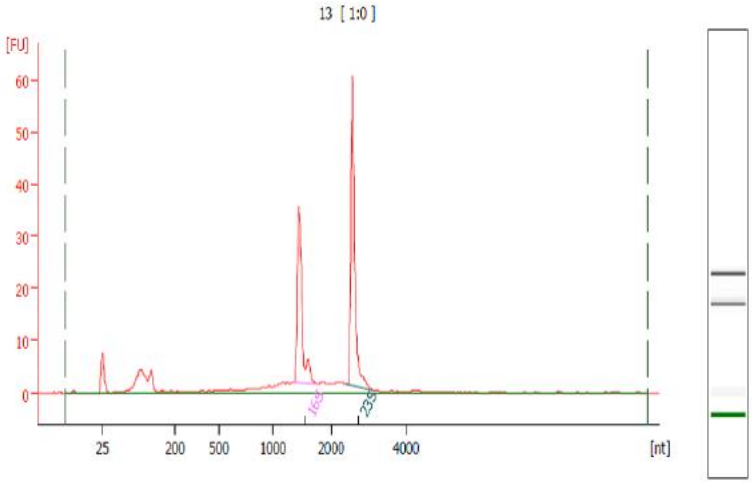
Overall Results for sample 12 : 12

RNA Area:	127.6	rRNA Ratio [23s / 16s]:	1.4
RNA Concentration:	58 ng/µl	RNA Integrity Number (RIN):	9.1 (B.02.08)

Fragment table for sample 12 : 12

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,353	1,800	27.6	21.6
23S	2,424	3,244	39.3	30.7

(13) Sample name: 13



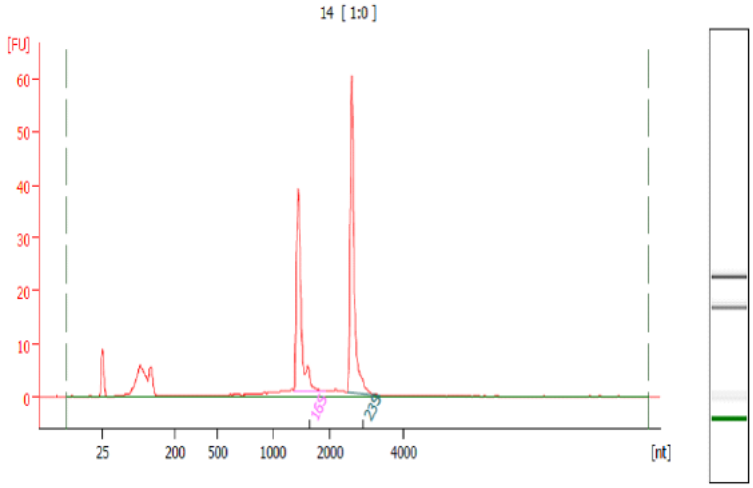
Overall Results for sample 7 : 13

RNA Area:	178.6	rRNA Ratio [23s / 16s]:	1.2
RNA Concentration:	110 ng/µl	RNA Integrity Number (RIN):	9.2 (B.02.07)

Fragment table for sample 7 : 13

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,362	1,780	42.1	23.6
23S	2,402	3,141	52.6	29.4

(14) Sample name: 14



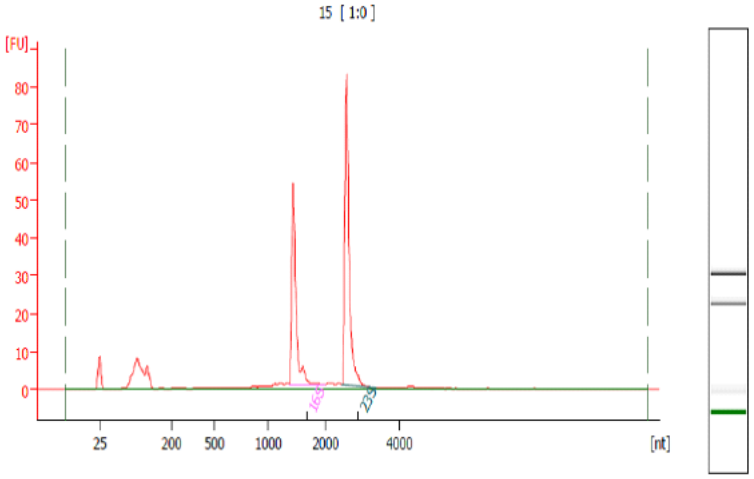
Overall Results for sample 8 : 14

RNA Area:	170.5	rRNA Ratio [23s / 16s]:	1.2
RNA Concentration:	105 ng/µl	RNA Integrity Number (RIN):	9.8 (B.02.07)

Fragment table for sample 8 : 14

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,376	1,920	45.6	26.7
23S	2,478	3,331	55.2	32.4

(15) Sample name: 15



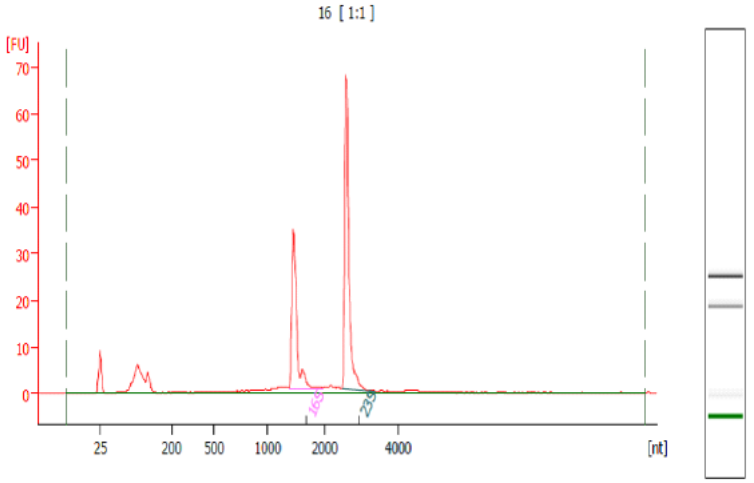
Overall Results for sample 11 : 15

RNA Area:	217.8	rRNA Ratio [23s / 16s]:	1.3
RNA Concentration:	134 ng/µl	RNA Integrity Number (RIN):	9.9 (B.02.07)

Fragment table for sample 11 : 15

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,368	2,013	58.4	26.8
23S	2,459	3,381	78.6	36.1

(16) Sample name: 16



Overall Results for sample 12 : 16

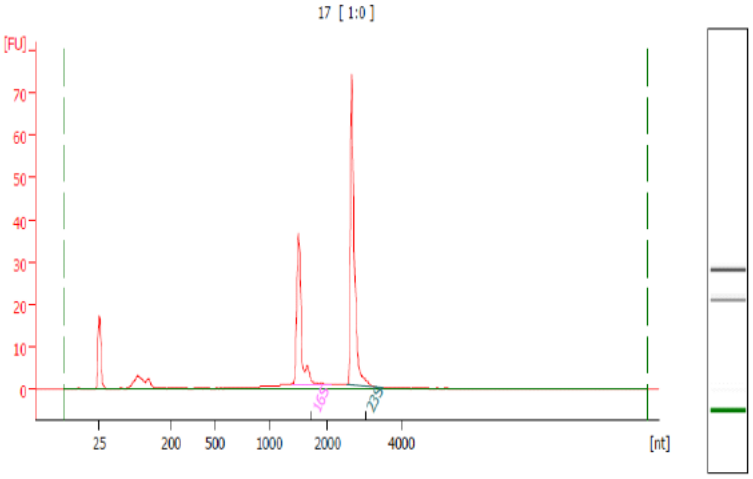
RNA Area:	177.4	rRNA Ratio [23s / 16s]:	1.5
RNA Concentration:	109 ng/µl	RNA Integrity Number (RIN):	9.5 (B.02.07)

Fragment table for sample 12 : 16

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,378	1,957	41.6	23.4
23S	2,475	3,365	63.1	35.5

(17) Sample name: 17

(17) Sample name: 17



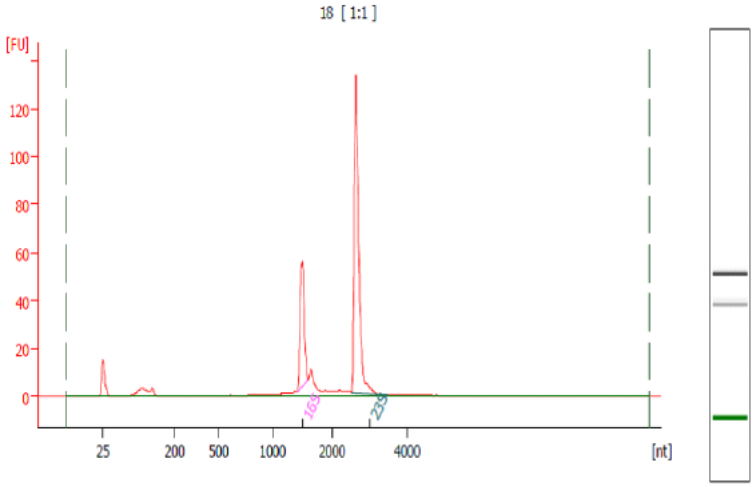
Overall Results for sample 1 : 17

RNA Area:	159.7	rRNA Ratio [23s / 16s]:	1.6
RNA Concentration:	64 ng/µl	RNA Integrity Number (RIN):	9.7 (B.02.07)

Fragment table for sample 1 : 17

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,409	2,099	43.6	27.3
23S	2,541	3,489	68.0	42.6

(18) Sample name: 18



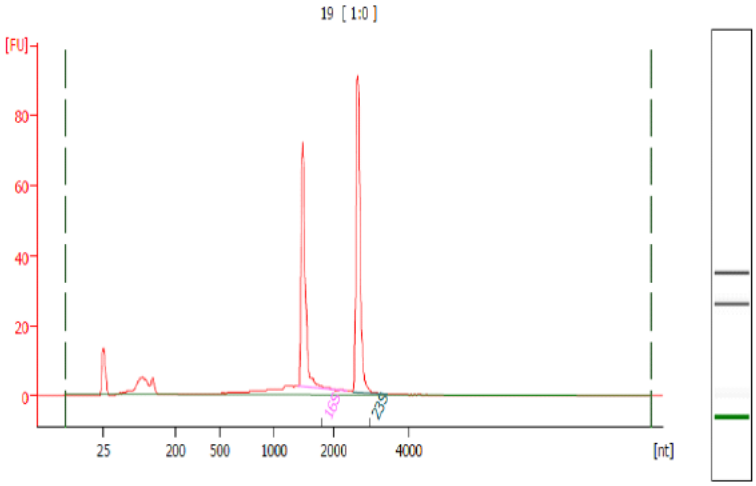
Overall Results for sample 2 : 18

RNA Area:	281.2	rRNA Ratio [23s / 16s]:	2.4
RNA Concentration:	112 ng/µl	RNA Integrity Number (RIN):	9.4 (B.02.07)

Fragment table for sample 2 : 18

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,414	1,595	55.9	19.9
23S	2,528	3,471	132.2	47.0

(19) Sample name: 19



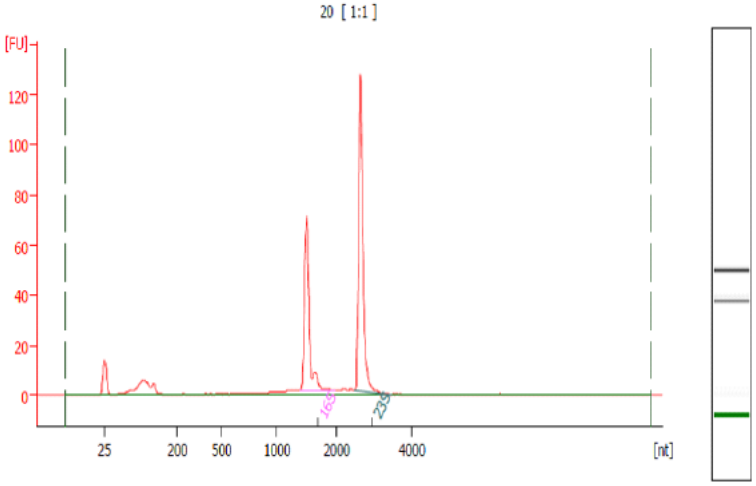
Overall Results for sample 3 : 19

RNA Area:	232.5	rRNA Ratio [23s / 16s]:	1.0
RNA Concentration:	93 ng/µl	RNA Integrity Number (RIN):	9.5 (B.02.07)

Fragment table for sample 3 : 19

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,407	2,372	76.7	33.0
23S	2,529	3,439	77.8	33.5

(20) Sample name: 20



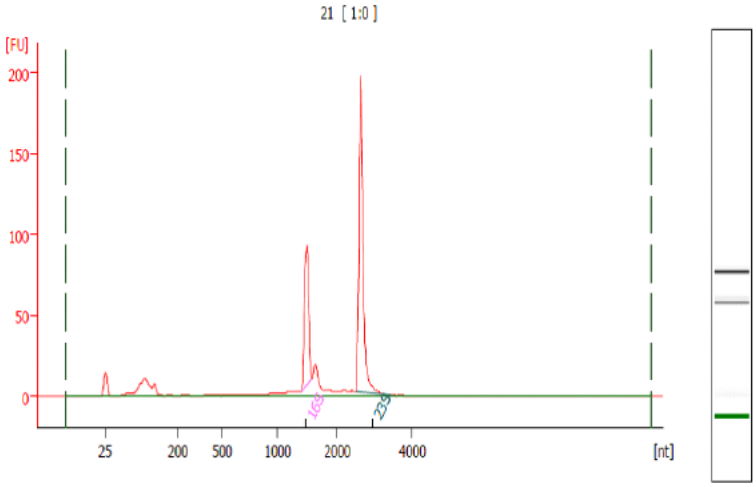
Overall Results for sample 4 : 20

RNA Area:	292.0	rRNA Ratio [23s / 16s]:	1.5
RNA Concentration:	117 ng/µl	RNA Integrity Number (RIN):	9.9 (B.02.07)

Fragment table for sample 4 : 20

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,385	1,991	79.8	27.3
23S	2,485	3,436	117.7	40.3

(21) Sample name: 21



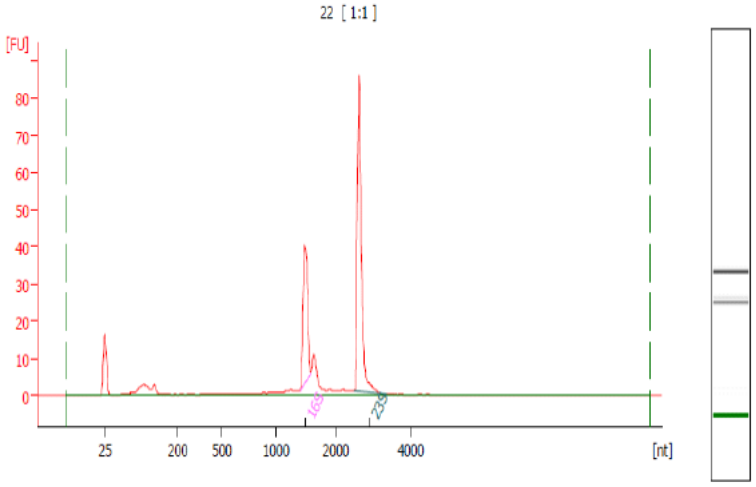
Overall Results for sample 5 : 21

RNA Area:	456.0	rRNA Ratio [23s / 16s]:	2.1
RNA Concentration:	182 ng/µl	RNA Integrity Number (RIN):	9.4 (B.02.07)

Fragment table for sample 5 : 21

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,405	1,584	85.4	18.7
23S	2,516	3,405	180.1	39.5

(22) Sample name: 22



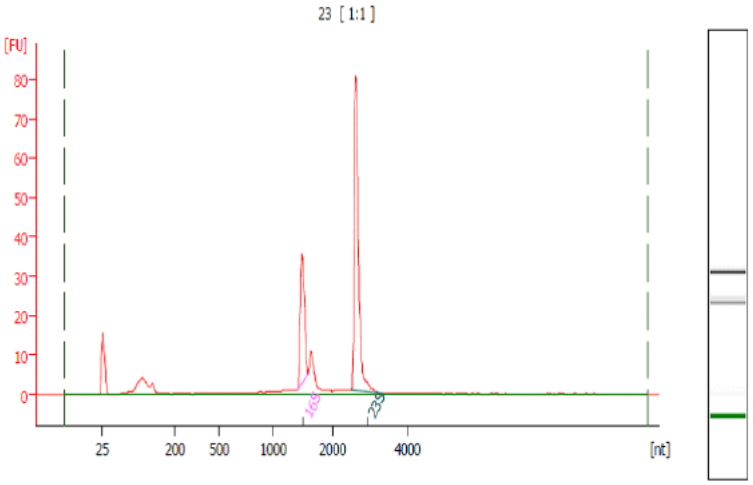
Overall Results for sample 6 : 22

RNA Area:	192.1	rRNA Ratio [23s / 16s]:	1.9
RNA Concentration:	77 ng/µl	RNA Integrity Number (RIN):	9.4 (B.02.07)

Fragment table for sample 6 : 22

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,403	1,592	39.8	20.7
23S	2,515	3,343	76.6	39.9

(23) Sample name: 23



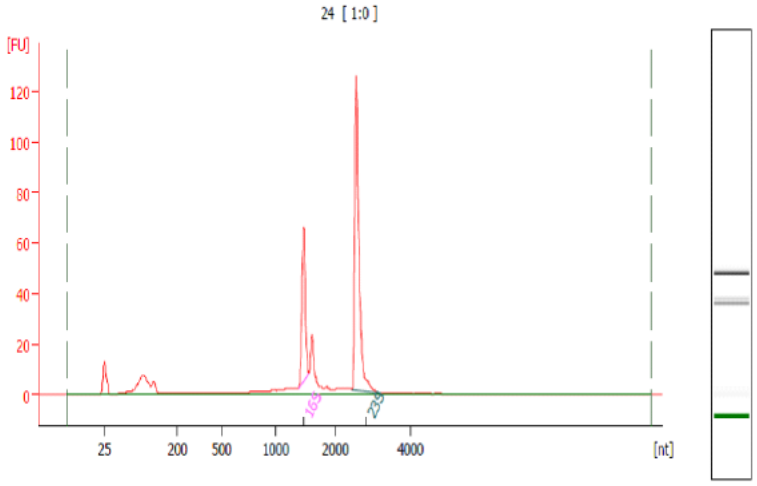
Overall Results for sample 7 : 23

RNA Area:	181.8	rRNA Ratio [23s / 16s]:	2.2
RNA Concentration:	73 ng/µl	RNA Integrity Number (RIN):	9.3 (B.02.07)

Fragment table for sample 7 : 23

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,410	1,589	33.0	18.2
23S	2,498	3,328	73.0	40.2

(24) Sample name: 24



Overall Results for sample 8 : 24

RNA Area:	318.1	rRNA Ratio [23s / 16s]:	2.2
RNA Concentration:	127 ng/µl	RNA Integrity Number (RIN):	9.3 (B.02.07)

Fragment table for sample 8 : 24

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
16S	1,380	1,559	52.6	16.5
23S	2,451	3,203	115.0	36.2

Appendix 5

Data Production

1 Data Production

After sequencing ,the raw reads were filtered.Data filtering includes removing adapter sequences, contamination and low-quality reads from raw reads.Next, we get the statistics of data production.Table 1-1 shows statistical results after data treatment.

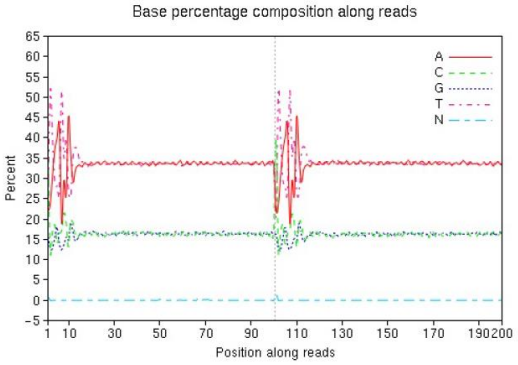
Table 1-1 Reads statistics results

No.	Sample Name	Read length(bp)	Clean Reads	Clean bases	Q20(%)	GC(%)
1	1	100	11051456	1105145600	98.83;98.39	32.75
2	10	100	10935652	1093565200	98.87;97.77	33.23
3	11	100	11022866	1102286600	98.87;97.84	33.48
4	12	100	11105056	1110505600	99.15;98.48	33.62
5	2	100	11079800	1107980000	99.16;98.32	34.88
6	3	100	10950598	1095059800	99.05;98.13	33.65
7	4	100	10979936	1097993600	99.01;98.25	33.42
8	5	100	11052270	1105227000	99.05;98.31	33.37
9	6	100	11069556	1106955600	99.15;98.40	33.57
10	7	100	11059688	1105968800	99.06;98.32	33.04
11	8	100	10889276	1088927600	98.86;97.65	33.12
12	9	100	11020102	1102010200	99.04;98.29	33.21
1	13	100	10098222	1009822200	98.84;97.16	33.22
2	14	100	10952742	1095274200	98.95;97.64	32.94
3	15	100	10953326	1095332600	99.00;97.34	32.95
4	16	100	10950280	1095028000	98.96;97.61	32.77
5	17	100	10933592	1093359200	99.02;97.65	32.80
6	18	100	10929202	1092920200	98.98;97.68	32.71
7	19	100	10900270	1090027000	98.85;97.52	32.19
8	20	100	10893438	1089343800	98.94;97.59	33.50
9	21	100	10041548	1004154800	98.91;97.48	33.02
10	22	100	10933810	1093381000	98.94;97.74	32.76
11	23	100	10858750	1085875000	98.95;97.40	32.92
12	24	100	10993076	1099307600	99.00;97.61	32.92

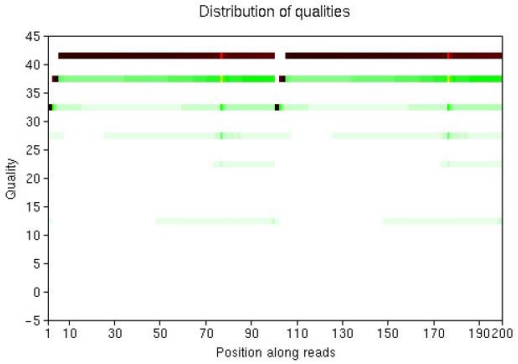
Data Quality Control

2 Data Quality Control

The distribution of base percentage and quality along reads in data filtering are shown in Figure below

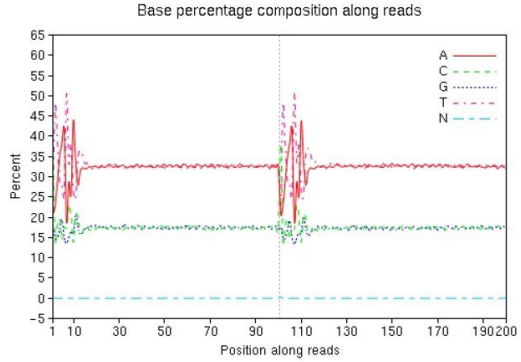


a)Base percentage distribution along reads after filtering

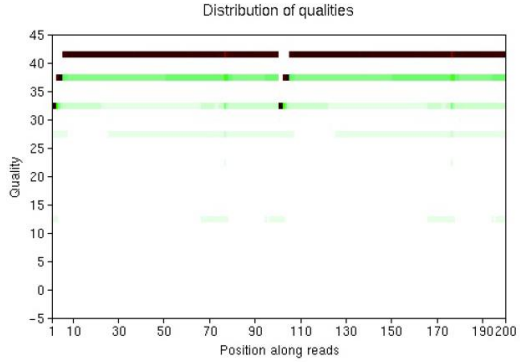


b)Distribution of qualities along reads after filtering

Quality control of sample 1

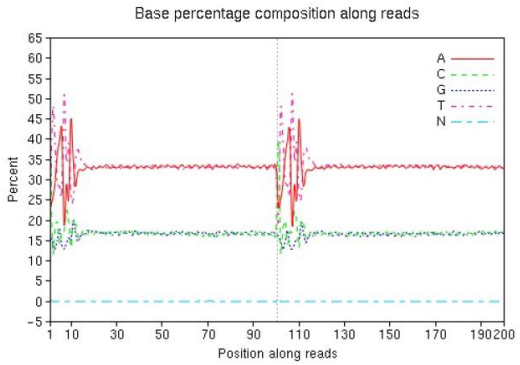


a)Base percentage distribution along reads after filtering

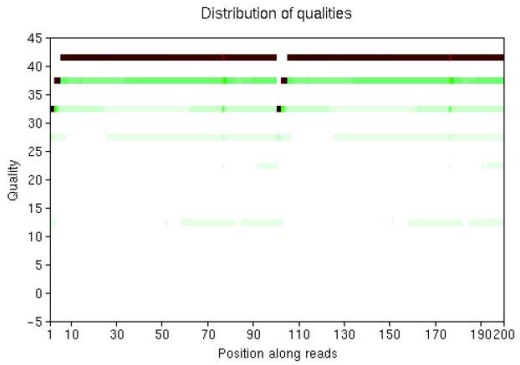


b)Distribution of qualities along reads after filtering

Quality control of sample 2

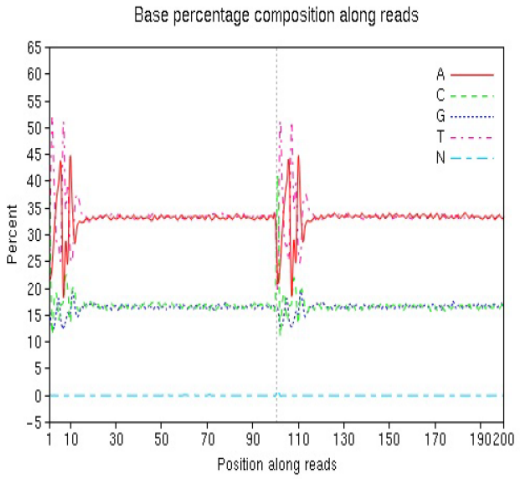


a)Base percentage distribution along reads after filtering

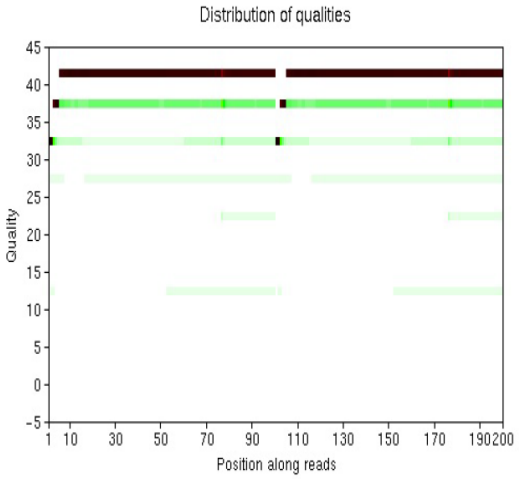


b)Distribution of qualities along reads after filtering

Quality control of sample 3

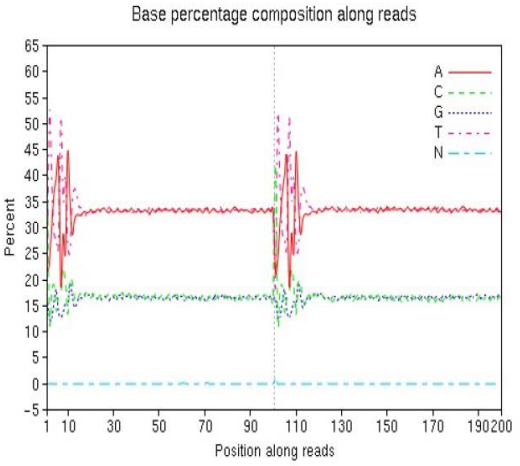


a)Base percentage distribution along reads after filtering

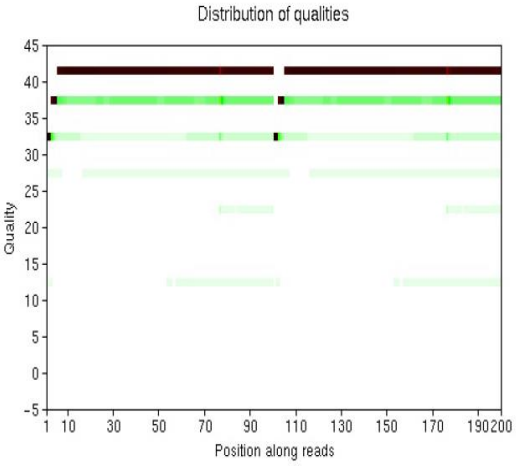


b)Distribution of qualities along reads after filtering

Quality control of sample 4

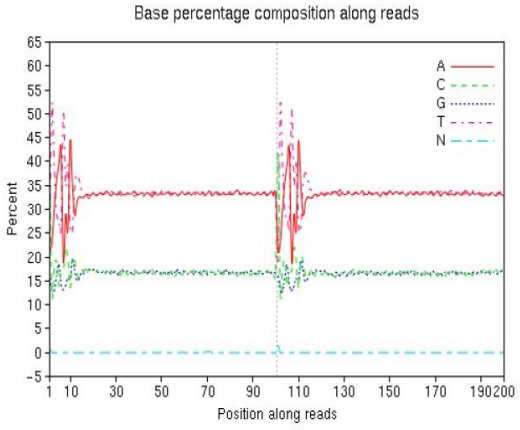


a)Base percentage distribution along reads after filtering

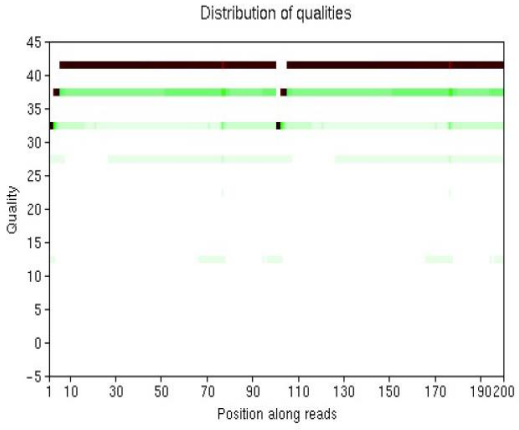


b)Distribution of qualities along reads after filtering

Quality control of sample 5

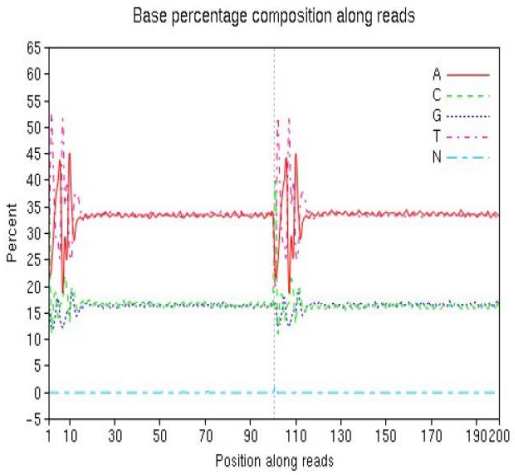


a)Base percentage distribution along reads after filtering

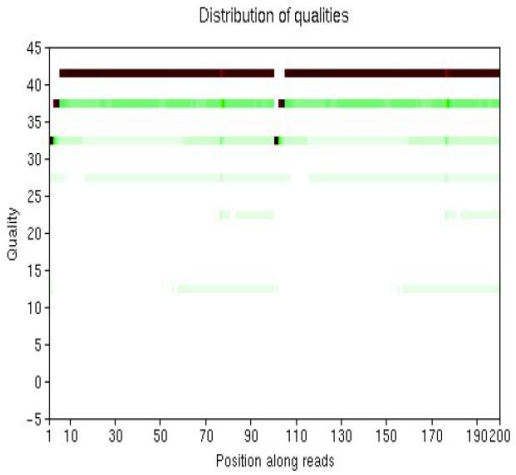


b)Distribution of qualities along reads after filtering

Quality control of sample 6

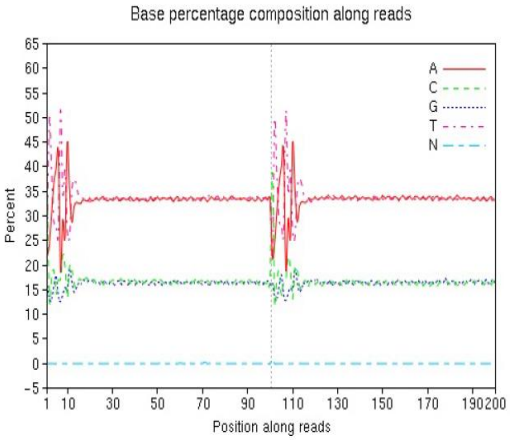


a)Base percentage distribution along reads after filtering

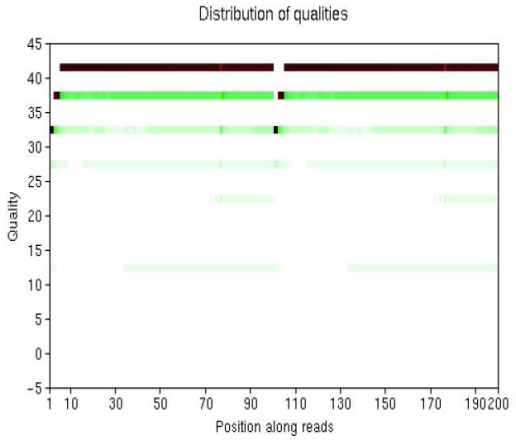


b)Distribution of qualities along reads after filtering

Quality control of sample 7

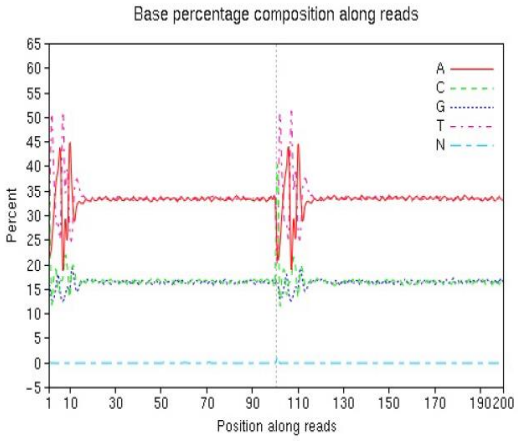


a)Base percentage distribution along reads after filtering

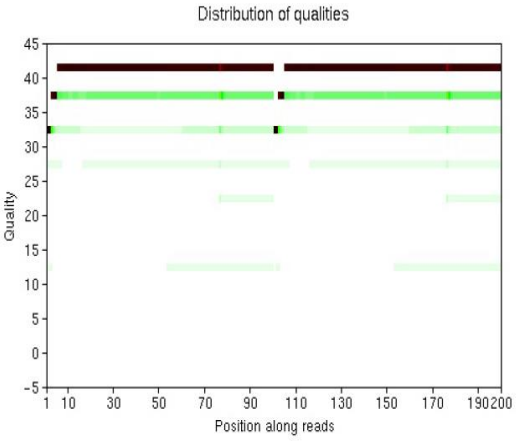


b)Distribution of qualities along reads after filtering

Quality control of sample 8

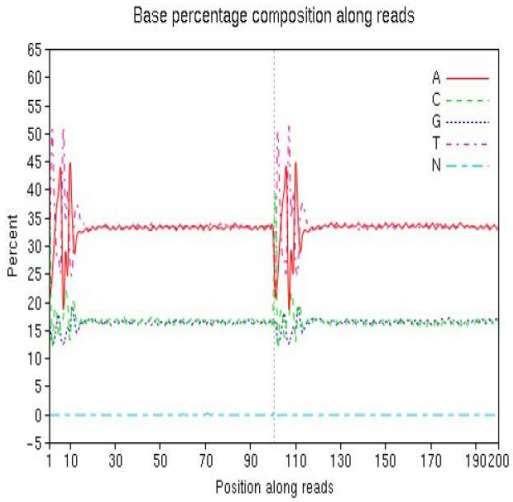


a)Base percentage distribution along reads after filtering

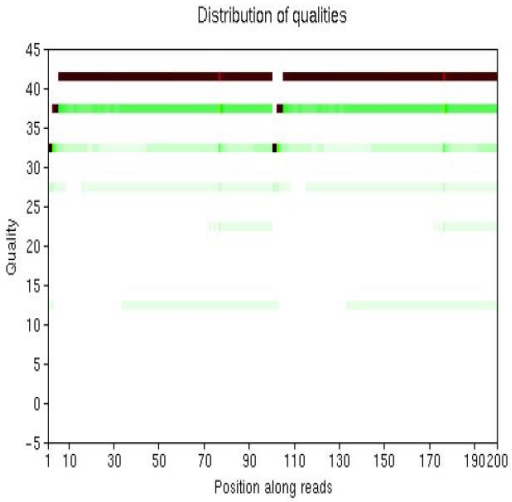


b)Distribution of qualities along reads after filtering

Quality control of sample 9

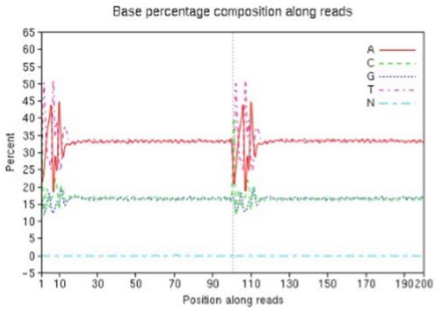


a)Base percentage distribution along reads after filtering

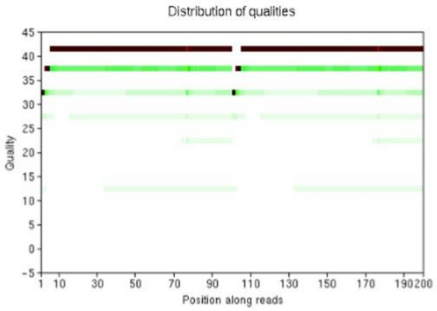


b)Distribution of qualities along reads after filtering

Quality control of sample 10

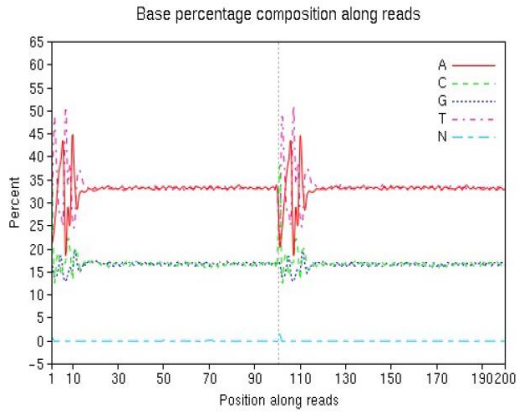


a)Base percentage distribution along reads after filtering

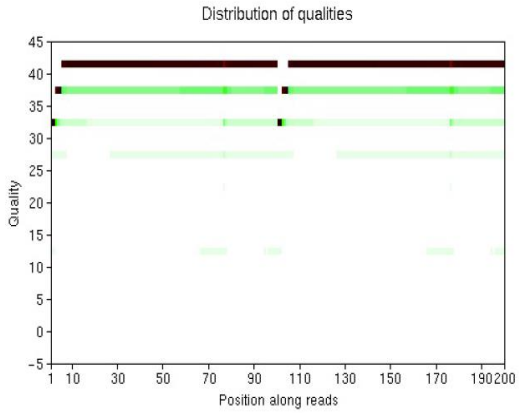


b)Distribution of qualities along reads after filtering

Quality control of sample 11

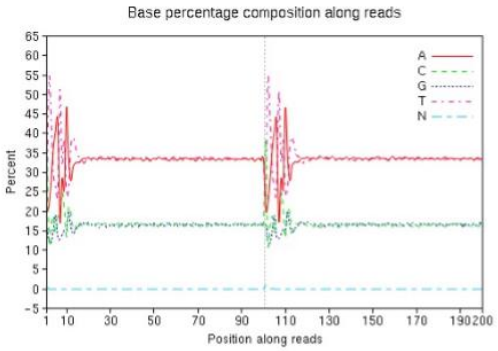


a)Base percentage distribution along reads after filtering

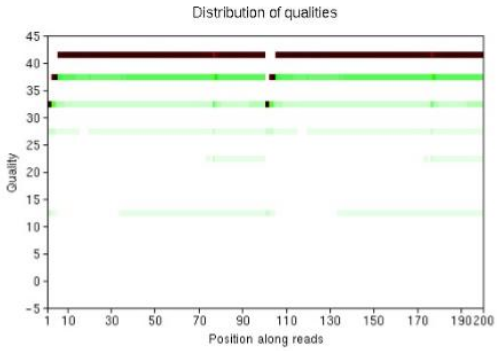


b)Distribution of qualities along reads after filtering

Quality control of sample 12

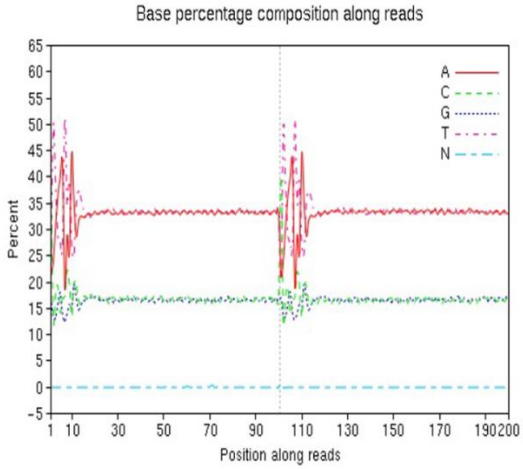


a)Base percentage distribution along reads after filtering

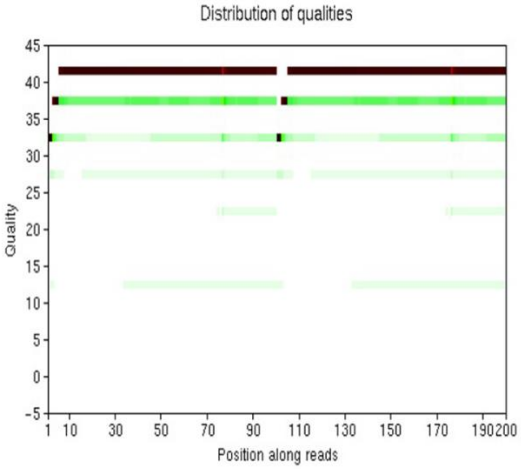


b)Distribution of qualities along reads after filtering

Quality control of sample 13

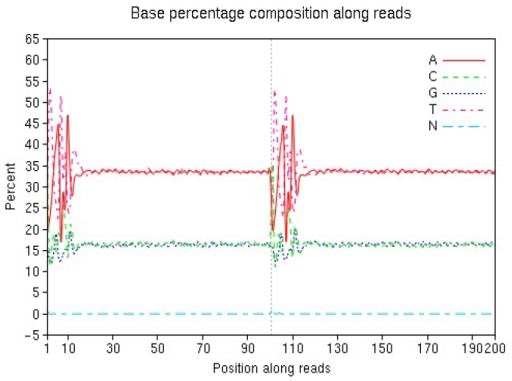


a)Base percentage distribution along reads after filtering

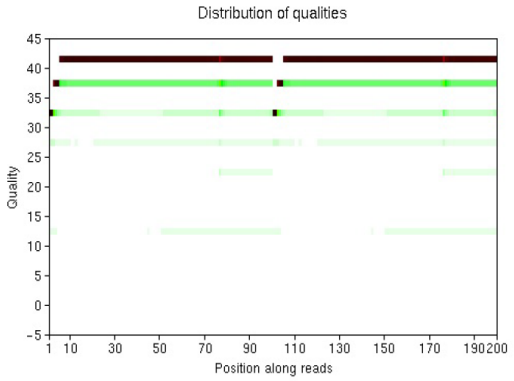


b)Distribution of qualities along reads after filtering

quality control of sample 14

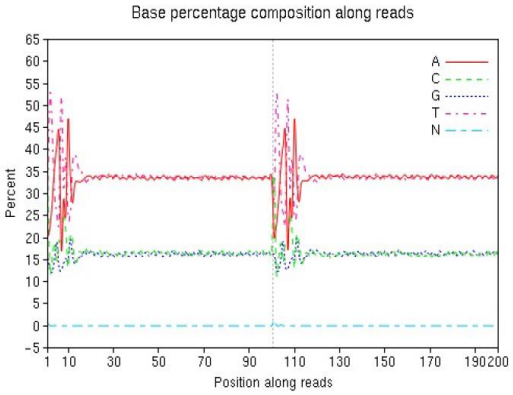


a)Base percentage distribution along reads after filtering

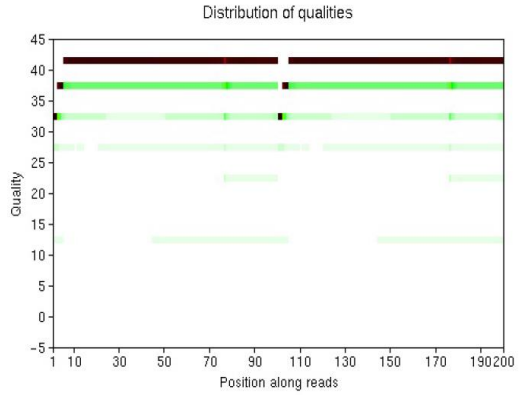


b)Distribution of qualities along reads after filtering

Quality control of sample 15

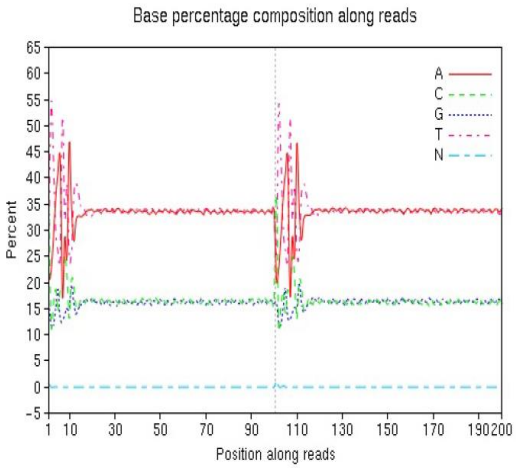


a)Base percentage distribution along reads after filtering

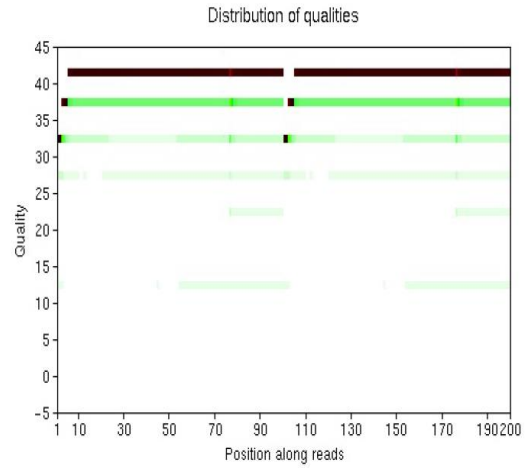


b)Distribution of qualities along reads after filtering

Quality control of sample 16

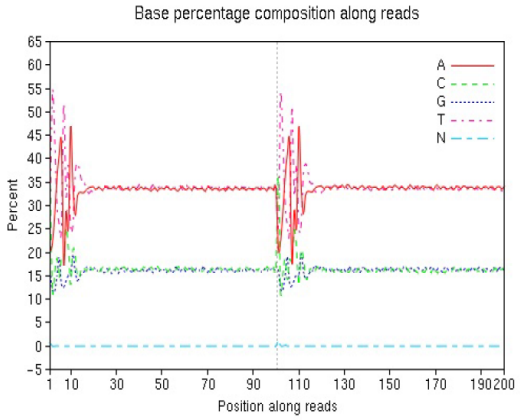


a)Base percentage distribution along reads after filtering

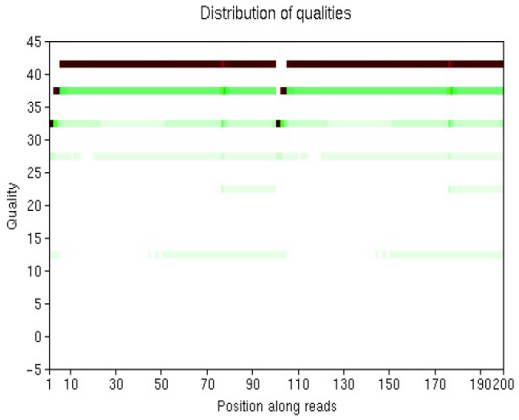


b)Distribution of qualities along reads after filtering

Quality control of sample 17

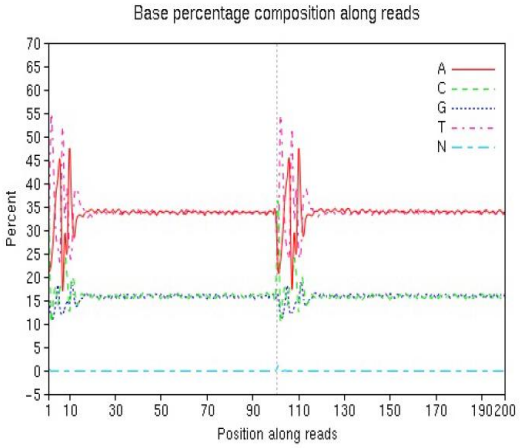


a)Base percentage distribution along reads after filtering

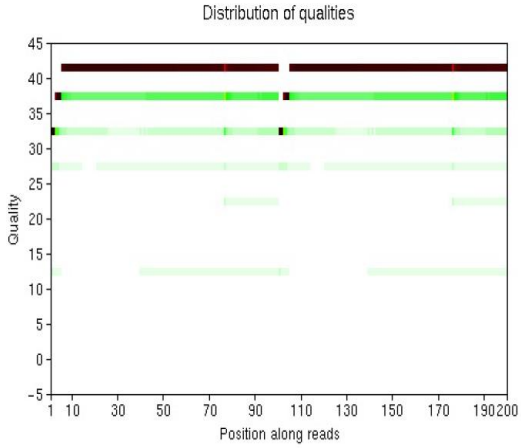


b)Distribution of qualities along reads after filtering

Quality control of sample 18

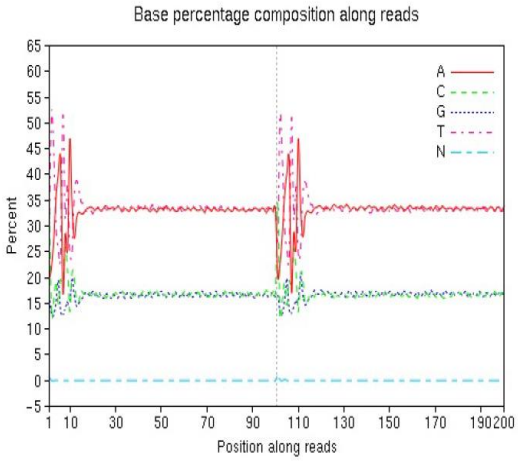


a)Base percentage distribution along reads after filtering

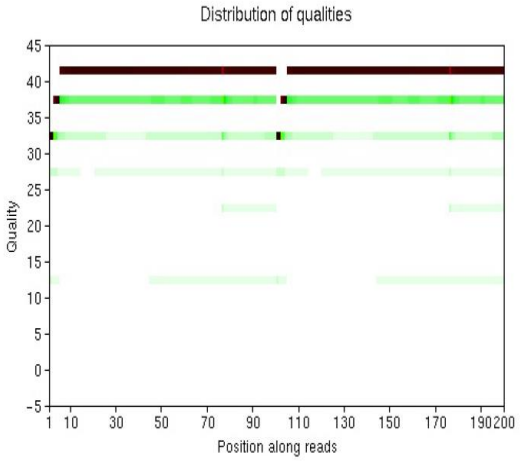


b)Distribution of qualities along reads after filtering

Quality control of sample 19

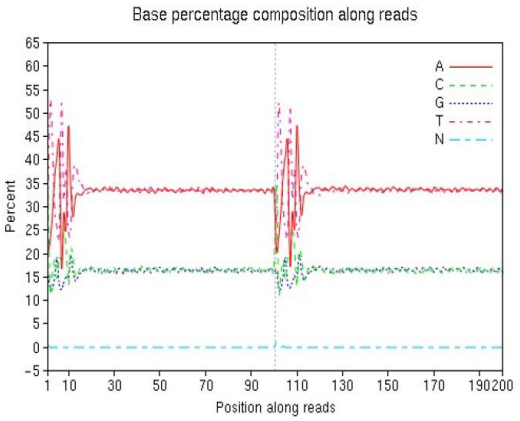


a)Base percentage distribution along reads after filtering

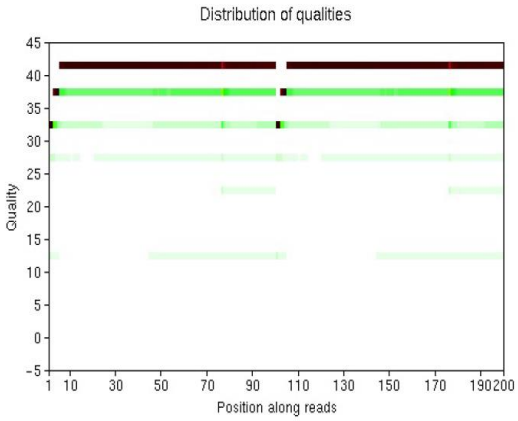


b)Distribution of qualities along reads after filtering

Quality control of sample 20

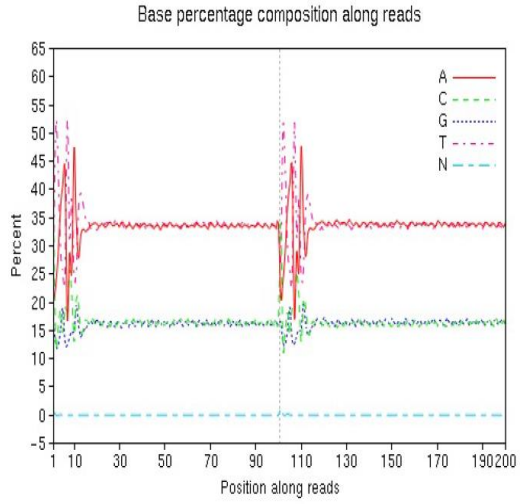


a)Base percentage distribution along reads after filtering

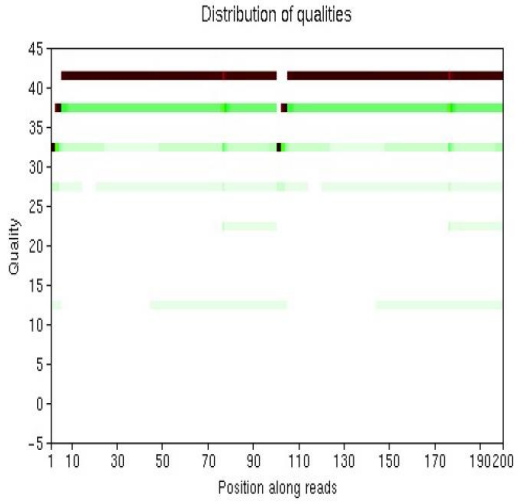


b)Distribution of qualities along reads after filtering

Quality control of sample 21

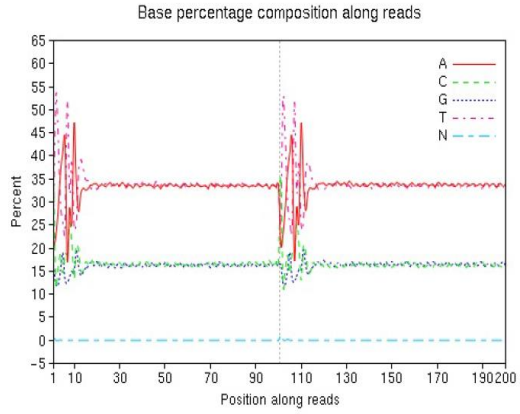


a)Base percentage distribution along reads after filtering

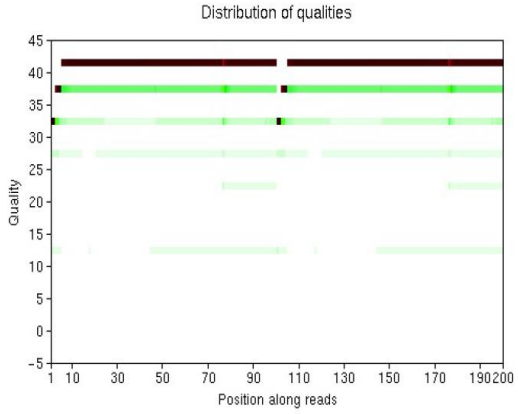


b)Distribution of qualities along reads after filtering

Quality control of sample 22

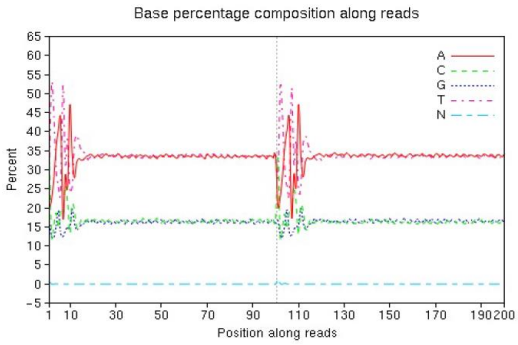


a)Base percentage distribution along reads after filtering

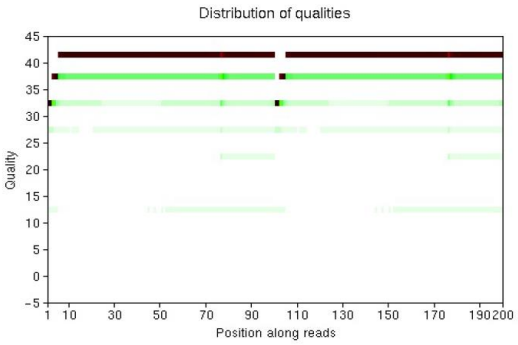


b)Distribution of qualities along reads after filtering

Quality control of sample 23



a)Base percentage distribution along reads after filtering



b)Distribution of qualities along reads after filtering

Quality control of sample 24

Help Document

1 Pipeline of Experiments

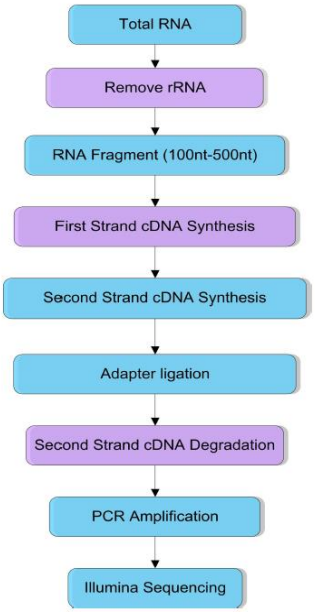


Figure 1-1 Experiment pipeline of Prokaryotic Strand-Specific Transcriptome

rRNA was removed with kit after total RNA was collected from prokaryote. Fragmentation buffer was added for interrupting mRNA to short fragments. Taking these short fragments as templates, Random hexamer-primer were used to synthesize the first-strand cDNA. The second-strand cDNA was synthesized using buffer, dATPs, dGTPs, dCTPs, dUTPs, RNase H and DNA polymerase I respectively after removing dNTPs. Short fragments were purified with QiaQuick PCR extraction kit and resolved with EB buffer for end reparation and adding poly (A). After that, the short fragments were connected with sequencing adapters. Then, the UNG enzyme was used to degrade the second-strand cDNA, and the product was purified by MiniElute PCR Purification Kit before PCR amplification. At last, the library could be sequenced using HiSeq^{TM}4000 or other sequencer when necessary.

2 Raw sequence data

The sequencing images were translated into text-based sequencing data via base calling software as FASTQ documents. FASTQ documents are defined as original raw data provided for users, including the detailed read sequences and sequencing quality information. The format of FASTQ file is listed as below:

FASTQ file sequenced by Illumina HiSeq^{TM}4000 or other sequencer using Phred +33 quality system.
@SIM:1:FCX:1:15:6329:1045 1:N:0:ATCCGA
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>##=><9=AAAAAAAAA9#:<#<<<<????#=

Or FASTQ file sequenced by Illumina HiSeq^{TM}2000 or other sequencer using Phred +64 quality system.
@A80GVTABXX:4:1:2587:1979#ACAGTGAT/1
NTTTGATATGTGTGAGGACGTCTGCAGCGTCACCTTTATCGGCCATGGT
+
BTTMKZXUUUddddddddddddddddddddddddddddaddddddd^WYYU

The first line in FASTQ format is description automatically generated by the sequence analyzer. The second line is sequenced bases and each base corresponds with a sequencing quality value on the fourth line. The base's quality score is equal to the ASCII value minus 33 or 64 depends on the quality system. e.g. The ASCII value of A is 65, then its base quality value is 63-33=32; the ASCII value of c is 99, then its base quality value is 99-64=35.Table 2-1 demonstrates the relationship between Illumina HiSeq^{TM} sequencer sequencing error rate and the sequencing quality value. Specifically, if the sequencing error rate is denoted as$E$, Illunima HiSeq^{TM} sequencer base quality value is denoted as sQ , the relationship is as follows:

$$sQ = -10\log_{10} E$$

Table 2-1 Relationship between Illumina HiSeq^{TM} sequencer error rate and sequencing quality

Sequencing error rate	Sequencing quality value	Character of Phred +33 quality system	Character of Phred +64 quality system
5%	13	.	M
1%	20	5	T
0.1%	30	?	^

Column description for data production

Header	Description
Sample Name	Sample name that identify each sample
Insert Size (bp)	The length of sequencing fragment (bp)
Read length (bp)	The length of the total reads (bp)
Clean Reads	Total clean reads number
Clean bases	Total clean bases number
Q20 (%)	The number of nucleotide with quality higher than 20/nucleotide (clean read1,read2)
GC (%)	GC number / nucleotide (clean read1,read2)

Appendix 6

Tables A to F show the top 20 up-regulated and down-regulated genes with their products for both R20291 and φCDHS1 during infection at 0, 10, 20, 30, 40 and 50 min. The φCDHS1 genes are highlighted in yellow. Expression level changes considered significant if they differed at least twofold (Log2 fold change) and at *P* value of <0.01.

No.	locus_tag/Gene	FoldChange(Infected vs. control 0 min)	Product/function
1	CDHS1_00037	654.211	Cro/C1-type transcriptional regulator
2	CDHS1_00046	564.626	hypothetical protein
3	CDHS1_00049	561.716	hypothetical protein
4	CDHS1_00040	519.891	hypothetical protein
5	CDHS1_00038	504.148	hypothetical protein
6	CDHS1_00045	495.471	hypothetical protein
7	CDHS1_00039	470.415	hypothetical protein
8	CDHS1_00047	466.066	single-stranded DNA-binding protein
9	CDHS1_00041	423.605	replication terminator-like protein
10	CDHS1_00035	325.126	DNAse/helicase
11	CDHS1_00042	273.94	hypothetical protein
12	CDHS1_00044	268.932	hypothetical protein
13	CDHS1_00028	162.17	hypothetical protein
14	CDR20291_2456	20.0979	nucleotide binding
15	CDR20291_1626	11.8031	sodium-dependent phosphate transport
16	CDR20291_2119	10.4262	sodium-dependent phosphate transport
17	CDR20291_2118	4.83674	ATP binding
18	*hymA*	4.55331	iron, sulfur cluster binding
19	CDR20291_3014	4.41467	membrane
20	*hymB*	3.81349	iron, sulfur cluster binding
21	CDR20291_1888	-6.5172	membrane,integral component of membrane
22	CDR20291_0191	-6.90089	membrane,integral component of membrane
23	CDR20291_1107	-7.41948	membrane,integral component of membrane
24	CDR20291_0318	-7.50064	metabolic process
25	CDR20291_2417	-7.84032	sequence-specific DNA binding
26	CDR20291_1334	-8.32649	conserved hypothetical protein
27	CDR20291_0949	-8.38665	membrane,integral component of membrane
28	*fldX*	-8.42604	electron carrier activity
29	CDR20291_1472	-9.33038	peptidoglycan turnover
30	CDR20291_1184	-9.88288	hypothetical protein
31	CDR20291_3277	-10.5225	membrane,integral component of membrane
32	CDR20291_0665	-11.8302	glycerophospholipid metabolic process
33	CDR20291_0317	-12.2253	regulation of transcription, DNA-templated
34	CDR20291_1265	-12.3782	drug transmembrane transport
35	CDR20291_1546	-12.6894	transporter activity
36	CDR20291_1935	-15.3289	hypothetical protein
37	CDR20291_1934	-17.6984	hypothetical protein
38	CDR20291_1751	-21.3311	membrane,integral component of membrane
39	CDR20291_1279	-21.434	membrane
40	CDR20291_1889	-28.5816	acyl-CoA metabolic process

Table B.

No.	locus_tag/Gene	FoldChange(Infected vs. control 10 min)	Product/function
1	CDHS1_00012	721.281	hypothetical protein
2	CDHS1_00047	612.555	Cro/C1-type transcriptional regulator
3	CDHS1_00037	586.981	single-stranded DNA-binding protein
4	CDHS1_00002	566.739	terminase large subunit
5	CDHS1_00041	487.13	replication terminator-like protein
6	CDHS1_00039	470.491	hypothetical protein
7	CDHS1_00042	460.395	hypothetical protein
8	CDHS1_00020	455.887	hypothetical protein
9	CDHS1_00040	454.529	hypothetical protein
10	CDHS1_00013	449.029	hypothetical protein
11	CDHS1_00011	414.62	minor capsid protein
12	CDHS1_00010	400.269	hypothetical protein
13	CDHS1_00017	386.067	tail tape measure
14	CDHS1_00044	379.648	hypothetical protein
15	CDHS1_00022	374.96	tail protein
16	CDHS1_00051	367.136	integrase
17	CDHS1_00049	365.803	hypothetical protein
18	CDHS1_00050	352.071	sigma70/sigmaF-like protein
19	CDHS1_00019	350.366	tail endopeptidase
20	CDHS1_00014	347.124	hypothetical protein
21	CDR20291_3359	-8.72486	conserved hypothetical protein
22	CDR20291_1073	-9.26481	hypothetical protein
23	CDR20291_0319	-10.3825	putative membrane protein
24	CDR20291_0364	-10.9479	putative membrane protein
25	CDR20291_0320	-11.1369	hypothetical protein
26	*mtlA*	-11.5234	mannitol-specific IIbc component
27	CDR20291_0363	-11.9947	Radical SAM-superfamily protein
28	CDR20291_2417	-12.5656	conserved hypothetical protein
29	*mtlR*	-12.6454	putative transcription antiterminator
30	CDR20291_3387	-13.2019	conserved hypothetical protein
31	*mtlD*	-14.6155	mannitol-phosphate-dehydrogenase
32	CDR20291_2867	-14.9361	putative aminotransferase
33	*mtlF*	-16.1598	mannitol-specific IIa component
34	CDR20291_1750	-20.4967	binding protein
35	CDR20291_1751	-25.9699	permease protein
36	CDR20291_2418	-26.27	putative membrane protein
37	CDR20291_1889	-31.2186	putative acyl-CoA thioesterase
38	CDR20291_1938	-38.0721	permease protein
39	CDR20291_2419	-40.6191	putative aminotransferase
40	CDR20291_1752	-330.743	permease protein

Table C.

No.	locus_tag/Gene	FoldChange(Infected vs. control 20 min)	Product/function
1	CDHS1_00023	722.602	holin
2	CDHS1_00019	711.131	Distal tail protein
3	CDHS1_00025	546.103	Endolysin
4	CDHS1_00031	510.54	Hypothetical protein
5	CDHS1_00021	505.268	tail fiber protein
6	CDHS1_00002	489.217	terminase large subunit
7	CDHS1_00020	472.23	Hypothetical protein
8	CDHS1_00032	442.065	DNA-binding protein
9	CDHS1_00005	436.93	Hypothetical protein
10	CDHS1_00042	397.429	Hypothetical protein
11	CDHS1_00028	391.181	Hypothetical protein
12	CDHS1_00022	385.525	tail fiber protein
13	CDHS1_00013	381.488	Hypothetical protein
14	CDHS1_00050	380.02	sigma70/sigmaF-like protein
15	CDHS1_00007	375.468	scaffolding protein
16	CDHS1_00010	371.915	Hypothetical protein
17	CDHS1_00051	366.062	integrase
18	CDHS1_00018	359.421	tail protein
19	CDHS1_00040	349.339	Hypothetical protein
20	CDHS1_00011	347.509	minor capsid protein
21	CDR20291_2050	-8.76834	putative transcriptional regulator
22	CDR20291_0203	-9.28835	conserved hypothetical protein
23	CDR20291_2416	-9.32831	conserved hypothetical protein
24	CDR20291_1107	-9.46141	permease protein
25	CDR20291_1546	-10.5254	permease protein
26	CDR20291_2867	-10.814	putative aminotransferase
27	CDR20291_1575	-12.3802	putative membrane protein
28	CDR20291_1889	-15.2629	putative acyl-CoA thioesterase
29	CDR20291_1184	-16.1539	hypothetical protein
30	CDR20291_0319	-16.7031	putative membrane protein
31	CDR20291_0364	-21.0257	putative membrane protein
32	CDR20291_0363	-24.3725	Radical SAM protein
33	CDR20291_3194	-27.576	putative glycosyl transferase
34	CDR20291_2417	-31.4166	conserved hypothetical protein
35	CDR20291_3154	-43.4484	hypothetical protein
36	CDR20291_1751	-49.3662	permease protein
37	CDR20291_0979	-63.5015	hypothetical protein
38	CDR20291_1752	-63.7285	permease protein
39	CDR20291_2419	-82.2712	putative aminotransferase
40	CDR20291_2418	-83.2952	putative membrane protein

Table D.

No.	locus_tag/Gene	FoldChange(Infected vs. control 30 min)	Product/function
1	lCDHS1_00032	1318.59	DNA-binding protein
2	CDHS1_00025	980.483	Endolysin
3	CDHS1_00023	979.053	holin
4	CDHS1_00001	866.272	terminase small subunit
5	CDHS1_00045	834.868	hypothetical protein
6	CDHS1_00013	823.313	hypothetical protein
7	CDHS1_00012	781.766	hypothetical protein
8	CDHS1_00031	779.161	Hypothetical protein
9	CDHS1_00002	723.079	terminase large subunit
10	CDHS1_00050	669.938	sigma70/sigmaF-like protein
11	CDHS1_00007	641.889	scaffolding protein
12	CDHS1_00040	637.818	Hypothetical protein
13	CDHS1_00010	625.585	Hypothetical protein
14	CDHS1_00021	622.64	tail fiber protein
15	CDHS1_00011	600.556	minor capsid protein
16	CDHS1_00019	592.162	Distal tail protein
17	CDHS1_00018	574.821	tail protein
18	CDHS1_00020	555.707	Hypothetical protein
19	CDHS1_00014	517.953	Hypothetical protein
20	CDHS1_00051	512.887	integrase
21	CDR20291_2104	-7.30782	putative subunit of oxidoreductase
22	CDR20291_0203	-7.49629	conserved hypothetical pr.
23	CDR20291_1753	-8.36869	utative uncharacterized pr.
24	*fhuB*	-8.46628	permease protein
25	CDR20291_2867	-8.84382	putative aminotransferase
26	CDR20291_1752	-9.35524	permease protein
27	CDR20291_1750	-9.89837	atp-binding protein
28	CDR20291_1691	-10.1105	nitrite and sulfite reductase subunit
29	CDR20291_2416	-11.062	conserved hypothetical protein
30	CDR20291_1889	-12.5392	putative acyl-CoA thioesterase
31	CDR20291_0979	-13.6647	hypothetical protein
32	CDR20291_1692	-15.8439	pyridine n-disulfide oxidoreductase
33	CDR20291_1546	-17.0769	permease protein
34	CDR20291_1334	-17.2296	conserved hypothetical protein
35	CDR20291_1751	-21.1816	permease protein
36	CDR20291_0364	-24.3586	putative membrane protein
37	CDR20291_0363	-26.0195	Radical SAM-superfamily protein
38	CDR20291_2417	-33.6534	conserved hypothetical protein
39	CDR20291_2419	-81.7977	putative aminotransferase
40	CDR20291_2418	-83.178	putative membrane protein

Table E.

No.	locus_tag/Gene	FoldChange(Infected vs. control 40 min)	Product/function
1	CDHS1_00005	1091.17	hypothetical protein
2	CDHS1_00010	1076.3	Hypothetical protein
3	CDHS1_00025	957.978	Endolysin
4	CDHS1_00021	818.306	tail fiber protein
5	CDHS1_00004	799.91	minor capsid protein
6	CDHS1_00013	789.05	hypothetical protein
7	CDHS1_00023	770.38	holin
8	CDHS1_00019	768.284	Distal tail protein
9	CDHS1_00007	752.144	scaffolding protein
10	CDHS1_00001	726.61	terminase small subunit
11	CDHS1_00012	721.443	hypothetical protein
12	CDHS1_00017	710.088	tail tape measure
13	CDHS1_00014	680.081	Hypothetical protein
14	CDHS1_00011	655.499	minor capsid protein
15	CDHS1_00018	646.67	tail protein
16	CDHS1_00020	640.583	Hypothetical protein
17	CDHS1_00024	627.541	acetylmuramoyl-alanine amidase
18	CDHS1_00051	619.045	integrase
19	CDHS1_00022	598.937	tail protein
20	CDHS1_00002	587.468	terminase large subunit
21	CDR20291_1546	-5.76995	permease protein
22	CDR20291_1888	-6.06205	putative membrane protein
23	CDR20291_2905	-6.36114	glycosyl hydrolase
24	CDR20291_0020	-7.22069	transcriptional regulator
25	CDR20291_0724	-7.74541	putative membrane protein
26	CDR20291_2532	-7.87691	sporulation sigma E factor peptidase
27	CDR20291_1750	-8.78868	putative lantibiotic abc transporter
28	CDR20291_0300	-9.78877	putative biotin synthase
29	CDR20291_3277	-10.3145	putative exported protein
30	CDR20291_1889	-12.0696	putative acyl-CoA thioesterase
31	CDR20291_2417	-14.9121	conserved hypothetical protein
32	CDR20291_1752	-16.4006	permease protein
33	CDR20291_1691	-19.8677	nitrite and sulfite reductase subunit
34	CDR20291_0364	-20.8629	putative membrane protein
35	CDR20291_0363	-21.1467	Radical SAM-superfamily protein
36	CDR20291_1692	-25.4408	pyridine n-disulfide oxidoreductase
37	CDR20291_0728	-28.1149	hydroxymethylglutaryl-CoA lyase
38	CDR20291_2418	-40.4521	putative membrane protein
39	CDR20291_2419	-42.8241	putative aminotransferase
40	CDR20291_1751	-137.447	permease protein

Table F.

No.	locus_tag/Gene	FoldChange(Infected vs. control 50 min)	Product/function
1	CDHS1_00021	1343.49	tail fiber protein
2	CDHS1_00025	1144.53	Endolysin
3	CDHS1_00011	971.929	minor capsid protein
4	CDHS1_00001	921.066	terminase small subunit
5	CDHS1_00022	907.514	tail protein
6	CDHS1_00007	885.546	scaffolding protein
7	CDHS1_00023	852.48	holin
8	CDHS1_00012	848.368	hypothetical protein
9	CDHS1_00013	818.554	hypothetical protein
10	CDHS1_00050	815.101	sigma70/sigmaF-like protein
11	CDHS1_00024	783.44	putative endolysin
12	CDHS1_00010	777.711	Hypothetical protein
13	CDHS1_00019	775.422	Distal tail protein
14	CDHS1_00002	750.84	terminase large subunit
15	CDHS1_00051	717.438	integrase
16	CDHS1_00020	620.057	Hypothetical protein
17	CDHS1_00018	618.126	tail protein
18	CDHS1_00017	616.396	tail tape measure
19	CDHS1_00015	592.191	Hypothetical protein
20	CDHS1_00014	585.389	Hypothetical protein
21	CDR20291_0602	-6.11481	putative membrane protein precursor
22	*fhuG*	-6.3018	permease protein
23	*fhuB*	-7.49698	permease protein
24	CDR20291_1750	-7.53438	atp-binding protein
25	CDR20291_1744	-7.58244	ite-specific recombinase
26	CDR20291_0203	-8.29159	conserved hypothetical protein
27	CDR20291_0319	-9.43268	putative membrane protein
28	CDR20291_1889	-9.62194	putative acyl-CoA thioesterase
29	CDR20291_2417	-10.2041	conserved hypothetical protein
30	CDR20291_3207	-10.5323	permease protein
31	CDR20291_2419	-11.8593	putative aminotransferase
32	CDR20291_2497	-12.195	conserved hypothetical protein
33	CDR20291_0364	-12.3357	putative membrane protein
34	CDR20291_2867	-12.4293	putative aminotransferase
35	CDR20291_0363	-13.4737	Radical SAM-superfamily protein
36	CDR20291_1546	-13.7981	permease protein
37	CDR20291_2418	-14.4336	putative membrane protein
38	CDR20291_1751	-22.587	permease protein
39	CDR20291_1691	-26.8781	nitrite and sulfite reductase subunit
40	CDR20291_1692	-28.5576	pyridine n-disulfide oxidoreductase

Appendix 7

Table showing the oligonucleotides used to detect contamination of phage CDHS1. 6.2.1.2.4 Section.

Name	Forward and reverse primers	Target	Expected product size (bp)	Reference
F R	5'-TTGAGCGATTTACTTCGGTAAAGA-3' 5'- CCATCCTGTACTGGCTCACCT-3'	*C. difficile* 16s rDNA	157 bp	(Rinttilä et al., 2004)
8F 1391R	5'- AGAGTTTGATCCTGGCTCAG -3' 5'- GACGGGCGGTGTGTRCA -3'	Universal bacterial 16s rDNA	1-1.8 kb	(Turner et al., 1999)

Appendix 8

Table. Primers used for real-time RT-PCR analysis of *Galleria* genes

Genes	Forward primer	Reverse primer
Moricin	5'-GCG ATC ATT GCC CTC TTT AT-3'	5'-AGT GCC TTC TGT TTT TAA TGT GTT C-3'
Gloverin	5'-ATG GGA ACA GCA ATC ACC TC-3'	5'-TGC CTT GTG CAG ATA TTT CG-3'
Ecdysteroid 22-kinase	5'-CGA TCA CTG TAT CCA AGT CTG G-3'	5'-TGG GCA CAA CTA ATC TCC TTG-3'
Ecdysteroid-regulated protein	5'-GTT ATA GTG CTG CTG GCG AGT-3'	5'-TGC ACC TCC GTC ACT AAA GAT-3'
Juvenile hormone-inducible	5'-CGT CGC TAC AAT ATG AAA CTC G-3'	5'-AAC CTC CTT CGT CCA ACA AAT-3'
Juvenile hormone binding protein 1	5'-GTG TCG CAA TCT CTC TTC CTG-3'	5'-GCG ATG TAT GGT TCA AGT CGT-3'
Juvenile hormone binding protein 2	5'-TGG AGC ATT ATG AAG CTG GTC-3'	5'-TCG ACT ATC CTG CAA GCT GTT-3'
Juvenile hormone binding protein 3	5'-GTC GCC CGA GCT ACT TAC AAT-3'	5'-TGC AGC TGA TAT CGT CTT CAA T-3'
Juvenile hormone binding protein 4	5'-TGT TCA CGT CAA TCT GTC CAA-3'	5'-AGC ACT TCT GTC TCA CCG AAA-3'
Juvenile hormone epoxide hydrolase 1	5'-CAC TGG TGG TGG AAG ACA AAT-3'	5'-TCT CCG AGT CCA AGA TGA GAA-3'
Juvenile hormone epoxide hydrolase 2	5'-CAG TCT CCA CCC TGG ATG TAA-3'	5'-CGA AGT CAT TCC TCT GCT CAC-3'
Juvenile hormone esterase	5'-TAC GGT ATT TGC ACG GGT TTA-3'	5'-TAT TTG GAT ACG GCG ATG AAG-3'
GME-string_Contig_704.0	5'-GAG GTG CCT GGT CAG AAT GT-3'	5'-TGG GGT CAC TAT TCG CTT TC-3'
GME-string_Contig_233.0	5'-CGT GAC GAG TGA GAT CGT TG-3'	5'-CGG TTG TTG TCT CCG TTA CA-3'
Housekeeping genes		
Actin	5'-ATC CTC ACC CTG AAG TAC CC-3'	5'-CCA CAC GCA GCT CAT TGT A-3'
18S r-RNA	5'-ATG GTT GCA AAG CTG AAA CT-3'	5'-TCC CGT GTT GAG TCA AAT TA-3'

Chapter 9

Chapter 9 Bibliography

Abbasifar, R., Kropinski, A. M., Sabour, P. M., Chambers, J. R., MacKinnon, J., Malig, T. & Griffiths, M. W. 2014. Efficiency of bacteriophage therapy against Cronobacter sakazakii in *Galleria mellonella* (greater wax moth) larvae. *Archives of virology,* 159**,** 2253-2261.

Abdelnour, A., Arvidson, S., Bremell, T., Ryden, C. & Tarkowski, A. 1993. The accessory gene regulator (agr) controls *Staphylococcus aureus* virulence in a murine arthritis model. *Infection and immunity,* 61**,** 3879-3885.

Abedon, S. T., Herschler, T. D. & Stopar, D. 2001. Bacteriophage latent-period evolution as a response to resource availability. *Appl Environ Microbiol,* 67**,** 4233-41.

Abedon, S. T. & Thomas-Abedon, C. 2010. Phage therapy pharmacology. *Current pharmaceutical biotechnology,* 11**,** 28-47.

Ackermann, H. W. 1998. Tailed bacteriophages: the order Caudovirales. *Advances in virus research. Elsevier.*

Ackermann, H. W. 2009. Phage classification and characterization. *Bacteriophages.* Springer.

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B. & Moreno, R. F. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science,* 252**,** 1651-1656.

Adams, M. H. 1959. Bacteriophages. *Bacteriophages.*

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

Aksyuk, A. A., Leiman, P. G., Kurochkina, L. P., Shneider, M. M., Kostyuchenko, V. A., Mesyanzhinov, V. V. & Rossmann, M. G. 2009. The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. *The EMBO journal,* 28**,** 821-829.

Albers, S., Verena & Meyer, B. H. 2011. The archaeal cell envelope. *Nature Reviews Microbiology,* 9**,** 414.

Alghoribi, M. F., Gibreel, T. M., Dodgson, A. R., Beatson, S. A. & Upton, M. 2014. *Galleria mellonella* infection model demonstrates high lethality of ST69 and ST127 uropathogenic *E. coli*. *PloS one,* 9**,** e101547.

Alyousef, A. 2013. *Identification and characterisation of lysin enzymes as potential therapeutics for the treatment of Clostridium difficile.* Ph.D. thesis thesis, University of Cardiff.

Aperis, G., Fuchs, B. B., Anderson, C. A., Warner, J. E., Calderwood, S. B. & Mylonakis, E. 2007. *Galleria mellonella* as a model host to study infection by the *Francisella tularensis* live vaccine strain. *Microbes and infection,* 9**,** 729-734.

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y. & Wishart, D. S. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic acids research,* 44**,** W16-W21.

Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. 1998. The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion. *Infection and immunity,* 66**,** 1000-1007.

Arya, R. & Princy, S. A. 2013. An insight into pleiotropic regulators Agr and Sar: molecular probes paving the new way for antivirulent therapy. *Future microbiology,* 8**,** 1339-1353.

Ashelford, K. E., Fry, J. C., Bailey, M. J., Jeffries, A. R. & Day, M. J. 1999. Characterization of six bacteriophages of *Serratia liquefaciens* CP6 isolated from the sugar beet phytosphere. *Appl. Environ. Microbiol.,* 65**,** 1959-1965.

Autret, N., Raynaud, C., Dubail, I., Berche, P. & Charbit, A. 2003. Identification of the agr locus of *Listeria monocytogene*s: role in bacterial virulence. *Infection and immunity,* 71**,** 4463-4471.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M. & Kubal, M. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC genomics,* 9**,** 75.

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

Bacci, S., Molbak, K., Kjeldsen, M. K. & Olsen, K. E. 2011. Binary toxin and death after *Clostridium difficile* infection. *Emerging infectious diseases,* 17**,** 976.

Barbut, F., Decre, D., Lalande, V., Burghoffer, B., Noussair, L., Gigandon, A., Espinasse, F., Raskine, L., Robert, J. & Mangeol, A. 2005. Clinical features of *Clostridium difficile*-associated diarrhoea due to binary toxin (actin-specific ADP-ribosyltransferase)-producing strains. *Journal of medical microbiology,* 54**,** 181-185.

Barnoy, S., Gancz, H., Zhu, Y., Honnold, C. L., Zurawski, D. V. & Venkatesan, M. M. 2017. The *Galleria mellonella* larvae as an in vivo model for evaluation of Shigella virulence. *Gut microbes,* 8**,** 335-350.

Baron, S. & Peterson, J. W. 1996. Bacterial Pathogenesis.

Barth, H., Aktories, K., Popoff, M. R. & Stiles, B. G. 2004. Binary bacterial toxins: biochemistry, biology, and applications of common *Clostridium* and *Bacillus* proteins. *Microbiology and Molecular Biology Reviews,* 68**,** 373-402.

Bartlett, J. G. 1994. *Clostridium difficile*: history of its role as an enteric pathogen and the current state of knowledge about the organism. *Clinical Infectious Diseases,* 18**,** S265-S272.

Bartlett, J. G. 2008. Historical perspectives on studies of *Clostridium difficile* and *C. difficile* infection. *Clinical infectious diseases,* 46**,** S4-S11.

Bartlett, J. G., Chang, T. W., Gurwith, M., Gorbach, S. L. & Onderdonk, A. B. 1978. Antibiotic-associated pseudomembranous colitis due to toxin-producing clostridia. *New England journal of medicine,* 298**,** 531-534.

Bartlett , J. G., Chang , T. W., Gurwith , M., Gorbach , S. L. & Onderdonk , A. B. 1978. Antibiotic-Associated Pseudomembranous Colitis Due to Toxin-Producing Clostridia. *New England Journal of Medicine,* 298**,** 531-534.

Batinovic, S., Wassef, F., Knowler, S. A., Rice, D. T., Stanton, C. R., Rose, J., Tucci, J., Nittami, T., Vinh, A. & Drummond, G. R. 2019. Bacteriophages in Natural and Artificial Environments. *Pathogens,* 8**,** 100.

Belcaid, M., Bergeron, A. & Poisson, G. The evolution of the tape measure protein: units, duplications and losses. BMC bioinformatics, 2011. BioMed Central, S10.

Bella, DS., Ascenzi, P., Siarakas, S., Petrosillo, N. & di Masi, A. 2016a. *Clostridium difficile* Toxins A and B: Insights into Pathogenic Properties and Extraintestinal Effects. *Toxins (Basel),* 8.

Bella, D. S., Ascenzi, P., Siarakas, S., Petrosillo, N. & di Masi, A. 2016b. *Clostridium difficile* toxins A and B: insights into pathogenic properties and extraintestinal effects. *Toxins,* 8**,** 134.

Bianco, M., Fedele, G., Quattrini, A., Spigaglia, P., Barbanti, F., Mastrantonio, P. & Ausiello, C. M. 2011. Immunomodulatory activities of surface-layer proteins obtained from epidemic and hypervirulent *Clostridium difficile* strains. *Journal of medical microbiology,* 60**,** 1162-1167.

Bingley, P. & Harding, G. 1987. *Clostridium difficile* colitis following treatment with metronidazole and vancomycin. *Postgraduate medical journal,* 63**,** 993-994.

Bishai, W. 1988. Bacteriophage gene products that cause human disease. *The bacteriophages.,* 2**,** 683-724.

Blair, K. M., Turner, L., Winkelman, J. T., Berg, H. C. & Kearns, D. B. 2008. A molecular clutch disables flagella in the *Bacillus subtilis* biofilm. *science,* 320**,** 1636-1638.

Bolger-Munro, M., Cheung, K., Fang, A. & Wang, L. 2013. T4 bacteriophage average burst size varies with *Escherichia coli* B23 cell culture age. *J Exp Microbiol Immunol,* 17**,** 115-119.

Borody, T. J. & Campbell, J. 2012. Fecal microbiota transplantation: techniques, applications, and issues. *Gastroenterology Clinics,* 41**,** 781-803.

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

Borriello, S. 1998. Pathogenesis of Clostridium difficile infection. *The Journal of antimicrobial chemotherapy,* 41**,** 13-19.

Borriello, S., Davies, H., Kamiya, S., Reed, P. & Seddon, S. 1990. Virulence factors of Clostridium difficile. *Reviews of infectious diseases,* 12**,** S185-S191.

Borriello, S., Davies, H. A. & Barclay, F. E. 1988. Detection of fimbriae amongst strains of Clostridium difficile. *FEMS microbiology letters,* 49**,** 65-67.

Bouillaut, L., Dubois, T., Sonenshein, A. L. & Dupuy, B. 2015. Integration of metabolism and virulence in *Clostridium difficile*. *Research in microbiology,* 166**,** 375-383.

Boyd, E. F. 2012. Bacteriophage-encoded bacterial virulence factors and phage–pathogenicity island interactions. *Advances in virus research.* Elsevier.

Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology,* 34**,** 525.

Brazier, J. 1998a. The diagnosis of Clostridium difficile-associated disease. *The Journal of antimicrobial chemotherapy,* 41**,** 29-40.

Brazier, J. 1998b. The diagnosis of *Clostridium difficile* -associated disease. *The Journal of antimicrobial chemotherapy,* 41**,** 29-40.

Brouwer, M. S., Allan, E., Mullany, P. & Roberts, A. P. 2012. Draft genome sequence of the nontoxigenic *Clostridium difficile* strain CD37. Am Soc Microbiol.

Brown, T. A. 2006. *Genomes 3,* London; New York, N.Y: Garland Science.

Brussow, H., Canchaya, C. & Hardt, W.-D. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiology and molecular biology reviews,* 68**,** 560-602.

Brussow, H. & Hendrix, R. W. 2002. Phage genomics: small is beautiful. *Cell,* 108**,** 13-16.

Bryan, D., El-Shibiny, A., Hobbs, Z., Porter, J. & Kutter, E. M. 2016. Bacteriophage T4 infection of stationary phase *E. coli*: Life after log from a phage perspective. *Frontiers in microbiology,* 7**,** 1391.

Buermans, H. & Den Dunnen, J. 2014. Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1842**,** 1932-1941.

Buffie, C. G. & Pamer, E. G. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology,* 13**,** 790.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W. & Shipley, G. L. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry,* 55**,** 611-622.

Calabi, E., Ward, S., Wren, B., Paxton, T., Panico, M., Morris, H., Dell, A., Dougan, G. & Fairweather, N. 2001. Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Molecular microbiology,* 40**,** 1187-1199.

Calap, D. & Martínez, D. 2018. Bacteriophages: protagonists of a post-antibiotic era. *Antibiotics,* 7**,** 66.

Cao, D., Ji, W., Yan, Y., Wang, H., Sun, J. & Lu, C. 2014. Effect of Stx2-encoding phage on the motility and gene expression involved in moving of *Escherichia coli* lysogen. *Acta microbiologica Sinica,* 54**,** 737-745.

Cao, Z., Zhang, J., Niu, Y. D., Cui, N., Ma, Y., Cao, F., Jin, L., Li, Z. & Xu, Y. 2015. Isolation and characterization of a "phiKMV-like" bacteriophage and its therapeutic effect on mink hemorrhagic pneumonia. *PLoS One,* 10**,** e0116571.

Cardarelli, L., Lam, R., Tuite, A., Baker, L. A., Sadowski, P. D., Radford, D. R., Rubinstein, J. L., Battaile, K. P., Chirgadze, N. & Maxwell, K. L. 2010. The crystal structure of bacteriophage HK97 gp6: defining a large family of head–tail connector proteins. *Journal of molecular biology,* 395**,** 754-768.

Carman, R. J., Stevens, A. L., Lyerly, M. W., Hiltonsmith, M. F., Stiles, B. G. & Wilkins, T. D. 2011. *Clostridium difficile* binary toxin (CDT) and diarrhea. *Anaerobe,* 17**,** 161-165.

Carroll, K. C. 2011. Tests for the diagnosis of *Clostridium difficile* infection: the next generation. *Anaerobe,* 17**,** 170-174.

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

Castillo, D., Kauffman, K., Hussain, F., Kalatzis, P., Rørbo, N., Polz, M. F. & Middelboe, M. 2018. Widespread distribution of prophage-encoded virulence factors in marine Vibrio communities. *Scientific reports,* 8**,** 9973.

Catalao, M. J., Gil, F., Moniz-Pereira, J., Sao-Jose, C. & Pimentel, M. 2013. Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS microbiology reviews,* 37**,** 554-571.

Cerquetti, M., Molinari, A., Sebastianelli, A., Diociaiuti, M., Petruzzelli, R., Capo, C. & Mastrantonio, P. 2000. Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microbial pathogenesis,* 28**,** 363-372.

Champion, O. L., Cooper, I. A., James, S. L., Ford, D., Karlyshev, A., Wren, B. W., Duffield, M., Oyston, P. C. & Titball, R. W. 2009. *Galleria mellonella* as an alternative infection model for *Yersinia pseudotuberculosis*. *Microbiology,* 155**,** 1516-1522.

Chandrasekaran, R. & Lacy, D. B. 2017. The role of toxins in *Clostridium difficile* infection. *FEMS Microbiol Rev,* 41**,** 723-750.

Chanishvili, N. 2012. Phage therapy-history from Twort and d'Herelle through Soviet experience to current approaches. *Advances in virus research.* Elsevier.

Chatterjee, A. N. 1969. Use of bacteriophage-resistant mutants to study the nature of the bacteriophage receptor site of *Staphylococcus aureus*. *Journal of Bacteriology,* 98**,** 519-527.

Chen, Y.-L., Brand, A., Morrison, E. L., Silao, F. G. S., Bigol, U. G., Malbas, F. F., Nett, J. E., Andes, D. R., Solis, N. V. & Filler, S. G. 2011. Calcineurin controls drug tolerance, hyphal growth, and virulence in *Candida dubliniensis*. *Eukaryotic cell,* EC. 00310-10.

Chen, Y., Golding, I., Sawai, S., Guo, L. & Cox, E. C. 2005. Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. *PLoS Biol,* 3**,** e229.

Chu, Y. & Corey, D. R. 2012. RNA sequencing: platform selection, experimental design, and data interpretation. *Nucleic acid therapeutics,* 22**,** 271-274.

Clabots, C., Lee, S., Gerding, D., Mulligan, M., Kwok, R., Schaberg, D., Fekety, R. & Peterson, L. 1988. *Clostridium difficile* plasmid isolation as an epidemiologic tool. *European Journal of Clinical Microbiology and Infectious Diseases,* 7**,** 312-315.

Claro, T., Daniels, S. & Humphreys, H. 2014. Detecting *Clostridium difficile* spores from inanimate surfaces of the hospital environment: which method is best? *Journal of clinical microbiology,* 52**,** 3426-3428.

Clements, A. C., Magalhães, R. J. S., Tatem, A. J., Paterson, D. L. & Riley, T. V. 2010. *Clostridium difficile* PCR ribotype 027: assessing the risks of further worldwide spread. *The Lancet infectious diseases,* 10**,** 395-404.

Clokie, M. R. 2009. Quantification of host and phage mRNA expression during infection using real-time PCR. *Bacteriophages.* Springer.

Clokie, M. R. & Kropinski, A. M. 2009. *Bacteriophages*: Springer.

Clooten, J., Kruth, S., Arroyo, L. & Weese, J. S. 2008. Prevalence and risk factors for *Clostridium difficile* colonization in dogs and cats hospitalized in an intensive care unit. *Veterinary microbiology,* 129**,** 209-214.

Coleman, d. C., Sullivan, D. J., Russel, R. J., Arbuthnott, J. P., Carey, B. F. & Pomeroy, H. M. 1989. *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of β-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. *Microbiology,* 135**,** 1679-1697.

Collins, M., Lawson, P., Willems, A., Cordoba, J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H. & Farrow, J. 1994. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *International Journal of Systematic and Evolutionary Microbiology,* 44**,** 812-826.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L. & Zhang, X. 2016. A survey of best practices for RNA-seq data analysis. *Genome biology,* 17**,** 13.

Connell, I., Agace, W., Klemm, P., Schembri, M., Mărild, S. & Svanborg, C. 1996. Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proceedings of the National Academy of Sciences,* 93**,** 9827-9832.

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

Costa, V., Angelini, C., De Feis, I. & Ciccodicola, A. 2010. Uncovering the complexity of transcriptomes with RNA-Seq. *BioMed Research International,* 2010.

Cotter, G., Doyle, S. & Kavanagh, K. 2000. Development of an insect model for the in vivo pathogenicity testing of yeasts. *FEMS Immunology & Medical Microbiology,* 27**,** 163-169.

Crobach, M., Dekkers, O., Wilcox, M. & Kuijper, E. 2009. European Society of Clinical Microbiology and Infectious Diseases (ESCMID): data review and recommendations for diagnosing *Clostridium difficile*-infection (CDI). *Clinical microbiology and infection,* 15**,** 1053-1066.

d'Herelle, F. 2007. On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. 1917. *Research in microbiology,* 158**,** 553.

Dailey, D., Kaiser, A. & Schloemer, R. H. 1987. Factors influencing the phagocytosis of *Clostridium difficile* by human polymorphonuclear leukocytes. *Infection and immunity,* 55**,** 1541-1546.

Dapa, T., Leuzzi, R., Ng, Y. K., Baban, S. T., Adamo, R., Kuehne, S. A., Scarselli, M., Minton, N. P., Serruto, D. & Unnikrishnan, M. 2013. Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. *Journal of bacteriology,* 195**,** 545-555.

Darkoh, C. & DuPont, H. L. 2017. The accessory gene regulator-1 as a therapeutic target for *Clostridium difficile* infections. Taylor & Francis.

Darkoh, C., DuPont, H. L., Norris, S. J. & Kaplan, H. B. 2015. Toxin synthesis by *Clostridium difficile* is regulated through quorum signaling. *MBio,* 6**,** e02569-14.

Darkoh, C., Odo, C. & DuPont, H. L. 2016. Accessory gene regulator-1 locus is essential for virulence and pathogenesis of *Clostridium difficile*. *MBio,* 7**,** e01237-16.

Davies, E. V., Winstanley, C., Fothergill, J. L. & James, C. E. 2016. The role of temperate bacteriophages in bacterial infection. *FEMS microbiology letters,* 363.

Davies, H. & Borriello, S. 1990. Detection of capsule in strains of *Clostridium difficile* of varying virulence and toxigenicity. *Microbial pathogenesis,* 9**,** 141-146.

Deepak, S., Kottapalli, K., Rakwal, R., Oros, G., Rangappa, K., Iwahashi, H., Masuo, Y. & Agrawal, G. 2007. Real-time PCR: revolutionizing detection and expression analysis of genes. *Current genomics,* 8**,** 234-251.

Delisle, A. L., Barcak, G. J. & Guo, M. 2006. Isolation and expression of the lysis genes of *Actinomyces naeslundii* phage Av-1. *Appl. Environ. Microbiol.,* 72**,** 1110-1117.

Deneve, C., Janoir, C., Poilane, I., Fantinato, C. & Collignon, A. 2009. New trends in *Clostridium difficile* virulence and pathogenesis. *International journal of antimicrobial agents,* 33**,** S24-S28.

Desbois, A. P. & Coote, P. J. 2012. Utility of greater wax moth larva (*Galleria mellonella*) for evaluating the toxicity and efficacy of new antimicrobial agents. *Advances in applied microbiology.* Elsevier.

Dewey, J. S., Savva, C. G., White, R. L., Vitha, S., Holzenburg, A. & Young, R. 2010. Micron-scale holes terminate the phage infection cycle. *Proceedings of the National Academy of Sciences,* 107**,** 2219-2223.

Diago-Navarro, E., Chen, L., Passet, V., Burack, S., Ulacia-Hernando, A., Kodiyanplakkal, R. P., Levi, M. H., Brisse, S., Kreiswirth, B. N. & Fries, B. C. 2014. Carbapenem-resistant *Klebsiella pneumoniae* exhibit variability in capsular polysaccharide and capsule associated virulence traits. *The Journal of infectious diseases,* 210**,** 803-813.

Diep, D. B., Håvarstein, L., Nissen-Meyer, J. & Nes, I. 1994. The gene encoding plantaricin A, a bacteriocin from *Lactobacillus plantarum* C11, is located on the same transcription unit as an agr-like regulatory system. *Appl. Environ. Microbiol.,* 60**,** 160-166.

Dingle, K. E., Griffiths, D., Didelot, X., Evans, J., Vaughan, A., Kachrimanidou, M., Stoesser, N., Jolley, K. A., Golubchik, T. & Harding, R. M. 2011. Clinical *Clostridium difficile*: clonality and pathogenicity locus diversity. *PloS one,* 6**,** e19993.

Dobzhansky, T. 1950. Mendelian populations and their evolution. *American Naturalist,* 401-418.

Donskey, C. J., Kundrapu, S. & Deshpande, A. 2015. Colonization versus carriage of *Clostridium difficile*. *Infectious Disease Clinics,* 29**,** 13-28.

Doval, G. & Raaij, V. M. 2012. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proceedings of the National Academy of Sciences,* 109**,** 9390-9395.

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

Doyle, C. K., Labruna, M. B., Breitschwerdt, E. B., Tang, Y.-W., Corstvet, R. E., Hegarty, B. C., Bloch, K. C., Li, P., Walker, D. H. & McBride, J. W. 2005. Detection of medically important *Ehrlichia* by quantitative multicolor TaqMan real-time polymerase chain reaction of the dsb gene. *The Journal of molecular diagnostics,* 7**,** 504-510.

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

Dubberke, E. R. 2010. Prevention of healthcare-associated *Clostridium difficile* infection: what works? *Infection Control & Hospital Epidemiology,* 31**,** S38-S41.

Duffy, C. & Feiss, M. 2002. The large subunit of bacteriophage λ's terminase plays a role in DNA translocation and packaging termination. *Journal of molecular biology,* 316**,** 547-561.

Dumyati, G., Stevens, V., Hannett, G. E., Thompson, A. D., Long, C., Maccannell, D. & Limbago, B. 2012. Community-associated *Clostridium difficile* infections, Monroe County, New York, USA. *Emerg Infect Dis,* 18**,** 392-400.

Duplessis, M., Russell, W. M., Romero, D. A. & Moineau, S. 2005. Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. *Virology,* 340**,** 192-208.

Dupuy, B., Govind, R., Antunes, A. & Matamouros, S. 2008. *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. *Journal of medical microbiology,* 57**,** 685-689.

Dupuy, B. & Matamouros, S. 2006. Regulation of toxin and bacteriocin synthesis in *Clostridium* species by a new subgroup of RNA polymerase σ-factors. *Research in microbiology,* 157**,** 201-205.

Dupuy, B. & Sonenshein, A. L. 1998. Regulated transcription of *Clostridium difficile* toxin genes. *Molecular microbiology,* 27**,** 107-120.

Dy, R. L., Richter, C., Salmond, G. P. & Fineran, P. C. 2014. Remarkable mechanisms in microbes to resist phage infections. *Annual review of virology,* 1**,** 307-331.

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

Edwards-Ingram, L., Gitsham, P., Burton, N., Warhurst, G., Clarke, I., Hoyle, D., Oliver, S. G. & Stateva, L. 2007. Genotypic and physiological characterization of *Saccharomyces boulardii*, the probiotic strain of *Saccharomyces cerevisiae*. *Applied and environmetal microbiology,* 73**,** 2458-2467.

Edwards, A. N., Karim, S. T., Pascual, R. A., Jowhar, L. M., Anderson, S. E. & McBride, S. M. 2016. Chemical and Stress Resistances of *Clostridium difficile* Spores and Vegetative Cells. *Front Microbiol,* 7**,** 1698.

Eggertson, L. 2004. Quebec strikes committee on *Clostridium difficile*. *Canadian Medical Association Journal,* 171**,** 123-123.

Elliott, B., Androga, G. O., Knight, D. R. & Riley, T. V. 2017. *Clostridium difficile* infection: Evolution, phylogeny and molecular epidemiology. *Infect Genet Evol,* 49**,** 1-11.

Ellis, E. L. & Delbrück, M. 1939. The growth of bacteriophage. *The Journal of general physiology,* 22**,** 365-384.

Erickson, D. L., Russell, C. W., Johnson, K. L., Hileman, T. & Stewart, R. M. 2011. PhoP and OxyR transcriptional regulators contribute to *Yersinia pestis* virulence and survival within *Galleria mellonella*. *Microbial pathogenesis,* 51**,** 389-395.

Eriksson, S., Lucchini, S., Thompson, A., Rhen, M. & Hinton, J. C. D. 2003. Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Molecular Microbiology,* 47**,** 103-118.

Evans, B. & Rozen, D. 2012. A *Streptococcus pneumoniae* infection model in larvae of the wax moth *Galleria mellonella*. *European journal of clinical microbiology & infectious diseases,* 31**,** 2653-2660.

Fallico, V., Ross, R. P., Fitzgerald, G. F. & McAuliffe, O. 2011. Genetic response to bacteriophage infection in *Lactococcus lactis* reveals a four-strand approach involving induction of membrane stress proteins, D-alanylation of the cell wall, maintenance of proton motive force, and energy conservation. *J Virol,* 85**,** 12032-42.

Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I. & Herskovits, A. A. 2015. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nature Reviews Microbiology,* 13**,** 641.

Fekety, R. 1997. Guidelines for the diagnosis and management of *Clostridium difficile*-associated diarrhea and colitis. *American Journal of Gastroenterology,* 92.

Ferro, T. A. F., Araujo, J. M. M., Pinto, B. L. S., dos Santos, J. S., Souza, E. B., da Silva, B. L. R., Colares, V. L., Novais, T. M. G., Filho, C. M. B., Struve, C., Calixto, J. B., Monteiro-Neto, V., da Silva, L. C. N. & Fernandes, E. S. 2016. Cinnamaldehyde inhibits *staphylococcus aureus* virulence factors and protects against infection in a *galleria mellonella* model. *Frontiers in Microbiology,* 7.

Fischer, S., Kittler, S., Klein, G. & Glünder, G. 2013. Microplate-Test for the Rapid Determination of Bacteriophage-Susceptibility of *Campylobacter* Isolates-Development and Validation. *PLoS One,* 8**,** e53899.

Fogg, P. C., Colloms, S., Rosser, S., Stark, M. & Smith, M. C. 2014. New applications for phage integrases. *Journal of molecular biology,* 426**,** 2703-2716.

Ford, E. B. 1965. *Genetic polymorphism*: MIT Press Cambridge, Massachusetts.

Fortier, L.-C. 2018. Bacteriophages contribute to shaping *clostridioides* (*Clostridium*) *difficile* species. *Frontiers in microbiology,* 9.

Fortier, L.-C. & Sekulovic, O. 2013. Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence,* 4**,** 354-365.

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

Freeman, V. J. 1951. Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *Journal of bacteriology,* 61**,** 675.

Frobisher , M. & Brown, J. H. 1927. Transmissible Toxicogenicity of Streptococci. *Bulletin of the Johns Hopkins Hospital,* 41**,** 167-173.

Frye, J. G., Porwollik, S., Blackmer, F., Cheng, P. & McClelland, M. 2005. Host gene expression changes and DNA amplification during temperate phage induction. *Journal of bacteriology,* 187**,** 1485-1492.

Fuchs, B. B., O'Brien, E., El Khoury, J. B. & Mylonakis, E. 2010. Methods for using *Galleria mellonella* as a model host to study fungal pathogenesis. *Virulence,* 1**,** 475-482.

Garneau, J. R., Sekulovic, O., Dupuy, B., Soutourina, O., Monot, M. & Fortier, L.-C. 2018. High prevalence and genetic diversity of large phiCD211 (phiCDIF1296T)-like prophages in *Clostridioides difficile*. *Appl. Environ. Microbiol.,* 84**,** e02164-17.

Gaulton, T., Misra, R., Rose, G., Baybayan, P., Hall, R., Freeman, J., Turton, J., Picton, S., Korlach, J. & Gharbia, S. 2015. Complete genome sequence of the hypervirulent bacterium *Clostridium difficile* strain G46, ribotype 027. *Genome Announc.,* 3**,** e00073-15.

Gauss, P., Park, K., Spencer, T. E. & Hacker, K. J. 1994. DNA helicase requirements for DNA replication during bacteriophage T4 infection. *Journal of bacteriology,* 176**,** 1667-1672.

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

Gerding, D. N., Johnson, S., Rupnik, M. & Aktories, K. 2014. *Clostridium difficile* binary toxin CDT: mechanism, epidemiology, and potential clinical importance. *Gut microbes,* 5**,** 15-27.

Geric, B., Rupnik, M., Gerding, D. N., Grabnar, M. & Johnson, S. 2004. Distribution of *Clostridium difficile* variant toxinotypes and strains with binary toxin genes among clinical isolates in an American hospital. *Journal of medical microbiology,* 53**,** 887-894.

Giron, J. A., Torres, A. G., Freer, E. & Kaper, J. B. 2002. The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Molecular microbiology,* 44**,** 361-379.

Goh, S., Chang, B. J. & Riley, T. V. 2005a. Effect of phage infection on toxin production by *Clostridium difficile*. *Journal of medical microbiology,* 54**,** 129-135.

Goh, S., Ong, P. F., Song, K. P., Riley, T. V. & Chang, B. J. 2007. The complete genome sequence of *Clostridium difficile* phage ϕC2 and comparisons to ϕCD119 and inducible prophages of CD630. *Microbiology,* 153**,** 676-685.

Goh, S., Riley, T. V. & Chang, B. J. 2005b. Isolation and characterization of temperate bacteriophages of *Clostridium difficile*. *Applied and environmental microbiology,* 71**,** 1079-1083.

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

Gorbach, S., Barza, M., Giuliano, M. & Jacobus, N. 1988. Colonization resistance of the human intestinal microflora: testing the hypothesis in normal volunteers. *European Journal of Clinical Microbiology and Infectious Diseases,* 7**,** 98-102.

Gorski, A., Borysowski, J., Międzybrodzki, R. & Weber-Dąbrowska, B. 2007. Bacteriophage: genetics and molecular biology.

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

Govind, R., Fitzwater, L. & Nichols, R. 2015. Observations on the Role of TcdE Isoforms in *Clostridium difficile* Toxin Secretion. *J Bacteriol,* 197**,** 2600-9.

Govind, R., Fralick, J. A. & Rolfe, R. D. 2006. Genomic organization and molecular characterization of *Clostridium difficile* bacteriophage ΦCD119. *Journal of bacteriology,* 188**,** 2568-2577.

Govind, R., Vediyappan, G., Rolfe, R. D., Dupuy, B. & Fralick, J. A. 2009. Bacteriophage-mediated toxin gene regulation in *Clostridium difficile*. *Journal of virology,* 83**,** 12037-12045.

Grant, C., Konkel, M. E., Cieplak, W. & Tompkins, L. S. 1993. Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in

nonpolarized and polarized epithelial cell cultures. *Infection and immunity,* 61**,** 1764-1771.

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

Grunberg-Manago, M. 1999. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annual review of genetics,* 33**,** 193-227.

Grundling, A., Manson, M. D. & Young, R. 2001. Holins kill without warning. *Proceedings of the National Academy of Sciences,* 98**,** 9348-9352.

Guttman, B., Raya, R. & Kutter, E. 2005. Basic phage biology. *Bacteriophages: Biology and applications,* 4.

Hacker, J. & Ott, M. 1989. Genetic regulatory mechanisms of bacterial virulence. *Immunitat und Infektion,* 17**,** 199-205.

Hafiz, S. & Oakley, C. 1976. *Clostridium difficile*: isolation and characteristics (Plate VIII). *Journal of medical microbiology,* 9**,** 129-136.

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

Hamm, E. E., Voth, D. E. & Ballard, J. D. 2006. Identification of *Clostridium difficile* toxin B cardiotoxicity using a zebrafish embryo model of intoxication. *Proc Natl Acad Sci U S A,* 103**,** 14176-81.

Hammond, G. A. & Johnson, J. L. 1995. The toxigenic element of *Clostridium difficile* strain VPI 10463. *Microbial pathogenesis,* 19**,** 203-213.

Hanlon, G. W. 2007. Bacteriophages: an appraisal of their role in the treatment of bacterial infections. *International journal of antimicrobial agents,* 30**,** 118-128.

Harding, C. R., Schroeder, G. N., Collins, J. W. & Frankel, G. 2013. Use of *Galleria mellonella* as a model organism to study *Legionella pneumophila* infection. *Journal of visualized experiments: JoVE*.

Hargreaves, K. R. 2012. *Isolation and characterisation of bacteriophages infecting environmental strains of Clostridium Difficile.* Department of Infection, Immunity and Inflammation.university of Leicester.

Hargreaves, K. R. & Clokie, M. R. 2014. *Clostridium difficile* phages: still difficult? *Frontiers in microbiology,* 5**,** 184.

Haslam, S., Ketley, J., Mitchell, T., Stephen, J., Burdon, D. & Candy, D. 1986. Growth of *Clostridium difficile* and production of toxins A and B in complex and defined media. *Journal of medical microbiology,* 21**,** 293-297.

Hatfull, G. F., Cresawn, S. G. & Hendrix, R. W. 2008. Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. *Research in microbiology,* 159**,** 332-339.

He, M., Sebaihia, M., Lawley, T. D., Stabler, R. A., Dawson, L. F., Martin, M. J., Holt, K. E., Seth-Smith, H. M., Quail, M. A. & Rance, R. 2010. Evolutionary dynamics of Clostridium difficile over short and long time scales. *Proceedings of the National Academy of Sciences,* 107**,** 7527-7532.

Heller, K. & Braun, V. 1979. Accelerated adsorption of bacteriophage T5 to *Escherichia coli* F, resulting from reversible tail fiber-lipopolysaccharide binding. *Journal of bacteriology,* 139**,** 32-38.

Hendrix, R. W., Lawrence, J. G., Hatfull, G. F. & Casjens, S. 2000. The origins and ongoing evolution of viruses. *Trends in microbiology,* 8**,** 504-508.

Hendry, G. & Fitz, J. P. 1974. Characteristics of φT, the temperate bacteriophage carried by *Bacillus megaterium* 899a. *Journal of virology,* 13**,** 494-499.

Henn, M. R., Sullivan, M. B., Stange-Thomann, N., Osburne, M. S., Berlin, A. M., Kelly, L., Yandava, C., Kodira, C., Zeng, Q. & Weiand, M. 2010. Analysis of high-throughput sequencing and annotation strategies for phage genomes. *PLoS One,* 5**,** e9083.

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

Hernandez-Doria, J. D. & Sperandio, V. 2018. Bacteriophage transcription factor Cro regulates virulence gene expression in enterohemorrhagic *Escherichia coli*. *Cell host & microbe,* 23**,** 607-617. e6.

Hinton, D. M. 2010. Transcriptional control in the prereplicative phase of T4 development. *Virology journal,* 7**,** 289.

Hopman, N., Keessen, E., Harmanus, C., Sanders, I., Van Leengoed, L., Kuijper, E. & Lipman, L. 2011. Acquisition of *Clostridium difficile* by piglets. *Veterinary microbiology,* 149**,** 186-192.

Horgan, M., O'Sullivan, O., Coffey, A., Fitzgerald, G. F., van Sinderen, D., McAuliffe, O. & Ross, R. P. 2010. Genome analysis of the *Clostridium difficile* phage ΦCD6356, a temperate phage of the Siphoviridae family. *Gene,* 462**,** 34-43.

Hsu, B. B., Way, J. C. & Silver, P. A. 2019. Stable neutralization of virulent bacteria using temperate phage in the mammalian gut. *BioRxiv,* 794222.

Hubert, B., Loo, V. G., Bourgault, A.-M., Poirier, L., Dascal, A., Fortin, É., Dionne, M. & Lorange, M. 2007. A portrait of the geographic dissemination of the *Clostridium difficile* North American pulsed-field type 1 strain and the epidemiology of *C. difficile*-associated disease in Quebec. *Clinical infectious diseases,* 44**,** 238-244.

Huelsenbeck, S. C., May, M., Schmidt, G. & Genth, H. 2009. Inhibition of cytokinesis by *Clostridium difficile* toxin B and cytotoxic necrotizing factors—reinforcing the critical role of RhoA in cytokinesis. *Cell motility and the cytoskeleton,* 66**,** 967-975.

Hughes, K. T. & Mathee, K. 1998. The anti-sigma factors. *Annual Reviews in Microbiology,* 52**,** 231-286.

Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M. & Von Eichel-Streiber, C. 1997. Transcription analysis of the genes tcdA-E of the pathogenicity locus of *Clostridium difficile*. *European journal of biochemistry,* 244**,** 735-742.

Hyman, P. & Abedon, S. T. 2009. Practical methods for determining phage growth parameters. *Bacteriophages.* Springer.

Hyman, P. & van Raaij, M. 2018. Bacteriophage T4 long tail fiber domains. *Biophysical reviews,* 10**,** 463-471.

Illumina. 2011. RNA-Seq Data Comparison with Gene Expression Microarrays: A cross-platform comparison of differential gene expression analysis. San Diego: Illumina.

Ionel, A., Velazquez-Muriel, J. A., Luque, D., Cuervo, A., Caston, J. R., Valpuesta, J. M., Martin-Benito, J. & Carrascosa, J. L. 2011. Molecular rearrangements involved in the capsid shell maturation of bacteriophage T7. *Journal of Biological Chemistry,* 286**,** 234-242.

Janezic, S., Marín, M., Martín, A. & Rupnik, M. 2015. A new type of toxin A-negative, toxin B-positive *Clostridium difficile* strain lacking a complete tcdA gene. *Journal of clinical microbiology,* 53**,** 692-695.

Jangi, S. & Lamont, J. T. 2010. Asymptomatic colonization by *Clostridium difficile* in infants: implications for disease in later life. *Journal of pediatric gastroenterology and nutrition,* 51**,** 2-7.

Jank, T., Belyi, Y. & Aktories, K. 2015. Bacterial glycosyltransferase toxins. *Cellular microbiology,* 17**,** 1752-1765.

Janka, J. & O'Grady, N. P. 2009. *Clostridium difficile* infection: current perspectives. *Current opinion in critical care,* 15**,** 149-153.

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

Janzon, L., Löfdahl, S. & Arvidson, S. 1989. Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to the accessory gene regulator (agr) of *Staphylococcus aureus*. *Molecular and General Genetics MGG,* 219**,** 480-485.

Jiang, H., Vilcinskas, A. & Kanost, M. R. 2010. Immunity in lepidopteran insects. *Invertebrate immunity.* Springer.

Joshi, L. T., Welsch, A., Hawkins, J. & Baillie, L. 2017. The effect of hospital biocide sodium dichloroisocyanurate on the viability and properties of *Clostridium difficile* spores. *Letters in applied microbiology,* 65**,** 199-205.

Jump, R. L., Pultz, M. J. & Donskey, C. J. 2007. Vegetative *Clostridium difficile* survives in room air on moist surfaces and in gastric contents with reduced acidity: a potential mechanism to explain the association between proton pump inhibitors and *C. difficile*-associated diarrhea? *Antimicrobial agents and chemotherapy,* 51**,** 2883-2887.

Karasawa, T., Maegawa, T., Nojiri, T., Yamakawa, K. & Nakamura, S. 1997. Effect of arginine on toxin production by *Clostridium difficile* in defined medium. *Microbiology and immunology,* 41**,** 581-585.

Karlsson, S., Burman, L. G. & Åkerlund, T. 1999. Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids. *Microbiology,* 145**,** 1683-1693.

Karlsson, S., Dupuy, B., Mukherjee, K., Norin, E., Burman, L. G. & Åkerlund, T. 2003. Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. *Infection and immunity,* 71**,** 1784-1793.

Kasman, L. M., Kasman, A., Westwater, C., Dolan, J., Schmidt, M. G. & Norris, J. S. 2002. Overcoming the phage replication threshold: a mathematical model with implications for phage therapy. *Journal of virology,* 76**,** 5557-5564.

Kasper, D. L. 1976. The polysaccharide capsule of *Bacteroides fragilis* subspecies fragilis: immunochemical and morphologic definition. *Journal of Infectious Diseases,* 133**,** 79-87.

Keen, E. C. 2012. Paradigms of pathogenesis: targeting the mobile genetic elements of disease. *Frontiers in cellular and infection microbiology,* 2**,** 161.

Kelly, C. P. & LaMont, J. T. 1998. *Clostridium difficile* infection. *Annual review of medicine,* 49**,** 375-390.

Kelly, C. P. & LaMont, J. T. 2008. *Clostridium difficile*—more difficult than ever. *New England Journal of Medicine,* 359**,** 1932-1940.

Kim, G. H., Kim, J., Pai, H. & Kang, J. O. 2014. Comparison of supplemented brucella agar and modified *Clostridium difficile* agar for antimicrobial susceptibility testing of *Clostridium difficile*. *Annals of laboratory medicine,* 34**,** 439-445.

Kim, J., Smathers, S. A., Prasad, P., Leckerman, K. H., Coffin, S. & Zaoutis, T. 2008. Epidemiological features of *Clostridium difficile*-associated disease among inpatients at children's hospitals in the United States, 2001–2006. *Pediatrics,* 122**,** 1266-1270.

King, A. M., Lefkowitz, E., Adams, M. J. & Carstens, E. B. 2011. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*: Elsevier.

Kirov, S. M., Castrisios, M. & Shaw, J. G. 2004. *Aeromonas* flagella (polar and lateral) are enterocyte adhesins that contribute to biofilm formation on surfaces. *Infection and Immunity,* 72**,** 1939-1945.

Klempner, M. S. 1984. Interactions of polymorphonuclear leukocytes with anaerobic bacteria. *Reviews of infectious diseases,* 6**,** S40-S44.

Klingler, P., Metzger, P., Seelig, M., Pettit, P., Knudsen, J. & Alvarez, S. 2000. *Clostridium difficile* infection: risk factors, medical and surgical management. *Digestive Diseases,* 18**,** 147-160.

Klopot, A., Zakrzewska, A., Lecion, D., Majewska, J. M., Harhala, M. A., Lahutta, K., Kaźmierczak, Z., Łaczmański, Ł., Kłak, M. & Dąbrowska, K. 2017. Real-Time

qPCR as a method for detection of antibody-neutralized phage particles. *Frontiers in microbiology,* 8**,** 2170.

Knapp, O., Benz, R. & Popoff, M. R. 2016. Pore-forming activity of clostridial binary toxins. *Biochimica et Biophysica Acta (BBA)-Biomembranes,* 1858**,** 512-525.

Knaus, U. G., Hertzberger, R., Pircalabioru, G. G., Yousefi, S. P. M. & Branco dos Santos, F. 2017. Pathogen control at the intestinal mucosa–H2O2 to the rescue. *Gut microbes,* 8**,** 67-74.

Knight, D. R., Elliott, B., Chang, B. J., Perkins, T. T. & Riley, T. V. 2015. Diversity and evolution in the genome of *Clostridium difficile*. *Clinical microbiology reviews,* 28**,** 721-741.

Koch, G., Nadal-Jimenez, P., Cool, R. H. & Quax, W. J. 2014. Assessing *Pseudomonas* virulence with nonmammalian host: *Galleria mellonella*. *Pseudomonas Methods and Protocols.* Springer.

Krishna, M., Powell, N. & Borriello, S. 1996. Cell surface properties of *Clostridium difficile*: haemagglutination, relative hydrophobicity and charge. *Journal of medical microbiology,* 44**,** 115-123.

Kristensen, N. P., Scoble, M. J. & Karsholt, O. 2007. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa,* 1668**,** 699-747.

Kroghsbo, N. M., Mærkedahl, R. B. & Svenningsen, S. L. 2013. A quorum-sensing-induced bacteriophage defense mechanism. *MBio,* 4**,** e00362-12.

Kropinski, A. M. 2006. Phage therapy--everything old is new again. *Canadian Journal of Infectious Diseases and Medical Microbiology,* 17**,** 297-306.

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

Kuijper, E., Coignard, B., Tüll, P., *difficile,* E. S. G. f. *C.*, States, E. M., Prevention, t. E. C. f. D. & Control 2006. Emergence of *Clostridium difficile*-associated disease in North America and Europe. *Clinical microbiology and infection,* 12**,** 2-18.

Kutateladze, á. & Adamia, R. 2008. Phage therapy experience at the Eliava Institute. *Médecine et maladies infectieuses,* 38**,** 426-430.

Kutter, E. & Sulakvelidze, A. 2004. *Bacteriophages: biology and applications*: CRC Press.

Kwok, C. S., Arthur, A. K., Anibueze, C. I., Singh, S., Cavallazzi, R. & Loke, Y. K. 2012. Risk of *Clostridium difficile* infection with acid suppressing drugs and antibiotics: meta-analysis. *The American journal of gastroenterology,* 107**,** 1011.

Kyne, L., Hamel, M. B., Polavaram, R. & Kelly, C. P. 2002. Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*. *Clinical Infectious Diseases,* 34**,** 346-353.

Laanto, E., Bamford, J. K., Laakso, J. & Sundberg, L.-R. 2012. Phage-driven loss of virulence in a fish pathogenic bacterium. *PLoS One,* 7**,** e53157.

LaFee, S. & Buschman, H. 2017. Novel phage therapy saves patient with multidrug-resistant bacterial infection. *UC San Diego News Center*.

Lamont, I., Richardson, H., Carter, D. & Egan, J. 1993. Genes for the establishment and maintenance of lysogeny by the temperate coliphage 186. *Journal of bacteriology,* 175**,** 5286-5288.

Lamont, J. T. 2002. Theodore E. Woodward Award. How bacterial enterotoxins work: insights from in vivo studies. *Transactions of the American Clinical and Climatological Association,* 113**,** 167.

Landelle, C., Verachten, M., Legrand, P., Girou, E., Barbut, F. & Buisson, C. B. 2014. Contamination of healthcare workers' hands with *Clostridium difficile* spores after caring for patients with *C. difficile* infection. *Infection Control & Hospital Epidemiology,* 35**,** 10-15.

Lange, A., Beier, S., Huson, D. H., Parusel, R., Iglauer, F. & Frick, J.-S. 2018. Genome sequence of *Galleria mellonella* (greater wax moth). *Genome Announc.,* 6**,** e01220-17.

Lappalainen, T., Sammeth, M., Friedlander, M. R., AC't Hoen, P., Monlong, J., Rivas, M. A., Gonzalez-Porta, M., Kurbatova, N., Griebel, T. & Ferreira, P. G. 2013. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature,* 501**,** 506.

Lawley, T. D. & Walker, A. W. 2013. Intestinal colonization resistance. *Immunology,* 138**,** 1-11.

Leffler, D. A. & Lamont, J. T. 2015. *Clostridium difficile* infection. *N Engl J Med,* 372**,** 1539-48.

Legendre, M., Audic, S., Poirot, O., Hingamp, P., Seltzer, V., Byrne, D., Lartigue, A., Lescot, M., Bernadac, A. & Poulain, J. 2010. mRNA deep sequencing reveals 75 new genes and a complex transcriptional landscape in Mimivirus. *Genome research,* 20**,** 664-674.

Leskinen, K., Blasdel, B., Lavigne, R. & Skurnik, M. 2016. RNA-sequencing reveals the progression of phage-host interactions between φR1-37 and *Yersinia enterocolitica*. *Viruses,* 8**,** 111.

Li, C., Harmanus, C., Zhu, D., Meng, X., Wang, S., Duan, J., Liu, S., Fu, C., Zhou, P. & Liu, R. 2018. Characterization of the virulence of a non-RT027, non-RT078 and binary toxin-positive *Clostridium difficile* strain associated with severe diarrhea. *Emerging microbes & infections,* 7**,** 1-11.

Lin, D. M., Koskella, B. & Lin, H. C. 2017. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World journal of gastrointestinal pharmacology and therapeutics,* 8**,** 162.

Liu, S., Wang, X., Sun, F., Zhang, J., Feng, J., Liu, H., Rajendran, K. V., Sun, L., Zhang, Y., Jiang, Y., Peatman, E., Kaltenboeck, L., Kucuktas, H. & Liu, Z. 2013a. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. *Physiol Genomics,* 45**,** 462-76.

Liu, X., Jiang, H., Gu, Z. & Roberts, J. W. 2013b. High-resolution view of bacteriophage lambda gene expression by ribosome profiling. *Proceedings of the National Academy of Sciences,* 110**,** 11928-11933.

Livak, K. J. & Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D. & Darnell, J. 2000. Viruses: structure, function, and uses. *Molecular Cell Biology. 4th edition.* WH Freeman.

Loh, J. M., Adenwalla, N., Wiles, S. & Proft, T. 2013. *Galleria mellonella* larvae as an infection model for group A streptococcus. *Virulence,* 4**,** 419-428.

Luo, Z.-H., Yu, Y.-P., Jost, G., Liu, W.-H., Huang, X.-L. & Gu, L. 2016. Characterization of two bacteriophages for specific treatment of biofilm formed by aVibriosp. isolated from an abalone farm. *Aquaculture Research,* 47**,** 3964-3972.

Lwoff, A., Horne, R. & Tournier, P. A system of viruses. Cold Spring Harbor Symposia on Quantitative Biology, 1962. Cold Spring Harbor Laboratory Press, 51-55.

Lyerly, D. M., Krivan, H. C. & Wilkins, T. D. 1988. *Clostridium difficile*: its disease and toxins. *Clinical microbiology reviews,* 1**,** 1-18.

Lyon, S. A., Hutton, M. L., Rood, J. I., Cheung, J. K. & Lyras, D. 2016. CdtR regulates TcdA and TcdB production in *Clostridium difficile*. *PLoS pathogens,* 12**,** e1005758.

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

Mahony, D., Bell, P. & Easterbrook, K. 1985. Two bacteriophages of *Clostridium difficile*. *Journal of clinical microbiology,* 21**,** 251-254.

Mandlik, A., Swierczynski, A., Das, A. & Ton-That, H. 2008. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends in microbiology,* 16**,** 33-40.

Mani, N. & Dupuy, B. 2001. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Proceedings of the National Academy of Sciences,* 98**,** 5844-5849.

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

Mansour, N. M. 2017. Bacteriophages are natural gift, could we pay further attention! *journal-food-microbiology,* 1**,** 22-24.

Martin, G., Lazaro, J. M., Méndez, E. & Salas, M. 1989. Characterization of the phage phi 29 protein p5 as a single-stranded DNA binding protein. Function in phi 29 DNA-protein p3 replication. *Nucleic acids research,* 17**,** 3663.

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

Martinez, M. R., Wiedmann, M., Ferguson, M. & Datta, A. R. 2017. Assessment of *Listeria monocytogenes* virulence in the *Galleria mellonella* insect larvae model. *PloS one,* 12**,** e0184557.

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

Matsuda, T., Freeman, T. A., Hilbert, D. W., Duff, M., Fuortes, M., Stapleton, P. P. & Daly, J. M. 2005. Lysis-deficient bacteriophage therapy decreases endotoxin and inflammatory mediator release and improves survival in a murine peritonitis model. *Surgery,* 137**,** 639-646.

Mayer, M. J., Narbad, A. & Gasson, M. J. 2008. Molecular characterization of a *Clostridium difficile* bacteriophage and its cloned biologically active endolysin. *Journal of bacteriology,* 190**,** 6734-6740.

Mazaheri Nezhad Fard, R., Barton, M. & Heuzenroeder, M. 2011. Bacteriophage-mediated transduction of antibiotic resistance in enterococci. *Letters in applied microbiology,* 52**,** 559-564.

McDonald, L. C., Coignard, B., Dubberke, E., Song, X., Horan, T. & Kutty, P. K. 2007. Recommendations for surveillance of *Clostridium difficile*–associated disease. *Infection Control,* 28**,** 140-145.

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

McFarland, L. V., Mulligan, M. E., Kwok, R. Y. & Stamm, W. E. 1989. Nosocomial acquisition of *Clostridium difficile* infection. *New England journal of medicine,* 320**,** 204-210.

McGettigan, P. A. 2013. Transcriptomics in the RNA-seq era. *Current opinion in chemical biology,* 17**,** 4-11.

McGowan, A., Lalayiannis, L., Sarma, J., Marshall, B., Martin, K. & Welfare, M. 2011. Thirty-day mortality of *Clostridium difficile* infection in a UK National Health Service Foundation Trust between 2002 and 2008. *Journal of Hospital Infection,* 77**,** 11-15.

McGrath, S. & van Sinderen, D. 2007. *Bacteriophage: genetics and molecular biology*: Horizon Scientific Press.

Meader, E., Mayer, M. J., Gasson, M. J., Steverding, D., Carding, S. R. & Narbad, A. 2010. Bacteriophage treatment significantly reduces viable *Clostridium difficile* and prevents toxin production in an in vitro model system. *Anaerobe,* 16**,** 549-54.

Meader, E., Mayer, M. J., Steverding, D., Carding, S. R. & Narbad, A. 2013. Evaluation of bacteriophage therapy to control *Clostridium difficile* and toxin production in an in vitro human colon model system. *Anaerobe,* 22**,** 25-30.

Meessen-Pinard, M., Sekulovic, O. & Fortier, L.-C. 2012. Evidence of in vivo prophage induction during *Clostridium difficile* infection. *Appl. Environ. Microbiol.,* 78**,** 7662-7670.

Meir, M., Grosfeld, T. & Barkan, D. 2018. Establishment and validation of *Galleria mellonella* as a novel model organism to study *Mycobacterium abscessus* infection, pathogenesis, and treatment. *Antimicrobial agents and chemotherapy,* 62**,** e02539-17.

Mele, M., Ferreira, P., Reverter, F., DeLuca, D., Monlong, J., Sammeth, M., Young, T., Goldmann, J., Pervouchine, D. & Sullivan, T. 2015. Human genomics. The human transcriptome across tissues and individuals.. 2015;(6235): 660-5. *Science,* 348.

Merabishvili, M., Pirnay, J.-P., Verbeken, G., Chanishvili, N., Tediashvili, M., Lashkhi, N., Glonti, T., Krylov, V., Mast, J. & Van Parys, L. 2009. Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *PloS one,* 4**,** e4944.

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

Metcalf, D., Sharif, S. & Weese, J. S. 2010. Evaluation of candidate reference genes in *Clostridium difficile* for gene expression normalization. *Anaerobe,* 16**,** 439-43.

Metzker, M. L. 2010. Sequencing technologies—the next generation. *Nature reviews genetics,* 11**,** 31.

Miles, A. A., Misra, S. & Irwin, J. 1938. The estimation of the bactericidal power of the blood. *Epidemiology & Infection,* 38**,** 732-749.

Millan, B., Park, H., Hotte, N., Mathieu, O., Burguiere, P., Tompkins, T. A., Kao, D. & Madsen, K. L. 2016. Fecal microbial transplants reduce antibiotic-resistant genes in patients with recurrent *Clostridium difficile* infection. *Clinical Infectious Diseases,* 62**,** 1479-1486.

Miyajima, F., Roberts, P., Swale, A., Price, V., Jones, M., Horan, M., Beeching, N., Brazier, J., Parry, C. & Pendleton, N. 2011. Characterisation and carriage ratio of *Clostridium difficile* strains isolated from a community-dwelling elderly population in the United Kingdom. *PloS one,* 6**,** e22804.

Mojardin, L. & Salas, M. 2016. Global Transcriptional Analysis of Virus-Host Interactions between Phage varphi29 and *Bacillus subtilis*. *J Virol,* 90**,** 9293-304.

Moncrief, J. S., Barroso, L. A. & Wilkins, T. D. 1997. Positive regulation of *Clostridium difficile* toxins. *Infection and immunity,* 65**,** 1105-1108.

Monot, M., Boursaux-Eude, C., Thibonnier, M., Vallenet, D., Moszer, I., Medigue, C., Martin-Verstraete, I. & Dupuy, B. 2011. Reannotation of the genome sequence of *Clostridium difficile* strain 630. *Journal of medical microbiology,* 60**,** 1193-9.

Monot, M., Eckert, C., Lemire, A., Hamiot, A., Dubois, T., Tessier, C., Dumoulard, B., Hamel, B., Petit, A. & Lalande, V. 2015. *Clostridium difficile*: new insights into the evolution of the pathogenicity locus. *Scientific reports,* 5**,** 15023.

Monteiro, R., Pires, D. P., Costa, A. R. & Azeredo, J. 2019. Phage therapy: going temperate? *Trends in microbiology,* 27**,** 368-378.

Montfort-Gardeazabal, J. M. 2017. *Investigation of the effect of insertional mutation in agr and sigB loci in Clostridium difficile R20291*. University of Nottingham.

Morin, R. D., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T. J., McDonald, H., Varhol, R., Jones, S. J. & Marra, M. A. 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques,* 45**,** 81-94.

Mowlds, P., Coates, C., Renwick, J. & Kavanagh, K. 2010. Dose-dependent cellular and humoral responses in *Galleria mellonella* larvae following β-glucan inoculation. *Microbes and infection,* 12**,** 146-153.

Mukherjee, K., Altincicek, B., Hain, T., Domann, E., Vilcinskas, A. & Chakraborty, T. 2010. *Galleria mellonella* as a model system for studying Listeria pathogenesis. *Appl. Environ. Microbiol.,* 76**,** 310-317.

Mukherjee, K., Hain, T., Fischer, R., Chakraborty, T. & Vilcinskas, A. 2013. Brain infection and activation of neuronal repair mechanisms by the human pathogen *Listeria monocytogenes* in the lepidopteran model host *Galleria mellonella*. *Virulence,* 4**,** 324-332.

Murphy, F. A., Fauquet, C. M., Bishop, D. H., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. 2012. *Virus taxonomy: classification and nomenclature of viruses*: Springer Science & Business Media.

Mutz, K.-O., Heilkenbrinker, A., Lönne, M., Walter, J.-G. & Stahl, F. 2013. Transcriptome analysis using next-generation sequencing. *Current opinion in biotechnology,* 24**,** 22-30.

Mylonakis, E., Moreno, R., El Khoury, J. B., Idnurm, A., Heitman, J., Calderwood, S. B., Ausubel, F. M. & Diener, A. 2005. *Galleria mellonella* as a model system to study *Cryptococcus neoformans* pathogenesis. *Infection and immunity,* 73**,** 3842-3850.

Naaber, P., Smidt, I., Štšepetova, J., Brilene, T., Annuk, H. & Mikelsaar, M. 2004. Inhibition of *Clostridium difficile* strains by intestinal *Lactobacillus* species. *Journal of medical microbiology,* 53**,** 551-554.

Nakayama, J., Cao, Y., Horii, T., Sakuda, S., Akkermans, A. D., De Vos, W. M. & Nagasawa, H. 2001. Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. *Molecular microbiology,* 41**,** 145-154.

Nale, J. 2013. *Isolation and Characterisation of Temperate Bacteriophages of the Hypervirulent Clostridium Difficile 027 Strains.* PhD. Thesis, University of Leicester.

Nale, J. Y., Shan, J., Hickenbotham, P. T., Fawley, W. N., Wilcox, M. H. & Clokie, M. R. 2012. Diverse temperate bacteriophage carriage in *Clostridium difficile* 027 strains. *PLoS One,* 7**,** e37263.

Nale, J. Y., Spencer, J., Hargreaves, K. R., Buckley, A. M., Trzepinski, P., Douce, G. R. & Clokie, M. R. J. 2016. Bacteriophage Combinations Significantly Reduce *Clostridium difficile* Growth *In Vitro* and Proliferation *In Vivo*. *Antimicrobial Agents and Chemotherapy,* 60**,** 968-981.

Nelson, N. J. 2001. Microarrays have arrived: gene expression tool matures. *Journal of the National Cancer Institute,* 93**,** 492-494.

Nicot, N., Hausman, J.-F., Hoffmann, L. & Evers, D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of experimental botany,* 56**,** 2907-2914.

Nilsson, A. S. 2014. Phage therapy--constraints and possibilities. *Ups J Med Sci,* 119**,** 192-8.

Norman, K., Harvey, R., Scott, H., Hume, M., Andrews, K. & Brawley, A. 2009. Varied prevalence of *Clostridium difficile* in an integrated swine operation. *Anaerobe,* 15**,** 256-260.

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

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. 1984. Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea. *Science,* 226**,** 694-696.

O'Connor, J. R., Johnson, S. & Gerding, D. N. 2009. *Clostridium difficile* infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology,* 136**,** 1913-1924.

O'Neil, H. S. & Marquis, H. 2006. *Listeria monocytogenes* flagella are used for motility, not as adhesins, to increase host cell invasion. *Infect Immun,* 74**,** 6675-81.

O'Sullivan, L., Bolton, D., McAuliffe, O. & Coffey, A. 2019. Bacteriophages in food applications: From foe to friend. *Annual review of food science and technology,* 10**,** 151-172.

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

Olsen, R. J., Watkins, M. E., Cantu, C. C., Beres, S. B. & Musser, J. M. 2011. Virulence of serotype M3 Group A *Streptococcus* strains in wax worms (*Galleria mellonella* larvae). *Virulence,* 2**,** 111-119.

Onderdonk, A. B., Johnston, J., Mayhew, J. W. & Gorbach, S. L. 1976. Effect of dissolved oxygen and Eh and *Bacteroides fragilis* during continuous culture. *Applied and environmental microbiology,* 31**,** 168-172.

Onderdonk, A. B., Kasper, D. L., Cisneros, R. L. & Bartlett, J. G. 1977. The capsular polysaccharide of *Bacteroides fragilis* as a virulence factor: comparison of the pathogenic potential of encapsulated and unencapsulated strains. *Journal of Infectious Diseases,* 136**,** 82-89.

Orlova, E. 2012. Bacteriophages and their structural organisation. *Bacteriophages.* InTech.

Ormonde, P. & Rstedt, P. O. T., R. 2000. Role of Motility in Adherence to and Invasion of a Fish Cell Line by *Vibrio anguillarum*. *JOURNAL OF BACTERIOLOGY,* 182.

Osmundson, J., Dewell, S. & Darst, S. A. 2013. RNA-Seq reveals differential gene expression in *Staphylococcus aureus* with single-nucleotide resolution. *PloS one,* 8**,** e76572.

Ozsolak, F. & Milos, P. M. 2011. RNA sequencing: advances, challenges and opportunities. *Nature reviews genetics,* 12**,** 87.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics,* 40**,** 1413.

Parasion, S., Kwiatek, M., Gryko, R., Mizak, L. & Malm, A. 2014. Bacteriophages as an alternative strategy for fighting biofilm development. *Pol. J. Microbiol,* 63**,** 137-145.

Paredes, S., D., Cofre-Araneda, G., Brito-Silva, C., Pizarro-Guajardo, M. & Sarker, M. R. 2012. *Clostridium difficile* Spore-Macrophage Interactions: Spore Survival. *PLoS One,* 7.

Park, Y.-J., Park, C. E., Hong, S.-J., Jung, B. K., Ibal, J. C., Park, G.-S. & Shin, J.-H. 2017. The complete mitochondrial genome sequence of the greater wax moth *Galleria mellonella* (Insecta, Lepidoptera, Pyralidae): sequence and phylogenetic analysis comparison based on whole mitogenome. *Mitochondrial DNA Part B,* 2**,** 714-715.

Patel, M. R. 2011. *Characterisation of a putative quorum sensing system in Colstridium difficile.* University of Nottingham.

Paul, V. D., Sundarrajan, S., Rajagopalan, S. S., Hariharan, S., Kempashanaiah, N., Padmanabhan, S., Sriram, B. & Ramachandran, J. 2011. Lysis-deficient phages as novel therapeutic agents for controlling bacterial infection. *BMC microbiology,* 11**,** 195.

Pepin, J., Valiquette, L., Alary, M.-E., Villemure, P., Pelletier, A., Forget, K., Pépin, K. & Chouinard, D. 2004. *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *Cmaj,* 171**,** 466-472.

Pereira, T., de Barros, P., Fugisaki, L., Rossoni, R., Ribeiro, F., de Menezes, R., Junqueira, J. & Scorzoni, L. 2018. Recent advances in the use of *galleria mellonella* model to study immune responses against human pathogens. *Journal of Fungi,* 4**,** 128.

Perelle, S., Gibert, M., Bourlioux, P., Corthier, G. & Popoff, M. R. 1997. Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. *Infection and immunity,* 65**,** 1402-1407.

Peterfreund, G. L., Vandivier, L. E., Sinha, R., Marozsan, A. J., Olson, W. C., Zhu, J. & Bushman, F. D. 2012. Succession in the gut microbiome following antibiotic and antibody therapies for *Clostridium difficile*. *PLoS One,* 7**,** e46966.

Pham, T. A. N., Clare, S., Goulding, D., Arasteh, J. M., Stares, M. D., Browne, H. P., Keane, J. A., Page, A. J., Kumasaka, N. & Kane, L. 2014. Epithelial IL-22RA1-mediated fucosylation promotes intestinal colonization resistance to an opportunistic pathogen. *Cell host & microbe,* 16**,** 504-516.

Phothichaisri, W., Ounjai, P., Phetruen, T., Janvilisri, T., Khunrae, P., Singhakaew, S., Wangroongsarb, P. & Chankhamhaengdecha, S. 2018. Characterization of Bacteriophages Infecting Clinical Isolates of *Clostridium difficile*. *Front Microbiol,* 9**,** 1701.

Pietu, G., Mariage-Samson, R., Fayein, N.-A., Matingou, C., Eveno, E., Houlgatte, R., Decraene, C., Vandenbrouck, Y., Tahi, F. & Devignes, M.-D. 1999. The Genexpress IMAGE knowledge base of the human brain transcriptome: a prototype integrated resource for functional and computational genomics. *Genome research,* 9**,** 195-209.

Pires, D. P., Vilas Boas, D., Sillankorva, S. & Azeredo, J. 2015. Phage Therapy: a Step Forward in the Treatment of *Pseudomonas aeruginosa* Infections. *J Virol,* 89**,** 7449-56.

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

Poranen, M. M., Ravantti, J. J., Grahn, A. M., Gupta, R., Auvinen, P. & Bamford, D. H. 2006. Global changes in cellular gene expression during bacteriophage PRD1 infection. *J Virol,* 80**,** 8081-8.

Pothoulakis, C. & Lamont, J. T. 2001. Microbes and microbial toxins: paradigms for microbial-mucosal interactions II. The integrated response of the intestine to *Clostridium difficile* toxins. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* 280**,** G178-G183.

Potuckova, L., Kelemen, G. H., Findlay, K. C., Lonetto, M. A., Buttner, M. J. & Kormanec, J. 1995. A new RNA polymerase sigma factor, σF is required for the late stages of morphological differentiation in *Streptomyces* spp. *Molecular microbiology,* 17**,** 37-48.

Pribnow, D. 1975. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences,* 72**,** 784-788.

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

Raffestin, S., Dupuy, B., Marvaud, J. C. & Popoff, M. R. 2005. BotR/A and TetR are alternative RNA polymerase sigma factors controlling the expression of the neurotoxin and associated protein genes in *Clostridium botulinum* type A and *Clostridium tetani*. *Molecular microbiology,* 55**,** 235-249.

Ramarao, N., Nielsen-Leroux, C. & Lereclus, D. 2012. The insect *Galleria mellonella* as a powerful infection model to investigate bacterial pathogenesis. *Journal of visualized experiments: JoVE*.

Ramirez-Vargas, G., Goh, S. & Rodriguez, C. 2018. The novel phages phiCD5763 and phiCD2955 represent two groups of big plasmidial siphoviridae phages of *Clostridium difficile*. *Frontiers in microbiology,* 9**,** 26.

Rankin, D. J., Rocha, E. P. & Brown, S. P. 2011. What traits are carried on mobile genetic elements, and why? *Heredity,* 106**,** 1.

Rao, M. S., Van Vleet, T. R., Ciurlionis, R., Buck, W. R., Mittelstadt, S. W., Blomme, E. A. & Liguori, M. J. 2018. Comparison of RNA-seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. *Frontiers in genetics,* 9**,** 636.

Ravantti, J. J., Ruokoranta, T. M., Alapuranen, A. M. & Bamford, D. H. 2008. Global transcriptional responses of *Pseudomonas aeruginosa* to phage PRR1 infection. *J Virol,* 82**,** 2324-9.

Refardt, D. 2012. Real-time quantitative PCR to discriminate and quantify lambdoid bacteriophages of *Escherichia coli* K-12. *Bacteriophage,* 2**,** 98-104.

Reidl, J. & Mekalanos, J. J. 1995. Characterization of Vibrio cholerae bacteriophage K139 and use of a novel mini-transposon to identify a phage-encoded virulence factor. *Molecular microbiology,* 18**,** 685-701.

Rejasse, A., Gilois, N., Barbosa, I., Huillet, E., Bevilacqua, C., Tran, S., Ramarao, N., Arnesen, L. S. & Sanchis, V. 2012. Temperature-dependent production of various PlcR-controlled virulence factors in *Bacillus weihenstephanensis* strain KBAB4. *Applied and Environmental Microbiology,* AEM. 07446-11.

Riggs, M. M., Sethi, A. K., Zabarsky, T. F., Eckstein, E. C., Jump, R. L. & Donskey, C. J. 2007. Asymptomatic carriers are a potential source for transmission of epidemic and nonepidemic *Clostridium difficile* strains among long-term care facility residents. *Clinical infectious diseases,* 45**,** 992-998.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-140.

Rocha, D. J., Santos, C. S. & Pacheco, L. G. 2015. Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. *Antonie Van Leeuwenhoek,* 108**,** 685-93.

Rosenwasser, S., Ziv, C., Van Creveld, S. G. & Vardi, A. 2016. Virocell metabolism: metabolic innovations during host–virus interactions in the ocean. *Trends in microbiology,* 24**,** 821-832.

Rupnik, M., Brazier, J. S., Duerden, B. I., Grabnar, M. & Stubbs, S. L. 2001. Comparison of toxinotyping and PCR ribotyping of *Clostridium difficile* strains and description of novel toxinotypes. *Microbiology,* 147**,** 439-447.

Rupnik, M., Wilcox, M. H. & Gerding, D. N. 2009. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nature Reviews Microbiology,* 7**,** 526-536.

Russel, M. 1995. Moving through the membrane with filamentous phages. *Trends in microbiology,* 3**,** 223-228.

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

Rydman, P. S. & Bamford, D. H. 2003. Identification and mutational analysis of bacteriophage PRD1 holin protein P35. *Journal of bacteriology,* 185**,** 3795-3803.

Sakaguchi, Y., Hayashi,T., Kurokawa, Ken., Nakayama, K., Oshima, K.,, Fujinaga, Y., Ohnishi, M. & Ohtsubo, E., Hattori, M. and Oguma, K. 2005. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Roy Curtiss, Arizona State University,* 102**,** 17472–17477.

Salamitou, S., Ramisse, F. o., Brehélin, M., Bourguet, D., Gilois, N., Gominet, M., Hernandez, E. & Lereclus, D. 2000. The plcR regulon is involved in the

opportunistic properties of *Bacillus thuringiensis* and *Bacillus cereus* in mice and insects. *Microbiology,* 146**,** 2825-2832.

Salmond, G. P. & Fineran, P. C. 2015. A century of the phage: past, present and future. *Nature Reviews Microbiology,* 13**,** 777.

Sára, M. & Sleytr, U. B. 2000. S-layer proteins. *Journal of bacteriology,* 182**,** 859-868.

Scaria, J., Mao, C., Chen, J.-W., McDonough, S. P., Sobral, B. & Chang, Y. 2013. Differential stress transcriptome landscape of historic and recently emerged hypervirulent strains of *Clostridium difficile* strains determined using RNA-seq. *PloS one,* 8**,** e78489.

Schachtele, C. F., De Sain, C. V., Hawley, L. A. & Anderson, D. L. 1972. Transcription during the development of bacteriophage φ29: production of host-and φ29-specific ribonucleic acid. *Journal of virology,* 10**,** 1170-1178.

Schaller, H., Gray, C. & Herrmann, K. 1975. Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. *Proceedings of the National Academy of Sciences,* 72**,** 737-741.

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

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

Sekulovic, O. & Fortier, L. C. 2015. Global transcriptional response of *Clostridium difficile* carrying the CD38 prophage. *Appl Environ Microbiol,* 81**,** 1364-74.

Sekulovic, O., Meessen-Pinard, M. & Fortier, L. C. 2011. Prophage-stimulated toxin production in *Clostridium difficile* NAP1/027 lysogens. *J Bacteriol,* 193**,** 2726-34.

Sell, T. L., Schaberg, D. R. & Fekety, F. R. 1983. Bacteriophage and bacteriocin typing scheme for *Clostridium difficile*. *Journal of clinical microbiology,* 17**,** 1148-1152.

Shan, J., Ramachandran, A., Thanki, A. M., Vukusic, F. B., Barylski, J. & Clokie, M. R. 2018. Bacteriophages are more virulent to bacteria with human cells than they are in bacterial culture; insights from HT-29 cells. *Scientific reports,* 8**,** 5091.

Sharkey, F. H., Banat, I. M. & Marchant, R. 2004. Detection and quantification of gene expression in environmental bacteriology. *Appl. Environ. Microbiol.,* 70**,** 3795-3806.

Sharma, C. M. & Vogel, J. 2014. Differential RNA-seq: the approach behind and the biological insight gained. *Curr Opin Microbiol,* 19**,** 97-105.

Shen, X., Li, M., Zeng, Y., Hu, X., Tan, Y., Rao, X., Jin, X., Li, S., Zhu, J. & Zhang, K. 2012. Functional identification of the DNA packaging terminase from *Pseudomonas aeruginosa* phage PaP3. *Archives of virology,* 157**,** 2133-2141.

Simango, C. 2006. Prevalence of *Clostridium difficile* in the environment in a rural community in Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 100**,** 1146-1150.

Skurnik, M. & Strauch, E. 2006. Phage therapy: facts and fiction. *International journal of medical microbiology,* 296**,** 5-14.

Slater, J., Gregson, L., Denning, D. & Warn, P. 2011. Pathogenicity of *Aspergillus fumigatus* mutants assessed in *Galleria mellonella* matches that in mice. *Medical mycology,* 49**,** S107-S113.

Slonczewski, J., Foster, J. W. & Gillen, K. M. 2013. *Microbiology: an evolving science*: John Wiley & Sons.

Slonczewski, J. L. & Foster, J. W. 2013. *Microbiology: An Evolving Science: Third International Student Edition*: WW Norton & Company.

Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. 2016. *Clostridium difficile* infection. *Nature reviews Disease primers,* 2**,** 16020.

Soengas, M. a. S., Gutierrez, C. & Salas, M. 1995. Helix-destabilizing activity of φ29 single-stranded DNA binding protein: effect on the elongation rate during strand displacement DNA replication. *Journal of molecular biology,* 253**,** 517-529.

Sorbara, M. T. & Pamer, E. G. 2018. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. *Mucosal immunology,* 1.

Sorensen, M. C. H., van Alphen, L. B., Harboe, A., Li, J., Christensen, B. B., Szymanski, C. M. & Brøndsted, L. 2011. The F336 bacteriophage recognizes the capsular phosphoramidate modification of Campylobacter jejuni NCTC11168. *Journal of bacteriology,* JB. 05276-11.

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

Spanier, J. G. & Cleary, P. 1980. Bacteriophage control of antiphagocytic determinants in group A streptococci. *Journal of Experimental Medicine,* 152**,** 1393-1406.

Speert, D. 2000. *Clostridium difficile*-Pathogen or pest? *Paediatrics & Child Health,* 5**,** 349-352.

Spinelli, S., Campanacci, V., Blangy, S., Moineau, S., Tegoni, M. & Cambillau, C. 2006. Modular Structure of the Receptor Binding Proteins of *Lactococcus lactis* Phages the RBP Structure of the temperate Phage TP901-1. *Journal of Biological Chemistry,* 281**,** 14256-14262.

Spurbeck, R. R., Stapleton, A. E., Johnson, J. R., Walk, S. T., Hooton, T. M. & Mobley, H. L. 2011. Fimbrial profiles predict virulence of uropathogenic *Escherichia coli* strains: contribution of ygi and yad fimbriae. *Infection and immunity,* 79**,** 4753-4763.

Squires, R. 2018. Bacteriophage therapy for management of bacterial infections in veterinary practice: what was once old is new again. *New Zealand veterinary journal,* 66**,** 229-235.

Stabler, He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., Lawley, T. D., Sebaihia, M., Quail, M. A. & Rose, G. 2009a. Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome biology,* 10**,** R102.

Stabler, R. A., Dawson, L. F., Valiente, E., Cairns, M. D., Martin, M. J., Donahue, E. H., Riley, T. V., Songer, J. G., Kuijper, E. J., Dingle, K. E. & Wren, B. W. 2012. Macro and micro diversity of *Clostridium difficile* isolates from diverse sources and geographical locations. *PLoS One,* 7**,** e31559.

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

Steffen, E. K. & Hentges, D. 1981. Hydrolytic enzymes of anaerobic bacteria isolated from human infections. *Journal of clinical microbiology,* 14**,** 153-156.

Stevenson, E., Minton, N. P. & Kuehne, S. A. 2015. The role of flagella in Clostridium difficile pathogenicity. *Trends in microbiology,* 23**,** 275-282.

Stewart, G. R., Wernisch, L., Stabler, R., Mangan, J. A., Hinds, J., Laing, K. G., Young, D. B. & Butcher, P. D. 2002. Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarraysa. *Microbiology,* 148**,** 3129-3138.

Stiles, B. G., Barth, H. & Popoff, M. R. 2016. *Clostridium perfringens* iota toxin: a successfully shared template for common enteric pathogens. *Microbial Toxins,* 1-20.

Strelau, E., Wagner, B., Wagner, M. & Karsch, W. 1989. Demonstration of capsules in *Clostridium difficile*. *Zentralbl Bakteriol Mikrobiol Hyg A,* 270**,** 456-61.

Stuart Ainsworth, Aldert Zomer, Jennifer Mahony, Douwe van Sinderena & 2013. Lytic Infection of *Lactococcus lactis* by Bacteriophages Tuc2009 and c2

Triggers Alternative Transcriptional Host Responses. *Applied and Environmental Microbiology* 79 4786–4798.

Sulakvelidze, A., Alavidze, Z. & Morris, J. G., Jr. 2001. Bacteriophage therapy. *Antimicrob Agents Chemother,* 45**,** 649-59.

Sullivan, M. B., Weitz, J. S. & Wilhelm, S. 2017. Viral ecology comes of age. *Environmental microbiology reports,* 9**,** 33-35.

Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A. & Parkhomchuk, D. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science,* 321**,** 956-960.

Sundriyal, A., Roberts, A. K., Shone, C. C. & Acharya, K. R. 2009. Structural basis for substrate recognition in the enzymatic component of ADP-ribosyltransferase toxin CDTa from Clostridium difficile. *J Biol Chem,* 284**,** 28713-9.

Surawicz, C. M., Brandt, L. J., Binion, D. G., Ananthakrishnan, A. N., Curry, S. R., Gilligan, P. H., McFarland, L. V., Mellow, M. & Zuckerbraun, B. S. 2013. Guidelines for diagnosis, treatment, and prevention of *Clostridium difficile* infections. *Am J Gastroenterol,* 108**,** 478-98; quiz 499.

Suttle, C. A. 2007. Marine viruses—major players in the global ecosystem. *Nature reviews microbiology,* 5**,** 801.

Takeda, Y., Folkmanis, A. & Echols, H. 1977. Cro regulatory protein specified by bacteriophage lambda. Structure, DNA-binding, and repression of RNA synthesis. *Journal of Biological Chemistry,* 252**,** 6177-6183.

Tan, K. S., Wee, B. Y. & Song, K. P. 2001. Evidence for holin function of tcdE gene in the pathogenicity of *Clostridium difficile*. *Journal of medical microbiology,* 50**,** 613-619.

Taslim, H. & Taslim, H. 2009. *Clostridium difficile* infection in the elderly. *Acta Med Indones,* 41**,** 148-51.

Tasteyre, A., Barc, M. C., Collignon, A., Boureau, H. & Karjalainen, T. 2001. Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. *Infect Immun,* 69**,** 7937-40.

Tedin, K. & Blasi, U. 1996. The RNA chain elongation rate of the λ late mRNA is unaffected by high levels of ppGpp in the absence of amino acid starvation. *Journal of Biological Chemistry,* 271**,** 17675-17686.

Thanki, A. M. 2016. *Development of a phage-based diagnostic test for the identification of Clostridium difficile.* PhD thesis, Loughborough University.

Thelestam, M. & Chaves-Olarte, E. 2000. Cytotoxic effects of the *Clostridium difficile* toxins. *Clostridium difficile.* Springer.

Theriot, C. M., Koenigsknecht, M. J., Carlson Jr, P. E., Hatton, G. E., Nelson, A. M., Li, B., Huffnagle, G. B., Li, J. Z. & Young, V. B. 2014. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nature communications,* 5**,** 3114.

Thomas, R. J., Hamblin, K. A., Armstrong, S. J., Müller, C. M., Bokori-Brown, M., Goldman, S., Atkins, H. S. & Titball, R. W. 2013. *Galleria mellonella* as a model system to test the pharmacokinetics and efficacy of antibiotics against *Burkholderia pseudomallei*. *International journal of antimicrobial agents,* 41**,** 330-336.

Thornton, B. & Basu, C. 2011. *Real-Time PCR (qPCR) Primer Design Using Free Online Software*: The International Union of Biochemistry and Molecular Biology.

Trevijano & Zaragoza, O. 2019. Immune response of *Galleria mellonella* against human fungal pathogens. *Journal of Fungi,* 5**,** 3.

Tsai, C. J.-Y., Loh, J. M. S. & Proft, T. 2016. *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. *Virulence,* 7**,** 214-229.

Tuohy, K. M., Probert, H. M., Smejkal, C. W. & Gibson, G. R. 2003. Using probiotics and prebiotics to improve gut health. *Drug discovery today,* 8**,** 692-700.

Twort, F. W. 1915. An investigation on the nature of ultra-microscopic viruses. *The Lancet,* 186**,** 1241-1243.

Ueno, H. & Yonesaki, T. 2004. Phage-induced change in the stability of mRNAs. *Virology,* 329**,** 134-141.

Vaishnavi, C. 2010. Diagnostic approach to *Clostridium difficile* infection. Springer.

Valiente, E., Bouche, L., Hitchen, P., Faulds-Pain, A., Songane, M., Dawson, L. F., Donahue, E., Stabler, R. A., Panico, M. & Morris, H. R. 2016. Role of glycosyltransferases modifying type B flagellin of emerging hypervirulent *Clostridium difficile* lineages and their impact on motility and biofilm formation. *Journal of Biological Chemistry,* 291**,** 25450-25461.

Vandecasteele, S., Peetermans, W., Merckx, R. & Van Eldere, J. 2001. Quantification of expression of *Staphylococcus epidermidis* housekeeping genes with Taqman quantitative PCR during in vitro growth and under different conditions. *Journal of bacteriology,* 183**,** 7094-7101.

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett Jr, D. E., Hieter, P., Vogelstein, B. & Kinzler, K. W. 1997. Characterization of the yeast transcriptome. *Cell,* 88**,** 243-251.

Viertel, T. M., Ritter, K. & Horz, H.-P. 2014. Viruses versus bacteria—novel approaches to phage therapy as a tool against multidrug-resistant pathogens. *Journal of Antimicrobial Chemotherapy,* 69**,** 2326-2336.

Vilcinskas, A. 2011. Anti-infective therapeutics from the Lepidopteran model host *Galleria mellonella*. *Current pharmaceutical design,* 17**,** 1240-1245.

Villarroel, J., Nielsen, M., Larsen, M. V. & Kilstrup, M. 2018. *Isolation and characterization of bacteriophages with therapeutic potential.* Technical University of Denmark (DTU).

Viswanathan, V., Mallozzi, M. & Vedantam, G. 2010. *Clostridium difficile* infection: An overview of the disease and its pathogenesis, epidemiology and interventions. *Gut microbes,* 1**,** 234-242.

Vogel, H., Altincicek, B., Glöckner, G. & Vilcinskas, A. 2011. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. *BMC genomics,* 12**,** 308.

Vohra, P. & Poxton, I. R. 2011. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. *Microbiology,* 157**,** 1343-1353.

Voth, D. E. & Ballard, J. D. 2005. *Clostridium difficile* toxins: mechanism of action and role in disease. *Clinical microbiology reviews,* 18**,** 247-263.

Vu, K. & Gelli, A. 2010. Astemizole and an analogue promote fungicidal activity of fluconazole against *Cryptococcus neoformans* var. grubii and *Cryptococcus gattii*. *Medical mycology,* 48**,** 255-262.

Wagner, P. L., Neely, M. N., Zhang, X., Acheson, D. W., Waldor, M. K. & Friedman, D. I. 2001. Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *Journal of bacteriology,* 183**,** 2081-2085.

Wagner, P. L. & Waldor, M. K. 2002. Bacteriophage control of bacterial virulence. *Infection and immunity,* 70**,** 3985-3993.

Waldor, M. K. & Mekalanos, J. J. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science,* 272**,** 1910-1914.

Waligora, A. J., Barc, M. C., Bourlioux, P., Collignon, A. & Karjalainen, T. 1999. *Clostridium difficile* cell attachment is modified by environmental factors. *Appl Environ Microbiol,* 65**,** 4234-8.

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

Walters, J. B. & Ratcliffe, N. 1983. Studies on the in vivo cellular reactions of insects: fate of pathogenic and non-pathogenic bacteria in *Galleria mellonella* nodules. *Journal of Insect Physiology,* 29**,** 417-424.

Wand, M. E., McCowen, J. W., Nugent, P. G. & Sutton, J. M. 2013. Complex interactions of *Klebsiella pneumoniae* with the host immune system in a *Galleria mellonella* infection model. *Journal of medical microbiology,* 62**,** 1790-1798.

Wang, I.-N., Smith, D. L. & Young, R. 2000. Holins: the protein clocks of bacteriophage infections. *Annual Reviews in Microbiology,* 54**,** 799-825.

Wang, J., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, J., Bao, Q., Kong, H. & Hu, S. 2005. Complete genome sequence of bacteriophage T5. *Virology,* 332**,** 45-65.

Wang, X. 2016. NGS Data Analysis Using Partek® Software. Partek.

Wang, Z., Gerstein, M. & Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics,* 10**,** 57.

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

Wassenaar, T. M. & Gaastra, W. 2001. Bacterial virulence: can we draw the line? *FEMS microbiology letters,* 201**,** 1-7.

Wayoff, M. & Jankowski, R. 1989. Virulence factors of *Haemophilus influenzae*. *Journal of chemotherapy (Florence, Italy),* 1**,** 26.

Weinbauer, M. G. 2004. Ecology of prokaryotic viruses. *FEMS microbiology reviews,* 28**,** 127-181.

Weld, R. J., Butts, C. & Heinemann, J. A. 2004. Models of phage growth and their applicability to phage therapy. *Journal of theoretical biology,* 227**,** 1-11.

White, R., Chiba, S., Pang, T., Dewey, J. S., Savva, C. G., Holzenburg, A., Pogliano, K. & Young, R. 2011. Holin triggering in real time. *Proceedings of the National Academy of Sciences,* 108**,** 798-803.

Wiegand, P., Nathwani, D., Wilcox, M., Stephens, J., Shelbaya, A. & Haider, S. 2012. Clinical and economic burden of *Clostridium difficile* infection in Europe: a systematic review of healthcare-facility-acquired infection. *Journal of Hospital Infection,* 81**,** 1-14.

Willey, J. 2014. Prescott's microbiology-/Joanne M. *Willey, Linda M. Sherwood, Christopher J. Woolverton. MacGraw-Hill, New York.*

Yamakawa, K., Karasawa, T., Ikoma, S. & Nakamura, S. 1996. Enhancement of *Clostridium difficile* toxin production in biotin-limited conditions. *Journal of medical microbiology,* 44**,** 111-114.

Yordpratum, U., Tattawasart, U., Wongratanacheewin, S. & Sermswan, R. W. 2011. Novel lytic bacteriophages from soil that lyse *Burkholderia pseudomallei*. *FEMS Microbiol Lett,* 314**,** 81-8.

Young, R. 1992. Bacteriophage lysis: mechanism and regulation. *Microbiology and Molecular Biology Reviews,* 56**,** 430-481.

Young, R. 2013. Phage lysis: do we have the hole story yet? *Current opinion in microbiology,* 16**,** 790-797.

Young, R., Wang, N. & Roof, W. D. 2000. Phages will out: strategies of host cell lysis. *Trends in microbiology,* 8**,** 120-128.

Zar, F. A., Bakkanagari, S. R., Moorthi, K. M. & Davis, M. B. 2007. A comparison of vancomycin and metronidazole for the treatment of *Clostridium difficile*-associated diarrhea, stratified by disease severity. *Clin Infect Dis,* 45**,** 302-7.

Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H. & O'Toole, G. A. 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *Journal of bacteriology,* 191**,** 210-219.

Zhao, H., Christensen, T. E., Kamau, Y. N. & Tang, L. 2013. Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. *Proceedings of the National Academy of Sciences,* 110**,** 8075-8080.

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K. & Liu, X. 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PloS one,* 9**,** e78644.

Zhao, X., Chen, C., Shen, W., Huang, G., Le, S., Lu, S., Li, M., Zhao, Y., Wang, J., Rao, X., Li, G., Shen, M., Guo, K., Yang, Y., Tan, Y. & Hu, F. 2016. Global Transcriptomic Analysis of Interactions between *Pseudomonas aeruginosa* and Bacteriophage PaP3. *Sci Rep,* 6**,** 19237.

Ziedaite, G., Daugelavicius, R., Bamford, J. K. & Bamford, D. H. 2005. The Holin protein of bacteriophage PRD1 forms a pore for small-molecule and endolysin translocation. *Journal of bacteriology,* 187**,** 5397-5405.

Zitzmann, J., Weidner, T. & Czermak, P. 2017. Optimized expression of the antimicrobial protein Gloverin from *Galleria mellonella* using stably transformed *Drosophila melanogaster* S2 cells. *Cytotechnology,* 69**,** 371-389.