Characterization of the late promoter pR' from Shiga toxin prophages in *Escherichia coli*

by

Lingxiao Zhang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Food Science and Technology

Department of Agricultural, Food and Nutritional Science

University of Alberta

Abstract

Shiga toxin producing *Escherichia coli* (STEC) strains are pathogens which frequently cause human intestinal diseases. The main virulence factor Shiga toxin (Stx) is encoded by the gene located on the Stx prophage and gene expression is affected by the diversity of Stx prophage genome.

Nineteen candidate STEC strains were selected according to the previous research. PCR was used, and the late promoter regions, designated as *p*R', were amplified from STEC strains. Contig gaps were filled by sequencing, and complete sequences of target *p*R' fragments were obtained. To confirm the diversity of the gene structure of the target *p*R' region, sequences were aligned, and phylogenetic trees were built based on the sequences of target *p*R' and *stx*, respectively. According to the phylogenetic analyses, *p*R' fragments were highly diverse compared to the *stx* subtypes they harbor.

To visualize Stx production, *p*R' fragments were fused with the *rfp* reporter gene DsRed constructed into pUC19. This reporter system was validated by inducing transformants *E. coli* O104:H4 11-3088 $\Delta stx::gfp::amp^r$ (P*rfp*::*chl^r*), *E. coli* O104:H4 11-3088 $\Delta stx::gfp::amp^r$ (P*p*1302::*rfp*::*chl^r*) and *E. coli* O157:H7 CO6CE900 (P*p*1302::*rfp*::*chl^r*) with mitomycin C and observed by microscopy.

Different combinations of *p*R' and target strains were made to determine the effect the late promoter *p*R' region conferred to Stx prophage induction. Not all prophages can be induced under the control of its parent prophage; no matter if they were controlled by the same or different prophage regulations regulatory systems, different *p*R' behaved diversely.

In conclusion, the diversity of late promoter region sequence structure affects the behavior of the Stx prophage induction.

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Dr. Michael Gänzle, for his many valuable suggestions and guidance during my research, and my study at the University of Alberta. I greatly appreciate his exceptional patience and encouragement. Dr. Gänzle is also acknowledged for providing much encouragement and insight toward my master study and research. He spent lots of time reviewing and helping improve this thesis.

I am also extremely grateful to my co-supervisor, Dr. Lynn McMullen, for her invaluable guidance, suggestions, and encouragement throughout my research, which is a great asset to my life. She really gave me a great deal of encouragement during my pregnancy.

I also want to express my thanks to Dr. David Simpson, for his untiring professional assistance through all the design and commissioning work throughout the research. Special thanks to Dr. Ryan Mercer, who gave me a lot of guidance, leading me to microbiology step by step.

I would like to thank Ms Heather Vandertol-Vanier for her help during my experimental work. I also like to thank Dr. Aja Rieger for her guidance during my experimental work at Faculty of Medicine and Dentistry Flow Cytometry Facility.

The development of this dissertation would not have been possible without the continuous love and unconditional support of my family, especially my grandparents. Their encouragement, patience and understanding are the most motivation for me to do the right thing. The love and support provided by my husband and our child are lovingly acknowledged.

 Finally, I would like to thank all my lab mates in lab 2-50 and numerous friends and colleagues.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

DNA	Deoxyribonucleic acid
FUA	Food Microbiology culture collection of University of Alberta
Gb3	Globotriaosylceramide
Gb4	Globotetraosylceramide
GFP	Green fluorescent protein
eGFP	Enhanced green fluorescent protein
HC	Hemorrhagic colitis
HGT	Horizontal gene transfer
HUS	Hemolytic uremic syndrome
LB	Luria-Bertani
PCR	Polymerase chain reaction
RFP	Red fluorescent protein
RNA	Ribonucleic acid
Stx	Shiga toxin
LEE	Locus of enterocyte effacement
A/E	Attaching and effacing (A/E) lesion
EHEC	Enterohemorrhagic *Escherichia coli*
EAEC	Enteroaggregative *Escherichia coli*
FP	Fluorescent protein
bp	Base pair
MMC	Mitomycin C

1   Introduction

1.1   *E. coli* and pathogenic *E. coli*.

Escherichia coli is a Gram-negative bacterium of the family *Enterobacteriaceae*. Typically, *E. coli* colonizes the gastrointestinal tract of an animal after birth, and coexists with its host through its lifetime. *E. coli* in the phylogenetic group B2 VIII clone with an O81 serotype appear to be human host specific, and some B1 strains, which carry the *hly* gene and exist in distinct serotypes, are specific to the animals (Escobar-Paramo et al. 2006; Clermont et al. 2008).

Pathogenic *E. coli* strains possess virulence factors that can cause diseases or death. Usually, the virulence genes are mobilized into different strains and thus frequently create novel combinations of virulence factors and new pathogenic strains of *E. coli*. The typical clinical symptoms caused by pathogenic *E. coli* include: enteric/diarrheal disease, urinary tract infections (UTIs) and sepsis/meningitis (Kesavan et al. 2015). There are six well-described *E. coli* pathogens: enteropathogenic *E. coli* (EPEC), Enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), which is the same as *Shigella*, and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper 1998; van den Beld and Reubsaet 2012).

1.1.1   Shiga toxin-producing *E. coli* (STEC).

Shiga toxin-producing *E. coli* (STEC), which was first described as verocytotoxin-producing *E. coli* (VTEC), are a group of *E. coli* that can cause human diseases from watery or bloody diarrhea to fatal complications, such as hemorrhagic colitis (HC) and hemolytic uremic syndrome (HUS) (Karmali et al. 1985). The presence of shiga toxin genes 1 or 2 (*stx1*, *stx2*), which are often a part of a bacteriophage genome, qualifies an *E. coli* as a STEC strain. STEC have been isolated from fresh products including plant tissues, animals as well as contaminated water source. Domestic ruminants, especially cattle (both meat and dairy), are the main reservoir of STEC strains, and thus infections can be caused by ingestion of food and water, or directly contacting the infected animals or their feces (Naylor, Gally, and Low 2005; Blanco et al. 2004; Berger et al. 2010).

Transmission of STEC is fecal to oral and the infectious dose of STEC is very low. An investigation of beef patty contamination of *E. coli* O157:H7 between November 1992 and February 1993 suggested that the highest level of contamination had only 15 organisms per gram (Tuttle et al. 1999). Another study about an outbreak caused by STEC O111:NM in dry fermented sausage also showed a small dose of only one organism per 10 gram tested samples caused illness (Paton et al. 1996).

1.1.2 Serotypes, O157 and non-O157 STECs.

The earliest sorting scheme for investigating bacteria strains, which is still widely used, is based on the serology method. Serotypes include O (lipopolysaccharide), K (capsular), H (flagellar antigens), whereas serogroups refer to O antigen only (Nataro and Kaper 1998; Kaper, Nataro, and Mobley 2004) . The focus was first placed on EHEC serotype O157:H7, the first serotype associated with severe illness in human, and thus has been investigated in detail for its gene structure and molecular mechanisms. However, non-O157 STEC, such as O104 and O26, also contribute to human illness (Croxen et al. 2013). Previous phylogenetic analyses showed that, other than having a homologous relationship with other STEC strains, the distribution of *E. coli* O104 more closely resembles commensal strains (Hao et al. 2012). Moreover, STEC strains with different serotypes are randomly located in different phylogenetic groups. This indicates that the main cause for obtaining or losing the *stx* gene is not because of bacterial strain evolution, but because of the sporadic gene transfer and recombination.

1.1.3 Epidemiology.

As a public health problem, STEC has been focused on throughout the world. The program According to the report provided by FoodNet, 465 STEC O157:H7 and 807 non-O157 cases happened in 2015 in the United States, which means there was a per capita rate of 0.95/100,000 and 1.65/100,000 acquired illnesses in the United States, respectively (CDC 2017). Hospitalization rates due to these cases for STEC O157:H7 (38.5%) was almost 2.5-fold higher than that for STEC non-O157 (15.5%). The case-fatality rate for O157:H7 was 0.65%, while it was 0.12% for non-O157 STEC. Although incidence of O157 in 1996-1998, 2006-2008, 2011-2014 has dropped 44%, 30%, 15%, respectively, when compared with 2015, the incidence rate of non-O157 has increased 41% when compared with 2011-2014. The incidence rate in Canada, according to the annual summary reported by the National Enteric Surveillance Program (NESP) in 2015, was 472 cases (1.33/100,000 population) for O157:H7, a slight decline compared to earlier years. However, the incidence rate for non-O157 was 131 cases (0.5/100,000) in 2013, which was an increase from 0.12 cases/100,000 in 2008 (NESP 2015). It should be noted that the non-O157 incidence rates increased in both countries in the past several years. Children, especially at the age less than five, are vulnerable to HUS infection by STEC strains. In the developing country Argentina, which has the highest worldwide incidence of HUS, STEC infections in children may be due to undercooked beef, and exposure to farm animals and their environment.

1.1.4   Enterohemorrhagic *E. coli* (EHEC).

Pathogenic *E. coli* have a wide range of tissue habitats for colonization. The colonization factors include: plasmids, transposons, pathogenicity islands, and mobile elements from lysogenic bacteriophage (Croxen et al. 2013). EHEC is a subpopulation of STEC, and was used to describe the pathogen that can cause HC and HUS. Typically, EHEC strains refers to the LEE-positive strains which have a locus of enterocyte effacement (LEE). The colonization strategy of EHEC is to provoke a histopathologic effect on the intestinal epithelial cells, known as attaching and effacing (A/E) lesion, which is the typical colonization behavior in EPEC strains (McDaniel et al. 1995). The gene *eaeA* (*E. coli* attaching and effacing), which is from the LEE island encodes an adhesin that contributes to the A/E lesion formation. However, some strains that are LEE-negative have been described as EHEC in recent years, such as serotypes O91:H21, O104:H4, and O113:H21, since they also cause HC and HUS (Kaper, Nataro, and Mobley 2004). *E. coli* O113:H21 may employ an unknown host cell mechanism to compensate for the absence of LEE. Adhesion was reported to rely on the activity of Rho GTPases. Adhesion of *E. coli* O91:H21 is thought to be related to *hns* which is responsible for production of hemolysin and alkaline (Scott, Melton-Celsa, and O'Brien 2003; Luck et al. 2005). *E. coli* O104:H4 has adhesion mechanisms of enteroaggregative *E. coli,* which will be discussed in the next paragraph.

1.1.5   Enteroaggregative *E. coli* (EAEC).

Compared to EHEC, enteroaggregative *E. coli* (EAggEC or EAEC) have a very different strategy for colonization and infection. Adhesion is by means of enteroaggregative fimbriae and forming a thick biofilm-like aggregates on the surface of the host cells, and thus adhere tightly to the epithelial layer (Ahmed et al. 2012).

Besides STEC O157 and the famous non-O157 "Big 6", *E. coli* O104:H4 also has been linked to food in the past few years. There was an outbreak in the Republic of Georgia in 2009 (Beutin et al. 2012). In 2011, O104:H4 caused an outbreak which occurred in central Europe and infected almost 4,000 people mainly in Germany. More than 22% (900 cases) developed HUS resulting in 54 deaths, and later, a small outbreak followed in France (Karch et al. 2012). The source of infection in Germany and France was due to the consumption of fenugreek sprouts imported from Egypt.

STEC O104:H4 can be considered as a hybrid of EHEC and EAEC. The genome of O104:H4 harbors virulence genes from both EHEC and EAEC and expresses various phenotypes for each pathotype, such as the production of Shiga toxin (Stx) 2a, which is the main virulence factor of EHEC, and the formation of A/E lesions for culturing on human intestinal epithelial cells, which is the typical colonization method for EAEC strains (Bielaszewska et al. 2011). The distribution

of age and gender in the 2011 outbreak shows an apparent imbalance, the number of adult patient exceeded the children and young women were more susceptible to this strain (Frank et al. 2011). A proposed reason for the unbalanced gender distribution may be the different frequencies of fruit and vegetable consumption between women and men (Perez 2002).

1.2   Shiga toxins.

The main virulence factor of STEC is Shiga toxins. Shiga toxin (Stx) was first reported by Kiyoshi Shiga in his work about the dysentery bacillus *Shigella dysenteriae* (*S. dysenteriae*) in 1898 (Croxen et al. 2013; Shiga 1898). Konowalchuk et al. (1977) discovered a cytotoxic factor that can kill cultured Vero cells, which is a cell line derived from African green monkey kidney epithelium, and therefore named this toxin "Verotoxin". *E. coli* strains that can produce this toxin were called Verotoxin-producing *E.coli* (VTEC) (Konowalchuk, Speirs, and Stavric 1977). Later, VTEC was found to be associated with HC and HUS. Furthermore, as toxins purified from STEC *E. coli* strains were confirmed to be the same as the one produced by VTEC, it was eventually believed that VTEC and STEC are the same (Riley et al. 1983; O'Brien et al. 1982).

When these toxins are produced by *S. dysenteriae*, they were written as Shiga toxins (Stxs), while the other two forms of toxins produced by STEC strains are named as Shiga toxin 1 (Stx1) and Shiga toxin 2 (Stx2). In the late 1980s, protein structure analysis showed that genes responsible for encoding Stx1 in *E. coli* are present in *S. dysenteriae* type 1, which means their *stx* genes are identical (Seidah et al. 1986; De Grandis et al. 1987; Strockbine et al. 1988). However, the structure of Stx2 is different to that of Stx and Stx1, and cannot be neutralized by the antibodies against Stx and Stx1 (Strockbine et al. 1988; Jackson et al. 1987). The subtypes of Stx family and their similarities compared to Stx are shown in Table 1.1(Bergan et al. 2012). In *E. coli*, all Stx subtypes can be genotyped by specific primers using PCR method (Scheutz et al. 2012; Friedrich et al. 2002).

Table 1.1 Variants and sequence similarity to Stx produced by *S. dysenteriae* among different Stx subtypes (Adapted from Bergan et al., 2012).

Toxin	A subunit homology	B subunit homology
Stx	100%	100%
Stx1a	99%	100%
Stx1c	97%	97%
Stx1d	94%	92%
Stx2a	55%	57%
Stx2b	94%	89%
Stx2c	100%	97%
Stx2d	99%	97%
Stx2e	94%	87%
Stx2f	71%	83%
Stx2g	96%	94%

The structure of Stx belong to the AB_5 protein toxin family, which means the Stx proteins have five identical B subunits (7.7 kDa) and one subunit A (~32.2 kDa). Subunit A consists of two parts, an enzymatically active A1 (~27.5 kDa) and a small A2 (~27.5 kDa) for attaching B moiety (Garred, van Deurs, and Sandvig 1995). Five non-toxic subunit B form a ring structure that provides a central pore for the C-terminus of the subunit A to attach (Stein et al. 1992). Each B subunit has three receptor binding sites for the trisaccharide moiety of the glycosphingolipid globotriaosylceramide (Gb3), thus every AB5 toxin can interact with up to 15 Gb3 molecules and result in a strong binding to the cell surface (Ling et al. 1998). Stx2e was identified as the major virulence factor that causes pig STEC infection in addition to human, which is due to the different binding mechanism, Stx2e binds to globotertraosylceramide (Gb4) (Beutin et al. 2008; Keusch et al. 1995). Humans can be infected by Stx2a but STEC strain which harbors *stx2a* persists in cattle without causing disease due to the lack of Gb3 in the intestine of adult cattle.

When the toxin enters the endoplasm, the A2 fragment is cleaved off by the protease furin, but still attach with fragment A1, until the toxin is exposed to the endoplasmic reticulum (ER) lumen (Garred, van Deurs, and Sandvig 1995; Fagerquist and Sultan 2010; Garred et al. 1997).

Using Vero monkey kidney cells to examine protein synthesis inhibition, Fuller et al. that Stx2a is more potent than Stx1 (Fuller et al. 2011). Moreover, among all Stx2 subtypes, stx2a and Stx2d are more potent than Stx2b and Stx2c, whereas Stx2b and Stx2c have similar potencies to Stx1. Cytotoxic test for Stxs revealed that the toxic activity is to inactivate 60S ribosomal subunit and

thus reduce protein synthesis (Reisbig, Olsnes, and Eiklid 1981; Brown, Rothman, and Doctor 1980). Specifically, the Stxs have an RNA N-glycosidase activity that can remove a particular adenosine from the 3' region of the 28S rRNA (Endo et al. 1988; Saxena, O'Brien, and Ackerman 1989).

1.3   Stx phage.

1.3.1   Stx phage gene structure and its diversity.

At the end of the 1990s, the confirmation of the complete sequence of Stx phages revealed that Stx phages belong to the lambdoid phage group (Makino et al. 1999; Matsushiro et al. 1999b; Plunkett et al. 1999). The phages, which have a high lifestyle and genome similarity with lambda phage are always mentioned as "lambdoid" phages. The name "lambloid" is not only due to the similarities in the lifestyles and gene structure, but also these two types of phages can realize a functional recombination to produce a hybrid phage (Kaiser and Jacob 1957), and has expanded to referring a phage with the same functional gene order as lambda phage (Popa et al. 1991).

Lambda phage was first found from a laboratory *E. coli* strain K-12 after UV irradiation (Lederberg 1951). Lambda and lambdoid phages have been studied in detail from the genetic structure to their physical mechanisms for decades. Lifestyles of lambdoid phage, lytic and lysogenic, and the essential genes for determining the lifestyles, such as *cI*, *cro*, *N*, and *Q*, were confirmed and studied by molecular techniques like random mutant hunting (Campbell 1961; Parkinson 1968; Eisen et al. 1970). With the finding of various operators, functional genes and signal pathways, the Stx phage detailed genetic map was constructed (Figure 1.1). Although the main organization of Stx phage genome is structurally similar to other lambdoid phage, and additionally sequences of *stx* genes are always conserved, the DNA sequences of specific function genes or regulators in Stx phage regulation region are usually highly diverse among phages.

Once the Stx phage DNA enters the host cell, there are two possible life cycles: lysogenize into the host bacterial chromosome or perform a lytic development and lyse the cell. Horizontal gene transfer happens during the recombination of phage DNA and bacterial chromosome. While the phage developed a mechanism for infecting the host bacteria, their hosts also developed an antiviral system to limit the integration. This immune system is a cluster of regularly interspaced short palindromic repeats (CRISPR) located in the bacterial genome which are obtained from previous phage infection, and thus protect bacteria when the phage with the same sequence infects again (Brouns et al. 2008).

According to the gene structure map of the Stx phage (Figure 1.1) (Wegrzyn, Licznerska, and Wegrzyn 2012; Casjens and Hendrix 2015), the whole genome can be divided into three parts,

immediate early transcription, delayed early transcription, and late transcription, based on the role they play in the transcription during lytic growth (Friedman and Gottesman 1983).

The *stx* genes located in the late transcription region, downstream of *Q* gene are under the control of the promoter *p*R' and upstream of the lysis cassette, so it is thought to be related to the lysis life cycle and phage replication (Ogura et al. 2015). In *S. dysenteriae* type 1, the Stx phage lacks several genes required for encoding the phage particles, it is therefore considered as a part of the host genome (McDonough and Butterton 1999; Unkmeir and Schmidt 2000). STEC strains can carry more than one *stx* gene with different gene subtypes. Moreover, it has been reported that strains can carry three or more *stx* subtypes (Kruger, Lucchesi, and Parma 2011; Bertin et al. 2001)

Previous studies revealed that Stx phages have the same general genomic organization with lambda phage, which includes key genes, such as *cI*, *cro*, *N*, *Q*, *O*, and *P*. However, there was a high degree of sequence diversity. The plasticity of phage genome is considered to be high, which confers a broad range to the size of Stx phage genome, ranging from 29.7 Kb to 68.7 Kb and mostly larger than 60 Kb. However, by assembling sequence reads from several *stx2a*-converting phage genomes, Yin et al. reported in 2015 that they obtained two phages with merely ~16 Kb size genomes in their study (Yin et al. 2015).
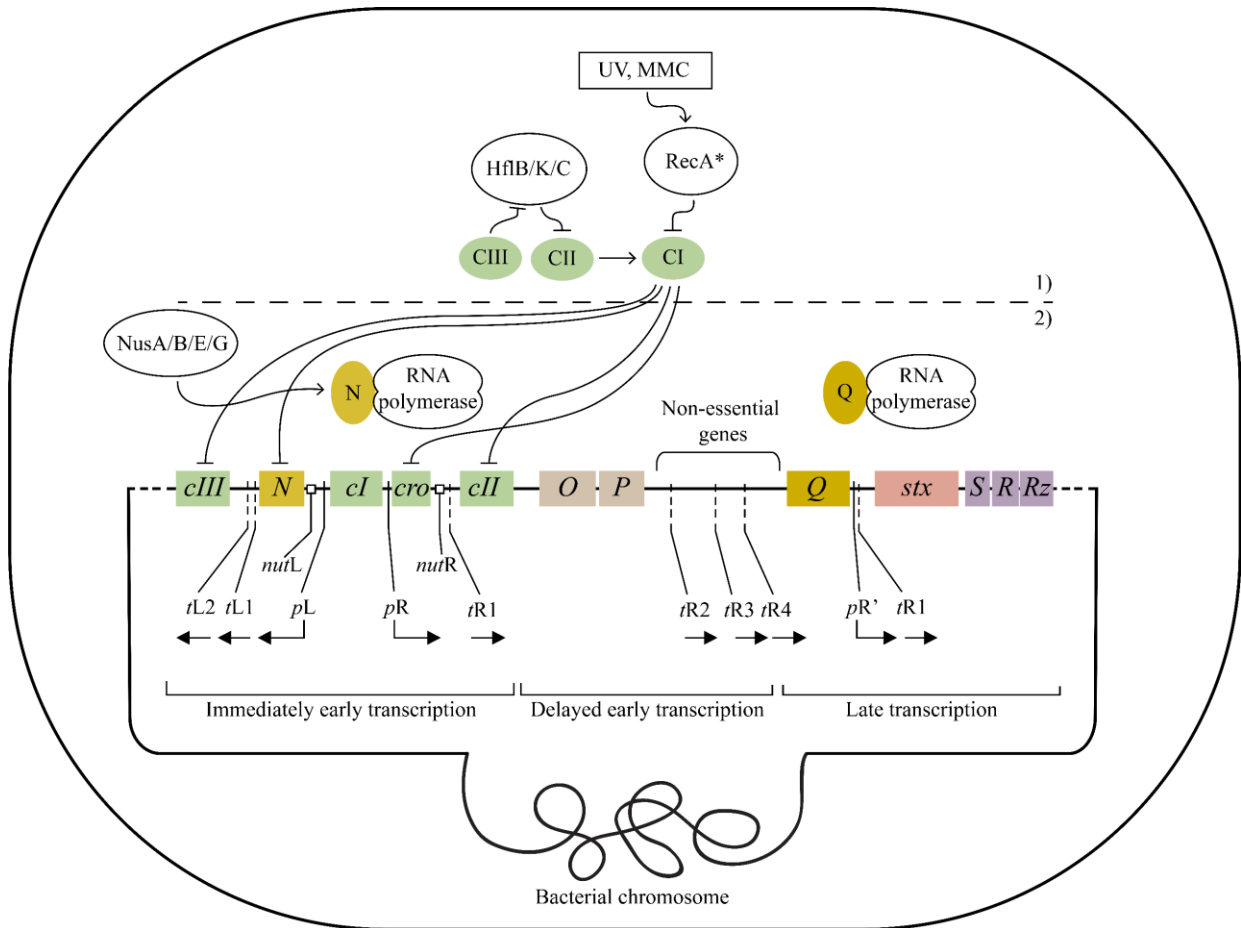
Figure 1.1 The The schematic of Stx prophage gene structure and typical induction mechanism. The round rectangle outside the whole figure indicates the bacterial cell. Rectangles with italic form inside indicate the genes on prophage; colored ovals are phage proteins; lined ovals are host proteins; lined rectangular is environmental conditions. 1) Interactions between host proteins and phage proteins during the lysogenic state and switch from lysogenic to lysis; 2) Interactions between host proteins and phage proteins when enter lysis state. In lysogenization state, proteins CI, CII and CIII establish and maintain the lysogenization state. Environmental stresses, like UV and Mitomysin C (MMC) stimulate host protein RecA to cleave off CI, and then phage switch state to lysis. With the help of protein N and Q, RNA polymerase read through the phage genome, and finally synthesize the Stx. → represents the transcription direction; → represents the activation; ⊣ represents the repression. This figure is edited from previous publications (Wegrzyn, Licznerska, and Wegrzyn 2012; Casjens and Hendrix 2015).

1.3.2 Genes for lysogenization.

As a transcriptional repressor, the CI protein is encoded by the gene *cI* which located in the immediate early transcript, and functions in the maintenance of the lysogeny cycle and immunity to superinfection (Thomas and Lambert 1962; Ptashne 1967). Once the lysogenic state is established, CI represses the main lytic promoters *p*R and *p*L and activates its own promoter *p*M, which are all located in the immediate early transcript. Although the lysogenic state only requires the CI protein to maintain, the establishment of lysongeny needs more factors, specifically the CII and CIII proteins, which are encoded by *cII* and *cIII* located on the right and left operon of the *cI* gene, respectively. CII protein stimulates *cI* transcription from *p*E promoter, which is functionally like the CI protein acting as the transcriptional activator of its own promoter *p*M, and thus ensures the high CI concentrations and establishment of the lysogenic state (Court, Oppenheim, and Adhya 2007). The host HflBC protease complex can rapidly degrade CII protein and thus destroy the lysogenic state, however, the protein CIII can inhibit this action by competitive binding on HflB (Kobiler, Rokney, and Oppenheim 2007). The Cro protein, which is also at the center of the lysis-lysogenization decision, performed as a dimer, to recognize operator DNA and repress transcription (Jia et al. 2005).

However, when we zoom in to the DNA sequence level, there are many differences can be observed in the immediate early transcription. Even with the same *stx* subtype, *cI* and *cro* can only share lower than 50% amino acid sequence identity among different Stx phage sequence types (Yin et al. 2015). As a typical *stx1*-converting phage, H-19B phage share a conserved gene organization as other lambdoid phage in the region containing *cI* and *cro*. However, sequence analysis revealed that sequences in the right operators do not resemble operator sequences of any other lambdoid phages and homology are even lesser with lambda corresponding region (Neely and Friedman 1998a).

1.3.3 Switch of the lysis-lysogenization decision: RecA-dependent and -independent induction.

The Stx expression and cell lysis are the results of Stx prophage induction and the switch from lysogenizaiton to lysis is linked to the bacterial SOS response. The SOS response is a ubiquitous response to DNA damage which triggered by the exposure of single strand DNA (ssDNA). In the case of the switch of the phage life cycle, the autoproteolysis of CI repressor is mediated by the activated host protein RecA*. Accumulating of the ssDNA at the replication fork stimulate RecA to RecA*, which happens after it binds to the ssDNA fragment, and then the activated protein RecA* triggers the autocleavage of the CI repressor, and thus starts the expression of *p*R, *p*L and

9

downstream excision of the prophage (Wegrzyn, Licznerska, and Wegrzyn 2012). Typical SOS inducers under standard laboratory conditions which include UV irradiation and mitomycin C are believed to trigger prophage induction. Many antibiotics that were used for curing the *E. coli* infection, such as kanamycin, phosphonomycin, and norfloxacin, can also provoke a phage induction (Matsushiro et al. 1999a; Los et al. 2009). The reason for this is because these antibiotics are inhibitors for the bacteria DNA gyrase, and thus cause an accumulation of ssDNA which capable of activating the RecA protein. On the other hand, factors emerging in the natural habitat of *E. coli* can also cause the lysis-lysogenizaiton switch. Hydrogen peroxide can also trigger phage induction. The mechanism of H_2O_2-mediated phage induction may include the action of the host protein OxyR, which play a role as a stimulus of prophage maintenance under the condition of oxidative stress (Glinkowska et al. 2010).

While SOS response is the main cause for Stx prophage induction, previous study also suggested that prophage induction may be occurred through a RecA-independent way. High concentration of salt (NaCl), such as 200 mM, which can cause phage 434 induction, is due to RecA-independent pathway (Shkilnyj and Koudelka 2007). The mechanism of this induction may be the impairment of CI binding to the DNA sequence, which is due to the topologic structure change of protein CI or DNA caused by the high salt concentration. It was a meaningful discovery since the salt stress condition is always found in natural environment. EDTA can also cause RecA-independent induction by its chelating properties (Imamovic and Muniesa 2012). Citrate and other chelating agent can also induce Stx phage by the same properties. The mechanism of EDTA induction is disrupting the bacterial outer membrane by chelation of Mg^{2+}, and thus pH value plays a important role in terms of chelate complex induction.

1.3.4 Genes for lysis.

After the prophage enters the lysis life cycle, RNA polymerase starts to produce RNA products. At the same time, regulatory proteins expressed from prophage genome participate into this process, form complexes with the RNA polymerase to facilitate it read through terminators on phage operators, and finally lead the cell lysis. Typically, there are three phases during transcription: initiation, elongation, and termination. Bacterial RNA polymerase contains a core enzyme structure which consists of subunit compositions $\alpha_2\beta\beta'\omega$, in charge of random, non-specific initiation and elongation, and an initiation factor σ, which is required for RNA polymerase holoenzyme ($\alpha_2\beta\beta'\omega\sigma$), in charge of sequence-specific interaction with promoter DNA (Vorobiev et al. 2014). In *E. coli*, principal bacterial factor σ is σ^{70}, which contains five conserved regions: σR1, σR2, σR3, σR4, and the σR3/σR4 linker. Among these five regions, σR2

and σR4 mediate sequence-specific interactions with the promoter -10 and -35 element, respectively.

1.3.4.1   Transcription antitermination.

The mechanism for regulating the efficiency of RNA production is called transcription antitermination. The major regulatory proteins are N and Q. N is responsible for helping the RNA polymerase to read through the early transcripts, whereas protein Q is for the late transcript. N-mediated antitermination is completed by a large nucleoprotein complex, which formed by N protein, RNA polymerase, and several host-encoded regulatory proteins, such as NusA, NusB, NusE, and NusG (Wegrzyn, Licznerska, and Wegrzyn 2012). With the help of this complex, RNA polymerase passes through terminators on the left and right operons, starts to transcribe lysis-related genes. Moreover, RNA polymerase and N protein are not combined to each other directly but via a *nut* (N utilization) site on the phage DNA located on both sides of the gene *N*. Detailly, N protein binds to the *boxB* sequence of the *nut* site in the left and right operons, respectively, and also binds to NusA protein and RNA polymerase (Mogridge et al. 1998; Mishra et al. 2013). Following N-mediated antitermination, protein Q, which is encoded by the gene *Q*, is expressed for helping the RNA polymerase to read through terminator *t*R' on the late transcript region and to allow the genes on the late operon to be expressed. In contrast to the antitermination of protein N, Q protein can provide stable modification of RNA polymerase in vitro alone, which only need an associated *qut* sequence on the DNA strand for recognition (Roberts et al. 1998). The transcriptions, which happene in this stage, are completed by the protein Q, *qut* site, holoenzyme, and NTPs and ions. The *qut* site contains the -35 and -10 elements of the late promoter *p*R', a Q binding element (QBE) which located between the -35 and -10 elements of *p*R' or partially overlap these elements. While the binding between protein Q and the QBE DNA site, the interaction between σR4 and the tip of the RNA polymerase β subunit known as the "flap" is required for σR4 binding to the -10 element (Kuznedelov et al. 2002).

Amonst antiterminator *Q* genes, Q_{933} which was found in phage 933W, is often related with higher *stx* expression. Another phage *Q* which identified from phage 21, share only 36% predicted amino acid homology with Q_{933}. Sequencing of three SF O157 strains showed that the sequence of their *Q* genes was identical or similar to the $Q_{O111:H-}$ gene observed in *E. coli* O111:H- strain AP010960, but different to Q_{933} or Q_{21} (Haugum et al. 2012).

Stx2-encoding phage 933W, which isolated from EHEC O157:H7 strain EDL933, was considered to be a mosaic of different phages since it concieves a common phage backbone while possesses a divergent arrangement of ORFs and genetic elements to its relative lambdoid phages (Plunkett et al. 1999). Three tRNA sequences were found in the region between gene *Q* and the

stx gene in phage 933W. Amongst these three, tRNA2 has an anticodon started by U, which is rarely found in *E. ocli* tRNAs. Since tRNA genes provide sequence for integration homologues recombination (Schmidt et al. 1997), this change may provide 933W phage a restrict recognition to a specific species during recombination.

1.3.4.2   DNA replication and its initiation.

The Stx phage harbors two replication genes, *O* and *P*. The replication initiates at *ori* region, and proceeds bidirectionally. During the first several tens of circles, phage genomes are generated by circle-to-circle replication (θ replication mode), whereas the rest replications switch to rolling circle mode, which is also known as σ mode, to generate the concatemer DNA for packaging the new virions (Narajczyk et al. 2007).

The understanding of phage DNA replication may be of bio-medical importance, since the genes coding for Shiga toxins are located immediately behind the replication cassette in the phage genome, and thus could be affected by the divergent of replication cassette.

During DNA replication, four O protein dimers compose an "O-some" complex and bind to the *ori* region, which conceives four sequence repeats and lies inside the *O* gene. Additionally, protein P and host protein DnaB helicase also participate into the formation of the replication complex, and finally form a *ori*:O:P:DnaB preprimosomal complex. Protein P and DnaB are functioned as O protein protectors from the degradation of host ClpP/ClpX-mediated proteolysis (Zylicz et al. 1998). Moreover, host DnaK-DnaJ-GrpE chaperone system participate in activating the preprimosomal complex and evoke more host replication factors to participate in the replication initiation (Learn et al. 1997). Moreover, the *ori* region can also be activated by the promoter *p*R by an unknown the mechanism (Olszewski et al. 2014). Notably, not all the lambdoid phages have protein P. the prophage N15 exists in bacteria as a linear plasmid molecule with its end closed. The only gene required for phage replication is *repA*, which encodes a protein have a function of helicase and thus independent to the host DnaB helicase but only need the host DnaG primase (Mardanov and Ravin 2006). Some other λ phage like P22, P protein is replaced by a DnaB type helicase type protein encoded by phage, which binds to the ssDNA in the presence of ATP (Wickner 1984).

H-19B phage genome contains nearly identical *O* and *P* genes with lambda phage. However, there are 39 additional nucleotides located in the middle of the H-19B *O* gene, which cause two additional direct repeats, and thus confer to the *O* gene the ability to encode a six O protein-binding sites, but not four. It is believed that *stx2a*-converting *E. coli* strains from serotype O157:H7 can often cause serious human disease. However, according to the type of replication proteins, Ogura et al. subtyping Stx2a phages of O157 into four different groups (Ogura et al.

2015). By subsequent examination of the toxin production of each group, the results revealed that only strains belonging to two of these groups performed highly Stx2 production level, whereas the toxin producing ability in the rest two groups were way lower than them, even if all of these strains harbor *stx2a*-converting prophage.

1.3.4.3   Host cell lysis.

The lysis cassette includes *S105*, *S107*, *R*, *Rz* and *Rz1*, which is located behind the *stx* gene. R protein is for cleaving the cell wall (Harris et al. 1967), which functioned as a murein transglycosidase in lambda phage (Bienkowska-Szewczyk, Lipinska, and Taylor 1981). In phage P22, this endolysin is considered to be a lysozyme by examining its crystal structure using X-ray crystalanalysis (Mooers and Matthews 2006). The release of protein R is under the control of protein S105 and S107, which are two holing proteins encoded by the same reading frame but different start codons. Holins are small membrane proteins accumulate in the cell membrane and help the endolysin to disrupt the inner membrane and break down the cell wall (Wang, Smith, and Young 2000). S107 is two amino acids longer than S105 and function as lysis inhibitor, whereas the shorter one, S105, function as a lysis stimulator by forming large holes in the membrane. Because of these two proteins opposite functions, they play a role in determining the time of lysis. Additionally, there are two other proteins Rz and Rz1 participate in the cell lysis process. Rz and Rz1 are inner membrane protein and outer membrane lipoprotein, respectively, and typically work as a complex, known as "spanin", since this complex always span through the whole periplasm (Young 2014). Generally, the gene organization inside this cassette is *S*, *R*, *Rz*. However, ORFs often insert into this cassette between these typical genes. In phage 933W, there is an additional ORF inserted between *R* and *Rz*, which can produce a protein with 34.2% identity with the phage P22 Ant antirepressor (Plunkett et al. 1999). It was considered all the *stx* gene were immediately upstream of *S* and *R* genes. However, in phage H-19B, there is a ~3 Kb DNA sequence separates the toxin genes and the lysis cassette (Neely and Friedman 1998b).

1.3.5   Interaction between Stx prophage and its host.

Besides the host HflBKC affects the stability of CII, the host physiological state might also influence the lysis/lysogeny decision by the following several pathways. The phages are more likely to be in the lysogeny state when the concentration of guanosine tetraphosphate (ppGpp), which is a nucleotide synthesized by the host bacteria in response to amino acid or carbon source starvation, is in a higher level (Slominska, Neubauer, and Wegrzyn 1999). Moreover, the cAMP, an alarmone produced by the host responses to the poor growth conditions, can also stimulate the CII activity via down-regulates the activity of HflB protease. RNase III activates *N* gene translation (Wilson et al. 2002). The first gene product N protein from the *p*L operon forms a

complex named *N* leader RNA, to help RNA polymerase read through terminators in the early transcript. The *N* leader RNA contains a *nut* sit, an RNase III-sensitive hairpin and an *N* ribosome-binding site. RNase III stimulates *N* gene translation more than 200-fold by preventing N-mediated translation repression of *N* gene expression. RNase III can also regulate the rate of *cIII* translation initiation by binding to the alternative structures region of the mRNA (Altuvia et al. 1991). The host integration host factor (IHF), which is a heterodimeric DNA-binding and -bending protein, modulates the phage proteins CII and CIII levels, and *N* gene translation (Giladi et al. 1998; Wilson et al. 2002). As positive regulators, the host proteins DnaA and SeqA, stimulate the transcription from *p*R to *p*L and bind downstream of the transcription start point, and thus affect the lysis-lysogeny decision (Wegrzyn, Licznerska, and Wegrzyn 2012). Finally, host protein RecA which gene expression is caused by DNA damage initiates a CI cleavage and thus cause the start of the phage lysis cycle.

Although phage proteins CII, CIII and proteins like N, O lead the prophage to different life cycles, all of them can be degraded within minutes after being made to make sure the life cycles can be switched quickly and maintain stable after establishment (Kobiler, Oppenheim, and Herman 2004; Casjens and Hendrix 2015).

1.3.6 Other factors affecting Stx phage diversity.

Other factors that can affect the toxin production and toxicity include the virulence plasmid which can encode various virulence factors. Large plasmids pO157 and pO113 are virulence plasmids are often carried by LEE-positive and LEE-negative STEC isolates. The virulence plasmid pO113 (~166 Kb) which was found in STEC O113:H21 isolate EH41, contributes to the EHEC hemolysin (Ehx) which is encoded by the *ehxA* gene located on pO113 (Newton et al. 2009). Similarly, pO157 carries the same *ehxA* gene in O157:H7 isolates.

Insertion sites of Stx phages are different with other lambdoid phages. Generally, phages insert within tRNA genes (Campbell 2003), whereas Stx phages insertion sites are frequently genes from basic genetic equipment of bacterial chromosome. Additionally, insertion sites of Stx phages are highly diverse. Most common insertion sites in LEE-positive STECs are *wrbA,* encoding for a tryptophan repressor protein, *yecE* for a transcriptional regulator, *yehV* with unknown function, *argW* for tRNA-Arg, *ssrA* for a tmRNA, and *prfC* for peptide chain release factor 3 (Ogura et al. 2007). Insertion sites in the genomes of LEE-negative STEC isolates are often different when compared with the LEE-positive strains, which include *ynfH*, *yecE*, *serU*, *yciD*, and *potC* (Steyert et al. 2012). The integration may cause physiologic change of the *stx*-converting bacteria (Plunkett et al. 1999). Additionally, when this gene integration occurs at a

specific functional genomic site, it may cause a gene breakage and thus trigger a gene function lost or some novel gene generation.

The DNA sequence variation existing in the Stx genome, especially in the late regulation region, lead to many genes with poorly known roles and many hypothetical proteins, and thus may lead to the difference in phage induction, *stx* gene expression and toxin production (Smith et al. 2012).

1.3.7   Bacteria stochastic phenotype variation.

A culture of microorganism population is considered as homogeneous has become a history. The current research indicates that, even in an ordinary laboratory culture, the composition is consisted of heterogeneous subpopulations. Closely related species, or even in pure culture derived from a single colony, may exhibit marked differences in biochemistry and behavior. When encounter environmental changes or under different conditions, these subpopulations react differentially (Rainey et al. 2011).

Bacteria generate variable offspring, which is called stochastic phenotype switching, is in case of being maladapted in the present environment, and increase the possibility of survival in different environment in the future. The selective agents of this risk-reducing strategy are exclusion rules and population bottlenecks which are act in tandem during stochastic phenotype switching(Libby and Rainey 2011).

Since small portion of STEC can kill unicellular predators or human leukocytes while the cell lysis, and thus can benefit the rest (Los et al. 2012). Single-cell level analysis provides a good way of characterizing the switch of lysis-lysogenization decision when combined with fluorescent reporter genes.

1.4   Flow cytometry.

1.4.1   Single-cell level analysis.

The techniques that can analyze cell individually include fluorescent and confocal microscopy, scanning and image cytometry, and flow cytometry (Shapiro 2000). The tasks for measuring microorganisms in single-cell level, is not only to count the cell numbers or the intracellular complexity, but also to characterize the growth, metabolism and clinical situations, and even detect, count and characterize each of several organisms in a mixed population.

Most instruments for single-cell analysis make optical measurements. Although measuring the light absorption is the easiest way for characterizing individual microorganisms or partials, the resolution limit makes the optical measurements difficult or even impossible. Electronic coulter counter is a type of electronic detector which record the electrical change while a cell or particle passes through the detection point. It can be used to count bacteria cell number and even viruses (Kubitschek 1958).

By exciting the fluorescent reagent from the sample, it is possible to measure the particles under the resolution limit of optical microscopes by means of an electronic detector, such as a photomultiplier tube (PMT), photodiode, or charge-coupled device (CCD). Based on the light scattering, small virous particles were first detected by a ultra-microscope in 1920s. The first modern flow cytometer appeared in the late 1940s and early 1950s, detected bacteria with a size of 0.6 μm in aerosols (Gucker, O'Konski, and et al. 1947).

Color, intensity, and polarization are the three qualities of light. Although the first two qualities of light, color and intensity, are usually considered into the flow cytometry measurement, the polarization are more and more concerned since the intensity may depend on the degree and direction of polarization (Asbury, Uy, and van den Engh 2000). The light polarization may especially be considered when measuring cells of eukaryotes, since the cell walls are usually birefringent.

1.4.2   Basic mechanics in flow cytometry.

Since Mack Fulwyler invented the first current type flow cytometer in 1965 (Fulwyler 1965), flow cytometry has been developed, refined and widely used in many fields, such as clinical diagnosis and laboratory researches. As the name 'flow' implies, cell measurement happens in a flowing stream.

According to the light scattered from the fluorescence emitted from dyes, fluorophore-conjugated antibodies and fluorescent proteins, flow cytometry techniques offers a transiently and simultaneously analysis in single-cell level, providing cell information such as cell size, cellular complexity (Shapiro 2003).

There are three main parts to conform a typical flow cytometer: fluidics, optics and electronics systems. The fluidics system is for making the cell or micrometric particle pass through the interrogation point in a single fine line. Sample contains cells or particles must be prepared in liquid. Ideally, by adjusting the laminar flow exerted by a surrounding sheath stream, cells pass through the center of the core stream one at a time. The optics system contains a light source which commonly is 488 nm light source, Argon ion laser. When cell passes through the laser at the interrogation point, it will reflect or scatter at all angles. Among all these angles, forward scatter (FSC) is the amount of light that scattered at acute angles, which indicates the size of the cell. The light scattered at large angles (90°), which is called side scatter (SSC), is caused by granularity and structural complexity inside the cell. The electronics system is for filtering unwanted wavelength, and at the same time, amplifying and converting light signal into voltage pulse, and finally converted to numbers by the analog signal processing electronics. The FSC light is firstly detected by a photodiode or by a PMT, and then converts into voltage pulse, the

magnitude of the voltage pulse is proportional to the cell size. The same as FSC, once SSC light comes out the cell, it is subsequently focused through a set of lenses and collected by separate detectors which are located 90 degrees from the laser's path. The side scattered light is then converted into voltage pulse proportional to the amount of light.

In environmental or food industrial microbiology, samples of microorganisms may from various source (e.g., water, beverages, and food). Some of the sources have strong background auto-fluorescence, and thus characterization by light scattering appears not enough to accomplish the task. In this situation, fluorescent biological reagents and sample pre-staining become important to the development of microbiological assays (Veal et al. 2000). Beside the fluorescent dye, the application of fluorescent proteins in flow cytometry provides a way for monitoring the gene expression or protein localization (Chalfie et al. 1994).

1.5   Fluorescent proteins.

Fluorescence is a form of luminescence. To make an atom or a molecule to emit fluorescence, there should be an excited light source which has a wavelength shorter than or equal to the wavelength of the emitted light. When excited by the light source, the electron of the molecule absorbs the energy and raise from the ground state to a higher level, which is known as an excited state. The fluorescence occurs when the electron loses the absorbed energy by light emission, the process between energy absorption and light emission is called fluorescence lifetime. As the discovery and development of fluorescent proteins, fluorescent reporter gene technology is widely used in various aspects and applications.

Fluorescent proteins (FPs), which are mainly from *Cnidaria* (Yarbrough et al. 2001), are widely used in observing living cells and tissues. The maturation of FPs includes two consecutive steps: protein folding and chromophore formation. FP tertiary structures are highly conserved and organized by a caged β-barrel (formed by 11-stranded β-sheets) and an internal distorted helix (Day and Davidson 2009). Contrast to other biosynthesized pigments, FPs barrel structure can modify its own internal amino acids, and emit light without any external cofactors and substrates except molecular oxygen. The FPs are stable in a wide range of pH because of the β-barrel, that the protein denaturation and chromophore degradation can only be observed when the pH values lower than 2 or above 12 (Piatkevich and Verkhusha 2011).

1.5.1   GFP

Green fluorescent protein (GFP) was first isolated and characterized from jellyfish *Aequorea victoria* (Morise et al. 1974). The cDNA structure of GFP revealed in 1992 enables researchers to construct various gene expression vectors by using it as a reporter (Prasher et al. 1992).

GFP is a 27 kDa (238 amino acids) protein. The chromophore group is formed by three amino acid residues of the helix at positions 65-67 (Ser-Tyr-Gly in *Aequorea victoria* GFP), which provides a large, plane conjugated π-system for absorbing and emitting lights within the visible range (Chudakov et al. 2010). Among these three amino acids, Tyr-66 and Gly-67 are conserved among all natural GFP-like proteins.

By introducing random polar amino acid substitutions flanking the chromophore core Ser-65-Tyr-66-Gly-67, the mutant of GFP, enhanced GFP (eGFP), have a 35-fold higher brightness than the wild type GFP, and a shift of the excitation from 395 nm to 490 nm. Thus, compared to wild-type GFP, it is better for standard 488 nm laser and fluorescein filter sets (Cormack, Valdivia, and Falkow 1996; Heim, Cubitt, and Tsien 1995).

1.5.2 DsRed

Among these color variants, red fluorescent protein (RFP) is ideal for multi-color imaging with GFP, since their narrow emission spectra differ sufficiently for them to be imaged in distinct fluorescent channels. RFP, which derived from nonbioluminescent *Anthozoa* species, belongs to a family of GFP homologue proteins. RFP has a similar tetrameric/monomeric protein structure as GFP. Its amino acids tend to alternate between hydrophobic and hydrophilic along the β-strand. The inside core of the protein and all key secondary structure elements are conserved when compared with GFP. The length of N terminus is the same as its of GFP (Matz et al. 1999).

DsRed or drFP583 is derived from coral *Discosoma* sp. It is a 28kDa wild-type spontaneously fluorescent protein which has excitation and emission maxima at 554 and 586 nm, respectively (Matz et al. 1999). Also, it can be excited by 488 nm laser, which enables it to be used in laser-based confocal microscope or flow cytometers (Hawley et al. 2001). DsRed has a similar topology to GFP, but has a red-shifted spectrum which is mainly due to a chromophore with a more extensive conjugated π-system (Wall, Socolich, and Ranganathan 2000). The chromophore of DsRed is formed by the residues 66-68, Gln-Tyr-Gly, which are homologous to the chromophore-forming residues of GFP, 65-67 (Ser-Tyr-Gly). Typically, all GFP-like proteins have a more or less pronounced tendency to oligomerize (Chudakov et al. 2010). Among all these fluorescent proteins, DsRed has a strong oligomerization which will raise a concern when it is fused with a host protein, since it can disturb the function and localization of the protein. However, this drawback may not affect when it is used as a reporter in gene expression (Bevis and Glick 2002).

In summary, since the isolating and characterization of the first FP, GFP, several decades ago, a wide palette of different color fluorescent proteins have been cloned out and biotechnologically developed for imaging detection. Combined with the application of flow cytometry, the FP

technology provides a possibility for sorting various cells, detecting protein-protein activities, and monitoring gene expression.

1.6 Hypothesis and objectives

This research aimed to test the hypothesis that different sequence structures in the prophage late promoter region contribute to the difference in Stx prophage induction.

The specific objectives were:

1) to analyze the gene structures of the late transcript in all the candidate STEC strains.

2) to construct a promoter screening vector (promoter-*rfp* fusion reporter system), which can be controlled by the native regulation transcript.

3) using the reporter system, to determine phage induction level resulted by the interaction between different promoters and different prophage regulations by flow cytometric analysis.

2   Methods and Materials

2.1   Bacterial strains and growth conditions

2.1.1   Bacterial strains and plasmids

Nineteen STEC strains examined in this study were obtained from the laboratory strain collection, which were collected from human, cattle, clinical incidence or unknown origin. According to previous study (Mercer, Zheng, Garcia-Hernandez, Ruan, Ganzle, et al. 2015), 16 strains were reported as EHEC strain which contain a virulence factor encoding gene *eae*. Mutant strain *E. coli* O104:H4 strain 11-3088 $\Delta stx::gfp::amp^r$ was obtained in an earlier study from the knockout of the *stx* gene and construct of GFP reporter gene and an ampicillin resistant gene. Strain *E. coli* DH5α was used during the molecular cloning process for carrying the plasmids.

The pDsRed-Express vector (Clontech, Mountain View, CA, USA) was purchased from Clontech, storaged in Tris-HCl and EDTA buffer, and was transformed into strain *E. coli* Top 10 in an earlier study. Strain *E. coli* Top 10 carrying the plasmid pUC19 was obtained from lab strain collection.

All bacterial strain used in this study are listed in Table 2.1.

Table 2.1 Strains and plasmids used in this study.

FUA	Accession Numbers	Strains	Description	Origin
1308	LDYN00000000	*E. coli* O26:H11 05-6544	*stx1 eae*	Human
1312	LDZZ00000000	*E. coli* O121:H19 03-2832	*stx2 eae*	Human
1313	LEAA00000000	*E. coli* O121:NM 03-4064	*stx2 eae*	Human
1307	LEAB00000000	*E. coli* O145:NM 03-6430	*stx1 eae*	Human
1303	LEAD00000000	*E. coli* O157:H7 1935	*stx1 stx2 eae*	Human
1399	LEAE00000000	*E. coli* O157:H7 CO6CE900	*stx2 eae*	Clinical
1401	LEAF00000000	*E. coli* O157:H7 CO6CE1353	*stx1 stx2 eae*	Clinical
1398	LEAG00000000	*E. coli* O157:H7 CO6CE1943	*stx1 stx2 eae*	Clinical
1400	LEAH00000000	*E. coli* O157:H7 CO6CE2940	*stx2 eae*	Clinical
1305	LEAI00000000	*E. coli* O157:H7 CO283	*stx1 stx2 eae*	Cattle
1306	LEAJ00000000	*E. coli* O157:H7 E0122	*stx2 eae*	Cattle
1304	LEAK00000000	*E. coli* O157:H7 LCDC7236	*stx1 stx2 eae*	Human
1402	LECF00000000	*E. coli* O103:H25 338	*stx1 eae*	Clinical
1302	LECH00000000	*E. coli* O104:H4 11-3088	*stx2*[d]	Human
1403	LECI00000000	*E. coli* O111:NM 583	*stx1 eae*	Clinical
1316	LECJ00000000	*E. coli* O111:NM PARC447	*stx1 stx2 eae*	Unknown
1309	LECK00000000	*E. coli* O113:H4 09-0525	*stx1 stx2*	Unknown
1311	LECM00000000	*E. coli* O45:H2 05-6545	*stx1 eae*	Human
1310	LECN00000000	*E. coli* O76:H19 09-0523	*stx1 stx2*	Unknown
1004		*E. coli* DH5α	For carrying plasmids	
1293		*E. coli* Top 10	Plasmid pUC19 for cloning	
1029		*E. coli* Top10	Plasmid pRFP	
		E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp*	*stx* gene knocked out	

(Mercer, Zheng, Garcia-Hernandez, Ruan, Gänzle, et al. 2015)

2.1.2 Growth conditions, media and antibiotics.

Strains were grown in Luria-Bertani (LB) medium (BD, Fisher Scientific, Ottawa, CA), with antibiotic selection when required, at 37 °C with agitation at 200 rpm. LB agar plates were made with LB broth to which 1.5 % granulated agar (BD, Fisher Scientific, Ottawa, CA) was added. Based on the different antibiotic-resistant genes, positive clones were selected on LB agar with antibiotic selection.

Strain stocks were stored at -80 °C in a 20 % glycerol (v/v). Inoculations were done by using sterilized plastic loops (INO-LOOP, Daigger Scientific, USA) from a single colony or with metal loop from -80 °C strain stocks.

Antibiotics were dissolved in HPLC water or 100 % ethanol as the properties required at the requisite stock concentration and filter sterilized, and stored at -20 °C for up to a year. The working concentrations and solvents of antibiotics are shown in Table 2.2.

Table 2.2 Antibiotic concentrations and solvents.

Antibiotic	Solvent	Stock solution (g/L)	Working concentration (mg/L)
Ampicillin	Water	100	50
Chloramphenicol	Ethanol	25	25

2.2 Sequence analysis and primer design.

2.2.1 Sequence analysis.

To determine the regulation region of the *stx* gene, which is located in different contigs with the toxin gene, contig assembly was done as the first step of sequence analysis. In addition, there were STEC strains harboring more *stx* genes while *stx* genes and their regulation regions were located in different contigs. Contig pairing was performed using polymerase chain reaction (PCR) to pair the toxin gene with its correct upstream regulation.

For contig assembly and pairing, contigs with the *stx* gene in this study were first submitted to the National Center for Biotechnology Information (NCBI) Nucleotide BLAST database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Among the resulting sequences, sequence which had 100% identity with the query was selected as the reference sequence. The complete reference sequences were downloaded from the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nucleotide) according to their accession numbers. Information of the reference sequences are shown in Table 2.3.

Table 2.3 Reference sequences for contig assembly and pairing.

FUA numbers	Contig numbers, *stx* subtypes	Reference strain	Accession number
FUA 1302	125, *stx2d*	*E. coli* strain GB089	CP013663
FUA 1303	77, *stx1*	*E. coli* O157:H7 SRCC1675	CP015023
	130, *stx2a*	*E. coli* O157:H7 SRCC1675	CP015023
FUA 1305	15, *stx1*	*E. coli* O157:H7 strain 8368	CP017444
	33, *stx2a*	*E. coli* O157:H7 strain F422	AP012531
FUA 1307	130, *stx1*	*E. coli* O26 FHI20	LM995817
FUA 1308	15, *stx1*	*E. coli* O26:H11 strain 11368	AP010953
FUA 1309	23, *stx1c*	*E. coli* O113 FHI30	LM995905
FUA 1310	33, *stx1c*	*E. coli* NA FHI92	LM997172
FUA 1311	159, *stx1*	*E. coli* FORC_028	CP012693
FUA 1312	129, *stx2*	*E. coli* O121 FHI83	LM996896
FUA 1313	186, *stx2*	*E. coli* O121 FHI83	LM996896
FUA 1316	108, *stx1*	*E. coli* O157:H7 strain 8368	CP017444
	180, *stx2a*	*E. coli* O157:H7 strain 8368	CP017444
FUA 1398	105, *stx1*	*E. coli* O157:H7 SRCC1675	CP015023
	131, *stx2*	*E. coli* O157:H7 SRCC1675	CP015023
FUA 1399	28, *stx2a*	*E. coli* O157:H7 FRIK944	CP016625
	179, *stx2a*	*E. coli* O157:H7 F765	AP012534
FUA 1400	150, *stx2a*	*E. coli* O157:H7 strain 8368	CP017444
FUA 1401	60, *stx1*	*E. coli* O157:H7 SRCC1675	CP015023
	99, *stx2*	*E. coli* O157:H7 SRCC1675	CP015023
FUA 1403	7, *stx1*	*E. coli* O111:H- 1639/77	AJ304858

Because some of the contigs of genome sequences used in this study still contained gaps, the complete sequences were obtained after Sanger DNA sequencing in the cloning step. Toxin gene searching, sequence alignment and primer designing were all performed in Geneious (v.9.1., Biomatters, Auckland, New Zealand).

2.2.2   *E. coli* FUA 1304 *stx* genotyping and *E. coli* FUA 1399 contig pairing.

Since there was no *stx* gene annotated in the gene file of *E. coli* O157:H7 LCDC7236 (FUA 1304), PCR was performed to search the carriage of *stx* gene by using the primer pairs shown in Table 2.4, to target the *stx1* subtype by generating a 282 bp size product and the *stx2* subtype by synthesizing a 584 bp product.

Table 2.4 Primers for *E. coli* FUA 1304 *stx* gene genotyping.

Primer	Sequence (5'-3')	Target gene	Reference
KS7	5'-CCCGGATCCATGAAAAAAACATTATTAATAGC -3'	*stx1*	(Schmidt et al. 1994)
KS8	5'- CCCGAATTCAGCTATTCTGAGTCAACG -3'		
LP43	5'- ATCCTATTCCCGGGAGTTTACG -3'	*stx2*	(Cebula, Payne, and Feng 1995)
LP44	5'- GCGTCATCGTATACACAGGAGC -3'		

The strain *E. coli* O157:H7 CO6CE900 (FUA 1399) has two *stx2a* genes and upstream or downstream regions of both genes are located in different contigs (Figure 2.1). However, since both *stx* genes have the same sequence, it cannot be paired by analyzing the regulation region. Therefore, contig pairings were performed by PCR using four primers: LEAE 28-182 F (5'-CCCGGACAGGCGTAATACTC-3'), LEAE 28 R (5'- TGGCATCCTCCTCAGCTACT-3'), LEAE 104 F (5'- TTCGGCCTGGTACCACCGAA-3'), LEAE 179 R-1 (5'-TGTCCGTAGCGTCAAAGCAG-3'). Through four types of combinations between different forward and reverse primers, different size products can be synthesized and used to determine which upstream regulation should be combined with which downstream region.
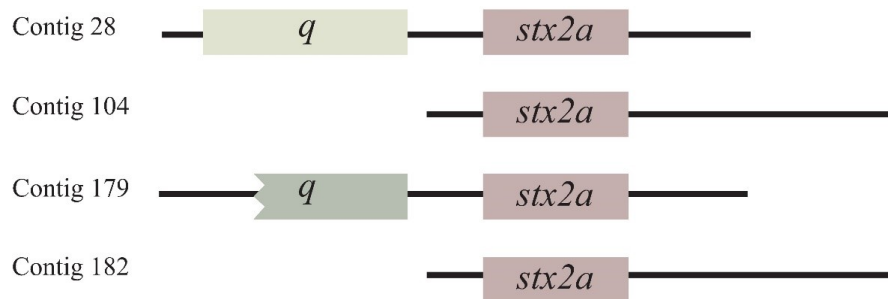
Figure 2.1 Sequence structures of the four contigs containing the regulation regions or toxin genes. Black lines are gene sequences, rectangular boxes are regulation genes and toxin genes.

2.2.3 Nomenclature of promotor constructs.

The region between the *stx* gene and the *q* gene contains the late promoters (*p*R') which is immediately responsible for starting the *stx* gene expression. To determine the exact location of this *p*R' site, *E. coli* O157:H7 strain RIMD 0509952 (accession number: AP000400) was chosen as the reference *p*R' sequence. Finally, the target fragment was determined to be the region started from the last 42 bp of Q protein and ended by the first 39 bp of the *stx* gene to make sure every strain's *p*R' can be included, and was named as *p*R' directly due to the existence of the *p*R' promoter, i.e. target *p*R' fragment from FUA 1302 in plasmid pUC19 written as P*p*1302.

Target fragments from strains harbouring only one *stx* gene were denoted by the strain FUA number, whereas fragments from strains with more than one *stx* gene were denoted by the FUA number and the abbreviation of the *stx* subtype. Target fragments from FUA 1399, which harbors two *stx2a* genes, were denoted by the FUA number and the contig number.

2.2.4 Phylogenetic analysis.

Phylogenetic trees of late regulation region and *stx* genes were generated by Geneious (Biomatters, Auckland, New Zealand) according to the user's manual. To generate the phylogenetic tree, sequences of the target fragment were first aligned using MUSCLE (Edgar 2004) in "Multiple alignment" tool, Geneious. Results of alignment were then used to build the tree. Parameters "Tree build Method" and "Resampling Method" were set as "Neighbor-Joining" and "Bootstrap", respectively, while the rest of the parameters were set to default values.

2.3 Procedures for construction and transformation of plasmids.

To obtain the *p*R'::*rfp*::*chl^r* fusion reporter system, target fragment *p*R', *rfp*, and *chl^r* were amplified from candidate EHEC strains, plasmid pDsRed, and plasmid pKD3, respectively. Three fragments were then ligated together and transformed into the vector pUC19.

The whole procedure for cloning and construction of the *p*R'::*rfp*::*chl^r* fusion is shown in Figure 2.2.
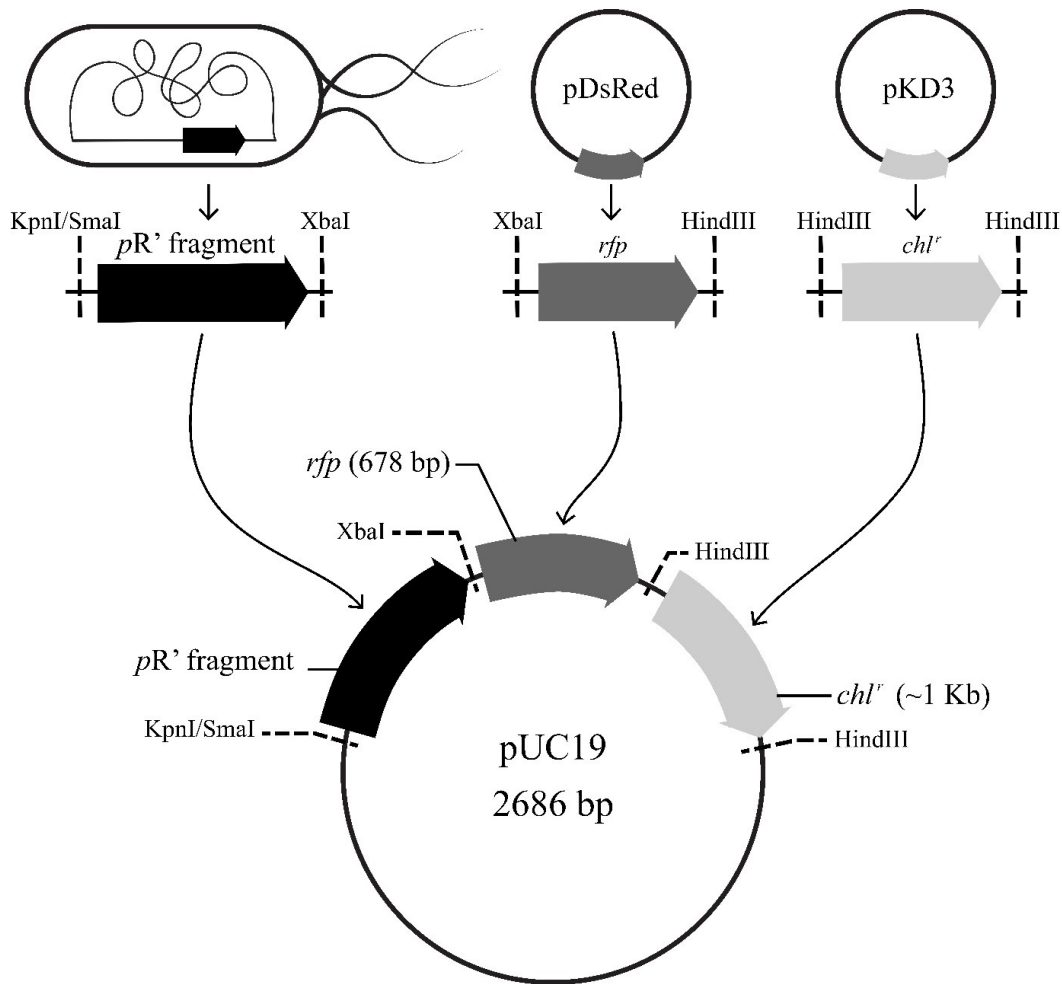
Figure 2.2 The schematic of P*p*R'::*rfp*::*chl^r* construction. Arrows with direction indicate the transcription orientation. Black arrow represents target *p*R' fragment; dark gray is the *rfp* fragment; light gray is the chloramphenicol resistance gene. Dash lines indicates the restriction sites, *p*1402 used restriction enzymes SmaI/XbaI, since sequence of *p*1402 contains the restriction site KpnI. Target fragments *p*R' and *rfp* were transformed into pUC19 vector, followed by a *chl^r* fragment for positive screening.

2.3.1   Amplification of promotor regions for cloning of promotor fusions to RFP

To obtain the target fragment, genomic DNA was isolated from 500 µL overnight culture using a Wizard® Genomic DNA Purification Kit (Promega, Canada) according to the manufacturer's instruction, and was then used as the template in PCR.

To amplify the target fragment with restriction sites from the genomic DNA, PCR conditions were set up as following: initial denaturation step at 94 °C for 1 min 50 s; denaturation step at 94 °C for 15 s, annealing at 54 or 55 °C (depending on primer Tm values) for 30 s and extension at 68 °C for 30~180 s, depending on different fragment sizes (1 minute/kb). No final elongation step was used. The PCR conditions for primers and positive cloning validation were: initial denaturation step at 94 °C for 5 min; denaturation, annealing, extension steps were at 94 °C for 45 s, 54 or 55 °C (depending on primer Tm values) for 30 s, and at 72 °C for 30~180 s (depending on different fragment sizes), respectively. The final elongation step was at 72 °C for 10 min. Denaturation, annealing and extension steps were set into a cycle and repeated 30 times. All PCR amplicons were maintained in thermal cycler at 4 °C after reactions, before separation in an agarose gel or stored at -20 °C. The enzyme for synthesizing target fragments from genomic DNA was Platinum™ *Taq* DNA Polymerase High Fidelity (Fisher Scientific, Ottawa, CA), and for PCR validation was Invitrogen™ *Taq* DNA Polymerase, recombinant (Fisher Scientific, Ottawa, CA).

The products of the PCR reaction were electrophoresis at 100 volts for 60 min on a 1% agarose gel to check the product sizes and visualized by a UV transillumination. The amplified fragments were purified using a GeneJET Gel Extraction Kit (Fisher Scientific, Ottawa, CA) according to the manufacturer's instructions. The purification products were quantified using the NanoDrop™ One/One^C Microvolume UV-Vis Spectrophotometer (NanoDrop, Madison, USA) by measuring light absorption at 260 nm.

2.3.2   Primers for generating target fragments.

Primers designed to generate restriction enzyme cutting sites at the end of each target fragment are listed in Table 2.5. Primer pairs RFP F-2 and RFP R-5 were used to amplify the red fluorescent protein gene *rfp*.

Table 2.5 Primers used for obtaining *p*R' and *rfp* fragments

Primer	Sequence (5'-3') [a]	Restriction site
LP F1-1	5'-CGGGAA<u>GGTACC</u>ACCTCTGTATTTTATCAG-3'	KpnI
LP R1-3	5'-GGGCCG<u>TCTAGA</u>AAAGAAAAAAGTTAGCAC-3'	XbaI
LP F2-2	5'-ATTAGT<u>CCCGGG</u>CTTGGATTTATTGATGGT-3'	SmaI
LP R3-2	5'-ATAACG<u>TCTAGA</u>TAACAGGCACAGTACCCA-3'	XbaI
LP F3-2	5'-AGC<u>GGTACC</u>AAAAACCGGAAACGTGTA-3'	KpnI
LP F4-1	5'-TGCGTA<u>GGTACC</u>AGCGTCTATAATTGTATG-3'	KpnI
LP R4-2	5'-GCATTA<u>TCTAGA</u>CAACAGGCACAGTATCCA-3'	XbaI
RFP F-2	5'- CTGATA<u>TCTAGA</u>ATGGCCTCCTCCGAG -3'	XbaI
RFP R-5	5'- ATCTGT<u>AAGCTT</u>CTACAGGAACAGGTGGT-3'	HindIII

[a] Restriction enzyme sites are underlined.

2.3.3 Cloning target fragment *p*R' and construction of promoter – reporter vector.

To obtain the promoter – receptor vector, all the plasmids were harvested from 5 mL overnight culture by using a QIAprep Spin Miniprep Kit (QIAGEN, Toronto, CA).

The standard cloning techniques was first used to fill contig gaps and obtain the complete sequence of the target fragment. For cloning the target *p*R' fragment into pUC19, 1 µg purified amplicons were subsequently digested with KpnI and XbaI, except the target fragment from strain O103:H25 338 (FUA 1402), which was digested by SmaI and XbaI. For enzymes KpnI and XbaI, reaction mixture was incubated at 37 °C for 1 h and then inactivated at 80 °C for 5 min and 65 °C for 15 min. For enzymes SmaI and XbaI, mixture was incubated at 37 °C for 1 h and inactivated at 65 °C for 20 min. To generate the same cutting site for ligating the *rfp* after the target fragment, 1 µg fragment of *rfp* was digested with XbaI and HindIII by incubated at 37 °C

for 1 h and heated at 80 °C for 10 min and 65 °C for 10 min to inactivate the enzyme. Plasmid pUC19 was also digested by the same enzymes at the same time.

The plasmid constructs with the target fragments were submitted to the Molecular Biology Service Unit (MBSU) at University of Alberta, and verified by Sanger DNA sequencing. Sequencing results were first checked and edited by Chromas (Technelysium, South Brisbane, Australia), and then aligned in Geneious.

To construct the pUC19 (*p*R'::*rfp*) fusion, the digestion products containing target *p*R' were loaded into a 1% agarose gel and run at 100 volts to separate the plasmids and the target fragments. The bands of the target fragments were cut out of the gel by using a scalpel. Gel containing the target fragment was purified using a GeneJET Gel Extraction Kit (Fisher Scientific, Ottawa, CA) according to the manufacturer's instructions. Subsequently, purified target fragments were ligated with the pRFP respectively using the Rapid DNA Ligation Kit (Fisher Scientific, Ottawa, CA). The ligation was performed in a 3:1 molar ratio of vector and insert by using a Rapid DNA Ligation Kit (Fisher Scientific, Ottawa, CA). Ligation mixture was incubated at room temperature for 2 to 3 h and then directly used in downstream transformation. Determination of correct clones were performed by digesting plasmids using the corresponding restriction enzymes.

Since the target mutant strain *E. coli* O104:H4 strain 11-3088 Δ*stx*::*gfp*::*amp*^{r} has ampicillin resistance, a chloramphenicol resistant gene (*chl*^{r}) was introduced into the *p*R'::*rfp* construct, positioned after the *rfp* gene, with the restriction site of HindIII at its both ends for positive cloning screening. The primers amplifying *chl*^{r} gene and for screening the positive cloning are shown in Table 2.6.

Table 2.6 Primers for cloning and checking chloramphenicol resistant gene.

Primer	Sequence (5'-3') [a]	Reference
priming site 2 HindIII [b] [d]	5'- ACGTC<u>AAGCTT</u>ATGGGAATTAGCCATGGTCC -3'	(Datsenko and Wanner 2000)
priming site 1 HindIII [b]	5'- ACTGA<u>AAGCTT</u>GTGTAGGCTGGAGCTGCTTC -3'	(Datsenko and Wanner 2000)
Chl HindIII F2 [c] [d]	5'- ACTATT<u>AAGCTT</u>TTCGGCGCGCCTACCTGT -3'	This study
Chl HindIII R3 [c]	5'- GCCTGT<u>AAGCTT</u>CGGAATAGGAACTTCATT -3'	This study
M13/pUC R [d]	5'- AGCGGATAACAATTTCACACAGG -3'	This study

[a] Restriction enzyme sites are indicated by underline.

[b] Primers amplifying *chl^r* gene for inserting into the P*p*R'*::rfp* constructs except *p*1402.

[c] Primers amplifying *chl^r* gene for inserting into the P*p*1402*::rfp* construct.

[d] Primers for checking the direction of the inserted *chl^r* gene.

2.3.4 Chemical transformation and electroporation.

For chemical transformation, *E. coli* DH5α cultures were grown overnight in 25 mL LB broth and then diluted 1:100 into 100 mL LB broth and grown at 37 °C with 200 rpm agitation the next morning. When OD_{600} reached 0.4~0.6, the 100 mL culture was divided into four portions and put on ice for 10 min. Each portion of the cultures was harvested by centrifugation at 5000 rpm for 2 min, resuspended in 15 mL 80 mM $MgCl_2$-20 mM $CaCl_2$ and chilled on ice for 30 min. After 30 min incubation on ice, the cultures were centrifuged at 5000 rpm for 5 min and resuspended with 500 µL 0.1 M $CaCl_2$-15% glycerol. Cells were distributed 50 µL per aliquot, and used immediately or stored at -80 °C.

Chemical transformation was done by mixing 5 µL of the ligation reaction mixture with 50 µL DH5α chemical competent cells. After 30 min incubation on ice, the sample was heat shocked in a 42 °C water bath for 30 s, and then immediately transferred the sample to the ice for 5 min. LB broth (950 µL) were added and the bacteria were allowed to recover for 1 h by growing at 37 °C with 200 rpm agitation in a shaker.

Electroporation was used to transform the plasmids into wildtype EHEC strains. Bacterial cultures were grown overnight to recover in LB broth with antibiotics, diluted to an OD_{600} of 0.05 in 5 mL of fresh LB broth and grown at 37 °C with 200 rpm agitation until OD_{600} reached 0.4~0.6. Cells at OD_{600} 0.4~0.6 were put on ice for 15 min and centrifuged at 4 °C for 5 min, and washed three times with 1 mL precooled 10% glycerol with centrifuging at 4 °C for 1 min between each time. After the last centrifuging, cells were resuspended by 500 µl precooled 10% glycerol.

Electroporation was done in Gene Pulser®/MicroPluler™ Electroporation cuvettes (Bio-Rad, CA). Cuvettes were chilled on ice for at least 15 min to cool thoroughly before use. Plasmid DNA (80~100 ng) was added into 50 µl competent cells and then transferred into the electroporation cuvette. The cell and plasmid mixtures were electroporated using a Bio-Rad Gene Pulser™ pulse generator at 25 µFD, 200 Ω, 2.5kV. Electroporated cells were transferred into 950 µl room temperature LB broth and recovered at 37 °C with 200 rpm agitation for 1 h. Recovered cultures and their 1:10, 1:100 dilutions were spread on selective LB agar plates and incubated overnight at 37 °C.

2.3.5   Screening for positive clones.

For screening of successful transformation of target fragments, colonies from selective media were grown in 5 mL LB broth with antibiotic, and then plasmid DNA was isolated and digested by using restriction enzymes and checked by agarose gel electrophoresis. To determine the direction of *chl*^{r} gene, PCR was carried out on the plasmid from the positive clones using shown in Table 2.6

2.4   Reporter system validation.

To validate the fluorescence gene fusion reporter system, the reporter constructs were validated by visualizing the cell image under the Axio Imager microscope (ZEISS, Toronto, CA), combined with multi-channel fluorescence imaging, UV-lamp mbq 52ac and filter for Rhodamine and GFP. Cells were grown in LB with 0.5 µg/mL final concentration MMC (M0503-2MG, Millipore Sigma, CA) for 4 h, and then observed with a 10× or 40× objective lens and a 10× ocular. Pictures were captured by an AxioCam M1m 385 camera and viewed by Axio Vision software (v.4.8.2.0, ZEISS, Toronto, CA).

2.5   Phenotype analysis.

2.5.1   Bacterial strains and growth conditions.

E. coli O104:H4 strain 11-3088 Δ*stx*::*gfp*::*amp*^{r} with P*p*R'::*rfp*::*chl*^{r} constructs and selected EHEC strains with P*p*R'::*rfp*::*chl*^{r} construct used in toxin expression analysis are listed in Table 2.7.

Single colonies from each strain were inoculated in LB broth with antibiotics at 37 °C with agitation at 200 rpm overnight (12 h~16 h). The following morning, OD_{600} measurements were taken and cultures diluted to a uniform OD_{600} reading of 0.05 in separate 15 mL screw capped tubes. Each strain was subcultured into two tubes, one was taken as untreated control and to check OD_{600}, and the other one was for MMC induction.

OD_{600} values were checked from 1 mL culture in a 1.5 mL disposable plastic cuvette (Fisher Scientific, Ottawa, CA) with an Ultrospec 100 pro Visible spectrophotometer (Biochrom, Ltd., USA).

Table 2.7 P*p*R'-*rfp* transformed strains used in phenotype analysis.

Strains	Constructs
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1302::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1303-s1::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1303-2a::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1306::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1309-1c::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1309-2d::*rfp*::*chl^r*
E. coli O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*	P*p*1311::*rfp*::*chl^r*
FUA 1302, *E. coli* O104:H4 11-3088	P*p*1302::*rfp*::*chl^r*
FUA 1303, *E. coli* O157:H7 1935	P*p*1302::*rfp*::*chl^r*
FUA 1303, *E. coli* O157:H7 1935	P*p*1303-s1::*rfp*::*chl^r*
FUA 1303, *E. coli* O157:H7 1935	P*p*1303-2a::*rfp*::*chl^r*
FUA 1311, *E. coli* O45:H2 05-6545	P*p*1302::*rfp*::*chl^r*
FUA 1311, *E. coli* O45:H2 05-6545	P*p*1311::*rfp*::*chl^r*
FUA 1399, *E. coli* O157:H7 CO6CE900	P*p*1302::*rfp*::*chl^r*
FUA 1399, *E. coli* O157:H7 CO6CE900	P*p*1399-28::*rfp*::*chl^r*
FUA 1399, *E. coli* O157:H7 CO6CE900	P*p*1399-79::*rfp*::*chl^r*

2.5.2 Determination of the temperature required for inactivation of strains of *E. coli*

To minimize the bias caused by cell lysis, a heat inactivation was performed after MMC induction. The treatment condition was determined by the maturation temperature of DsRed.

The red fluorescent protein DsRed has a maturation time of up to 48 h; to arrest phage maturation after expression of the late genes, cells were inactivated at a temperature that does not interfere with protein folding and maturation. Strains classified as highly heat resistant exhibited a decimal reduction time (*D*-values) of more than 6 min at 60 °C (D$_{60}$-value), while most *E. coli* strains have a D$_{60}$-value less than 1 min (Liu et al. 2015).

A time course experiment was performed to confirm the time for heat inactivation before the cell number decreased dramatically. When cells reached an OD$_{600}$ of 0.4~0.6, cultures were separated into two portions, with 600 µl each portion. One portion was treated with MMC at a final concentration of 0.5 µg/mL and the controls were cultures without MMC. Both portions were incubated at 37 °C with 200 rpm shaking. Aliquots of 600 µl were collected every 30 min from 3 h to 5.5 h after induction, and then heated at 60 °C for 15 min using a thermal cycler. Heated

cultures were then incubated at 37 °C for 7 h and 27 h, and incubated at 4 °C for 27 h before measurement of fluorescent cell population. Fluorescence signals obtained from those three treatments were compared to confirm the heating time and the incubation time.

The inoculation and incubation to determine the expression of the fluorescent protein under different promoter regulation are the same as the steps in heat inactivation. Cultures were induced 0.5 µg/mL MMC and incubated at 37 °C. After 4.5 h after induction, 100 µL cultures were heated at 60 °C for 15 min in a thermal cycler. Cultures were then incubated at 37 °C for 27 h before flow cytometry analysis.

2.5.3 Flow cytometry analysis.

Expressions of fluorescent protein were studied 27 h after MMC induction (22.5 h after heating inactivation). Samples were resuspended by pipetting and then diluted 1:100 in 1 mL 1×PBS (pH 7.4). A LSRFortessa™ X-20 cell analyzer (Biosciences, Mississauga, CA) was used to perform the cell analysis. Fluorescence was excited with a 488 nm Argon ion laser and followed by a 530/30 - 575/26 nm bandpass filters, and finally detected by side scatter detectors and forward scatter detector. Data was recorded by BD FACSDIVA™ software (BD Biosciences, CA) and analyzed by FlowJo (BD Biosciences, CA). Single cell population was defined by selecting the cell population located along the diagonal of the dot plot. Gating was set as "cells of favorite" 100% of the singlets (Y: SSC-A; X: FSC-A), since some of the fluorescent cells were filamented and thus had different SSC-A/FSC-A value versus the rest.

2.5.4 Statistical analysis

The experiments were repeated at least three separate times (biological replicates). Statistical analysis was performed with SigmaPlot (v.12.5., Systat Software Inc., UK) using one-way analysis of variance (ANOVA). A p-value of ≤0.05 was considered statistically significant.

3 Results

3.1 Confirmation of the *stx* genotype of *E. coli* FUA 1304 and pairing of contigs with fragments of the two Stx prophages in the genome of *E. coli* FUA 1399.

According to a previous study, *E. coli* FUA 1304 carries a *stx* gene in its bacterial genome. However, this *stx* gene was not found in the *E. coli* FUA 1304 gene file. To confirm the presence of *stx* in *E. coli* FUA 1304 genome, a *stx* gene screening was performed by PCR. Agarose gel electrophoresis of PCR products from *E. coli* FUA 1304 revealed that there is no *stx* prophage inserted in *E. coli* FUA 1304 genome (Figure 3.1A).

The PCR amplification carried out in *E. coli* FUA 1399 was to confirm the correct contig pairs, since its two *stx2a* genes are located in contigs with their regulation regions. According to the electrophoresis of *E. coli* FUA 1399 PCR products amplified by different primer pairs, only when LEAE 28-182 F is paired with LEAE 28 R, and LEAE 104 F is paired with LEAE 179 R-1, can amplicons be produced (Figure 3.1B). This result showed that *E. coli* FUA 1399 harbors two *stx2a* genes, which are under different regulations controls. Thus, this was the only strain used in this study which has two same *stx2a* genes.
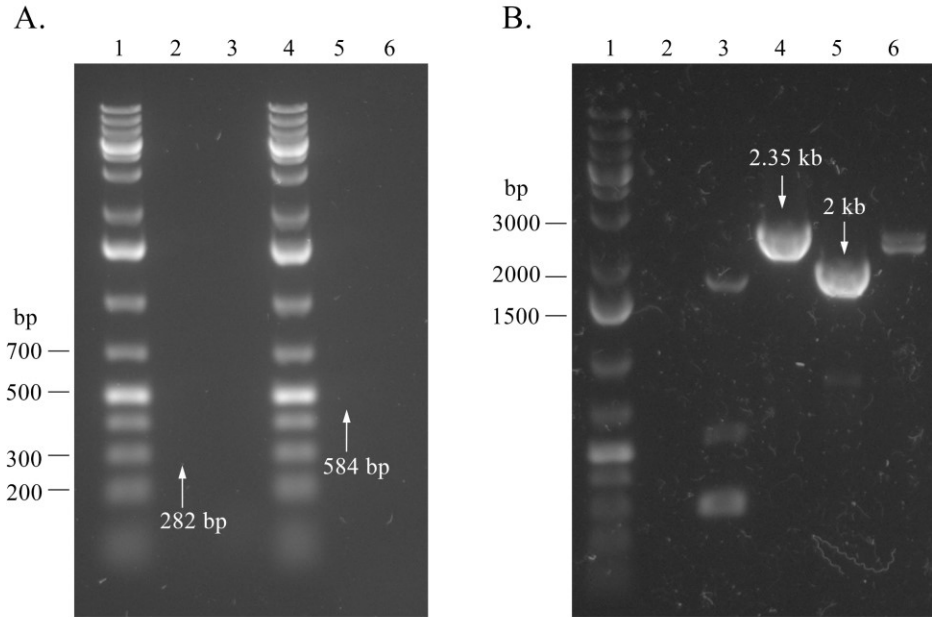
Figure 3.1 Agarose gel electrophoresis of *E. coli* FUA 1304 *stx* genotyping and *E. coli* FUA 1399 contig pairing. (A) PCR products of *E. coli* FUA 1304 synthesized by primer pairs confirming *stx1* and *stx2*, respectively. Lane 1 and 4 are 1 Kb plus DNA ladders; Lane 2 is the product synthesized by primer pairs KS7 and KS8; Lane 4 is the PCR product amplified by primer pairs LP43 and LP44; Lane 3 and Lane 6 are negative controls. (B) PCR products synthesized by four combinations of the following two forward and reverse primers: LEAE 28-182 F, LEAE 28 R, LEAE 104 F, and LEAE 179 R-1. Lane 1 is the 1 Kb plus DNA ladder; Lane 2 is the negative control; Lane 3 – 4 are PCR products amplified by LEAE 28-182 F/ LEAE 179 R-1, LEAE 28-182 F/ LEAE 28 R, LEAE 104 F/ LEAE 179 R-1, LEAE 104 F/ LEAE 28 R, respectively.

3.2   Diversity of late regulation region.

Strains analyzed in this study were isolated from cattle, human, clinical and unknown sources (Liu et al. 2015). There were 26 sequences of either target *p*R' regions or *stx* genes from 17 STEC strains aligned via Geneious Multiple Alignment tool (Figure 3.2).

The comparison of *p*R' regions and *stx* genes revealed that although the same subtype *stx2* genes share overall sequence similarities to each other, the late regulation *p*R' regions were more divergent even if they are from the same *stx* gene subtype (Figure 3.2A).

Most of the *stx2* sequence differences in the *p*R' regions were caused by single nucleotide changes and not the insertion of a whole flanking region. The target *p*R' fragments from FUA 1306 and FUA 1399, contig. 28 had more diverse nucleotide changes than the rest. The structure of two *p*R' sequences from *stx2d* prophages were more different compared with the *p*R' region from *stx2a* prophages. This indicates that the same *stx2* subtype may have different toxin expression and production levels due to sequence diversity of their regulation region. However, in *stx1* subtypes, although the prophage from FUA 1402 harbors a large *p*R' region, sequences were overall more conserved compared to the region in *stx2*.

3.3   Phylogenetic analysis of target *p*R' fragments of and *stx* genes.

To determine the genomic relationship, A phylogenetic analysis was performed based on the result of alignment by the program Geneious. The gene structures of *p*R' regulation region were more diverse in Stx2 prophages than in Stx1 prophages (Figure 3.3A). The result of the *stx* gene phylogenetic analysis revealed that the *stx* genes with the same subtype were located in the same clade (Figure 3.3B). Although *stx1* and *stx1c* located into two separate clades, genes belonging to the *stx2* subtypes were all in the same branch.

Stx2 is reported to cause more severe human disease (Friedrich et al. 2002). The *stx2a* gene located in FUA 1306 and FUA 1399 contig no. 28 share similar *stx* gene sequences with other *stx2a* (Figure 3.3B), however, these two prophages possess *p*R' regions which differ from other Stx2a prophages (Figure 3.3A). Thus, it is interesting to determine if they have different toxin production level compared to the other Stx2a prophages. The *stx2d* gene possessed by FUA 1309 and FUA 1310 exhibited a high phylogenetic similarity with *stx2a* prophages, while the similarities were lower when comparing their *p*R' regions. Moreover, difference in *p*R' regions were observed even in strains carrying identical *stx* genes, *e.g. E. coli* FUA 1309 and FUA 1310.

Figure 3.2 The sequence comparison of target *p*R' regions and *stx* genes. Consensus shown on top. Sequence identities are colored by green, yellow, and red, which represent the residue at the position is the same across all sequences, less than complete identity and very low identity, respectively. Identical sequences are colored in gray, while different residues are colored in different colors based on different nucleobase type. Dash line means no nucleotide in that position. Black boxes in the sequences are gaps produced during sequencing. (A) Alignment of the 26 target *p*R' fragments which start from the last 42 bp of *q* gene to the first 13 amino acid of *stx* gene; (a) and (b) are two groups which the internal candidates share similar gene structures between each other. (B) Alignment of 26 *stx* genes, which include the sequences encoded subunit A and B. The figure is provided in high resolution for large scale printing or viewing.

37
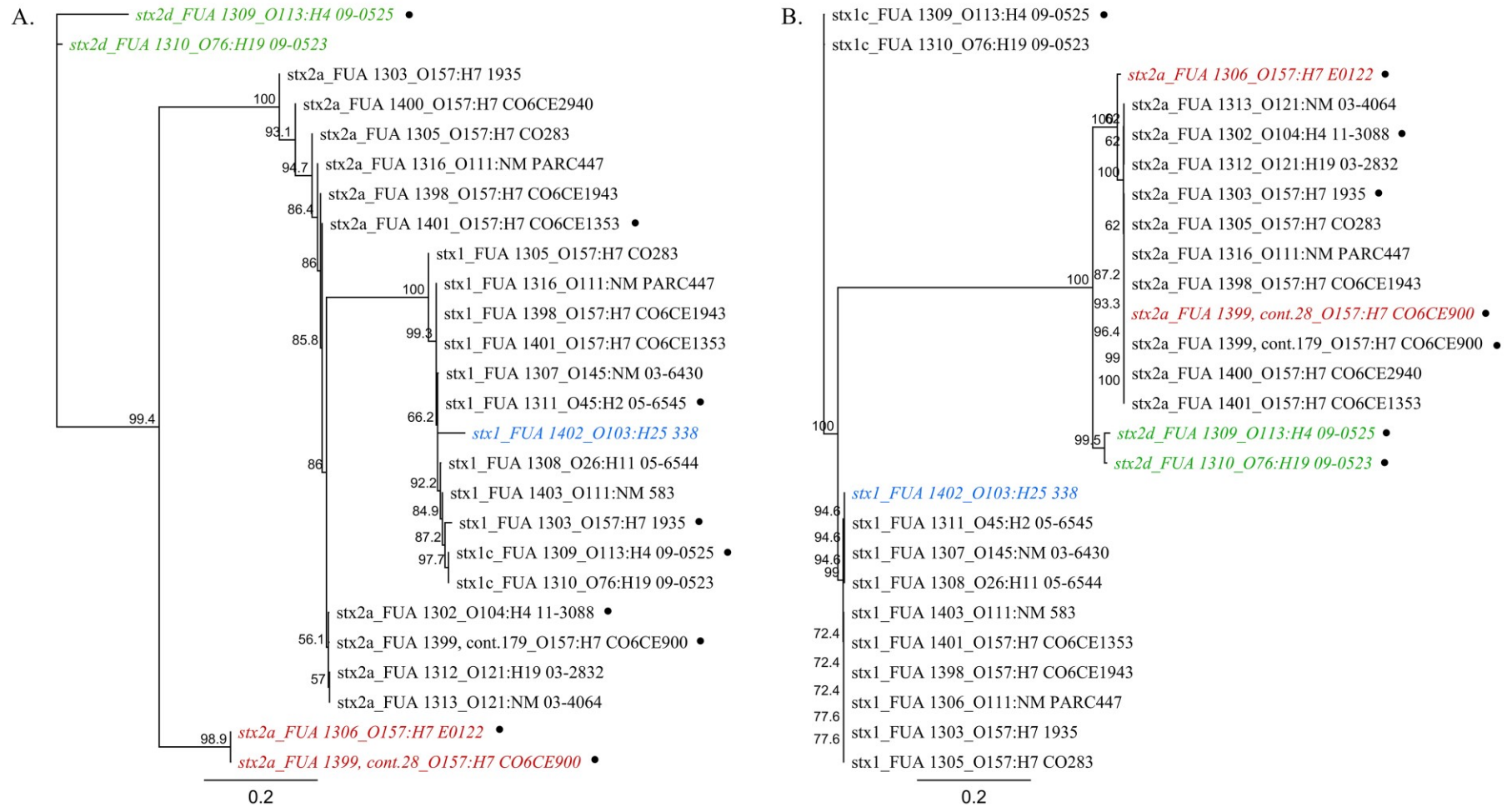
Figure 3.3 Phylogenetic tree from target _p_R' fragments and _stx_ gene sequences. Based on 26 sequences from 17 STEC strains investigated in this study. Trees were generated in Geneious, using the defaulted Tamura-Nei model with the tree build method of Neighbor-Joining. The reliabilities of the internal branches were assessed using bootstrapping with 1,000 pseudo-replicates. The scale bars represent the number of the substitution per site. Strains which have significant phylogenetic differences between the target pR' region and stx gene are highlighted by colors and dots. (A) Phylogenetic tree generated by aligning the regulation region between gene q and stx, which started from the last 42 bp of q and stopped at the thirteenth amino acid of Stx. (B) Phylogenetic tree generated by comparing the stx genes, which including both subunit A and B.

3.4 Construction of P*p*R'::*rfp*::*chl^r* transcriptional fusion, and transformation into target strains.

The promoter region directly upstream of *stx* gene started from the *qut* site, which is located immediately behind *Q* gene (Casjens and Hendrix 2015). To make sure all the late regulating elements were included, promotors were amplified to start from the last 42 bp of *Q* gene and stop at the thirteenth amino acid of Stx. A total of 16 target *p*R' fragments (Figure A.1; Table A.1) were selected according to gene analysis, cloned into the pUC19 (*rfp::chl^r*) vector, and subsequently cloned into DH5α for plasmid maintenance (Figure A.2). Afterward, these 16 P*p*R'::*rfp*::*chl^r* vectors were transformed into different target strains for two different purposes: 1) To study toxin production under the control from the same regulation system, they were transformed into the mutant strain *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*; 2) To compare toxin production under the control of different regulation systems, the *p*R' constructs were transformed into their native strains and *E. coli* O104:H4 strain 11-3088 Δ*stx*::*gfp*::*amp^r*.

3.5 Reporter system validation.

The expression of red and green fluorescent proteins by transformants of *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (P*p*1302::*rfp*::*chl^r*) and *E. coli* O157:H7 CO6CE900 (P*p*1302::*rfp*::*chl^r*) were examined under a microscope (Figure 3.4). E*. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* was used as control. Cell cultures were induced with 0.5 µg/mL MMC. After 4 h incubation at 37 °C, filamentation was observed in all three strains, which indicates that all three strains can be induced by 0.5 µg/mL MMC. In the absence of P*p*1302::*rfp*::*chl^r*, the control strain, *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*, only showed green fluorescence, whereas in strain *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (P*p*1302::*rfp*::*chl^r*), both green and red fluorescence can be observed. Red fluorescence was observed only in transformants *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (P*p*1302::*rfp*::*chl^r*) and *E. coli* O157:H7 CO6CE900 (P*p*1302::*rfp*::*chl^r*), which harbor the *p*R'::*rfp* transcriptional constructs. It can be concluded that the target regulation fragment *p*R' in pUC19 can be controlled by the host prophage regulation system. Moreover, a higher fluorescent cell population was observed in *E. coli* O157:H7 CO6CE900 (P*p*1302::*rfp*::*chl^r*) when compared with *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (P*p*1302::*rfp*::*chl^r*).
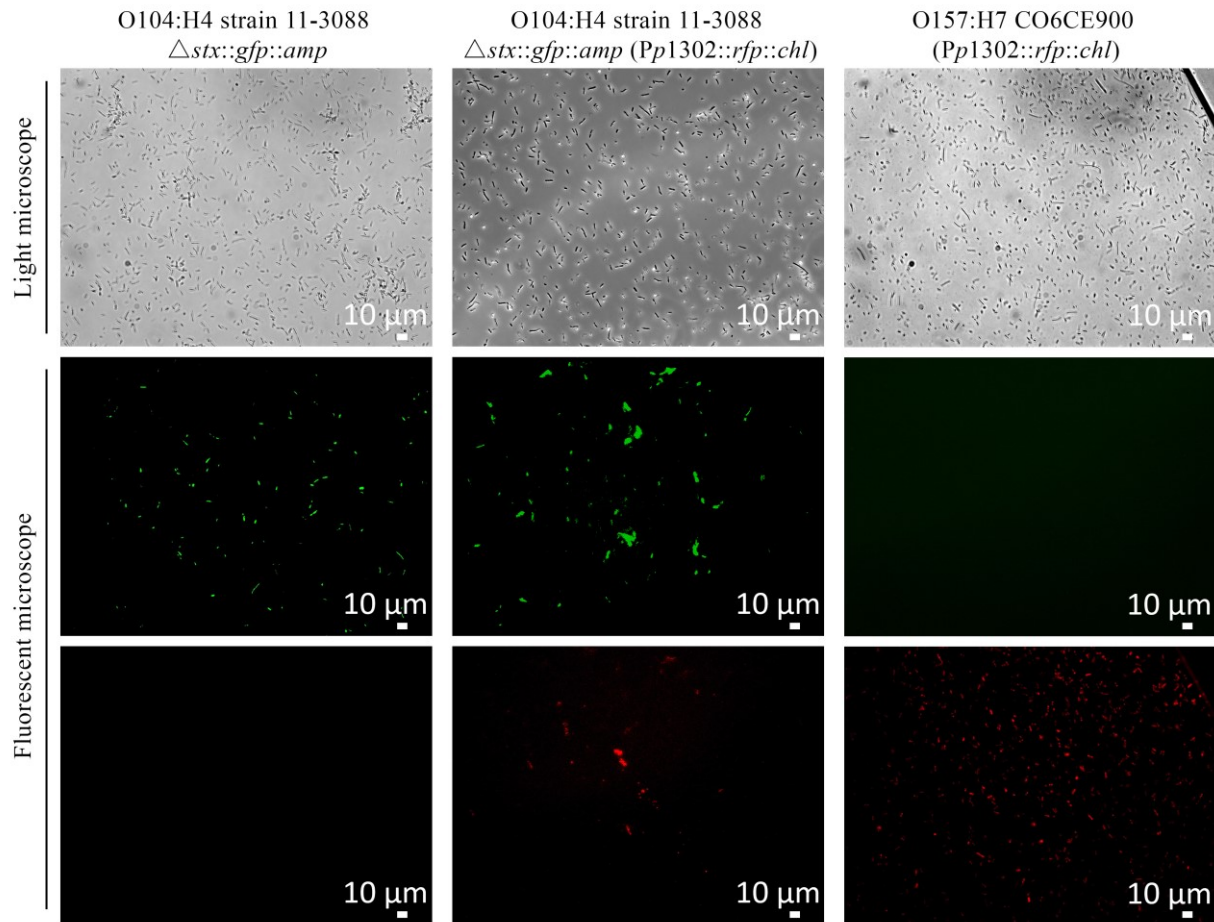
Figure 3.4 Images of the transformant STEC (*p*R':*:rfp*::*chl^r*) after MMC induction visualized by light and fluorescent microscopy (400× magnification). To validate the *p*R':: *rfp*::*chl^r* reporter system, construct *p*1302::*rfp*::*chl^r* was transformed into *E. coli* O104:H4 11- 3088 Δ*stx*::*gfp*::*amp^r* and *E. coli* O157:H7 CO6CE900, respectively. MMC induction was performed 4.5 h before microscopy observation. Images of the green or red fluorescent cells were detected by fluorescent microscope.

3.6   Efficiency of phage induction.

3.6.1   Determination of heat inactivation time.

Due to the lysis of the host cell, induction by MMC causes a decrease in culture density. Because DsRed matures slowly, cell lysis interfered with the quantification of DsRed expression by flow cytometry. To minimize the bias produced by cell lysis, cells were killed at 60°C to stop lysis. Typically, cell counts starts to decrease dramatically 3 to 5 h after induction with MMC, and the cell environment is not necessary for fluorescent protein folding and maturation (Macdonald, Chen, and Mueller 2012). Thus, a time course experiment was performed with *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (P*p*1302::*rfp*::*chl^r*) and MMC induction to determine the heating time for thermal inactivation of cells. The population quantification was done by flow cytometric analysis (Figure 3.9).

The slow development of DsRed limits the intensity of its signal, especially in fast growing organisms, such as yeast (Baird, Zacharias, and Tsien 2000). At room temperature, DsRed protein shows half of its maximal fluorescence after approximately 27 h and requires more than 48 h to reach > 90% of maximal fluorescence. Although it is slower in maturation than other fluorescent proteins, it is resistant to pH extremes and photobleaching. It was observed that when samples were heated 4.5 h after induction, RFP levels were more consistent compared with other sampling times (Figure 3.5). When comparing the incubation temperatures, samples incubated at 37 °C showed higher RFP intensity than at 4 °C. Moreover, the differential of RFP levels was more pronounced between 7 and 27 h after induction. It was determined based on the above results that, although higher RFP levels were observed when heated 5 h after induction, 4.5 h after induction was chosen as the thermal inactivation time and incubation time was 27 h at 37 °C. Flow cytometry data was obtained as a dot plot; gating of flow cytometry data was set to include 99.5% of cells of the negative control analyzed on the same day (**Error! Reference source not found.**).
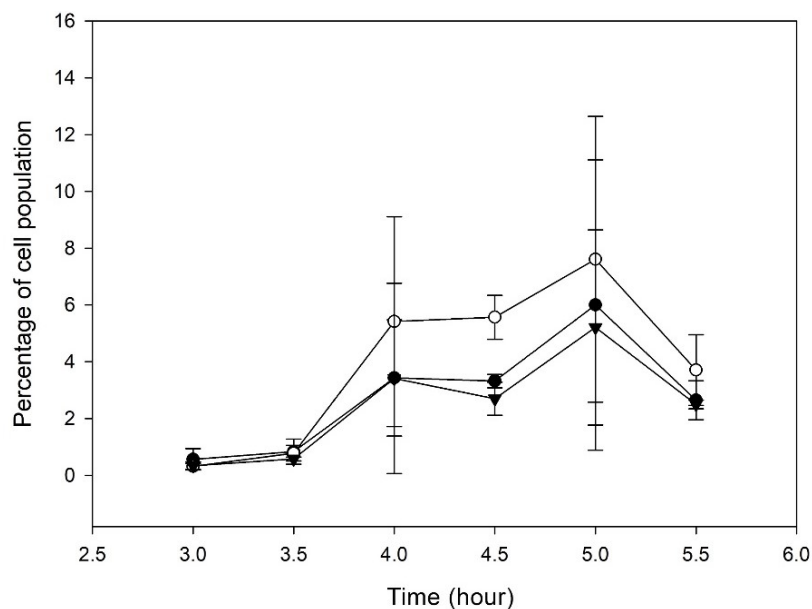
Figure 3.5 Percentage of cell population expressing RFP after induction with mitomycin C.
To reduce the bias of cell population caused by cell lysis, a time course experiment was performed to determine the time of heat inactivation. Samples were heated at different time after MMC induction. Y axis is the percentage of fluorescent cell population, X axis is the different time point for heat inactivation after induction. *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* (Pp1302::*rfp*::*chl^r*) cells were incubated in different conditions and then incubated at 37 °C for 7 h (●), 4 °C for 27 h (▼), 37 °C for 27 h (○), respectively.
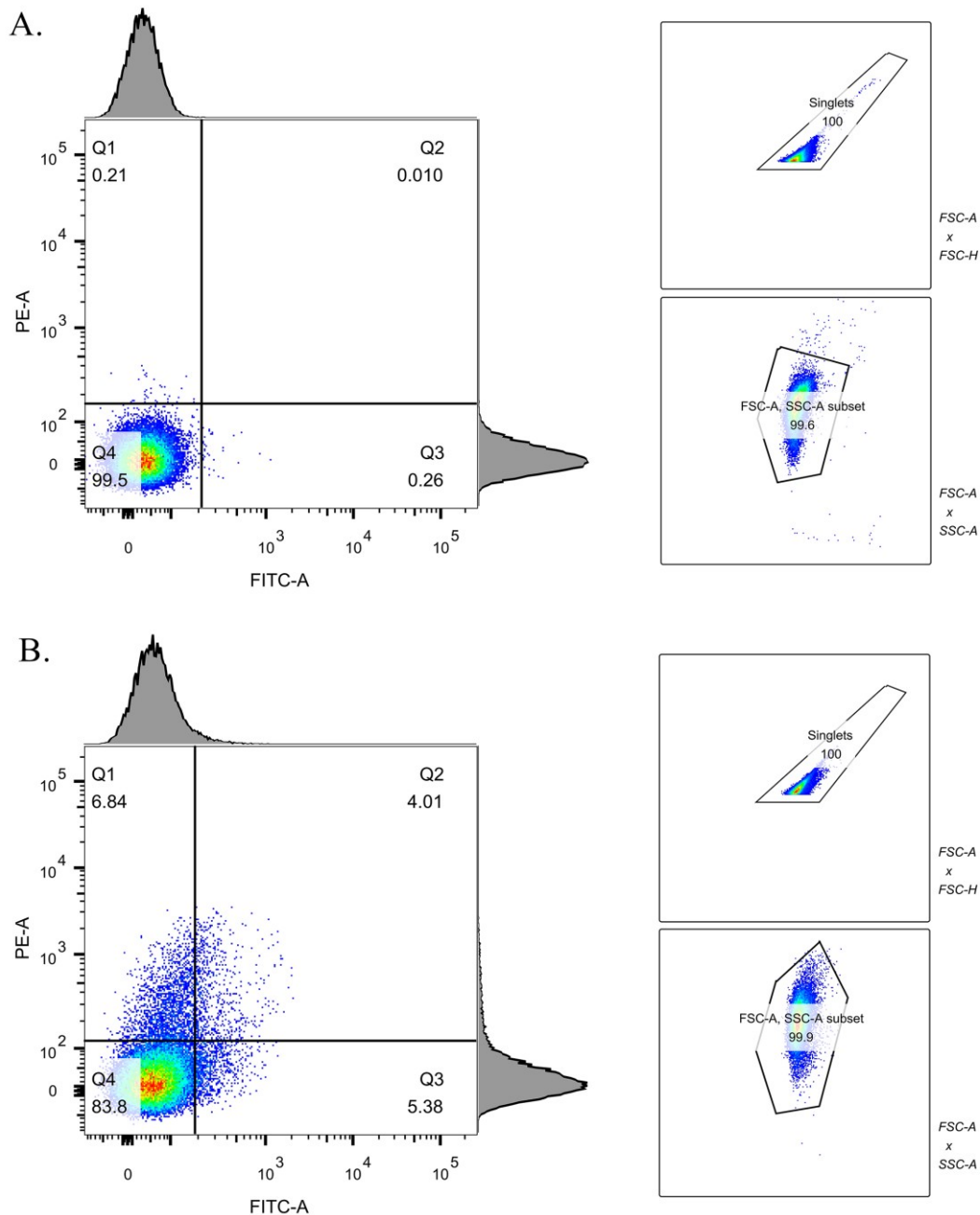
Figure 3.6 The gating of _E. coli_ O104:H4 11-3088 Δ_stx_::_gfp_::_amp_[r] (_p_1302::_rfp_::_chl_[r]) with or without MMC induction. Rectangular gates graphs are cell populations divided based on the fluorescence signal which the cell possessed: Q1 to Q4 represent the cell populations showing RFP, both RFP and GFP, GFP, and FP negative, respectively, with the percentage numbers under them. Small graphs beside the four rectangular gates graphs are the gating strategy for bacterial populations: "FSC-A × FSC-H" defined the single cells; "FSC-A × SSC-A" defined the final cell population for investigating. Gating of flow cytometry data was set to include > 99.5% of cells of the negative control analyzed on the same day. Panel A: Dot plot of the negative control without MMC induction. Panel B: Dot plot of the sample induced with MMC for 4.5 h.

3.6.2 Comparison of different promotors in the same reference strain, *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*.

To determine if foreign *p*R' are induced by the *E. coli* O104:H4 11-3088 prophage, fluorescent protein expression was quantified at the single-cell level by flow cytometer analysis 27 h after MMC induction. Based on the results of sequence alignment, the fusion constructs chosen for this experiment were P*p*1302::*rfp*::*chl^r*, P*p*1303-s1::*rfp*::*chl^r*, P*p*1303-2a::*rfp*::*chl^r*, P*p*1306::*rfp*::*chl^r*, P*p*1309-1c::*rfp*::*chl^r*, P*p*1309-2d::*rfp*::*chl^r*, P*p*1311::*rfp*::*chl^r*, and P*rfp*::*chl^r* was taken as control. The expressions of promoter constructs varied greatly (Figure 3.7). Under the control of the same *E. coli* O104:H4 11-3088 prophage, promoter constructs *p*1302::*rfp*::*chl^r*, *p*1303-2a::*rfp*::*chl^r* and *p*1306::*rfp*::*chl^r* had larger red fluorescent cell populations, followed by *p*1309-2d::*rfp*::*chl^r and p*1311::*rfp*::*chl^r*. Contrary to those transformants, promoter constructs *p*1303-s1::*rfp*::*chl^r* and *p*1309-1c::*rfp*::*chl^r* showed little ability to be promoted by the host prophage transcriptional regulation. Meanwhile, GFP expression among transformants was not different (Figure 3.8). This suggests that the expression of the chromosomal *gfp*, which represents the expression of the host *stx*, is not influenced by the plasmid.
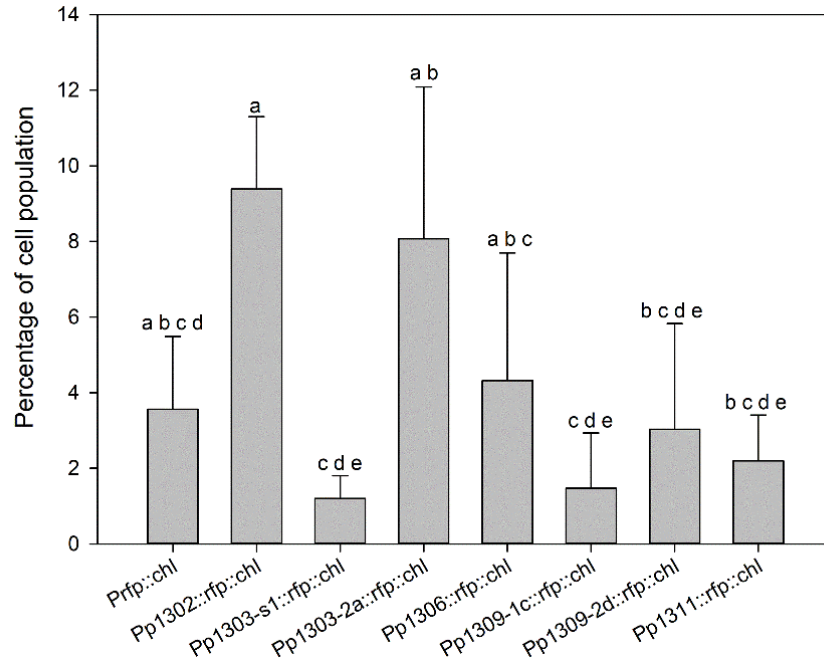
Figure 3.7 The red fluorescent cell populations of *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*
(P*p*R':*:rfp*::*chl^r*) transformants after MMC induction. Raw data was obtained from flow
cytometric analysis. The percentage of fluorescent cell population are shown as mean ± standard
deviations of quadruplicate independent experiments. Bars share a common superscript are not
significantly different ($p \leq 0.05$).
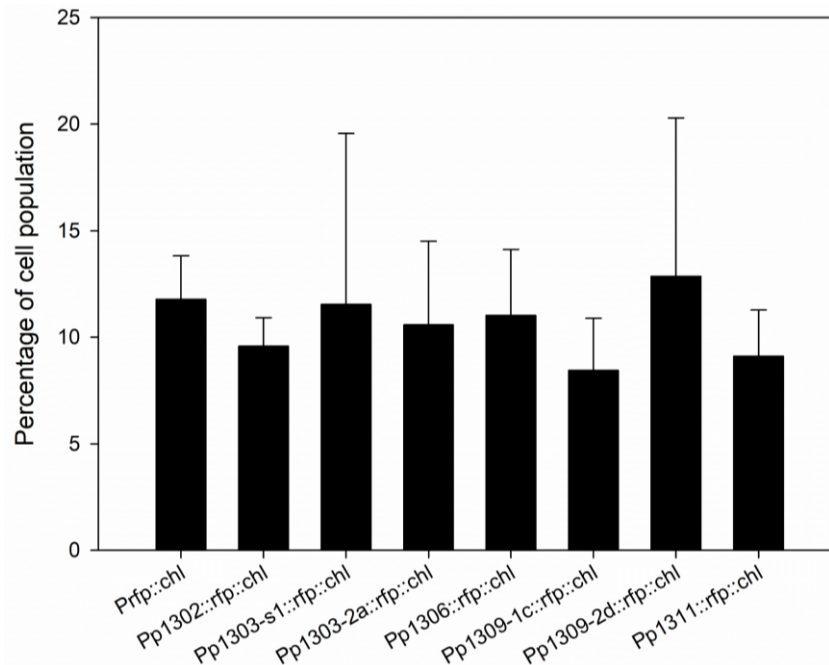
Figure 3.8 The green fluorescent cell populations of *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r*
(P*p*R':*:rfp*::*chl^r*) transformants induced by MMC. Raw data was obtained from flow
cytometric analysis. The percentage of fluorescent cell population are shown as mean ± standard
deviations of quadruplicate independent experiments.

3.6.3 Induction levels of STEC (*p*R':*:rfp::chl^r*) transformants.

The behaviors of *p*R' region under the control of different regulations were investigated by flow cytometry. The vector P*p*1302::*rfp*::*chl^r*, P*p*1303-s1::*rfp*::*chl^r*, P*p*1303-2a::*rfp*::*chl^r*, P*p*1311::*rfp*::*chl^r*, P*p*1399-28::*rfp*::*chl^r*, P*p*1399-79::*rfp*::*chl^r*, from strain FUA 1302, 1303, 1311, 1309, were transformed into their native strain. The vector P*rfp*::*chl^r* was chosen as control and transformed into strain FUA 1302. For strains possessing two more *stx* prophages, the toxin expression of each *stx* prophage can be quantified individually. Strains were induced by 0.5 µg/mL MMC for 4.5 h before heat inactivation and cell culture were quantified 27 h after MMC induction.

The P*p*1302::*rfp*::*chl^r* vector in different host strains behaved differently. In *E. coli* FUA 1303, the percentage of red fluorescent cell population was similar to P*p*1302::*rfp*::*chl^r* in its native host *E. coli* FUA 1302. However, the red fluorescent cell population of induced *E. coli* FUA 1311 (P*p*1302::*rfp*::*chl^r*) was higher than that in the native. Notably, it showed a significantly larger red fluorescent cell population when *E. coli* FUA 1399 (P*p*1302::*rfp*::*chl^r*) was induced than *E. coli* FUA 1302 (P*p*1302::*rfp*::*chl^r*).

The percentage of fluorescent cells in *E. coli* FUA 1311 (P*p*1311::*rfp*::*chl^r*) was lower than *E. coli* FUA 1311 (P*p*1302::*rfp*::*chl^r*). Similarly, the induction of P*p*1302::*rfp*::*chl^r* by the prophages of *E. coli* FUA 1303 lead to a larger RFP positive population when compared to that of the native P*p*1303-s1::*rfp*::*chl^r* and P*p*1303-2a::*rfp*::*chl^r*. Although prophages in *E. coli* FUA 1303 share the same regulation systems, induction efficiencies were different: the percentages of red fluorescent cell population of *E. coli* FUA 1303 (P*p*1303-2a::*rfp*::*chl^r*) was higher than *E. coli* FUA 1303 (P*p*1303-s1::*rfp*::*chl^r*). Variation of induction efficiency was also observed in *E. coli* FUA 1399 transformants. Noteworthy, induced population in *E. coli* FUA 1399 (P*p*1399-79::*rfp*::*chl^r*) cells was notably larger than *E. coli* FUA 1399 (P*p*1399-28::*rfp*::*chl^r*).
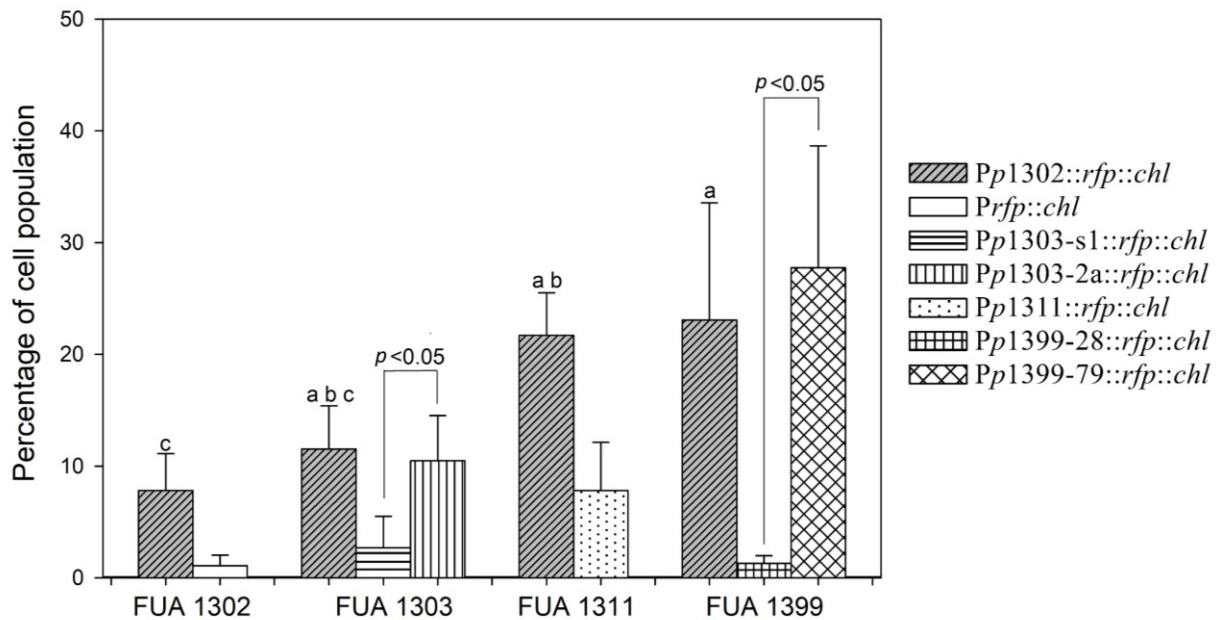
Figure 3.9 Comparison of the percentage of cell population of same promotors in different strains and promotors in their parent strains. The pR'::rfp::chl^r constructs, which cloned from the target strains, were transformed back into their parent strains. At the same time, the construct p1302::rfp::chl^r was transformed into those target strains and also its parent strain *E. coli* FUA 1302 O104:H4. Transformants were induced by MMC. In the graph, four bar groups represent four target strains, bars represent different pR' constructs; the very left bar in each group represents Pp1302::rfp::chl^r carries the pR' region from *E. coli* O104:H4 11-3088 (FUA 1302). The legend aside the graph shows the detail of each vector construct. Bars with the same pattern that do not share a common superscript differ significantly. The percentage of fluorescent cell population are shown as mean ± standard deviations of quadruplicate independent experiments ($p \leq 0.05$).

4 Discussion

Horizontal gene transfer (HGT) is thought to be a major stimulus for evolution of bacterial pathogens (Fortier and Sekulovic 2013). As a common element for gene exchange, bacterial phages exhibit a highly diversity in their genome, and considered as a vector for virulence factor transfer among bacteria, changing their host from non-toxic to pathogenic (Brüssow, Canchaya, and Hardt 2004). Among STECs, *E. coli* O157:H7 is the most common food- and waterborne pathogen. Non-STEC strains from other serotypes, such as *E. coli* O104:H4, a strain of EAEC, have been converted into STEC by gaining the Stx prophages. The emergence of non-O157 STEC challenged public health because they escape detection with tools aiming to detect *E. coli* O157 (Manning et al. 2008; Karch et al. 2012; Muniesa et al. 2012; Hao et al. 2012).

I hypothesized that it is the prophage genome diversity results in behavioral variations in phage induction efficiencies among STEC strains. Previous investigations of Stx genome diversity were mainly focused on functional analysis of open reading frames, insertion sites and virulence plasmids (Strauch et al. 2008; Ohnishi et al. 2002; Ahmed et al. 2012; Tozzoli et al. 2014) but not on investigating the late promoter *p*R' region (Harley and Reynolds 1987). To investigate if the diversity of *p*R' region affects Stx induction, sequences of late promoter region *p*R' of STECs from different origins were analyzed. Two phylogenetic trees were developed according to the late regulation region and the *stx* gene subtypes, respectively. The result of phylogenetic analyses revealed that there is not an apparent phylogenetic correlation between the *stx* gene subtype and Stx prophage late promoter region, indicating that the Stx prophage late region and *stx* gene are genetically heterogeneous. The sequence alignment and phylogenetic analysis also revealed that the *p*R' site, which is located in the target *p*R' region, is identical to the sequence of the reference *p*R' (from AP000400) (Makino et al. 1999) in most of the strains of *E. coli* that were analysed. However, the sequence of *p*R' in *E. coli* FUA 1402, FUA1399-28, FUA1306, FUA1309-2d, FUA 1310-2d was not similar to the reference *p*R', even if searching of the *p*R' site was extended into part of the upstream gene *Q*. As the most important antiterminator gene in the late transcriptional regulation, *Q* has been investigated in many previous studies (Wagner et al. 2001; Ahmad and Zurek 2006), and Q_{933} is considered as a marker of high Stx production (Lejeune et al. 2004). *E. coli* O157:H7 harboring *stx2* under control of Q_{21} rather than Q_{933} may exhibit a Stx2-negative phenotype (Koitabashi et al. 2006). However, in this study, prophages in both *E. coli* FUA 1302 and *E. coli* FUA 1311 harbor the typical *p*R' site and the highly conserved Q_{933}, while induction efficiencies of P*p*1302::*rfp*::*chl*^{r} and P*p*1311::*rfp*::*chl*^{r} were different under the control of the *E. coli* FUA 1311 prophage. It is likely that the determinant of induction efficiency in the late

transcript is not merely the type of Q and pR' site, but also sequence diversity in the late promoter "pR' region".

The cloning vector backbone used in this study was pUC19, which is known as a high copy number plasmid commonly used in molecular biology studies. Inserts slow plasmid replication due to the metabolic burden it brings to the bacteria cell; therefore, the highest copy number was reached with empty plasmids (Lin-Chao and Bremer 1986). Thus, it can be concluded that copy number may affect the production of fluorescent protein, which is reflected by fluorescence intensity. However, the proportion of positive cell are not affected. Promotors used for vector constructs range from ~0.4 to ~2.9 Kb. Pp1303-2a::rfp::chl^r contains a larger insert (~1 Kb) than Pp1303-s1::rfp::chl^r (~0.6 Kb), and thus is supposed to produce less plasmids than Pp1303-s1::rfp::chl^r. However, the proportion of RFP positive cells of *E. coli* O104:H4 11-3088 (Pp1303-2a::rfp::chl^r) was larger than *E. coli* O104:H4 11-3088 (Pp1303-s1::rfp::chl^r). Likewise, the data of the green fluorescent cell population showed that the green fluorescent cell population in the negative control, which carry the smallest size vector, was no statistically difference with the rest transformants, indicating that the GFP fluorescence was not affected by the RFP fluorescence (Hawley et al. 2004; Hawley et al. 2001) or plasmid vector propagation. Therefore, the copy number of the vector does affect the measurement of the proportion of cells with induced prophage. It can be concluded that, by transforming into different STEC strains, the pR'::rfp::chl^r reporter fusion system reflects the regulations from different sources of prophage genomes, no matter from the parent prophage of the pR' or from other foreign prophages.

Bacterial behavior is commonly assessed in bulk (Shimizu et al. 2011; Los et al. 2009), which can merely obtain their performance at the population level. However, stochastic switching between different phenotypic states is a ubiquitous phenomenon prokaryotic cells including bacterial pathogens (King and Masel 2007; Bull 1987). This cannot be observed by colonies or cultures even if they were derived from a single bacterial clone (Moxon et al. 1994). Flow cytometry allows to study the bacterial physiological responses in single-cell level (Shapiro 2000). In this study, microscopy and flow cytometry were employed to visualize and to assess the induction efficiencies of *stx* gene at a single-cell level. Moreover, by fusing promotor regions with different fluorescent proteins, GFP and RFP, the behaviors of chromosomal and foreign promoters under the same regulator's control were distinguished through multi-color detection simultaneously (Oi 1982). Although microscope can also distinguish single cells from cell aggregates, flow cytometry simplifies the process and avoids human bias during measuring. Moreover, flow cytometry is also an ideal way to provide us physiological information, like DNA content and distributions (Shapiro 2003).

Thermal death time of highly heat resistant strains of *E. coli* can be more than 6 min at 60 °C (Hauben et al. 1997; Dlusskaya, McMullen, and Gänzle 2011), However, all strains of *E. coli* that were used in this study are heat inactivated in less than 1 min at 60 °C (Mercer, Zheng, Garcia-Hernandez, Ruan, Ganzle, et al. 2015). In addition, red fluorescent proteins are stable at temperature up to 70-75 °C (Atta-ur-Rahman 2011). Therefore, the thermal inactivation condition was set at 60 °C for 15 min to killed cells thoroughly before flow cytometry.

In the microscopy observation, the percentages of red fluorescent cell population appeared to be different when the P*p*1302::*rfp*::*chl^r* were under the control of different host regulators. The higher fluorescent cell population was observed in *E. coli* O157:H7 FUA 1399 (P*p*1302::*rfp*::*chl^r*) which carries a high infectious marker Q_{933} (Plunkett et al. 1999) on its prophage 1399-79. The observation implicated that Q_{933} may be not only link to high Stx production level (Lejeune et al. 2004), but also high induction efficiency. Then, P*p*1302::*rfp*::*chl^r* was transformed into *E. coli* FUA 1303, *E. coli* FUA 1311, *E. coli* FUA 1399 and also its host strain *E. coli* FUA 1302. Compared with the behavior P*p*1302::*rfp*::*chl^r* in its parent strain *E. coli* FUA 1302, results showed remarkably higher induction levels when P*p*1302::*rfp*::*chl^r* in *E. coli* FUA 1311 and *E. coli* FUA 1399 (an *E. coli* O157:H7 carrying Q_{933}). *E. coli* O104:H4 with Q_{933} showed higher Stx mRNA level than *E. coli* O157:H7 (Q_{933}) (Olavesen et al. 2016), other factors may participate in Stx production to facilitate the low induction level of O104:H4 to express high level of Stx.

Since the *stx* gene is located immediately downstream of the late region of the prophage genome, the genomic diversity of the phage regulation region is thought to be critical in determining the level of *stx* expression (Wagner, Acheson, and Waldor 1999). The same Stx2 subtype is produced to various levels because of the prophage diverse genomic structure (Neupane et al. 2011); however the genetic determinants for these differences were not determined. The phylogenetic analyses of *p*R' region and *stx* genes determined in this study showed that although *p*1302::*rfp*::*chl^r* and *p*1309-2d::*rfp*::*chl^r* were highly homogeneous in terms of their *stx* gene, their *p*R' region distributed in different phylogenetic subclades. The efficiency of phage induction, which is a critical factor in terms of Stx production, is considered to be correlated with diversity of phage late gene promoter (Los et al. 2009). I examined the regulation of the same host chromosomal prophage to both native and foreign *p*R' promoter regions. Strain *E. coli* O104:H4 11-3088 Δ*stx*::*gfp*::*amp^r* carrying P*p*1302::*rfp*::*chl^r*, P*p*1306::*rfp*::*chl^r*, and P*p*1303-2a::*rfp*::*chl^r* performed higher *stx* expression ability, while P*p*1309-2d::*rfp*::*chl^r* behaved similar to P*p*1303-s1::*rfp*::*chl^r* and P*p*1309-1c::*rfp*::*chl^r* with a lower induction rate. It demonstrated that the same prophage transcriptional regulation might behave differently when interacts with different downstream promoters and thus may result in different Stx production level.

Transformation of the pR'::rfp::chl^r constructs in different target STEC strains was based on the method used in previous study (Aertsen, Van Houdt, and Michiels 2005). In *E. coli* FUA 1311, the percentage of fluorescent cells of *E. coli* FUA 1311 (Pp1311::rfp::chl^r) was lower than *E. coli* FUA 1311 (Pp1302::rfp::chl^r), which indicate that the stimulation of *E. coli* FUA 1311 prophage to the foreign promoter p1302 was more intense than to its own promoter p1311. O157:H- (Q_{O111}) and O121:H? (Q_{933}) showed comparable Stx mRNA level with *E. coli* EDL933 (Olavesen et al. 2016). The higher stimulation of prophage in *E. coli* FUA 1311 (Q_{933}, O45:H2) by p1302 provides evidence of different induction efficiency of different prophage promotors in strains carrying a Q_{933} antiterminator. Finally, *E. coli* FUA 1399 harbors two *stx2a* prophages but only p1399-79 contributed to Stx production. The higher stimulation ability of p1302 in *E. coli* FUA 1399 than in its native *E. coli* FUA1302 demonstrated that it is possible to generate a highly infectious pathogen from moderately infectious strains through gene recombination. The weak gene expression stimulation of p1399-28 could be explained by different Q gene sequence (Koitabashi et al. 2006) or the interaction between Q and other factors that influence Stx production (Olavesen et al. 2016). The results also indicate that not all the *stx* genes can be expressed in its own host strain, which suggest the pathogenicity of a STEC strain may not be determined by all the prophages it carries in some cases. The construct Pp1302::rfp::chl^r showed higher induction level in *E. coli* FUA 1399 indicated that the combination of different phage regulation region and different pR' could result various induction behaviors.

Divergence in phage genomes may lead the evolution direction of host bacteria, directly or indirectly. Moreover, by horizontally acquiring transferable virulence determinants or integral components from pathogenic strain through phage transduction, phages modify their host bacteria into new pathogens and consequently cause severe diseases like HC and HUS (Boyd and Brussow 2002; O'Brien et al. 1984; Karmali et al. 1985; Huang, Friesen, and Brunton 1987). However, the evolutionary pressure for bacteria-phage interaction does not aim to the human disease. As a component of the gut microbiome, the gut mucus line provides *E. coli* a nice condition to live (Chang et al. 2004). However, the mucus of the gut slough off constantly and excrete *E. coli* population with the feces into the environment. Unlike the stable environment inside of vertebrate hosts, the conditions of the external environment are fluctuating and may impose stress. Ubiquitous bacteriophages and protists are major predators of bacteria (Fuhrman 1999; Sherr and Sherr 1987). However, phages also cooperate with bacteria to gain a mutual benefit: lysogenizing bacterial chromosome adds functional genes for the benefit of the host, while prophages can rely on their host for replication (Canchaya et al. 2003). To escape from the predation of the protists, bacteria evolved many strategies (Matz and Kjelleberg 2005), like toxin

production, which includes prophage encoded toxins. Upon ingestion of the bacterium by its protest predator, it releases toxin inside the predator and kill it to protect the remaining bacteria population (Matz et al. 2004). This proposed "altruism" strategy (Los et al. 2012) was also shown in the current study: even the highest induction efficiency was less than 40% of the total population. The solution that STEC exerting belongs to this toxin-producing strategy. The highly diverse genomes of Shiga-toxin phages facilitate the production of Stx in different environmental conditions, and thus help bacteria defend against diverse predators (Arnold and Koudelka 2014; Miki and Jacquet 2008; Van Elsas et al. 2011). Human neutrophils produce H_2O_2 when bacteria are recognized as alien (Wagner, Acheson, and Waldor 2001), which is similar to the way *Tetrahymena* uses to kill its prey, and also may lead to STEC prophage induction. Thus, the STEC infections of human are thought to be a coincidence of evolution.

As a food-borne pathogen, this accidental evolutionary result has a deep impact on human society. Ruminants, and in particular cattle, are the major host of STEC (Nguyen and Sperandio 2012). Since the host of STEC, beef cattle, are also a main meat resource, there are many ways to transmit STEC from their main hosts to human, such as food contamination from beef processing plant, animal-to-human contact (Yatsuyanagi, Saito, and Ito 2002; European Food Safety, European Centre for Disease, and Control 2013). The infections of STEC lead to not only mild diarrhea, but also fatal complications. In addition, STEC contaminated beef products also bring economic losses which are the result of recalls. Current interventions include physical interventions, such as thermal and high-pressure interventions, and chemical approaches like UV, and acid interventions, which have been proven to prevent the transmission of STEC via food efficiently (Erickson and Doyle 2007). However, outbreaks of STEC infections still happen frequently because of the contamination of meat with STECs (Hussein and Bollinger 2005). The persistence and spread of STECs in meat industry are mainly due to the spread of Stx prophages among ruminants and their high mutability to generate new *E. coli* pathogens. Thus, understanding the link between genomic diversity of Stx prophages and Stx production, the transduction ability to convert a non-STEC strain into a new STEC strain and induction efficiency of a *stx*-converting *E. coli* to produce Stx are fundamental solutions on preventing STEC contamination in beef industry and STEC infection for public health.

This study provided evidence that the induction levels varied with different combinations of transcriptional regulation and the promoter region. This result implies that, in addition to the bacterial diversity, phage diversity is a determinant of the diversity of STEC virulence. Since the gene exchange by phage is not often restricted by species boundaries (Pajunen et al. 2001), it is easy to convert a non-toxic *E. coli* into an STEC. Among the findings of the current research,

different induction efficiencies of $p1302::rfp::chl^r$ in different *stx*-converting *E. coli* strains may partially explain the different clinical outcomes after infection with the same STEC strain. In addition, the immune system of different individuals may differ in the capacity to induce Stx prophages, and the diversity of the human intestinal microbiota (Eckburg et al. 2005) may result in different levels of transduction of commensal *E. coli* with Stx phages. Furthermore, combined with the evidence about the sensitivity of *E. coli* strains to transduction of Stx prophages, the assessment of the induction efficiency of a Stx promoter under different prophage regulation may provide novel insights about how transduction and gene transfer happens among non-virulent *E. coli* strains and how STEC strains affect the generation of new *E. coli* pathogens.

References

Aertsen, A., R. Van Houdt, and C. W. Michiels. 2005. 'Construction and use of an stx1 transcriptional fusion to *gfp*', *FEMS Microbiol Lett*, 245: 73-7.

Ahmad, A., and L. Zurek. 2006. 'Evaluation of the anti-terminator *Q933* gene as a marker for *Escherichia coli* O157:H7 with high Shiga toxin production', *Curr Microbiol*, 53: 324-8.

Ahmed, S. A., J. Awosika, C. Baldwin, K. A. Bishop-Lilly, B. Biswas, S. Broomall, P. S. Chain, O. Chertkov, O. Chokoshvili, S. Coyne, K. Davenport, J. C. Detter, W. Dorman, T. H. Erkkila, J. P. Folster, K. G. Frey, M. George, C. Gleasner, M. Henry, K. K. Hill, K. Hubbard, J. Insalaco, S. Johnson, A. Kitzmiller, M. Krepps, C. C. Lo, T. Luu, L. A. McNew, T. Minogue, C. A. Munk, B. Osborne, M. Patel, K. G. Reitenga, C. N. Rosenzweig, A. Shea, X. Shen, N. Strockbine, C. Tarr, H. Teshima, E. van Gieson, K. Verratti, M. Wolcott, G. Xie, S. Sozhamannan, and H. S. Gibbons. 2012. 'Genomic comparison of *Escherichia coli* O104:H4 isolates from 2009 and 2011 reveals plasmid, and prophage heterogeneity, including shiga toxin encoding phage stx2', *PLoS One*, 7: e48228.

Altuvia, S., D. Kornitzer, S. Kobi, and A. B. Oppenheim. 1991. 'Functional and structural elements of the mRNA of the *cIII* gene of bacteriophage lambda', *J Mol Biol*, 218: 723-33.

Arnold, J. W., and G. B. Koudelka. 2014. 'The Trojan Horse of the microbiological arms race: phage-encoded toxins as a defence against eukaryotic predators', *Environ Microbiol*, 16: 454-66.

Asbury, C. L., J. L. Uy, and G. van den Engh. 2000. 'Polarization of scatter and fluorescence signals in flow cytometry', *Cytometry*, 40: 88-101.

Atta-ur-Rahman, FRS. 2011. *Studies in Natural Products Chemistry: Bioactive Natural Products* (Elsevier Science).

Baird, G. S., D. A. Zacharias, and R. Y. Tsien. 2000. 'Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral', *Proc Natl Acad Sci U S A*, 97: 11984-9.

Bergan, Jonas, Anne Berit Dyve Lingelem, Roger Simm, Tore Skotland, and Kirsten Sandvig. 2012. 'Shiga toxins', *Toxicon*, 60: 1085-107.

Berger, C. N., S. V. Sodha, R. K. Shaw, P. M. Griffin, D. Pink, P. Hand, and G. Frankel. 2010. 'Fresh fruit and vegetables as vehicles for the transmission of human pathogens', *Environ Microbiol*, 12: 2385-97.

Bertin, Y., K. Boukhors, N. Pradel, V. Livrelli, and C. Martin. 2001. 'Stx2 subtyping of Shiga toxin-producing *Escherichia coli* isolated from cattle in France: detection of a new Stx2 subtype and correlation with additional virulence factors', *J Clin Microbiol*, 39: 3060-5.

Beutin, L., J. A. Hammerl, E. Strauch, J. Reetz, R. Dieckmann, Y. Kelner-Burgos, A. Martin, A. Miko, N. A. Strockbine, B. A. Lindstedt, D. Horn, H. Monse, B. Huettel, I. Muller, K. Stuber, and R. Reinhardt. 2012. 'Spread of a distinct Stx2-encoding phage prototype

among *Escherichia coli* O104:H4 strains from outbreaks in Germany, Norway, and Georgia', *J Virol*, 86: 10444-55.

Beutin, L., U. Kruger, G. Krause, A. Miko, A. Martin, and E. Strauch. 2008. 'Evaluation of major types of Shiga toxin 2E-producing *Escherichia coli* bacteria present in food, pigs, and the environment as potential pathogens for humans', *Appl Environ Microbiol*, 74: 4806-16.

Bevis, B. J., and B. S. Glick. 2002. 'Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed)', *Nat Biotechnol*, 20: 83-7.

Bielaszewska, Martina, Alexander Mellmann, Wenlan Zhang, Robin Köck, Angelika Fruth, Andreas Bauwens, Georg Peters, and Helge Karch. 2011. 'Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study', *The Lancet Infectious Diseases*, 11: 671-76.

Bienkowska-Szewczyk, K., B. Lipinska, and A. Taylor. 1981. 'The *R* gene product of bacteriophage lambda is the murein transglycosylase', *Mol Gen Genet*, 184: 111-4.

Blanco, M., J. E. Blanco, A. Mora, G. Dahbi, M. P. Alonso, E. A. Gonzalez, M. I. Bernardez, and J. Blanco. 2004. 'Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from cattle in Spain and identification of a new intimin variant gene (*eae-xi*)', *J Clin Microbiol*, 42: 645-51.

Boyd, E. F., and H. Brussow. 2002. 'Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved', *Trends Microbiol*, 10: 521-9.

Brüssow, Harald, Carlos Canchaya, and Wolf-Dietrich Hardt. 2004. 'Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion', *Microbiology and Molecular Biology Reviews*, 68: 560-602.

Brouns, S. J., M. M. Jore, M. Lundgren, E. R. Westra, R. J. Slijkhuis, A. P. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin, and J. van der Oost. 2008. 'Small CRISPR RNAs guide antiviral defense in prokaryotes', *Science*, 321: 960-4.

Brown, J. E., S. W. Rothman, and B. P. Doctor. 1980. 'Inhibition of protein synthesis in intact HeLa cells by *Shigella dysenteriae* 1 toxin', *Infect Immun*, 29: 98-107.

Bull, JJ. 1987. 'Evolution of phenotypic variance', *Evolution*, 41: 303-15.

Campbell, A. 1961. 'Sensitive mutants of bacteriophage lambda', *Virology*, 14: 22-32.

Campbell, Allan. 2003. 'Prophage insertion sites', *Research in Microbiology*, 154: 277-82.

Canchaya, Carlos, Caroline Proux, Ghislain Fournous, Anne Bruttin, and Harald Brüssow. 2003. 'Prophage genomics', *Microbiology and Molecular Biology Reviews*, 67: 238-76.

Casjens, Sherwood R., and Roger W. Hendrix. 2015. 'Bacteriophage lambda: Early pioneer and still relevant', *Virology*, 479–480: 310-30.

CDC. 2017. 'Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet 2015 Surveillance Report (Final Data)', *CDC, Atlanta, Georgia: U.S.*

Cebula, T. A., W. L. Payne, and P. Feng. 1995. 'Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR', *J Clin Microbiol*, 33: 248-50.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher. 1994. 'Green fluorescent protein as a marker for gene expression', *Science*, 263: 802-5.

Chang, D. E., D. J. Smalley, D. L. Tucker, M. P. Leatham, W. E. Norris, S. J. Stevenson, A. B. Anderson, J. E. Grissom, D. C. Laux, P. S. Cohen, and T. Conway. 2004. 'Carbon nutrition of *Escherichia coli* in the mouse intestine', *Proc Natl Acad Sci U S A*, 101: 7427-32.

Chudakov, D. M., M. V. Matz, S. Lukyanov, and K. A. Lukyanov. 2010. 'Fluorescent proteins and their applications in imaging living cells and tissues', *Physiol Rev*, 90: 1103-63.

Clermont, O., M. Lescat, C. L. O'Brien, D. M. Gordon, O. Tenaillon, and E. Denamur. 2008. 'Evidence for a human-specific *Escherichia coli* clone', *Environ Microbiol*, 10: 1000-6.

Cormack, B. P., R. H. Valdivia, and S. Falkow. 1996. 'FACS-optimized mutants of the green fluorescent protein (GFP)', *Gene*, 173: 33-8.

Court, D. L., A. B. Oppenheim, and S. L. Adhya. 2007. 'A new look at bacteriophage lambda genetic networks', *J Bacteriol*, 189: 298-304.

Croxen, M. A., R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska, and B. B. Finlay. 2013. 'Recent advances in understanding enteric pathogenic *Escherichia coli*', *Clin Microbiol Rev*, 26: 822-80.

Datsenko, K. A., and B. L. Wanner. 2000. 'One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products', *Proc Natl Acad Sci U S A*, 97: 6640-5.

Day, R. N., and M. W. Davidson. 2009. 'The fluorescent protein palette: tools for cellular imaging', *Chem Soc Rev*, 38: 2887-921.

De Grandis, S., J. Ginsberg, M. Toone, S. Climie, J. Friesen, and J. Brunton. 1987. 'Nucleotide sequence and promoter mapping of the *Escherichia coli* Shiga-like toxin operon of bacteriophage H-19B', *J Bacteriol*, 169: 4313-9.

Dlusskaya, E. A., L. M. McMullen, and M. G. Gänzle. 2011. 'Characterization of an extremely heat-resistant *Escherichia coli* obtained from a beef processing facility', *Journal of Applied Microbiology*, 110: 840-49.

Eckburg, Paul B., Elisabeth M. Bik, Charles N. Bernstein, Elizabeth Purdom, Les Dethlefsen, Michael Sargent, Steven R. Gill, Karen E. Nelson, and David A. Relman. 2005. 'Diversity of the Human Intestinal Microbial Flora', *Science*, 308: 1635-38.

Edgar, R. C. 2004. 'MUSCLE: multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Res*, 32: 1792-7.

Eisen, H., P. Brachet, L. Pereira da Silva, and F. Jacob. 1970. 'Regulation of repressor expression in lambda', *Proc Natl Acad Sci U S A*, 66: 855-62.

Endo, Y., K. Tsurugi, T. Yutsudo, Y. Takeda, T. Ogasawara, and K. Igarashi. 1988. 'Site of action of a Vero toxin (VT2) from *Escherichia coli* O157:H7 and of Shiga toxin on eukaryotic ribosomes. RNA N-glycosidase activity of the toxins', *Eur J Biochem*, 171: 45-50.

Erickson, M. C., and M. P. Doyle. 2007. 'Food as a vehicle for transmission of Shiga toxin-producing *Escherichia coli*', *J Food Prot*, 70: 2426-49.

Escobar-Paramo, P., A. Le Menac'h, T. Le Gall, C. Amorin, S. Gouriou, B. Picard, D. Skurnik, and E. Denamur. 2006. 'Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates', *Environ Microbiol*, 8: 1975-84.

European Food Safety, Authority, Prevention European Centre for Disease, and Control. 2013. 'The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks in 2011', *EFSA Journal*, 11: n/a-n/a.

Fagerquist, C. K., and O. Sultan. 2010. 'Top-down proteomic identification of furin-cleaved alpha-subunit of Shiga toxin 2 from *Escherichia coli* O157:H7 using MALDI-TOF-TOF-MS/MS', *J Biomed Biotechnol*, 2010: 123460.

Fortier, L. C., and O. Sekulovic. 2013. 'Importance of prophages to evolution and virulence of bacterial pathogens', *Virulence*, 4: 354-65.

Frank, C., D. Werber, J. P. Cramer, M. Askar, M. Faber, M. an der Heiden, H. Bernard, A. Fruth, R. Prager, A. Spode, M. Wadl, A. Zoufaly, S. Jordan, M. J. Kemper, P. Follin, L. Muller, L. A. King, B. Rosner, U. Buchholz, K. Stark, and G. Krause. 2011. 'Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany', *N Engl J Med*, 365: 1771-80.

Friedman, David I., and Max Gottesman. 1983. *Lytic Mode of Lambda Development*.

Friedrich, A. W., M. Bielaszewska, W. L. Zhang, M. Pulz, T. Kuczius, A. Ammon, and H. Karch. 2002. '*Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms', *J Infect Dis*, 185: 74-84.

Fuhrman, J. A. 1999. 'Marine viruses and their biogeochemical and ecological effects', *Nature*, 399: 541-8.

Fuller, C. A., C. A. Pellino, M. J. Flagler, J. E. Strasser, and A. A. Weiss. 2011. 'Shiga toxin subtypes display dramatic differences in potency', *Infect Immun*, 79: 1329-37.

Fulwyler, M. J. 1965. 'Electronic separation of biological cells by volume', *Science*, 150: 910-1.

Garred, O., E. Dubinina, A. Polesskaya, S. Olsnes, J. Kozlov, and K. Sandvig. 1997. 'Role of the disulfide bond in Shiga toxin A-chain for toxin entry into cells', *J Biol Chem*, 272: 11414-9.

Garred, O., B. van Deurs, and K. Sandvig. 1995. 'Furin-induced cleavage and activation of Shiga toxin', *J Biol Chem*, 270: 10817-21.

Giladi, H., S. Koby, G. Prag, M. Engelhorn, J. Geiselmann, and A. B. Oppenheim. 1998. 'Participation of IHF and a distant UP element in the stimulation of the phage lambda PL promoter', *Mol Microbiol*, 30: 443-51.

Glinkowska, M., J. M. Los, A. Szambowska, A. Czyz, J. Calkiewicz, A. Herman-Antosiewicz, B. Wrobel, G. Wegrzyn, A. Wegrzyn, and M. Los. 2010. 'Influence of the *Escherichia coli oxyR* gene function on lambda prophage maintenance', *Arch Microbiol*, 192: 673-83.

Gucker, F. T., Jr., C. T. O'Konski, and et al. 1947. 'A photoelectronic counter for colloidal particles', *J Am Chem Soc*, 69: 2422-31.

Hao, W., V. G. Allen, F. B. Jamieson, D. E. Low, and D. C. Alexander. 2012. 'Phylogenetic incongruence in *E. coli* O104: understanding the evolutionary relationships of emerging pathogens in the face of homologous recombination', *PLoS One*, 7: e33971.

Harley, C. B., and R. P. Reynolds. 1987. 'Analysis of *E. coli* promoter sequences', *Nucleic Acids Research*, 15: 2343-61.

Harris, A. W., D. W. Mount, C. R. Fuerst, and L. Siminovitch. 1967. 'Mutations in bacteriophage lambda affecting host cell lysis', *Virology*, 32: 553-69.

Hauben, K J, D H Bartlett, C C Soontjens, K Cornelis, E Y Wuytack, and C W Michiels. 1997. '*Escherichia coli* mutants resistant to inactivation by high hydrostatic pressure', *Applied and Environmental Microbiology*, 63: 945-50.

Haugum, K., B. A. Lindstedt, I. Lobersli, G. Kapperud, and L. T. Brandal. 2012. 'Identification of the anti-terminator *qO111:H)-* gene in Norwegian sorbitol-fermenting *Escherichia coli* O157:NM', *FEMS Microbiol Lett*, 329: 102-10.

Hawley, T. S., D. J. Herbert, S. S. Eaker, and R. G. Hawley. 2004. 'Multiparameter flow cytometry of fluorescent protein reporters', *Methods Mol Biol*, 263: 219-38.

Hawley, T. S., W. G. Telford, A. Ramezani, and R. G. Hawley. 2001. 'Four-color flow cytometric detection of retrovirally expressed red, yellow, green, and cyan fluorescent proteins', *Biotechniques*, 30: 1028-34.

Heim, Roger, Andrew B. Cubitt, and Roger Y. Tsien. 1995. 'Improved green fluorescence', *Nature*, 373: 663-64.

Huang, A., J. Friesen, and J. L. Brunton. 1987. 'Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in *Escherichia coli*', *J Bacteriol*, 169: 4308-12.

Hussein, H. S., and L. M. Bollinger. 2005. 'Prevalence of Shiga toxin-producing *Escherichia coli* in beef', *Meat Sci*, 71: 676-89.

Imamovic, Lejla, and Maite Muniesa. 2012. 'Characterizing RecA-Independent Induction of Shiga toxin2-Encoding Phages by EDTA Treatment', *PLoS One*, 7: e32393.

Jackson, Matthew P., Roger J. Neill, Alison D. O'Brien, Randall K. Holmes, and John W. Newland. 1987. 'Nucleotide sequence analysis and comparison of the structural genes for Shiga-like toxin I and Shiga-like toxin II encoded by bacteriophages from *Escherichia coli* 933', *FEMS Microbiology Letters*, 44: 109-14.

Jia, H., W. J. Satumba, G. L. Bidwell, 3rd, and M. C. Mossing. 2005. 'Slow assembly and disassembly of lambda Cro repressor dimers', *J Mol Biol*, 350: 919-29.

Kaiser, A. D., and F. Jacob. 1957. 'Recombination between related temperate bacteriophages and the genetic control of immunity and prophage localization', *Virology*, 4: 509-21.

Kaper, J. B., J. P. Nataro, and H. L. Mobley. 2004a. 'Pathogenic *Escherichia coli*', *Nat Rev Microbiol*, 2: 123-40.

Kaper, James B., James P. Nataro, and Harry L. T. Mobley. 2004b. 'Pathogenic *Escherichia coli*', *Nat Rev Micro*, 2: 123-40.

Karch, Helge, Erick Denamur, Ulrich Dobrindt, B. Brett Finlay, Regine Hengge, Ludgers Johannes, Eliora Z. Ron, Tone Tønjum, Philippe J. Sansonetti, and Miguel Vicente. 2012. 'The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak', *EMBO Molecular Medicine*, 4: 841-48.

Karmali, M. A., M. Petric, C. Lim, P. C. Fleming, G. S. Arbus, and H. Lior. 1985. 'The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. 1985', *J Infect Dis*, 189: 556-63.

Keusch, G. T., M. Jacewicz, D. W. Acheson, A. Donohue-Rolfe, A. V. Kane, and R. H. McCluer. 1995. 'Globotriaosylceramide, Gb3, is an alternative functional receptor for Shiga-like toxin 2e', *Infect Immun*, 63: 1138-41.

King, Oliver D., and Joanna Masel. 2007. 'The evolution of bet-hedging adaptations to rare scenarios', *Theoretical Population Biology*, 72: 560-75.

Kobiler, O., A. B. Oppenheim, and C. Herman. 2004. 'Recruitment of host ATP-dependent proteases by bacteriophage lambda', *J Struct Biol*, 146: 72-8.

Kobiler, O., A. Rokney, and A. B. Oppenheim. 2007. 'Phage lambda CIII: a protease inhibitor regulating the lysis-lysogeny decision', *PLoS One*, 2: e363.

Koitabashi, Tsutomu, Varaporn Vuddhakul, Son Radu, Tadaaki Morigaki, Norio Asai, Yoshitsugu Nakaguchi, and Mitsuaki Nishibuchi. 2006. 'Genetic Characterization of *Escherichia coli* O157:H7/– Strains Carrying the *stx2* Gene but Not Producing Shiga Toxin 2', *Microbiology and Immunology*, 50: 135-48.

Konowalchuk, J., J. I. Speirs, and S. Stavric. 1977. 'Vero response to a cytotoxin of *Escherichia coli*', *Infect Immun*, 18: 775-9.

Kruger, A., P. M. Lucchesi, and A. E. Parma. 2011. 'Verotoxins in bovine and meat verotoxin-producing *Escherichia coli* isolates: type, number of variants, and relationship to cytotoxicity', *Appl Environ Microbiol*, 77: 73-9.

Kubitschek, H. E. 1958. 'Electronic counting and sizing of bacteria', *Nature*, 182: 234-5.

Kuznedelov, K., L. Minakhin, A. Niedziela-Majka, S. L. Dove, D. Rogulja, B. E. Nickels, A. Hochschild, T. Heyduk, and K. Severinov. 2002. 'A role for interaction of the RNA polymerase flap domain with the sigma subunit in promoter recognition', *Science*, 295: 855-7.

Learn, B. A., S. J. Um, L. Huang, and R. McMacken. 1997. 'Cryptic single-stranded-DNA binding activities of the phage lambda P and Escherichia coli DnaC replication initiation proteins facilitate the transfer of *E. coli* DnaB helicase onto DNA', *Proc Natl Acad Sci U S A*, 94: 1154-9.

Lederberg, Esther. 1951. '" Lysogenicity in *Escherichia coli* K-12"'.

Lejeune, J. T., S. T. Abedon, K. Takemura, N. P. Christie, and S. Sreevatsan. 2004. 'Human *Escherichia coli* O157:H7 genetic marker in isolates of bovine origin', *Emerg Infect Dis*, 10: 1482-5.

Libby, E., and P. B. Rainey. 2011. 'Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching', *Proc Biol Sci*, 278: 3574-83.

Lin-Chao, S., and H. Bremer. 1986. 'Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*', *Mol Gen Genet*, 203: 143-9.

Ling, H., A. Boodhoo, B. Hazes, M. D. Cummings, G. D. Armstrong, J. L. Brunton, and R. J. Read. 1998. 'Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3', *Biochemistry*, 37: 1777-88.

Liu, Y., A. Gill, L. McMullen, and M. G. Ganzle. 2015. 'Variation in heat and pressure resistance of verotoxigenic and nontoxigenic *Escherichia coli*', *J Food Prot*, 78: 111-20.

Los, J. M., M. Los, A. Wegrzyn, and G. Wegrzyn. 2012. 'Altruism of Shiga toxin-producing *Escherichia coli*: recent hypothesis versus experimental results', *Front Cell Infect Microbiol*, 2: 166.

Los, J. M., M. Los, G. Wegrzyn, and A. Wegrzyn. 2009. 'Differential efficiency of induction of various lambdoid prophages responsible for production of Shiga toxins in response to different induction agents', *Microb Pathog*, 47: 289-98.

Luck, Shelley N., Vicki Bennett-Wood, Rachael Poon, Roy M. Robins-Browne, and Elizabeth L. Hartland. 2005. 'Invasion of epithelial cells by locus of enterocyte effacement-negative enterohemorrhagic *Escherichia coli*', *Infect Immun*, 73: 3063-71.

Macdonald, Patrick J., Yan Chen, and Joachim D. Mueller. 2012. 'Chromophore maturation and fluorescence fluctuation spectroscopy of fluorescent proteins in a cell-free expression system', *Analytical Biochemistry*, 421: 291-98.

Makino, K., K. Yokoyama, Y. Kubota, C. H. Yutsudo, S. Kimura, K. Kurokawa, K. Ishii, M. Hattori, I. Tatsuno, H. Abe, T. Iida, K. Yamamoto, M. Onishi, T. Hayashi, T. Yasunaga, T. Honda, C. Sasakawa, and H. Shinagawa. 1999. 'Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak', *Genes Genet Syst*, 74: 227-39.

Manning, S. D., A. S. Motiwala, A. C. Springman, W. Qi, D. W. Lacher, L. M. Ouellette, J. M. Mladonicky, P. Somsel, J. T. Rudrik, S. E. Dietrich, W. Zhang, B. Swaminathan, D. Alland, and T. S. Whittam. 2008. 'Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks', *Proc Natl Acad Sci U S A*, 105: 4868-73.

Mardanov, A. V., and N. V. Ravin. 2006. 'Functional characterization of the repA replication gene of linear plasmid prophage N15', *Res Microbiol*, 157: 176-83.

Matsushiro, A., K. Sato, H. Miyamoto, T. Yamamura, and T. Honda. 1999a. 'Induction of prophages of enterohemorrhagic *Escherichia coli* O157:H7 with norfloxacin', *J Bacteriol*, 181: 2257-60.

Matsushiro, Aizo, Koki Sato, Hiroshi Miyamoto, Tadashi Yamamura, and Takeshi Honda. 1999b. 'Induction of Prophages of Enterohemorrhagic *Escherichia coli* O157:H7 with Norfloxacin', *J Bacteriol*, 181: 2257-60.

Matz, C., P. Deines, J. Boenigk, H. Arndt, L. Eberl, S. Kjelleberg, and K. Jurgens. 2004. 'Impact of violacein-producing bacteria on survival and feeding of bacterivorous nanoflagellates', *Appl Environ Microbiol*, 70: 1593-9.

Matz, C., and S. Kjelleberg. 2005. 'Off the hook--how bacteria survive protozoan grazing', *Trends Microbiol*, 13: 302-7.

Matz, M. V., A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov, and S. A. Lukyanov. 1999. 'Fluorescent proteins from nonbioluminescent Anthozoa species', *Nat Biotechnol*, 17: 969-73.

McDaniel, T. K., K. G. Jarvis, M. S. Donnenberg, and J. B. Kaper. 1995. 'A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens', *Proc Natl Acad Sci U S A*, 92: 1664-8.

McDonough, M. A., and J. R. Butterton. 1999. 'Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1', *Mol Microbiol*, 34: 1058-69.

Mercer, R. G., J. Zheng, R. Garcia-Hernandez, L. Ruan, M. G. Ganzle, and L. M. McMullen. 2015. 'Genetic determinants of heat resistance in *Escherichia coli*', *Front Microbiol*, 6: 932.

Mercer, Ryan G., Jinshui Zheng, Rigoberto Garcia-Hernandez, Lifang Ruan, Michael G. Gänzle, and Lynn M. McMullen. 2015. 'Genetic determinants of heat resistance in *Escherichia coli*', *Front Microbiol*, 6: 932.

Miki, T., and S. Jacquet. 2008. 'Complex interactions in the microbial world: underexplored key links between viruses, bacteria and protozoan grazers in aquatic environments', *Aquatic Microbial Ecology*, 51: 195-208.

Mishra, S., S. Mohan, S. Godavarthi, and R. Sen. 2013. 'The interaction surface of a bacterial transcription elongation factor required for complex formation with an antiterminator during transcription antitermination', *J Biol Chem*, 288: 28089-103.

Mogridge, J., P. Legault, J. Li, M. D. Van Oene, L. E. Kay, and J. Greenblatt. 1998. 'Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein', *Mol Cell*, 1: 265-75.

Mooers, B. H., and B. W. Matthews. 2006. 'Extension to 2268 atoms of direct methods in the ab initio determination of the unknown structure of bacteriophage P22 lysozyme', *Acta Crystallogr D Biol Crystallogr*, 62: 165-76.

Morise, H., O. Shimomura, F. H. Johnson, and J. Winant. 1974. 'Intermolecular energy transfer in the bioluminescent system of *Aequorea*', *Biochemistry*, 13: 2656-62.

Moxon, E. R., P. B. Rainey, M. A. Nowak, and R. E. Lenski. 1994. 'Adaptive evolution of highly mutable loci in pathogenic bacteria', *Curr Biol*, 4: 24-33.

Muniesa, Maite, Jens A. Hammerl, Stefan Hertwig, Bernd Appel, and Harald Brüssow. 2012. 'Shiga Toxin-Producing *Escherichia coli* O104:H4: a New Challenge for Microbiology', *Applied and Environmental Microbiology*, 78: 4065-73.

Narajczyk, M., S. Baranska, A. Wegrzyn, and G. Wegrzyn. 2007. 'Switch from theta to sigma replication of bacteriophage lambda DNA: factors involved in the process and a model for its regulation', *Mol Genet Genomics*, 278: 65-74.

Nataro, J. P., and J. B. Kaper. 1998. 'Diarrheagenic *Escherichia coli*', *Clin Microbiol Rev*, 11: 142-201.

Naylor, S. W., D. L. Gally, and J. C. Low. 2005. 'Enterohaemorrhagic *E. coli* in veterinary medicine', *Int J Med Microbiol*, 295: 419-41.

Neely, M. N., and D. I. Friedman. 1998a. 'Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin', *Gene*, 223: 105-13.

———. 1998b. 'Functional and genetic analysis of regulatory regions of coliphage H-19B: location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release', *Mol Microbiol*, 28: 1255-67.

NESP. 2015. 'National Enteric Surveillance Program (NESP) annual summary 2013', *The National Microbiology Laboratory (NML) and Centre for Food-borne, Environmental and Zoonotic Infectious Diseases (CFEZID), Public Health Agency of Canada and Provincial Public Health Microbiology Laboratories, Winnipeg, Canada.*

Neupane, Mahesh, Galeb S. Abu-Ali, Avishek Mitra, David W. Lacher, Shannon D. Manning, and James T. Riordan. 2011. 'Shiga toxin 2 overexpression in *Escherichia coli* O157:H7 strains associated with severe human disease', *Microbial Pathogenesis*, 51: 466-70.

Newton, Hayley J., Joan Sloan, Dieter M. Bulach, Torsten Seemann, Cody C. Allison, Marija Tauschek, Roy M. Robins-Browne, James C. Paton, Thomas S. Whittam, Adrienne W. Paton, and Elizabeth L. Hartland. 2009. 'Shiga Toxin–producing *Escherichia coli* Strains Negative for Locus of Enterocyte Effacement', *Emerging Infectious Diseases*, 15: 372-80.

Nguyen, Y., and Vanessa Sperandio. 2012. 'Enterohemorrhagic E. coli (EHEC) pathogenesis', *Front Cell Infect Microbiol*, 2: 90.

O'Brien, A. D., G. D. LaVeck, M. R. Thompson, and S. B. Formal. 1982. 'Production of *Shigella dysenteriae* type 1-like cytotoxin by *Escherichia coli*', *J Infect Dis*, 146: 763-9.

O'Brien, A. D., J. W. Newland, S. F. Miller, R. K. Holmes, H. W. Smith, and S. B. Formal. 1984. 'Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea', *Science*, 226: 694-6.

Ogura, Y., S. I. Mondal, M. R. Islam, T. Mako, K. Arisawa, K. Katsura, T. Ooka, Y. Gotoh, K. Murase, M. Ohnishi, and T. Hayashi. 2015. 'The Shiga toxin 2 production level in enterohemorrhagic *Escherichia coli* O157:H7 is correlated with the subtypes of toxin-encoding phage', *Sci Rep*, 5: 16663.

Ogura, Y., T. Ooka, Asadulghani, J. Terajima, J. P. Nougayrede, K. Kurokawa, K. Tashiro, T. Tobe, K. Nakayama, S. Kuhara, E. Oswald, H. Watanabe, and T. Hayashi. 2007. 'Extensive genomic diversity and selective conservation of virulence-determinants in enterohemorrhagic *Escherichia coli* strains of O157 and non-O157 serotypes', *Genome Biol*, 8: R138.

Ohnishi, M., J. Terajima, K. Kurokawa, K. Nakayama, T. Murata, K. Tamura, Y. Ogura, H. Watanabe, and T. Hayashi. 2002. 'Genomic diversity of enterohemorrhagic *Escherichia coli* O157 revealed by whole genome PCR scanning', *Proc Natl Acad Sci U S A*, 99: 17043-8.

Oi, V.T., Glazer, A.N., Stryer, L. 1982. 'Fluorescent phycobiliprotein conjugates for analyses of cells and molecules', *The Journal of Cell Biology*, 93: 981-86.

Olavesen, Kristoffer K., Bjørn-Arne Lindstedt, Inger Løbersli, and Lin T. Brandal. 2016. 'Expression of Shiga toxin 2 (Stx2) in highly virulent Stx-producing *Escherichia coli* (STEC) carrying different anti-terminator (*q*) genes', *Microbial Pathogenesis*, 97: 1-8.

Olszewski, P., A. Szambowska, S. Baralska, M. Narajczyk, G. Wegrzyn, and M. Glinkowska. 2014. 'A dual promoter system regulating lambda DNA replication initiation', *Nucleic Acids Res*, 42: 4450-62.

Pajunen, M. I., S. J. Kiljunen, M. E. Soderholm, and M. Skurnik. 2001. 'Complete genomic sequence of the lytic bacteriophage phiYeO3-12 of Yersinia enterocolitica serotype O:3', *J Bacteriol*, 183: 1928-37.

Parkinson, J. S. 1968. 'Genetics of the left arm of the chromosome of bacteriophage lambda', *Genetics*, 59: 311-25.

Paton, A. W., R. M. Ratcliff, R. M. Doyle, J. Seymour-Murray, D. Davos, J. A. Lanser, and J. C. Paton. 1996. 'Molecular microbiological investigation of an outbreak of hemolytic-uremic syndrome caused by dry fermented sausage contaminated with Shiga-like toxin-producing *Escherichia coli*', *J Clin Microbiol*, 34: 1622-27.

Perez, C. E. 2002. 'Fruit and vegetable consumption', *Health Rep*, 13: 23-31.

Piatkevich, Kiryl D., and Vladislav V. Verkhusha. 2011. 'Guide to red fluorescent proteins and biosensors for flow cytometry', *Methods in cell biology*, 102: 431-61.

Plunkett, G., 3rd, D. J. Rose, T. J. Durfee, and F. R. Blattner. 1999. 'Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product', *J Bacteriol*, 181: 1767-78.

Popa, M. P., T. A. McKelvey, J. Hempel, and R. W. Hendrix. 1991. 'Bacteriophage HK97 structure: wholesale covalent cross-linking between the major head shell subunits', *J Virol*, 65: 3227-37.

Prasher, D. C., V. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier. 1992. 'Primary structure of the *Aequorea victoria* green-fluorescent protein', *Gene*, 111: 229-33.

Ptashne, M. 1967. 'Isolation of the lambda phage repressor', *Proc Natl Acad Sci U S A*, 57: 306-13.

Rainey, P. B., H. J. Beaumont, G. C. Ferguson, J. Gallie, C. Kost, E. Libby, and X. X. Zhang. 2011. 'The evolutionary emergence of stochastic phenotype switching in bacteria', *Microb Cell Fact*, 10 Suppl 1: S14.

Reisbig, R., S. Olsnes, and K. Eiklid. 1981. 'The cytotoxic activity of Shigella toxin. Evidence for catalytic inactivation of the 60 S ribosomal subunit', *J Biol Chem*, 256: 8739-44.

Riley, L. W., R. S. Remis, S. D. Helgerson, H. B. McGee, J. G. Wells, B. R. Davis, R. J. Hebert, E. S. Olcott, L. M. Johnson, N. T. Hargrett, P. A. Blake, and M. L. Cohen. 1983. 'Hemorrhagic colitis associated with a rare *Escherichia coli* serotype', *N Engl J Med*, 308: 681-5.

Roberts, J. W., W. Yarnell, E. Bartlett, J. Guo, M. Marr, D. C. Ko, H. Sun, and C. W. Roberts. 1998. 'Antitermination by bacteriophage lambda Q protein', *Cold Spring Harb Symp Quant Biol*, 63: 319-25.

Saxena, S. K., A. D. O'Brien, and E. J. Ackerman. 1989. 'Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into Xenopus oocytes', *J Biol Chem*, 264: 596-601.

Scallan, E. 2007. 'Activities, achievements, and lessons learned during the first 10 years of the Foodborne Diseases Active Surveillance Network: 1996-2005', *Clin Infect Dis*, 44: 718-25.

Scheutz, F., L. D. Teel, L. Beutin, D. Pierard, G. Buvens, H. Karch, A. Mellmann, A. Caprioli, R. Tozzoli, S. Morabito, N. A. Strockbine, A. R. Melton-Celsa, M. Sanchez, S. Persson, and A. D. O'Brien. 2012. 'Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature', *J Clin Microbiol*, 50: 2951-63.

Schmidt, H., H. Russmann, A. Schwarzkopf, S. Aleksic, J. Heesemann, and H. Karch. 1994. 'Prevalence of attaching and effacing *Escherichia coli* in stool samples from patients and controls', *Zentralbl Bakteriol*, 281: 201-13.

Schmidt, H., J. Scheef, C. Janetzki-Mittmann, M. Datz, and H. Karch. 1997. 'An *ileX* tRNA gene is located close to the Shiga toxin II operon in enterohemorrhagic *Escherichia coli* O157 and non-O157 strains', *FEMS Microbiology Letters*, 149: 39-44.

Scott, Maria E., Angela R. Melton-Celsa, and Alison D. O'Brien. 2003. 'Mutations in hns reduce the adherence of Shiga toxin-producing *E. coli* 091:H21 strain B2F1 to human colonic epithelial cells and increase the production of hemolysin', *Microbial Pathogenesis*, 34: 155-59.

Seidah, N. G., A. Donohue-Rolfe, C. Lazure, F. Auclair, G. T. Keusch, and M. Chretien. 1986. 'Complete amino acid sequence of Shigella toxin B-chain. A novel polypeptide containing 69 amino acids and one disulfide bridge', *J Biol Chem*, 261: 13928-31.

Shapiro, H. M. 2000. 'Microbial analysis at the single-cell level: tasks and techniques', *J Microbiol Methods*, 42: 3-16.

Shapiro, Howard M. 2003. "Practical flow cytometry." In. Hoboken, N.J. :: Wiley-Liss.

Sherr, Evelyn B, and Barry F Sherr. 1987. 'High rates of consumption of bacteria by pelagic ciliates', *Nature*, 325: 710-11.

Shiga, K. 1898. 'Über den Dysenteriebacillus (*Bacillus dysenteriae*)', *Zentbl. Bakteriol*, 24: 817-28.

Shimizu, Takeshi, Yuko Ohta, Hiroyasu Tsutsuki, and Masatoshi Noda. 2011. 'Construction of a novel bioluminescent reporter system for investigating Shiga toxin expression of enterohemorrhagic *Escherichia coli*', *Gene*, 478: 1-10.

Shkilnyj, P., and G. B. Koudelka. 2007. 'Effect of salt shock on stability of lambdaimm434 lysogens', *J Bacteriol*, 189: 3115-23.

Slominska, M., P. Neubauer, and G. Wegrzyn. 1999. 'Regulation of bacteriophage lambda development by guanosine 5'-diphosphate-3'-diphosphate', *Virology*, 262: 431-41.

Smith, Darren L., David J. Rooks, Paul C. M. Fogg, Alistair C. Darby, Nick R. Thomson, Alan J. McCarthy, and Heather E. Allison. 2012. 'Comparative genomics of Shiga toxin encoding bacteriophages', *BMC Genomics*, 13: 311.

Stein, P. E., A. Boodhoo, G. J. Tyrrell, J. L. Brunton, and R. J. Read. 1992. 'Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*', *Nature*, 355: 748-50.

Steyert, S. R., J. W. Sahl, C. M. Fraser, L. D. Teel, F. Scheutz, and D. A. Rasko. 2012. 'Comparative genomics and *stx* phage characterization of LEE-negative Shiga toxin-producing *Escherichia coli*', *Front Cell Infect Microbiol*, 2: 133.

Strauch, E., J. A. Hammerl, A. Konietzny, S. Schneiker-Bekel, W. Arnold, A. Goesmann, A. Puhler, and L. Beutin. 2008. 'Bacteriophage 2851 is a prototype phage for dissemination of the Shiga toxin variant gene *2c* in *Escherichia coli* O157:H7', *Infect Immun*, 76: 5466-77.

Strockbine, N. A., M. P. Jackson, L. M. Sung, R. K. Holmes, and A. D. O'Brien. 1988. 'Cloning and sequencing of the genes for Shiga toxin from *Shigella dysenteriae* type 1', *J Bacteriol*, 170: 1116-22.

Thomas, R., and L. Lambert. 1962. 'On the occurrence of bacterial mutations permitting lysogenization by clear variants of temperate bacteriophages', *Journal of Molecular Biology*, 5: 373-74.

Tozzoli, R., L. Grande, V. Michelacci, R. Fioravanti, D. Gally, X. Xu, R. La Ragione, M. Anjum, G. Wu, A. Caprioli, and S. Morabito. 2014. 'Identification and characterization of a peculiar vtx2-converting phage frequently present in verocytotoxin-producing *Escherichia coli* O157 isolated from human infections', *Infect Immun*, 82: 3023-32.

Tuttle, J., T. Gomez, M. P. Doyle, J. G. Wells, T. Zhao, R. V. Tauxe, and P. M. Griffin. 1999. 'Lessons from a large outbreak of *Escherichia coli* O157:H7 infections: insights into the

infectious dose and method of widespread contamination of hamburger patties', *Epidemiol Infect*, 122: 185-92.

Unkmeir, A., and H. Schmidt. 2000. 'Structural analysis of phage-borne *stx* genes and their flanking sequences in shiga toxin-producing *Escherichia coli* and *Shigella dysenteriae* type 1 strains', *Infect Immun*, 68: 4856-64.

van den Beld, M. J., and F. A. Reubsaet. 2012. 'Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*', *Eur J Clin Microbiol Infect Dis*, 31: 899-904.

Van Elsas, Jan Dirk, Alexander V Semenov, Rodrigo Costa, and Jack T Trevors. 2011. 'Survival of *Escherichia coli* in the environment: fundamental and public health aspects', *The ISME journal*, 5: 173-83.

Veal, D. A., D. Deere, B. Ferrari, J. Piper, and P. V. Attfield. 2000. 'Fluorescence staining and flow cytometry for monitoring microbial cells', *J Immunol Methods*, 243: 191-210.

Vorobiev, Sergey M., Yocheved Gensler, Hanif Vahedian-Movahed, Jayaraman Seetharaman, Min Su, Janet Y. Huang, Rong Xiao, Gregory Kornhaber, Gaetano T. Montelione, Liang Tong, Richard H. Ebright, and Bryce E. Nickels. 2014. 'Structure of the DNA-binding and RNA polymerase-binding region of transcription antitermination factor λQ', *Structure (London, England : 1993)*, 22: 488-95.

Wagner, P. L., D. W. Acheson, and M. K. Waldor. 1999. 'Isogenic lysogens of diverse shiga toxin 2-encoding bacteriophages produce markedly different amounts of shiga toxin', *Infect Immun*, 67: 6710-4.

———. 2001. 'Human neutrophils and their products induce Shiga toxin production by enterohemorrhagic *Escherichia coli*', *Infect Immun*, 69: 1934-7.

Wagner, P. L., M. N. Neely, X. Zhang, D. W. Acheson, M. K. Waldor, D. I. Friedman, Journal Article, Non-U.S. Gov't Research Support, U.S. Gov't Research Support, Non-P.H.S., U.S. Gov't Research Support, P.H.S., United States, and J Bacteriol. 2001 Mar;183(6):2081-5. 2001. 'Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain', *J Bacteriol*, 183: 2081-5.

Wall, M. A., M. Socolich, and R. Ranganathan. 2000. 'The structural basis for red fluorescence in the tetrameric GFP homolog DsRed', *Nat Struct Biol*, 7: 1133-8.

Wang, I. N., D. L. Smith, and R. Young. 2000. 'Holins: the protein clocks of bacteriophage infections', *Annu Rev Microbiol*, 54: 799-825.

Wegrzyn, G., K. Licznerska, and A. Wegrzyn. 2012. 'Phage lambda--new insights into regulatory circuits', *Adv Virus Res*, 82: 155-78.

Wickner, S. 1984. 'DNA-dependent ATPase activity associated with phage P22 gene 12 protein', *J Biol Chem*, 259: 14038-43.

Wilson, Helen R., Daiguan Yu, Howard K. Peters, Jian-guang Zhou, and Donald L. Court. 2002. 'The global regulator RNase III modulates translation repression by the transcription elongation factor N', *The EMBO Journal*, 21: 4154-61.

Yarbrough, D., R. M. Wachter, K. Kallio, M. V. Matz, and S. J. Remington. 2001. 'Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution', *Proc Natl Acad Sci U S A*, 98: 462-7.

Yatsuyanagi, J., S. Saito, and I. Ito. 2002. 'A case of hemolytic-uremic syndrome associated with shiga toxin 2-producing *Escherichia coli* O121 infection caused by drinking water contaminated with bovine feces', *Jpn J Infect Dis*, 55: 174-6.

Yin, S., B. Rusconi, F. Sanjar, K. Goswami, L. Xiaoli, M. Eppinger, and E. G. Dudley. 2015. '*Escherichia coli* O157:H7 strains harbor at least three distinct sequence types of Shiga toxin 2a-converting phages', *BMC Genomics*, 16: 733.

Young, R. 2014. 'Phage lysis: three steps, three choices, one outcome', *J Microbiol*, 52: 243-58.

Zylicz, M., K. Liberek, A. Wawrzynow, and C. Georgopoulos. 1998. 'Formation of the preprimosome protects lambda O from RNA transcription-dependent proteolysis by ClpP/ClpX', *Proc Natl Acad Sci U S A*, 95: 15259-63.

Appendix

**A.1. *p*R' fragment amplification**
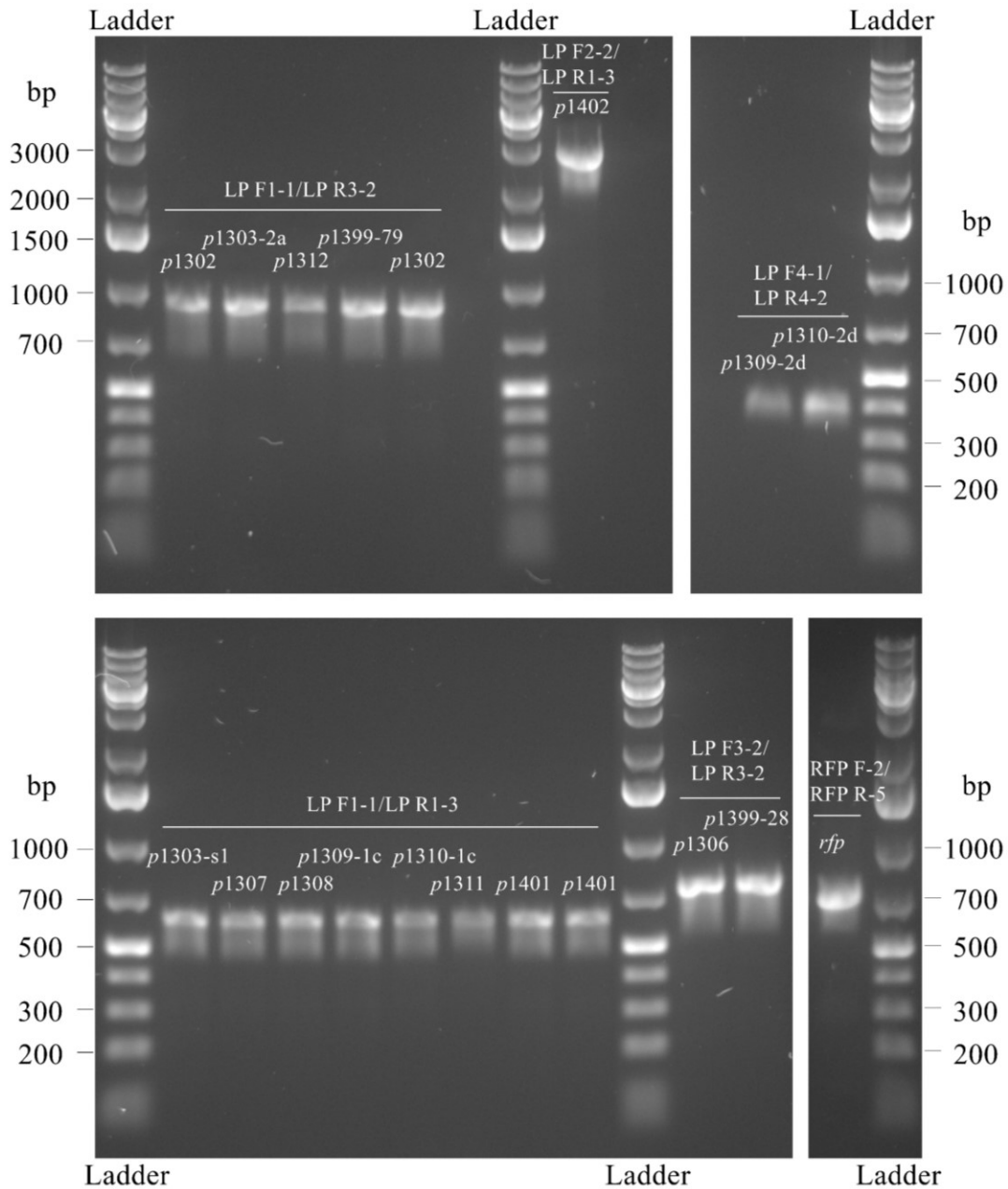
Figure A.1 PCR products of target fragment *p*R' from genomic DNA. Fragment names and primers were annotated on images in white color; molecular weights of ladder bands are shown aside the images.

A.2. Constructs and transformants obtained during molecular cloning.

Table A.1 Sixteen target *p*R' fragments used in cloning.

Target fragment	Fragment size (bp)	Toxin subtype
*p*1307	548	*stx1*
*p*1308	548	*stx1*
*p*1311	548	*stx1*
*p*1303-s1	548	*stx1*
*p*1403	522	*stx1*
*p*1309-1c	587	*stx1c*
*p*1310-1c	588	*stx1c*
*p*1402	2793	*stx1*
*p*1302	864	*stx2a*
*p*1303	865	*stx2a*
*p*1306	730	*stx2a*
*p*1312	864	*stx2a*
*p*1399-28	729	*stx2a*
*p*1399-79	864	*stx2a*
*p*1309-2d	378	*stx2d*
*Pp*1310-2d	418	*stx2d*

A.3. Digestion after transformed _p_R' vector into _E. coli_ O104 mutant strain, and into selected STEC target strains.
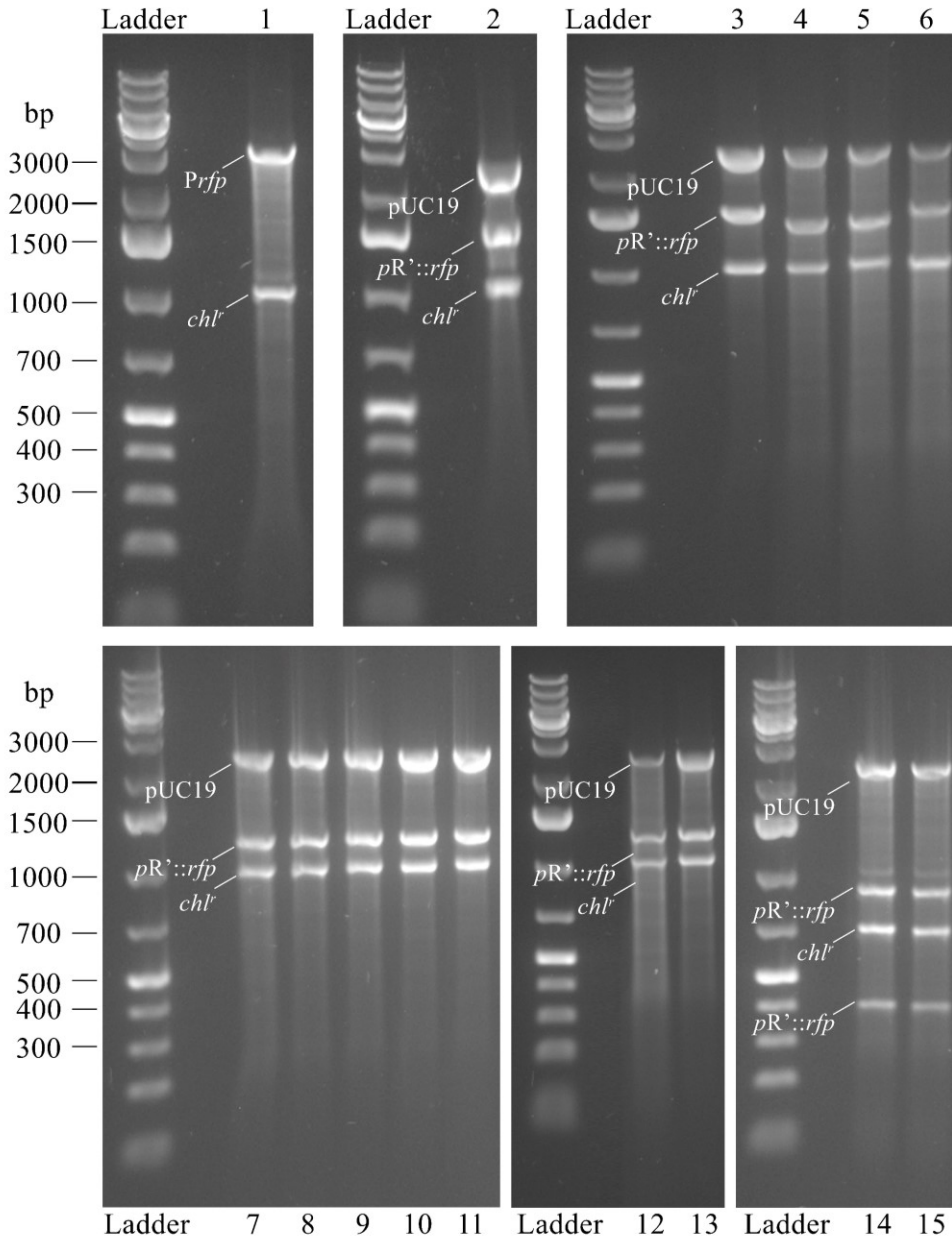
Figure A.2 Digestion for determining the positive clones. P*p*R'::*rfp*::*chl* constructs were transformed into DH5α and positive clones were checked by digestion, which were done before transformed into different target strains. The representations of each band were annotated in images.