Developing a Multiplex PCR Platform to Determine the Virulence Genes of *Escherichia coli* in Boulevard Lake, Thunder Bay, Ontario

A Thesis presented to

The Faculty of Graduate Studies

of

Lakehead University

by

Sultanah Albalawi

In partial fulfillment of requirements

for the degree of

Master of Science in Biology

Thunder Bay, Ontario, Canada

Contents

List of Tables

List of Figures

Acknowledgments

First, I would like to thank my supervisor Dr. Kam Tin Leung for his support of my research project. Without his guidance, and assistance during my scientific endeavors, this work would not have been a success. During my interaction with Dr. Leung, I have improved my knowledge and research skills during the past two and a half years. I would also wish to thank my committee members Dr. Heidi Schraft, Dr. Wensheng Qin, Ms. Elizabeth Pszczolko, for their time and support. I am also grateful to Ms Misung Yim for her assistance in conducting most of the experiments and for her training in use of various machines. Without her kind contribution, the success obtained in this project would not have been achieved. To all my student colleagues who helped me in one way or the other including Jocelyn Bel, Erica Tropea, Brian Hicks, Devon Prontack, Braden Peters, Nicholas Bel and John Winkler I say a big thank you all.

To my family members, particularly my dear husband who has been very supportive along my academic journey. Without his support and words of encouragement, I would not have achieved what I have achieved today. I cannot forget to thank my parents and sisters, I am most grateful to you for the support and encouragement that you gave me. During the difficult times, I always found solace and encouragement from you, which motivated me to move forward.

Finally, I am grateful to the Saudi Bureau in Ottawa for the support and funding that facilitated my study at Lakehead University.

Developing a Multiplex PCR Platform to Determine the Virulence Genes of *Escherichia coli* in Boulevard Lake, Thunder Bay, Ontario

Abstract of the thesis

Given the evidence that *E. coli* can establish in the periphytic community and be released back into the water column, it is important to understand the pathogenicity of this group of *E. coli* by examining the virulence genes that they possess. In this study, an optimized DNA extraction method and multiplex PCR platform were developed to detect the virulence genes of five major pathotypes of *E. coli* (enteropathogenic *E. coli*, ETEC; enteropathogenic *E. coli*, EPEC; shiga toxin-producing *E. coli*, STEC; enteroinvasive *E. coli*, EIEC; and uropathogenic *E. coli*, UPEC). The multiplex PCR platform was used to examine the virulence genes of *E. coli* isolates from the periphyton and lakewater samples from Boulevard Lake, Thunder Bay, Ontario. In addition, *E. coli* isolated from goose feces and sewer pumping stations around the lake were also tested.

Four DNA extraction methods were compared including the: (I) Fermentas Genomic DNA Extraction Kit, (II) XS Buffer method, (III) Chelex DNA extraction method, and (IV) Chelex+RNase method. The average amounts of DNA obtained by the four methods were 1.1, 14.9, 140.7 and 150.2 µg/10^9 cells, respectively. However, only the DNAs extracted by the XS Buffer method were able to be specifically amplified by their respective PCR primers. Therefore, the XS Buffer method was selected in this study.

Three sets of multiplex PCR primers were initially designed to target and amplify nine specific virulence genes of the five major pathotypes of *E. coli*. Since the multiplex primers Set 2 and 3 were not functioning properly, they were replaced by multiplex PCR

primers Set 4 and 5, respectively. The Set 1 primers contained the *hs* and *hl* primers that amplified the heat-stable enterotoxin (*hs*) and heat-labile enterotoxin (*hl*) genes of ETEC respectively, and *iron$_{EC}$* primers for the iron sequestering gene (*iron$_{EC}$*) of UPEC. The Set 1 multiplex primers successfully amplified the target genes either individually or simultaneously to produce specific DNA fragments at 170, 322 and 665 bp respectively. The Set 4 multiplex primers amplified the *ial* of EIEC, *bfp*A of EPEC and *hly* of UPEC to produce amplicons of 650, 324 and 1000 bp respectively. Finally, the Set 5 multiplex primers amplified the shiga-like toxin I and II genes (*stx*I and *stx*II) of STEC and *pap*A of UPEC successfully to produce amplicons of 150, 255 and 720 bp respectively.

The multiplex primers Set 1, 4 and 5 were used to determine the presence of the nine targeted virulence genes in 306 *E. coli* isolates isolated from the periphyton, lakewater, goose feces and sewage samples in or around Boulevard Lake. The percentage of *E. coli* isolates in the sewage (37 isolates), goose (38 isolates), periphyton (75 isolates) and lakewater (76 and 80 isolates collected in 2010 and 2014 respectively) samples that contained one or more virulence gene(s) were 48.5, 28.9, 2.6, 5.3 and 5.0 % respectively. The results indicate that the periphytons were likely to be the major source of *E. coli* in Boulevard Lake. Furthermore, with the exception of two potential diarrheagenic *E. coli* isolates from the sewage (with *hl* gene) and 2014 lakewater (with *bfp*A gene) samples, all the virulence gene-positive isolates belonged to the uropathogenic *E. coli* pathotype.

Chapter 1. Literature Review on Pathogenic and Environmental *Escherichia coli*

1.1 *Escherichia coli*

In 1885 *Escherichia coli* was first identified by a German-Austrian pediatrician Theodor Escherich. He discovered this organism through studies of the intestinal flora of infants and called it *Bacterium coli* (Rubino, Cappuccinelli and Kelvin, 2011; Percival et al., 2014). In 1895, the bacterium was reclassified as *Bacillus coli*. The name, *Bacterium coli* was vastly used until 1919 and later on Castellani and Chalmers reclassified the bacteria under the genus *Escherichia* and named the type species *Escherichia coli (*Priest and Barker, 2009).

Escherichia coli (*E. coli*) are rod shaped (about 2.0 µm in length and 1 µm wide), Gram-negative, non-spore forming and facultative anaerobic bacteria. Through the action of peritrichous flagella, *E. coli* are usually motile (Reshes et al., 2008). *E. coli* ferment lactose to form lactic acid as well as produce gases such as CO_2 and H_2 at 37 and 44 °C. Most *E. coli* strains possess the β-galactosidase enzyme therefore, they yield a positive ortho-nitrophenyl-β-D-galactoside (ONPG) reaction (Basu et al., 2008; Cai et al., 2012). Moreover, as glucose-fermenting organisms, they exhibit a positive methyl red reaction*,* which indicates acid fermentation. *E. coli* is a group of bacteria largely comprised of non-pathogenic members, which is representative of the normal intestinal microflora of both human and animal populations (Osman et al., 2012). These strains can benefit their hosts by synthesizing vitamin K2 and preventing colonization of the intestine with pathogenic bacteria (Gamage et al., 2006)

E. coli are of fecal origin and found in the gastrointestinal tract of humans and other warm-blooded animals (Skurnik et al., 2008). They constitute about 0.1% of the gut flora, and fecal-oral transmission is the major route through which pathogenic strains of the bacteria are spread (Percival et al., 2014).

Some unique *E. coli* strains were found in both temperate and tropical soils (Byappanahalli, Roll and Fujioka, 2012). Algae and periphyton were also shown to house *E. coli* as secondary non-host habitat. A study also showed that *E. coli* could persist through winter in the periphyton communities even when the air temperature reached a low of -40°C (Ishii et al., 2006; Ksoll et al*.,* 2007). This group of environmental *E. coli* is able to live through various environmental conditions. This is mainly because of its genetic diversity, which increases its adaptability and resistance to the environment (Goto and Yan, 2011).

1.2 Pathogenic type of *E. coli*

Although the majority of *E. coli* strains are harmless intestinal inhabitants, some *E. coli* strains have acquired virulence genes and cause serious illness. The pathogenic strains can be divided into two groups, including the Diarrheagenic *E. coli* and the Extraintestinal pathogenic *E. coli* (ExPEC). The Diarrheagenic *E. coli* strains are divided into six virotypes, which include the Enterotoxigenic *E. coli* (ETEC), Enteroinvasive *E. coli* (EIEC), Enteropathogenic *E. coli* (EPEC), Enterohaemorrhagic *E. coli* (EHEC), Enteroaggregative *E. coli* (EAEC) and Diffuse adhering *E. coli* (DAEC) (Osman et al., 2012; Chen et al., 2014). The Extraintestinal pathogenic *E. coli* (ExPEC) are a group of pathogens that is capable of causing urinary tract infections (classified as

Uropathogenic *E. coli* or UPEC), sepsis and meningitis (Wang et al., 2013 and Tan et al., 2012). The six virotypes of diarrheagenic pathogenic strains are distinguished from each other based on the differences between their surface antigens (such as lipopolysaccharide, flagellar and capsular antigens) and virulence factors (VFs). The expression of these virulence factors (VFs) manifests as symptoms related to diarrhea. The ExPEC strains have a more diverse group of VFs that allow them to colonize areas outside of the intestinal tract (Tan et al., 2012).

1.3 Enterotoxigenic *E. coli* (ETEC)

ETEC strains are one of the major causes of acute infectious diarrhea disease among children in developing countries (Wang et al., 2015). This watery diarrhea is accompanied with abdominal pain, malaise, nausea, and vomiting. Approximately 10^6 bacterial cells are required to cause infection (Qadri et al., 2005). ETEC is the most common cause of 'Traveller's Diarrhea' in developing countries, which attributes to about 30-50% of the cases reported (Wiedermann and Kollaritsch, 2006). In the United States and Ecuador, a survey has shown that there was a statistical relationship between the consumption of poor quality drinking water and infection of ETEC among children between ages 7-10 months (Croxen et al., 2013)

Osman et al (2012) demonstrated that ETEC, the second most prevalent pathogenic intestinal strains of *E. coli*, typically produce acute symptoms of diarrhea by the expression of either a heat labile enterotoxin (LT), a heat stable enterotoxin (ST) or both. Approximately 1/3 of ETEC strains produce only ST toxin, another 1/3 produce only LT toxin, while the remaining 1/3 produce both ST and LT toxins (Osman et al.,

2012). In addition to the production of enterotoxins, it also produces species-specific colonization factors, Type 1 pili and fimbria (CFA/I, CFA/II), which aids the infection of the ETEC. All ETEC strains contain plasmids that carry virulence genes. The virulence-associated genes on these plasmids can be detected by using gene probe techniques (Murray, Rosenthal and Pfaller, 2005)

To infect a host, the ETEC bacteria must adhere to the mucosal surface of the epithelial cells in the small intestine. The adhesion step is highly complicated. The pili of the ETEC help the bacterial cells to bind to the specific complex carbohydrate receptors on the epithelial cells. After adhesion, the ST and/or LT toxins secreted by the bacteria will cause damage on the host cells (Guevara et al., 2013; Croxen et al., 2013).

LT are subdivided into LT-I and LT-II. ETEC strains that are pathogenic to both humans and animals contain the LT-I. However, LT-II is usually found in strains that are pathogenic to animals (Rajkhowa, Hussain and Rajkhowa, 2009; Qadri et al., 2005). LT-I contains one protein A subunit and five identical protein B subunits. Subunits A and B have a molecular weight of 25000 and 115000 daltons, respectively (Percival et al., 2014). The five B subunits are important in binding the toxin to the epithelial cells whereas they are able to form a steady doughnut shape transmembrane structure with a central watery channel. The A subunit will activate the adenylate cyclase enzyme in the host cells and lead to a subsequent increase in intracellular cyclic AMP, which in turn disrupts sodium and chloride channels leading to an ion imbalance. As a result of these reactions, the intraluminal osmolarity increases (or a decrease in osmotic

potential in the lumen) and water is drawn into the gut. Loss of chloride ions and water into the intestine leads to the development of watery diarrhea (Hodges and Gill, 2010).

1.4 Enteropathogenic *Escherichia coli* (EPEC)

EPEC strains are frequently responsible for outbreaks of diarrhea in infants, and are the leading cause of death amongst children five years of age and younger in the developing world, despite the fact that they lack the genes coding for toxins normally required to elicit pathogenicity (Monaghan et al., 2012). Instead, EPEC strains rely on their ability to form attaching and effacing (A/E) lesions, which allow EPECs to adhere to the enterocyte membrane, facilitated by the membrane protein intimin, the 94-kDa outer membrane protein product of the *eaeA* gene (Pizarro-Cerdá and Cossart, 2006). This causes the rapid thinning of intestinal microvilli resulting in rapid secretion of a series of effector molecules into the host cells (Monaghan et al., 2012). These changes in the structure and thickness of the microvilli are also associated with changes in their actin cytoskeleton, and would alone be enough to cause the rapid loss of fluid and reduction in absorptive ability that is commonly associated with diarrhea. The wild type EPEC strains with the intimin protein are significantly more virulent than mutants that lack the intimin when tested in humans, and this indicates the important role of intimin's role as a virulence factor (Salyers and Whitt, 2002).

EPEC virulence genes are located on a 35-kb chromosomal pathogenicity island described as the locus of enterocyte effacement (LEE). Intimin protein is one of the essential proteins encoded by the pathogenicity island for the attachment of EPEC to the host epithelial layer (Franzin and Sircili, 2015). The EPEC diarrhea is a more

complicated disease than the ETEC diarrhea, and there are three stages involved in the lesions (Humphries & Armstrong, 2010). At first, the distinctive bundle-forming fimbriae enable the bacteria to attach to the enterocyte layer. In the second stage, signal transductions are stimulated in the enterocytes causing an increase of intracellular Ca^{2+} in the host cells, which leads to microvilli demise. In the third stage, the bacteria bind to the enterocyte membrane on the host cells. This process is arranged by the intimin protein. The intimin protein is also necessary for the actin adjustment, which leads to formation of the pedestal-like structure in the host cell (Schmidt and Hensel, 2006). These pedestal-like structures are important for EPEC pathogenesis.

The EPEC adherence factor (EAF) plasmid carries a gene encoding the subunit of bundle forming pili (BFPA) which is essential for the attachment of EPEC to the surface of the host cell. The bundle forming pili (BFP) exhibit a localized adherence (LA) pattern to HEp-2 cells (Bardiau, Szalo and Mainil, 2010). The mechanism of diarrhea caused by EPEC is not fully characterized, however, the bacterial attachment damages the absorptive surface causing the loss of absorptive power. This might be responsible for the diarrhea (Hodges and Gill, 2010).

1.5 Enterohemorrhagic *Escherichia coli* (EHEC)

Enterohemorrhagic *E. coli* (EHEC) is a subgroup of the STEC (shiga-like toxin producing *E. coli*) and is the major cause of food and water borne illnesses. Strains that make the shiga-like toxins (*Stx*) are called Shiga-like toxin producing *E. coli* (STEC) (Dini and De Urraza, 2010). The name EHEC is used to indicate strains that possess and express both the *stx* gene(s) and the pathogenicity island LEE. EHEC has been

described by its association with hemorrhagic colitis (HC) that is clinically distinguishable from shigellosis, and has different genotypic and phenotypic characteristics from EPEC (Croxen et al., 2013). *E. coli* O157:H7 is the most common serotype of EHEC, which causes serious human illnesses related to contaminated food and water. It should be noted, however, that there have been many studies focused on O157:H7 particularly in the molecular mechanisms of pathogenesis (Bielaszewska et al., 2011). In addition, there are other STEC strains that also cause significant human disease. Brooks et al (2005) demonstrated that serogroups O26 (22%), O111 (16%), O103 (12%), O121 (8%), O45 (7%) and O145 (5%) were the most common cause of non-O157 STEC infections that frequently occur among young persons during the summer season.

The major virulence factors which lead to the toxicity of EHEC are Shiga-like toxins (*Stx*). These toxins are reported as the common virulence factor for hemorrhagic colitis (HC) and hemolytic uremic syndrome (HUS). These *Stx* toxins are comprised of two main groups, *Stx*1 and *Stx*2 (Wani et al., 2009). The *Stx*1 consists of only a few variants, but there have been over 20 variations of the *Stx*2. Both *Stx*1 and *Stx*2 consisted of five B subunits and an A subunit. B subunits are responsible for binding to the host cell surfaces while the A subunit is essential for inhibition of protein synthesis. The EHEC genome also contains the LEE pathogenicity island (The locus of enterocyte effacement). It is a 35.6 kb cluster of genes that encodes the genes responsible for Attaching and Effacing (A/E) lesion phenotype (Hansen and Kaper, 2009).

The manner in which the EHEC *E. coli* debilitates the target cells is as follows. The *E. coli* will adhere to the mucosal lining of the gastrointestinal tract after being

ingested by the host. This is done through the fimbrial structure, which aids in adhesion. This allows for the toxins to be close enough to the cells to attack them. However, little is known about the fimbrial structure which is involved in the adhesion of EHEC (Gonyar and Kendall, 2013). Once the toxin is attached to the target cell, A subunit removes an adenine from 28S rRNA of the ribosome preventing protein synthesis which leads to cell death (Robins-Browne, 2013).

1.6 Enteroinvasive *Escherichia coli* (EIEC)

EIEC strains are typically implicated in food-borne infection cases of non-bloody diarrhea and are capable of breaching and multiplying within the colonic epithelial cells. This pathotype is mostly responsible for major food outbreaks in developed nations. From 10^6 to 10^{10} EIEC cells are required to cause an infection that is accompanied with watery diarrhea (Bando et al., 2010). Unlike most *E. coli* strains, EIEC are non-motile and do not ferment lactose. This group of *E. coli* resembles *Shigella* in terms of pathogenicity and both cause bacillary dysentery using a similar process (Croxen et al., 2013). The bacteria initially penetrate the intestinal epithelial cells (IEC) via transcytosis and gain access to the basolateral surface of the IEC. *E. coli* invades the intestinal epithelial cells, multiplies within the infected host cells and infects other nearby cells. This process depends on the *E. coli*'s ability to evade the macrophagic attacks and weaken the host's signaling pathways. Attachment of the bacteria to the host cell is facilitated by the interaction of the IpaBCD complex with the host's hyaluronic receptor CD44 and integrin respectively (Ud-Din and Wahid, 2014).

The EIEC and *Shigella spp*. are similar in antigenic as well as pathogenic traits. However, the adhesion process does not play a significant role for EIEC infection, which is a difference from the other *E. coli* pathotypes. Although the majority of the EIEC's VF genes are found on a large plasmid (220 kb) called pINV plasmid, little is known about the infection and pathogenic mechanism of the bacterium (Lan et al., 2004). The plasmid encodes a Type 3 secretion system (T3SS) which is essential for the bacterial invasion of the host cells. The toxins also destroy cells by binding to enterocytes of the large intestine leading to tissue damage and considerable inflammatory responses. There are no known animal reservoirs of EIEC; whereas infected persons are the main source of infection (Blackburn and McClure, 2002).

1.7 Enteroaggregative *Escherichia coli* (EAEC)

Enteroaggregative *Escherichia coli* (EAEC) were discribed for the first time in 1987, in Lima, Peru. It was discovered in a child who suffered from acute diarrhea. Since then, this disease has been connected with persistent diarrhea for children who reside in regions where EAEC is endemic (Torres, 2010). It is linked to individuals suffering from the human Acquired Immune Deficiency Syndrome (AIDS) and a cause of diarrhea for travelers to third-world nations. Diarrhea caused by EAEC is persistent and it usually lasts for more than fourteen days. EAEC are the most common bacterial pathogen that can be identified from the majority of diarrheal stool samples in the United States. The main source of EAEC outbreaks has been connected to food contamination (Okhuysen & DuPont, 2010)

Cravioto et al (1979) found that EAEC strains adhere to HEp-2 cells differently than the EPEC strains. EPEC strains have a localized adherence pattern. The aggregative adherence property of EAEC cells has a distinct auto agglutination phenotype on the surface of the HEp-2 cells. Currently EAEC strains are defined as *E. coli* that adhere to HEp-2 cells in an aggregative adherence manner (AA) and do not produce LT or ST enterotoxin (Torres, Zhou & Kaper, 2004).

The pathogenesis of Enteroaggregative *Escherichia coli* is unclear, and further research is being carried out in the field. However, many virulence properties and the histopathology of EAEC have been described. A three-stage infection mechanism is involved in EAEC pathogenesis as described by (Croxen et al., 2013; Regua-Mangia et al., 2004), (i) adherence to the intestinal mucosa, (ii) production of enterotoxins and cytotoxins, and (iii) mucosal inflammation. EAEC adherence is related to both fimbrial (aggregative adhesion fimbria [AAF]) and afimbrial adhesins. The result of the adhesion is mucosal damage and diarrhea. In the second stage, the EAEC cells secrete cytotoxins and enterotoxins in the intestinal mucosa (Donnenberg, 2013) and these toxins are responsible for causing microvillus vesiculation, enlarged crypt openings and increased epithelial cell extrusion (Harrington, Dudley and Nataro, 2006). A 108-kDa cytotoxin has been identified Eslava et al (1998) and this protein belongs to an auto transporter family of proteins that is responsible for the enterotoxic activity of the bacteria. There is also a 220-kDa protein that is essential for increasing the intracellular calcium levels in the HEp-2 cells. The calcium presents as a second messenger of the EAEC toxins and is important in the loss of microvilli via membrane vesiculation (Weintraub, 2007). Finally, the EAEC cells are able to catalyze the release of

inflammatory mediators during the final stage. The inflammation is dependent on several factors such as the innate immune system of the host cell and the EAEC strain involved (Estrada-Garcia & Navarro-Garcia, 2012). EAEC has a characteristic of forming a mucuous biofilm that results in cytotoxic effects on the intestinal mucosa. The mucuous biofilm formation is thought to be related to the ability of the organisms to cause persistent colonization and diarrhea (Weintraub, 2007).

1.8 Diffusely Adherent Escherichia coli (DAEC)

Diffusely Adherent *E. coli* (DAEC) are widely considered as a diarrheagenic category of *E. coli,* and numerous potential virulence factors have been studied in the past few years. Despite numerous studies being carried out on DAEC, the results have been inconsistent. The pathogenic bacteria are linked to diarrhea, and like EAEC, it is more prevalent in developing or third world nations. DAEC can be identified through their unique pattern of adherence on the surface of epithelial cells. The bacteria will cover the whole surface of the Hep-2 cells in a scattered pattern. It is still not clear if all DAEC infections will cause diarrhea because some studies show that the presence of DAEC in children sometimes may not produce the expected diarrhea symptom (Donnenberg, 2013).

The pathogenic mechanism of DAEC is not fully understood, but in most cases, it is associated with watery diarrhea that can be persistent in young children from the age of 18 months to 5 years (Kaper et al., 2004; Servin, 2005). Outbreaks of diarrhea are very common in Brazil, and in most cases, they are related to DAEC. Some studies

have also shown that DAEC infections are mostly associated with young children (Ochoa et al., 2009).

Bilge et al (1989) have illustrated that surface fimbria are responsible for the diffuse adherence (DA) phenotype of DAEC. The essential genes for encoding the fimbria are located on either the bacterial chromosome or a plasmid. DA phenotype is associated with 100-kDa outer membrane proteins (OMP) in the serotype 0126: H27 (designated as AIDA-I). Another adhesion factor that encodes the DA phenotype in DAEC is F1845 and it is essential for the adherence of the bacterial cells to the entire surface of the epithelial HEp-2 cells.

1.9 Extraintestinal pathogenic *E. coli* (ExPEC)

Extraintestinal pathogenic *E. coli* cause urinary tract infection (UTI), septicemia and neonatal meningitis. UTI is most frequently caused by uropathogenic strains of *E. coli* (UPEC strains) and the majority of the infections are in women (Manges et al., 2008). Men also have urinary tract infections, but the percentage of male infections is much lower. The occurrence of ExPEC-induced diseases increases with patient's age. Therefore, an increase in the elderly population worldwide may mean an increase in ExPEC infections (Smith et al., 2007). The uropathogenic group is responsible for 7 million urinary tract infections in the United States reported annually. In terms of morbidity and mortality, ExPEC poses a significant risk to public health and the medical costs can be more than a billion dollars each year (Johnson and Russo, 2002; Smith et al., 2007).

Diarrhea-associated *E. coli* strains do not generally cause extraintestinal diseases. Also, diarrhea is not normally induced by those strains that cause extraintestinal illnesses. ExPEC strains are either harmless as part of the normal flora in human intestinal tracts or serious pathogens when they enter the urinary tract, bloodstream, or cerebrospinal fluid (Bien, Sokolova and Bozko, 2012). ExPEC possesses specific combinations of virulence traits that enable the bacteria to invade, colonize, and induce disease in organs outside of the gastrointestinal tract. Its ability to colonize the host is an important step in causing urinary tract infections.

Adherence is the preliminary step in bacterial colonization and is necessary for pathogenesis. For ExPEC, there are important adhesion factors that work together to promote successful bacterial colonization. Type 1 pili are essential for colonizing the bladder (Melican et al., 2011). They bind the mannose residues on the surface of transitional uroepithelial cells on bladder glycoproteins followed by invasion of uroepithelial cells (Sivick and Mobley, 2009). Another important adhesion factor is P fimbriae which allow the binding between the bacteria and the epithelial cells. P fimbriae are only associated with the UPEC strains and are not present in non-pathogenic *E. coli* strains. The pyelonephritis-associated pili genes (*pap* genes) are responsible for encoding proteins that are involved in the synthesis and assembly of P fimbriae (Lane & Mobley, 2007). Afimbrial adhesins a group of other adhesion factors produced by UPEC strains that do not produce P fimbriae, and they do not have fimbrillar structure (AFAI, AFAIII).

Colonization of the bladder causes an inflammatory response that is responsible for the symptoms of an acute urinary tract infection. UPEC strains also produce

exotoxins. They excrete a hemolysin protein, which is generally called alpha-hemolysin (HlyA) (Schmidt and Hensel, 2004). HlyA is part of a group of proteins called RTX toxins, which contain repeats of a nine-amino acid sequence. RTX toxins' function is to create pores in eukaryotic cell membranes. *E. coli* strains that do not produce alpha-hemolysin are less virulent than the *hly*A producing strains (Los et al., 2013). There are some other toxins produced by UPEC such as the cytotoxic necrotizing factor 1 (CNF-1) that acts as a killer for the epithelial cells. Iron acquisition is another important virulence factor of UPEC. Iron is necessary for UPEC to cause infection, therefore, UPEC synthesize siderophore-based iron acquisition systems and use them to get the iron away from the host-iron bound protein (Wiles, Kulesus and Mulvey, 2008).

Most of the virulence genes of UPEC are located on the chromosome. These include the pili subunit, exotoxins and mobility genes of the bacteria. Some of these virulence genes are present in DNA segments called pathogenicity islands, which usually possess different guanine and cytosine content from the rest of the bacterial genome (Lloyd, Rasko and Mobley, 2007). The pathogenicity islands of the UPEC are about 25 to 100 kbp and they control the virulence properties (bacterial pathogenesis) of the bacteria. It is possible for one UPEC strain to possess more than one pathogenicity island. *PapA* gene is an important virulence gene of UPEC. It is located in the chromosome and is required for the synthesis of P fimbriae. However, the process of the UPEC strains to enter the intestine is still unknown (Jacobsen et al., 2008). The major UPEC virulence genes are listed in Table 1.1.

Table 1.1 Major virulence factors of Uropathogenic *E. coli* (UPEC)

Virulence genotype	Gene encodes	Reference
$Iron_{EC}$	Siderophore receptor	
Hly	Hemolycin	
*Cnf*1	Cytotoxic necrosis factor 1	(Chapman et al., 2006)
*Pap*C	Genes of P fimbriae operon	
*Pap*A	Pyelonephritis-associated pili	

1.10 Detection of pathogenic *E. coli*

The presence of pathogenic *E. coli* in water or food is a threat and concern to public health worldwide. In developing countries, outbreaks are due to poor sanitation and insufficient water treatment systems, and in the developed countries threats are due to growing antimicrobial resistance and inadequate detection measures (Kulpakko et al., 2014; Silva & Domingues 2014; de Man et al., 2014; Vale et al., 2009). *E. coli* is one of the most common causes of pathogenic infections and is a major causative agent of diarrhea and urinary tract infections (UTI) (Kulpakko et al., 2014). Therefore, detection of pathogenic *E. coli* is an important step to protect the safety of our water resources (Molaee et al., 2015).

Preventing outbreaks of pathogenic *Escherichia coli* requires early detection of the pathogens. Therefore, various methods have been developed for *E. coli* detection in drinking and recreation waters (Percival & Williams 2014; Muniesa et al., 2006). These include the culturing method, the molecular method and immunological based methods. These methods rely on the counting of viable bacteria, genomic DNA analysis and antigen-antibody interactions, respectively (Lazcka et al., 2007).

The culture-based methods are the traditional approach for bacterial detection and they have many limitations. Therefore, it is essential to employ other techniques because culturing methods are time consuming compared to the molecular-based methods (Espy et al., 2006). In addition, viable but non-culturable bacteria coming from environmental sources are not detectable using culture-based methods, which results in false-negative analysis of *E. coli* in water (Liu et al., 2008). These limitations affect public health safety leading increases in health care costs and in costs spent in water quality control management (Silva and Domingues, 2015). Several culture and colony counting methods have been reported for detecting *E. coli* O157:H7 in water (Ngwa et al., 2012; Maheux et al., 2012). Maheux et al (2014) tested three different chromogenic culture-based methods, MI agar, Chromocult and DC media to detect *E. coli* colonies in water in order to assess microbiological water quality. They were able to detect *E. coli* colonies using MI agar medium, while the *E. coli* counts using Chromocult and DC media were different when the non-*E. coli* background concentrations were higher (i.e. the CFU count exceeded 10^3 CFU/100ml). Sen et al (2011) developed an effective culture enrichment-qPCR method to detect viable *E. coli* O157: H7 in drinking water. But his work mainly relied upon the use of PCR for the detection of pathogenic *E. coli*. Thus, culture based method alone cannot distinguish the pathotypes of *E. coli* in absence of other complementary techniques such as biochemical typing and molecular methods. Additionally, these techniques are time consuming (Sen et al., 2011).

Recently, molecular-based methods for pathogen detection have significantly increased the speed and specificity in identifying pathogens (Horakova et al., 2008).

Molecular techniques are being refined constantly to insure the reliability and sensitivity of the detection methods (Girones et al., 2010).

Polymerase chain reaction (PCR) is a technique that relies on nucleic acid amplification. Developed in 1983 by Kary Mullis, this technique amplifies a fragment of DNA generating millions copies (Lazcka et al., 2007). Different PCR methods have been explored to detect *E. coli* and Diarrheagenic *E. coli* in water samples. These include (i) singleplex PCR, (ii) multiplex PCR and (iii) quantitative PCR (qPCR) (Liu et al., 2012). Using these techniques with specific probes and primers leads to a rapid detection with a high level of sensitivity and specificity that bring forth significant information regarding the existence and quantity of classical and emergent pathogens in water. Therefore, PCR-based methods are widely used to assess and monitor water quality by screening a large group of pathogens in contaminated water (Girones et al., 2010).

Multiplex PCR (mPCR) is a PCR-based technique that detects several different DNA targets at the same time, this process utilizes multiple primers to amplify genomic DNA regions coding for specific genes of the targeted bacteria (Souza et al., 2013; Touron et al., 2005). Molecular techniques are considered to be more efficient in detecting and quantifying many pathogens and indicator bacteria in the environment because many of these pathogens and bacteria are not culturable in microbiological growth media available to date. Besides their ability to determine the source of fecal contamination in water, they are also useful in assessing water quality because qPCR techniques allow quantitative estimation of pathogens' concentrations in water (Girones et al., 2010; Albinana-Gimenez et al., 2009). However, there are some limitations using

PCR-based techniques. First of all, it can be only used for targeting known genes (Smith & Osborn, 2008). Furthermore, the major challenge for performing PCR technology on environmental samples is to remove PCR inhibitors, such as humic substances, from the DNA extract. Therefore, the PCR protocols are required to be optimized and standardized for specific environmental samples. Nevertheless, over the recent past, different researchers have made use of multiplex PCR for determining and distinguishing specific groups of pathogenic *E. coli*. A multiplex PCR method was designed by Horakova et al (2008a) to detect *E. coli* isolated from water samples. Using four target genes *uidA*, *lacZ*, *lacY* and *cyd,* they were able to differentiate *E. coli* from other bacteria. Botkin et al (2012) reported the use of mPCR for the consistent amplification of genes specific to the prototype of enterohemorrhagic *E. coli* , O157:H7 strain EDL933 (*lpfA1-3*, *lpfA2-2*, *stx1*, *stx2*, and*eae*-γ), and enteropathogenic *E. coli*, strain O127:H6 E2348/69 (*eae*-α, *lpfA1-1*, and *bfpA*) (Botkin et al., 2012). In another instance, (Guion et al., 2008) developed a novel real-time fluorescence-based multiplex PCR technique for the detection of various types of *E. coli* strains which caused diarrhea (Guion et al., 2008). Similarly, Tobias et al (2012) reported a two-step multiplex PCR method for the identification of diarrheagenic *E. coli* (DEC) strains including enteroaggregative, enterotoxigenic, enteropathogenic and enterohemorrhagic types. In his approach, he used mPCR to determine DEC by targeting *cdv432*, *lt, sth, stp, eae*, *bfp*, *stx*1, and *stx*2 genes (Tobias et al*.,* 2012).

1.11 *E. coli* as a fecal indicator organism

Fecal contamination is the leading cause of water-borne diseases in many parts of the world. It can be detected using two principal methods, microbial and chemical indicators (Madigan & Brock, 2012). Microbial tests involve testing for primary organisms contained in the human or animal gastrointestinal tract such as coliform bacteria, *E. coli* and enterococci tests (Wade et al., 2003). Initially, presence of fecal coliform bacteria was the most commonly used fecal indicator but has since been replaced by *E. coli* and *Enterococcus* because these two are better indicators. Chemical indicators involve testing for chemicals produced by humans, including those that pass through the human GIT, or those associated with sewage (Isobe et al., 2004). An example of chemical indicator is the fecal sterol, coprostanol, a product of bacterial degradation of cholesterol in human GIT. The amount of coprostanol sterol in humans is greater than in animals. Therefore, this method can be used to discriminate fecal contamination between humans and that of animals (Glassmeyer, 2005).

Microbial fecal indicator methods involve the use of specific indicator organisms. The primary indicator organisms include coliform bacteria and fecal streptococci (or fecal enterococci). The U.S. Public Health Service suggested the use of coliform bacteria as an indicator. Two tests were developed, the total coliform and the fecal coliform tests (Bower et al., 2005). Coliform bacteria are a large group of organisms that belong to the *Enterobacteriaceae* family. The coliforms include Gram-negative, aerobic, and facultative anaerobic microorganisms. The coliform group includes *Escherichia coli*, *Enterobacter*, *Klebsiella,* and *Citrobacter* species. The group is largely found in human and animal fecal matter (Wheeler Alm, Burke and Spain, 2003). Consequently, coliform tests have been used to detect fecal contaminations in the environment. However, the

coliform group contains organisms of fecal and non-fecal origin. Fecal coliforms are different from the previously discussed coliform group. Fecal coliform bacteria are thermotolerant and lactose fermenters at approximately 44.5°C. Fecal coliforms are a better fecal indicator than the total coliforms because they have a higher correlation to fecal contaminations (Sadowsky & Whitman, 2011). The criteria for fecal indicator organisms are as follows (Madigan et al., 2012):

a. The indicator should be consistently and exclusively associated with the source of pathogens.

b. It should be present only with the existence of the pathogens and lacking in uncontaminated samples.

c. The numbers should be greater than that of the pathogens.

d. It should be equally resistant as the associated pathogens to the environmental stressors and disinfection.

e. It should not multiply in the environment and should be non-pathogenic.

f. It should be readily detectable using rapid and inexpensive methods.

Because of its prevalence in the digestive tracts of the vast majority of warm-blooded animals, *E. coli* acts as an invaluable indicator organism to determine fecal contamination of water supplies (Bower et al., 2005). It has been used as a fecal indicator for over a century worldwide. It can easily be distinguished from other coliform groups because it has β-glucuronidase and lacks urease. Furthermore, methods used to test *E. coli* are simple and inexpensive. *E. coli* meets the criteria required for indicator organisms and studies have also shown that it is a much better fecal indicator than both the total coliform and fecal coliform fecal indicators.

The presence of *E. coli* population densities of greater than 100 colony forming units (CFUs) per 100 mL in recreational water, as indicated by the Canadian Recreational Water Quality Guidelines, is a sufficient cause to issue water quality advisories. The presence of *E. coli* in drinking water means that water is contaminated by fecal materials and is unsafe for drinking. For drinking water tests, no *E. coli* should be detectable in 100 ml of water (0 CFU / 100ml).

Although *E. coli* is considered to be the most commonly used indicator for fecal contamination in water, recent studies have shown that *E. coli* can survive and replicate in the environment. In a study conducted by (Ksoll et al., 2007), several strains of *E. coli* were isolated from epilithic periphyton samples. This indicated that these organisms had become naturalized in the environment. Persistence of the naturalized *E. coli* in the environment brings its use as a reliable fecal indicator into doubt. Therefore, false positive tests may occur in some cases.

1.12 Naturalized *E. coli*

Based on recent studies, *E. coli* can survive and grow in the environment. Moreira et al (2011) measured the cell densities of *E. coli* residing in the epilithic periphyton of Boulevard Lake, Chippewa Park Beach and Billy Lake, Ontario. They also found that the periphytic *E. coli* populations were continually present in all these locations.

E. coli persistence is dependent on their survival rate, their capability to acquire nutrients and to produce adhesion structures that aid in attachment to surfaces. This is the primary reason that *E. coli* can survive in soil, manure, irrigation water (van Elsas et al., 2011). They also take shelter inside the plant so that it is not easily detached from

31

the plant surface due to rain or irrigation. Due to their ability to survive on various surfaces, *E. coli* is able to further thrive and move to human hosts. In natural environments, *E. coli* can survive if that there are enough resources available. *E. coli* populations were able to survive in soil even though the temperature exceeded 30°C, and the soil was usually nutrient poor (van Elsas et al., 2011).

There are various abiotic conditions that help *E. coli* to survive in the environment such as temperature, and availability of water. Temperature is very important for *E. coli* survival (Blaustein et al., 2013), as *E. coli* prefer the homeostatic conditions of a host body, as opposed to the fluctuating temperature occurring in the environment. The availability of water also determines the extent to which *E. coli* survives and thrives. There are two extreme conditions under which *E. coli* survival decreases, (i) extremely low still water content, and (ii) when the soil is saturated which causes an aerobic condition.

E. coli has been studied from numerous sources, such as soil, sand, temperate lake water, algae and freshwater periphyton (Moreira et al., 2011). Periphytic *E. coli* strains are found on shoreline rocks and are able to survive through the harshness of winter. Planktonic *E. coli* are free living and don't adhere to a substrate, but rather move with the flow of the water. Although they do not attach to a substrate they produce extracellular polymeric substances, which could help them to survive in the environment. Throughout the literature it has been shown that the planktonic *E. coli* are the least prevalent *E. coli* because they are lacking protective structures and stress-survival mechanisms. On the other hand, biofilm *E. coli* are found in low nutrient environments of temperate freshwater lakes. They are capable of forming biofilms under temperate

32

conditions and this capacity to form biofilms is related to their strategy towards mitigating issues related with environmental stress (Meshram et al., 2012). Surface proteins such as conjugative pili, type I fimbriae and curli are involved in adhesion to surfaces. This ability to form biofilm and curli allows the *E. coli* to persist in the periphyton of the temperate freshwater lakes. (Vogeleer et al., 2014; Meshram et al., 2012).

The fact that the naturalized *E. coli* can survive in such variable conditions, compromises *E. coli* as an indicator of fecal contamination in water. Given that naturalized *E. coli* can be abundant in the environment and only limited information about their pathogenicity can be found in the literature, it is important to determine the potential pathogenicity of this group of *E. coli*.

1.13 Thesis Objectives

E. coli is used as a fecal indicator for both drinking water and recreational water. It has been documented that the *E. coli* population in Boulevard Lake, Thunder Bay, Ontario increases above the recommended limit of 100 CFU/100 ml lakewater several times every summer causing swimming advisories. In previous studies on Boulevard Lake, it was shown that there were three possible direct sources of *E. coli*, which were geese feces, sewage contamination, and periphyton. To date, only a few studies have looked at the pathogenicity of the *E. coli* isolated in recreation all waters and among these studies, none has examined the pathogenicity of the periphytic population of *E. coli*. I hypothesize that certain sub-populations of the *E. coli* released from geese and the sewage into the water are likely to be pathogenic, but the periphytic *E. coli*

population in Boulevard Lake is non-pathogenic. Furthermore, the *E. coli* detected in the water of Boulevard Lake should reflect its source(s) of contamination. In order to determine the pathogenicity of the *E. coli* in this study, a multiplex PCR assay has been developed to determine specific virulence genes of various pathotypes (or virotypes) of *Escherichia coli*. The specific research objectives of this study are as follows:

1. To develop a simple and effective DNA extraction method for *E. coli* culture samples.

2. To develop a PCR assay to detect various virotypes of pathogenic *E. coli.*

3. To develop a multiplex PCR assay to detect various virulence genes of pathogenic *E. coli*.

4. To determine the pathogenicity of *E. coli* isolates isolated from the periphyton, goose and sewage samples by the multiplex PCR assay.

5. To determine the pathogenicity of *E. coli* isolated from the water samples collected at Boulevard Lake, Thunder Bay, Ontario.

Chapter 2. Developing a Multiplex PCR Platform to Determine the Virulence Genes of *Escherichia coli* in Boulevard Lake, Thunder Bay, Ontario

2.1 Introduction

Pathogenic *E. coli* represent a major public health issue worldwide. Based on the virulence factors expressed by these bacteria and the type of gastrointestinal disease they cause, diarrheagenic *E. coli* have been classified into six virotypes, including enterohemmorhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic, *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), Enteroaggregative *E. coli* (EaggEC) and diffusely adherent *E. coli* (DAEC). Among these six types, the first four are the majors cause of *E. coli*-linked diarrheal diseases (Sullivan et al., 2006). Extraintestinal pathogenic *E. coli* is another group of *E. coli* pathogens that cause urinary tract infections, meningitis and septicemia (Jafari et al., 2012). The group that causes urinary tract infection is classified as uropathogenic *E. coli* (UPEC).

One of the most common methods to determine the pathogenicity of an *E. coli* isolate has been to determine the virulence genes that it carries. Each pathogenic group of *E. coli* possesses specific virulence factors for pathogenicity and these virulence genes can be used as biomarkers to determine the pathogenicity of *E. coli* isolates. Some studies have shown that a DNA-based approach (utilizing multiplex-PCR, MP-PCR) is a very reliable method for identifying and differentiating pathogenic strains of *E. coli*. MP-PCR is a modification of the polymerase chain reaction where a DNA sample is amplified using multiple pairs of primers to rapidly detect multiple genetic biomarkers or targets simultaneously (Markoulatos et al., 2002; Rappelli et al., 2001). Numerous MP-

PCR assays have been developed to differentiate diarrheagenic *E. coli* from other pathogens (such as diarrheagenic *Shigella* species) that cause similar disease symptoms. MP-PCR has also been used to determine the presence of virulence genes (or pathogenicity) in pathogenic *E. coli* (López-Saucedo et al., 2003).

Boulevard Lake is a popular recreational lake in Thunder Bay, Ontario and it attracts a large number of swimmers, especially during the summer. But in many instances, swimming advisories are issued for the Boulevard Lake beach due to high levels of *E. coli* in its water. So the focus of our research is to determine the sources and pathogenic types of *E. coli* present in this lake. From our past work, we concluded that the *E. coli* population observed in Boulevard Lake was due to three possible sources, the periphytic *E. coli* communities, the goose droppings and the human sewage contaminations. Among the three sources, the periphytic *E. coli* community contributed to about 50-80% of the *E. coli* population in the water column (Yang, 2012). Because the standard recreational water *E. coli* testing by the public health authority does not include determining the pathogenicity of the bacteria, the pathogenicity of the *E. coli* populations in both the periphyton communities and the water column of Boulevard Lake were not clear. Therefore, our present work focused on determining the pathogenicity of *E. coli* isolated from the lakewater, periphytic, goose and sewage samples and the following objectives were developed to achieve the goal of this study systematically.

1. To develop a rapid and economical DNA extraction and purification method to recover DNA from *E. coli* samples.

2. To Develop a set of virulence gene biomarkers and their PCR primers to determine the different virotypes of pathogenic *E. coli*.

3. To develop and optimize a multiplex PCR method to detect the virulence gene biomarkers of *E. coli*.

4. To determine the pathogenicity of *E. coli* isolated from the lake water, periphyton, goose fecal and raw sewage samples by the optimized multiplex PCR assay.

2.2 Materials and Methods

2.2.1 Bacterial strains & storage

E. coli isolates used in this study included various known and unknown type strains and isolates. Five pathotypes of *E. coli* were included in this study. They were the STEC (STEC 920004, STEC 920026), ETEC (ETEC 05, ETEC 505), EIEC (EIEC 0136, EIEC 0164), EPEC (EPEC 055, EPEC 2348), and UPEC (UPEC 25922). A non-pathogenic *E. coli* SY327 was included as a negative control for the virulence genes' PCR primers (Table 2.1). Furthermore, a total of 306 unknown *E. coli* isolates were obtained from various sources including the periphyton (75 isolates), human sewage (37 isolates), goose faeces (38 isolates) and lakewater (156 isolates) samples.

The water and periphytic samples were obtained from three locations of Boulevard Lake, Thunder Bay, Ontario, Canada. The first two sites (site1 and site 2) chosen for sampling were adjacent to the main beach whereas the third sampling site was at a rocky shore located 500 m downstream of the main beach (site 3). The water and periphytic samples were taken on April 13, May 24, June 28, July 28, August 25, September 28, and November 18 for 2010. For the water samples, three sterile 1L

Nalgene bottles were used at each site to collect water approximately one foot under the surface. To avoid disturbance to the periphyton and the sediment, water samples were collected at locations away from the shore and downstream of the periphyton samples where cross-contaminations between the water samples and the periphytic samples were minimal. For the periphyton samples, sampling was carried out by removing a submerged rock (three rocks from each site) located approximately 30-40 cm below the water's surface. However, the rock was shaken gently prior to removal in order to eliminate loose sediments that are attached to its surface. Once taken out of water, the rock was placed on shore with its surface to be sampled facing up. Substratum surface area was quantified by placing a sterile, square rubber template (10 cm x 10 cm) on the rock. The periphyton samples within the area marked by the square template were scraped off by sterile spatula and suspended in 50 ml of sterile phosphate-buffered saline (PBS; 8.00 g NaCl, 0.20 g KCl, 1.44g Na_2HPO_4, 0.24 g KH_2PO_4, 1L distilled H_2O,pH adjusted to 7.4). The lakewater and periphyton samples were drawn through a sterile 47-mm mixed cellulose ester filter (Fisherbrand water testing membrane filter, pore size 0.45μm; Thermo Fisher Scientific, Whitby, ON, Canada) using a sterile pneumatic pump funnel filtration apparatus. The filter was placed face up on Differential Coliform Agar (Oxoid Limited, Basingstoke, England) and incubated overnight at 37°C. Blue colonies were presumptively identified as *E. coli*. In addition to the 76 *E. coli* isolates obtained in 2010, an addition of 80 *E. coli* isolates that were isolated from the lakewater samples collected at the Boulevard Lake beach by the Thunder Bay district health unit and submitted for testing to the Thunder Bay Public

Health Laboratory between July 2 and August 25, 2014 were also examined in this study.

Sewage *E. coli* isolates were isolated from raw sewage samples collected monthly from July to September in 2010 at the Thunder Bay Sewage Treatment and Water Pollution Control Plant. Geese were considered a potential source of *E. coli* as they were observed around the lake throughout the summer season. Goose fecal samples were collected from the periphery of Boulevard Lake main beach area from July to September in 2010. *E. coli* bacteria were isolated from the fecal samples by the same filtering method used for the lakewater and periphyton samples. Thirty-eight goose isolates were analyzed in the study.

Individual presumptive *E. coli* colonies from the Differential Coliform Agar were streaked onto mFC agar (membrane Fecal Coliform agar. Becton, Dickson and Company, Spark, MD, USA) and incubated at 37 ^{o}C for overnight in an incubator (Isotemp 205 incubator, Fisher Scientific) for growth. All the *E. coli* isolates were grown in Luria-Bertani (LB) broth (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) at 37 ^{o}C and stored at -80 ^{o}C using 25% glycerol (v/v) as a cryoprotectant. Table 2.1 shows the positive and negative control *E. coli* strains included in our study.

2.2.2 DNA extraction methods

Individual *E. coli* strains were streaked onto Luria Bertani agar (LBA) from frozen stock and incubated at 37 °C overnight. From the respective plates, a single colony was picked and cultured in 5ml of TSB (Tryptic Soy Broth. Becton, Dickinson and Company, Sparks, MD, USA) and incubated overnight at 37 ^{o}C under aerobic condition in a rotary

shaker (Innova 4430 Incubator Shaker, New Brunswick Scientific) at 150 rpm. Cell density of each isolate was measured using a UV-Vis spectrophotometer at 600nm. Extraction of DNA was carried out using four different methods to make a comparative study between the methods and the best method was selected based on the quantity and quality of DNA obtained. The four methods were: (I) Extraction using Genomic DNA Kit; (II) XS buffer (Potassium ethyl xanthogenate) method; (III) Chelex DNA extraction method; and (IV) Chelex+ RNase method. All the methods are briefly described below. One ml of each *E. coli* cell culture sample (at a cell density of OD_{600nm} 1) was harvested and its DNA was extracted and purified by the following methods. For each extraction method three replications were applied for every *E. coli* sample.

2.2.2.1 Protocol I. Genomic DNA extraction and purification kit

DNA extraction was carried out using the Genomic DNA Extraction and Purification Kit (Fermentas, Waltham, MA, USA) according to the manufacturer's instructions with some adjustments.

a.   One ml of *E. coli* cell culture sample was centrifuged for 5 minutes at 14,000 Xg and the supernatant was discarded. The cell pellet was re-suspended in 1 ml of TE buffer (Tris-EDTA buffer) then centrifuged for 2 mins at 14,000 Xg. This step was repeated one more time.

b.   After washing, the cell pellet was suspended with 400 µl of lysis solution and incubated at 65 ^{o}C for 5 mins.

c.   Six hundred µl of chloroform was immediately added to this mixture, gently mixed to form an emulsion and centrifuged at 14,000 Xg for 2 mins.

d.   Precipitation solution was then prepared by mixing 720 µl of sterile deionized water with 80 µl of 10x precipitation solution that was supplied with the kit.

e.   After centrifugation in step c, upper aqueous phase containing the DNA was transferred to a new tube and 800 µl of freshly prepared precipitation solution (as in step d) was added to it. The solution was then mixed by several inversions at room temperature (for 1-2 mins) and then centrifuged at 14,000 Xg for 2 mins.

f.   Supernatant was then discarded and the DNA pellet that was settled in the bottom was dissolved in 100 µl of 1.2 M NaCl solution by gentle vortexing.

g.   Two µl of 10mg/ml Rnase were added to the sample mixed by vortexing then incubated for 10 min at 37 ^{o}C.

h.   DNA was then precipitated by adding 300 µl of cold ethanol and the sample was chilled for 10 min at -20 ^{o}C. The DNA precipitate was spun down for 3-4 min at 14,000 Xg.

i.   The ethanol was poured off and the DNA pellet was washed two times with 1ml of 70% of cold ethanol and the DNA pellet was then dissolved in 100 µl of sterile deionized water by gentle vortexing.

j.   DNA concentration was quantified using spectrophotometer and also visually using gel illuminator after electrophoresis in agarose gel.

2.2.2.2 Protocol II. Chelex DNA extraction and purification method

The Chelex 100 grade resin was purchased from (Bio-Rad, Mississauga, ON, Canada). Chelex extraction method was adopted from (Lamallerie *et al*., 1992) with modifications. Briefly, the steps were as follows:

a.   One ml of *E. coli* cell culture sample was centrifuged for 5 minutes at 14,000 Xg and the supernatant was discarded. The cell pellet was re-suspended in 1 ml of sterile double distilled water.

b.   The washing step "a" was repeated one more time and the cell suspension was centrifuged again at 14,000 Xg for 5 min to obtain the cell pellet.

c.   To the above cell pellet, 100 µl of 12% Chelex solution and 10 µl of 10 mg/ml protease K were added, vortexed and incubated at $55^{o}C$ in a water bath for 30 minutes and then in boiling water for 8 minutes.

d.   This was followed by centrifugation at 14,000 Xg for 5 min and the supernatant was transferred to a new sterile 1.5 ml microfuge tube for later analysis.

2.2.2.3 Protocol III. Chelex+Rnase extraction and purification method

This method of DNA extraction was similar to the Chelex method as described above. The only modification was after step "c" in Protocol II. In step "d", 2 µl of Rnase was added and incubated at $37\ ^{o}C$ for one hour. The sample was heated at $70\ ^{o}C$ for 15 minutes, centrifuged at 14,000 Xg for 5 min and the supernatant was transferred to a new sterile 1.5 ml microfuge tube for later analysis.

2.2.2.4 Protocol IV. XS DNA extraction and purification method

The DNA extraction and purification protocol used for this method was adopted from (Tillett and Neilan, 2000) with modifications. It consisted of the following steps:

a.   One ml of the *E. coli* cell culture at a cell density of OD_{600nm} 1 was centrifuged in a sterile 1.5 ml micro centrifuge tube for 5 min at 14,000 Xg. The cell pellet was recovered by discarding the spent medium.

b.   The pellet was then washed with 1 ml of double distilled water, vortexed and spun at 14,000 Xg for 5 min. This step was repeated one more time and the supernatant was discarded.

c.   To the cell sample obtained in step b, 800 µl of freshly prepared XS buffer (1% potassium ethyl xanthogenate, 100 mmol I^{-1} Tris-HCl (pH 7.4), 20 mmol I^{-1} EDTA (pH 8.0), 1% sodium dodecyl sulfate, and 800 mmol I^{-1} ammonium acetate) were added and mixed by inverting the tubes several times. The tube was then incubated at 70 $^{\circ}$C for 1 hour.

d.   After incubation, the tube was then inverted twice and then placed on ice for 30 minutes. This was followed by centrifugation at 14,000 Xg for 10 minutes. The supernatant was then transferred to a new sterile Eppendorf tube containing 750 µl of isopropanol and left overnight at -30 $^{\circ}$C for DNA to precipitate. The DNA precipitation was then centrifuged at 14,000 Xg for 10 minutes to obtain the DNA pellet.

e.   The DNA pellet was further washed twice with 1 ml of 70% ethanol, air-dried and finally resuspended in 100 µl of sterile double-distilled water.

2.2.2.5 DNA quantification

        After the completion of DNA extractions by the above four methods, DNA concentrations were determined using a NanoDrop uv-vis spectrophotometer (Thermo scientific, model 2000, Wilmington-DE-USA). Five µl of DNA sample extracted from

each strain were transferred onto the bottom pedestal of the NanoDrop machine and the cover of the NanoDrop was closed. The concentration of the DNA samples was measured in ng/µl unit. Absorbance of the dsDNA was measured at 260 nm and the 260 nm/280 nm ratio was used to assess the purity of DNA. Agarose gel electrophoresis was carried out to determine the quality of the DNA samples.

2.2.3 Agarose gel electrophoresis protocol

This method is used to separate and analyze biological macromolecules (such as nucleic acids and proteins) based on their size and charge. In this process, the migration of DNA fragments takes place from the "– ve" terminal of the electrophoresis chamber to the "+ ve" terminal. Larger DNA fragments travel slower than the smaller fragments. At the end of an electrophoretic run, a series of DNA bands is observed across the gel. The pattern of DNA is then analyzed using an UV-illuminator where DNA fragments fluoresce due to the presence of a DNA stain, such as ethidium bromide. Gel electrophoresis consists of following steps:

For this project, agarose gel electrophoresis was used to visualize the genomic DNA extracts of *E. coli* samples. It was also used to separate and visualize the amplified DNA fragment(s) obtained after the PCR assay. Initially, 0.5 g of agarose powder (Fisher Scientific, Ottawa, ON, Canada) was weighed and poured into a 125 ml Erlenmeyer flask. To this powder, 50 ml of 1X TAE (Tris-acetate-EDTA) buffer was added and the agarose suspension was heated in a dual wave microwave oven (General Electric, Fairfield, CT. USA) for 1.5 minutes until the agarose was completely dissolved. Five µl of ethidium bromide solution (10 g/ml) was added to the agarose

solution. The mixture was gently swirled and left to cool for 1 minute. The gel solution containing ethidium bromide was then poured into a gel tray (Bio-Rad). A gel comb was then placed in the gel to create an appropriate number of gel wells. The gel was then left for approximately 30 minutes to solidify. After 30 minutes, the gel tray was placed into a Bio-Rad gel electrophoresis chamber with wells placed at the negative end.
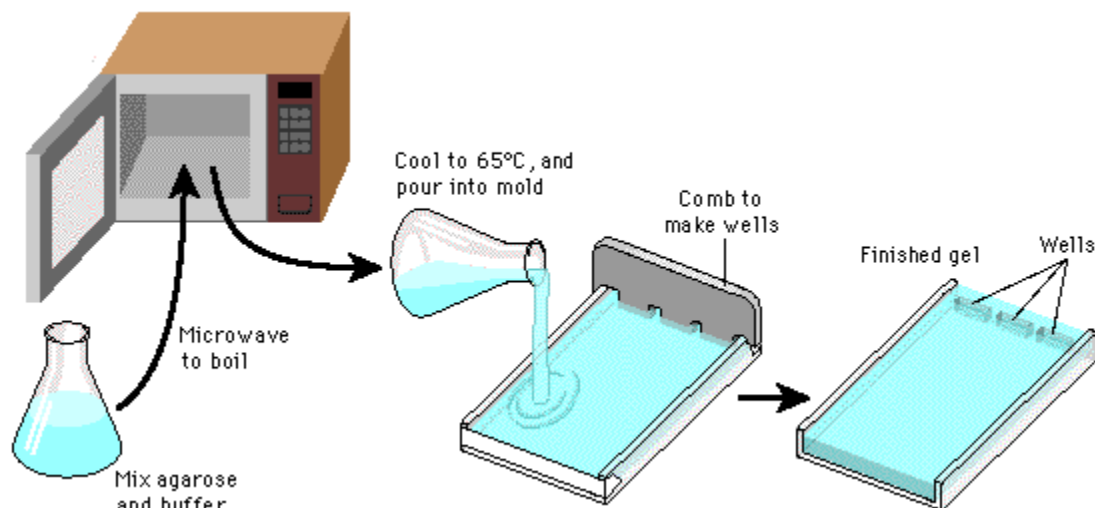
Figure 1.1 Showing preparation of agarose gel

To the electrophoresis chamber, 1X TAE buffer (40 mM Tris, 1 mM EDTA-Na$_2$, 20 mM acetic acid) was added in sufficient amount to submerge the whole gel completely. This was followed by the removal of the comb. Three µl of the GeneRuler 1Kb Plus DNA ladder (0.1 µg/ µl, Fermentas) was added into the first and last wells of the gel. Ten µl of the extracted DNA sample were mixed with 2 µl of 6X loading dye (Thermo Scientific, Waltham, MA, USA) on a parafilm strip and the mixture was loaded into the corresponding well of the gel. The DNA samples were allowed to run by turning on the electrophoresis system (100 Volt, 40 minutes). Once the DNA fragments were separated completely, the apparatus was turned off. The gel tray was then placed in a

Gel Doc XR+ System (Bio-Rad, Mississauga, ON, Canada) and the image of the DNA samples was obtained and analysed. The size of each DNA fragment of the sample was determined by comparing to the 1 kb plus DNA ladder (marker).

2.2.4 PCR analysis

To optimize the PCR detection assay to detect the virulence gene biomarkers of various pathotypes of *E. coli*, single-plex PCR analyses of the biomarkers were performed with specific pairs of PCR primers (Table 2.2). These biomarkers are: shiga-like toxin genes (*slt*) I and II, plasmid encoded invasion associated genetic loci (*ial*), plasma encoded enteroadherence factor gene (*eaf*), heat stable and heat labile enterotoxin genes (*hs* and *hl*), alpha-hemolysin gene (*hly*A), P-fimbrial adhesion genes *pap*A, iron acquisition gene (*iron$_{EC}$*). *E. coli* SY327 was used as negative control of the pathogenic *E. coli*. The purpose of this analysis was to determine the quality of the genomic DNA extracts, the specificity of the primers and the pathogenicity of the *E. coli* isolates examined in this study. The twelve primer pairs used in this study (Table 2.2) were synthesized by Sigma Life Science (Oakville, ON, Canada). The PCR reaction mixtures included the following ingredients: 5 μL of 2 mM dNTP mix, 5 μL of 25 mM MgCl$_2$, 5 μL of 10X *Taq* DNA polymerase buffer, 1 μL of *Taq* DNA polymerase (1U/μl), 0.2 μM of each of forward and reverse primers, 1 μL of genomic DNA (about 150 ηg DNA), and sterile UV-treated double distilled water to bring the final volume to 50 μL. The PCR was conducted in a BioRad MJ Mini Thermal Cycler MMO 10885 (Fisher Thermo Scientific). The protocol was performed in three stages: first, one cycle for initial denaturation at 95°C for five minutes; second, 34 cycles of 95°C for 1 minute for

denaturation, 55°C for 1 minute for annealing, and 72°C for 1 minute for extension; third, 72°C for 10 minutes and followed by 4°C as a holding temperature until the samples were removed from the thermocycler. The PCR products were visualized using 1% agarose gel electrophoresis as described above.

2.2.5 Multiplex PCR

In this study, five sets of multiplex PCR primers were developed for the detection of nine virulence genes of five pathogenic types of *E. coli* (STEC, ETEC, EPEC, EIEC and UPEC) (Table 2.3 and Table 2.4). Each set of multiplex primers consisted of three pairs of primers, with Set 1 targeting *hs* and *hl* genes of ETEC and *IroN$_{EC}$* of UPEC; Set 2 targeting *ial* of EIEC, *eaf* of EPEC and *hly* of UPEC; and Set 3 targeting shiga-like toxin I and II genes of STEC and *pap*A of UPEC (Table 2.3). Because Set 2 and 3 were not effective in amplifying all the target virulence genes, Set 4 and 5 were designed to replace them respectively (Table 2.4). The optimized multiplex-PCR primer sets (Set 1, 4 and 5) were used to determine the pathogenicity of the 306 *E. coli* isolates isolated from the Boulevard Lake water, periphyton, goose feaces and raw sewage samples. Three pairs of primers were grouped into a set of multiplex-PCR primers to allow the simultaneous detection of three different target virulence genes in a single reaction (Table 2.4). Individual primer pairs were grouped together based on the size of their amplicons. The sizes of amplicons within each set of multiplex-PCR reaction should be easily distinguished from each other and the three pair of primers must not interfere with each other. Each multiplex PCR reaction was 50 μl in volume. The reaction mixtures contained three primer pairs (*stx*I, *stx*II and *pap*A) targeting STEC and UPEC

respectively; (ial, *bfp*A and *hly*) targeting EIEC, EPEC and UPEC, and (*hs*, *hl* and *iron$_{EC}$*) target ETEC and UPEC. The optimal concentrations of primers and reagents used were determined. Each multiplex PCR reaction mixture contained 5 µL of 2 mM dNTP mix, 5 µL of 25 mM MgCl$_2$, 5 µL of 10X *Taq* DNA polymerase buffer, 1 µL *Taq* DNA polymerase (1 unit/µL), 0.2 µM each of forward and reverse primers, 1 µL of genomic DNA (about 150 ηg DNA), and sterile UV-treated sterile double distilled water to bring the final volume to 50 µL. The parameters and cycles of the multiplex PCR reactions were the same as the single-plex PCR reaction protocol described previously.

2.2.6 Statistical analysis

DNA extraction methods were conducted in three replications and the data were analyzed by SigmaPlot 12 statistic program (Systat Software Inc., San Jose, CA, USA) using a one-way analysis of variance (ANOVA). The Tukey's Multiple Range Test was used to compare the concentrations of DNA extracted by the four DNA extraction methods.

2.3 Results

2.3.1 DNA extraction methods and images

In this study four methods of DNA extraction were performed in order to determine an efficient DNA extraction method suitable for PCR detection of the virulence genetic markers of *E. coli*. The four methods included the Fermentas Genomic DNA Purification Kit, the XS Buffer method, the Chelex Extraction method and the Chelex Extraction + RNase method. The four methods were performed on each reference bacterial strain three times and the purified DNA samples were quantified using a NanoDrop

spectrophotometer to determine the amount of DNA extracted from each strain. Concentration of the extracted DNA was measured in ng/µl. Table 2.5 illustrates the performances of the four DNA extraction methods. There were significant differences ($p < 0.05$) between the four methods, but the differences among strains within individual methods were not significant ($p > 0.05$).The highest average amount of DNA extracted was obtained by the Chelex+RNase method. The effectiveness of the other three methods was ranged from the greatest to the least amount in the following order, the Chelex, the XS Buffer and the DNA purification kit (Fermentas) methods. The average DNA concentrations obtained by the four methods were 1501.5, 1407.2, 148.7 and 10.7 ηg/µl, respectively. The differences in DNA yield using the four methods can be clearly observed in the gel images in Figure 2.1. Although the use of the Chelex and the Chelex+RNase methods produced a large amount of DNA, electrophoresis analysis showed that the DNA extracts obtained by these methods were substantially sheared. Furthermore, the Chelex method did not remove cellular RNAs from the DNA extracts. On contrast, the XS Buffer method produced DNA extracts that formed compact and distinct DNA bands at about 20 kbp. No DNA bands were observed in the agarose gel electrophoresis assay from the DNA extracts obtained by the Fermentas DNA Extraction and Purification Kit.

2.3.2 Single-plex PCR assay to determine the quality of DNA extracts

To determine the quality of the DNA extracted from each method and its ability to be successfully amplified, a single-plex PCR was performed with DNA obtained from each method. Eight strains of *E. coli* representing five pathotypes (i.e. STEC, ETEC,

EPEC, EIEC, and UPEC) were used as positive controls and the *E. coli* sy327 strain was also used as a positive control for the 16S rDNA primers and negative control for all virulence primers. The virulence gene biomarkers and their PCR primers used in this analysis are shown in Table 2.2. The biomarkers are as follow: *sltI*, *sltII*, *hl*, *hs*, *ial*, *eaf*, *papA*, *papC*, *hly*, *iron$_{EC}$* and *kpsII*. The amplicon sizes in base pairs are: 350, 262, 322, 170, 390, 293, 720, 200, 1000, 665 and 336 respectively. Table 2.6 shows the results of the single-plex PCR assay. Gel images for the PCR products are displayed in Figure 2.2. It can be clearly seen that (i) not all the targeted sequences in all methods were successfully amplified and (ii) not all the primers in all methods were specific for the target gene. The XS Buffer was the only method that produced all positive PCR amplicons at the right sizes. For the other three methods, they produced negative reactions and/or non-specific amplifications for some strains. Figure 2.2 illustrates how the amplification of the target genes is different between the four methods. In the DNA kit method, only 7 primer pairs out of 12 showed amplification. The others showed either no amplification or non-specific reactions that did not correspond to the expected size range of the amplicons. Similarly, in the Chelex and the Chelex+RNase methods some non-specific reactions were detected. On the other hand, in the XS Buffer method all targeted strains tested positive for their specific biomarkers. The DNA fragments on the gel also looked distinct and the molecular sizes of the amplicons can easy be estimated by comparing to the 1kb plus DNA ladder. Because the DNA extracts of the XS Buffer method provided the most reliable outcomes of the PCR assay, it was selected to be the DNA extraction method in this study.

2.3.3 Specificity of the PCR primers

The specificity of the PCR primers was evaluated by testing the twelve primer pairs with the DNA extracts (extracted by the XS Buffer method) of eleven strains of *E. coli* and five non-*E. coli* bacterial species (Table 2.7). As shown in Table 2.7, all primers were specific for their own specific virulence gene biomarkers and did not amplify other *E. coli* virulence genes. In addition, the *E. coli* biomarker PCR primers did not amplify DNA extracted from any of the non-*E. coli* species. As expected, the 16S rDNA primer was positive for all bacterial species tested.

2.3.4 Multiplex PCR

2.3.4.1 Preliminary sets of multiplex PCR primers

A multiplex PCR assay was developed for the simultaneous detection of three specific virulence gene biomarkers (Table 2.3). Preliminary screening of the three primers sets showed that only Set 1 (*hs*, *hl*, *iron*$_{EC}$) could amplify the three specific genetic markers specifically (Table 2.8; Figure 2.3).

For the Set 2 primers, the *hly* genetic marker was successfully amplified in both the single-plex and multiplex PCR reactions (lane 4 and 8, Figure 2.4). Although strain EIEC0164 should be positive for the *ial* genetic marker (390 bp), the multiplex PCR showed only a weak and non-specific band at about 250 bp (Figure 2.4). Similarly, the multiplex PCR reactions on the *eaf* virulence factor positive control (EPEC055) were inconsistent. The expected 293 bp amplicon showed up only in about 50% of the replication trials (lane 6 and 7, Figure 2.4). When the DNAs of all the three positive

controls were used as the targets of the multiplex PCR reaction, only one band (at 293 bp, *eaf* amplicon) was observed in the gel (Lane 9, Figure 2.4).

E. coli O157:H7 strain STEC920004 contains both shiga-like toxin I and II genes (lane 2 and 3, Figure 2.4). However, only the *sltII* was amplified in the multiplex PCR reaction using the Set 3 multiplex primers (lane 5, Figure 2.5). Even when all the three target genes (shiga-like toxin I & II genes and *pap* A) were present in the Set 3 multiplex PCR reaction, only the shiga-like toxin II genetic marker were amplified (lane7, Figure 2.5).

2.3.4.2 Optimized sets of multiplex PCR primers

Since the MP-PCR of the Set 2 and Set 3 primers did not amplify some positive controls (Table 2.8), new sets of primers, Set 4 and 5, were designed to replace the ineffective Primer Sets 2 and 3, respectively (Table 2.4). In Set 4, the new *ial* and *bfpA* primers yielded a 650 bp and a 324 bp amplicons, respectively, in both the single-plex and multiplex PCR reactions (Table 2.9 and Figure 2.6). In addition, Set 4 primers successfully amplified the three targets (*ial, bfpA* and *hly*) either individually (lanes 5, 6 and 7, Figure 2.6) or as a mixture of all three targets (lane 8, Figure 2.6).

Set 5 was composed of *stx*I and *stx*II primers that amplified a 150 bp and a 255 bp DNA fragments of the shiga-like toxin I and II genes, respectively (Table 2.4). Although *slt*I and *slt*II were replaced by the new *stx*I and *stx*II primers, the *pap*A primers remained unchanged (Table 2.4). The *stx*I and *stx*II primers amplified the *stx*I and *stx*II genes specifically in single-plex PCR reactions, respectively (lanes 2 and 3, Figure 2.7). In addition, the new combination of the primer pairs in Set 5 showed success in

amplifying the three targets (*stx*I, *stx*II and *pap*A) either individually (lanes 5 and 6, Figure 2.7) or as a mixture of all three targets (lane 7, Figure 2.7).

2.3.5 Virulence genes of *E. coli* isolated from sewage, goose faeces, periphyton and lakewater

After the 3 sets of multiplex primers (Set 1, 4 and 5) were optimized, they were used to screen 306 *E. coli* isolated from sewage (37 isolates), goose faeces (38 isolates), periphyton (75 isolates) samples collected in 2010 and lakewater samples in 2010 and 2014 (76 and 80 isolates respectively). The results are summarized in Table 2.10. The percentages of isolates that contain one or more virulence factors in the sewage, goose faeces and periphyton samples were 48.5, 28.9 and 2.6%, respectively. For the isolates collected in 2010 and 2014, 5.3 and 5.0% of the isolates contained at least one virulence factor, respectively. For the sewage sample, 14 out of the 15 virulence factor-positive isolates belonged to the UPEC and 1 isolate contained the heat-labile enterotoxin gene of ETEC (Table 2.10). Out of the 14 UPEC positive isolates, 11 of them contained only one virulence factor (*pap*A or *iron$_{EC}$*) and 3 with two virulence factors (*pap*A and *iron$_{EC}$*) (Table 2.10 and 2.11). For the goose isolates, only UPEC were detected. Among the UPEC positive isolates, 10 of the 11 isolates had only one virulence factor (*iron$_{EC}$* or *hly*A) and only 1 isolate possessed two virulence factors (*iron$_{EC}$*, *pap*A). No diarrheagenic *E. coli* was detected in the periphyton samples and the 2 UPEC isolates (out of 75 isolates) detected contained either just the *iron$_{EC}$* virulence factor or all three virulence factors (*iron$_{EC}$*, *hly*A and *pap*A). The percentages of virulence factor-positive *E. coli* in the lakewater samples in 2010 and 2014 were similar

at 5.3 and 5.0%, respectively (Table 2.10). For the 2010 samples, all isolates contained only the $iron_{EC}$ virulence factor (Table 2.11). For the 2014 samples, there were 3 UPEC-positive and 1 EPEC-positive isolates. Two of the 3 UPEC-positive *E. coli* contained all three virulence factors and 1 isolate had only the $iron_{EC}$ virulence factor (Table 2.11).

2.4   Discussion

Recently, *E. coli* have been found in numerous tropical and temperate aquatic environments, soils, lakes and periphytons (Byappanahalli et al., 2006; Gordon and Cowling, 2003; Van Elsas et al., 2011). *E. coli* is able to attach to surfaces by producing filamentous structures from the cells and persist in biofilm or periphytic communities. In Boulevard Lake, Thunder Bay, Ontario, three major sources (including goose, human sewage and periphyton) contributed to the increased level of *E. coli* population in its water (Yang, 2012). Because the persistence of *E. coli* in recreational water represents a significant risk to human health (Cabral, 2010), it is important to differentiate the pathogenic strains from the commensal population (Croxen et al., 2013). The purpose of this study was to identify the virulence genes of the *E. coli* population in the periphytic community and the lakewater of Boulevard Lake, and to compare their virulence gene patterns to that of the goose feces and the human sewage samples.

Currently, there are numerous protocols available for the isolation of DNA from bacterial cells (Chen et al., 2010; Yuan et al., 2012; Carrigg et al., 2007). In this study, four different DNA extraction methods were tested. Some criteria used for choosing the optimal DNA extraction method in this study included the quality and quantity of the extracted DNA and the suitability of the DNA for PCR amplification (Cankar et al., 2006;

Ahmed et al., 2009). The four DNA extraction methods were: the XS Buffer method, the Fermentas Genomic DNA kit, the Chelex and the Chelex+RNase methods. Among these methods, XS Buffer was selected based on the quality and quantity of DNA obtained by this method. When the extracted DNA was quantified by the NanoDrop spectrophotometer 2000 (Table 2.5), it was found that the highest concentration of DNA was obtained by the Chelex+RNase method (about 150 μg/10^9 *E. coli* cells). On the contrary, the Fermentas DNA extraction kit gave the lowest concentration of DNA with about 1 μg/10^9 *E. coli* cells. In order to purify DNA from a crude bacterial DNA extract, it is important to precipitate and separate the DNA from other soluble cellular components (Maarit-Niemi et al., 2001). The most common method to purify DNA is ethanol precipitation. The lack of DNA precipitation step associated with both the Chelex and the Chelex+RNase protocols in this study may explain the high DNA yield of these two methods. However, the DNA obtained from these methods might contain high quantity of other cellular impurities such as proteins and polysaccharides. The presence of these impurities in the DNA extracts caused an overestimation of DNA concentrations and a reduction of PCR effectiveness. The impurities also resulted in smearing of the DNA fragments when visualized in agarose gel (Bertrand et al., 2005). The low yield of the Fermentas Genomic DNA Extraction Kit may be caused by the inadequate DNA precipitation step of the protocol. Generally, DNA precipitation by ethanol requires at least 1.5 hours (Köchl, Niederstaetter and Parson, 2005). However, only 5 minutes of precipitation was recommended by the Fermentas protocol.

The XS Buffer method yielded a moderate amount of DNA with about 15 μg/10^9 *E. coli* cells. Although the yield was lower than the two Chelex methods, it was

55

significantly higher than the Fermentas DNA Kit. The XS Buffer method also gave cleaner and sharper electrophoresis DNA bands than the Chelex methods. Although the DAN band quality of the XS Buffer method was similar to the DNA Kit method, all the DNA bands produced by the XS Buffer were clearly visible in the agarose gel, but only some of the DNA extracts obtained with the DNA Kit were visible. Therefore, the XS Buffer was the best method compared to the other three methods.

The quality of the DNA extracts produced by the four methods was evaluated by a singleplex PCR testing. Using samples of DNA template that were uncleaned and had inadequate amounts of DNA would result in low amplification efficiency (Glenn et al., 2005; Demeke and Jenkins., 2010). In agreement to these studies, the DNA template from the two Chelex methods led to non-specific amplification in the PCR analysis (Figure 2.2). In addition, the low concentration of the DNA extracts recovered by the Genomic Kit method caused negative amplification for some samples (Figure 2.2). On the other hand, the XS Buffer method showed clean and specific PCR bands for all the *E. coli* positive controls (Figure 2.2). The PCR testing further confirms our UV absorbance and agarose gel electrophoresis analyses that the XS Buffer method is the best DNA extraction method in comparison to the other three methods.

The singleplex PCR primers initially used to test the quality of the DNA extracts were shown to be specific to their respective target virulence genes (Table 2.2). However, when the singleplex primers were combined for a multiplex PCR platform (Set 1: *hl*, *hs*, $iron_{EC}$; Set 2: *ial*, *eaf*, *hly*; and Set 3: *slt*I, *slt*II, *pap*A. Table 2.3) to detect three virulence genes simultaneously, Set 2 and 3 primers were not able to amplify some of the target genetic factors (Table 2.8). *Slt*I and *slt*II in Set 3 could not be combined

together because the multiplex primer did not amplify the *slt*I virulence gene (Figure 2.5). Similarly, the *ial* and *eaf* primers in Set 2 produced non-specific amplification and/or negative results to their target virulence genes. Interactions between the primers in the multiplex PCR primer mixture may be the cause of failure for Set 2 and 3 primers. Yang et al (2006), Kubista et al (2007) and Pelt-Verkuil et al (2008) showed that binding between the primers negatively affected the binding of the primers to and the amplification of their specific target genetic markers in a PCR reaction. New multiplex primer sets, Set 4 and 5, were designed to replace the ineffective Set 2 and 3, respectively. In Set 4, a different fragment of the *ial* gene was targeted to produce a 650 bp amplicon. In addition, the *eaf* target in Set 2 was replaced by the *bfp*A virulence gene of the EPEC virotype. Consequently, the interactions between the primers of the EIEC, EPEC and UPEC were eliminated. For the Set 5 multiplex primers, *stx*1 and *stx*2 primers replaced the *slt*I and *slt*II primers to amplify the shiga-like toxin I and II, respectively. The *stx*1 and *stx*2 primers, which amplified the *stx*I and II genes to produce a 150 and 255 bp amplicon respectively, were adopted from Lopez-Saucedo et al (2003). The *stx*1 and *stx*2 primers did not interact with the *pap*A primers and successfully amplified the *stx*I, *stx*II, and *pap*A genes simultaneously. Therefore, Set 1, 4 and 5 multiplex primers were used in this study to determine the virulence genes of *E. coli* isolates isolated from the sewage, goose faeces, periphyton and lakewater samples.

Multiplex PCR has been used to detect virulent genes associated with pathogenic *E. coli* (Fatemeh et al., 2014; Kim, & Ihm, 2010; Mata et al., 2004; Molina et al., 2015). However, most of the studies are related to diarrheagenic *E. coli*. Fagan et al

(1998) described a Multiplex PCR method that was capable of detecting Shiga toxins 1 and 2 (*stx*1 and *stx*2), intimin (*eae*A), and enterohemolysin A (*hly*A) in 444 fecal samples derived from healthy and clinically affected cattle, sheep, pigs and goats. In this study, we designed a multiplex platform that consist of 3 sets of multiplex PCR primers that detected both the diarrheagenic and extraintestinal pathogenic *E. coli,* including the STEC, ETEC, EPEC, EIEC and UPEC.

The multiplex PCR platform was used to determine the pathogenicity of *E. coli* isolated from the lakewater (156 isolates) of Boulevard Lake (in 2010 and 2014) and the three major sources of *E. coli* in the lakewater including the periphyton (75 isolates), goose faeces (38 isolates) and raw sewage samples (37 isolates). Comparing the three sources, the percentage of potential *E. coli* pathogen (i.e. *E. coli* that contains one or more virulence gene(s)) was the highest in the sewage samples with 48.5%. Only 1 out of the 37 (2.7%) sewage *E. coli* was a potential diarrheagenic *E. coli* and the rest of the virulent isolates were potential UPEC. This agrees with Anastasi et al (2012) that most of the pathogenic sewage *E. coli* were UPEC. In this study, leakage of the municipal sewer collection system around the Boulevard Lake area could cause fecal contamination to the lake water.

While human sewage is a leading contributor of pathogenic *E. coli* in surface water systems, fecal waste from birds, especially waterfowls such as geese are also an important contributor of *E. coli* contamination (Ksoll et al., 2007). In the current study, 28.9% of the goose *E. coli* isolates were positive for at least one virulence gene and only UPEC isolates were identified. But other studies (Chandran and Mazumder, 2014;

Kullas et al., 2002) have shown that diarrheagenic *E. coli* strains such as ETEC, STEC, EPEC and EIEC could also be isolated from goose feces.

For the periphytic *E. coli*, only 2.6% of them (2 of 75 isolates) contained the UPEC virulence genes, one isolate with only the $Iron_{EC}$ gene and the other one with all three UPEC virulence genes. This observation suggests that the periphytic *E. coli* population was very different from those in the raw sewage and goose fecal samples. This conclusion is supported by the findings of Yang (2012) that the ERIC-PCR DNA fingerprinting profiles of the Boulevard Lake periphytic *E. coli* were closely clustered and were significantly different from the sewage and goose *E. coli* populations.

The lakewater *E. coli* isolates from 2010 (76 isolates) and 2014 (80 isolates) were tested by our multiplex PCR platform and the results were remarkably similar with 5.3 and 5.0 % of them carrying at least one virulence gene respectively. Based on the percentage of virulence positive *E. coli* in the lakewater samples, it is not likely that either the goose feces or the sewage were the major source of *E. coli* in the lake water. However, the percentages were much closer to the periphytic samples with 2.6% positive to the multiplex PCR analysis. Furthermore, isolates that contained three UPEC virulence genes were only detected in the periphyton and lakewater samples. This suggests that the periphytic population was the major contributor of *E. coli* in the water of the Boulevard Lake.

The distribution patterns of virulence genes in the lakewater, periphyton, sewage and goose samples were also compared (Figure 2.8). However, none of the four *E. coli* communities displayed the same or similar virulence gene distribution pattern. Furthermore, even the 2010 and 2014 lakewater *E. coli* samples showed different

distribution patterns of the virulence genes. This may be attributed to the low number and percentage of virulence positive *E. coli* in the periphyton and lakewater samples. One diarrheagenic virulence gene (*bfp*A) was detected in the lakewater samples. Since diarrheagenic genes were only detected in the fecal samples, but not in the periphyton samples, it is indicative that fecal contaminations might also contribute to the *E. coli* population in the lakewater. Yang (2012) showed that the periphytic *E. coli* community contributed to the majority of the *E. coli* population (ranged from 50-80% through the summer of 2010) in the water column of Boulevard Lake, however, fecal *E. coli* from sewage and goose samples were also detected. This further supports findings of this study that the percentage of virulence genes in the lakewater *E. coli* were similar to the periphytic *E. coli,* but fecal *E. coli* from geese and/or sewage may also contribute to the virulence positive *E. coli* in Boulevard Lake.

Conclusion and Future work

In conclusion, the XS buffer method was the cheapest and most effective method for extracting DNA from *E. coli*. In addition, three sets of multiplex PCR primers were designed and they were successful in amplifying nine specific virulence genes of the ETEC, EPEC, EIEC, STEC and UPEC. The results of the multiplex PCR assay showed that the percentage of *E. coli* isolates in the sewage (37 isolates), goose (38 isolates), periphyton (75 isolates) and lakewater (76 and 80 isolates collected in 2010 and 2014 respectively) samples that contained of one or more virulence gene(s) were 48.5, 28.9, 2.6, 5.3 and 5.0 % respectively. This indicates that the periphytons were likely to be the major source of *E. coli* in Boulevard Lake. Furthermore, with the exception of two

potential diarrheagenic *E. coli* isolates from the sewage and 2014 lakewater samples, all the virulence gene-positive isolates belonged to the uropathogenic *E. coli* pathotype.

UPEC virulence genes were the most predominant virulence genes in all the *E. coli* isolates tested. UPEC possesses a wide range of virulence factors, and there is no clear indication in the literature regarding the combination of virulence genes that will be most likely to determine the pathogenicity of the bacteria. Therefore, further research to determine the relationships between the combinations of virulence genes and the bacteria's pathogenicity is necessary. Furthermore, some virulence factors that help the survival in the environment such as the attachment proteins are important for the *E. coli* to establish and naturalize in aquatic environments. Therefore, it is important to determine and characterize these factors in order to understand the persistence and risks of *E. coli* in the environment.

Table 2.1 Strains of STEC, EPEC, EIEC, ETEC and UPEC used to monitor the specificity of PCR analysis.

Pathotype	Serotype and/or strain	Source
STEC (Shiga-toxins *E. coli*)	O157:H7 Strain 920004	Dr. C. Gyles, University of Guelph, Guelph, ON
	O157:H7 Strain 920026	
ETEC (enterotoxigenic *E. coli*)	Strain 07	
	Strain 505	
EPEC (enteropathogenic *E. coli*)	Strain 2348	Dr. B. Ciebin, Ministry of Health, Etobicoke, ON
	Strain 055	
EIEC (enteroinvasive *E. coli*)	Strain 0164	
	Strain 0136	
UPEC (uropathogenic *E. coli*)	ATCC 25922	ATCC (American Type Culture Collection)
Non-pathogenic *E. coli*	SY327	Dr. D. Cuppels, Agriculture and Agri-Food Canada, London, ON

Table 2.2 Primer sequences used for PCR amplification

Primer pair	Sequences	Target gene (Bacteria)	PCR amplicon (bp)	References
SltI-F SltI-R	5' ACCTCACTGACAGTCTGTGG 3' 5' TCTGCCGGACACATAGAAGGAAA 3'	Shiga like toxins (STEC)	350	(Rappelli et al., 2001)
SltII-F SltII-R	5' ACTGTCTGAAACTGCTCCTGTG 3' 5' TTATTTTTATAACGGGCCTGTTCGC 3'	Shiga like toxins (STEC)	262	
Hl-F Hl-R	5' TCTCTATGTGCACAC GGA GC 3' 5' CCA TAC TGA TTG CCG CAA T 3'	Heat labile enterotoxin (ETEC)	322	
Hs-F Hs-R	5' TCTTTC CCCTCTTTTAGTCAGTC 3' 5' CCAGCACAGGCAGGATTA C 3'	heat stable enterotoxin (ETEC)	170	
Ial -F Ial –R	5' TTTCTGGATGGTATGGTGAGG 3' 5' CACGCTGGTTGTCAATAATGCT 3'	Invasion associated loci (EIEC)	390	
Eaf-F Eaf-R	5' ACGCTTGGAGTGATCGAACG 3' 5' TGCCAACACAGCTTGTCAGAA 3'	plasma encoded enteroadherence factor (EPEC)	293	
PapA-F PapA-R	5' ATG GCA GTG GTG TCT TTT GGT G 3' 5' CGT CCC ACC ATA CGT GCT CTT C 3'	Pyelonephritis-associated pili (UPEC)	720	(Chapman et al.,2006)
PapC-F PapC-R	5'GTGGCAGTATGAGTAATGACCGTTA 3' 5' ATATCCTTTCTGCAGGGATGCAATA 3'	Genes of Pfimbriae operon (UPEC)	200	
HlyA-F HlyA-R	5' CATCTCTGGTTGGTGCACCGGTA 3' 5' AACTTGTCGGCACGCGTGGTC 3'	Hemolysin toxins (UPEC)	1000	
IronEC-F IronEC-R	5' AAGTCAAAGCAGGGGTTGCCCG 3' 5' GAC GCC GAC ATT AAG ACG CAG 3'	Siderophore receptor (UPEC)	665	
KpsII-F KpsII-R	5' GCGCATTTGCTGATACTGTTG 3' 5' CATCCAGACGATAAGCATGAGCA 3'	PapEF (UPEC)	336	

Table 2.3 Preliminary Multiplex-PCR primer sets

Primer set		sequences (5'-3') F/R[a]	Size (bp)
Set1	Hs	TCTTTCCCCTCTTTTAGTCAGTC/ CCA GCA CAG GCA GGA TTA C	170
	Hl	TCTCTATGTGCACACGGAGC/ CCA TAC TGA TTG CCG CAA T	322
	Iron$_{EC}$	AAGTCAAAGCAGGGGTTGCCCG/ GAC GCC GAC ATT AAG ACG CAG	665
Set2	Ial	TTTCTGGATGGTATGGTGAGG/ CACGCTGGTTGTCAATAATGCT	390
	Eaf	ACGCTTGGAGTGATCGAACG/ TGCCAACACAGCTTGTCAGAA	293
	Hly	CATCTCTGGTTGGTGCACCGGTA/ AACTTGTCGGCACGCGTGGTC	1000
Set3	Sltl	ACCTCACTGACAGTCTGTGG/ TCTGCCGGACACATAGAAGGAAA	350
	Sltll	ACTGTCTGAAACTGCTCCTGTG/ TTATTTTTATAACGGGCCTGTTCGC	262
	PapA	ATGGCAGTGGTGTCTTTTGGTG/ CGT CCC ACC ATA CGT GCT CTTC	720

a F = Forward primer, R = Reverse primer

Table 2.4 An optimized Multiplex-PCR primer sets

Primer set		sequences (5'-3') F/R	Size (bp)	Reference
Set 1	Hs	TCTTTCCCCTCTTTTAGTCAGTC/ CCAGCACAGGCAGGATTAC	170	(Rappelli et al., 2001)
	Hl	TCTCTATGTGCACACGGAGC/ CCATACTGATTGCCGCAAT	322	
	Iron$_{EC}$	AAGTCAAAGCAGGGGTTGCCCG/ GACGCCGACATTAAGACGCAG	665	(Chapman et al.,2006)
Set 4	Ial	GGTATGATGATGATGAGTCCA/ GGAGGCCAACAATTATTTCC	650	(López-Saucedo et al., 2003)
	BfpA	GCCGCTTTATCCAACCTGGTA/ TGCTGGACCTACATTTAATTCC	324	
	Hly	CATCTCTGGTTGGTGCACCGGTA/ AACTTGTCGGCACGCGTGGTC	1000	(Chapman et al.,2006)
Set 5	StxI	CTGGATTTA ATGTCGCATAGTG/ AGAACGCCCACTGAGATCATC	150	(López-Saucedo et al., 2003)
	StxII	GGCACTGTCTGA AACTGCTCC/ TCG CCAGTTATCTGA CATTCTG	255	
	PapA	ATGGCAGTGGTGTCTTTTGGTG/ CGTCCCACCATACGTGCTCTTC	720	(Chapman et al.,2006)

Table 2.5 Concentration of DNA extracted by the four DNA extraction methods[D,E]

Strain	DNA kit		XS buffer		Chelex		Chelex+Rnase	
	DNA (ng/µl)[a]	260/280 ratio	DNA (ng/µl)	260/280 ratio	DNA (ng/µl)	260/280 ratio	DNA (ng/µl)	260/280 ratio
	Average(Sd)[b]	Average(Sd)	Average(Sd)	Average(Sd)	Average(Sd)	Average(Sd)	Average(Sd)	Average(Sd)
STEC 920004	13.4 (±10.4)	1.9 (±0.3)	80.2 (±13.2)	2.4 (±1.4)	1323.5 (±343.6)	1.9 (±0.2)	1413.4 (±286.9)	1.8 (±0.1)
STEC 920026	9 (±9.6)	2.6 (±0.9)	113.6 (±87.4)	1.7 (±0.05)	1434.1 (±781.4)	1.8 (± 0.0329)	1610.4 (±663.5)	1.8 (±0.1)
ETEC 07	5.5 (±5)	2.5 (±1.0)	144.1 (±48.5)	1.8 (±0.2)	1543.9 (±628.6)	2.0 (±0.1)	1682.7 (±570.6)	1.9 (±0.1)
ETEC 505	11.2 (±1.6)	1.10 (±0.019)	380.3 (±347.3)	1.5 (±0.1)	1124.6 (±581.1)	1.8 (±0.1)	1140.4 (±825.3)	1.9 (±0.1)
EPEC 2348	5.2 (±3.1)	2.4 (±0.5)	108.6 (±60.6)	2 (±0.4)	1145.8 (±222.7)	2 (±0.2)	1507.1 (±73.1)	1.9 (±0.1)
EIEC 0164	6.1 (±2.7)	2.1 (±0.216)	115.7 (±37.1)	1.7 (±0.2)	858.1 (±72.4)	2 (±0.1)	961.1 (±234.4)	1.9 (±0.1)
UPEC 25922	22.9 (±28.3)	2.2 (±0.4)	155.2 (±80.7)	1.7 (±0.1)	2438.8 (±976.6)	1.10 (±0.03)	2283.4 (±1001.6)	1.9 (±0.03)
E. coli sy327	12.4 (±7.2)	2.11 (±0.13)	92.2 (±81.2)	1.6 (±0.3)	1389 (±475.2)	1.9 (±0.1)	1413.5 (±325.2)	2.0 (±0.03)
Average	10.7	2.1	148.7	1.8	1407.2	1.8	1501.5	1.9
P Value[c]	0.631	0.693	0.234	0.6	0.126	0.009	0.276	0.056

[a] sample extracted from one ml of OD 1 bacterial culture. 1 ml contains about 1.535×10^9 cells. DNA quantified using 5 µl out of 100 µl DNA sample.

[b] sd = standard deviation with (n=3)

[c] P values of the concentration of DNA extracts within individual methods are >0.05, indicating that there are no significant difference within methods.

[d] P value of the concentration of DNA extracts between the four methods is <0.001

[e] statistical analysis was done by anova one way

Table 2.6 PCR amplification of DNA extracted by DNA purification kit, XS buffer, Chelex and Chelex+Rnase methods

Strain	DNA kit	XS buffer	Chelex	Chelex+ Rnase
STEC 920004 (*Slt*I)	-	+	NSR[a]	NSR
STEC 920026 (*Slt*II)	+	+	+	+
ETEC 07 (*Hl*)	+	+	+	+
ETEC 505 (*Hs*)	+	+	+	+
EPEC 2348 (*Eaf*)	+	+	+	+
EIEC 0164 (*Ial*)	+	+	+	+
UPEC 25922 (*Pap*A)	-	+	-	-
UPEC 25922 (*Pap*C)	+	+	+	+
UPEC 25922 (*Hly*)	-	+	+	+
UPEC 25922 (*Iron*$_{EC)}$	+	+	NSR	+
UPEC 25922 (*Pap*EF)	-	+	+	NSR
E. coli sy327 (16S universal)	-	+	+	+

[a] NSR= non-specific amplification (size of the amplicon was different from the expected target DNA fragment).

Table 2.7 Determine the specificity of the singleplex PCR primers on DNA samples extracted by the XS buffer method

Strain	*Hs*	*Hl*	*Eaf*	*Slt*I	*Slt*II	*Ial*	*Pap*A	*Hly*	*Iron$_{EC}$*	*16S* universal
	170 bp	322 bp	293 bp	350 bp	262 bp	390 bp	720 bp	1000 bp	665 bp	193 bp
STEC 920004	-	-	-	+	+	-	-	-	-	+
STEC 920026	-	-	-	+	+	-	-	-	-	+
ETEC 07	-	+	-	-	-	-	-	-	-	+
ETEC 505	+	-	-	-	-	-	-	-	-	+
EPEC 055	-	-	+	-	-	-	-	-	-	+
EPEC 2348	-	-	+	-	-	-	-	-	-	+
EIEC 0164	-	-	-	-	-	+	-	-	-	+
EIEC 0136	-	-	-	-	-	+	-	-	-	+
UPEC 25922	-	-	-	-	-	-	+	+	+	+
E. coli sy327	-	-	-	-	-	-	-	-	-	+
E. coli 11775	-	-	-	-	-	-	-	-	-	+
Micrococcus luteas	-	-	-	-	-	-	-	-	-	+
Alcaligenes faecalis	-	-	-	-	-	-	-	-	-	+
Baciluus megaterium	-	-	-	-	-	-	-	-	-	+
Psudomonas fluorescence	-	-	-	-	-	-	-	-	-	+
Lactobacillus casei	-	-	-	-	-	-	-	-	-	+

Table 2.8 Multiplex PCR result of set1, 2 and 3 on Positive controls; Heat stable toxin (*hs*), 170 bp; heat labile toxin (*hl*), 322 bp; catechole siderophore receptor ($iron_{EC}$), 665 bp; Invasion associated loci (*ial*), 390 bp; plasma encoded entero adherence factor (*eaf*), 293 bp; Hemolysin (*hly*), 1000 bp; Shiga toxin 1(*slt*I), 350 bp; Shiga toxin 2 (*slt*II), 262 bp; Pyelonephritis-associated pili (*pap*A), 720 bp

Strain	Singleplex primers			Multiplex Primers
Set 1	*hs*	*hl*	$iron_{EC}$	*hs*, *hl* and $iron_{EC}$
ETEC07	-	+	-	+*hl*
ETEC505	+	-	-	+*hs*
UPEC25922	-	-	+	+*Iron$_{EC}$*
ETEC07, ETEC505, UPEC25922	+	+	+	+*hl*, +*hs*, + $iron_{EC}$
Set 2	*ial*	*eaf*	*hly*	*ial*, *eaf* and *hly*
EIECO136	+	-	-	+NSR[a]
EPECO55	-	+	-	-
UPEC25922	-	-	+	+*hly*
EIECO136, EPECO55 and UPEC25922	+	+	+	+*eaf*
Set 3	*slt*I	*slt*II	*pap*A	*slt*I, *slt*II and *pap*A
STEC920004	+	+	-	+*slt*II
UPEC25922	-	-	+	+*pap*A
STEC920004 and UPEC25922	+	+	+	+*slt*II

[a] NSR= non-specific reaction

Table 2.9 Optimized Multiplex-PCR result on Positive controls; Heat stable toxin (*hs*), 170 bp; heat labile toxin (*hl*), 322 bp; catechol siderophore receptor (*iroN$_{EC}$*), 665 bp; Invasion associated loci(*ial*), 650 bp; Bundle Forming Pili (*bfp*A), 324 bp; Hemolysin (*hly*), 1000 bp; Shiga toxin 1(*Stx 1*), 150 bp; Shiga toxin 2(Stx 2), 255 bp; Pyelonephritis-associated pili (*PapA*), 720 bp

Strain	Singleplex primers			Multiplex Primers
Set 1	*hs*	*hl*	*iron$_{EC}$*	*hs*, *hl* and *iron$_{EC}$*
ETEC07	-	+	-	+*hl*
ETEC505	+	-	-	+*hs*
UPEC25922	-	-	+	+ *iron$_{EC}$*
ETEC07, ETEC505, UPEC25922	+	+	+	+*hl*, +*hs*, + *iron$_{EC}$*
Set 4	*ial*	*bfp*A	*hly*	*ial*, *bfp*A and *hly*
EIECO136	+	-	-	+*Ial*
EPECO55	-	+	-	+*bfp*A
UPEC25922	-	-	+	+*hly*
EIECO136, EPECO55 and UPEC25922	+	+	+	+*ial*, +*bfp*A, +*hly*
Set 5	*stx*1	*stx*2	*pap*A	*stx*1, *stx*2 and *pap*A
STEC920004	+	+	-	+*stx*1, +*stx*2
UPEC25922	-	-	+	+*pap*A
STEC920004 and UPEC25922	+	+	+	+*stx*1, +*stx*2, +*pap*A

Table 2.10 Percentage of *E. coli* isolates tested positive with *E. coli* virulence genes

Source	% of isolates contain one virulent factor	% of isolates contain two virulent factor	% of isolates contain three virulent factor	% of isolates contain one or more virulent factor	% of isolates contain 0 virulent factor
Sewage(37 isolates)	32.4% (12strains)	8.1% (3 strains)	0	48.5% (15 strains)	51.5% (22 strains)
Goose (38 isolates)	26.3% (10 strains)	2.6% (1 strain)	0	28.9% (11 strains)	71% (27 strains)
Periphytic (75 isolates)	1.3% (1 strain)	0	1.3% (1 strain)	2.6% (2 strains)	97.3% (73 strains)
Planktonic 2010 (76 isolates)	5.3% (4 strains)	0	0	5.3% (4 strains)	94.7% (72 strains)
Planktonic 2014 (80 isolates)	2.5% (2 strains)	0	2.5% (2 strains)	5% (4 strains)	95% (76 strains)

Table 2.11 Distribution of virulence genes within the virulence positive *E. coli* population

| Source | Extraintestinal *E. coli* | | | Diarrheagenic *E. coli* | | | | | |
| | UPEC | | | STEC | | ETEC | | EPEC | EIEC |
	iron$_{EC}$	*hly*A	*pap*A	*stx* 1	*stx* 2	*hs*	*hl*	*bfp*A	*ial*
Sewage (37)	10.80%	0%	35%	0%	0%	0%	2.70%	0%	0%
Goose (38)	10.50%	18.40%	2.60%	0%	0%	0%	0%	0%	0%
Periphytic (75)	2.60%	1.30%	1.30%	0%	0%	0%	0%	0%	0%
Planktonic 2010 (76)	5.30%	0%	0%	0%	0%	0%	0%	0%	0%
Planktonic 2014 (80)	3%	2%	2%	0%	0%	0%	0%	1.25%	0%

Figure 2.1 Gel images of DNA samples extracted by the four DNA extraction methods. Lane 1&10, 1kb plus DNA ladder; lane2, STEC920004; Lane3, STEC920026; Lane4, ETEC07; Lane5, ETEC505; Lane6, EPEC2348; Lane7, EIEC0164; Lane8, UPEC25922; Lane9, *E. coli* SY327
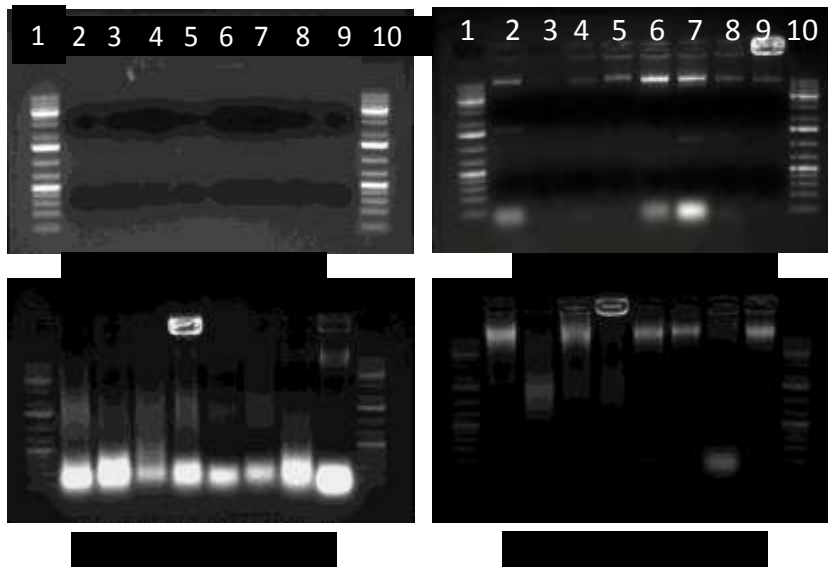
Figure 2.2 PCR images of DNA samples extracted by the four DNA extraction methods. Lane1 & 8, 1kb plus DNA ladder; (A) Lane2, STEC920004(*slt*I, 350bp); Lane3, STEC920026(*slt*II, 262bp); Lane4, ETEC07(*hl*, 322bp); Lane5, ETEC505(*hs*, 170bp); Lane6, EIECO164(*ial*, 390bp); Lane7, EPEC2348(*eaf*, 293bp); (B) UPEC, Lane2, (*pap*A, 720bp); Lane3, (*pap*EF, 336bp); Lane4, (*hly*, 1000bp); lane5, (*pap*C, 200bp); Lane6, (*iron*$_{EC}$, 665bp); Lane7, SY327(16s, 556bp)

A **B**

1   gDNA kit

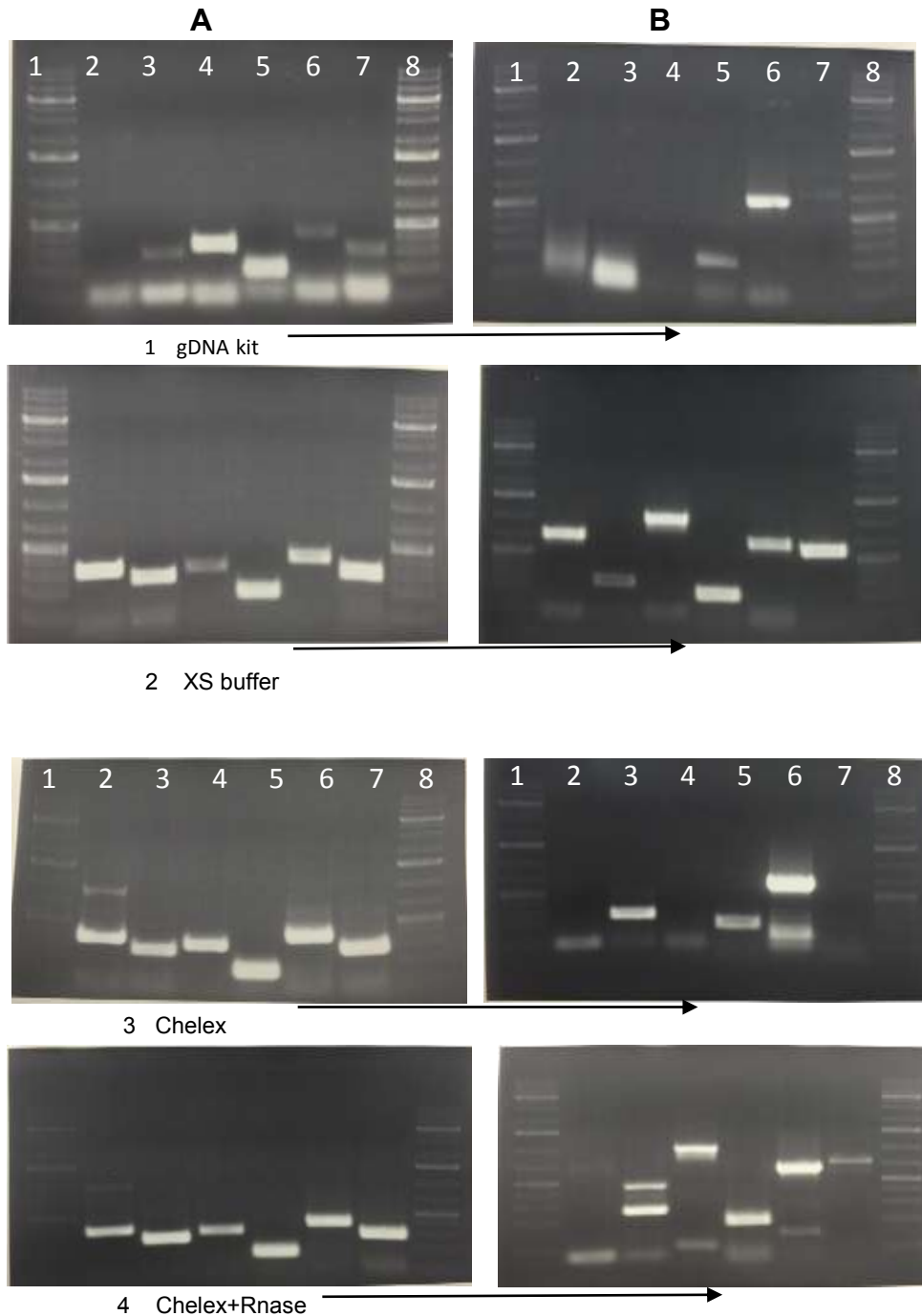2   XS buffer

3   Chelex

4   Chelex+Rnase

74

Figure 2.3 PCR products of singleplex and multiplex PCR reactions with set 1 primers on positive controls; Lane1 & 9, 1kb plus DNA ladder; Lane2 to 4, singleplex of ETEC505(*hs*, 170bp), ETEC07(*hl*, 322bp) and UPEC25922(*iron*$_{EC}$, 665bp); Lane5 to 8, MP-PCR of ETEC505(*hs*, 170bp), ETEC07(*hl*, 322bp), UPEC25922(*iron*$_{EC}$, 665bp) and MP-PCR reaction with all three target DNA (*hs*,170bp; *hl*, 322bp; *iron*$_{EC}$, 665bp)
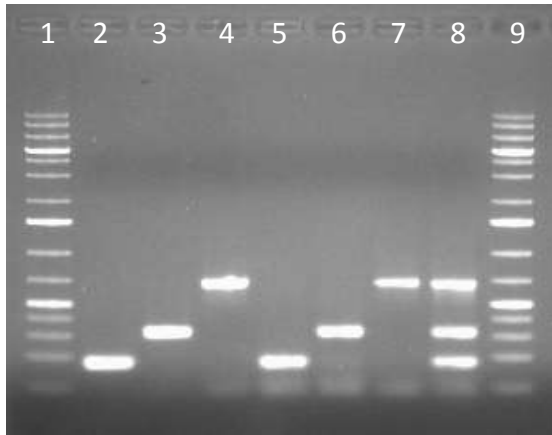
Figure 2.4 PCR products of singleplex and multiplex PCR reactions with set 2 primers on positive controls; Lane 1&10 1kb plus DNA ladder; Lane2 to 4, singleplex of EIEC0136(*ial*. 390bp), EPEC055(*eaf*. 293bp) and UPEC25922(*hly*. 1000bp); Lane5 to 9, MP-PCR of EIEC0136(*ial*, 390bp), Lane6&7, EPEC055(*eaf*, 293bp), UPEC 25922(*hly*, 1000bp) and MP-PCR reaction with all three target DNA (*ial*. 390bp, *eaf*. 293bp, *hly*. 1000bp)
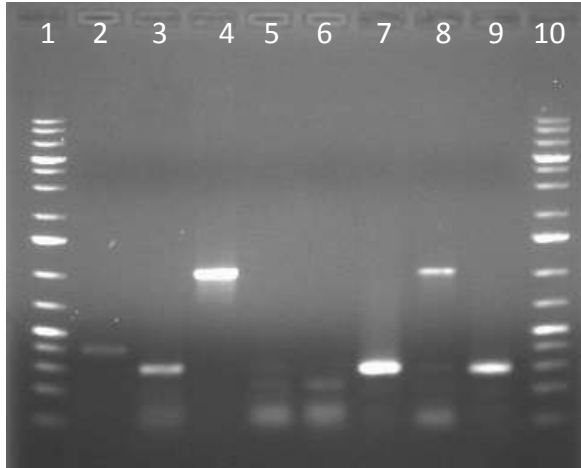
Figure 2.5 PCR products of singleplex and multiplex PCR reactions with set 3 primers on positive controls; Lane1&8, 1kb plus DNA ladder; Lane2 to 4, singleplex PCR of STEC920004(*slt*I3. 350bp), STEC920004(*slt*II4. 262bp) and UPEC25922(*pap*A. 720bp); Lane5 to 7, MP-PCR of STEC920004(*slt*I, 350bp; *slt*II, 262bp), UPEC25922 (*pap*A, 720bp) and MP-PCR reaction with all three target DNA (*slt*I, 350bp; *slt*II, 262bp; *pap*A, 720bp)

Figure 2.6 PCR products of singleplex and multiplex PCR reactions with set 4 primers on positive controls; Lane 1&9, 1kb plus DNA ladder; Lane2 to 4, singleplex PCR of EIEC0136(*ial*. 650bp), EPEC055(*bfp*A. 324bp) and UPEC25922(*hly*. 1000bp); Lane5 to 8, MP-PCR of EIEC0136(*ial*, 650bp), EPEC055(*bfp*A, 324bp), UPEC25922(*hly*, 1000bp) and MP-PCR reaction with all three target DNA (*ial*, 650bp; *bfp*A, 324bp; *hly*, 1000bp)
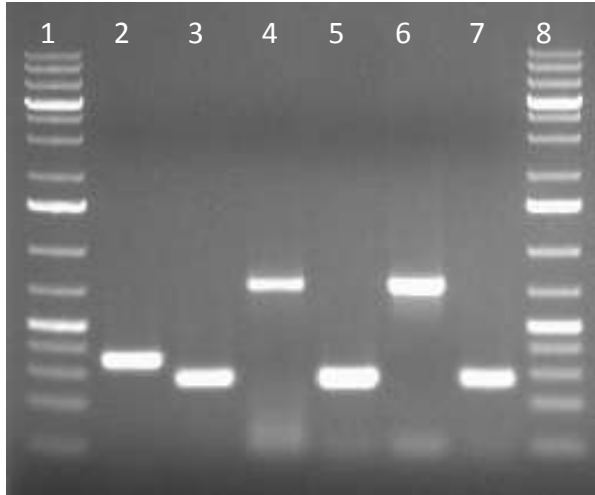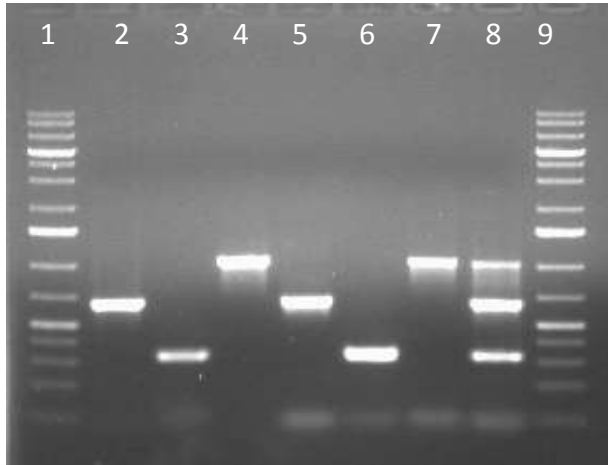
Figure 2.7 PCR products of singleplex and multiplex PCR reactions with set 5 primers on positive controls; Lane1&8, 1kb plus DNA ladder; Lane2 to 4 singleplex PCR of STEC920004(*stx*1. 150bp), STEC920004(*stx*2. 255bp) and UPEC25922(*pap*A. 720bp); Lane5 to 7 MP-PCR of STEC920004(*stx*1, 150bp; *stx* 2, 255bp), UPEC25922(*pap*A, 720bp) and MP-PCR reaction with all three target DNA (*stx*1, 150bp; stx2, 255bp; *pap*A, 720bp)

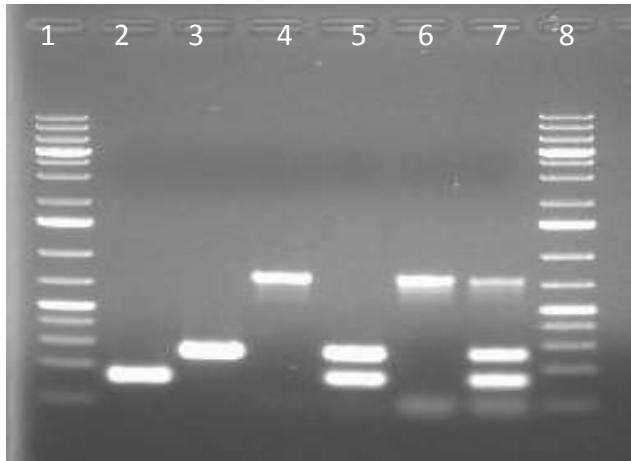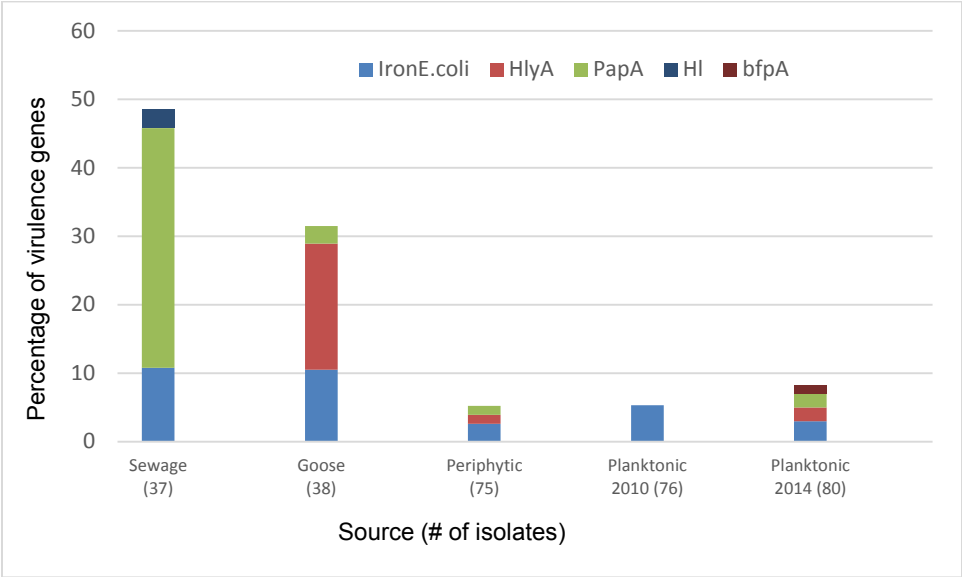Figure 2.8 Percentage distribution of virulence genes within the *E. coli* populations.

References

Ahmed, I., Islam, M., Arshad, W., Mannan, A., Ahmad, W. and Mirza, B. (2009). High-quality plant DNA extraction for PCR: an easy approach. *J Appl Genet*, 50(2), pp.105-107.

Albinana-Gimenez, N., Clemente-Casares, P., Calgua, B., Huguet, J., Courtois, S. and Girones, R. (2009). Comparison of methods for concentrating human adenoviruses, polyomavirus JC and noroviruses in source waters and drinking water using quantitative PCR. *Journal of Virological Methods*, 158(1-2), pp.104-109.

Anastasi, E., Matthews, B., Stratton, H. and Katouli, M. (2012). Pathogenic *Escherichia coli* Found in Sewage Treatment Plants and Environmental Waters. *Applied and Environmental Microbiology*, 78(16), pp.5536-5541.

Anderson, M., Whitlock, J. and Harwood, V. (2006). Diversity and Distribution of *Escherichia coli* Genotypes and Antibiotic Resistance Phenotypes in Feces of Humans, Cattle, and Horses. *Applied and Environmental Microbiology*, 72(11), pp.6914-6922.

Bando, S., Moreno, A., Albuquerque, J., Amhaz, J., Moreira-Filho, C. and Martinez, M. (2010). Expression of bacterial virulence factors and cytokines during in vitro macrophage infection by enteroinvasive *Escherichia coli* and Shigella flexneri: a comparative study. *Mem. Inst. Oswaldo Cruz*, 105(6), pp.786-791.

Bardiau, M., Szalo, M. and Mainil, J. (2010). Initial adherence of EPEC, EHEC and VTEC to host cells. *Vet. Res.*, 41(5), p.57.

Basu, A., Samanta, D., Bhattacharya, A., Das, A., Das, D. and DasGupta, C. (2008). Protein folding following synthesis in vitro and in vivo: Association of newly synthesized protein with 50S subunit of E. coli ribosome. *Biochemical and Biophysical Research Communications*, 366(2), pp.592-597.

Berthe, T., Ratajczak, M., Clermont, O., Denamur, E. and Petit, F. (2013). Evidence for Coexistence of Distinct *Escherichia coli* Populations in Various Aquatic Environments and Their Survival in Estuary Water. *Applied and Environmental Microbiology*, 79(15), pp.4684-4693.

Bertrand, H., Poly, F., Van, V., Lombard, N., Nalin, R., Vogel, T. and Simonet, P. (2005). High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction. *Journal of Microbiological Methods*, 62(1), pp.1-11.

Bielaszewska, M., Mellmann, A., Zhang, W., Kock, R., Fruth, A., Bauwens, A., Peters, G. and Karch, H. (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(9), pp.671-676.

Bien, J., Sokolova, O. and Bozko, P. (2012). Role of Uropathogenic *Escherichia coli* Virulence Factors in Development of Urinary Tract Infection and Kidney Damage. *International Journal of Nephrology*, 2012, pp.1-15.

Bilge, S., Clausen, C., Lau, W., & Moseley, S. (1989). Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *Journal of Bacteriology*, *171*(8), pp.4281–4289.

Blackburn, C. and McClure, P. (2002). *Foodborne pathogens*. Boca Raton, FL: CRC Press.

Blaustein, R., Pachepsky, Y., Hill, R., Shelton, D. and Whelan, G. (2013). *Escherichia coli* survival in waters: Temperature dependence. *Water Research*, 47(2), pp.569-578.

Botkin, D., Galli, L., Sankarapani, V., Soler, M., Rivas, M. and Torres, A. (2012). Development of a Multiplex PCR Assay for Detection of Shiga Toxin-Producing *Escherichia coli*, Enterohemorrhagic E. coli, and Enteropathogenic E. coli Strains. *Front. Cell. Inf. Microbio.*, 2.

Bower, P., Scopel, C., Jensen, E., Depas, M. and McLellan, S. (2005). Detection of Genetic Markers of Fecal Indicator Bacteria in Lake Michigan and Determination of Their Relationship to *Escherichia coli* Densities Using Standard Microbiological Methods. *Applied and Environmental Microbiology*, 71(12), pp.8305-8313.

ducing *Escherichia coli* Infections in the United States, 1983–2002. *The Journal of Infectious Diseases*, 192(8), pp.1422-1429.

Byappanahalli, M., Roll, B. and Fujioka, R. (2012). Evidence for Occurrence, Persistence, and Growth Potential of *Escherichia coli* and Enterococci in Hawaii^|^rsquo;s Soil Environments. *Microbes and Environments*, 27(2), pp.164-170.

Byappanahalli, M., Whitman, R., Shively, D., Sadowsky, M. and Ishii, S. (2006). Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. *Environ Microbiol*, 8(3), pp.504-513.

Cankar, k., Stebih, D., Dreo, T., Zel, J. and Gruden, K. (2006). Critical points of DNA quantification by real-time PCR - effects of DNA extraction method and sample matrix on quantification of genetically modified organisms. *BMC Biotechnology*, 6(37).

Carson, C., Shear, B., Ellersieck, M. and Asfaw, A. (2001). Identification of Fecal *Escherichia coli* from Humans and Animals by Ribotyping. *Applied and Environmental Microbiology*, 67(4), pp.1503-1507.

Chandran, A. and Mazumder, A. (2014). Occurrence of Diarrheagenic Virulence Genes and Genetic Diversity in *Escherichia coli* Isolates from Fecal Material of Various Avian Hosts in British Columbia, Canada. *Applied and Environmental Microbiology*, 80(6), pp.1933-1940.

Chapman, T., Wu, X., Barchia, I., Bettelheim, K., Driesen, S., Trott, D., Wilson, M. and Chin, J. (2006). Comparison of Virulence Gene Profiles of *Escherichia coli* Strains Isolated from Healthy and Diarrheic Swine. *Applied and Environmental Microbiology*, 72(7), pp.4782-4795.

Chen, H., Rangasamy, M., Tan, S., Wang, H. and Siegfried, B. (2010). Evaluation of Five Methods for Total DNA Extraction from Western Corn Rootworm Beetles. *PLoS ONE*, 5(8), p.e11963.

Chen, Y., Chen, X., Zheng, S., Yu, F., Kong, H., Yang, Q., Cui, D., Chen, N., Lou, B., Li, X., Tian, L., Yang, X., Xie, G., Dong, Y., Qin, Z., Han, D., Wang, Y., Zhang, W., Tang, Y. and Li, L. (2014). Serotypes, genotypes and antimicrobial resistance patterns of human diarrheagenic *Escherichia coli* isolates circulating in southeastern China. *Clinical Microbiology and Infection*, 20(1), pp.52-58.

Cravioto, A., Gross, R., Scotland, S. and Rowe, B. (1979). An adhesive factor found in strains of *Escherichia coli* belonging to the traditional infantile enteropathogenic serotypes. *Current Microbiology*, 3(2), pp.95-99.

Croxen, M., Law, R., Scholz, R., Keeney, K., Wlodarska, M. and Finlay, B. (2013). Recent Advances in Understanding Enteric Pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*, 26(4), pp.822-880.

de Lamballerie, X., Zandotti, C., Vignoli, C., Bollet, C. and de Micco, P. (1992). A one-step microbial DNA extraction method using "Chelex 100" suitable for gene amplification. *Research in Microbiology*, 143(8), pp.785-790.

Demeke, T. and Jenkins, G. (2009). Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Anal Bioanal Chem*, 396(6), pp.1977-1990.

Dini, C. and De Urraza, P. (2010). Isolation and selection of coli phages as potential biocontrol agents of enterohemorrhagic and Shiga toxin-producing *E. coli* (EHEC and STEC) in cattle. *Journal of Applied Microbiology*, 109(3), pp.873-887.

Donnenberg, M. (2013). *Escherichia coli*. Amsterdam: Elsevier/Academic Press.

Elnifro, E., Ashshi, A., Cooper, R. and Klapper, P. (2000). Multiplex PCR: Optimization and Application in Diagnostic Virology. *Clinical Microbiology Reviews*, 13(4), pp.559-570.

Espy, M., Uhl, J., Sloan, L., Buckwalter, S., Jones, M., Vetter, E., Yao, J., Wengenack, N., Rosenblatt, J., Cockerill, F. and Smith, T. (2006). Real-Time PCR in Clinical Microbiology: Applications for Routine Laboratory Testing. *Clinical Microbiology Reviews*, 19(3), pp.595-595.

Estrada-Garcia, T. and Navarro-Garcia, F. (2012). Enteroaggregative *Escherichia coli* pathotype: a genetically heterogeneous emerging foodborne enteropathogen. *FEMS Immunology & Medical Microbiology*, 66(3), pp.281-298.

Fagan, P., Hornitzky, M., Bettelheim, K., & Djordjevic, S. (1999). Detection of Shiga-Like Toxin (*stx1* and *stx2*), Intimin (*eaeA*), and Enterohemorrhagic *Escherichia coli* (EHEC) Hemolysin (EHEC *hlyA*) Genes in Animal Feces by Multiplex PCR. *Applied and Environmental Microbiology*, *65*(2), pp. 868–872.

Fatemeh, D., Reza, Z., Mohammad, A., Salomeh, K., Reza, A., Hossein, S., Maryam, S., Azam, A., Mana, S., Negin, N., Reza, K. and Saeed, F. (2014). Rapid detection of coliforms in drinking water of Arak city using multiplex PCR method in comparison with the standard method of culture (Most Probably Number). *Asian Pacific Journal of Tropical Biomedicine*, 4(5), pp.404-409.

Franzin, F. and Sircili, M. (2015). Locus of Enterocyte Effacement: A Pathogenicity Island Involved in the Virulence of Enteropathogenic and Enterohemorrhagic *Escherichia coli* Subjected to a Complex Network of Gene Regulation. *BioMed Research International*, 2015, pp.1-10.

Gamage, S., Patton, A., Strasser, J., Chalk, C. and Weiss, A. (2006). Commensal Bacteria Influence *Escherichia coli* O157:H7 Persistence and Shiga Toxin Production in the Mouse Intestine. *Infection and Immunity*, 74(3), pp.1977-1983.

Girones, R., Ferrus, M., Alonso, J., Rodriguez-Manzano, J., Calgua, B., de Abreu Correa, A., Hundesa, A., Carratala, A. and Bofill-Mas, S. (2010). Molecular detection of pathogens in water - The pros and cons of molecular techniques. *Water Research*, 44(15), pp.4325-4339.

Glassmeyer, S., Furlong, E., Kolpin, D., Cahill, J., Zaugg, S., Werner, S., Meyer, M. and Kryak, D. (2005). Transport of Chemical and Microbial Compounds from Known Wastewater Discharges: Potential for Use as Indicators of Human Fecal Contamination. *Environmental Science & Technology*, 39(14), pp.5157-5169.

Gonyar, L. and Kendall, M. (2013). Ethanolamine and Choline Promote Expression of Putative and Characterized Fimbriae in Enterohemorrhagic *Escherichia coli* O157:H7. *Infection and Immunity*, 82(1), pp.193-201.

Gordon, D. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology*, 149(12), pp.3575-3586.

Goto, D. and Yan, T. (2011). Genotypic Diversity of *Escherichia coli* in the Water and Soil of Tropical Watersheds in Hawaii. *Applied and Environmental Microbiology*, 77(12), pp.3988-3997.

Guevara, C., Luiz, W., Sierra, A., Cruz, C., Qadri, F., Kaushik, R., Ferreira, L. and Gomez-Duarte, O. (2013). Enterotoxigenic *Escherichia coli* CS21 pilus contributes to adhesion to intestinal cells and to pathogenesis under in vivo conditions. *Microbiology*, 159(Pt_8), pp.1725-1735.

Guion, C., Ochoa, T., Walker, C., Barletta, F. and Cleary, T. (2008). Detection of Diarrheagenic *Escherichia coli* by Use of Melting-Curve Analysis and Real-Time Multiplex PCR. *Journal of Clinical Microbiology*, 46(5), pp.1752-1757.

Hansen, A. and Kaper, J. (2009). Hfq affects the expression of the LEE pathogenicity island in enterohaemorrhagic *Escherichia coli*. *Molecular Microbiology*, 73(3), pp.446-465.

Harrington, S., Dudley, E. and Nataro, J. (2006). Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiology Letters*, 254(1), pp.12-18.

Hodges, K. and Gill, R. (2010). Infectious diarrhea. *Gut Microbes*, 1(1), pp.4-21.

Horakova, K., Mlejnkova, H. and Mlejnek, P. (2008). Specific detection of *Escherichia coli* isolated from water samples using polymerase chain reaction targeting four genes: cytochrome bd complex, lactose permease, β-d-glucuronidase, and β-d-galactosidase. *Journal of Applied Microbiology*, 105(4), pp.970-976.

Horáková, K., Mlejnková, H. and Mlejnek, P. (2008). Evaluation of methods for isolation of DNA for polymerase chain reaction (PCR)-based identification of pathogenic bacteria from pure cultures and water samples. *Water Science & Technology*, 58(5), p.995.

Humphries, R. and Armstrong, G. (2010). Sticky situation: localized adherence of enteropathogenic *Escherichia coli* to the small intestine epithelium. *Future Microbiology*, 5(11), pp.1645-1661.

Ishii, S. and Sadowsky, M. (2008). *Escherichia coli* in the Environment: Implications for Water Quality and Human Health. *Microbes and Environments*, 23(2), pp.101-108.

Ishii, S., Ksoll, W., Hicks, R. and Sadowsky, M. (2006). Presence and Growth of Naturalized *Escherichia coli* in Temperate Soils from Lake Superior Watersheds. *Applied and Environmental Microbiology*, 72(1), pp.612-621.

Isobe, K., Tarao, M., Chiem, N., Minh, L. and Takada, H. (2004). Effect of Environmental Factors on the Relationship between Concentrations of Coprostanol and Fecal Indicator Bacteria in Tropical (Mekong Delta) and Temperate (Tokyo) Freshwaters. *Applied and Environmental Microbiology*, 70(2), pp.814-821.

Jacobsen, S., Stickler, D., Mobley, H. and Shirtliff, M. (2008). Complicated Catheter-Associated Urinary Tract Infections Due to *Escherichia coli* and Proteus mirabilis. *Clinical Microbiology Reviews*, 21(1), pp.26-59.

Journal of Laboratory and Clinical Medicine, 139(3), pp.155-162.

kaper, J., Nataro, J. and Mobley, H. (2004). Pathogenic *Escherichia coli*. *Nat Rev Microbiol*, 2(123-140).

Kim, H., Kim, J. and Ihm, C. (2010). The usefulness of multiplex PCR for the identification of bacteria in joint infection. *Journal of Clinical Laboratory Analysis*, 24(3), pp.175-181.

Köchl S, Niederstaetter H, Parson W. (2005). DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. *Methods in Molecular Biology*, 297, pp.13–30.

Ksoll, W., Ishii, S., Sadowsky, M. and Hicks, R. (2007). Presence and Sources of Fecal Coliform Bacteria in Epilithic Periphyton Communities of Lake Superior. *Applied and Environmental Microbiology*, 73(12), pp.3771-3778.

Kullas, H., Coles, M., Rhyan, J. and Clark, L. (2002). Prevalence of *Escherichia coli* serogroups and human virulence factors in faeces of urban Canada geese (Branta canadensis). *International Journal of Environmental Health Research*, 12(2), pp.153-162.

Kulpakko, J., Kopra, K. and Hanninen, P. (2015). Time-resolved fluorescence-based assay for rapid detection of *Escherichia coli*. *Analytical Biochemistry*, 470, pp.1-6.

Lan, R., Alles, M., Donohoe, K., Martinez, M. and Reeves, P. (2004). Molecular Evolutionary Relationships of Enteroinvasive *Escherichia coli* and Shigella spp. *Infection and Immunity*, 72(9), pp.5080-5088.

Lane, M. and Mobley, H. (2007). Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney International*, 72(1), pp.19-25.

Large, T., Walk, S. and Whittam, T. (2005). Variation in Acid Resistance among Shiga Toxin-Producing Clones of Pathogenic *Escherichia coli*. *Applied and Environmental Microbiology*, 71(5), pp.2493-2500.

Lazcka, O., Campo, F. and Munoz, F. (2007). Pathogen detection: A perspective of traditional methods and biosensors. *Biosensors and Bioelectronics*, 22(7), pp.1205-1217.

Liu, J., Gratz, J., Maro, A., Kumburu, H., Kibiki, G., Taniuchi, M., Howlader, A., Sobuz, S., Haque, R., Talukder, K., Qureshi, S., Zaidi, A., Haverstick, D. and Houpt, E. (2012). Simultaneous Detection of Six Diarrhea-Causing Bacterial Pathogens with an In-House PCR-Luminex Assay. *J. Clin. Microbiol.*, 50(1), pp.98-103.

Liu, Y., Gilchrist, A., Zhang, J. and Li, X. (2008). Detection of Viable but Nonculturable *Escherichia coli* O157:H7 Bacteria in Drinking Water and River Water. *Applied and Environmental Microbiology*, 74(5), pp.1502-1507.

LIU, Y., PAN, C., XIA, W., ZENG, G., ZHOU, M., LIU, Y., KE, J. and HUANG, C. (2008). Simultaneous removal of Cr(VI) and phenol in consortium culture of Bacillus sp. and Pseudomonas putida Migula (CCTCC AB92019). *Transactions of Nonferrous Metals Society of China*, 18(4), pp.1014-1020.

Lloyd, A., Rasko, D. and Mobley, H. (2007). Defining Genomic Islands and Uropathogen-Specific Genes in Uropathogenic *Escherichia coli*. *Journal of Bacteriology*, 189(9), pp.3532-3546.

Los, F., Randis, T., Aroian, R. and Ratner, A. (2013). Role of Pore-Forming Toxins in Bacterial Infectious Diseases. *Microbiology and Molecular Biology Reviews*, 77(2), pp.173-207.

Maarit Niemi, R., Heiskanen, I., Wallenius, K. and Lindström, K. (2001). Extraction and purification of DNA in rhizosphere soil samples for PCR-DGGE analysis of bacterial consortia. *Journal of Microbiological Methods*, 45(3), pp.155-165.

Madigan, M. and Brock, T. (2012). *Brock biology of microorganisms*. San Francisco, Calif.: Pearson.

Maheux, A., Dion-Dupont, V., Bisson, M., Bouchard, S. and Rodriguez, M. (2014). Detection of *Escherichia coli* colonies on confluent plates of chromogenic media used in membrane filtration. *Journal of Microbiological Methods*, 97, pp.51-55.

Manges, A., Tabor, H., Tellis, P., Vincent, C. and Tellier, P. (2008). Endemic and Epidemic Lineages of *Escherichia coli* that Cause Urinary Tract Infections. *Emerg. Infect. Dis.*, 14(10), pp.1575-1583.

Mata, A., Gibello, A., Casamayor, A., Blanco, M., Dominguez, L. and Fernandez-Garayzabal, J. (2004). Multiplex PCR Assay for Detection of Bacterial Pathogens Associated with Warm-Water Streptococcosis in Fish. *Applied and Environmental Microbiology*, 70(5), pp.3183-3187.

Melican, K., Sandoval, R., Kader, A., Josefsson, L., Tanner, G., Molitoris, B. and Richter-Dahlfors, A. (2011). Uropathogenic *Escherichia coli* P and Type 1 Fimbriae Act in Synergy in a Living Host to Facilitate Renal Colonization Leading to Nephron Obstruction. *PLoS Pathog*, 7(2), p.e1001298.

Mendes Silva, D. and Domingues, L. (2015). On the track for an efficient detection of *Escherichia coli* in water: A review on PCR-based methods. *Ecotoxicology and Environmental Safety*, 113, pp.400-411.

Meyer, K., Appletoft, C., Schwemm, A., Uzoigwe, J. and Brown, E. (2005). Determining the source of fecal contamination in recreational waters. *Journal of environmental health*, 68(1), (pp. 25-30).

Molaee, N., Abtahi, H., Ghannadzadeh, M., Karimi, M. and Ghaznavi-Rad, E. (2015). Application of Reverse Transcriptase -PCR (RT-PCR) for rapid detection of viable *Escherichia coli* in drinking water samples. *Journal of Environmental Health Science and Engineering*, 13(1).

Molina, F., Lopez-Acedo, E., Tabla, R., Roa, I., Gomez, A. and Rebollo, J. (2015). Improved detection of *Escherichia coli* and coliform bacteria by multiplex PCR. *BMC Biotechnology*, 15(1), p.48.

Moreira, S., Brown, A., Ha, R., Iserhoff, K., Yim, M., Yang, J., Liao, B., Pszczolko, E., Qin, W. and Leung, K. (2011). Persistence of *Escherichia coli* in freshwater periphyton: biofilm-forming capacity as a selective advantage. *FEMS Microbiol Ecol*, 79(3), pp.608-618.

Muniesa, M., Jofre, J., Garcia-Aljaro, C. and Blanch, A. (2006). Occurrence of *Escherichia coli* O157:H7 and Other Enterohemorrhagic *Escherichia coli* in the Environment . *Environmental Science & Technology*, 40(23), pp.7141-7149.

Murray, P. (2005). *Medical microbiology*. 2nd ed. St. Louis: Mosby.

Navarro-García, F., Eslava, C., Villaseca, J. M., López-Revilla, R., Czeczulin, J. R., Srinivas, S., Cravioto, A. (1998). In Vitro Effects of a High-Molecular-Weight Heat-Labile Enterotoxin from Enteroaggregative *Escherichia coli*. *Infection and Immunity*, *66*(7), pp.3149–3154.

Norton, E., Branco, L. and Clements, J. (2015). Evaluating the A-Subunit of the Heat-Labile Toxin (LT) As an Immunogen and a Protective Antigen Against Enterotoxigenic *Escherichia coli* (ETEC). *PLOS ONE*, 10(8), p.e0136302.

Ochoa, T., Ecker, L., Barletta, F., Mispireta, M., Gil, A., Contreras, C., Molina, M., Amemiya, I., Verastegui, H., Hall, E., CI Periurban Areas in Lima, Peru. *Clinical Infectious Diseases*, 49(11), pp.1694-1702.

Okhuysen, P. and DuPont, H. (2010). Enteroaggregative *Escherichia coli* (EAEC): A Cause of Acute and Persistent Diarrhea of Worldwide Importance. *The Journal of Infectious Diseases*, 202(4), pp.503-505.

Osman, K., Elhariri, M., Mustafa, A. and AbdElhamed, G. (2012). Identification of serotypes and virulence markers of *Escherichia coli* isolated from human stool and urine samples in Egypt. *Indian Journal of Medical Microbiology*, 30(3), p.308.

Pelt-Verkuil, E., Belkum, A. and Hays, J. (2008). *Principles and technical aspects of PCR amplification*. [Dordrecht]: Springer.

Percival, S., Yates, M., Williams, D., Chalmers, R. and Gray, N. (2014). *Microbiology of waterborne diseases*.

Pizarro-Cerda, J. and Cossart, P. (2006). Bacterial Adhesion and Entry into Host Cells. *Cell*, 124(4), pp.715-727.

Priest, F. and Barker, M. (2009). Gram-negative bacteria associated with brewery yeasts: reclassification of Obesumbacterium proteus biogroup 2 as Shimwellia pseudoproteus gen. nov., sp. nov., and transfer of Escherichia blattae to Shimwellia blattae comb. nov. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, 60(4), pp.828-833.

Qadri, F., Svennerholm, A., Faruque, A. and Sack, R. (2005). Enterotoxigenic *Escherichia coli* in Developing Countries: Epidemiology, Microbiology, Clinical Features, Treatment, and Prevention. *Clinical Microbiology Reviews*, 18(3), pp.465-483.

Rajkhowa, S., Hussain, I. and Rajkhowa, C. (2009). Detection of heat-stable and heat-labile enterotoxin genes of *Escherichia coli* in diarrhoeic faecal samples of mithun (Bos frontalis) calves by polymerase chain reaction. *Journal of Applied Microbiology*, 106(2), pp.455-458.

Rappelli, P., Maddau, G., Colombo, M., Fiori, P. and Cappuccinelli, P. (2001). Development of a set of multiplex PCR assays for the simultaneous indentification of enterotoxigenic, enteropathogenic, enterohemorchagic and enteroinvasive *Escherichia coli*. *Microbiologica*, 24(77-83).

Regua-Mangia, A., Gomes, T., Vieira, M., Andrade, J., Irino, K. and Teixeira, L. (2004). Frequency and characteristics of diarrhoeagenic *Escherichia coli* strains isolated from children with and without diarrhoea in Rio de Janeiro, Brazil. *Journal of Infection*, 48(2), pp.161-167.

Reshes, G., Vanounou, S., Fishov, I. and Feingold, M. (2008). Cell Shape Dynamics in *Escherichia coli*. *Biophysical Journal*, 94(1), pp.251-264.

Robins-Browne, R. (2013). PATHOGENIC *ESCHERICHIA COLI*-AN UPDATE. *Pathology*, 45, p.S50.

Rubino, S., Cappuccinelli, P. and Kelvin, D. (2011). *Escherichia coli* (STEC) serotype O104 outbreak causing haemolytic syndrome (HUS) in Germany and France. *J Infect Dev Ctries*, 5(06).

Russo, T. (2003). Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*, 5(5), pp.449-456.

Sadowsky, M. and Whitman, R. (2011). *The fecal bacteria*. Washington, DC: ASM Press.

Salyers, A. and Whitt, D. (2002). *Bacterial pathogenesis*. 2nd ed. Washington, D.C.: ASM Press.

Schmidt, H. and Hensel, M. (2006). Pathogenicity Islands in Bacterial Pathogenesis. *Clinical Microbiology Reviews*, 19(1), pp.257-257.

Sen, K., L. Sinclair, J., Boczek, L. and Rice, E. (2011). Development of a Sensitive Detection Method for Stressed E. coli O157:H7 in Source and Finished Drinking Water by Culture-qPCR. *Environmental Science & Technology*, 45(6), pp.2250-2256.

Servin, A. (2005). Pathogenesis of Afa/Dr Diffusely Adhering *Escherichia coli*. *Clinical Microbiology Reviews*, 18(2), pp.264-292.

Sivick, K. and Mobley, H. (2009). Waging War against Uropathogenic *Escherichia coli*: Winning Back the Urinary Tract. *Infection and Immunity*, 78(2), pp.568-585.

Skurnik, D., Bonnet, D., Bernde-Bauduin, C., Michel, R., Guette, C., Becker, J., Balaire, C., Chau, F., Mohler, J., Jarlier, V., Boutin, J., Moreau, B., Guillemot, D., Denamur, E., Andremont, A. and Ruimy, R. (2008). Characteristics of human intestinal *Escherichia coli* with changing environments. *Environmental Microbiology*, 10(8), pp.2132-2137.

Smith, C. and Osborn, A. (2009). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiology Ecology*, 67(1), pp.6-20.

Smith, J., P. Fratamico, and N. Gunther. (2007). Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathogen Dis*. 4(2), pp.134-163.

Tan, C., Tang, X., Zhang, X., Ding, Y., Zhao, Z., Wu, B., Cai, X., Liu, Z., He, Q. and Chen, H. (2012). Serotypes and virulence genes of extraintestinal pathogenic *Escherichia coli* isolates from diseased pigs in China. *The Veterinary Journal*, 192(3), pp.483-488.

Tobias, J. and Vutukuru, S. (2012). Simple and rapid multiplex PCR for identification of the main human diarrheagenic *Escherichia coli*. *Microbiological Research*, 167(9), pp.564-570.

Torres, A. (2010). *Pathogenic Escherichia coli in Latin America*. [Place of publication not identified]: Bentham e Books.

Torres, A., Zhou, X. and Kaper, J. (2004). Adherence of Diarrheagenic *Escherichia coli* Strains to Epithelial Cells. *Infection and Immunity*, 73(1), pp.18-29.

Ud-Din, A. and Wahid, S. (2014). Relationship among Shigella spp. and enteroinvasive *Escherichia coli* (EIEC) and their differentiation. *Brazilian Journal of Microbiology*, 45(4), pp.1131-1138.

van Elsas, J., Semenov, A., Costa, R. and Trevors, J. (2011). Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *The ISME Journal*, 5(2), pp.367-367.

Vogeleer, P., Tremblay, Y., Mafu, A., Jacques, M. and Harel, J. (2014). Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing *Escherichia coli*. *Front. Microbiol.*, 5.

Wade, T., Pai, N., Eisenberg, J. and Colford, J. (2003). Do U.S. Environmental Protection Agency Water Quality Guidelines for Recreational Waters Prevent Gastrointestinal Illness? A Systematic Review and Meta-analysis. *Environ Health Perspect*, 111(8), pp.1102-1109.

Wang, F., Yang, Q., Kase, J., Meng, J., Clotilde, L., Lin, A. and Ge, B. (2013). Current Trends in Detecting Non-O157 Shiga Toxin-Producing *Escherichia coli* in Food. *Foodborne Pathogens and Disease*, 10(8), pp.665-677.

Wang, X., Wang, J., Sun, H., Xia, S., Duan, R., Liang, J., Xiao, Y., Qiu, H., Shan, G. and Jing, H. (2015). Etiology of Childhood Infectious Diarrhea in a Developed Region of China: Compared to Childhood Diarrhea in a Developing Region and Adult Diarrhea in a Developed Region. *PLOS ONE*, 10(11), p.e0142136.

Wani, S., Hussain, I., Fayaz, I., Mir, M. and Nishikawa, Y. (2009). Subtype analysis of stx1, stx2 and eae genes in Shiga toxin-producing *Escherichia coli* (STEC) and typical and atypical enteropathogenic E. coli (EPEC) from lambs in India. *The Veterinary Journal*, 182(3), pp.489-490.

Weintraub, A. (2007). Enteroaggregative *Escherichia coli*: epidemiology, virulence and detection. *Journal of Medical Microbiology*, 56(1), pp.4-8.

Wheeler Alm, E., Burke, J. and Spain, A. (2003). Fecal indicator bacteria are abundant in wet sand at freshwater beaches. *Water Research*, 37(16), pp.3978-3982.

Wiedermann, U. and Kollaritsch, H. (2006). Vaccines against traveler's diarrhoea and rotavirus disease - a review. *Wiener klinische Wochenschrift*, 118(S3), pp.2-8.

Wiles, T., Kulesus, R. and Mulvey, M. (2008). Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Experimental and Molecular Pathology*, 85(1), pp.11-19.

Yang, J., 2013. Source tracking of *Escherichia coli* in a freshwater lake in Northwestern Ontario (Boulevard Lake, Thunder Bay). Thunder Bay, Ontario: Lakehead University.

Yang, M., Tsoi, P., Li, C. and Zhao, J. (2006). Analysis of interactions of template/primer duplexes with T7 DNA polymerase by oligonucleotide microarray. *Sensors and Actuators B: Chemical*, 115(1), pp.428-433.

Yuan, S., Cohen, D., Ravel, J., Abdo, Z. and Forney, L. (2012). Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLoS ONE*, 7(3), p.e33865.