Source Tracking of *Escherichia coli* in a Freshwater Lake in Northwestern Ontario (Boulevard Lake, Thunder Bay)

A Thesis

Presented to

The Faculty of Graduate Studies

Of

Lakehead University

by

Jeremy Yang

In partial fulfillment of requirements

For the degree of

Master of Science in Environmental Engineering

Thunder Bay, Ontario, Canada

Abstract of the Thesis

Escherichia coli is regularly used as a fecal indicator bacteria (FIB) in recreational waters but its persistence in the environment brings its use into doubt. A secondary FIB known as *Bacteroides* has been under a lot of research lately. Because *Bacteroides* can be measured with quantitative PCR (qPCR) techniques easily, this makes it an effective FIB to detect fecal contamination. Boulevard Lake in Thunder Bay, Ontario experiences instances of elevated levels of *E. coli* throughout the summer seasons. For the 2011 season the *Bacteroides* 16S rDNA markers were monitored and compared with the *E. coli* population. As both of these FIB are found in feces, influxes of fecal contamination would see increases in both populations. The planktonic *E. coli* population densities exceeded the Canadian Recreational Water Quality Guidelines of 2.30 log CFU $100ml^{-1}$ two times throughout the summer season. These were measured at 2.86 and 2.38 log CFU 100 ml^{-1} on July 21, 2011 and September 2, 2011, respectively. The *Bacteroides* biomarkers did have any significant increases during these peak periods of *E. coli* with $p > 0.05$. This would suggest that the increased levels of *E. coli* may not have been due to fecal contaminants. Further investigations with a microbial source tracking approach will provide insights to the potential source(s) of *E. coli* in Boulevard Lake. For both 2010 and 2011, the planktonic *E. coli* population at Boulevard Lake was monitored. A microbial source tracking (MST) library was created using REP-PCR (Repetitive Extragenic Palindromic-Polymerase Chain Reaction) to DNA-fingerprint *E. coli* isolated from farm animals (horse, sheep, pig, and chickens), human sewage, geese, and periphyton. The farm animals were raised on a small farm located approximately 3.5 km upstream of and about 200 m from the river running into Boulevard Lake. Any fecal contaminants from the farm would have to be due to

runoff from precipitation, but between 15.0 to 50.0 % of the planktonic *E. coli* was found to be similar to the *E. coli* isolates of the farm animals when no precipitation was observed. The farm animals were therefore not likely to be the sources of *E. coli* blooms in the lake and were taken out of the MST library. When the planktonic *E. coli* was compared to the periphyton, goose, and human sources, the majority of the isolates belonged to the periphytic communities during the *E. coli* peaks at 57.6, 65.8, and 38.7 % on June 28, 2010, August 25, 2010, and September 2, 2011, respectively. During the non-peak periods, the majority of the planktonic *E. coli* was also found to belong to the periphyton (35.0 – 80.0 %). The geese were seen to be the predominant source of *E. coli* in the water (50%) on August 9, 2011 but the total planktonic *E. coli* densities were only log 1.59 CFU 100 ml^{-1}, which did not exceed the maximum limits recommended by the Guidelines for Canadian Recreational Water Quality. Overall, the increased levels of *E. coli* seen in Boulevard Lake were not coming primarily from fecal contamination but were attributed to the periphytic *E. coli* community.

Table of Contents

List of Tables

List of Figures

Chapter 1 – Literature Review

1.0 Introduction

Fecal contaminations of water have always been a large area of concern. Locating the origin of the source of the pollutants is an arduous and strenuous task. Contamination due to fecal matters is a threat to human health and is a problem world-wide (Tallon *et al.* 2005). Inputs of fecal matter into the environment can also contribute to problems such as eutrophication in lakes (Paerl *et al.* 2003) as well as an influx of fecal pathogens into the waters. These include pathogenic bacteria *(Shigella, Salmonella*, and *Campylobacter*), viruses (norovirus, hepatitis A), and protozoa (*Cryptosporidium* and *Giardia*) (Ishii and Sadowsly 2008). Areas affected include the human health, the environment and the economy if beaches and popular spots along the water are closed. In 2000, a well in Walkerton, Ontario was contaminated with cattle manure from a nearby farm and was not properly treated. This resulted in more than 2,300 people experiencing gastroenteritis, 65 were hospitalized, 27 developed haemolytic uraemic syndrome, and 7 deaths occurred (Hrudey *et al.* 2003). From 2007 to 2008, there were 134 reported bacterial waterborne outbreaks in recreational waters in 38 states in the U.S and Puerto Rico. During those outbreaks, 13,966 bacterial infection cases resulted. The National Resources Defense Council conducted an epidemiology study and found that approximately 10% of beachgoers report getting sick after swimming in the Great Lakes. 62 percent of the beach closings/advisories in the U.S. were due to unknown sources of contamination in 2008. Also in 2008, Wisconsin had 578 closing or advisory notices for beaches, California had 1,003, and Illinois had 396. Almost all of these cases were due to high levels of bacterial indicators (Dorfman 2009).

A case study along the Lake Michigan freshwater beach shows losses from $1274-37,030 for each day of closure (Rabinovici *et al.* 2004). Approximately 85% of all U.S. tourism comes from coastal states (Dorfman 2009). There are a large number of beaches being closed every year and enormous revenues are lost. Are the bacterial indicators always indicative of fecal contamination? Storm water and sewage can be attributed to some of these numbers but most of the time the cause is unknown (Parker *et al.* 2010). From 1971-2000 three-quarters of the outbreaks in recreational water were unidentified (Craun *et al.* 2005). This thesis will look into discovering the source of these unknown bacterial contaminations. A fast and reliable bacterial identification method will ensure recreational waters are closed only when the contaminating bacteria are harmful to the human population.

1.1 Detection of Fecal Contamination

Fecal contamination can be detected by chemical or microbial approaches. Chemical compounds that are typically only used by humans can be potential candidates to detect human fecal contamination. Examples include carbamazepine, coprostanol, diphenhydramine, and caffeine (Isobe *et al.* 2004; Glassmeyer *et al.* 2005). In California, fecal steroids were used to source the fecal indicator bactieria. This was done by using fecal steroid ratios in the sewage and comparing them to an area with sources of fecal indicator bacteria. Differences were found in the ratios, implying that the sewage was not the source of fecal steroids and therefore may not have been the source of fecal indicator bacteria (Noblet et al 2004). Some techniques used to detect chemical compounds include mass spectrometry (MS), gas chromatography (GC), liquid

chromatography (LC), and high performance liquid chromatography (HPLC) (Piocos and de la Cruz 2000; Isobe *et al.* 2004; Noblet *et al.* 2004; Glassmeyer *et al.* 2005).

Indicator organisms are utilized as the primary method in measuring fecal contamination in the water. These are used because it is not feasible to test every pathogen, such as *Salmonella, Shigella, Yersinia, S. aureus,* and, *C. botulinum,* as most are difficult and time-consuming to detect and culture (Sadowsky and Whitman 2011). A group of organisms known as the coliform bacteria have been popularly used to indicate fecal contamination of water. This is due to the fact that they inhabit the intestinal tract in high numbers. They also generally live longer than disease-causing bacteria so an absence of coliform bacteria can indicate that the water is safe. The coliforms are defined as facultatively anaerobic, gram negative, non-spore-forming, rod-shaped bacteria that produce gas upon lactose fermentation within 48 hours at 35°C (Madigan *et al.* 2012). The coliform group includes organisms of fecal and non-fecal origin; therefore a more restrictive definition is needed to refine the group to fecal origin. This group is known as the fecal coliforms, and they are different from the total coliforms in that they ferment lactose and produce acid and gas at 44.5°C within 24 hours (Maier *et al.* 2009). The total coliforms are a broader range of bacteria that can be found in nature and are usually used to test for drinking water to ensure safety. The fecal coliforms are more fecal specific in origin and are used abundantly in ensuring the safety of recreational waters (U.S. Environmental Protection Agency). Bacteria primarily found in feces and fall under the term 'fecal coliform' include *E. coli, Klebsiella, Enterobacter and Citrobacter*, though some may still originate outside feces (Tallon *et al*. 2005; Sadowsky and Whitman 2011).

Criteria for an ideal indicator organism are as follows: (Maier *et al.* 2009)

- The organism should be useful for all types of water
- The organism should be present whenever enteric pathogens are present
- The organism should have a reasonably longer survival time than the hardiest enteric pathogen
- The organism should not grow in water (environment)
- The testing method should be easy to perform
- The density of the indicator organism should have some direct relationship to the degree of fecal pollution
- The organism should be a member of the intestinal microflora of warm-blooded animals

There is not a single indicator organism which would be sufficient in fulfilling all these criteria; therefore using more than one indicator would be ideal (Tyagi *et al.* 2006). There are five major fecal indicators that are used often. These fecal indicator organisms are *Bacteroides, Bifidobacterium, Clostridium perfringens, Escherichia coli,* and *Enterococcus* (Sadowsky and Whitman 2011). *Bacteroides, Bifidobacterium, and Clostridium perfringens* are obligate anaerobes which are difficult to culture. *Clostridium perfringens* have varying levels among different animals and therefore may inaccurately indicate fecal contaminants. *Bifidobacterium* was found to have low cell densities in fecal materials. Therefore, it may not always correlate with the fecal pollution and health risks (Lamendella *et al.* 2008). Other than being difficult to culture, *Bacteroides* is a good fecal indicator bacterium because it is found in high amounts and amongst different host species in feces (Sadowsky and Whitman 2011). *E. coli* and

Enterococcus are facultative anaerobes which make them easy to culture. When looking for human contamination, *Enterococcus* is a good indicator because its host range is limited to mostly humans. Chicken and Geese have been found to have low levels of *Enterococcus* and it can also be found abundantly in the environment (Wheeler *et al.* 2002; Alm *et al.* 2003; Sadowsky and Whitman 2011). The U.S. Environmental Protection Agency reviewed many cases of gastro-intestinal illnesses and found that *E. coli* was a far more reliable fecal indicator in freshwater than *Enterococcus* (Wade *et al.* 2003).

1.2 *Escherichia coli*

Escherichia coli is a gram negative, rod shaped, facultative anaerobic bacterium that is usually found in the gastrointestinal tract. It can be classified into 3 groups which include commensal, diarrheagenic, as well as extraintestinal. The commensal *E. coli* is the most common type which normally lives in the gastrointestinal tract of warm blooded animals (Sadowsky and Whitman 2011). Most strains of *E. coli,* like the commensal groups are harmless but there are some virulent types. The diarrheagenic strains can cause diseases such as diarrhea, hemorrhagic colitis, hemolytic uremic syndrome, inflammatory colitis, and dysentery. The extraintestinal strains can cause urinary tract infections, septicemia, and neonatal meningitis (Clermont *et al.* 2000; Sadowsky and Whitman 2011).

E. coli is used as a fecal indicator because methods to test for it are relatively inexpensive, simplistic and they fulfill many of the criteria previously stated. *E. coli* can live from 4 – 12 weeks in water, depending on environmental conditions (Edberg *et al.* 2000)**.** Tropical and temperate soils were shown to house *E. coli* as well as secondary non-host habitats,

which include sand, algae, and periphyton (Byappanahalli and Fujioka 1998; Whitman *et al.* 2003; Ishii *et al.* 2006; Ksoll *et al.* 2007). *E. coli* can be shown to live through various stresses and it is well documented that it can survive through temperatures well below freezing (Ksoll *et al.* 2007). The ability for *E. coli* to survive through all different types of environmental stress could be due to the fact it has a high genetic diversity, as a higher genetic diversity tends to increase adaptability as well as resistance (Goto and Yan 2011).

As *E. coli* is used as a fecal indicator, the maximum limit of *E. coli* in recreation water is approximately 200 CFU / 100ml as set by the Canadian Recreational Water Quality Guidelines. When levels are higher than these thresholds, beach areas are usually closed. *E. coli* can be found in densities of $>10^7$ CFU/ gram of feces in mammals. It was found that *E. coli* had higher numbers in mammals as opposed to birds. Thus, larger amounts of bird feces are required to be deposited compared to mammalian feces in order to create similar *E. coli* counts (Gordon and Cowling 2003).

E. coli can be distinguished quite easily because 95% have β-glucuronidase activity which can convert BCIG (5-bromo-4-chloro-3-indolyl-beta-D-glucuronide) to produce a visible blue colour (Watkins *et al.* 1988). *E. coli* is also a good fecal indicator as it provides a relatively efficient method to culture and differentiate from other bacteria. This rapid method is the reason that *E. coli* is being utilized as a fecal indicator not only in North America but worldwide.

1.3 Naturalized *E. coli*

Recent studies have shown that *E. coli* can also survive naturally in the environment. Ishii *et al.* (2006) have shown that *E. coli* can grow in non-enriched and non-sterile soil. These were

found in northern temperate soils in Lake Superior watersheds. The presence of the bacteria living in this soil may invalidate the use of *E. coli* as a fecal indicator.

The survival of *E. coli* in the environment seems unlikely due to the various stresses. Variations in temperature, pH, as well as salinity levels make it difficult for *E. coli* to survive (Conner and Kotrola 1995; Sadowsky and Whitman 2011). Solar Radiation was also found to inactivate *E. coli* and can be looked at as a major environmental stress as well (Muela *et al.* 2000). Recently though, *E. coli* was found to not only survive but to grow in tough conditions in the environment. In tropical regions such as Hawaii, *E. coli* were shown to be able to survive in the soil in small, but significant populations (Byappanahalli and Fujioka 1998). Alternatively, northern temperate soils from Lake Superior watersheds have also shown the presence of naturalized *E. coli* populations (Ishii *et al.* 2006).

Overall, *E. coli* appears to be able to survive and grow in various climates. Ksoll *et al.* (2007) found that *E. coli* isolated from the periphyton at the beginning of the winter season could be isolated again in the coming spring. This shows that the *E. coli* could persist through the winter in the periphyton communities even when the air temperature reached a low of -40°C. Even off shore, *E. coli* was observed to survive in *Cladophora* mats. It is possible that these algal blooms act as a secondary habitat to the fecal indicator bacteria (Whitman *et al.* 2003). *E. coli* was also found to thrive in the environment in free living coliform blooms in Australian Lakes (Power *et al.* 2005).

E. coli can survive and replicate in the environment (soil, sands, algae, and periphyton), therefore there is a need to differentiate the environmental and fecal *E. coli* (Goto and Yan 2011). These data support the idea that the environmental *E. coli* should be grouped into a source

of its own when tracking down the source of the contamination. How does *E. coli* live in these environments? The cells can adhere to particles as well as each other in what are known as biofilms, which can sustain the bacteria in the environment even under nutrient deprivation.

1.4 Biofilms

E. coli as well as other bacteria can survive on rocks, soil and other moist environments. Bacteria seem to respond to environmental stresses by initiating biofilm formation (O'toole *et al*. 2000). These formations can consist of algae, fungi, protists, as well as bacteria known collectively as periphyton. Outside of the natural environment, biofilms are known to form as dental plaque as well as in the drinking water distribution systems (O'Toole *et al.* 2000). Bacteria living in these communities are physiologically different from bacteria in the planktonic water. They excrete slimy material known as extracellular polymeric substances (EPS) to form the internal structure. This occurs after it establishes a foothold on a surface (Stewart and Franklin 2008; Maier *et al.* 2009). The EPS can help the biofilms become resistant to antibiotics as well as detergents (Stewart and Franklin 2008).

Environmental stress such as nutrient availability can cause the bacteria to initiate biofilm formation. Biofilms are a stable point in a biological cycle that includes initiation, maturation, maintenance, and dissolution (O'Toole *et al.* 2000). To start off, micro colonies will begin to form and layers of cells begin to thicken in order of cell growth. The rate of growth for biofilms is dependent on factors that include flow rate, nutrient content, and temperature (Palmer and White 1997). Initiation of biofilm attachment is limited by the texture or roughness of the substratum, the hydrophobicity as well as the conditioning film. As the biofilms mature they can

be slowed down when temperatures begin to lower (Donlan 2002). Mature biofilms form complex architectural structures and consist of large quantities of EPS. After the biofilms mature and reach their maximum growth they go into their dispersal phase. The mechanisms for dispersion are not quite understood but it was found that the biofilms will disperse when nutrients start to diminish or the daughter cells begin to shed from actively growing cells (O'Toole *et al.* 2000; Donlan 2002). After the biofilms disperse the bacteria can colonize new areas and initiate the attachment phase. Research on the functions and regulatory pathways are still being conducted on the release of bacteria from the biofilm (O'Toole *et al.* 2000).

It is possible for the bacteria in these biofilms to become dispersed in the environment through natural occurrences. Flood disturbances have been shown to have an impact on the periphytic bacteria in streams (Biggs 1995) and may cause the bacteria to break off the surface it is adhering to. In the middle of summer when temperatures are high and the biofilms have grown to saturation in lakes, the bacteria would most likely be broken off from the biofilms into the water if exposed to a larger flow rate (Costerton *et al.* 1999). This could potentially raise the counts of fecal indicator bacteria in the water over the maximal level allowed. Lee *et al*. (2006) measured levels of fecal indicator bacteria from the sediments at a beach after storm occurrences. They found that elevated levels of *E. coli* could be coming from the sediments as a result of the storm.

1.5 *Bacteroides*

Members of the *Bacteroides* genus are gram negative, obligate anaerobic bacteria. They are restricted to warm blooded mammals and are not known to survive well in the environment

(Sadowsky and Whitman 2011). *Bacteroides* is known to be resistant to kanamycin, neomycin, penicillin, aminoglycosides, and beta-lactam antibiotics. Its growth in the environment is limited mainly due to protozoa grazing (Bell *et al.* 2009). Other environmental factors include bacteriophage infection, pH, initial cell concentration and cell aggregate size. Having large initial cell concentrations and cell aggregate sizes allows the *Bacteroides* populations to survive for a longer period of time. It was also found that the *Bacteroides* could survive longer at lower temperatures and higher salinity (Okabe and Shimazu 2007; Balleste and Blanch 2010). Balleste and Blanch (2010) found that environmental *Bacteroides* strains in the sewage system survived better than the laboratory strains, but they still had a higher die-off rate compared to fecal coliforms.

Bacteroides species are found in high abundance in feces and account for about 30% of the total fecal isolates (Sadowsky and Whitman 2011). Since *Bacteroides* spp. do not survive well in freshwater due to being obligate anaerobes, they would make a good indicator for recent fecal contamination (Balleste and Blanch 2010). They are also less likely to reproduce once in the environment as well. Key characteristics to fecal indicators include broad geographical stability, ability to be tested easily and having high numbers in the host animal (Maier *et al.* 2009)**.** *Bacteroides* spp. has also been shown to have high host specificity which may be due to differences in host animal digestive systems (Layton *et al.* 2006). The detection of *Bacteroides* can generally be correlated to the *E. coli* fecal indicator in the presence of fecal pollution, because both types of bacteria can be found in high amounts in feces (Bower *et al.* 2005). Recently, molecular tools for detecting *Bacteroides* 16S rDNA genes have been quite successful and will be discussed in the library-independent methods of microbial source tracking.

1.6 Microbial Source Tracking

Microbial source tracking methods have been an effective way of determining the source of fecal pollution. Sources of fecal contamination can include human waste, farm animals, wildlife, and waterfowl. Lands impacted with manure were found to enable pathogenic bacteria to survive longer, possibly due to more nutrients. Environmental sources that occur naturally can be found in riparian soil, sediments, and beach sands. The growth of fecal coliforms and *E. coli* were also found to persist within the periphytic communities over the winter season (Ksoll *et al.* 2007). Ksoll *et al.* (2007) isolated *E. coli* strains from the periphytic communities throughout the year and their similarity strongly indicates that they persist over the winter.

Determining the source of contamination can prepare for detection of potential pathogens that are associated with the animal. Pathogenic bacteria *(Shigella, Salmonella*, and *Campylobacter*), viruses (norovirus, hepatitis A), and protozoa (*Cryptosporidium* and *Giardia*) can come from various host sources (Ishii and Sadowsky 2008). Once the source is determined, a possible solution to the problem can then be assessed.

Microbial source tracking methods can be library or non-library based. The library based or library-dependent technique requires isolating the required indicator bacteria from many fecal sources and classifying them by phenotype or genotype (Anderson *et al.* 2005). Phenotypic characterizations usually include carbon utilization and antibiotic resistance (Konopka *et al.* 1998; Sayah *et al.* 2005). The use of antibiotics to source human feces stemmed from its likelihood to have higher antibiotic resistance. Bacteria of the same species tend to have differences in their use of carbon and nitrogen for growth and survival and thus can be differentiated based on these properties (Simpson *et al.* 2002). A large disadvantage in creating

an antibiotic phenotypic library is that the isolates have to show antibiotic resistance or they will not be classified properly (Maier *et al.* 2009). The major disadvantage of the carbon utilization method is its inability to be measured quantitatively and will lead to an underestimation of the carbon utilization diversity in a bacterial population (Seurinck *et al*. 2005).

Genotypic characterizations which include Pulsed-field gel electrophoresis (PFGE), Restriction fragment length polymorphism (RFLP), Random amplified polymorphic DNA (RAPD) assay, Amplified fragment length polymorphism (AFLP), and Repetitive DNA elements polymerase chain reaction (Rep-PCR) techniques have been found to be more accurate than phenotypic characterizations due to less variability (Ishii and Sadowsky 2008). Once these libraries are made they can predict the source of the isolates from the water. The major problem with these libraries is that they are restricted to regions where the water bodies have a limited number of potential sources. They are also limited to the specific geographical regions they are tested for as there is variability among host distribution of bacterial strains (Moore *et al.* 2005).

Non-library based or library-independent techniques are a quick method to detect specific host/source organisms. The absence of creating a library saves time and money but the host specific markers need to be tested thoroughly in the target population. These tests are rapid, require minimal analysis and do not require culturing (Meays *et al.* 2004; Maier *et al.* 2009). Tests that differentiate human fecal pollution from bovine fecal pollution using *Bacteroides* have been found to be successful (Scott *et al.* 2002; Lee *et al.* 2010). Host specific 16S rDNA markers have been discovered for *Bacteroides*. These gene markers have been found for humans, bovine, and pigs as well as a gene marker for total *Bacteroides* (Layton *et al.* 2006; Okabe *et al.* 2007). As genetic markers are still being developed for source typing, it must be used with caution and further validation is needed. The survival and distribution of the molecular markers in the aquatic

environment has to be tested in the specific area that it is used for (Maier *et al.* 2009). Quantitative PCR techniques can detect fecal contamination with the total *Bacteroides* or ALLBac genetic markers (Jeong *et al.* 2010). Since the quantitative PCR method detects the 16S rDNA copy number of *Bacteroides*, it does not matter if the *Bacteroides* is living or not. Therefore, using this analysis can result in false-positive signals of *Bacteroides* due to dead microbial cells but increases in fecal matter will certainly cause an increase in copies of DNA markers. The PCR assays are sensitive to contamination, which can inhibit the detection of the genetic markers of *Bacteroides* (Sadowsky and Whitman 2011). The use of the quantitative PCR technique gives more information than the regular PCR which only detects the presence or absence of the specific molecular marker. Quantitative PCR techniques can quantify the amount of target marker and indicate influxes of fecal contamination rather than just the presence of it.

There are two main quantitative PCR methods. The first one uses a fluorescent dye called SYBR Green. SYBR Green binds to doubled stranded DNA (dsDNA) and the increased fluorescent activity correlates with an increase in dsDNA (Ramakers *et al.* 2003). The second method also uses a fluorescent technique known as a Taqman probe. A hybridization probe is required which is labeled with two different fluorescent dyes. The first dye is known as a reporter dye (e.g. FAM) and the second is known as a quencher dye (e.g. TAMRA). As long as the probe is intact, fluorescent emissions from the reporter dye will be absorbed by the quencher. The emissions will only be adsorbed when they are in close proximity. In the PCR the dual labeled probe hybridizes to the target sequence. When the DNA polymerase starts its extension the fluorescent hybridization probe will be cleaved by the 5'-3' exonuclease activity of the DNA polymerase. This causes the probe to be cleaved and the fluorescent emission of the reporter dye is no longer in close proximity to the quencher. Therefore it can no longer adsorb the emissions

which result in an increased fluorescent emission of the reporter dye (Gibson *et al.* 1996; Heid *et al.* 1996). The amplification of the target DNA will cause an increase in the cleaving of the dual labeled probe and thus a brighter emission. The SYBR Green binds to any double stranded DNA including non-specific amplifications. This will result in false positives but this is resolved by the melting curve analysis. The Taqman Probe will only detect specific amplifications products. When looking for a different gene to amplify, a new probe will need to be synthesized for the unique target sequence. This makes the SYBR Green the more efficient method of the two when it may take too long to synthesize a new probe.

Both library-dependent and library-independent methods have their advantages and disadvantages. Genotypic or phenotypic library-based methods are easy to perform and highly reproducible. Major problems come from having to produce a large isolate database and these can only be used for a specific geography (Maier *et al*. 2009). The library-independent methods are more rapid than the library based methods as no cultivation is necessary. The survival of the molecular markers in the aquatic environments need more research and these techniques are currently applicable to a limited number of host groups (Maier *et al.* 2009).

1.7 Source Typing Methods

There are many different DNA based typing methods, which include Pulsed-field gel electrophoresis (PFGE), Restriction fragment length polymorphism (RFLP), Random amplified polymorphic DNA (RAPD) assay, Amplified fragment length polymorphism (AFLP), and Repetitive DNA elements polymerase chain reaction (Rep-PCR) to name a few (Olive and Bean 1999). PFGE allows the separation of large fragments of DNA (10 – 800 kb) by switching the

direction of the current in the agarose gel. The large fragments eventually separate from each other and a distinct pattern can be compared (Tenover *et al.* 1995). RFLP uses restriction enzymes to digest the chromosomal DNA of prokaryotic or eukaryotic organisms and the fragments are then separated on an agarose gel. The separated DNA can be transferred to a membrane by blotting. The DNA fragments that are specific to the probes are then hybridized and then can be visualized (Oliver and Bean 1999). RAPD uses primers with about 10 nucleotides in a PCR with low annealing temperatures. The RAPD assay uses the genomic DNA, but problems arise when there is contaminant DNA as the amplification targets 10 arbitrary short oligonucleotide primers and there is no knowledge of the target DNA sequence (Hadrys *et al*. 1992; Lynch and Milligan 1994; Olive and Bean 1999). AFLP is a DNA fingerprinting technique that can work with different types of genomes including plants, fungi, animals and bacteria. The DNA fragments generated by a restriction enzyme digestion are amplified and the pattern can be compared (Vos *et al.* 1995; Olive and Bean 1999). The AFLP technique has also been found to differentiate pathogenic *E. coli* from nonpathogenic strains (Leung *et al.* 2004). Lastly, Rep-PCR has become a popular technique to compare similarities between strains. The approach amplifies the regions between repetitive DNA elements present in the bacteria genomes to create a distinct pattern (de Bruijn 1992; Versalovic *et al.* 1994; Olive and Bean 1999). All the aforementioned methods have their own use in DNA fingerprinting, but the Rep-PCR was found to be the most cost effective and efficient method to differential bacterial strains (Olive and Bean 1999).

1.8 Differentiation of Animal, Human and Periphytic *E. coli* Using Rep-PCR

Intergenic repeated sequences can be found on the genomic DNA in both eukarytic as well as bacterial cells. They are known to be highly conserved and these repeats were found to constitute approximately 5% of the bacterial genome (Ussery *et al.* 2004). Some processes to differentiate genomes based on these repeats are known as REP (Repetitive Extragenic Palindromic), ERIC (enterobacterial repetitive intergenic consensus), and BOX PCR, all of which fall under the category of Rep-PCR. This procedure uses primers specific to the intergenic repeated sequences and amplifies DNA fragments between them. REP, ERIC and BOX all have different intergenic repeated sequences. Rep-PCR has been able to differentiate bacterial strains based on their genomic DNA (de Bruijn 1992). A distinct genotypic DNA fingerprint can be retrieved from this method and used to compare different sources.

Dombek *et al.* (2000) used REP-PCR to differentiate chickens, sheep, pigs, geese, cows, humans and ducks. Along with a discriminant analysis technique, REP-PCR was shown to successfully differentiate all the source groups. In 2009, Kon *et al.* (2009) demonstrated the ability to differentiate *E. coli* with REP-PCR between animal hosts as well as the environmentally adapted strains of *E. coli* found in the interstitial waters from a beach.

There is sufficient evidence that strains of *E. coli* can establish and grow in the environment. The use of the REP-PCR methods may be able to distinguish naturalized populations from specific host source populations and paint a more accurate picture of fecal contaminants in the waters.

1.9 Thesis Objectives

E. coli is widely used as a fecal indicator and any increases in its population will be assumed to be coming from fecal contamination. Boulevard Lake in Thunder Bay, Ontario experiences instances of high levels of *E. coli* throughout the summer seasons and shuts down its beach whenever the *E. coli* reaches the maximum limits. The public generally assumes that these elevated levels of *E. coli* are coming from geese feces. Our previous studies on Boulevard Lake showed that *E. coli* could persist in the periphytic communities (Moreira *et al.* 2012). We propose that the high levels of *E. coli* in Boulevard Lake are coming from the naturalized *E. coli* in the periphytic communities. We will use quantitative PCR (qPCR) and microbial source tracking (MST) techniques to validate this hypothesis.

The long-term goal of this study is to review and test out two different molecular techniques to differentiate and source fecal contamination. Quantitative PCR (qPCR) will be used to find *Bacteroides* specific molecular markers in the freshwater to indicate fecal contaminants. Microbial source tracking (MST) will be used to differentiate *E. coli* isolates from various host sources to create a DNA fingerprint library. Boulevard Lake will be used as a case study and the DNA fingerprint library will be used to compare the *E. coli* isolates in the water and determine the source(s) of the high bacterial levels. The specific research objectives are as follows:

1. Validate the use of *Bacteroides* ALLBac primers with fecal samples from different animals and various aquatic samples with and without fecal contamination.
2. Monitor the *Bacteroides* molecular markers over the 2011 season at Boulevard Lake.
3. Monitor the *E. coli* population over the 2010 and 2011 season at Boulevard Lake.
4. Compare the use of *Bacteroides* and *E. coli* and their effectiveness as fecal indicator bacteria.

5. Create an *E. coli* REP-PCR DNA fingerprint library from possible sources around and upstream of Boulevard Lake.
6. Determine the source of the planktonic *E. coli* in Boulevard Lake by the REP-PCR DNA fingerprinting method.

1.10 References

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

Anderson, K.L., Whitlock, J.E. and Harwood, V.J. (2005) Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Appl Environ Microbiol* **71**, 3041-3048.

Balleste, E. and Blanch, A.R. (2010) Persistence of Bacteroides species populations in a river as measured by molecular and culture techniques. *Appl Environ Microbiol* **76**, 7608-7616.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R. and Sayler, G.S. (2009) Factors influencing the persistence of fecal Bacteroides in stream water. *J Environ Qual* **38**, 1224-1232.

Biggs, B.J.F. (1995) The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. *Freshwater Biol* **33**, 419-438.

Bower, P.A., Scopel, C.O., Jensen, E.T., Depas, M.M. and Mclellan, S.L. (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. *Appl Environ Microbiol* **71**, 8305-8313.

Byappanahalli, M.N. and Fujioka, R.S. (1998) Evidence that tropical soil environment can support the growth of *Escherichia coli*. *Water Sci Technol* **38**, 171-174.

Clermont, O., Bonacorsi, S. and Bingen, E. (2000) Rapid and simple determination of the Escherichia coli phylogenetic group. *Appl Environ Microbiol* **66**, 4555-4558.

Costerton, J.W., Stewart, P.S. and Greenberg, E.P. (1999) Bacterial biofilms: a common cause of persistent infections. *Science* **284**, 1318-1322.

de Bruijn, F.J. (1992) Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of Rhizobium meliloti isolates and other soil bacteria. *Appl Environ Microbiol* **58**, 2180-2187.

Donlan, R.M and Costerton, J.W. (2000) Biofilms: survival mechanisms and clinically relevant microorganisms. *Clin Microbiol Rev* **15**, 167-193.

Dorfman, M., and K.S. Rosselot. 2009. Testing the waters: a guide to water quality at vacation beaches. San Francisco: Natural Resources Defense Council.

Edberg, S.C., Rice, E.W., Karlin, R.J. and Allen, M.J. (2000) Escherichia coli: the best biological drinking water indicator for public health protection. *Symp Ser Soc Appl Microbiol* **(29)**, 106S-116S.

Gibson, U.E., Heid, C.A. and Williams, P.M. (1996) A novel method for real time quantitative RT-PCR. *Genome Res* **6**, 995-1001.

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. *Environ Sci. Technol* **39**, 5157-5169.

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

Goto, D.K. and Yan, T. (2011) Genotypic diversity of *Escherichia coli* in the water and soil of tropical watersheds in Hawaii. *Appl Environ Microbiol* **77**, 3988-3997.

Hadrys, H., Balick, M. and Schierwater, B. (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol Ecol* **1**, 55-63.

Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M. (1996) Real time quantitative PCR. *Genome Res* **6**, 986-994.

Hrudey, S.E., Payment, P., Huck, P.M., Gillham, R.W. and Hrudey, E.J. (2003) A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. *Water Sci Technol* **47**, 7-14.

Ishii, S., Ksoll, W.B., Hicks, R.E. and Sadowsky, M.J. (2006) Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Appl Environ Microbiol* **72**, 612-621.

Ishii, S. and Sadowsky, M.J. (2008) *Escherichia coli* in the environment: implications for water quality and human health. *Microbes Environ* **23**,101-108.

Isobe, K.O., Tarao, M., Chiem, N.H., Minh, L.Y. 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. *Appl Environ Microbiol* **70**, 814-821.

Jeong, J.Y., Park, H.D., Lee, K.H., Hwang, J.H and Ka, J.O. (2010) Quantitative analysis of human- and cow-specific 16S rRNA gene markers for assessment of fecal pollution in river waters by real-time PCR. *J Microbiol Biotechnol* **20**, 245-253.

Kon, T., Weir, S.C., Howell, E.T., Lee, H. and Trevors, J.T. (2009) Repetitive element (REP)-polymerase chain reaction (PCR) analysis of *Escherichia coli* isolates from recreational waters of southeastern Lake Huron. *Can J Microbiol* **55**, 269-276.

Konopka, A., Oliver, L., Turco Jr, R.F. (1998) The use of carbon substrate utilization patterns in environmental and ecological microbiology. *Microbial Ecol* **35**, 103-115.

Ksoll, W.B., Ishii, S., Sadowsky, M.J. and Hicks, R.E. (2007) Presence and sources of fecal coliform bacteria in epilithic periphyton communities of Lake Superior. *Appl Environ Microbiol* **73**, 3771-3778.

Lamendella, R., Santo Domingo, J.W., Kelty, C. and Oerther, D.B. (2008) *Bifidobacteria* in feces and environmental waters. *Appl Environ Microbiol* **74**, 575-584.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler, G. (2006) Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Appl Environ Microbiol* **72**, 4214-4224.

Lee, C.M., Lin, T.Y., Lin, C.C., Kohbodi, G.A., Bhatt, A., Lee, R. and Jay, J.A. (2006) Persistence of fecal indicator bacteria in Santa Monica Bay beach sediments. *Water Res* **40**, 2593-2602.

Lee, D.Y., Weir S.C., Lee, H. and Trevors, J.T. (2010) Quantitative identification of fecal water pollution sources by TaqMan real-time PCR assays using Bacteroidales 16S rRNA genetic markers. *Appl Microbiol Biotechnol* **88**, 1373-1383.

Leung, K.T., Mackereth, R., Tien, Y.C. and Topp, E. (2004) A comparison of AFLP and ERIC-PCR analyses for discriminating Escherichia coli from cattle, pig and human sources. *FEM Microbiol Ecol* **47**, 111-119.

Lynch, M. and Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. *Mol Ecol* **3**, 91-99.

Madigan, M.T., Martinko, J.M., Stahl, D.A. and Clark, D.P. (2012) *Brock biology of microorganisms.* pp.541. 13th ed. San Francisco: Pearson Prentice Hall.

Maier, R.M., Pepper, I.L. and Gerba, C.P. (2009) *Environmental microbiology*. pp.485-486. 2nd ed. Burlington: Elsevier.

Meays, C.L., Broersma, K., Nordin, R. and Mazumder, A. (2004) Source tracking fecal bacteria in water: a critical review of current methods. *J Environ Manage* **73**, 71-79.

Moore, D.F., Harwood, V.J., Ferguson, D.M., Lukasik, J., Hannah, P., Getrich, M. and Brownell, M. (2005) Evaluation of antibiotic resistance analysis and ribotyping for identification of faecal pollution sources in an urban watershed. *J Appl Microbiol* **99**, 618-628.

Moreira, S., Brown, A., Ha, R., Iserhoff, K., Yim, M., Yang, J., Liao, B., Pszczolko, E., Qin, W. and Leung, K.T. (2012) Persistence of *Escherichia coli* in freshwater periphyton: biofilm-forming capacity as a selective advantage. *FEMS Microbiol Ecol* **79**, 608-618.

Muela, A., Garcia-Bringas, J.M., Arana, I.I. and Barcina, I.I. (2000) The Effect of Simulated Solar Radiation on Escherichia coli: The Relative Roles of UV-B, UV-A, and Photosynthetically Active Radiation. *Microb Ecol* **39**, 65-71.

Noblet, J.A., Young, D.L., Zeng, E.Y. and Ensari, S. (2004) Use of fecal steroids to infer the sources of fecal indicator bacteria in the Lower Santa Ana River Watershed, California: sewage is unlikely a significant source. *Environ Sci Technol* **38**, 6002-6008.

Okabe, S., Okayama, N., Savichtcheva, O. and Ito, T. (2007) Quantification of host-specific Bacteroides-Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Appl Microbiol Biotechnol* **74**, 890-901.

Okabe, S. and Shimazu, Y. (2007) Persistence of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. *Appl Microbiol Biotechnol* **76**, 935-944.

Olive, D.M. and Bean, P. (1999) Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol* **37**, 1661-1669.

O'Toole, G., Kaplan, H.B. and Kolter, R. (2000) Biofilm formation as microbial development. *Annu Rev Microbiol* **54**, 49-79.

Paerl, H.W., Dyble, J., Moisander, P.H., Noble, R.T., Piehler, M.F., Pinckley, J.L., Steppe, T.F., Twomey, L. and Valdes, L.M. (2003) Microbial indicators of aquatic ecosystem change: current applications to eutrophication studies. *FEMS Microbiol Ecol* **46**, 233-246.

Palmer, R.J. JR. and White, D.C. (1997) Developmental biology of biofilms: implications for treatment and control. *Trends Microbiol* **5**, 435-440.

Parker, J.K, McIntyre, D. and Noble, R.T. (2010) Characterizing fecal contamination in stormwater runoff in coastal North Carolina, USA. *Water Res* **44**, 4186-4194.

Piocos, E.A. and dela Cruz, A.A. Solid phase extraction and high performance liquid chromatography with photodiode array detection of chemical indicators of human fecal contamination in water. *J Liq Chrom & Rel Technol* **23**, 1281-1291.

Power, M.L., Littlefield-Wyer, J., Gordon, D.M., Veal, D.A. and Slade, M.B. (2005) Phenotypic and genotypic characterization of encapsulated Escherichia coli isolated from blooms in two Australian lakes. *Environ Microbiol* **7**, 631-640.

Rabinovici, S.J., Bernknopf, R.L., Wein, A.M., Coursey, D.L. and Whitman, R.L. (2004) Economic and health risk trade-offs of swim closures at a Lake Michigan beach. *Environ Sci Technol* **38**, 2737-2745.

Ramakers, C., Ruijter, J.M., Deprez, R.H. and Moorman, A.F. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**, 62-66.

Sadowsky, M.J. and Whitman, R.L. (2011) *The Fecal Bacteria.* Washington, DC: ASM Press.

Seurinck, S., Verstraete, W. and Siciliano, S.D. (2005) Microbial source tracking for identification of fecal pollution. *Rev Environ Sci BioTechnol* **4**, 19-37.

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R. and Lukasik, J. (2002) Microbial source tracking: current methodology and future directions. *Appl Environ Microbiol* **68**, 5796-5803.

Simpson, J.M., Santo Domingo, J.W. and Reasoner, D.J. (2002) Microbial source tracking: state of the science. *Environ Sci Technol* **36**, 5279-5288.

Stewart, P.S. and Franklin, M.J. (2008) Physiological heterogeneity in biofilms. *Nat Rev Microbiol* **6**, 199-210.

Sayah, R.S., Kaneene, J.B., Johnson, Y. and Miller, RA. (2005) Patterns of antimicrobial resistance observed in Escherichia coli isolates obtained from domestic- and wild-animal fecal samples, human septage, and surface water. *Appl Environ Microbiol* **71**, 1394-1404.

Tallon, P., Magajna, B., Lofranco, C. and Leung, K.T. (2005) Microbial indicators of faecal contamination: a current perspective. *Water Air Soil Pollut* **166**, 139-166.

Tenover, F.C., Arbeit, R.D., Goering, R.V., Mickelsen, P.A., Murray, B.E., Persing, D.H. and Swaminathan, B. (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* **33**, 2233-2239.

The Minister of National Health and Welfare. 1992. Guidelines for Canadian recreational water quality. Canadian Government Publishing Centre Supply and Services Canada, Ottawa, Canada.

Tyagi, V.K., Chopra, A.K., Kazmi, A.A. and Kumar, A. (2006) Alternative microbial indicators of faecal pollution: current perspective. *Iran J Environ Health Sci Eng* **3**, 205-216.

U.S. Environmental Protection Agency. *Water: monitoring & assessment. 5.11 Fecal bacteria*. http://water.epa.gov/type/rsl/monitoring/vms511.cfm. Accessed February 2012.

Ussery, D.W., Binnewies, T.T., Gouveia-Oliveira, R., Jarmer, H. and Hallin, P.F. (2004) Genome update: DNA repeats in bacterial genomes. *Microbiology* **150**, 3519-3521.

Versalovic, J., Schneider, M., de Bruijn, F.J. and Lupski, J.R. (1994) Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol* **5**, 25-40.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kulper, M. and Zabeau M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**, 4407-4414.

Wade, T.J., Pai, N., Eisenberg, J.N.S. and Colford Jr, J.M. (2003) Do US Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. *Environ Health Perspect* **111**, 1102-1109.

Watkins, W.D., Rippey, S.R., Clavet, C.R., Kelley-Reitz, D.J. and Burkhardt III, W. (1988) Novel compound for identifying Escherichia coli. *Appl Environ Microbiol* **54**, 1874-1875.

Wheeler, A.L., Hartel, P.G., Godfrey, D.G., Hill, J.L. and Segars, W.I. (2002) Potential of Enterococcus faecalis as a human fecal indicator for microbial source tracking. *J Environ Qual* **31**, 1286-1293.

Whitman, R.L., Shively, D.A., Pawlik, H., Nevers, M.B. and Byappanahalli, M.N. (2003) Occurrence of *Escherichia coli* and *Enterococci* in cladophora (chlorophyta) in nearshore water and beach sand of Lake Michigan. *Appl Environ Microbiol* **69**, 4714-4719.

Chapter 2 - Monitoring *Bacteroides* 16S rDNA Biomarker Levels in a Freshwater Lake with Incidence of *Escherichia coli* Blooms

2.0 Abstract

Escherichia coli and *Bacteroides* populations were monitored and compared at a freshwater lake. Standard BCIG-Differential Medium was used to monitor the planktonic and periphytic *E. coli* densities and a qPCR method to quantify the *Bacteroides* 16S rDNA biomarker in the lake water samples. Two times throughout the 2011 summer season the planktonic *E. coli* population densities exceeded the Canadian Recreational Water Quality Guidelines of 2.30 log CFU 100 ml^{-1}. These were 2.86 and 2.38 log CFU 100 ml^{-1} on July 21 and September 2, respectively. Despite a gradual increase in the *Bacteroides* biomarker density from spring to fall, no significant changes were observed during the two peak periods of *E. coli*. Furthermore, heterotrophic counts of the water samples did not show significant changes during the two *E. coli* peaks. Therefore, the evidence did not support substantial fecal contamination during the two *E. coli* blooms. A stable and persistent population of periphytic *E. coli* was observed at the site. It increased and leveled in July and August with a density averaging 2.78 log CFU 100 cm^{-2}. We propose that the periphytic *E. coli* population might be a source of the *E. coli* blooms during the summer season.

2.1 Introduction

Fecal contamination of water is a large area of concern. Contamination due to fecal matter is a threat to human health and is a problem worldwide (Tallon *et al*. 2005). It can also

lead to changes in the ecosystem due to eutrophication (Paerl *et al*. 2003). Fecal contamination does not only affect the environment and human health but also the economy. Closures of public beaches have been shown to have large economic impacts on businesses in the surrounding area (Rabinovici *et al*. 2009).

Escherichia coli have been used as fecal indicator bacteria (FIB) worldwide to determine the presence of fecal pollution in recreational waters. Generally, it is believed that *E. coli* grows exclusively in warm-blooded animals and is not able to replicate in the environment (Maier *et al*. 2009; Tallon *et al*. 2005). Furthermore, the presence of the bacterium indicates a correlation to fecal contamination (Edge and Hill 2004; Mclellan 2004; Kon *et al*. 2009).

Recently, *E. coli* has been found to persist in some natural environments which include beach sands, soil, and algae (Byappanahalli *et al*. 1998; Whitman *et al*. 2003; Ishii *et al*. 2006; Ksoll *et al*. 2007). Also, a study of two Australian lakes found coliform blooms with the same three *E. coli* strains in different geographical regions (Power *et al*. 2005). This showed that it may be possible for *E. coli* to persist in the environment. Ksoll *et al*. (2007) also found that *E. coli* survived in the periphytic communities over the winter months in Duluth Harbour, Minnesota. These findings place the reliability of using *E. coli* as the sole FIB under doubt. Therefore, additional fecal indicators should be used to validate the detection of fecal contaminants.

Other than *E. coli*, several major groups of bacteria can be used as indicators for fecal contamination. These include *Clostridium perfringens*, *Enterococcus*, *Bifidobacterium* and *Bacteroides*. *Clostridium perfringens* has varying levels among different animals and therefore may not be a good secondary FIB to *E. coli* (Yost *et al.* 2011). *Enterococcus* can be found

abundantly in the environment (Alm *et al*. 2003; Whitman *et al*. 2003). *Bifidobacterium* was found to have low cell densities in fecal materials. Therefore, it may not always correlate with the fecal pollution and health risks (Lamendella *et al*. 2008). *Bacteroides* is a good fecal indicator bacterium because it is found in high amounts and amongst different host species in feces (Yost *et al.* 2011). *Bacteroides* is an obligate anaerobe and its ability to grow in the environment is limited due to protozoan grazing, temperature, pH, initial cell concentrations as well as oxygen (Okabe *et al*. 2007; Bell *et al*. 2009; Yost *et al.* 2011). Due to its high amounts in feces, the detection of *Bacteroides* can generally be correlated to the *E. coli* fecal indicator in the presence of fecal pollution (Bower *et al*. 2005).

Quantitative PCR (qPCR) is a non-cultivation method that has been used to detect and quantify *Bacteroides* in water samples (Dick and Fields 2004). Since it is difficult to culture *Bacteroides*, qPCR is an ideal non-cultivation approach. Quantitative PCR is used to estimate the amount of *Bacteroides* 16S rDNA genetic biomarkers. Specific primers have been developed to detect the presence of *Bacteroides* (Dick and Fields 2004).

As *E. coli* is being scrutinized as a fecal indicator bacterium, our research will use *Bacteroides* as a secondary fecal indicator bacterium to verify the presence of fecal contamination. Every year, Boulevard Lake in Thunder Bay, Ontario, Canada has incidents of elevated levels of *E. coli* in the water causing beach closures. In our previous study, it was shown that an *E. coli* population had persisted within the periphytic community in Boulevard Lake (Moreira *et al*. 2012). Therefore, it is uncertain if the high levels of *E. coli* during the *E. coli* blooms in Boulevard Lake are indicative of fecal contamination. The objective of this study is to use *Bacteroides* as a secondary fecal indicator to evaluate the reliability of using *E. coli* as the sole FIB to determine fecal contamination at the Boulevard Lake beach.

2.2 Materials and Methods

2.2.1 Collection of periphyton and lake water samples

Water and periphytic samples were taken at three locations in Boulevard Lake (48°27′34″N and 89°12’26″W) in Thunder Bay, Ontario, Canada. Site 1 and 2 were taken adjacent to the main beach. Site 3 was taken on a rocky shore approximately 500m downstream of the main beach. Two sterile 1L Nalgene bottles were used at each site to collect water approximately one foot under the surface. At each site, a rock submerged 30-40 cm below the water's surface was selected. Prior to removal from the water, the rock was gently shaken 10 times to remove loose sediment from its surface. The rock was placed on shore with the surface to be sampled facing up. A sterile, square rubber template delineating an area of 10 cm x 10 cm was placed on the rock and used to quantify the substratum surface. All material was scraped from within the area marked by the template using a 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, 0.24g KH_2PO_4, 1 L distilled H_2O, pH adjusted to 7.4). Three rocks were selected from each site for enumeration.

2.2.2 Enumeration of planktonic and periphytic *E. coli*

For enumeration of planktonic *E. coli*, 20 mL of lake water was 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* and used to compute the overall population at that time point. A similar technique was used for the enumeration of periphytic *E. coli*, except 10 mL of the periphyton sample suspended in PBS was filtered. Samples were taken on May 4, June 2, July 6, July 22, August 9, August 18, September 2, September 9, October 7, October 19, October 20 and November 10 of 2011. Additional planktonic *E. coli* data points between June 27 and August 22 were provided by the regional Public Health Laboratory (PHL) in Thunder Bay, Ontario. The PHL used a standard basal mFC-BCIG *E. coli* detection method (Ciebin *et al.* 1995), which uses an incubation temperature of 43-44 °C, to enumerate the number of *E. coli* in the beach water samples obtained from Boulevard Lake.

2.2.3 Heterotrophic background bacteria of planktonic and periphytic samples

Heterotrophic background bacteria were enumerated from both the lake water and the periphytic samples. Both the water and periphytic samples collected in the manner previously described were serial diluted and spread-plated onto R2A agar (Becton, Dickinson, and Co.). The plates were incubated for 48 h at 30 °C and the total number of colony forming units were counted and recorded.

2.2.4 Environmental conditions and physicochemical analysis of water

Immediately after the set of water samples was collected, a small volume was used to determine the pH (Fisher Scientific Accument Basic AB15, Toronto, ON, Canada) of the water

at each site. Furthermore, the chemical composition of the water samples collected during June, August, and October was determined with the aid of the Lakehead University Environmental Laboratory (LUEL) (Table 2-1). The concentrations of pertinent chemicals were recorded in mg l^{-1}. These chemical analyses were performed immediately after each sampling.

Analyses of the nitrate, nitrite, phosphate, sulphate, and chloride concentrations of the water samples were carried out using ion chromatography with the Dionex ICS1100 System. An AS14 Analytical Column of size 4 x 250 mm and an AG14 Guard Column of size 4 x 50 mm were also used, with both having a particle size of 9 μm. The columns were washed with a carbonate and bicarbonate eluent under high pressure.

Dissolved organic carbon (DOC) was determined for each sample using an Autoanalyser Skalar SAN System. Samples were diluted with a 30 mmol l^{-1} solution of sulphuric acid and washed with pure oxygen and nitrogen gas in order to remove any inorganic carbonate. They were then digested with a solution containing 45 mmol l^{-1} potassium persulphate and 90 mmol l^{-1} sodium tetraborate. After digestion, the samples were exposed to an ultraviolet source which caused persulphate free radicals to form which interacted with the organic carbon to form carbon dioxide and water. A colour reagent was then prepared containing 0.875 mmol l^{-1} sodium carbonate, 0.875 mmol l^{-1} sodium hydrogen carbonate, 0.5 ml phenolphthalein and 2 ml Triton X-100 in 1000 ml distilled water. After a gas dialysis was performed with the digested samples, the separated carbon dioxide was exposed to the colour reagent and the absorbance of the gas was measured at 550 nm. Finally, a standard curve was used to quantify the DOC concentration of the samples.

The heavy metal concentrations of the water samples, as well as those for calcium, sodium, magnesium and potassium, were determined using inductively coupled plasma atomic emission spectroscopy (ICP-AES). In the cases of the heavy metal analyses, samples were digested using a CEM Mars 5 microwave, dissolved in nitric acid and concentrated five-fold before analysis. For all of these tests, a Varian Vista Pro CCD Simultaneous ICP-OES and CETAC ASX-510 Auto Sampler were used.

The data for the precipitation and the daily high and low temperatures were obtained from the National Climate Data and Information Archive (www.climate.weatheroffice.gc.ca).

2.2.5 PCR detection of *Bacteroides*

Prior to monitoring the level of *Bacteroides* 16S rDNA biomarker in Boulevard Lake, the specificity of the AllBac primer set (AllBac296f, 5'-GAGAGGAAGGTCCCCCAC-3' and AllBac412r, 5'-CGCTACTTGGCTGGTTCAG-3') (Layton *et al*. 2006) was confirmed using a conventional PCR method with genomic DNA of twelve animal fecal samples, eight non-*Bacteroides* bacteria as negative controls, and a *Bacteroides fragilis* genomic DNA sample (ATCC 25285D-5) as a positive control (Cedarlane, Burlington, ON) (Fig 2-2). The cow, duck, sheep, horse, rabbit, goat and chicken fecal samples were provided by the Gammondale Farm located in Thunder Bay, Ontario; pig fecal samples by the Sandy Acres Farm, Thunder Bay; human fecal samples by students in our laboratory; dog fecal samples by domestic dog owners in Thunder Bay; and goose and deer fecal samples collected from the vicinity of Boulevard Lake. The eight negative controls were ATCC type strains including *Enterococcus faecalis* ATCC 29212, *Bacillus subtilis* ATCC 6633, *Escherichia coli K-12 strain MG1655*, *Serratia marcescens*

ATCC 8100, *Pseudomonas fluorescens* ATCC 49838, *Streptococcus latics* ATCC 11454, *Staphylococcus aureus* ATCC 25923 and *Salmonella typhimurium* ATCC 14028.

Genomic DNA from overnight pure culture was extracted using a sterile XS buffer (Tillett *et al*. 2000) containing 1% potassium ethyl xanthogenate, 100 mmol l^{-1} Tris-HCl (pH 7.4), 20 mmol l^{-1} EDTA (pH 8.0), 1% sodium dodecyl sulfate, and 800 mmol l^{-1} ammonium acetate. Harvested cells were suspended in the XS buffer and incubated at 70^{o}C for 15 minutes. They were then placed on ice for 30 minutes, after which they were centrifuged at 24,100 x g for 10 minutes. The supernatant was transferred to a new sterile 1.5 ml microcentrifuge tube and mixed with one volume of 100% isopropanol for DNA precipitation and washed with 70% ethanol. DNA pellets formed by centrifugation at 24,100 x g for 10 minutes were left to air-dry and subsequently resuspended in sterile ddH_2O for the polymerase chain reaction.

Genomic DNA from the fecal samples was extracted using the UltraClean Fecal DNA Isolation Kit (MoBio Laboratories, Inc. Carlsbad, CA) according to the manufacturer's guide. Briefly, 0.25 g of a fecal sample was mixed with 550 µl of the Bead Solution in the Dry Bead Tubes provided by the DNA Isolation Kit. Sixty µl of the solution S1 and 200 µl of the solution IRS were added. After vortexing for 30 seconds, the tubes were placed in *FastPrep* (Qbiogene, Inc. Carlsbad, CA) twice at speed 4 for 20 seconds to disrupt the cells instead of using the MO BIO Vortex Adapter tube holder recommended by the protocol. The tubes were centrifuged at 10,000 x g for 30 seconds and the supernatant was transferred to a new sterile tube containing 250 µl of solution S2. The mixture was then incubated at 4°C for 5 minutes, after which it was centrifuged at 10,000 x g for 1 minute again. Avoiding the pellet, 450 µl of supernatant was taken and mixed with 900 µl of solution S3 and the mixture was vortexed. Next, this mixture was loaded onto the spin filter and centrifuged at 10,000 x g for 1 minute. The spin filter was washed

with 300 µl of solution S4 by centrifugation for 20 seconds at 10,000 x g after discarding the flow through. Then, the spin filter was carefully placed in a new clean tube and 50 µl of sterile ddH_2O was added to release the DNA from the filter. Suspended DNA in ddH_2O was collected by centrifugation at 10,000 x g for 30 seconds.

One µl Genomic DNA (approximately 10 ng DNA) was added to a PCR reaction mixture containing 0.2 mmol l^{-1} of each dNTP, 2.5 mmol l^{-1} $MgCl_2$, 1×PCR buffer, 1.0 U of *Taq* polymerase and 0.3 µmol l^{-1} of each primer. The protocol consisted of an initial denaturation at 94°C for 3 minutes, followed by 45 cycles, each consisting of stages at 95°C for 20 seconds, 62°C for 20 seconds, 72°C for 20 seconds , and a final extension at 72°C for 3 minutes. The PCR products were analyzed by electrophoresis in 2% agarose gel containing TAE buffer (40 mmol l^{-1} Tris-HCl, 20 mmol l^{-1} acetic acid and 1 mmol l^{-1} EDTA) and ethidium bromide (1µg l^{-1}).

2.2.6 Cloning of the *Bacteroides fragilis* 16S rDNA biomarker and qPCR

The 16S rDNA of *Bacteroides fragilis* was amplified with the primer set Bac32F and Bac708R (Bac32F, 5'- AACGCTAGCTACAGGCTT and Bac708, 5'- CAATCGGAGTTCTTCGTG)(Bernhard and Field 2000) and the PCR product was ligated into a cloning vector (pGEM) to produce a standard plasmid clone containing a single copy of the target 16S rDNA fragment of *Bacteroides fragilis*. PCR was performed to amplify the target 16S rRNA gene fragment using a touch-down temperature protocol consisting of 5 min at 94°C, followed by 10 cycles at 94°C for 15 s, 65°C for 45 s (decreasing 1°C per cycle) and 72°C for 60s, followed by 30 cycles consisting of 94°C for 15s, 55°C for 45s, and 70°C for 60s, ending with a final extension time of 10 minutes at 72°C (Layton *et al.* 2006). The PCR product was

cloned into the pGEM vector using the pGEM®-T Easy Vector Systems (Promega, Madison, WI) and transformed into a competent *E.coli* JM109 cell culture using the Transform Aid™ Bacterial Transformation Kit (Fermentas, Burlington, ON). The transformants with the desired 16S rDNA clone fragment were selected on LB plates containing 100 μg l^{-1} of ampicillin. The plasmid containing the rDNA fragment was extracted and digested with *EcoRI* to confirm presence of the insert. Plasmid DNA containing *Bacteroides* 16S rRNA genes was extracted using GeneJET™ Plasmid Miniprep Kit (Fermentas, Burlington, ON) according to the manufacturer's instructions and the plasmid DNA concentration was determined using the NanoDrop2000 Spectrophotometer (NanoDrop Products, Wilmington, DE). A series of 10-fold dilutions of the plasmid DNA extract was prepared in sterile ddH_2O and used as a template for each reaction of the standard curve. Another set of 10-fold dilutions was prepared in lake water to determine possibility of PCR inhibition by the lake water. Real-time quantitative PCR for the standard curve was performed as follows. Four μl of each plasmid DNA dilution sample was added to 21 μl of qPCR mixture containing 12.5 μl of Maxima® SYBR Green qPCR Master Mix (2X) (Fermentas, Burlington, ON), 2.5 mmol l^{-1} $MgCl_2$ and 0.3 μmol l^{-1} of AllBac296f and AllBac412r. PCR amplification protocol consisted of an initial 50°C for 2 minutes followed by a denaturation at 94°C for 3 minutes. After the initial denaturation step, the target DNA was amplified by 45 PCR cycles consisting of 95°C for 20 s, 62°C for 20 s, and 72°C for 20 s. A melting curve analysis was performed at the end of the PCR to distinguish nonspecific amplification including primer dimers from the targeted PCR product. The qPCR amplification was performed using the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad, Mississauga, ON). For the standard curve, each concentration (copies of rDNA per reaction) was plotted against mean threshold cycle values.

2.2.7 Comparing levels of *Bacteroides* biomarker and *E. coli* cell density in known fecal contaminated samples

The levels of *Bacteroides* biomarker and the densities of *E. coli* were compared in water samples collected from a sewage treatment plant, a body of water where fecal contaminants can be seen (Lake Tamblyn) and a residential well. A sterile ddH_2O water sample was used as a negative control for both the *Bacteroides* and *E. coli* assays. The raw sewage samples were collected from the Thunder Bay Sewage Treatment and Water Pollution Control Plant. Lake Tamblyn (Thunder Bay, Ontario) is a small artificial water body approximately 50 x 75 m in area located in the Thunder Bay campus of Lakehead University, Ontario, Canada. The lake is suspected of being contaminated with fecal materials generated by Canada geese that reside by its bank because goose droppings have been found along the shoreline the lake. The well water samples were collected from a residential well located in a rural neighbourhood on Cavar Road in the southwest region of Thunder Bay.

For DNA extraction, five hundred ml of a water sample (either ddH_2O, well water or Lake Tamblyn water) was drawn through a sterile 47-mm mixed cellulose ester filter (Fisherbrand water testing membrane filter, pore size 0.45 μm; Thermo Fisher Scientific, Whitby, ON, Canada) using a sterile pneumatic pump funnel filtration. For the sewage samples, 50 ml was filtered because volumes in excess of this amount clogged the filter. Then, the filter was placed in a tube for DNA extraction. DNA was extracted from each filter by the UltraClean Fecal DNA Isolation Kit (MoBio Laboratories, Inc. Carlsbad, CA) as described earlier. Four μl (20-30 ng DNA) of each DNA extract was used in the qPCR assay developed in this study. The

concentrations of *Bacteroides* 16S rDNA biomarker in the water samples were determined using the qPCR biomarker calibration curve.

The number of *E. coli* in the water samples was enumerated using the aforementioned method for *E. coli* level monitoring in Boulevard Lake. In this experiment, *E. coli* densities in the ddH_2O and well water samples were determined by passing 500 ml of each sample through a sterile 0.45 µl membrane filter (47-mm diameter).

2.2.8 Real time quantitative polymerase chain reaction assay to monitor the level of *Bacteroides* spp. in Boulevard Lake

Duplicate water samples were collected from each of the three sites at Boulevard Lake. One hundred ml of each water sample were filtered for DNA extraction as described previously. Quantitative PCR was performed on the samples as previously described to determine the concentration of *Bacteroides* biomarker in the lake water from June 2 to November 10 of 2011.

2.2.9 Statistical analysis

A one-way ANOVA analysis was performed to compare the means between the *E. coli* densities or *Bacteroides* biomarker densities throughout the sampling period in 2011. The F-test was used for multiple comparisons among the samples and the Tukey's Multiple Range Test was used to compare the difference between individual samples. IBM SPSS Statistics Version 19 was used for the analysis.

2.3 Results

2.3.1 Population dynamics of *E. coli* and heterotrophic background bacteria

The planktonic *E. coli* was found in low concentrations in spring from May 4, 2011 to June 2, 2011 at an average of 0.82 log CFU 100 ml^{-1} (Fig 2-1A). During the summer season, the background planktonic *E. coli* population was about 1.72 log CFU 100 ml^{-1}. However, the *E. coli* cell density at the Boulevard Lake beach peaked and exceeded the maximum limit of *E. coli* recommended by the Guidelines for Canadian Recreational Quality on July 21 and September 2 at 2.86 and 2.38 log CFU 100 ml^{-1}, respectively. In the fall, the planktonic *E. coli* density again decreased to an average of 1.36 log CFU 100 ml^{-1} between October 7 and November 10. Generally, the largest planktonic *E. coli* populations happen in the summer season with smaller populations seen in spring and fall. Planktonic *E. coli* densities at the Boulevard Lake beach were significantly different during the period of this study, with $p=0.00$. Further analysis by Tukey's Range Test showed that the *E. coli* peak density that appeared on July 21 and September 2 were significantly different from most of the background *E. coli* densities during the study period with p values ranged between 0.00-0.01.

The periphytic *E. coli* on the other hand showed a more steady increase in its population with no real spikes seen. For late spring, the periphytic *E. coli* population remained steady around 0.39 log CFU 100 cm^{-2}. In the summer from July 6 to August 18 there was an increased population of periphytic *E. coli* averaging about 2.78 log CFU 100 cm^{-2}. There is a maximum population seen on August 9 of 3.08 log CFU 100 cm^{-2}. During the summer, the elevated levels of periphytic *E. coli* were not significantly different from each other from July 6 to September 2 ($p = 0.06$). From the end of summer into the fall season, the periphytic *E. coli* population

decreased to 1.83 log CFU 100 cm^{-2} on September 2 and averaged a low of 0.99 log CFU 100 cm^{-2} from October 19 to November 10. Despite the planktonic *E. coli* peaks on July 21 and September 2, no significant increase of the periphytic *E. coli* population was observed on these two days.

The planktonic heterotrophic background bacteria population remained relatively constant from spring to the beginning of summer (Fig 2-1B). The populations averaged 5.38 log CFU 100 ml^{-1} from May 4 to July 6. From the middle to the end of summer the populations increased to an average of 5.84 log CFU 100 ml^{-1} from July 22 to September 2. During the fall season the planktonic heterotrophic bacteria decreased to 5.01 and 4.26 log CFU 100 ml^{-1} in October 7 and November 10, respectively. There was no significant differences among the planktonic heterotrophic populations with $p=0.18$.

The periphytic heterotrophic bacterial population remained similar in the spring into the middle of summer averaging 6.85 log CFU 100 cm^{-2} from May 4 to August 9 with the exclusion of July 22 which had a population of 6.03 log CFU 100 cm^{-2}. Nearing the end of summer and into the fall season the periphytic heterotrophic bacteria began to decrease to 6.12 log CFU 100 cm^{-2} on September 2, and 5.86 log CFU 100 cm^{-2} and 5.79 log CFU 100 cm^{-2} for October 7 and November 10, respectively. It should be noted that the periphytic heterotrophic counts in October and November were significantly lower than those recorded during the initial spring and summer sampling dates, barring the decrease in July. It may be possible that the temperature and precipitation could play a role in these populations as heterotrophic bacteria are influenced by the temperature (Pomeroy and Wiebe 2001). The overall F-test had a $p=0.00$. Even though there was a drop in the population on July 22, it was not significantly different from the previous two

months. The Tukey's Range Test for heterotrophic populations between July 22 and May 4 and June 2 had p values of 0.07 and 0.10 respectively.

2.3.2 Temperature, precipitation, and water

The air temperature, water temperature and precipitation were recorded from May 4, 2011 to November 10, 2011 (Fig 2-1C). The average day and night-time air temperatures in the spring were 17.5°C and 10.3°C, respectively. The day-time air temperature steadily increased into the summer averaging 25.6°C in July and 25.7°C in August. The day-time air temperature reached a maximum of 33.4°C in mid-summer and decreased from September onwards to a low of -8.8°C on November 4, 2011. The water temperature at the time of sampling on May 4, 2011 was 3.3°C and slowly rose into the summer with a maximum of 24.5°C on August 9, 2011. After August the water temperature slowly decreased and eventually reached 1.2°C on November 10, 2011. Precipitation in the spring from May 4 to June 20 had a maximum of 21 mm on May 23. The beginning of summer, on June 22 saw a large precipitation of 44.2 mm and two weeks later had another large rainfall of 50.9 mm on July 5. The month of July saw sporadic rainfalls and August saw very little rainfall. The fall months had little precipitation also but a maximum of 23.2 mm on October 14.

Dissolved organic carbon (DOC), nitrates and phosphates are known to cause eutrophication (Ryther and Dunstan 1971). Therefore, the water in Boulevard Lake was tested three times throughout the season (Table 2-1) for various amounts of chemicals and metal ions. In general, there is no recommended guideline for the concentration level of DOC in recreational water. However, a study by Abril *et al.* (2002) showed that the DOC levels of some high water

quality estuaries were between 2.5 to 6.8 mg l^{-1}. The DOC in Boulevard Lake showed no increases from spring to fall and it was actually found to decrease slightly from 8.55 to 8.30 mg l^{-1} from June 2 to August 9, respectively. There was a 28% decrease of DOC on October 7 in comparison to DOC detected in the lake water sampled on August 9. The amount of phosphates in the water at all three tested times were below the detectable limit of 0.04 mg l^{-1}. The nitrate levels were 0.06, 0.02 and 0.2 mg l^{-1} for June 2, August 9, and October 7 respectively. Therefore, both the P and N levels were far below the eutrophic level of P and N at 84.4 and 1875 mg l^{-1}, respectively (Wetzel 1983). The pH levels in the water also showed little change ranging between 6.7 and 7.3 from May 4 to November 10.

2.3.3 AllBac PCR primers

The specificity of the AllBac primers was evaluated on genomic DNA extracted from 12 different animal feces and eight non-*Bacteroides* bacterial species before applying these primers to quantify the *Bacteroides* 16S rDNA biomarker on the Boulevard Lake water samples. Genomic DNA of *Bacteroides fragilis* ATCC 25285D-5 was used as a positive control (Fig 2-2). All fecal samples, including human, goose, cow, duck, sheep, horse, rabbit, goat, chicken, pig, deer and dog, were tested positive for the AllBac primer set showing a 120 bp amplicon. This had the same molecular size as the amplicon of the *Bacteroides fragilis* ATCC 25285D-5 DNA positive control. On the other hand, *Enterococcus faecalis* ATCC 29212, *Bacillus subtilis* ATCC 6633, *Escherichia coli K-12*, *Serratia marcescens* ATCC 8100, *Pseudomonas fluorescens* ATCC 49838, *Streptococcus latics* ATCC 11454, *Staphylococcus aureus* ATCC 25923 and *Salmonella typhimurium* ATCC14028 were tested negative by the AllBac primers (Fig 2-2).

2.3.4 Cloning *Bacteroides* markers and generating the standard curve

The primer set Bac32F and Bac807R yielded an approximate 800 bp amplicon from the *Bacteroides fragilis* ATCC 25285D-5 genomic DNA. The PCR product was successfully cloned into the pGEM vector and transformed into *E. coli* JM109. The range of the copy number from 1.56×10^0 to 1.56×10^9 was generated in both ddH_2O and lake water and each standard dilution were added to the quantitative real-time PCR-mix as template. After performing quantitative real time PCR, Ct values were plotted against the logarithm of the target gene copy numbers (Fig 2-3). The standard curve of the *Bacteroides* biomarker showed strong linearity (r^2=0.996) ranging from 1.56×10^1 to 1.56×10^8 copies of the biomarker. The detection limit in ddH_2O was 15.6 copies of the target biomarker (in a 25 µl reaction mixture) since no amplification was detected in the reaction containing 1.56×10^0 copies of the biomarker. The standard curve from the dilution series in lake water was similar to the standard curve generated by the standard plasmid in sterile ddH_2O and showed strong linearity (r^2=0.997) between 1.56×10^2 and 1.56×10^8 copies of the biomarker. It indicated that copy numbers of template DNA in lake water would not be underestimated due to inhibition by lake water in that range. However the detection limit in lake water was higher and it could be due to the inhibition by the lake water.

2.3.5 Correlation of *Bacteroides* vs *E. coli*

To evaluate the possibility of AllBac primers to predict fecal contamination in aquatic environments, a correlation between the quantification of *Bacteroides* biomarkers by the real-time PCR and the number of *E. coli* in sewage, Lake Tamblyn water, well water and distilled

water was evaluated. The concentrations of the *Bacteroides* biomarker in the sewage, Lake Tamblyn water and well water samples were 7.31, 5.72 and 2.64 log copies of biomarkers 100 ml^{-1}, respectively. The *Bacteroides* biomarker concentration in ddH_2O was well below the detection limit. The *E. coli* cell densities in the sewage and the Lake Tamblyn water samples were 5.08 and 3.20 log CFU 100 ml^{-1}, respectively. The well water samples had an average *E. coli* cell density of 0.066 CFU 100 ml^{-1} and no *E. coli* was detected in the distilled water samples. A low background level of *Bacteroides* biomarkers were detected in well water although only one culturable *E. coli* was detected in one out of the three 500-ml samples. However, the correlation plot between the numbers of *Bacteroides* 16S rDNA biomarkers and the average number of *E. coli* revealed a high R^2 value of 0.956 (Fig 2-4). These results demonstrated that the AllBac primers could be applicable to detect fecal pollution in aquatic environments.

2.3.6 *Bacteroides* in Boulevard Lake

The concentrations of the *Bacteroides* biomarker were monitored in Boulevard Lake from June 4 to November 10. In late spring to early summer the target copy numbers were 3.25, 3.36 and 3.2 log copies 100 ml^{-1} for June 4, June 23 and July 6, respectively. During the summer months they increased slightly from July 22 to September 2 averaging 3.65 log copies 100 ml^{-1} for the four sampling dates in between. In the fall, the *Bacteroides* biomarker copy numbers rose slightly to 4.07, 4.52 and 4.01 log copies 100 ml^{-1} for October 7, October 20 and November 10, respectively. The overall trend for the *Bacteroides* population can be seen as a gradual increase with an overall F-test with $p=0.00$. However, the Tukey's Range Test showed that the density of

the *Bacteroides* biomarkers detected on July 22 was not significantly different from the rest of the samples ($p = 0.09$-1.00), except for October 20 ($p = 0.00$). It was also found that on September 2, the *Bacteroides* population was not significantly different from the rest of the sampling dates with $p > 0.05$.

2.4 Discussion

This study was undertaken to compare and contrast the population dynamics of the two fecal indicator bacteria, *E. coli* and *Bacteroides,* in Boulevard Lake, Thunder Bay, Ontario. *E. coli* has been the main fecal indicator bacteria used worldwide. The Canadian and U.S. Government Guidelines set maximum levels of 200 and 126 CFU 100 ml^{-1}, respectively (The Ministry of National Health and Welfare 1992; Dufour and Ballentine 1986). Boulevard Lake is a man-made lake within the Current River Watershed and the water levels and flow are controlled by a dam in the southeast end of the lake. This lake is used mainly for recreational activities which include swimming, canoeing and kayaking. The Regional Public Health Laboratory of Thunder Bay regularly tests the water from mid-June to the end of August and the Thunder Bay District Health Unit will post advisory warnings when *E. coli* levels are above the Guidelines for Canadian Recreational Water Quality.

The planktonic *E. coli* density increased with the temperature initially but it had two large peaks throughout the summer season. As *E. coli* is used as a fecal indicator, the large increases were presumed to be due to fecal contamination. In recreational water, fecal contamination has been found to come from humans, birds, wildlife, agriculture as well as sewage. For instance, the beaches of Hamilton Harbour along Lake Ontario, Canada, showed elevated levels of *E. coli*. It

was found that this was predominantly due to contamination by bird feces (Edge and Hill 2004). In recreational waters of southeastern Lake Huron, high levels of *E. coli* were attributed to agricultural runoff of manure (Kon *et al*. 2009). Beaches along Lake Michigan also were found to have elevated levels of *E. coli* resulting from gull feces (Mclellan 2004). Other Lake Michigan beaches near the Indiana coast were found to have high *E. coli* levels associated with human waste contamination (Liu *et al.* 2006). There are a few potential sources of fecal contamination in Boulevard Lake. Canada geese can regularly be seen in and around the waters during the summer. On the North-side of Boulevard Lake, there is a children's small animal farm remotely located approximately 3.5 km upstream. Also on the East-side of the Lake, there is a residential subdivision which could be a potential source of sewage contamination through storm drainage and/or leakages.

Bacteroides are found abundantly in fecal contamination (Madigan *et al*. 2012) and their ability to persist in the environment is limited by many factors including warm temperatures (Okabe and Shimazu 2007); therefore when there is fecal matter in water, a correlation between fecal coliforms and *Bacteroides* can be observed (Savichtcheva *et al*. 2007). In waters where there was no evidence of fecal contamination, *Bacteroides* could still be detected at low levels (Bower *et al*. 2005). A water quality study in the Netherlands detected *Bacteroides* in drinking water but they were environmental strains of non-fecal origin (van der Wielen and Medema 2010). Figure 2-4 also shows that low levels of *Bacteroides* could be detected in well water when there was no conclusive evidence of fecal contamination. Detection of *Bacteroides* biomarkers can be overstated because the qPCR technique can identify dead cells as well as live cells, whereas the enumeration of *E. coli* could be understated due to viable but non-culturable cells. The levels of *Bacteroides* in well water were 2.65 log copies 100 ml^{-1} whereas high fecal

polluted areas were measured at approximately 7.31 and 5.72 log copies 100 ml^{-1} for sewage and Lake Tamblyn, respectively. The concentrations of *Bacteroides* in Boulevard Lake from June 2 to November 10 ranged from 3.25 to 4.52 log copies 100 ml^{-1} which were well below the sewage and Lake Tamblyn levels. Throughout the 2011 season, there are also no correlated increases in the *Bacteroides* population during the *E. coli* spikes on July 21 and September 2. The planktonic heterotrophic bacteria remained steady throughout the spring and summer indicating no major influx of nutrients or contaminants (Fig 2-1B). Table 2-1 also showed no major increases in DOC, nitrates or phosphates throughout the season. This evidence supports the notion that fecal contamination may not be the primary source of the two *E. coli* blooms in Boulevard Lake. If the *E. coli* blooms are not coming from fecal contamination then they must be coming from another source.

To the best of our knowledge *E. coli* and *Bacteroides* have not yet been monitored and compared over an entire season. We expected to see a correlation among these two fecal indicator bacteria but this was not the case. The *E. coli* in the periphyton showed a steady increase in its population correlating with the temperature in the water (Fig 2-1C). Recent studies have found that in certain cases, *E. coli* can grow naturally in the environment (Byappanahalli *et al*. 1998; Whitman *et al*. 2003; Ishii *et al*. 2006; Ksoll *et al*. 2007). A persistent population of *E. coli* has been shown to be established in the periphyton in Boulevard Lake (Moreira *et al*. 2012). In summer, the periphytic *E. coli* density (per area) in Boulevard Lake was as high as 1.2×10^3 CFU 100 cm^{-2}. Considering that the thickness of the periphyton samples used in this study were no more than 0.5 mm, the periphytic *E. coli* cell density per volume (assuming the periphyton thickness was 0.5 mm) could be as high as 2.4×10^4 CFU 100 cm^{-3} or ml, which was 33X and 100X of the two peak planktonic *E. coli* densities on July 21 and September 2, respectively. With

the high density of periphytic *E. coli* inhabiting on the rocks and possibly in the sediments of Boulevard Lake, the bacteria will multiply and be released into the water column when favorable environmental conditions prevail. As our *Bacteroides* data not supporting any major fecal input during the two *E. coli* peaks, we propose that the periphytic *E. coli* population could be a major source of the *E. coli* blooms in Boulevard Lake during the summer of 2011. Further studies with microbial source tracking techniques and detailed environmental and physicochemical analyses will determine the validity of our hypothesis.

2.5 References

Abril, G., Nogueira, M., Etcheber, H., Cabecadas, G., Lemaire, E. and Brogueira, M.J. (2002) Behaviour of organic carbon in nine contrasting European estuaries. *Estuar Coast Shelf Sci* **54,** 241-262.

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

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R. and Sayler, G.S. (2009) Factors influencing the persistence of fecal *Bacteroides* in stream water. *J Environ Qual* **38**, 1224-1232.

Bernhard, A.E. and Field, K.G. (2000) A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Appl Environ Microbiol* **66**, 4571-4574.

Bower, P.A., Scopel, C.O., Jensen, E.T., Depas, M.M. and Mclellan, S.L. (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. *Appl Environ Microbiol* **71**, 8305-8313.

Byappanahalli, M.N. and Fujioka, R.S. (1998) Evidence that tropical soil environment can support the growth of *Escherichia coli*. *Water Sci Technol* **38**, 171-174.

Ciebin, B.W., Brodsky, M.H., Eddington, R., Horsnell, G., Choney, A., Palmateer, G., Ley, A., Joshi, R. and Ahears, G. (1995) Comparative evaluation of modified m-FC and m-TEC

media for membrane filter enumeration of *Escherichia coli* in water. *Appl Environ Microbiol* **61**, 3940-3942.

Dick, L.K. and Field, K.G. (2004) Rapid estimation of numbers of fecal *Bacteroidetes* by use of a quantitative PCR assay for 16S rRNA genes. *Appl Environ Microbiol* **70**, 5695-5697.

Dufour, A.P. and Ballentine, R.K. (1986) Ambient water quality criteria for bacteria. *U.S. Environmental Protection Agency*.

Edge, T.A. and Hill, S. (2005) Occurrence of antibiotic resistance in *Escherichia coli* from surface waters and fecal pollution sources near Hamilton, Ontario. *Can J Microbiol* **51**, 501-505.

Ishii, S., Ksoll, W.B., Hicks, R.E. and Sadowsky, M.J. (2006) Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Appl Environ Microbiol* **72**, 612-621.

Ishii, S. and Sadowsky, M.J. (2008) *Escherichia coli* in the environment: implications for water quality and human health. *Microbes Environ* **23**,101-108.

Kon, T., Weir, S.C., Howell, E.T., Lee, H. and Trevors, J.T. (2009) Repetitive element (REP)-polymerase chain reaction (PCR) analysis of *Escherichia coli* isolates from recreational waters of southeastern Lake Huron. *Can J Microbiol* **55**, 269-276.

Ksoll, W.B., Ishii, S., Sadowsky, M.J. and Hicks, R.E. (2007) Presence and sources of fecal coliform bacteria in epilithic periphyton communities of Lake Superior. *Appl Environ Microbiol* **73**, 3771-3778.

Lamendella, R., Santo Domingo, J.W., Kelty, C. and Oerther, D.B. (2008) *Bifidobacteria* in feces and environmental waters. *Appl Environ Microbiol* **74**, 575-584.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler, G. (2006) Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Appl Environ Microbiol* **72**, 4214-4224.

Liu, L., Phanikumar, M.S., Molloy, S.L., Whitman, R.L., Shively, D.A., Nevers, M.B., Schwab, D.J. and Rose, J.B. (2006) Modeling the transport and inactivation of *Escherichia coli* and *Enterococci* in the near-shore region of Lake Michigan. *Environ Sci Technol* **40**, 5022-5028.

Madigan, M.T., Martinko, J.M., Stahl, D.A. and Clark, D.P. (2012) *Brock biology of microorganisms.* pp.541. 13th ed. San Francisco: Pearson Prentice Hall.

Maier, R.M., Pepper, I.L. and Gerba, C.P. (2009) *Environmental microbiology*. pp.485-486. 2nd ed. Burlington: Elsevier.

McLellan, S.L. (2004) Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. *Appl Environ Microbiol* **70**, 4658-4665.

Moreira, S., Brown, A., Ha, R., Iserhoff, K., Yim, M., Yang, J., Liao, B., Pszczolko, E., Qin, W. and Leung, K.T. (2012) Persistence of *Escherichia coli* in freshwater periphyton: biofilm-forming capacity as a selective advantage. *FEMS Microbiol Ecol* **79**, 608-618.

Okabe, S. and Shimazu, Y. (2007) Persistence of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. *Appl Microbiol Biotechnol* **76**, 935-944.

Paerl, H.W., Dyble, J., Moisander, P.H., Noble, R.T., Piehler, M.F., Pinckley, J.L., Steppe, T.F., Twomey, L. and Valdes, L.M. (2003) Microbial indicators of aquatic ecosystem change: current applications to eutrophication studies. *FEMS Microbiol Ecol* **46**, 233-246.

Power, M.L., Littlefield-Wyer, J., Gordon, D.M., Veal, D.A. and Slade, M.B. (2005) Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. *Environ Microbiol 7*, 631-640.

Rabinovici, S.J., Bernknopf, R.L., Wein, A.M., Coursey, D.L. and Whitman, R.L. (2004) Economic and health risk trade-offs of swim closures at a Lake Michigan beach. *Environ Sci Technol* **38**, 2737-2745.

Ryther, J.H. and Dunstan, W.M. (1971) Nitrogen, phosphorus, and eutrophication in the coastal marine environment. *Science* **171**, 1008-1013.

Savichtcheva, O., Okayama, N. and Okabe, S. (2007) Relationships between *Bacteroides* 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. *Water Res* **41**, 3615-3628.

Tallon, P., Magajna, B., Lofranco, C. and Leung, K.T. (2005) Microbial indicators of faecal contamination: a current perspective. *Water Air Soil Pollut* **166**, 139-166.

The Minister of National Health and Welfare. 1992. Guidelines for Canadian recreational water quality. Canadian Government Publishing Centre Supply and Services Canada, Ottawa, Canada.

Tillett, D. and Neilan, B.A. (2000) Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J Phycol* **36**, 251-258.

Van der Wielen, P.W.J.J. and Medema, G. (2010) Unsuitability of quantitative *Bacteroides* 16S rRNA gene assays for discerning fecal contamination of drinking water. *Appl Environ Microbiol* **76**, 4876-4881.

Wetzel, R.G. (1983) *Limnology*. pp. 292-295. 2nd ed. Philadelphia: Saunders College Publishing.

Whitman, R.L., Shively, D.A., Pawlik, H., Nevers, M.B. and Byappanahalli, M.N. (2003) Occurrence of *Escherichia coli* and *Enterococci* in cladophora (chlorophyta) in nearshore water and beach sand of Lake Michigan. *Appl Environ Microbiol* **69**, 4714-4719.

Yost, C.K., Diarra, M.S. and Topp, E. (2011) Animals and humans as sources of fecal indicator bacteria. In *The Fecal Bacteria* ed. Sadowsky, M.J. and Whitman, R.L. pp. 67-91. Washington, DC: ASM Press.

Figure Legend

Figure 2-1 – A) Population dynamics of planktonic *E. coli* (●), periphytic *E. coli* (○), and *Bacteroides* (▲). 'Counts' refer to DNA copy number for the *Bacteroides* and CFUs for the *E. coli*. The planktonic *E. coli* and *Bacteroides* use 100 ml and the periphytic *E. coli* uses 100 cm^2. B) Population dynamics of planktonic (●) and periphytic (○) heterotrophic bacteria. 'Counts' refer to CFUs. Planktonic uses 100 ml and periphytic uses 100 cm^2. C) High and low temperatures refer to air temperature. High and low temperatures are represented by a broken and solid black line, respectively. Water temperature (●) refers to average water temperature during sampling. Precipitation is represented by a solid grey line.

Figure 2-2 - PCR amplifications were performed on the genomic DNA extracts of: 1, human feces; 2, goose feces; 3, cow feces; 4, duck feces; 5, sheep feces; 6, horse feces; 7, rabbit feces; 8, goat feces; 9, chicken feces; 10, pig feces; 11, deer feces; 12, dog feces; 13, *Bacteroides fragilis* ATCC 25285D-5; 14, *Enterococcus faecalis* ATCC 29212; 15, *Bacillus subtilis* ATCC 6633; 16, *Escherichia coli K-12*; 17, *Serratia marcescens* ATCC 8100; 18, *Pseudomonas fluorescens* ATCC 49838; 19, *Streptococcus latics* ATCC 11454; 20, *Staphylococcus aureus* ATCC 25923; 21, *Salmonella typhimurium* ATCC 14028; 22, ddH_2O and M, 100-bp DNA ladder marker.

Figure 2-3 - Standard curve of real-time AllBac PCR assays for quantification of general *Bacteroides* markers. AllBac assay in ddH_2O (●) with $r^2 = 0.996$ and AllBac assay in lake water (▽) with $r^2 = 0.997$.

Figure 2-4 - *E. coli* and *Bacteroides* correlation with slope = 1.17 and $r^2 = 0.911$. Tests were performed on sewage (●), Lake Tamblyn water (○), well water (▼) and ddH_2O (△).

Table 2-1 – Boulevard Lake water analysis

	Concentrations (mg l^{-1})		
Substrate	June 2, 2011	August 9, 2011	October 7, 2011
Dissolved Organic Carbon	8.55±0.35	8.3±0.28	5.95±0.07
Chloride (IC)	4.66±0	6.28±0.01	15.98±0.04
Nitrate NO_3-N [IC]	0.06±0	0.02±0	0.2±0
Sulphate (SO_4) [IC]	3.76±0.03	2.94±0.01	<DL
Total Sulfur	1.39±0	1.19±0.02	1.82±0.02
Calcium	9.59±0.04	15.39±0.01	19.6±0.37
Potassium	0.63±0.02	0.62±0.01	0.96±0.03
Magnesium	2.84±0.01	4.28±0.03	5.64±0.11
Sodium	3.01±0.01	4.32±0.01	5.68±0.11
Total Aluminum	0.18±0.02	0.02±0	0.12±0.01
Total Barium	0.02±0	0.02±0	0.03±0
Total Copper	0.01±0	0.01±0	0.01±0
Total Iron	0.37±0.02	0.26±0	0.34±0.01
Total Manganese	0.03±0	0.04±0	0.09±0.01
Total Strontium	0.02±0	0.03±0	0.04±0
Total Zinc	0.01±0	<DL	<DL

Nitrite NO_2-N (IC), Total Arsenic, Total Beryllium, Total Cadmium, Total Cobalt, Total, Chromium, Total Molybdenum, Total Nickel, Total Lead, Total Titanium, Total Vanadium, and Phosphate (PO_4-P) by IC all have a maximum detectable limit less than 0.04 mg l^{-1}.

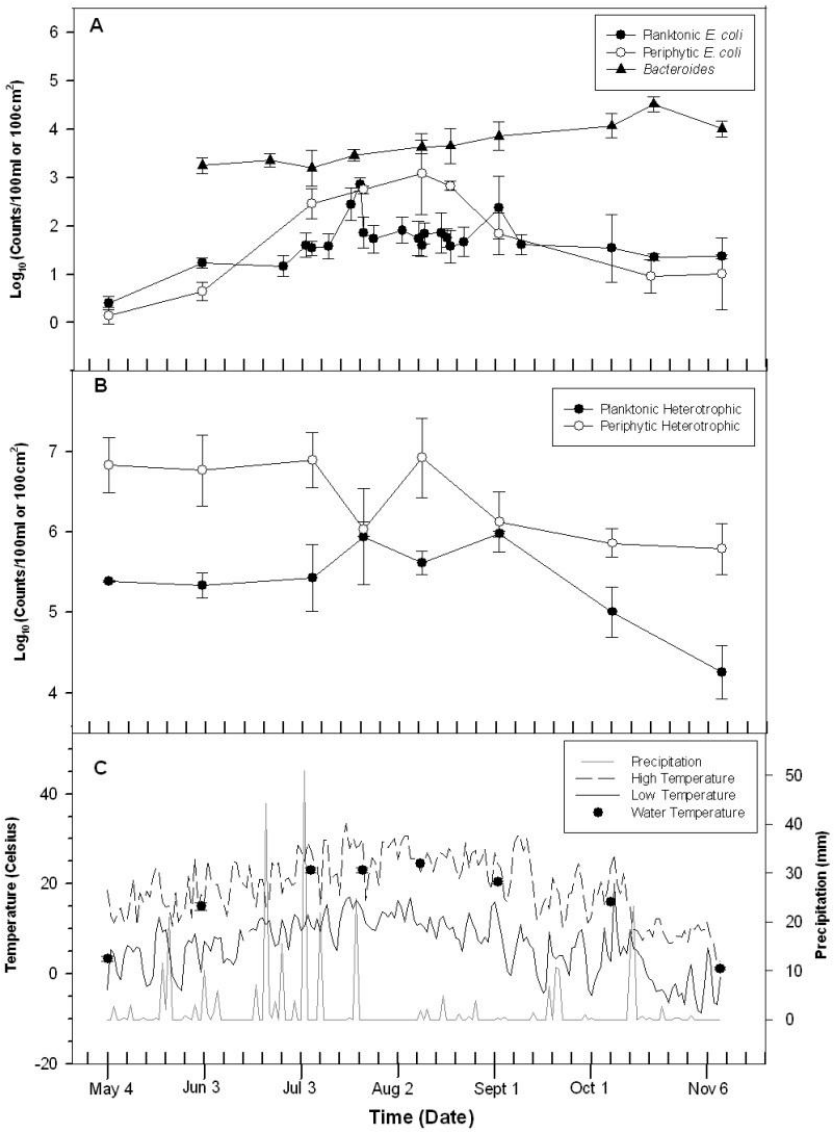
A
Planktonic E. coli
Periphytic E. coli
Bacteroides
Log10 (Counts/100ml or 100cm2)
B
Planktonic Heterotrophic
Periphytic Heterotrophic
Log10 (Counts/100ml or 100cm2)
C
Precipitation
High Temperature
Low Temperature
Water Temperature
Temperature (Celsius)
Precipitation (mm)
May 4
Jun 3
Jul 3
Aug 2
Sept 1
Oct 1
Nov 6
Time (Date)

Figure 2-1

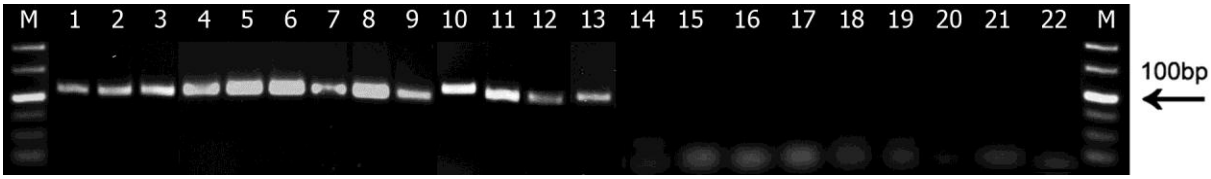

Figure 2-2

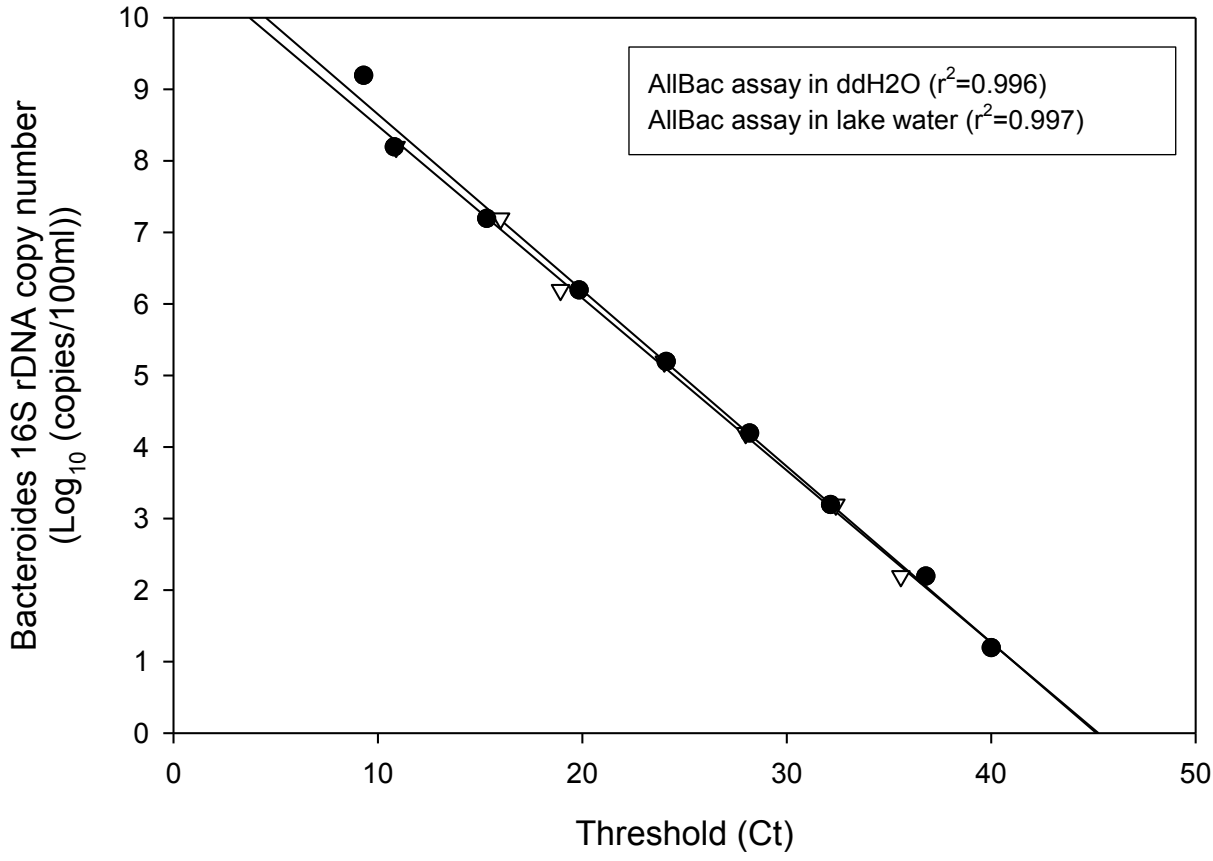
AllBac assay in ddH2O (r^2=0.996)
AllBac assay in lake water (r^2=0.997)
Bacteroides 16S rDNA copy number
(Log_{10} (copies/100ml))
Threshold (Ct)
0
1
2
3
4
5
6
7
8
9
10
0
10
20
30
40
50

Figure 2-3

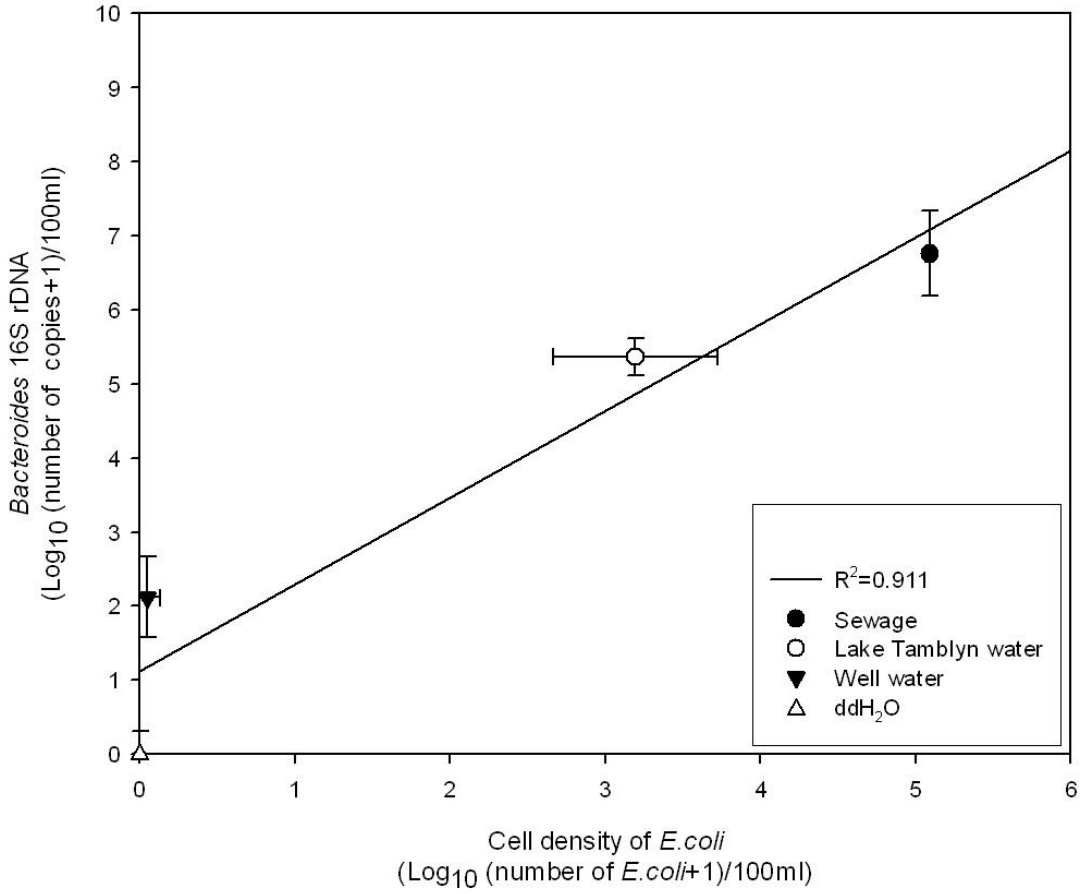
Bacteroides 16S rDNA
(Log10 (number of copies+1)/100ml)
Cell density of E.coli
(Log10 (number of E.coli+1)/100ml)
R2=0.911
Sewage
Lake Tamblyn water
Well water
ddH2O
0
1
2
3
4
5
6
7
8
9
10

Figure 2-4

Chapter 3 - Microbial Source Tracking using REP-PCR of *Escherichia coli* Blooms in a Northwestern Ontario Fresh Water Lake

3.0 Abstract

An increased level of *Escherichia coli* (> 2.30 log CFU 100 ml^{-1}) was detected at a freshwater lake multiple times over two summer seasons resulting in beach closures. A repetitive extragenic palindromic polymerase chain reaction (REP-PCR) DNA fingerprinting technique was performed on *E. coli* isolates to create a library based microbial source tracking (MST) library. The MST library consisted of farm animals (horse, sheep, pig, and chickens), human sewage, geese, and periphyton. The farm animals were located approximately 3.5 km upstream and about 200 m from the river leading into the Boulevard Lake. Any fecal pollution would be attributed to run-off due to precipitation, but no precipitation was present when high *E. coli* was found. Therefore the farm animals were not found to be a factor when using the MST techniques. Three planktonic *E. coli* peaks for June 28, 2010, August 25, 2010, and September 2, 2011, were analyzed with the MST library and found to belong primarily to the periphytic population at 57.58%, 65.79%, and 38.71%, respectively. The planktonic *E. coli* throughout the non-peak times were also mostly found to belong to the periphyton. On August 9, 2011, the geese were found to be the predominant source of *E. coli* in the water at 50%, but the total planktonic *E. coli* densities did not exceed the maximum limits recommended by the Guidelines for Canadian Recreational Water Quality. High levels of *E. coli* in the water can sometimes be mistaken for fecal contamination, therefore the use of *E. coli* as the primary fecal indicator bacteria (FIB) should be used with caution.

3.1 Introduction

Fecal contamination in recreational waters has always been a major issue. Some possible sources include human waste, wildlife, farm animals, and waterfowl. The types of potential pathogens in the water can be determined by the source of the contamination. Not all animals have the same pathogens. Therefore, finding the source of the fecal contaminant can determine how hazardous the water is. Some pathogens which pose a threat to human health include pathogenic bacteria *(Shigella, Salmonella*, and *Campylobacter*), viruses (norovirus, hepatitis A), and protozoa (*Cryptosporidium* and *Giardia*), which can come from humans and a variety of animals (Ishii and Sadowsly 2008).

Microbial source tracking (MST) is one method to determine the source of fecal contaminants. Many types of MST techniques exist today, which include library-dependent and library-independent techniques. Library-independent methods are typically quicker and can target host specific 16S rDNA markers or other specific genes for different types of bacteria, which can include *Bacteroides*, *E. coli*, *Enterococcus*, and *Bifidobacterium adolescentis* (Sadowsky and Whitman 2011). Studies for these techniques need to be validated for false-positives and false-negatives before use in the field. Use of host-specific markers also needs to be tested in the environment for sensitivity issues (Ishii and Sadowsky 2008). Library-dependent methods include phenotypic and genetic typing. Phenotypic typing typically involves antibiotic resistance and carbon utilization (Konopka *et al.* 1998; Sayah *et al.* 2005). Genetic variability can take place when the organism is exposed to antibiotics in a host or adapt to the environment outside of the animal gut. Bias can also arise due to plasmid transfers which carry multiple antibiotic resistance genes (Simpson *et al.* 2002). The genetic typing methods face less variability and thus are favoured. The most popular type of genetic typing is the rep-PCR DNA

fingerprinting technique. Rep-PCR has become popular and involves using the regions between the short intergenic repeated sequences of the genomic DNA of bacteria. Conservative primers specific to these repeated sequences can be used to create a unique pattern or fingerprint by amplifying the regions between the repeats. Isolates that are closely related were found to have similar patterns (de Bruijn 1992). Primers used to identify different repeats include REP (repetitive extragenic palindromic), BOX, and ERIC (enterobacterial repetitive intergenic consensus). The main problem with these libraries is geographical constraint; as they can only be used in the location where the samples were collected (Sadowsky and Whitman 2011).

Fecal Indicator Bacteria (FIB) are used to detect fecal contamination in water and they have very stringent criteria. FIB should (1) only grow in warm-blooded animals but not be able to replicate in the environment, (2) persist just slightly longer than the hardiest enteric pathogens in the environment and, most importantly, (3) indicate a correlation to fecal contamination and the presence of fecal pathogens (Maier *et al.* 2009). It is difficult for any species or group of FIB to satisfy all the criteria, but *E. coli* was thought to fulfill most of them. An important criterion is that the organism should not establish and grow in water (Maier *et al.* 2009). *E. coli* is a widely used fecal indicator organism, but studies of survival in the environment bring its use into doubt. *E. coli* has been found persistently in some environments including soil, sand, algae and periphyton (Byappanahalli and Fujioka 1998; Whitman *et al.* 2003; Ishii *et al.* 2006; Ksoll *et al.* 2007). In tropical regions such as Hawaii, *E. coli* were shown to be able to survive in the soil in small, but significant populations (Byappanahalli and Fujioka 1998). Alternatively, northern temperate soils from Lake Superior watersheds have also shown the presence of naturalized *E. coli* populations (Ishii *et al.* 2006). In near shore waters and off the beach sands of Lake Michigan, Whitman *et al.* (2003) found *Cladophora* mats to be a source for *E. coli*, and after

being dried for 6 months the *E. coli* was able to survive and replicate when the algae was rehydrated. As *E. coli* is able to persist in the environment there needs to be a method to differentiate these environmental strains from other sources. Rep-PCR has been used successfully in differentiating *E. coli* isolates from human, animal, as well as environmental sources (Dombek *et al.* 2000; Ishii *et al.* 2006; Kon *et al.* 2009). As naturalized *E. coli* can establish in the environment, it is important to identify them as potential sources.

Boulevard Lake, Ontario, Canada, experiences increased levels of *E. coli* (> 2.30 log CFU 100 ml^{-1}) which exceed the Canadian Recreational Water Quality Guidelines in the water multiple times throughout the summer season. These increased levels indicate fecal contamination, which can be a health hazard resulting in beach closures. We created a MST library to find the source of the increased levels of *E. coli.* A two year study involved obtaining samples from possible sources of *E. coli* in Boulevard Lake which consisted of periphyton from the rocks; human sewage; and chicken, goose, horse, sheep, and pig feces. The REP-PCR library based microbial source tracking technique was used and samples were taken from the Boulevard Lake vicinity. In our previous studies, the planktonic *E. coli* was hypothesized to belong to the periphytic *E. coli*. We will determine whether the high levels of *E. coli* are coming from fecal contamination or possibly the periphytic communities.

3.2 Materials and Methods

3.2.1 Collection of water, periphytic and fecal samples

Water and periphytic samples were taken at three locations in Boulevard Lake (48°27′34″N and 89°12’26″W) in Thunder Bay, Ontario, Canada. Site 1 and 2 were taken

adjacent to the main beach. Site 3 was taken on a rocky shore approximately 500m downstream of the main beach. Two sterile 1L Nalgene bottles were used at each site to collect the lake water approximately one foot under the surface. At each site, a rock submerged 30-40 cm below the water's surface was selected. Prior to removal from the water, the rock was gently shaken 10 times to remove loose sediment from its surface. The rock was placed on shore with the surface to be sampled facing up. A sterile, square rubber template delineating an area of 10 cm x 10 cm was placed on the rock and used to quantify the substratum surface. All material was scraped from within the area marked by the template using a 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, 0.24g KH_2PO_4, 1 L distilled H_2O, pH adjusted to 7.4). Three rocks were selected from each site for enumeration.

Goose feces were collected around the Boulevard Lake main beach area; horse, pig, sheep, and chicken fecal samples were obtained at Centennial Park Farm; and human *E. coli* isolates were isolated from sewage samples collected from the Thunder Bay Sewage Treatment and Water Pollution Control Plant. The geese were seen as a potential source as they are observed regularly throughout the summer season in and around the Lake. Centennial Park Farm is located approximately 3.5 km upstream of Boulevard Lake and about 200 m from the river. The farm consists of two horses, a dozen chickens, three pigs, and four sheep. All the samples were collected three times per year for two years. All the geese and farm animal fecal samples were collected using a sterile scoopula and sterile 50 ml centrifuge tubes (Thermo Fisher Scientific, Whitby, ON, Canada). All samples were transported to the laboratory on ice and analyzed within 12 hours.

The inputs of *E. coli* from geese, periphyton, and human sewage were considered as direct sources of contamination at Boulevard Lake. Fecal contamination from Centennial Park Farm is considered as an indirect source because the fecal runoff from the farm would need to travel to the river before reaching Boulevard Lake.

3.2.2 Enumeration and/or isolation of *E.* coli from water, periphyton and fecal samples

For enumeration of planktonic *E. coli*, 20 mL of lake water was 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* and used to compute the overall population at that time point. A total of about 40 *E. coli* isolates were isolated at each sampling date across the three sites. The isolates were maintained in Luria Bertani (LB) broth supplemented with 25% glycerol (v/v) and stored at -80 °C. A similar technique was used for the enumeration of periphytic *E. coli*, except 10 mL of the periphyton sample suspended in PBS was filtered. Four or Five presumptive *E. coli* isolates were randomly selected and isolated from each rock sample to a total of about 40 periphytic *E. coli* isolates at each sampling date.

Water and periphytic samples were taken on April 13, May 24, June 28, July 28, August 25, September 28, and November 18 for 2010; and May 4, June 2, July 6, July 22, August 9, August 18, September 2, September 9, October 7, October 19, October 20 and November 10 of 2011. Additional planktonic *E. coli* data points between May 31 and August 17 for 2010; and

June 27 and August 22 for 2011 were provided by the regional Public Health Laboratory in Thunder Bay, Ontario.

For the isolation of *E. coli* from animal feces, one gram of each fecal sample was added to 9 ml of sterile ddH_2O in a test tube for vortexing. The slurry was then serially diluted to 100-1000 times depending on the kinds of animal samples. Ten ml of the solution was taken and filtered and isolated using the aforementioned method. The same was done for the sewage samples but one ml of sewage sample was used instead of one gram. Approximately 40 *E. coli* isolates were taken from each animal at each sampling date.

3.2.3 Heterotrophic background bacteria in water samples

Heterotrophic background bacteria were enumerated from the lake water. The water samples collected in the manner previously described were serial diluted and spread-plated onto sterile R2A agar (Becton, Dickinson, and Co.). The plates were incubated for 48 h at 30 °C and the total number of colony forming units were counted and recorded.

3.2.4 Environmental conditions and physicochemical analysis of water

Immediately after the set of water samples was collected, a small volume was used to determine the pH (Fisher Scientific Accument Basic AB15, Whitby, ON, Canada) of the water at each site. Furthermore, the chemical compositions of the water samples collected on May 18, 2010, June 2, 2011, August, 9, 2011 and October 7, 2011 were determined with the aid of the Lakehead University Environmental Laboratory (LUEL). These chemical analyses were

performed immediately after each sampling and the concentrations of pertinent chemicals were averaged and recorded in mg l^{-1}: dissolved organic carbon, 7.40; K, 0.69; Mg, 3.85; Na, 3.80; Ca, 12.96; Cl^-, 7.71; NO_3^-, 0.06; Al, 0.10; Ba, 0.02; Cu, 0.01; Fe, 0.31; Mn, 0.04; S, 1.39; SO_4^{2-}, 3.89; Sr, 0.03; and Zn, 0.01. As, Be, Cd, Cr, Ni, NO_2^-, Pb and PO_4^{3-}-P were < 0.04 mg/L.

The analysis was performed by the LUEL using standard procedures from the 'Standard Methods for the Examination of Water and Wastewater' 21st Edition.

The data for the precipitation and the daily high and low temperatures were obtained from the National Climate Data and Information Archive (www.climate.weatheroffice.gc.ca).

3.2.5 DNA extraction

Individual *E. coli* strains were streaked onto Luria Bertani agar from frozen stock and incubated at 37 °C overnight. Individual colonies were selected and subsequently grown in Luria Bertani broth overnight at 37 °C with shaking at 150 rpm. After overnight growth, 1 ml of each cell culture was centrifuged for 3 min at 21 000 x g to pellet the bacteria. The supernatant was removed, and 800 µl of XS lysis buffer (1% w/v potassium ethyl xanthogenate, 100 mol l^{-1} Tris–HCl, 20 mmol l^{-1} EDTA, 1% w/v SDS, 800 mmol l^{-1} ammonium acetate) and 1 µl of 5 µg $µl^{-1}$ RNase (Promega,Madison, WI) were added to each sample (adapted from Tillett and Neilan, 2000). The samples were subsequently re-suspended and incubated for 1 h at 37 °C. Following incubation, the samples were placed in a 70 °C water bath for 15 min, mixed, and placed on ice for 30 min to precipitate cell debris. The cell debris was removed by centrifuging for 10 min at 21 000 x g. The supernatant was transferred to a new, sterile tube. The DNA in the supernatant was precipitated by adding 750 µl of isopropyl alcohol and then incubating at room temperature

for 10 min with regular inversions. The precipitated DNA was recovered by centrifuging at 21 000 x g for 10 min and removing the supernatant. The DNA was washed with 750 µl of 70% ethanol and stored in 100 µl of sterile ddH_2O at 4°C.

3.2.6 REP-PCR

REP primers designed by de Bruijn (1992) were used to target REP sequences within the bacterial genome. The sequences between the REP elements were amplified through PCR and separated using gel electrophoresis, generating unique banding patterns that could be used to compare the genetic diversity of the *E. coli* isolates. Each polymerase chain reaction had a final volume of 25 µl containing 2.5 mmol l^{-1} $MgCl_2$, 1x Taq buffer (Fermentas, Burlington, ON, Canada), 0.2 mmol l^{-1} of each deoxynucleoside triphosphates, 1 µmol l^{-1} Primer REP-2I (5′;-ICGICTTATCIGGCCTAC-3′), 1 µmol l^{-1} Primer REP-IR (5′-IIIICGICGICATCIGGG-3′), 1.25 units Taq polymerase, and 1 µl (Approximately 200 ng) of genomic DNA from extraction. The PCR parameters include an initial denaturation step of 6 min at 95 °C followed by 30 cycles of denaturation (1 min at 94 °C), annealing (1 min at 40 °C), and extension (8 min at 65 °C), with a final extension step of 16 min at 65 °C. The final PCR products were held at 4 °C (de Bruijn, 1992). After amplification, the amplified DNA fragments were separated using agarose gel electrophoresis (2% w/v agarose, 1x TAE buffer and 1 µg ml^{-1} ethidium bromide) for visualization under UV light. Five µl of 6x Fermentas loading dye was added to each reaction tube, giving a final volume of 30 µl. The first and last wells of the gel were loaded with 6 µl of 0.1 µg $µl^{-1}$ 1 kb plus DNA ladder (Fermentas). The remaining wells were loaded with 10 µl of sample. The PCR samples were separated by gel electrophoresis at 100 V for 60 min. Upon finishing, the bands were visualized using a Chemi Genius Bio-imaging system (Syngene,

Frederick, MD). An image of the gel was saved and imported into the Fingerprinting II Informatix software (Bio-Rad Laboratories Inc., Hercules, CA) for analysis.

3.2.7 Computer assisted REP-PCR analysis

The gel images were normalized using the Fingerprinting II Informatix software with respect to the 1 kb plus DNA ladder (Fermentas) and then converted to binary codes. A discriminant analysis was performed (IBM SPSS Statistics Version 19) to differentiate the periphyton, chicken, goose, horse, human, sheep, and pig *E. coli* sources and also to match up the planktonic *E. coli* to the sources. A one-way ANOVA analysis was performed to compare the means between the *E. coli* densities over the two seasons and to compare the means among the planktonic heterotrophic densities. The F-test was used for multiple comparisons among the samples and the Tukey's Multiple Range Test was used to compare the difference between individual samples. IBM SPSS Statistics Version 19 was used for the analysis.

3.3 Results

3.3.1 *E. coli* and heterotrophic bacteria population dynamics

Over the two seasons (April 13 to November 18 2010 and May 4 to November 10, 2011) the average planktonic *E. coli* populations in Boulevard Lake were 1.52 and 1.41 log CFU 100 ml^{-1} for 2010 and 2011, respectively, barring the spikes in the population (Fig 3-1A). In 2010, the *E. coli* populations had two peaks reaching 2.62 and 2.45 log CFU 100 ml^{-1} on June 28 and August 16, respectively. In 2011, a similar trend was seen with two recorded *E. coli* peaks of

2.86 and 2.38 log CFU 100 ml^{-1} on July 21 and September 2, respectively. Unfortunately, no *E. coli* isolates were taken for the MST study on July 21, 2011.

The periphytic *E. coli* also showed similar trends for 2010 and 2011. The first recorded sample for 2010 was on April 13 at 2.26 log CFU 100 cm^{-2}. The population slowly increased into the summer and plateaued between June 28, 2010 and August 25, 2010 at an average of 3.16 log CFU 100 cm^{-2} and slowly decreased afterwards. In 2011, a plateau was seen between July 22, 2011 and August 18, 2011 averaging 2.89 log CFU 100 cm^{-2}.

The planktonic heterotrophic bacteria generally remained steady throughout the 2010 and 2011 season (Fig 3-1A). Average cell densities for 2010 and 2011 were 5.25 and 5.71 log CFU 100 ml^{-1}, respectively.

3.3.2 Temperature, precipitation and water

During the *E. coli* blooms, precipitation was not seen three days before and at the sampling date except for June 28, 2010, which had a maximum level of 16.9 mm seen on June 27, 2010. The rest of the precipitation in 2010 and 2011 occurred when the *E. coli* blooms were not present (Fig 3-1B).

The temperatures in the air and water slowly increased into the spring, peaking in the summer, and decreased into the fall for 2010 and 2011 (Fig 3-1B). For the 2010 and 2011 season, the periphytic *E. coli* populations generally follow the mean water temperatures, increasing into the summer and decreasing into the fall season (Fig 3-1). The planktonic *E. coli*

populations (Fig 3-1A) typically rose into the summer and dropped in the fall season. Throughout the summer there were sporadic increases in the planktonic *E. coli* populations.

3.3.3 Host grouping

Representatives of The REP-PCR DNA profiles of sheep, chicken, pig, horse, goose, human, and periphytic *E. coli* isolates are shown in Figure 3-2. The size of the PCR products ranged between 300 to 8000 bp. The REP-PCR banding patterns of *E. coli* can generally be seen to be quite similar within host and different between hosts. The discriminant function analysis for the host sources were performed to determine the level of host groupings and can be seen in Figures 3-3 and 3-4. When the three direct host sources (periphyton, goose, human) were considered, the isolates grouped together very clearly. When four additional sources (chicken, horse, sheep, pig) were added, more overlaps were observed between the host groups and the grouping was less clear.

Table 3-1 and 3-2 show the isolates being grouped with three and seven host sources, respectively. In Table 3-1 more than 92 % of the periphyton, goose and human *E. coli* isolates are correctly classified into their host groups respectively. Table 3-2 shows the sheep, chicken, pig, horse, goose, human, and periphytic *E. coli* isolates are correctly classified between 67.1 and 81.5 %.

3.3.4 Source tracking of planktonic *E. coli* samples

The planktonic *E. coli* samples obtained from the lake water were analyzed by a discriminant function analysis (IBM SPSS 19) monthly from April 13, 2010 to November 18, 2010 and May 4, 2011 to November 10, 2011. To identify the source of the planktonic *E. coli* isolates, they were compared with *E. coli* isolated from the three direct sources (periphyton, goose and human) and all seven sources (periphyton, goose, human, chicken, horse, sheep, pig).

When the planktonic *E. coli* isolates were compared with the seven sources at the peak periods on June 28, 2010 and August 25, 2010, the majority were being classified into the periphytic *E. coli* group at 48.5 and 39.5 %, respectively (Table 3-3). Source Typing analyses were performed on samples collected on August 25, 2010 because *E. coli* was not isolated at the peak period of the *E. coli* bloom on August 16, 2010. Planktonic *E. coli* collected on the peak period of September 2, 2011 had a mixture of three main sources, which were periphyton, chicken, and goose at 25.8 %, 19.4 %, and 29.0 %, respectively. Even for the non-peak dates in 2010, approximately 22.7 to 60.0 % of the planktonic *E. coli* was still classified into the periphytic group. The rest of the isolates for 2010, in the non-peak period fell between 0 to 15 % for chicken, goose, horse, human, sheep and pig. Similar trends were observed in 2011 with 25.0 to 53.9 % being classified into the periphytic group. The exception was observed on August 9, 2011 where 40 % of the planktonic *E. coli* isolates were classified into the geese host group. When the planktonic *E. coli* isolates were compared with the three direct sources (periphyton, goose, and human) at the peak periods on June 28, 2010 and August 25, 2010, 57.6 and 65. 8 % of the isolates were classified as periphytic *E. coli*, respectively. In 2010, the non-peak planktonic *E. coli* isolates were also found to predominantly belong to the periphyton, ranging from 50.0 to 80.0%. In 2011, during the *E. coli* bloom on September 2, the *E. coli* isolates were more spread out among the three sources, with the majority still belonging to periphyton at

38.7%, and goose and human at 35.5 and 25.8%, respectively. The majority of the background planktonic *E. coli* isolates in 2011 mainly belonged to the periphytic group except for August 9, which had 50% of the *E. coli* isolates belonging to the geese. However, the density of *E. coli* in the water on August 9, 2011 was 1.59 log CFU 100 ml^{-1}, which was well below the safety level of the Canadian Recreational Water Quality Guidelines for *E. coli* (2.30 log CFU 100 ml^{-1}).

In Figure 3-1A, the Thunder Bay Regional Public Health Laboratory provided some of the planktonic *E. coli* data for the 2010 and 2011 seasons. When our data were taken on similar times as the regional Public Health Laboratory, the numbers were found to be consistently similar.

3.4 Discussion

Throughout the 2010 and 2011 seasons at Boulevard Lake, the water was observed to have increased levels of *E. coli* (> 2.30 log CFU 100 ml^{-1}) at various times throughout the summer season resulting in beach closures. Three resulting *E. coli* peaks were detected on June 28, 2010, August 25, 2010, and September 2, 2011 (Figure 3-1A). A fourth peak was recorded by the regional Public Health Laboratory on July 21, 2011 at 2.86 log CFU 100 ml^{-1}, but isolates were not obtained at this time-point for the MST study. Since *E. coli* is used as a fecal indicator bacterium, the resulting increases are perceived to be coming from fecal pollution. There are many sources of fecal pollution which include human sewage, farm animals, wildlife, and waterfowl (Jiang *et al.* 2007). Recent studies at the Bayfront Park beach in Hamilton, Ontario, found that bird feces was the prominent source of *E. coli* over the other sources (Edge and Hill 2005; Edge and Hill 2007). Another study enumerated fecal coliforms from ring-billed gulls and

Canada geese and found that the levels were significant enough to cause an impact on the water quality. Hundreds of fecal samples were collected from the two types of birds and the amount of fecal coliform bacteria was calculated. The impact the birds could have on the water was based on the average number of fecal coliforms per gram in the feces. These numbers were large enough to have a potential effect on the water quality, but this is relative to the number and type of birds, the time of day they roost, as well as the defecation rates (Alderisio and DeLuca 1999). This leads to our studies over the increased levels of *E. coli* at Boulevard Lake. The public attributes the geese population to be the major source of contamination as they can be found in and around the lake throughout the summer season. Our study will use microbial source tracking (MST) techniques to determine the source of the *E. coli* in the water.

It is generally believed that *E. coli* grows exclusively in warm-blooded animals and is not able to replicate in the environment (Tallon *et al.* 2005). Recent studies have shown that *E. coli* have been able to survive and multiply outside of the host organism. They have been shown to be able to persist in periphyton, soil, interstitial waters, sand, as well as algae (Byappanahalli and Fujioka 1998; Whitman *et al.* 2003; Ishii *et al.* 2006; Kon *et al.* 2007; Ksoll *et al.* 2007). Byappanahalli and Fujioka (1998) found that tropical soil can supply the proper nutrients for *E. coli* to grow in the environment. On the beach sands of Lake Michigan, Whitman et al (2003) was able to identify *Cladophora* as a source for *E. coli. E. coli* was able to survive on the sun-dried *Cladophora* mats over 6 months at 4°C and readily grew up on rehydration. Also, off the shores of Duluth, Minnesota, Ksoll *et al.* (2007) discovered *E. coli* to colonize periphyton and they were also likely to persist over the winter. Goto and Yan (2011) looked at the genotypic diversity of *E. coli* in the stream waters of a tropical watershed and concluded that the high spatial variations indicated non-point source. The soil also exhibited high temporal variations of

E. coli genotypes, which suggest a dynamic *E. coli* population. The use of *E. coli* as a fecal indicator bacterium in a tropical watershed should be used with caution and there is a need to differentiate fecal and environmental *E. coli* when monitoring water (Goto and Yan 2011). All these examples of environmental *E. coli* populations persisting in the environment can lead to doubts of using *E. coli* as a fecal indicator bacterium. To resolve this problem, the environmental strains of *E. coli* need to be classified as a separate source in our MST library. The source that will be used has to have some specific ecological niche that can allow the establishment of distinct populations of *E. coli*. Our study will be using the periphytic *E. coli* as the environmental source because persistent populations of *E. coli* have been identified in the periphyton samples collected from Boulevard Lake and other freshwater water bodies in northwestern Ontario (Moreira *et al*. 2012). Over the course of this two-year study, samples were obtained from the periphyton, lake water, human sewage, and fecal samples from geese, chicken, sheep, horse, and pig to use as our MST library.

A study by Moreira (2010) showed that the *E.* coli strains did not survive well in water samples obtained from Boulevard Lake. When nutrients were limited in the water, populations of *E. coli* were found to decrease by more than 100 times in four days. The average velocity of the water moving downstream at the sampling sites was approximately 0.075 m s^{-1}. Due to the water having a flow and the *E. coli* being unable to replicate in the water, the planktonic *E. coli* is presumed to be a transient population. The planktonic *E. coli* population in Boulevard Lake did not have a stable population and large increases and decreases occur during the summer seasons (Fig 3-1A). Therefore, the planktonic *E. coli* are likely to be a result from inputs of other *E. coli* sources. On the other hand, the periphytic *E. coli* showed a steady rise in its population, slowly increasing from spring to summer, plateauing in the summer, and then decreasing into the fall

season (Fig 3-1A). The *E. coli* in the periphytic communities can survive with limited nutrients and these periphytic communities have been found to persist in the environment (Ksoll *et al.* 2007; Moreira *et al.* 2012).

For Boulevard Lake, there are three possible direct sources of *E. coli*, which are geese, human, and periphyton. There is a residential subdivision on the east-side of the lake which could be a source of human *E. coli* if there is a sewage leak. For the MST library, when geese, human and periphyton were used, our discriminant analysis was able to differentiate and distinguish them with little overlap (Fig 3-3). In Table 3-1, > 92 % of the *E. coli* isolates are correctly classified into their respective host groups. When analyzed with the seven host sources (Table 3-2) there was more overlap seen and the periphyton, geese, and human *E. coli* sources experienced a decrease in its classification. This is probably due to a limited degree of diversity within the *E. coli* species.

Parker *et al.* (2010) found that during storms, runoff was able to increase the amount of fecal indicator bacteria in the water from fecal contaminants being washed out. Centennial Park Farm houses some farm animals and is approximately 3.5 km upstream of Boulevard Lake and 200 m from the river. It is possible that any runoff from this farm can contribute to fecal contamination in the river, and hence into the lake. Since the farm is more than 200 m from the river, any runoff would have to come from precipitation. Over the two seasons there are four observed *E. coli* density spikes on June 28, 2010, August 16, 2010, July 21, 2011 and September 2, 2011. During the four *E. coli* density spikes and 3 days beforehand, very little or no precipitation was observed. The small amount of precipitation seen during the peak *E. coli* densities in the water would lead to the idea that the farm is not directly impacting Boulevard Lake. However, the MST data in Figure 3-3 shows that the farm animals contribute between 32.3

to 44.7 % of the total *E. coli* isolates in the water during the *E. coli* peaks. Since there was little or no rainfall during the *E. coli* peaks it is most likely that the high numbers of planktonic *E. coli* classified under the farm animal categories are artifacts (Table 3-3), which come from overlaps between the isolates from the farm animals and the direct sources (Figure 3-4). Centennial Park Farm also has a limited number of animals (two horses, a dozen chickens, three pigs, and four sheep), which would correspond to less runoff of fecal matter into the river system. It is likely that the farm animals do not have a direct impact on Boulevard Lake and are erroneously used; therefore the three direct sources of periphyton, goose, and human *E. coli* will be used as the *E. coli* source library.

Table 3-4 shows the monthly comparisons of the *E. coli* groupings with the three main sources of human, geese and periphyton. It verifies that most of the planktonic *E. coli* isolates belong to the periphyton, which account for 50.0 to 80.0 % in 2010, and 35.0 to 72.2 % in 2011. The only exception was on August 9, 2011, where 50.0% of the planktonic *E. coli* populations belong to the geese. Despite 50.0% of the *E. coli* isolates belonging to geese on August 9, 2011, the planktonic *E. coli* population was only 1.59 log CFU 100 ml^{-1}, which was well below the Canadian Recreational Water Quality Guidelines of 2.30 log CFU 100 ml^{-1}. From this data, it is clear that the fecal input from the geese was not a major contributing factor of the *E. coli* blooms seen in Boulevard Lake throughout the summer seasons for 2010 and 2011. The majority of the *E. coli* isolates from the two seasons at Boulevard Lake predominantly belong to the periphytic group.

Potential causes for the increase of the periphytic *E. coli* to be released at Boulevard Lake could be due to many physical characteristics. A decrease of nutrients can start a cue for the biofilms to begin to release themselves (Hunt et al 2004). The periphyton can also be affected by

increases in flow rate and increased cell growth can lead to shedding of daughter cells (O'Toole *et al*. 2000). Under laminar flow conditions the bacteria tend to have lower tensile strength and can easily be washed into the water (Donland and Costerton 2002). At Boulevard Lake the flow is typically a low, laminar flow, which causes the periphytic *E. coli* to have low tensile strength and allows them to break off more easily when flow rates increase. We suggest that the large increase of planktonic *E. coli* may be attributed to the periphyton being released into the water and not to fecal contamination. Future studies on water quality should include MST studies specifically for the *E. coli* blooms to determine the direct cause of the increased level of bacteria.

3.5 References

Alderisio, K.A. and DeLuca, N. (1999) Seasonal enumeration of fecal coliform bacteria from the feces of ring-billed gulls (*Larus delawarensis*) and Canada Geese (*Branta Canadensis*). *Appl Environ Microbiol* **65,** 5628-5630.

APHA (American Public Health Association). (2005) *Standard methods for the examination of water and wastewater*. 21st edition. American Public Health Association, Washington , D.C., USA.

Bergey, E.A. (2008) Does rock chemistry affect periphyton accrual in streams?. *Hydrobiologia* **614,** 141-150.

Byappanahalli, M.N. and Fujioka, R.S. (1998) Evidence that tropical soil environment can support the growth of *Escherichia coli*. *Water Sci Technol* **38**, 171-174.

Carson C.A., Shear, B.L., Ellersieck, M.R. and Schnell, J.D. (2003) Comparison of ribotyping and repetitive extragenic palindromic-PCR for identification of fecal *Escherichia coli* from humans and animals. *Appl Environ Microbiol* **69,** 1836-1839.

de Bruijn, F.J. (1992) Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl Environ Microbiol* **58,** 2180-2187.

Dombek, P.E., Johnson, L.K., Zimmerley, S.T. and Sadowsky M.J. (2000) Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. *Appl Environ Microbiol* **66,** 2572-2577.

Donlan, R.M and Costerton, J.W. (2000) Biofilms: survival mechanisms and clinically relevant microorganisms. *Clin Microbiol Rev* **15**, 167-193.

Edge, T.A. and Hill, S. (2005) Occurrence of antibiotic resistance in *Escherichia coli* from surface waters and fecal pollution sources near Hamilton, Ontario. *Can J Microbiol* **51**, 501-505.

Edge, T.A. and Hill, S. (2007) Multiple lines of evidence to identify the sources of fecal pollution at a freshwater beach in Hamilton Harbour, Lake Ontario. *Water Res* **41**, 3585-3594.

Goto, D.K. and Yan, T. (2011) Genotypic diversity of *Escherichia coli* in the water and soil of tropical watersheds in Hawaii. *Appl Environ Microbiol* **77**, 3988-3997.

Hunt, S.M., Werner, E.M., Huang, B., Hamilton M.A. and Stewart, P.S. (2004) Hypothesis for the role of nutrient starvation in biofilm detachment. *Appl Environ Microbiol* **70**, 7418-7425.

Ishii, S., Ksoll, W.B., Hicks, R.E. and Sadowsky, M.J. (2006) Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. *Appl Environ Microbiol* **72**, 612-621.

Ishii, S. and Sadowsky, M.J. (2008) *Escherichia coli* in the environment: implications for water quality and human health. *Microbes Environ* **23**,101-108.

Jiang, S.C., Chu, W., Olson, B.H., He, J., Choi, S., Zhang, J., Le, J.Y. and Gedalanga, P.B. (2007) Microbial source tracking in a small southern California urban watershed indicates wild animals and growth as the source of fecal bacteria. *Appl Microbiol Biotechnol* **76**, 927-934.

Kon, T., Weir, S.C., Howell, E.T., Lee, H. and Trevors, J.T. (2009) Repetitive element (REP)-polymerase chain reaction (PCR) analysis of *Escherichia coli* isolates from recreational waters of southeastern Lake Huron. *Can J Microbiol* **55**, 269-276.

Konopka, A., Oliver, L., Turco Jr, R.F. (1998) The use of carbon substrate utilization patterns in environmental and ecological microbiology. *Microbial Ecol* **35**, 103-115.

Ksoll, W.B., Ishii, S., Sadowsky, M.J. and Hicks, R.E. (2007) Presence and sources of fecal coliform bacteria in epilithic periphyton communities of Lake Superior. *Appl Environ Microbiol* **73**, 3771-3778.

Maier, R.M., Pepper, I.L. and Gerba, C.P. (2009) *Environmental microbiology*. pp.485-486. 2nd ed. Burlington: Elsevier.

Moreira, S. 2010. Biofilm formation of environmental and enterohemorrhagic *Escherichia coli*. (2010) Masters Thesis, Lakehead University.

Moreira, S., Brown, A., Ha, R., Iserhoff, K., Yim, M., Yang, J., Liao, B., Pszczolko, E., Qin, W. and Leung, K.T. (2012) Persistence of *Escherichia coli* in freshwater periphyton: biofilm-forming capacity as a selective advantage. *FEMS Microbiol Ecol* **79**, 608-618.

O'Toole, G., Kaplan, H.B. and Kolter, R. (2000) Biofilm formation as microbial development. *Annu Rev Microbiol* **54**, 49-79.

Parker, J.K, McIntyre, D. and Noble, R.T. (2010) Characterizing fecal contamination in stormwater runoff in coastal North Carolina, USA. *Water Res* **44**, 4186-4194.

Sadowsky, M.J. and Whitman, R.L. (2011) *The fecal bacteria*. pp. 32-33, 79-80. Washington, DC: ASM Press.

Sayah, R.S., Kaneene, J.B., Johnson, Y. and Miller R. (2005) Patterns of antimicrobial resistance observed in *Escherichia coli* isolates obtained from domestic- and wild-animal fecal samples, human septage, and surface water. *Appl Environ Microbiol* **71**, 1394-1404.

Simpson, J.M., Santo Domingo, J.W. and Reasoner, D.J. (2002) Microbial source tracking: state of the science. *Environ Sci Technol* **36**, 5279-5288.

Tallon, P., Magajna, B., Lofranco, C. and Leung, K.T. (2005) Microbial indicators of faecal contamination: a current perspective. *Water Air Soil Pollut* **166**, 139-166.

The Minister of National Health and Welfare. 1992. Guidelines for Canadian recreational water quality. Canadian Government Publishing Centre Supply and Services Canada, Ottawa, Canada.

Tillett, D. and Neilan, B.A. (2000) Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J Phycol* **36**, 251-258.

Whitman, R.L., Shively, D.A., Pawlik, H., Nevers, M.B. and Byappanahalli, M.N. (2003) Occurrence of *Escherichia coli* and *Enterococci* in cladophora (chlorophyta) in nearshore water and beach sand of Lake Michigan. *Appl Environ Microbiol* **69**, 4714-4719.

Figure Legend

Figure 3-1 A) Population dynamics of planktonic *E. coli* (●), periphytic *E. coli* (○), planktonic heterotrophic bacteria (▲) and the Health Unit Planktonic *E. coli* (★). 'Counts' refer to CFUs for the *E. coli* and heterotrophic bacteria. The planktonic *E. coli*, planktonic heterotrophic bacteria and Health Unit planktonic *E. coli* use 100 ml and the periphytic *E. coli* uses 100 cm^2. B) High and low temperatures refer to air temperature. High and low temperatures are represented by a broken and solid black line, respectively. Water temperature (●) refers to average water temperature during sampling. Precipitation is represented by a solid grey line.

Figure 3-2 - REP-PCR of *E.* coli isolates of various animals. Lanes A and W contain an external 1kb Plus DNA ladder. Lanes B, C, and D contain sheep. Lanes E, F, and G contain chicken. Lanes H, I, and J contain pig. Lanes K, L, and M contain horse. Lanes N, O, and P contain goose. Lanes Q, R, and S contain human. Lanes T, U, and V contain periphyton.

Figure 3-3 – Discriminant function analysis of *E. coli* isolates belonging to the periphyton, goose and human samples.

Figure 3-4 - Discriminant function analysis of *E. coli* isolates belonging to the periphyton, goose and human, chicken, horse, sheep, and pig samples.

Table 3-1 – Discriminant function analysis of *E. coli* isolates belonging to the periphyton, goose and human sources

		Predicted Group Membership			Total
		Periphyton	Goose	Human	
% and (Count)	Periphyton	92.3(334)	5.5(20)	2.2(8)	100.0(362)
	Goose	4.3(6)	94.3(133)	1.4(2)	100.0(141)
	Human	5.0(6)	2.5(3)	92.4(110)	100.0(119)

Table 3-2 – Discriminant function analysis of *E. coli* isolates belonging to the periphyton, goose, human, chicken, horse, sheep and pig sources

		Predicted Group Membership							Total
		Periphyton	Chicken	Goose	Horse	Human	Sheep	Pig	
% and (Count)	Periphyton	67.1(243)	5.2(9)	7.2(26)	5.8(21)	5.2(19)	6.1(22)	3.3(12)	100.0(362)
	Chicken	6.9(10)	80.6(116)	1.4(2)	3.5(5)	.7(1)	4.2(6)	2.8(4)	100.0(144)
	Goose	7.1(10)	2.1(3)	78.7(111)	2.8(4)	2.8(4)	2.1(3)	4.3(6)	100.0(141)
	Horse	8.3(12)	6.9(10)	2.8(4)	72.4(105)	1.4(2)	6.9(10)	1.4(2)	100.0(145)
	Human	8.4(10)	1.7(2)	2.5(3)	3.4(4)	81.5(97)	.8(1)	1.7(2)	100.0(119)
	Sheep	7.0(9)	4.7(6)	.0(0)	6.3(8)	.0(0)	81.3(104)	.8(1)	100.0(128)
	Pig	8.8(11)	4.8(6)	2.4(3)	2.4(3)	.8(1)	4.8(6)	76.0(95)	100.0(125)

Table 3-3 – Discriminant Function Analysis of monthly planktonic *E. coli* isolates compared with the 7 Sources (periphyton, goose, human, chicken, horse, sheep and pig sources)

		Predicted Group Membership							
		Periphyton	Chicken	Goose	Horse	Human	Sheep	Pig	Total
% & (Count)	April 13 2010	42.86(3)	14.29(1)	14.29(1)	14.29(1)	0.00(0)	14.29(1)	0.00(0)	100(7)
	May 24 2010	22.73(5)	9.09(2)	4.55(1)	9.09(2)	31.82(7)	13.64(3)	9.09(2)	100(22)
	June 28 2010	**48.48(16)**	**0.00(0)**	**12.12(4)**	**18.18(6)**	**6.06(2)**	**9.09(3)**	**6.06(2)**	**100(33)**
	July 28 2010	51.43(18)	2.86(1)	14.29(5)	11.43(4)	0.00(0)	20.00(7)	0.00(0)	100(35)
	August 25 2010	**39.47(15)**	**5.26(2)**	**10.53(4)**	**15.79(6)**	**5.26(2)**	**10.53(4)**	**13.16(5)**	**100(38)**
	September 28 2010	60.00(12)	0.00(0)	10.00(2)	10.00(2)	15.00(3)	0.00(0)	5.00(1)	100(20)
	November 18 2010	54.55(6)	0.00(0)	9.09(1)	18.18(2)	9.09(1)	0.00(0)	9.09(1)	100(11)
	May 4 2011	53.85(7)	0.00(0)	7.69(1)	7.69(1)	15.38(2)	15.38(2)	0.00(0)	100(13)
	June 2 2011	27.03(10)	8.11(3)	10.81(4)	10.81(4)	13.51(5)	18.92(7)	10.81(4)	100(37)
	July 6 2011	30.56(11)	8.33(3)	11.11(4)	19.44(7)	13.89(5)	8.33(3)	8.33(3)	100(36)
	August 9 2011	25.00(10)	15.00(6)	40.00(16)	2.5(1)	10.00(4)	2.5(1)	5.00(2)	100(40)
	September 2 2011	**25.81(8)**	**19.35(6)**	**29.03(9)**	**3.23(1)**	**12.90(4)**	**3.23(1)**	**6.45(2)**	**100(31)**
	October 7 2011	25.00(5)	10.00(2)	10.00(2)	10.00(2)	15.00(3)	20.00(4)	10.00(2)	100(20)
	November 10 2011	33.33(5)	0.00(0)	6.67(1)	20.00(3)	13.33(2)	13.33(2)	13.33(2)	100(15)

Table 3-4 – Discriminant Function Analysis of monthly planktonic *E. coli* isolates compared with the 3 Sources (periphyton, goose and human)

		Predicted Group Membership			
		Periphyton	Goose	Human	Total
% & (Count)	April 13 2010	57.14(4)	28.57(2)	14.29(1)	100(7)
	May 24 2010	50.00(11)	13.64(3)	36.36(8)	100 (22)
	June 28 2010	**57.58(19)**	**15.15(5)**	**27.27(9)**	**100 (33)**
	July 28 2010	62.86(22)	20.00(7)	17.14(6)	100 (35)
	August 25 2010	**65.79(25)**	**21.05(8)**	**13.16(5)**	**100 (38)**
	September 28 2010	80.00(16)	10.00(2)	10.00(2)	100 (20)
	November 18 2010	72.73(8)	18.18(2)	9.09(1)	100 (11)
	May 4 2011	53.85(7)	30.77(4)	15.38(2)	100(13)
	June 2 2011	62.16(23)	24.32(9)	13.51(5)	100(37)
	July 6 2011	72.22(26)	13.89(5)	13.89(5)	100(36)
	August 9 2011	35.00(14)	50.00(20)	15.00(6)	100(40)
	September 2 2011	**38.71(12)**	**35.48(11)**	**25.81(8)**	**100(31)**
	October 7 2011	40.00(8)	20.00(4)	40.00(8)	100(20)
	November 10 2011	66.67(10)	6.67(1)	26.67(4)	100(15)

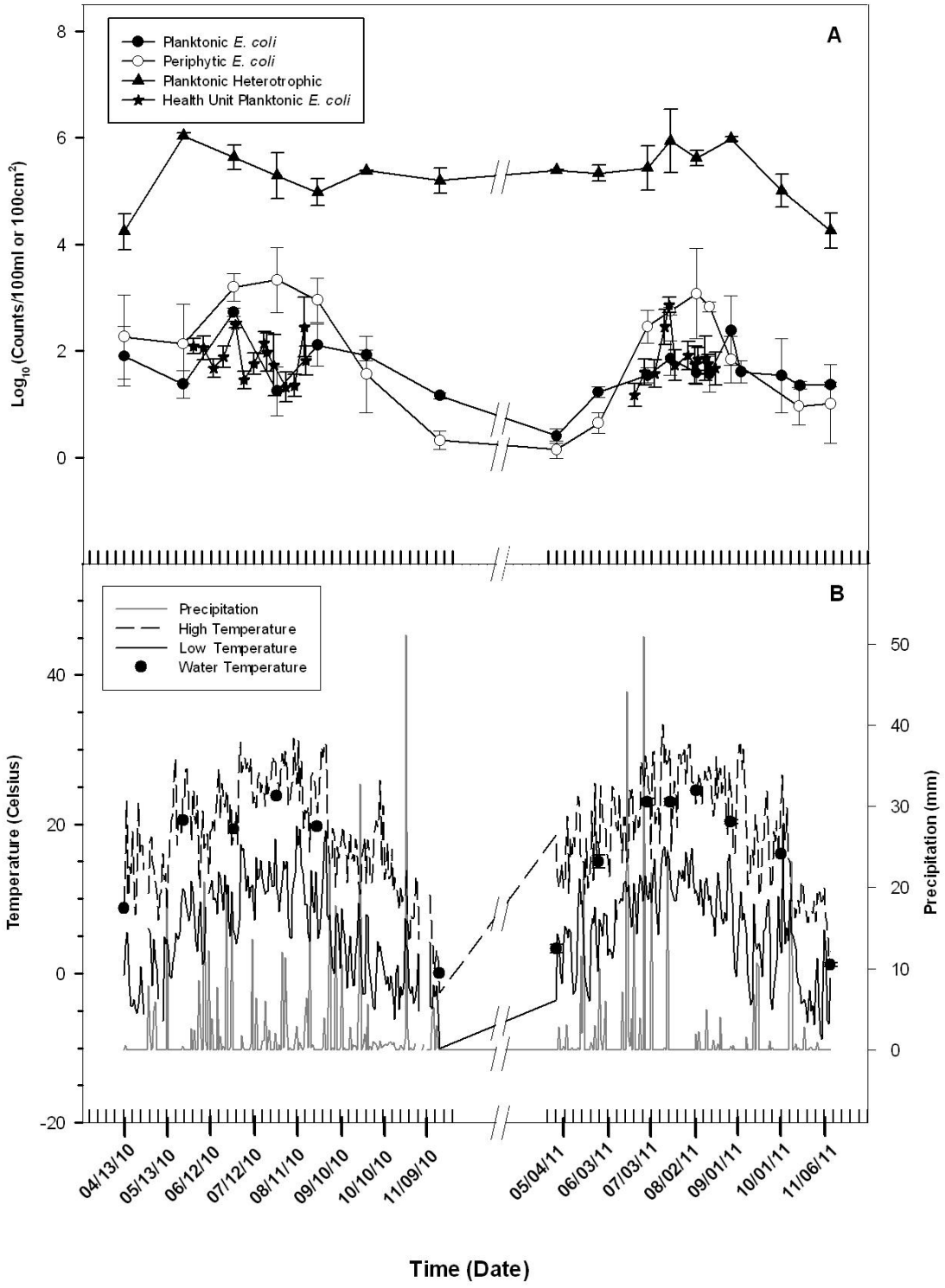
A
Planktonic E. coli
Periphytic E. coli
Planktonic Heterotrophic
Health Unit Planktonic E. coli
Log10 (Counts/100ml or 100cm2)
B
Precipitation
High Temperature
Low Temperature
Water Temperature
Temperature (Celsius)
Precipitation (mm)
04/13/10
05/13/10
06/12/10
07/12/10
08/11/10
09/10/10
10/10/10
11/09/10
05/04/11
06/03/11
07/03/11
08/02/11
09/01/11
10/01/11
11/06/11
Time (Date)

Figure 3-1

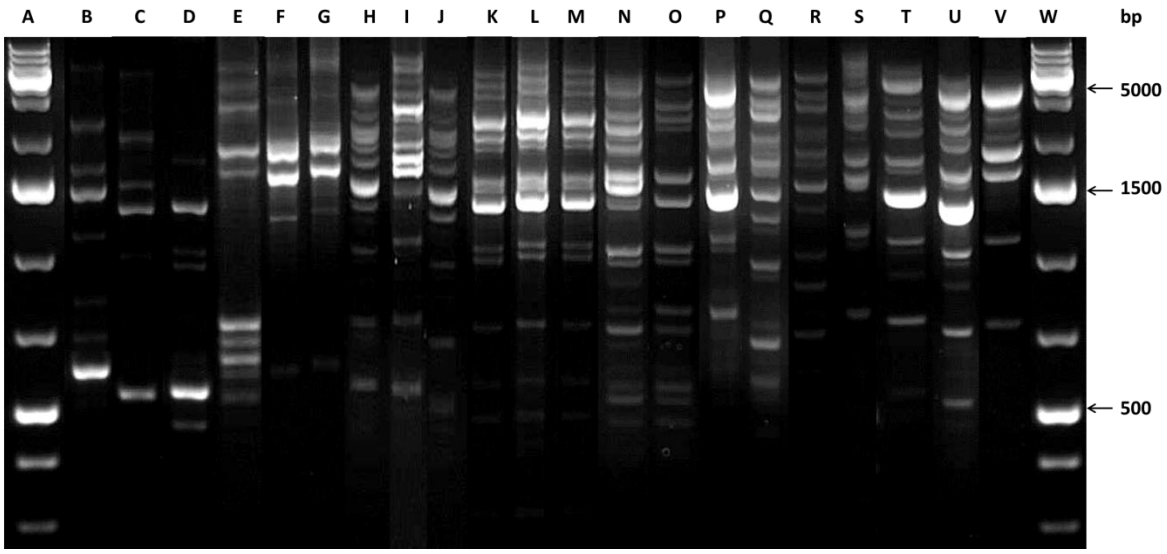

Figure 3-2

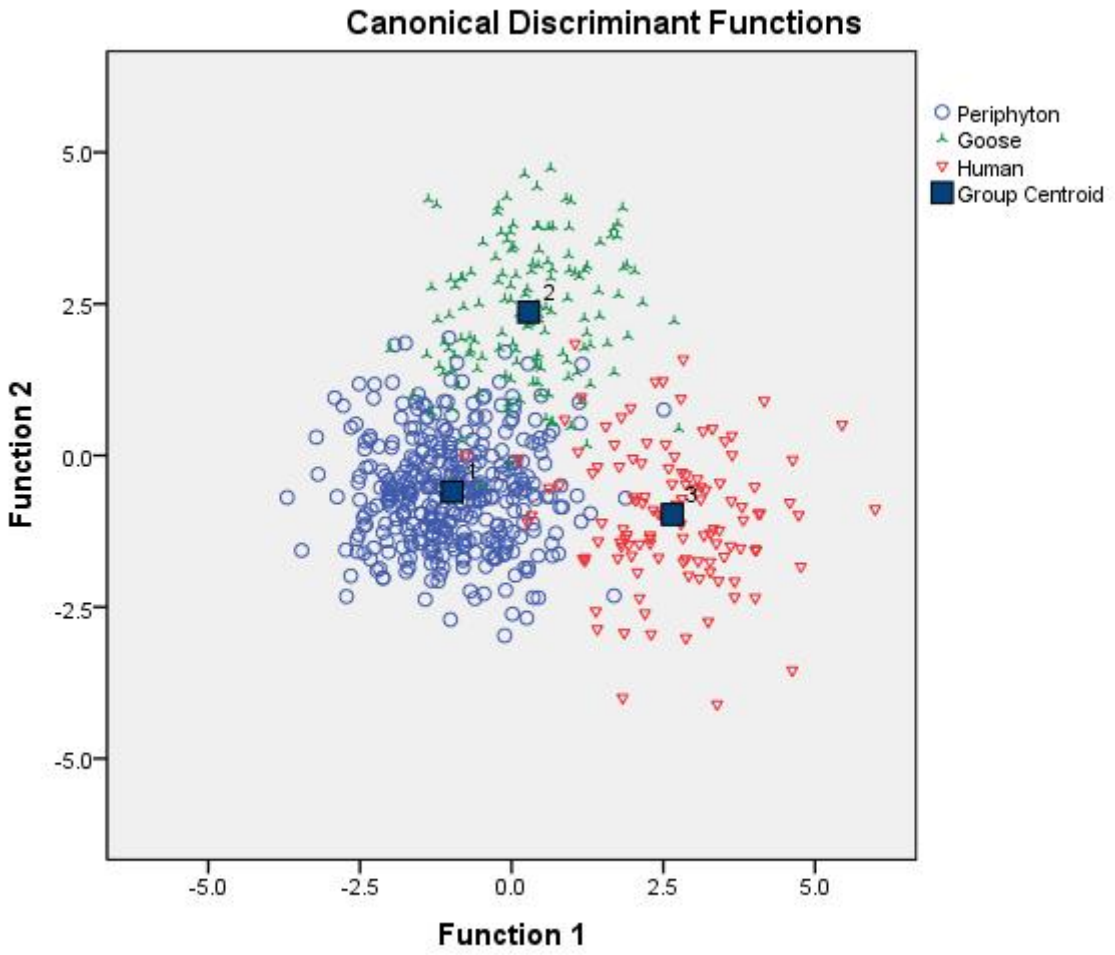
Canonical Discriminant Functions
Periphyton
Goose
Human
Group Centroid
5.0
2.5
0.0
-2.5
-5.0
Function 2
-5.0
-2.5
0.0
2.5
5.0
Function 1
1
2
3

Figure 3-3

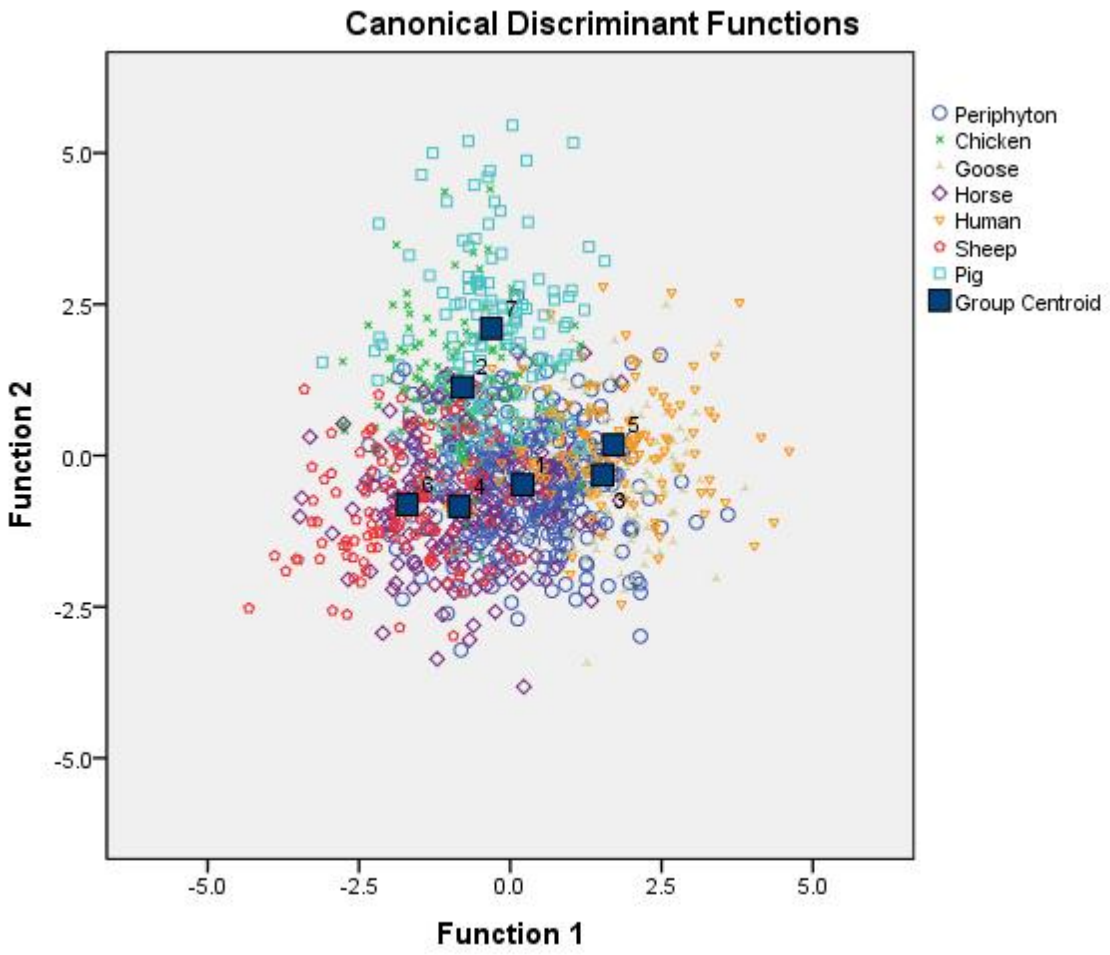
Canonical Discriminant Functions
Periphyton
Chicken
Goose
Horse
Human
Sheep
Pig
Group Centroid
Function 2
Function 1
5.0
2.5
0.0
-2.5
-5.0
-5.0
-2.5
0.0
2.5
5.0
1
2
3
4
5
6
7

Figure 3-4

Chapter 4 – Conclusions

Fecal contamination of recreational waters can have a large impact on human health, the environment, as well as the economy. *E. coli* is the most popular fecal indicator bacteria (FIB) to indicate fecal contamination. In Boulevard Lake, Thunder Bay, Ontario, most of the *E. coli* populations during the *E. coli* blooms was not directly contributed by fecal contaminants but was coming from the periphytic communities. Two lines of evidence support our conclusions. First, the *Bacteroides* 16S rDNA markers were monitored and compared with the *E. coli* populations in 2011. Both these fecal indicator bacteria were expected to have a correlated increase when fecal contamination occurred. There were two peaks in the planktonic *E. coli* population at 2.86 and 2.38 log CFU 100 ml^{-1} on July 21, 2011 and September 2, 2011, respectively. When the *Bacteroides* biomarkers were compared with these two *E. coli* peak periods, no correlated increases of the biomarkers were found. This would indicate that the increase of the *E. coli* population might not be due to fecal contamination.

The notion that *E. coli* blooms in Boulevard Lake were not a direct consequence of fecal contamination was also supported by our microbial source tracking (MST) data. The planktonic *E. coli* population was monitored for both 2010 and 2011 and increased levels in its population were observed numerous times. The REP-PCR of the planktonic *E. coli* isolates were compared to a REP-PCR *E. coli* library consisting of farm animals, (horse, sheep, pig, and chickens), human sewage, geese, and periphyton. It was determined that the farm animals did not have an impact on Boulevard Lake. Therefore, the three main sources used in the MST analysis were the human, geese, and periphyton *E. coli* isolates. The majority of the *E. coli* in the water during the peak periods was found to belong to *E. coli* in the periphytic communities at 57.6, 65.8, and 38.7

% on June 28, 2010, August 25, 2010, and September 2, 2011, respectively. The majority of the planktonic *E. coli* during the non-peak periods was also found to come from the periphytic *E. coli* (35.0 – 80.0 %). Overall, the *E. coli* in Boulevard Lake was found to primarily come from the periphytic *E. coli* throughout the 2010 and 2011 season.

The use of *E. coli* in conjunction with the *Bacteroides* biomarkers to indicate contamination proved to be an accurate technique and is recommended to indicate fecal contamination. We also suggest using MST techniques to monitor the bacterial population throughout the season and also to specifically follow the bacterial blooms to track their source.

Future studies on *E. coli* in the periphytic communities and the cause for their rapid proliferation and dispersal in Boulevard Lake should be performed. The pathogenicity of the planktonic *E. coli* in Boulevard Lake should also be tested to determine if it is safe for recreational activities. Dredging the lake could reduce the high *E coli* levels in the water column because it reduces the quantity of the periphyton as well as the amount of sediments being stirred by the water current. Furthermore, a deeper lake also means cooler temperature of the water column and the growth of the periphytic *E. coli* would, therefore, be slowed down. Even though the primary source of the *E. coli* in the water may not be due to fecal contaminants, the animals around Boulevard Lake should still be controlled. If the populations of geese are high enough, they may still contribute enough fecal matter to cause a bloom in the planktonic *E. coli* population.