Microbial Source Tracking in Two Nova Scotia Watersheds

by

Emma Stea

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
November 2013

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Halifax Water (HW) manages two source water systems, Lake Fletcher and the Musquodoboit River, where elevated levels (>200 CFU/100 mL) of *Escherichia coli* are common. Host-associated *Bacteroidales* order 16S rRNA gene-based microbial source tracking markers and human and cow mitochondrial DNA (mtDNA) markers were used to determine sources of fecal contamination in both watersheds. Human *Bacteroidales* markers were found at all sites except for the upper woodland site in the Lake Fletcher watershed. Higher rates of *Listeria monocytogenes, Campylobacter jejuni* and *E. coli* O157:H7 were detected in the Middle Musquodoboit (MM) watershed. Serogrouping of presumptive *L. monocytogenes* isolates showed the IIa (1/2a, 3a serovars) and IIb (1/2b, 3b serovars) serogroups were prevalent in both watersheds. Detection of *E. coli* weakly predicted *Campylobacter* and *Salmonella* occurrence in both watersheds (OR >1, p-value <0.001). Results suggest failing onsite wastewater systems are the leading cause of fecal pollution in both watersheds.

LIST OF ABBREVIATIONS USED

ARA	Antibiotic Resistance Analysis
BB	Bolton Broth
BD	Becton, Dickinson and Company
BMP	Beneficial Management Practices
BHI	Brain Heart Infusion
Bp	Base pair
BPW	Buffered Peptone Water
BSA	Bovine Serum Albumin
CCME	Canadian Council of the Ministers of the Environment
CDC	Center for Disease Control and Prevention
CFU	Colony Forming Units
CP	Collin's Park/Lake Fletcher
Ct	Threshold cycle
DNA	Deoxyribonucleic Acid
DGGE	Denaturing Gradient Gel Electrophoresis
DO	Dissolved Oxygen
dNTP	deoxynucleotides
dsDNA	double-stranded DNA
FIB	Fecal Indicator Bacteria
Ha	Hectare
HRM	Halifax Regional Municipality
HW	Halifax Water
IMS	Immunomagnetic separation
LB	Lysogeny Broth
LEB	*Listeria* Enrichment Broth
LD-MST	Library-Dependent Microbial Source Tracking
LI-MST	Library-Independent Microbial Source Tracking
LOD	Limit of Detection
LOQ	Limit of Quantification
MLST	Multi-Locus Sequence Typing
MM	Middle Musquodoboit
MPC	Magnetic Particle Concentrator
MPN	Most Probable Number
mS	milliSiemens
MST	Microbial Source Tracking
mtDNA	Mitochondrial DNA
NCBI	National Center for Biotechnology Information
NADH	Nicotinamide Adenine Dinucleotide
NTU	Nephelometric Turbidity Unit
NWRI	National Water Research Institute
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with Tween
PCR	Polymerase Chain Reaction
PFGE	Pulsed Field Gel Electrophoresis

PMA	Propidium Monoazide
qPCR	Quantitative PCR
rDNA	Ribosomal DNA
rRNA	Ribosomal RNA
REP-PCR	Repetitive Extragenic Palindromic PCR
RVS	Rappaport-Vassiliadis Soy Broth
SGAL	3,4-cyclohexenoesculetin-b-D-galacto- pyranoside
SR	Sampling Run
SWAT	Soil and Water Assessment Tool
SWP	Source Water Protection
TSB	Trypticase Soy Broth
TSS	Total Suspended Solids
UNG	Uracil-N Glycosylase
US EPA	United States Environmental Protection Agency
VBNC	Viable But Not Culturable
WHO	World Health Organization

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr. Lisbeth Truelstrup-Hansen and Dr. Rob Jamieson for their help and guidance throughout the project. I would also like to thank the following people for their involvement with this project: Richard Scott, Greg Piorkowski, Laura Purdue, Amy Jackson, Ali Hosein, Kristen Myette, Jenny Hayward, Danielle Finlayson-Bourque, Tristan Goulden, Justine Lywood, Evan Bridson-Pateman, Yannan Huang, Marcia English, Marta Piercey, Tricia Hingston, and Christina Ridley. I would also like to thank Barry, Anna and the Archibald family. Thank you to the Canadian Water Network, Halifax Water and NSERC for providing funding. Thank you to my family and friends, for helping me along the way.

CHAPTER 1 INTRODUCTION

1.1 THESIS CONTEXT

Drinking water systems in Canada withdraw water from surface water or groundwater, or a combination of both water sources. Surface water can easily become contaminated from fecal sources (such as agricultural runoff) while groundwater under the direct influence of surface water may also be at risk of contamination through infiltration from nearby surface sources. The majority of Canadians using a private water supply collect water from groundwater sources (Moffatt & Struck, 2011). Surface water sources, in contrast, are the main source of water supply for Canada's public water utilities (Moffatt & Struck, 2011).

Fecal pollution of aquatic ecosystems leading to contamination of irrigation water, and source water for drinking water treatment plants, can significantly impact human health (Edge et al., 2006). In 2000, a tragic outbreak of waterborne disease occurred in Walkerton Ontario, claiming the lives of seven Canadians (O'Connor, 2002). In the United States, there were 764 documented waterborne outbreaks associated with drinking water from 1971 to 2002, resulting in 575,457 cases of illness and 79 deaths (Blackburn et al., 2004).

To avoid waterborne disease outbreaks and boil water advisories, water utilities such as Halifax Water (HW) use a multi-barrier approach to prevent or reduce the entry of pathogenic bacteria such as *Escherichia coli* O157:H7 into drinking water. Source water protection is the first step in the multi-barrier approach, which also includes continuous monitoring and testing of source water. The goal of source water protection is

to maintain or improve the quality of raw drinking water before it reaches the treatment facility (Halifax Regional Municipality, 2013).

Halifax Water currently manages eight watersheds and three groundwater sources, supplying drinking water to over 79,000 households and business in the Halifax Regional Municipality (HRM, 2013). The source watersheds in Middle Musquodoboit and Collin's Park (CP)/Lake Fletcher managed by HW were examined in this study. Lake Fletcher is part of the Shubenacadie Canal, connecting the Bay of Fundy with Halifax Harbour. The Lake Fletcher sub watershed is characterized by 50% urban residential development, mostly surrounding the shore of Lake Fletcher. The sub watershed land area is approximately 1,408 ha, with 45% of land area presently forested, and the remaining 5% is used commercially. The community of Fall River, surrounding Lake Fletcher, is an area that will increase in residential developments in the future as several areas are slated for development. Lake Fletcher is also downstream of Lake MicMac and Lake Charles, which are surrounded by heavy residential developments. Lake Fletcher serves as the water supply for approximately 80 customer connections in the Collin's Park subdivision. Raw water is treated by nano-filtration and ultra-filtration, UV disinfection, and chlorination prior to distribution.

The Musquodoboit River is the source water supply for 96 households and businesses in the village of Middle Musquodoboit (MM). Upstream of the intake, the watershed land base consists of 70% forest cover, 10% is available for agriculture, and the remaining 20% is developed residentially or commercially. Raw water treatment processes include bank filtration at the intake, followed by nano-filtration and ultra-

filtration, UV disinfection and chlorination. The MM water supply area upstream of Halifax Water's water treatment plant covers a drainage area of approximated 33,284 ha.

Currently, Halifax Water's source water program does not provide insight into the sources of fecal contamination in these watersheds making it difficult to properly target watershed management activities. A number of pathogen sources including humans and livestock are present in both watersheds. Uncertainties surrounding the use of *E. coli* as a predictor of pathogen presence exist, making interpretation of the commonly observed elevated levels of the indicator organism *E. coli* (>200 CFU/100 mL) in Lake Fletcher and the Musquodoboit River difficult. In addition, the current sampling scheme provides little information with samples of the indicator organism collected on an infrequent basis. Additional tools are needed to identify sources of microbial pollution in these watersheds. This thesis strives to improve upon current monitoring and watershed management strategies by conducting an in-depth analysis of fecal contamination source(s) by use of microbial source tracking (MST) techniques.

1.2 Research Questions and Objectives

The following research questions were identified for this project in order to address source water management concerns in the two study watersheds:

1. What is the relationship between the presence and level of the fecal indicator organisms and specific bacterial pathogens within both study watersheds?

2. Do watershed characteristics and storm events affect the relationships between indicator organisms and bacterial pathogens?

3. Does the type of anthropogenic pressure (urban vs. rural) impact the diversity of pathogens present in the watersheds?

4. What are the relative contributions of different fecal contamination sources within both watersheds?

The specific objectives of this study were to:

1. Validate the HF183 Human *Bacteroidales* and HcytB/AcytB mitochondrial DNA markers for use in Nova Scotia

2. Characterize the relationships between *E. coli* and total coliform levels and the presence of pathogenic microorganisms *Campylobacter* spp., *Salmonella*, *Listeria monocytogenes* and *E. coli* O157:H7 within both study watersheds

2. Determine the relative contribution of different fecal contamination sources within both watersheds using human and bovine 16S rRNA gene-based *Bacteroidales* markers and human and bovine mitochondrial DNA markers

CHAPTER 2 LITERATURE REVIEW

2.1 MICROORGANISMS IN SURFACE WATER

Many different microorganisms may be present in surface water. These microorganisms can include bacteria, protozoa, and viruses. Some bacteria are of little concern to human health, while others have the ability to cause disease in humans and/or animals. Some microorganisms exist naturally in water, while others are a result of contamination from fecal sources. Disease-causing microorganisms are called pathogens, and they are of primary concern to individuals and organizations using surface water as a source of drinking water or for recreational, personal, and agricultural purposes. Water safety is a massive worldwide problem, with mortality due to waterborne disease exceeding 5 million annually (WHO, 2013).

2.1.1 Bacterial Pathogens

Examples of pathogenic bacteria of concern in Canada include *E. coli* O157:H7, *Salmonella*, *Campylobacter*, and *Listeria monocytogenes*. *E. coli* is a Gram-negative bacterium found in the digestive tract of warm blooded animals (Rosen, 2000). *E. coli* can be divided into four main phylogenetic groups (A, B1, B2 and D) (Clermont et al., 2000). The pathogenic *E. coli* can be further divided into six pathotypes based on clinical features and virulence characteristics (Cabral, 2010). *E. coli* O157:H7 is a member of the enterohaemorrhagic pathotypes, and has been responsible for waterborne and food borne outbreaks of disease (Petridis et al., 2002). Healthy cattle are the primary reservoir of *E. coli* O157:H7 (Rosen, 2000). This strain of *E. coli* causes abdominal pain, bloody

diarrhea, and hemolytic uremic syndrome in humans (Health Canada, 2006). In the United States, 52% of *E. coli* O157:H7 outbreaks were foodborne and 9% were waterborne between 1982-2002 (CDC, 2005). Additionally, individuals may be capable of transmitting the bacteria to others via the fecal-oral route (Rosen, 2000). As few as 10-100 cells can cause infection (Petridis et al., 2002), although the probability of falling ill from *E. coli* O157:H7 depends on the number of organisms ingested, the resistance of the person to this organism, and the health of the individual (Health Canada, 2006). Children, the elderly, and immunocompromised persons may be more susceptible to serious complications from *E. coli* O157:H7, as well as other pathogens (Rosen, 2000; Health Canada, 2006; Cabral, 2010). *E. coli* O157:H7 has also been documented to persist in the environment in soils, water and cattle effluents (Fremaux et al., 2008).

Members of the *Salmonella* genus are Gram-negative rod-shaped bacteria found in the intestines of a wide variety of animal hosts (Cabral, 2010). Approximately 1.2 million illness linked to *Salmonella* occur every year in the United States (CDC, 2011). *Salmonella* infections are often associated with consumption of food and water that was improperly cooked or disinfected (CDC, 2011). In 1993, a large outbreak of waterborne *Salmonella* Typhimurium in Gideon, Missouri resulted in 650 cases of illness and 7 deaths (Angulo et al., 1997). *Salmonella* causes typhoid and paratyphoid fever and gastroenteritis, especially in vulnerable individuals (Rosen, 2000). *Salmonella* can be divided into two species (*S. bongori* and *S. enterica*), where *S. enterica* consists of six subspecies and over 2.500 serotypes based on diversity of the O and H antigen surface structures (CDC, 2011). Infection by less than 1000 cells is sufficient to cause infection and isolates from environmental sources are typically non-Typhi or Paratyphi serovars

(Cabral, 2010). Similar to other pathogens discussed in this review, there is currently no maximum allowable concentration of this pathogen in drinking water (Health Canada, 2006).

Members of the *Campylobacter* genus are highly motile Gram-negative bacteria that thrive in the warm, anaerobic environment of the intestinal tract (OIE Terrestrial Manual, 2008). There are currently 17 species and 6 subspecies assigned to the genus *Campylobacter* (World Health Organization, 2011). *Campylobacter* has a wide host range including sheep, goats, dogs, cattle, and humans (Ogden et al., 2009). *Campylobacter jejuni* causes 80% of *Campylobacter*-related gastroenteritis in humans, although other species of *Campylobacter* (*C. lari, C. coli* and *C. upsaliensis*) also have the ability to cause illness (OIE Terrestrial Manual, 2008). Survival of *Campylobacter* spp. and *C. jejuni* in water is possible. *Campylobacter* spp. have been isolated from environments such as groundwater, rivers and lakes, and coastal waters (McElhany & Pillai, 2011). *Campylobacter* co-occurred with *E. coli* O157:H7 in the waterborne disease outbreak that happened in Walkerton, Ontario in May 2000 (Clark et al., 2003).

Although not listed as a pathogen of concern in surface water by Health Canada or the US Department of Agriculture, this study examines the prevalence of *Listeria monocytogenes* in an urban and agricultural watershed. *Listeria monocytogenes* is a widely known food borne pathogen that causes listeriosis (high fever, gastro intestinal symptoms, miscarriage) in vulnerable populations including pregnant women, children and the elderly (Government of Canada, 2012). High-risk foods for *L. monocytogenes* include soft cheese, deli meats, and unpasteurized milk (Ryser and Marth, 2007). Current data suggests only a high level of bacterium numbers in surface water could cause illness

(Sauders & Weidmann, 2007). Understanding *Listeria* ecology in the natural environment, however, is essential in understanding how this pathogen contaminates food products through transmission pathways between various reservoirs and environments. Few studies in the current literature have examined the ecology and characteristics of *Listeria monocytogenes* in a environmental mixed-use watershed setting (Lyautey et al., 2007).

Listeria monocytogenes is a Gram-positive bacterium that is able to tolerate high salt concentrations, various temperatures and acid stress (Sauders & Wiedmann, 2007). Members of the *Listeria* genus include both pathogenic and non-pathogenic species. Pathogenic species include *L. monocytogenes* and *L. ivanovii* (Guillet et al., 2010). Non-pathogenic species include *L. innocua*, *L. welshimeri*, *L. grayi*, *L. seeligeri*, *L. racourtiae*, and *L. marthii* (Orsi et al., 2011). *Listeria* is considered to be ubiquitous in the natural environment (Sauders &Wiedmann, 2007). *L. monocytogenes*, and *Listeria* spp., have been isolated from many environments including soil, water, animal feeds, sewage, food processing plants, and farm environments (Sauders & Wiedmann, 2007). *L. monocytogenes* was isolated from fecal samples of livestock, deer, moose, otter, raccoon and humans (biosolid and septic samples) in Ontario (Lyautey et al., 2007). In an United States study, *L. seeligeri* and *L. welshimeri* were significantly associated with natural environments, while *L. innocua* and *L. monocytogenes* were significantly associated with urban environments (Sauders et al., 2012)

Methods of characterizing *Listeria* spp. isolates to the species level, followed by serogrouping and further genotypic/serotyping analyses are important in food borne outbreaks and ecological studies. Selective and indicative agars such as Bio-Rad's Rapid

Lmono© agar can be used to identify isolates to the species level. The principle of this medium acts upon detection of phospholipase of *L. monocytogenes* and the inability of this species to metabolize xylose (Bio-Rad, 2005). *L. monocytogenes* can be divided into four main serogroups [IIa (1/2a, 3a serovars), IIb (I/2b, 3b serovars), IIc (1/2c, 3c serovars), and IVb (4b, 4d, 4e serovars)] which represent evolutionary complexes typified by a predominant serotype (Ward et al., 2010). Serogroup IVb accounts for 40% of sporadic human listeriosis cases (Ward et al., 2010). In the South Nation River watershed in Ontario, serogroups IIa and IVb dominated surface water samples (Lyautey et al., 2007). Differences in the somatic and flagellar antigens result in the division of *L. monocytogenes* into 13 different serotypes (Kérouanton et al., 2010). Four serovars, 1/2a, 1/2c, 1/2b, and 4b (part of serogroup IVb), cause over 98% of human listeriosis cases (Jacquet et al., 2002). Methods used to subtype *L. monocytogenes* include ribotyping, pulsed-field gel electrophoresis (PFGE), and multilocus-sequence based typing (MLST) (Sauders & Wiedmann, 2007).

2.1.2 Protozoan Pathogens

Protozoan pathogens of concern in aquatic systems include *Giardia* spp. and *Cryptosporidium parvum*. Both of these organisms can cause diarrhea in infected individuals, with worse effects in immunocompromized persons (Huang & White, 2006). Neither organism can reproduce outside the host (Health Canada, 2012). Both organisms can produce a chlorine-resistant waterborne cyst (*Giardia*) or oocyst (*Cryptosporidium*) (Rosen, 2000). *Giardia* was found in remote, pristine water in the Canadian North (Roach et al., 1993). *Giardia* is frequently found in rivers and lakes and in the intestinal tract of

humans, dogs, cats, bears, beavers, some birds and amphibians (Huang & White, 2006). *Giardia* is a flagellated protozoan with a motile trophozoite form and cyst form that can remain infectious for long periods (Rosen, 2000). *Cryptosporidium* can infect humans, cattle, sheep, horses, dogs, and wildlife like raccoons and mice (Health Canada, 2012)

2.1.3 Viral Pathogens

Viruses are important pathogens of concern in drinking water. Viruses typically range in size between 30-50 nanometers in diameter (Rosen, 2000). Viruses that can multiply in the gastrointestinal tract of humans are known as enteric viruses (Health Canada, 2011). Norovirus, Hepatitis A virus, adenoviruses, and rotaviruses are examples of enteric viruses (Rosen, 2000). Viruses are transmitted via the fecal oral route, and modes of transmission can include water, food, and person-to-person contact (Health Canada, 2011). The main symptom associated with enteric viruses is gastrointestinal illness (Rosen, 2000). For the majority of enteric viruses, the infectious dose is assumed to be low (Health Canada, 2011). Adenoviruses were routinely detected with other markers of human fecal pollution at coastal beaches in California (McQuaig et al., 2012). The multi-barrier approach is recommended to reduce or eliminate viruses in drinking water sources (Health Canada, 2011).

2.2 Measuring Microbial Contamination in Surface Water for Public Use

2.2.1 *E. coli* and Fecal Indicator Organisms

Non-pathogenic fecal indicator bacteria (FIB) have been used traditionally as an indicator of microbial water quality and safety. Levels of pathogens in water are often low, and distribution can be sporadic (Field & Samadpour, 2007). Direct detection and cultivation of pathogens and/or viruses in water can be costly and labour-intensive. In contrast, identification and cultivation of FIB in water is relatively inexpensive and simple to perform. An ideal indicator of fecal pollution should have the following characteristics: 1) exist in high numbers in feces; 2) not be pathogenic; 3) be simple and cost-effective to detect in water; 4) not be able to multiply in the environment; 5) exist in greater number than pathogens; and 6) have a slightly longer persistence than pathogens in the environment in order to indicate that pathogens were/may still be present (Cabral, 2010).

Total coliforms are a group of lactose fermenting fecal bacteria used as a water quality indicator. *E. coli* is a member of the fecal coliform group, a subgroup of total coliforms (Cabral, 2010). Fecal coliforms have the ability to ferment lactose at 44.5 ºC (Rosen, 2000). This ability to ferment lactose at 44.5 ºC is thought to indicate that fecal coliforms are foreign to the environment and therefore an indicator of recent fecal contamination (Rosen, 2000). The Gram-positive enterococci are a subgroup of the fecal streptococci and are also commonly found in the intestinal tracts of humans and other animals (Yost et al., 2011). Both the enterococci and fecal streptococci have been used as water quality indicators (Rosen, 2000). Methods to measure fecal indicator bacteria

(FIB) include membrane filtration, enzyme substrate tests, Colilert most probable number (MPN) methods, enzymatic methods based on *ß-galactosidase* and *ß-glucoronidase* activities, and polymerase chain reaction culture-independent methods (Edge & Boehm, 2011).

Many regulatory agencies have adopted guidelines for water safety based on levels of *E. coli* and/or FIB. In Canada, *E. coli* should not exceed 100 CFU/100 mL in irrigation water, 200 CFU/100 mL in recreational water and none detectable/100 mL in drinking water (Canadian Council for Ministers of the Enivronment, 1999). *E. coli* has become a primary FIB of interest, with studies linking elevated *E. coli* levels in recreational waters to increased risk of gastrointestinal illness (Marion et al., 2010). *E. coli* is considered the best indicator for presence of fecal matter from warm blooded animals in ground water and recreational waters, while total coliforms are considered useful as an indicator of treatment effectiveness in drinking water plants and groundwater (Edberg et al., 2000; Carraro-Colon et al., 2011).

2.2.2 Limitations of *E. coli* as a FIB

Issues concerning the ability of *E. coli* and other FIB to accurately indicate microbial water quality and safety exist. Most importantly, *E. coli* and other FIB do not identify sources of the fecal contamination in surface water (Field & Samadpour, 2007). These bacteria are found in the intestines of many animals (Carrero-Colon et al., 2011); therefore *E. coli* counts cannot provide evidence of a particular source of fecal pollution (such as human or livestock-related pollution).

Conflicting evidence concerning the ability of *E. coli* and other FIB to predict pathogen presence exists. Some studies suggest there is some ability of *E. coli* and other FIB to predict pathogen presence, especially *Salmonella* (Walters et al., 2011; McEgan et al., 2013). Other studies show no relationship between *E. coli*/FIB and pathogens (Ahmed et al., 2009; Drozd et al., 2012). Total coliforms and fecal coliforms showed no relationship or correlation with *Cryptosporidium* oocysts, *Giardia* cysts, and infectious enteric viruses at six wastewater reclamation facilities (Harwood et al., 2005). Lack of correlation between FIB and pathogens could be caused by differences in physiological and phylogenetic factors between the pathogens and indicator organisms (Harwood et al., 2013). Other studies have demonstrated *E. coli*/FIB persistence and/or regrowth in sediments in the environment, decreasing their usefulness as an indicator organism (Anderson et al., 2005; Litton et al., 2010; Byappanahalli & Ishii, 2011; Piorkowski et al., 2013). *E. coli* demonstrated a ubiquitous distribution (97% occurrence) in the green algae *Cladophora* in Lake Michigan during the summer months, indicating that algae may be another source of FIB in the environment (Whitman et al., 2003). Survival of *E. coli* in the environment is a complex process influenced by many factors such as temperature, presence of protozoan predators, and availability of nutrients (Carrero-Colon et al., 2011).

2.2.3 Origins of Microbial Source Tracking

In recent years, there has been a growing need for other methods that possess the capacity to rapidly identify microorganisms at the molecular level and identify fecal contamination sources. Identifying fecal contamination sources will result in better

source protection strategies by targeting specific sources of pollution (Edge et al., 2006; Stewart et al., 2007; Lee et al., 2011). Microbial source tracking (MST) is one method to determine sources of fecal contamination present in a particular watershed (Simpson et al., 2002). MST methods have been developed to address problems in source water for drinking, shellfish harvesting waters, and waters used recreationally. The MST approach compares the similarity of microorganisms/cells collected from aquatic ecosystems to microorganisms/cells collected from known fecal pollution sources in order to make inferences about the likely source of fecal contamination (NWRI Scientific Assessment Report, 2006). MST can be used to build a water quality monitoring toolbox that relies on both the traditional methods of *E. coli*/FIB testing as well as methods for targeting sources of fecal pollution in a particular area.

Recently, the city of Toronto used MST approaches at urban beaches. Results identified bird droppings as the main source of pollution in that area, not raw municipal sewage, as was initially expected. Antimicrobial resistance analysis and rep-PCR DNA fingerprinting of *E. coli* collected at the beach indicated that *E. coli* in sand and water samples were predominantly from bird droppings (Edge & Hill, 2007). Bird control measures have been put in place since the findings became available. These findings resulted in reduced beach closures near the city of Toronto (Environment Canada, 2008).

2.3 Microbial Source Tracking (MST) Approaches

2.3.1 Library-dependent MST

MST can be divided into two different categories: 1) library-dependent MST (LD-MST) and 2) library-independent MST (LI-MST). Both *Enteroccocus* spp. and *E. coli* are

indicator organisms, which have been widely used in LD-MST (Simpson et al., 2002). An example of LD-MST would involve the comparison of *E. coli* isolate fingerprints obtained from aquatic ecosystems with the DNA fingerprints of isolates in the library from known sources of pollution such as animal feces or wastewater in order to determine the likely source of the waterborne *E. coli* (Harwood et al., 2011). The established FIBs, *E. coli* and *Enterococcus* spp., are often chosen in LD-MST because they are relatively inexpensive and easy to culture in the laboratory (US EPA, 2005, NWRI Scientific Assessment Report, 2006).

Both phenotypic and genotypic methods can be used to classify isolates in LD-MST. Phenotypic methods use cellular or physiological comparisons between the isolates based on characteristics such as antibiotic resistance and less frequently, carbon utilization profiles (Simpson et al., 2002). Multiple antibiotic resistance methods test isolates of *E. coli* against different antibiotics to determine the likely source of fecal pollutions (Harwood et al., 2011). This method is based on the principal that humans, livestock, and wildlife have all been exposed to different antibiotics, and the fecal bacteria that they carry would differ in antibiotic resistance (Harwood et al., 2011). Some problems with resistance methods include: 1) spatial limitations of antibiotic resistance, 2) use of non-standardized antibiotic resistance test methods, and 3) wildlife coming into contact with animal feeds, causing isolates to gain resistance to livestock antibiotics (Field & Samadpour, 2007). Genotypic approaches, in contrast, are based on DNA fingerprinting methods. Various DNA cutting and amplifying techniques are used to obtain DNA fragments of different sizes from extracted DNA. Methods can include repetitive extragenic palindromic polymerase chain reaction (REP-PCR), pulsed field gel

electrophoresis (PFGE), ribotyping, random amplified DNA polymorphisms, and denaturing gradient gel electrophoresis (DGGE) (Simpson et al., 2002; Field & Samadpour, 2007; Harwood et al., 2011). The pattern, or fingerprint, of DNA fragments can then be visualized using gel electrophoresis (NWRI Scientific Assessment Report, 2006). LD-MST methods could require extensive sampling and hours of labour. Determining adequate sample size (library size) and spatial limitations of libraries (library results and conclusions are limited the study area) are two main areas of concern in LD-MST (Field & Samadpour, 2007).

2.3.2 Library-independent MST

Library-independent methods (LI-MST) use genotypic traits to identify sources of fecal pollution, but in contrast to LD-MST, do not rely on isolate collection. Typically, host-associated microbial DNA sequences are targeted in library independent analyses (US EPA, 2005). Specific marker genes are detected by polymerase chain reaction (Simpson et al., 2002). This approach removes a culture step, and permits detection of markers that could be difficult to detect using culture methods (Simpson et al., 2002). Molecular methods can be less time-consuming than culture-based techniques, but may require expensive reagents and equipment to perform. LI-MST techniques use markers based on bacteria or viruses that have adapted to the gut environment of particular host and/or eukaryotic (animal) cells that are shed by the host itself (Roslev & Bukh, 2011).

Fecal anaerobic bacteria, including the two important genera *Bifidobacterium* and *Bacteroides*, have host-associated distributions (Kreader, 1995) and are abundant in feces, however, were not used as indicators until the science of molecular detection

became readily available (Harwood et al., 2013). Traditionally, these fecal anaerobic bacteria were difficult to culture (Field & Samadpour, 2007). Bernhard & Field (2000a) were the first researchers to describe a PCR-based method for distinguishing human and cow fecal contamination without culturing indicator organisms. Two human-specific genetic markers were identified in fecal samples by amplifying 16S ribosomal DNA (rDNA) fragments from member of the genus *Bifidobacterium* and *Bacteroides-Prevotella* group (Bernhard & Field, 2000a). In the years since this advance in the field, *Bacteroidales* host-associated PCR based assays have been developed to detect feces from humans (Bernhard & Field, 2000b; Seurinck et al., 2005; Kildare et al., 2007), ruminants/cattle (Bernhard & Field, 200b; Reischer et al., 2006; Okabe et al., 2007; Shanks et al., 2008), horse (Dick et al., 2005b), deer/elk (Dick et al., 2005a), dogs (Dick et al., 2005a), geese (Fremaux et al., 2010), sheep (Lu et al., 2007), muskrat (Marti et al., 2011) and pigs (Okabe et al., 2007).

LI-MST methods have also been developed to detect host-associated viruses and mtDNA from cells shed by the host itself. Human and bovine adenoviruses were used to detect source-specific fecal pollution in Australia (Ahmed et al., 2008). Wolf et al. (2010) developed a multiplex method to detect human and animal enteric viruses, including norovirus genogroups (I, II, and III), porcine adenovirus, ovine adenovirus, and human adenovirus. mtDNA markers have been developed to detect fecal pollution from humans, sheep, horses, dogs, deer, and cows (Caldwell et al., 2007; Schill & Mathes, 2008; Baker-Austin et al., 2009; Baker-Austin et al., 2010; Kortbaoui et al., 2010). A multiplex assay has also been developed to detect mtDNA from eighteen common European mammals including the badger, mouse, rabbit and cats (Tobe & Linacre, 2008). The rationale

behind using host's own cells containing mtDNA for MST is based on feces containing large numbers of these cells shed from the host itself (Caldwell et al., 2007). In addition, human cells can contain approximately 100-2000 mtDNA copies/cell (Caldwell et al., 2007). LI-MST is not without problems or limitations. Limitations of library independent markers include: 1) lack of absolute host-specificity among microbial human and animal associated markers, 2) horizontal gene transfer of markers associated with virulence genes, 3) low or unknown levels of microbial markers in some host individuals or populations and 4) potential carryover of mtDNA from an animal (beef, pork) eaten then shed into human feces and non-fecal mtDNA sources (Roslev & Bukh, 2011). Host-specificity can be evaluated through marker specificity studies (described in section 2.3.5), but absolute host-specificity may not be a significant problem. Although the host-associated marker could cross-react with a small percentage of fecal samples from other species, the associated copy number value of the host-associated marker in the fecal sample would be very small and most likely below the assay detection limit.

2.3.3 PCR and qPCR Methods in MST and Pathogen Detection

PCR methods to rapidly amplify DNA are heavily used in both LD-MST and LI-MST methods and have an important role in pathogen detection in surface water samples. During a PCR reaction, template DNA (containing the target sequence) is mixed with deoxynucleotides (dNTPs), a DNA polymerase and primers. Primers are short segments of complimentary DNA that base pair with the template DNA upstream of the region of interest and serve as recruitment sites for the polymerase (NCBI, 2013). Conditions can be optimized so that primer binding is of high stringency, therefore increasing the

specificity of the assay. Cycles of denaturation, annealing and extension are repeated to amplify the target sequence. Visualizing the DNA following gel electrophoresis determines presence or absence of amplicons (NCBI, 2013). End-point PCR is qualitative and provides a presence/absence result for the host-associated marker or target.

Advances in DNA quantification technology led to the development of quantitative PCR technologies that allow quantification of copies of a particular target, such as a host-associated marker. The highly specific oligonucleotide probe in Taqman® qPCR is designed to hybridize within the target sequence (Reischl & Kochanowski, 1999). After the probe is cleaved during PCR due to activity of the Taq polymerase, fluorescence is released because of the separation of the fluorescent label and the quencher molecule (such as black hole quencher). Fluorescence is measured by the instrument. SYBR Green® is another commonly used fluorescent DNA binding dye in qPCR that binds all double–stranded DNA (Reischl & Kochanowski, 1999). The unknown amount of a particular target in a sample is determined in qPCR by running serial dilutions of a known amount of standard (often a plasmid containing the target sequence) in parallel to the samples of interest (Reischl & Kochanowski, 1999). qPCR can also be used to determine if a target is present (target amplifies during qPCR run) or absent (target does not amplify during qPCR run). Adequate positive and negative controls are important in all PCR and qPCR reactions. Taqman® qPCR methods have been developed to quantify or detect the presence of *E. coli* O157:H7 (Ibekwe et al., 2002). *Salmonella* (Cheng et al., 2008), *L. monocytogenes* (Rodríguez-Lázaro et al., 2004) and *Campylobacter* spp. (Lund, et al., 2004).

2.3.4 Limitations of PCR and qPCR Methods

qPCR and PCR methods have several limitations. PCR assays do not differentiate between live and dead microbial cells (Harwood et al., 2011). However, a method using a nucleic acid intercalating dye propidium monoazide (PMA), capable of penetrating dead cells and inhibiting PCR, inhibited PCR amplification in dead *E. coli* cells with over 99% efficiency (Taskin et al., 2011). PCR and qPCR analyses are also very sensitive to contamination in the laboratory especially if assay target DNA (such as human DNA) in present in PCR master mix preparation areas. PCR inhibitors such as humic acid can prevent or interfere with DNA amplification during the PCR reaction, and may be present in DNA extracted from soil and fecal samples. Differences in qPCR methods and culture based methods for the detection of a particular organism could be caused by 1) detection of dead cells by PCR methods and 2) changes in microbial communities in culture broths due to factors such as competition and the availability of nutrients and 3) cells in a viable but not culturable state (VBNC) due to environmental stresses, resulting in underestimation by culture-based methods (Oliver, 2005; Converse et al., 2009; Harwood et al., 2013). *L. monocytogenes*, *C. jejuni*, *E. coli*, and *Salmonella* have all been reported to enter into a VBNC state due to the exposure to natural stresses such as temperature changes, osmotic pressure, and oxygen concentrations (Oliver, 2005).

2.3.5 Important Criteria for LI-MST Method Validation

Validation of LI-MST markers begins with determining the sensitivity and specificity of the marker. Useful MST methods are sensitive to low initial starting values of a target and use a target that has a high enough level in feces so it can be diluted and

still be detected (Field & Samadpour, 2007). Sensitivity is defined as the ability of the assay (primers and/or probes sets) to detect the target. Specificity, in contrast, is the ability of the assay to discriminate between the target and other species (Fremaux et al., 2009). Sensitivity is also known as the true positive rate, and specificity could be known as the false positive rate (Harwood et al., 2013). Feces from a wide variety of animals common to the particular geographic area should be used in specificity and sensitivity studies. Determining if an assay is temporally and geographically stable is important, and many of the same assays have been tested in various parts of the world in LI-MST studies.

The limit of quantification (LOQ) is another important criterion for LI-MST studies. The LOQ is the fewest gene copies, or the least amount of gene copies in fecal material that can be accurately quantified, and this value is frequently an order of magnitude greater than the limit of detection (LOD). The LOD is defined as the lowest amount of target than can be detected (Harwood et al., 2013).

2.4 Universal, Human and Ruminant *Bacteroidales* and mtDNA markers for Detection of Fecal Pollution

2.4.1 Universal (AllBac) *Bacteroidales* and Universal mtDNA Marker Detection

Layton et al. (2006) were the first to develop a universal marker for the detection of *Bacteroides* from many animals. AllBac was tested on feces diluted in water and used to estimate the total amount of fecal contamination in water (Layton et al., 2006). Kildare et al. (2007) developed another 16S rRNA-based universal *Bacteroides* marker (BacUNI-

UCD) that was tested using various fecal samples. BacUNI was detected in all stool samples from humans, cats, dogs, seagulls, cows, and horses (Kildare et al., 2007). AllBac has since been used in many LI-MST studies as a marker of total fecal pollution (Park et al., 2010; Wyer et al., 2010; Mauffret et al., 2012; Eichmiller et al., 2013). AllBac's usefulness as a fecal indicator has been examined in relation to traditional FIB such as *E. coli*. Levels of *E. coli* were correlated with AllBac in other studies (Ridley et al., 2011; Mauffret et al., 2012). However, the AllBac marker may not be completely specific for feces. A recent study found large concentrations of AllBac in pristine alpine soil samples, suggesting a lack of specificity for feces (Vierheilig et al., 2012).

Assays have recently been developed to detect the host's own cells. This method differs from previous assays that detect the host indirectly through bacteria found in the intestine of the host. Kortbaoui et al. (2010) developed a universal mtDNA assay based on primer sets from consensus nucleic acid sequences found between human, ovine, bovine, and chicken mtDNA. In their study, universal mtDNA PCR combined with species-specific dot blot assays were used as a source tracking method for human, bovine, chicken, ovine and porcine pollution.

2.4.2 Human *Bacteroidales* and mtDNA Marker Detection

Many methods targeting the 16S rRNA gene of human associated *Bacteroides* have been established (Seurinck et al., 2005; Layton et al., 2006; Kildare et al., 2007; Reischer et al., 2007). "Human-associated *Bacteroidales*" is another name for the methods that target the bacterial 16S rRNA gene. HF183, a widely-used human-associated *Bacteroidales* marker assay, was first developed by Bernhard and Field in

2000, then modified with a new reverse primer targeting the human-associated sequence better suited for Sybr Green®-based qPCR in 2005 by Seurinck et al. The HF183 assay showed excellent specificity and sensitivity in many studies, with values ranging from 90-100% specificity and 90-100% sensitivity (Ahmed et al., 2008; Ahmed et al., 2009; Fremaux et al., 2009). Layton et al. (2006) developed a Taqman®-based qPCR human-associated *Bacteroidales* assay, HuBac that showed 100% sensitivity for human fecal samples, but low specificity (68%) when tested against animal samples. Another Taqman®-based human assay, BacH, developed by Reischer et al. (2007), was validated for in an agricultural watershed in Nova Scotia. In that study, however, BacH showed sensitivity problems (64% sensitivity) but had high specificity (91.9%) (Ridley et al., 2011).

Human mtDNA-based markers have recently been developed and validated in several studies. Excellent specificity and sensitivity over 90% were found in several studies for human mtDNA markers developed using either the cytochrome B gene or the mitochondrial NADH dehydrogenase subunit 5 (ND5) gene (Caldwell et al., 2007; Schill & Mathes, 2008).

2.4.3 Ruminant and Cow *Bacteroidales* and mtDNA Marker Detection

Host-associated assays were also developed to detect cow and ruminant fecal pollution in water. A Taqman®-based assay (BacR) was developed to detect ruminant fecal pollution (Reischer et al., 2006). Another Taqman®-based assay (CowM2) was developed to detect bovine fecal pollution by detecting a specific gene (often carried in a single copy per bacterial genome) encoding for a protein involved with host-bacterium

interaction (Shanks et al., 2008). In contrast, 16S rRNA genes are highly conserved and often carried in multiple copies per genome (Shanks et al., 2008). The CowM2 marker demonstrated a broad distribution among individual bovine samples (98% to 100%) (Shanks et al., 2008). Baker-Austin et al. (2010) developed a primer set (AcytB) that amplifies pig, sheep, and cow mtDNA sequence but uses a Taqman®-based probe specific for cattle only. Other probes were developed to detect the other animals (pigs, sheep).

2.4.4 Persistence of *Bacteroidales* and mtDNA Markers in the Environment

Bacteroidales and mtDNA based-markers may be successful predictors of recent fecal contamination in watersheds. However, some of these markers have demonstrated persistence in the environment, decreasing their usefulness as indicators of recent fecal pollution. AllBac decay curves showed a higher persistent population when compared with other targets such as HF183 and *E. coli* (Dick et al., 2010). When sediments were resuspended at the end of the experiment, the final concentration of the AllBac marker returned to approximately 50% of its original concentration, while the other markers remained at less than 1% of the original concentration (Dick et al., 2010). Environmental *Bacteroides* strains survived longer than cultivable *Bacteroides* strains in a river, but the period of their survival was shorter than that observed for fecal coliforms and enterococci (Ballesté & Blanch, 2010). Factors influencing the persistence of fecal *Bacteroides* in stream water include temperature, initial fecal concentrations and presence of other native microorganisms (Bell et al., 2007). Comparison of *Bacteroides* 16s rRNA genes/mL in

microcosms showed that stream water filtration (to remove native microorganisms and competition) followed by temperature, had the largest effects on gene persistence, with lower temperatures resulting in slower *Bacteroides* 16S rRNA gene decay (Bell et al., 2007). The mitochondrial HcytB was detected for up to 2 weeks at 4ºC (Baker-Austin et al., 2010). HF183, in contrast, was detected for up to 24 days at 4 ºC and up to 8 days at 28 ºC (Seurinck et al., 2005).

A Nova Scotian study examined the persistence of MST indicators in manure-amended agricultural soils. Soil type affected the persistence of the AcytB cow mtDNA markers, but had no effect on the decay of the ruminant-specific *Bacteroidales* marker (BacR) (Piorkowski, 2013). In contrast, the BacR marker decay rate was higher in soils that received a higher loading of liquid dairy manure, an effect not observed for the AcytB marker. Decay rates for the CowM2 marker could not be determined due to poor recovery of this marker in topsoil, reflecting its low initial concentration in manure (6.1 x 10^4 copies/g). Tambalo et al. (2012) found all of the *Bacteroidales* host-associated markers (human, cow, ruminant and horse) tested had a significantly shorter persistence than the conventional *E. coli* marker and were not detectable beyond 12 days. Concentrations of the ruminant (BacR) marker were well correlated with proximity to cattle operations in an urban and agricultural mixed-use prairie watershed (Tambalo et al., 2012).

2.4.5 Correlation of Marker Detection with Pathogens and FIB

Relationships between LI-MST markers, fecal indicator bacteria, and pathogens are very important in determining the usefulness of these markers as indicators of recent

fecal pollution. A moderate correlation was found between human, cow and pig 16S rRNA gene markers and fecal coliforms (r^2= 0.49) but no significant correlation was found between the human specific *Bacteroides* 16S rRNA gene marker and the concentrations of total or fecal coliforms (Okabe et al., 2007). Human-associated *Bacteroidales* correlated significantly with *E. coli* O157:H7, *Salmonella*, and enterotoxigenic *E. coli* in freshwater and wastewater treatment plants in Japan (Sivichtcheva et al., 2007). Walters et al. (2007) found that ruminant-specific *Bacteroidales* markers could predict *E. coli* O157:H7 and *Salmonella* in the Oldman River Basin in Alberta. The odds of detecting *Salmonella* in a sample when ruminant-specific markers were present were 2.5 times greater than the odds of detection when ruminant-specific markers were absent (95% CI from 1.46 to 4.18). The odds of detecting *E. coli* O157:H7 when CF128 (ruminant-specific marker) was present were approximately 37 times greater than when this marker was absent (95% CI from 4.65 to 298) (Walters et al., 2007). Fecal coliforms, enterococci, and universal *Bacteroidales* marker showed significant predictive ability for *Vibrio cholerae*, *Cryptosporidium*, and *Giardia* spp., while human *Bacteroidales* did not demonstrate any predictive ability for these pathogens (Schriewer et al., 2010).

More recent studies, however, demonstrate there may be little to no relationship between host-associated markers and pathogens. None of the host-associated *Bacteroidales* markers or *E. coli* were able to confidently predict *Campylobacter* or Shiga-toxin- gene positive *E. coli* in a prairie watershed (Fremaux et al., 2009). Human, ruminant, and pig host-associated *Bacteroidales* markers did not correlate with FIB or pathogens (*Salmonella* and *E. coli* O157:H7) in coastal California (Walters et al., 2013).

Human-specific *Bacteroidales* correlated much better with human mtDNA (R=0.62) in comparison to *E. coli* (R=0.33) in water sampled from an urban creek system in Ohio (Kapoor et al., 2013).

2.5 Sources of Contamination Driving the Need for MST Studies

2.5.1 Agricultural and Wildlife Impacts

Fecal contamination of aquatic ecosystems is a complicated problem. Fecal contamination in watersheds can lead to public health concerns, but also economic and ecological problems. Agricultural practices and wildlife sources are considered two of the most common sources of pollution in Canada (NWRI Scientific Assessment Report, 2006). The application of manure to agricultural fields and the treatment of livestock waste are areas of concern. Farmers that follow beneficial management practices, or BMPs, minimize environmental impacts and optimize plant growth. BMPs also decrease the nitrogen and phosphorus load from agricultural fields that can severely impact the ecology and environment of nearby streams (Agriculture Canada, 2013). BMPs include leaving crop residues on the soil surface during the winter, controlling manure deposition, maintaining vegetation in ditch banks, forming buffer zones close to streams and rivers, and shaping or seeding field edges to filter runoff as much as possible (Agriculture Canada, 2013).

Poor farming practices, storms and surface runoff can result in pathogen dispersal into nearby streams, rivers, and lakes. Water managers and communities need to be able to track sources of livestock fecal pollution quickly to prevent contamination of source water for drinking, irrigation, or recreation purposes. Water management challenges

include the management of waste produced by rearing of hogs, poultry, and cattle. In Canada, livestock densities and manure production varies. However, current trends point towards specialized larger farms that employ a small, immediate land base (NWRI Scientific Report, 2006).

Wildlife represents another challenge in managing microbial water quality in watersheds. Wildlife densities are highly variable, and densities can vary between seasons and years. MST methods have recently been developed to detect fecal pollution from wildlife of interest, including Canada geese (Fremaux et al., 2010), gulls and ducks (Green et al., 2012). Birds and wild animals were the major contributors to fecal bacterial pollution in a small southern Californian urban watershed (Jiang et al., 2007). Wildlife management measures may be needed especially in the case of gulls or geese, which have been linked to contamination at beaches in Canada (Environment Canada, 2008).

2.5.2 Municipal Effluents and Onsite-wastewater System Impacts

Other major contributors to microbial pollution in surface water include residential onsite wastewater systems, urban storm water, and sewer overflows. In the past, the impact of onsite wastewater systems may have been minimal with smaller rural populations. The connection between groundwater, streams, and lakes, may not have been an important concern in the past, however, this issue is important today with close to half of Nova Scotians using onsite wastewater systems (NS Wastewater Society, 2009). Wastewater leaving the drain field of a conventional onsite wastewater system moves through the unsaturated zone above the water table before eventually reaching the water table below. In some cases, the depth to water table is shallow and soils are

permeable, resulting in rapid recharge from septic systems into groundwater (Winter et al., 1998). Partially treated septic effluent reaching the water table can carry bacteria, viruses, and nitrogen; groundwater movement then moves these contaminants towards rivers or lakes (Winter et al., 1998). Septic systems that are properly designed, built, and maintained reduce and/or eliminate the majority of contaminants found in household wastewater. Regular maintenance and monitoring every 3-5 years is necessary and recommended to avoid septic tank failure and pollution (NS Wastewater Society, 2009).

Evidence of widespread human fecal contamination in urban storm water has been found (Sauer et al., 2011). Measures to reduce the amount of storm water, including harvesting roof water, minimizing impervious surfaces, and managing rainwater into bioretention areas, have been adopted in various cities across Canada (Marsalek & Schreier, 2008). Sewer overflows and leaking sewage infrastructure are also a particular concern in urban environments (WHO, 2004; NWRI Scientific Assessment Report, 2006).

2.6 MST and Preventing Pollution in Surface Water

2.6.1 Source Water Protection Plans

Both LD-MST and LI-MST studies can provide science-based evidence for corrective actions to improve microbial water quality in aquatic ecosystems and enhance source water protection. Source water protection plans (SWPs) are part of the multi-barrier approach to drinking water safety, where multiple barriers are put into place to control hazards. Steps of SWPs include 1) delineating the source water protection area; 2) creating an inventory of known and potential sources of contamination; 3) determining

the susceptibility of public drinking water sources to contaminants identified; 4) notifying the public about identified threats, and 5) implementing control measures to prevent or reduce risks from entering the drinking water supply (US EPA, 2002). Control measure examples include land acquisition and riparian buffer zones. Results from MST studies can drive implementation of specific control measures to reduce or prevent microbial risks (WHO, 2004).

2.6.2 Watershed Models as a Water Management Tool

Models can be used to improve source water protection and predict impacts of land-use changes on microbial water quality. The soil and water assessment tool (SWAT) has gained recognition as a robust, interdisciplinary watershed-modeling tool. SWAT is a basin-scale continuous-time model that operates on a daily time step and is designed to predict the impact of management and land use practices on water, sediment, nutrient and pesticide yields (Gassman et al., 2007). Major components of the model include weather, hydrology, soil temperature, nutrients, pesticide application rates, pathogens, and bacteria. Models such as SWAT, combined with watershed monitoring, might be useful in predicting contamination events and alerting downstream water users of potential hazards (WHO, 2004; Gassman et al., 2007).

CHAPTER 3 MATERIALS AND METHODS

3.1 STUDY BACKGROUND

Two distinct watersheds were chosen in this study to determine sources of fecal pollution in an urban area (Lake Fletcher watershed) and an agricultural area (Middle Musquodoboit watershed) where high levels of the indicator organism *E. coli* are common. Host-associated *Bacteroidales* markers (BacR, CowM2) were validated by Ridley et al. (2011) for use in Nova Scotia and were also used in this project using TaqMan® qPCR methodology. In addition, new mtDNA markers (one human, HcytB and one cow based, AcytB) were added to this study and evaluated for use in Nova Scotia. A widely used *Bacteroidales* marker for human fecal contamination (HF183) was chosen for this study based on its high sensitivity and reliability as a marker of human fecal pollution (Ahmed et al., 2008). This marker was quantified in samples using Sybr Green® qPCR chemistry. Further characterization of *Campylobacter* DNA, and *E. coli* and *L. monocytogenes* isolates using multiplex PCR methods provided greater insight to the diversity, ecology and potential sources of these bacteria in the environment. Quantifying markers of fecal pollution and understanding pathogen diversity in these two different watersheds will provide direction for future source water protection activities.

3.2 WATER SAMPLE SITES

Water samples were collected in 33 separate sampling events from five sites in Middle Musquodoboit (Figure 3.1) and five sites in Lake Fletcher (Figure 3.2) over an 18-month period. Middle Musquodoboit is a primarily agricultural watershed while the

Lake Fletcher watershed consists primarily of urban residential communities. Lake Fletcher provides the source water for the Collin's Park (CP) subdivision. In this thesis, water sampling locations in this watershed were designated by "CP". Sample locations in the Middle Musquodoboit watershed were designated by "MM". The sampling sites are described in detail in Table 3.1.

Table 3.1 Description of sampling sites in the MM and CP watersheds

Site	Description
MM1	Main river channel, near water treatment plant and agricultural activities
MM2	Sub-watershed, agricultural activities
MM3	Sub-watershed, dairy farm
MM4	Main river channel, upstream
MM5	Control location further upstream, also close to a working farm and a household
CP1	Lake Thomas run, flowing into Lake Fletcher, fast flows, wide channel
CP3	Bottom of Holland Brook, residential subdivision
CP4	Top of Holland Brook, forested area
CP5	Site below lift station that pumps wastewater into nearby wastewater treatment plant
CP6	Control location below Lizard Lake, one new home built nearby – area to be developed further in the future
CP7	New location added in 2013 just below A-lake, new subdivision including large homes with onsite wastewater systems

Samples were also collected a total of three times over the sampling period from the raw intake water at the Middle Musquodoboit and Collin's Park Water Treatment Plant (WTP). These samples underwent the same testing for markers and pathogens as other water samples described in Section 3.6. The raw intake water at the MM WTP is filtered through a sandy riverbank prior to entering a well. The raw intake water at the CP WTP flows through a pipe from the center of the lake.

In summer 2013, an additional site labeled CP7 was added to the project. CP7 is a stream just below A Lake (Figure 3.2). The stream sits in a recently developed area with large lot sizes and onsite wastewater systems.

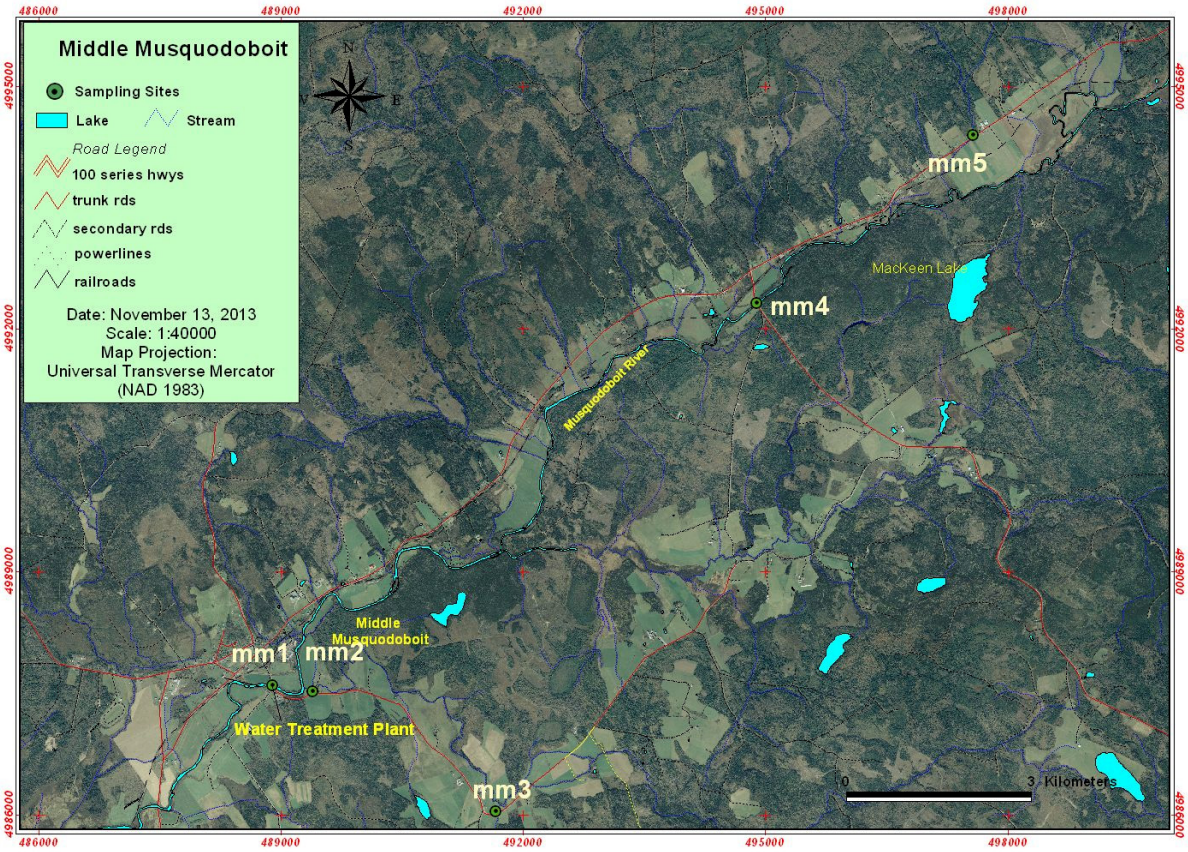

Figure 3.1 Middle Musquodoboit watershed and sampling locations (MM1-MM5)

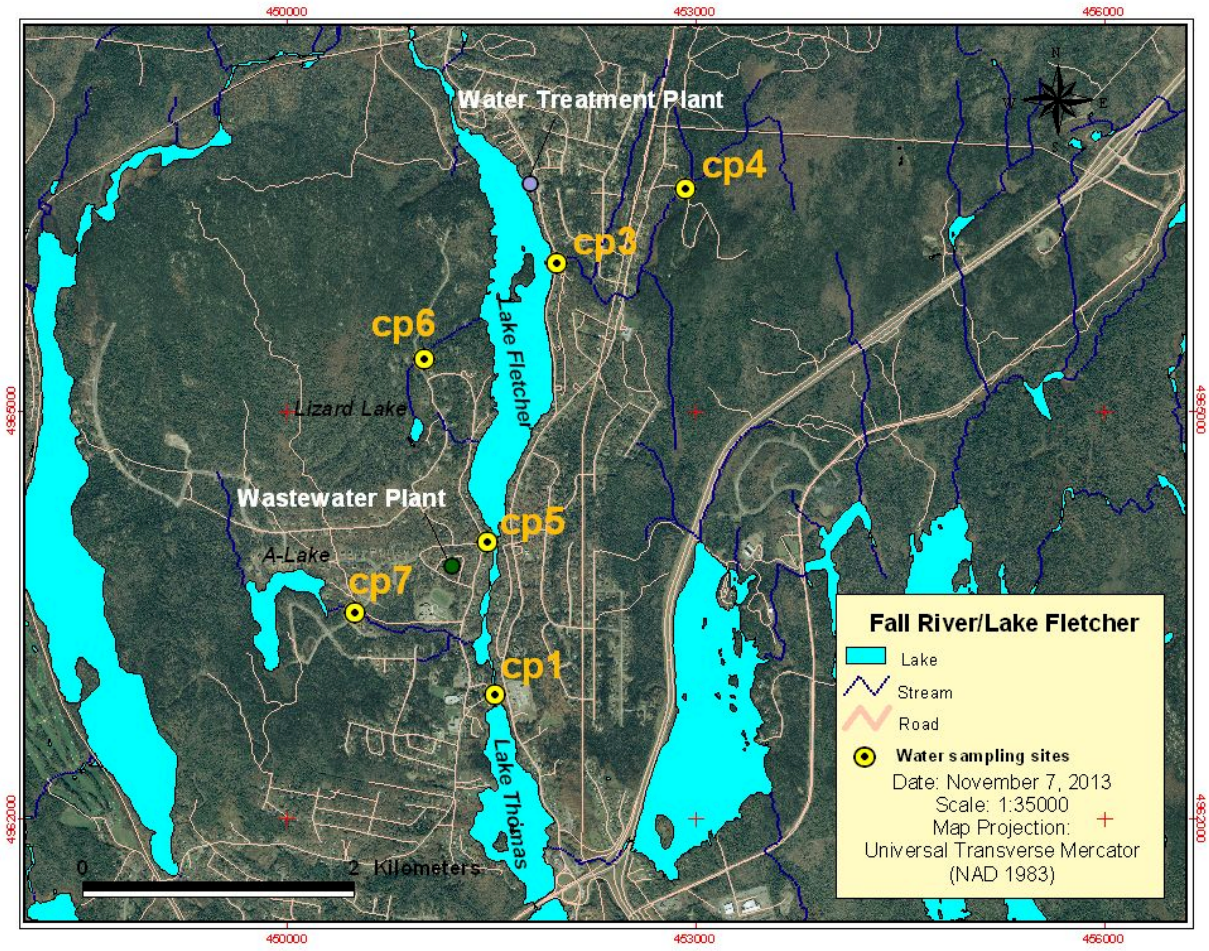

Figure 3.2. Lake Fletcher/Fall River watershed system and sampling locations (CP1-CP7)

3.3 Water Sample Collection Frequency

Water samples were collected at sampling sites monthly during the winter months (November-March) and biweekly during the spring, summer and fall. Sampling dates were sometimes adjusted to capture storm events defined as rainfall occurring in the past 24 hours with rainfall amounts over 20 mm. The sampling dates of storm events are listed below in Table 3.2.

Table 3.2 Storm events captured in MM and CP during 2012-2013 sampling periods

Sampling Date	Precipitation (mm)	Watershed and SR[a]	Watershed Storm Event
August 20th, 2012	20	MM SR #12	1
September 11th, 2012	85	MM SR #14	2
September 24th, 2012	20	MM SR #15	3
December 3rd, 2012	20	MM SR #19	4
May 14th, 2013	43	MM SR #26	5
July 2nd, 2013	20	MM SR #30	6
September 11th, 2012	85	CP SR #14	1
September 24th, 2012	20	CP SR #15	2
June 12th, 2013	20	CP SR #29	3

[a] – Sampling run

3.4 Fecal Samples (Point Source Sample Collection)

Wildlife fecal samples (Table 3.3) were used for marker validation (sensitivity/specificity) studies for the new HF183, HcyB and AcytB markers added to this project. All wildlife samples, with the exception of deer and cow fecal samples, had been collected in 2010 from the Shubenacadie Wildlife Park in Stewiacke, NS (Ridley, 2011). All fecal samples were transported on ice and returned to the lab for immediate DNA extraction using PowerSoil DNA Extraction Kit (MoBio Laboratories, Carlsbad, CA, USA) according to the manufacturer's instructions. DNA was stored at -20°C until further qPCR analysis.

Additonally, septic tank samples were collected from a home in Fall River in fall 2012. Raw sewage samples and treated effluent samples were collected from the Collin's Park wastewater treatment plant in 2012 and 2013. Septic tank, raw sewage, and effluent samples were not filtered prior to DNA extraction. These 500 mL sample bottles were shaken prior to centrifugation at 3200 x *g* for 10 minutes. Ten mL of the liquid sample were poured into a 15 mL falcon tube for centrifugation. The supernatant was decanted and 250 mg of resuspended pellet was used for DNA extraction using the PowerSoil

DNA extraction kit (MoBio). This DNA was subsequently used for human marker quantification and specificity/sensitivity studies. Three other 10 mL aliquots of the septic, raw sewage, and effluent samples were centrifuged using the same method described above in order to obtain 250 mg pellets that were then spiked into enrichment broths for pathogens (buffered peptone water [BPW], Bolton Broth [BB], and Listeria Enrichment Broth [LEB] for *Salmonella*/*E. coli* O157:H7, *Campylobacter* spp. and *Listeria* spp., respectively). The selective enrichment and DNA extraction protocol were then completed as described below in Section 3.6.

Table 3.3 Sewage and fecal samples collected for marker validation and sensitivity/specificity studies

Source(s)	**Samples**	**Number of Samples**
Collin's Park, Truro, NS	Septic tank	CP (3), Truro (5)
Collin's Park, NS	Raw sewage	6
Collin's Park, NS	Treated effluent	3
Halifax, NS	Human feces	10
Middle Musquodoboit, NS	Cow feces	24
Shubenacadie Wildlife Park, NS	Deer feces	4
Shubenacadie Wildlife Park, NS	Pig feces	3
Shubenacadie Wildlife Park, NS	Horse feces	2
Shubenacadie Wildlife Park, NS	Beaver feces	1
Shubenacadie Wildlife Park, NS	Bobcat feces	1
Shubenacadie Wildlife Park, NS	Chicken feces	1
Shubenacadie Wildlife Park, NS	Coyote feces	1
Halifax, NS	Dog feces	2
Shubenacadie Wildlife Park, NS	Fisher feces	1
Shubenacadie Wildlife Park, NS	Porcupine feces	1
Shubenacadie Wildlife Park, NS	Raccoon feces	1
Shubenacadie Wildlife Park, NS	Red fox feces	1
Shubenacadie Wildlife Park, NS	River otter feces	1
Shubenacadie Wildlife Park, NS	Skunk feces	1
Shubenacadie Wildlife Park, NS	Snowshoe hare feces	1

3.5 WATER QUALITY PARAMETERS

Several physical and chemical water quality parameters were measured on site at the time of sampling. Also, flow was determined monthly in the summer for specific sites (MM1 and CP3). These parameters, described in Table 3.4, are widely used in water research and may impact microbiological findings at the MM and CP Sites. In addition to on-site measurements (see below) of temperature, dissolved oxygen, pH and conductivity, total suspended solids (TSS) and turbidity were determines in the water samples as part of the general water processing protocol (Figure 3.3).

Table 3.4 Water quality and hydrologic parameters measured in the study

Parameter	Unit
Temperature	Celsius
Dissolved oxygen	mg/L and %
pH	n/a
Electrical conductivity	mS/cm
Total suspended solids (TSS)	grams
Turbidity	NTU
Flow	$meters^3/sec$
Stage measurement	meters
Precipitation	mm

Using a handheld 600R Sonde device (YSI, Yellow Springs, OH, USA), temperature (C), dissolved oxygen (DO) (mg/L and %), pH, and electrical conductivity (mS/cm) were measured and recorded at each sampling site during each sampling run. The Sonde device was calibrated for pH monthly and calibrated for DO in water saturated with ambient air before each sampling day's first measurement. The Sonde was placed downstream of water sampling locations and measurements were recorded only after they had stabilized.

In the absence of a weather station in each watershed, precipitation data was estimated using records from the Halifax International Airport, located 20 kilometers from Lake Fletcher and 45 kilometers from Middle Musquodoboit.

3.5.1 Flow Measurements

Stream flow was continuously determined at MM1 and CP3 using the stage-discharge method developed from manual flow measurements obtained using the velocity-area integration method. During summer 2012, manual base flow measurements were recorded. During late summer 2012 and summer 2013, manual flow measurements were additionally taken during storm events. Pressure transducers were installed at MM1 and CP3. The flowtracker current meter (SonTek/YSI. San Diego, CA, USA) or pygmy current meter (Gurley Precision Instruments, NY, USA) was used to measure velocity. Flow velocities were measured along with corresponding water depths at 5-10 intervals across the channel depending on the width of the channel. Stage measurements (the measurement from the top of the water level to a marked line near the shore) were taken from predetermined objects such as tops of rocks or bridges at each location to measure water level. These measurements, along with time of stage measurement collection, were related to concurrent pressure transducer measurements to develop stage-discharge relationships

3.5.2 Total Suspended Solids

Five hundred mL of sample water from each site was used to determine the total suspended solids (TSS in mg/L). A metal tin containing one pre-dried (at 103 °C) 934-

AH Whatman Glass Microfibre Filter (pore size 1.5 µm, Sigma-Aldrich, Oakville, ON, Canada) was pre-weighed per sample. The sample volume (500 mL) was filtered then the filter was folded and transferred to the metal tin. Once all samples were processed, all tins and filters were dried at 103 °C for 24 hours. Tins and filters were then weighed again and TSS was calculated.

3.5.3 Turbidity

Approximately 20 mL of water sample was loaded into the sample cell of a 2100AN IS Laboratory Turbidimeter (ISO, 115Vac, Hach, Mississauga, ON, Canada). Three readings of signal average in NTU were recorded per sample. A blank sample containing only distilled water was also measured and recorded.

3.6 WATER PROCESSING PROTOCOL

3.6.1 Water Filtration and Enrichment of Samples

At each sampling site, four litres of water were collected in sterile (washed and rinsed in 70% ethanol) sample bottles. The water samples were obtained by first rinsing the sterile bottles three times in the river/stream/lake water by letting the water bottle fill just below the water surface and then letting it fill one final time. Water from the sampling sites was processed in order MM1-MM5 followed by CP1-CP7.

Figure 3.3 presents the workflow in processing the water samples. Briefly, aliquots of 500 mL of sample water were filtered through a 0.45 µm pore size membrane filter (Millipore, Billerica, MA, USA). A Millipore filtration system with sterile plastic cups (Millipore, Billerica, MA, USA) was used to filter water. Tweezers were flame

sterilized using 100% ethanol and the metal Millipore filtration equipment was sterilized using 70% ethanol prior to beginning processing the samples and between each sampling site. A total of four aliquots were filtered onto separate filters and placed into enrichment broths or sample water as detailed in Table 3.5

Table 3.5 Pathogen and marker primary enrichment methods

Amount Filtered (mL)	**1st Enrichment Step (filter placement)**	**Target**	**Incubation temperature (ºC) and time (hours)**	**2nd Enrichment Step**
500	15 mL tube containing 10 mL sample site water	*Bacteroidales*/mtDNA quantification	4, 48	No
500	50 mL falcon tube containing 40 mL of supplemented Bolton Broth (Oxoid) containing 5% (v/v Laked Horse Blood (Oxoid)	*Campylobacter* spp.	42, 48 BD GasPak EZ Campy Container System, anaerobic environment	No
500	Glass tube containing 10 mL of Listeria Enrichment Broth (Oxoid)	*Listeria* spp.	37, 24	Yes – refer to Section 3.6.4
500	250 mL flask containing 100 mL of buffered peptone water (BD-Difco)	*E. coli* O157:H7, *Salmonella* spp.	37, 24	Yes – refer to Section 3.6.5

3.6.2 Sampling Controls

Aliquots of autoclaved distilled water in increments of 500 mL water were also aseptically filtered onto four 0.45 µm membrane size filters. These filters were added as

negative filter controls to each enrichment broth (BB, BPW, and LEB) using the same protocol described above and incubated along with samples. One filter was incubated at 4 ºC in 10 mL autoclaved distilled water along with samples for *Bacteroidale*s/mtDNA markers. In addition for each of the pathogen enrichments, one culture tube was incubated along with the samples containing only the media (BB, BPW, or LEB). During the DNA extraction step, a negative extraction control (250 µL of sterile water) was also processed, along with all other negative controls (filter, media, and for *E. coli* O157:H7 and *Salmonella* spp. immunomagnetic separation bead controls [see section 3.6.5]).

A sampling (field) control was also processed in summer 2013 to determine if samples were being contaminated during the sample water collection process itself. A batch of autoclaved distilled water was prepared and added in the lab prior to sample collection to two 2 L and two 500 mL sample bottles, which had been sterilized with ethanol using the standard operating procedure. The sample control bottles were closed, brought to the field, then returned to the lab and processed along with the other samples.

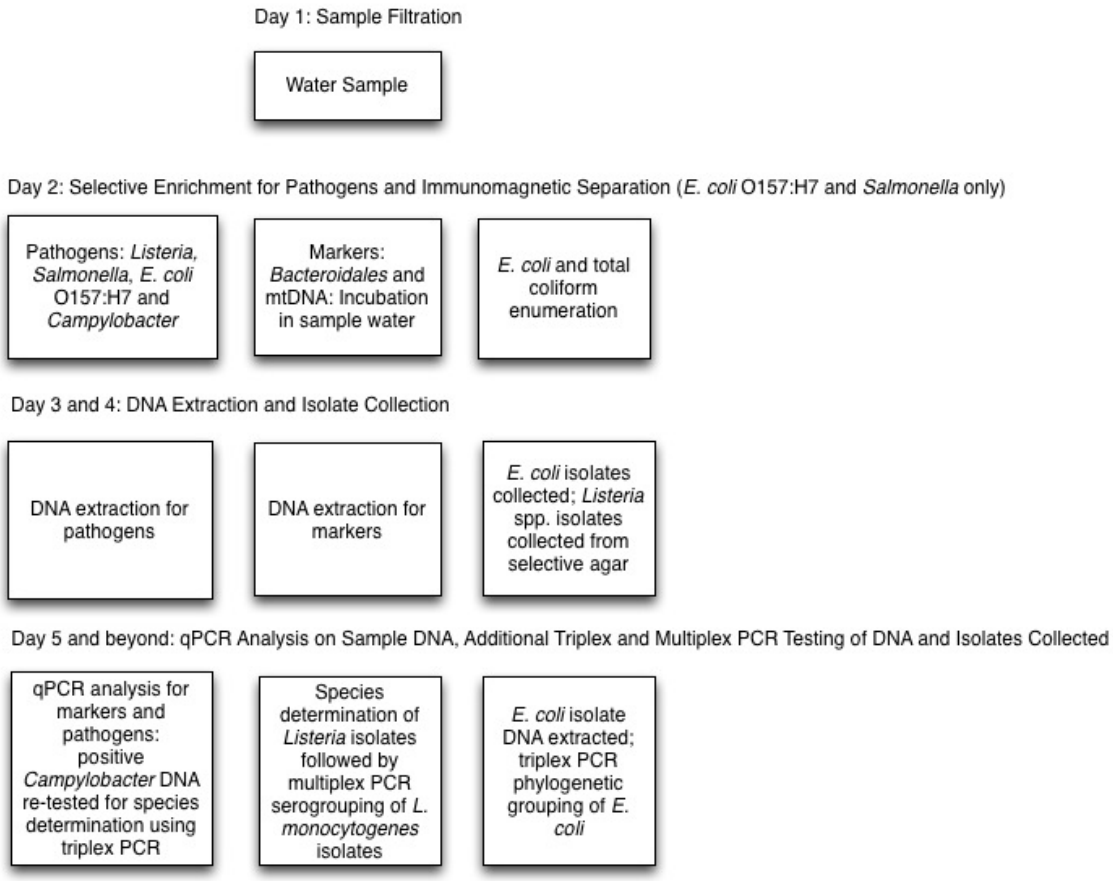

Figure 3.3 Workflow for the microbiological analysis of bacterial pathogens and *Bacteroidales* and mtDNA markers in water

3.6.3 *E. coli* and Total Coliforms Enumeration and Isolate Collection

E. coli and total coliforms were cultivated and enumerated using m-Coliblue broth (Millipore, Billerica, MA, USA). One hundred mL of sample water from each site was filtered through a 0.45 µm pore size membrane as described above. Filters were transferred aseptically to the absorbent pad containing 2 mL of m-Coliblue broth on the 37 mm petri dish (Millipore). Suitable 10-fold dilutions were made in 100 mL of phosphate buffered saline (PBS), filtered and plated. As a negative control, 100 mL of

sterile PBS was filtered along with the samples. Petri dishes were incubated at 37 °C for 24 hours. Blue colonies were counted as *E. coli* and the total number of red and blue colonies represented total coliforms. Results were converted to colony-forming units (CFU) in 100 mL sample water. Five *E. coli* colonies were aseptically removed using tweezers from each petri dish containing *E. coli* and incubated in 1 mL Brain Heart Infusion (BHI) broth (Neogen, Lansing, MI, USA) for 24 hours at 37 °C. Isolates were frozen at -20 °C in 20% glycerol (Fisher Scientific) for further analysis and grouping using triplex PCR.

3.6.4 *Listeria* spp. Enrichment

After 24 hours incubation, 1 mL from each tube (sample) of LEB enrichment broth was transferred to 10 mL Fraser (Oxoid) broth and incubated another 24 hours at 37°C. Another tube containing only media (Fraser) was also incubated along with the samples. Using the streak plate method, each sample of Fraser broth after incubation was transferred to PALCAM (Oxoid) agar for the isolation of *Listeria* species. PALCAM plates were incubated at 35 °C for 48 hours. After incubation, up to 8 colonies were collected from each sample (PALCAM plate) and incubated at 37 °C for 48 hours in 1 mL of BHI broth. Isolates were frozen in 20% glycerol at -80 °C after incubation. Isolates were further tested on Rapid Lmono Agar (Bio-Rad, Hercules, CA, USA) according to the manufacturer's instructions to identify the isolates to the species level. Presumptive *L. monocytogenes* isolates were further characterized using a multiplex PCR method (see section 3.7.4).

3.6.5 Immunomagnetic Separation of *Salmonella* spp. and *E. coli* O157:H7

All samples of BPW from Day 1 (including negative filter and negative media controls) were subjected to an immunomagnetic separation protocol, enriching for *E. coli* O157:H7 and *Salmonella* spp. separately. Following vortexing, *E. coli* O157:H7 and *Salmonella* were separated from the enrichment broth (BPW) using paramagnetic beads (Dynabeads anti-*E. coli* O157:H7 and Dynabeads anti-*Salmonella*, Invitrogen Life Technologies). Briefly, the following steps were carried out: Cells were pelleted from 10 mL of BPW enrichment broth by centrifugation at 3200 x *g* for 5 minutes, and resuspended in 500 µL of PBS-Tween (0.05% Tween – Fisher Scientific). This suspension was transferred to a 1.5 mL tube where 20 µL of dynabeads specific for *Salmonella* spp. or *E. coli* O157:H7 were added to each tube. Tubes were placed in the magnetic particle concentrator (MPC) (DynaMag2 Magnet, Life Technologies) without the magnetic base and rotated for 25 minutes at 1.5 rpm. After 25 minutes the magnetic base was replaced and the MPC was rotated for 3 minutes to concentrate the magnetic beads into a pellet against the side of the tube. Without disturbing the pellet, the supernatant was removed and 1 mL of PBS-T was added to the sample. The rotation and washing steps were repeated twice for each sample tube. After the final wash, the supernatant was removed and 1 mL of media from new culture tubes either containing 9 mL of RVS broth (for *Salmonella* spp.) or 9 mL of TSB (BD-Difco) with 20 ug/mL novobiocin (for *E. coli*) was used to resuspend the appropriate beads. Media and bead mixture were transferred back into the appropriate tubes and incubated overnight at 42°C.

Negative controls were added during this protocol, where 10 mL of Day 1 media control was subjected to both the *Salmonella* and *E. coli* O157:H7 separation protocols to control for any contamination originating from the beads added to the samples (negative bead control). In addition, 9 mL of Rappaport Vassiliadis Soya (RVS) Peptone Broth (Oxoid) and 9 mL of Tryptic Soy Broth (BD Difco) with 20 µg/mL novobiocin (MP Biomedicals) in separate tubes were incubated at 42 °C overnight along with other samples.

3.6.6. DNA Extractions

Pathogen enrichment broths were combined prior to DNA extraction. Two mL of each enrichment broth (BPW, TSB, RVS, Fraser and BB) were combined and centrifuged at 3200 x *g* for 10 minutes. Supernatant was removed and 250 µL of resuspended pellet was used for DNA extraction following manufacturer's instructions using the PowerSoil DNA extraction kit (MoBio Laboratories).

Bacteroidales/mtDNA sample tubes were vortexed for 3 minutes prior to DNA extraction to remove bacteria from the filter. Filters were removed from each tube aseptically and tubes were centrifuged at 3200 x *g* for 10 minutes. Supernatant was removed and 250 µL of the resuspended pellet was used for DNA extraction following manufacturer's instructions using the PowerSoil DNA extraction kit (MoBio Laboratories).

Zygem PrepGem DNA extraction kit (VWR International, Mississauga, ON, Canada) was used to extract DNA from all *E. coli* isolates from m-Coliblue plates. All DNA was stored at -20 °C until time of testing.

3.7 qPCR and PCR Methods

3.7.1 Composition of a PCR/qPCR Experiment

Each qPCR or PCR reaction contained the following components: Dnase-free water (Fisher Scientific, Ottawa, ON, Canada), master mix (Applied Biosystems Fast Advanced 2X, Applied Biosystems, Burlington, ON, Canada; Promega GoTaq qPCR 2X, Fisher Scientific, Ottawa, ON, Canada; New England Biolabs 2X, New England Biolabs, Ipswich, MA, USA), primers (Sigma-Aldrich, Tables 3.6-3.12), probes (Sigma-Aldrich, if Taqman® qPCR methodology was used, Tables 3.6-3.12) and template DNA (extracted from samples, standards, and positive controls).

3.7.2 Presence/Absence TaqMan qPCR Experiments

Presence/absence Taqman® probe-based qPCR experiments were used to determine if a sample contained one of the four pathogens of interest. The selective enrichment was carried out on filters capturing the bacteria present in 500 mL of the original sample, the presence or absence of the pathogens were theoretical determined per 500 mL sample. Positive control DNA (DNA extracted from the target pathogen of interest) was used in each presence/absence experiment for the four pathogens in this study. Control strains *Salmonella* Typhimurium (ATCC 14028), *E. coli* O157:H7 (Dalhousie culture collection), and *Listeria monocytogenes* (ATCC 19115) were grown overnight in BHI (Oxoid) at 37 °C. DNA was extracted from colonies using Zygem PrepGem DNA Extraction Kit. The Canadian Food Inspection Agency (Dartmouth, NS, Canada) kindly provided *Campylobacter jejuni, C. lari* and *C. coli* isolates. Individual colonies were grown for 48 hours at 37 °C in unsupplemented BB before DNA extraction

using MoBio Ultraclean Microbial DNA extraction kit (MoBio). Negative or no template controls (without any template DNA) were also used in every qPCR or PCR experiment containing DNA free water instead of template.

Tables 3.6-3.7 outline the qPCR protocols used for presence/absence experiments. The test result from each sample was given a value of 0 (absence or no amplification) or 1 (presence or amplification) for each pathogen of interest. Each Taqman® presence/absence qPCR experiment used a 23 µL total reaction volume with 4 µL template DNA.

Table 3.6 Primers used in presence/absence Taqman-based qPCR experiments

Primer Name	Sequence 5' to 3'	Annealing Temp ºC	Amplicon Size (bp)	Reference
EaeF	GTAAGTTACACTATAAAAGCACCGTCG			
EaeR	TCTGTGTGGATGGTAATAAATTTTTG	59	106	Ibekwe et al., 2002
EaeP	FAM-AAATGGACATAGCATCAGCATAATAGGCTTGCT-BHQ1			
InvAF	AACGTGTTTCCGTGCGTAAT			
InvAR	TCCATCAAA TTAGCGGAGGC	56	262	Cheng et al., 2008
InvAP	FAM-TGGAAGCGCTCGCATTGTGG-BHQ1			
CampF2	CACGTGCTACAATGGCATAT			
CampR2	GGCTTCATGCTCTCGAGTT	58	108	Lund et al., 2004
CampP2	FAM-CAGAGAACAATCCGAACTGGGACA-BHQ1			
HlyQF	CATGGCACCACCAGCATCT			
HlyQR	ATCCGCGTGTTTCTTTTCGA	56	64	Rodriguez-Lazaro et al., 2004
HlyQP	FAM-CGCCTGCAAGTCCTAAGACGCCA-TAMRA			

FAM – fluorescein, TAMRA -5(6) carboxytetramethylrhodamine, BHQ1 – Black hole quencher

Table 3.7 qPCR protocols and primer concentrations used in presence/absence experiments

Primer Name	qPCR Thermocycling Protocols	Final Concentration of Primer in Reaction (nM)	Target Organism/Gene
EaeF		150	
EaeR	95ºC for 6 min; 40 cycles of 95ºC for 20 sec, 55ºC for 30 sec, 72ºC for 40 sec	150	*E. coli* O157:H7, e*ae* gene
EaeP		50	
InvAF		300	
InvAR	95ºC for 6 min; 40 cycles of 95ºC for 15 sec, 60ºC for 30 sec	300	*Salmonella* spp., *invA* invasion gene
InvAP		100	
CampF2		300	
CampR2	95ºC for 6 min; 40 cycles of 95ºC for 15 sec, 60ºC for 1 min	300	*Campylobacter* spp., 16S rRNA
CampP2		100	
HlyQF		300	
HlyQR		300	
	95ºC for 10 min; 40 cycles of 95ºC for 20 sec, 56ºC for 30 sec, 72ºC for 1 min		*Listeria monocytogenes, hly* virulence gene
HlyQP		100	

3.7.3 Standard Curve qPCR Experiments

Standard curves using either Sybr Green® methodology (for HF183 assay only) or Taqman® methodology (for mtDNA and *Bacteroidales* markers) were created in order to be able to quantify the copy number of a particular marker in the samples. Extracted plasmid DNA standards containing the target sequences for the CowM2, AllBac, and BacR markers were provided by the Yost lab (Fremaux et al., 2010). Plasmid DNA standards were created for the remaining HcytB, AcytB, and HF183 markers using the TOPO TA cloning kit and its plasmid. Briefly, the PCR products were obtained from

appropriate templates (DNA extracted from human or cow feces) after use of the protocols listed in Tables 3.9-3.10. In the case of HF183, the Bac708R primer was used to create this standard. The PCR products were purified using Ultraclean Gelspin DNA purification kit (MoBio) and visualized on a 1.5% agarose gel containing GelRed (Biotium). The purified PCR product was then cloned using the Invitrogen Topo TA cloning kit PCR 2.1 following the manufacturer's instructions. Clones containing plasmids with inserts (PCR produced) were selected on SGAL kanamycin/LB agar (Sigma, Oakville, ON, Canada) in a blue-white screen according to cloning kit protocol. White *E. coli* colonies were picked from SGAL plates and subcultured in TSB with kanamycin (50 µg/ml, Sigma) overnight at 37 °C. The plasmid DNA from white colonies was extracted using Ultraclean Standard Mini Plasmid Prep kit (MoBio) and sent for sequencing at Nanuq (Genome Quebec, McGill University, Montreal, QC, Canada). The resulting sequences were trimmed using CLC Bio Workbench 6.9 software then entered into the NCBI Blast program.

The plasmid DNA standards were used to establish standard curves for each qPCR method. Essentially, 10-fold (10^{-2} to 10^{-9}) dilutions of the appropriate plasmid DNA standard were included in each assay, creating a standard curve and allowing quantification of copy numbers in sample DNA. Melt curve peaks were also analyzed for the HF183 marker (Sybr Green® chemistry). A correct melt curve peak for the HF183 assay was observed at 79.5 +/-0.5 °C. Concentrations (ng/µL of plasmid standard) was checked using an Implen Nanophotometer P330 (Westlake Village, CA, USA) each time a plasmid DNA standard was used. Copy numbers were subsequently determined for each dilution of the standard and were entered into the StepOne Software to generate a

standard curve. Copy number was calculated from concentration (in ng) of the plasmid DNA standard using the following equation:

Equation 3.7:

$$Copy\ number = \frac{(amount(ng) \times 6.022 \times 10^{23}\ number/mole)}{[length(bp) \times 10^{9}\ ng/g \times 650 g/mole(bp)]}$$

Each Taqman® based experiment used a 23 µL total reaction volume including 4 µL template DNA. The HF183 Sybr Green® assay used a 20 µL total reaction volume with 2 µL template DNA.

Base pair lengths of the plasmid DNA standards are listed below in Table 3.8:

Table 3.8 Total base pair lengths of plasmid standards used in qPCR standard curve experiments

Plasmid Standard	Length of plasmid + insert in bp
AllBac	4037
BacR	4024
CowM2	4023
AcytB	4056
HcytB	4056
HF183	4431

Table 3.9 Primers used for standard curve marker quantification experiments

Primer Name	Sequence 5' to 3'	Annealing Temp ºC	Amplicon Size (bp)	Reference
AllBacF	GAGAGGAAGGTCCCCCAC			
AllBacR		60	106	Layton et al., 2006
	CGCTACTTGGCTGGTTCAG			
AllBacP	FAM-CCATTGACCAA TATTCCTCACTGCTGCCT-BHQ1			
BacR F	GCGTATCCAACCTTCCCG			
BacR R		58	115	Reischer et al., 2006
	CATCCCCATCCGTTACCG			
BacR P	FAM-CTTCCGAAAGGGAGATT-BHQ1			
HF183F		59		Bernhard & Field, 2000
	ATCATGAGTTCACATGTCCG			
HF183R		53	82	Seurinck et al., 2005
	TACCCCGCCTACTATCTAATG			
Bac708 R		58	500	Bernhard & Field, 2000
	CAATCGGAGTTCTTCGTG			
CowM2 F	CGGCCAAATACTCCTGATCGT			
CowM2 R		60	92	Shanks et al., 2008
	GCTTGTTGCGTTCCTTGAGATAAT			
CowM2 P	FAM-AGGCACCTATGTCCTTTACCTCATCAACTACAGACA-BHQ1			
CowM2 IAC P	TET-TAGGAACAGGCGGCGACGA-BHQ1			
AcytB F	GCAATACACTACACATCTGACACAACAA			
AcytB R		60	125	Baker-Austin et al., 2010
	CAGATAAAAAATGATGCTCCGTTTG			
Cow Mito P	FAM-CTCCTCTGTTACCCATATCTGCCGAGACG-BHQ1			
HcytB F	CCTCCAAATCACCACAGGACTAT			
HcytB R		60	125	Baker-Austin et al., 2010
	CGTGAAGGTAGCGGATGATTC			
Human MitoP	FAM-CAATCGCCCACATCACTCGAGACGT-BHQ1			

Table 3.10 qPCR protocols and primer concentrations used for marker quantification

Primer Name	qPCR Thermocycling Protocols	Final Concentration of Primer in Reaction (nM)	Target Organism/Gene
AllBacF		300	
AllBacR	95ºC for 10 min; 40 cycles of 95ºC for 30 sec, 60ºC for 45 sec	300	All *Bacteroides*, 16S rRNA
AllBacP		100	
BacR F		50	
BacR R	95ºC for 6 min; 50 cycles of 95ºC for 15 sec, 55ºC for 15 sec, 70ºC for 45 sec	250	Ruminant *Bacteroidales*, 16S rRNA
BacR P		50	
HF183F	95ºC for 10 min; 40 cycles of 95ºC for 30 sec, 53ºC for 1 min, 60ºC for 1 min; melt curve analysis: 60ºC to 95ºC at 0.4ºC per minute	300	
HF183R		300	Human *Bacteroidales*, 16S rRNA
Bac708 R	95ºC for 4 min; 35 cycles of 95ºC for 30 sec, 58ºC for 1 min, 72ºC for 2 min; 72ºC for 10 min	300	Human *Bacteroidales*, 16S rRNA
CowM2 F		500	
CowM2 R	95ºC for 6 min; 50 cycles of 95ºC for 15 sec, 60ºC for 1 min	500	Cow *Bacteroidales*, host-bacterium interaction protein
CowM2 P		40	
CowM2 IAC P		40	
AcytB F		300	
AcytB R	95ºC for 10 min; 40 cycles of 95ºC for 15 sec, 60ºC for 1 min	300	Consensus (pig, sheep, cow) mtDNA, *cytochrome B*
Cow Mito P		100	Cow mtDNA, *cytochrome B*
HcytB F		300	
HcytB R	95ºC for 10 min; 40 cycles of 95ºC for 15 sec, 60ºC for 1 min	300	Human mtDNA, *cytochrome B*
Human MitoP		100	

3.7.4 Multiplex PCR Experiments for Further Characterization of Pathogens

E. coli triplex PCR was used to determine the phylogenetic grouping of *E. coli* isolates collected from water samples (Clermont et al., 2000). *Listeria* multiplex PCR was used to determine serogroup of *L. monocytogenes* isolates (Doumith et al., 2004). Grouping provided greater characterization of the bacteria found in the watershed and also provided insight into the source(s) of these bacteria in MM and CP watersheds. In addition, a triplex PCR was also run to further characterize *Campylobacter* spp. pathogen DNA first confirmed with Taqman® qPCR, which is a method developed by Lund et al. (2004) to detect 6 species of *Campylobacter*. A triplex PCR method (Khan & Edge, 2007) was then used to determine if the three main pathogenic species of *Campylobacter* were present. *C. jejuni, C. lari* and *C. coli* positive control band patterns were compared with unknown samples in *Campylobacter* triplex runs. The absence of amplification products in the triplex PCR assay indicated that other *Campylobacter* spp. such as *C. upsaliensis, C. hyoinstensalis,* or *C. helveticus* were present in the sample.

PCR experiments were conducted as laid out in the protocols listed in Tables 3.11-3.12. Each PCR reaction volume used a total volume of 25 µL. Colony PCR was used for *L. monocytogenes* serogrouping. Positive and negative controls were used in every PCR run. DNA from *E. coli* 25922 (ATCC), a member of the B2 group, was used a positive control for the *E. coli* triplex. PCR products were detected in 1.5% agarose (Sigma Aldrich) gels using GelRed (Biotium) under a UV light. Gels were run at 90V for 45 minutes using a BioRad PowerPac Basic Supply unit.

Table 3.11. Primers used for multiplex PCR experiments

Organism	Primer Name	Sequence 5' to 3'	Annealing Temp (ºC)	Amplicon Size (bp)	Reference
Campylobacter	J-UP	CTTAGATTTATTTTTATCTTTAACT			
	J-DN	ACTAAATGATTTAGTCTCA	46	349	Khan & Edge, 2007
	L-UP	CTTACTTTAGGTTTTAAGACC			
	L-DN	CAATAAAACCTTACTATCTC	46	279	
	C-UP	GAAGTATCAATCTTAAAAAGATAA			
	C-DN	AAATATATACTTGCTTTAGATT	46	72	
E. coli	ChuF	GACGAACCAACGGTCAGGAT			
	ChuR	TGCCGCCAGTACCAAAGACA	55	279	Clermont et al., 2000
	TspF	GAGTAATGTCGGGGCATTCA			
	TspR	CGCGCCAACAAAGTATTACG	55	152	
	YjaF	TGAAGTGTCAGGAGACGCTG			
	YjaR	ATGGAGAATGCGTTCCTCAAC	55	211	
L. monocytogenes	Lmo737F	AGGGCTTCAAGGACTTACCC			
	Lmo737R	ACGATTTCTGCTTGCCATTC	53	691	Doumith, et al., 2004
	Lmo1118 F	AGGGGTCTTAAATCCTGGAA			
	Lmo1118 R	CGGCTTGTTCGGCATACTTA	53	906	
	Orf2819F	AGCAAAATGCCAAAACTCGT			
	Orf2819R	CATCACTAAAGCCTCCCATTG	53	471	
	Orf2110F	AGTGGACAATTGATTGGTGAA			
	Orf2110R	CATCCATCCCTTACTTTGGAC	53	597	
	LIP1 (*prfA*F)	GATACAGAAACATCGGTTGGC			Dagostino et al., 2004
	LIP2a (*prfA*R)	GTGTAATCTTGATGCCATCAGG	53	274	Kérouanton et al., 2010

Table 3.12. PCR thermocycling programs and primer concentrations for multiplex PCR

Primer Name	PCR Thermocycling Protocol	Final Concentration of Primer in Reaction (nM)	Target Organism/Gene
J-UP		200	
J-DN	95ºC for 3 min; 35 cycles of 95ºC for 30 sec, 47.2ºC for 30 sec, 68ºC for 45 sec; 68ºC for 5 min	200	*C. jejuni* 16S–23S rDNA internal transcribed spacer
L-UP		200	
L-DN		200	*C. lari* 16S–23S rDNA internal transcribed spacer
C-UP		200	
C-DN		200	*C. coli* 16S–23S rDNA internal transcribed spacer
ChuF		200	
ChuR	95ºC for 4 min; 31 cycles of 95ºC for 30 sec, 55ºC for 30 sec, 68ºC for 45 sec; 68ºC for 5 min	200	*E. coli chuA* heme transport gene
TspF		200	
TspR		200	*E. coli* DNA fragment TSPE4.C2
YjaF		290	
YjaR		290	*E. coli yjaA* gene
Lmo737F		400	
Lmo737R	94ºC for 3 min; 40 cycles of 94ºC for 40 sec, 53ºC for 45 sec, 72ºC for 1min 15 sec; 72ºC for 7 min	400	*L. monocytogenes* serotypes 1/2a, 1/2c, 3a, 3c, *lmo0737* gene
Lmo1118F		400	
Lmo1118R		400	*L. monocytogenes* serotypes 1/2c and 3c, *lmo1118* gene
Orf2819F		400	
Orf2819R		400	*L. monocytogenes* serotypes 1/2b, 3b, 4b, *orf2819* gene
Orf2110F		400	
Orf2110R		400	*L. monocytogenes* serotypes 4b, 4d, 4e, *orf2110* gene
LIP1(*prfA* F)		200	
LIP2a(*prfA* R)		200	All *L. monocytogenes* serotypes, *prfA* gene

3.8 Sensitivity and Specificity Study

The sensitivity and specificity of the HF183, HcyB, and AcytB assays were determined by testing the assay primer sets against fecal sample DNA sourced from various animals and humans in Nova Scotia (Table 3.3). Table 3.10 outlines qPCR methodologies used to run the tests. Sensitivity is defined as the ability of the assay to detect the target. Specificity, in contrast, is the ability of the assay to discriminate between the target and other species.

Equation 3.2: $Sensitivity = a/ (a + c)$

Equation 3.3: $Specificity = d/ (b + d)$

a is the true positive (samples tested correctly for the marker of its own species) and *b* is the false positive (samples tested positively incorrectly for the marker of another species). *c* is the false negative (samples tested negatively incorrectly) while *d* is the true negative (samples tested negatively correctly) (Fremaux et al., 2009).

3.9 HCYTB and HF183 Marker Limit of Detection/Recovery Study

A limit of detection study was conducted using water samples from CP1, CP6, MM1, and MM5 during late summer 2012. These samples sites were chosen as representative samples of the variety of water from each watershed (one control site and one main body of water sampled). Fifty mL of raw sewage was filtered and used as the 10^0 standard. Four sets of 10-fold dilution tubes (10^{-1} to 10^{-8}) were prepared using sterile saline with the initial dilution consisting of 0.5 mL of raw sewage added to 4.5 mL of saline. Each 5 mL dilution tube was then added into 500 mL of sample water. One sample from each site was not seeded with any sewage, and therefore acted as the

background sample. The sample water was filtered according to the water processing protocol (*Bacteroidales*/mtDNA protocol) and DNA extraction protocols. The presence of the markers was tested using the qPCR methods described in Section 3.7.3. Marker copy number was calculated in the raw sewage standard and in each dilution tube. Overall recovery was also determined by comparing theoretical yield of copy number of marker seeded into the sample with actual yield of the marker in the sample in copies/100mL.

3.10 Statistical Analysis

SigmaPlot© software was used to create graphs of the data. SYSTAT© and R© software was used to perform statistical tests on the data. An overview of minimum and maximum concentrations of host-associated marker was provided by site. Time-series plots of *E. coli* concentrations, precipitation, flow and marker concentrations by site were created to visualize and understand trends over the entire sampling period.

Normality of data was evaluated using the Anderson-Darling and Shapiro-Wilk normality tests. Correlation analysis of *E. coli* counts with marker concentrations and variables such as turbidity and precipitation was completed using the Spearman rank-order method to determine if two variables were associated.

In addition, logistic regression of pathogen data with *E. coli* followed by other predictors such as *Bacteroidales* markers was performed. Logistic regression models the relationship between a binary response variable (in this case pathogen presence/absence) and one or more explanatory continuous variables (such as *E. coli*). Logistic regression uses the logit [ln (0) for pathogen absence, ln (1) for presence] of the odds as the response

variable. Odds are the ratio of the probability of occurrence of an event to that of non-occurrence. The odds ratio compares the probability of pathogen presence given *E. coli* detection with the probability of pathogen presence given no *E. coli* detection. In this study, the odds ratio was the most useful measure obtained from logistic regression. The odds ratio estimates the risk of pathogen presence (in this scenario) given *E. coli* as a risk factor. It also measures the strength of association between *E. coli* (or other predictors) and pathogen presence. Strong odds ration are above 3, while weak odds ratios are between 1-1.5. Logistic regression results were obtained using data from each watershed to determine if *E. coli*-pathogen relationships differ between watersheds.

A scatter plot matrix was used as a starting point to visualize the relationships between continuous variables such as *E. coli*, markers, and other environmental variables measured in this study. Multivariate analysis of *E. coli* concentrations, marker concentrations and environmental variables such as TSS, temperature and turbidity was completed in R© statistical software (http://www.r-project.org). Principal Components Analysis (PCA) was chosen in R© (statistical software) to analyze the multivariate data set to decrease redundancy in the dataset, to simplify the dataset by filtering noise, and to understand the underlying structure of the dataset. The important assumption of multivariate normality was tested in R prior to analysis of the correlation matrix.

Principal components are the principal axes ordered by the amount of variability that they account for in the original dataset. Coefficients are elements of eigenvectors (column vectors) and relate original variables to components. Scores are values of units on components. The amount of variance accounted for by a component is given by the

eigenvalue. Finally, the loading of the variable on a component is the correlation between the variable and the component. The loading explains how well each of the original variables relates to the new variables. First, the eigenvectors and eigenvalues were found. The number of components to use was decided based on the Scree diagram. The loadings were then examined and the scores were plotted on the bi-plot diagram

CHAPTER 4 RESULTS

4.1 OVERVIEW

Water samples were collected from five sites in Middle Musquodoboit (MM) and five sites in Collin's Park/Lake Fletcher (CP) between January 2012 and August 2013. Each watershed was sampled over an eighteen months period in 33 sampling runs resulting in the collection of 333 water samples. Six sampling runs in MM occurred during storm event conditions, while in Collin's Park/Lake Fletcher, 3 sampling runs took place during storm events. CP7 was added to the sampling scheme during the summer of 2013 at the request of Halifax Water to determine if the nearby new subdivision is impacting water quality. CP7 was sampled four times in total. All samples were tested by qPCR for the presence of four pathogens and concentration of MST markers. Samples that tested positive for *Campylobacter* spp. were analyzed by triplex PCR to determine if *C. jejuni, C. lari* and/or *C. coli* were present.

All water samples were also tested for the presence of *Listeria* spp. using a selective enrichment/agar approach. Isolates (1322) were collected and identified to the species level using the indicative Rapid Lmono agar. *L. monocytogenes* isolates (294) were then further classified into serogroups by triplex PCR. At the time of collection, water samples were tested for a number of common water quality parameters and concentrations of the fecal indicator organism *E. coli* and total coliform. *E. coli* isolates (727) were collected from water samples during 21 sampling runs between July 2012 and July 2013.

4.2 Marker Sensitivity and Specificity Study

Validity of using the HF183 (human *Bacterioidales*), HcytB (human mtDNA), and AcytB (bovine mtDNA) markers in Nova Scotia, where the markers had not previously been applied was first investigated. The sensitivity and specificity of the HF183, HcyB, and AcytB qPCR assays were determined by testing the assay primer sets against fecal sample DNA sourced from various fecal samples obtained from animals and humans in Nova Scotia (Table 3.3). Ridley (2011) had in a previous study validated AllBac, BacR, and CowM2 *Bacteroidales* markers for use in Nova Scotia. The sensitivity of these markers had been found to be 97.3%, 100% and 94.4%, respectively, with associated specificities of 93.9%, and 88.9%, 100% (Ridley et al., 2011). Results from the present HF183 and HcytB validation study revealed high sensitivity of these human markers (Table 4.1). In contrast, the cow cytochrome B marker (AcytB) showed slightly lower sensitivity (Table 4.1) with one false negative occurring in cow manure, and lower specificity with false positives results obtained from deer, fisher, fox, and river otter. One false positive out of 51 samples, was found in the HcytB study for cow manure, decreasing specificity of this marker. Four false positives (fisher, rabbit, raccoon and porcupine) occurred in the HF183 study, where 22 human and 30 non-human samples were analyzed, resulting in a specificity of 86.7% (Table 4.1).

Table 4.1 Sensitivity and specificity of HcytB, AcytB, and HF183 markers

Marker	**Sensitivity (%)**	**Specificity (%)**
Cow cytochrome B (AcytB)	90.9	78.9
Human cytochrome B (HcytB)	100.0	96.4
HF183 human *Bacteroidales*	100.0	86.7

4.3 Point Source Sample Marker Concentration

As part of HF183, HcytB and AcytB marker validation, fecal samples from respective hosts and point sources such as septic tanks, were analyzed for marker concentration by qPCR. The HF183 marker was found in large quantities (mean 8.28 x 10^{10} copies/g) in human feces and in raw sewage (mean 3.52 x 10^{11} copies/g) (Table 4.2). HcytB, in contrast, was found in lower quantities in human feces (mean 1.28 x 10^{7} copies/g) and raw sewage (mean 2.44 x 10^{8} copies/g). HcytB was absent in treated wastewater effluent (Table 4.2). Concentrations of AcytB were 2-4 orders of magnitude lower than CowM2 and BacR *Bacteroidales* markers in cow feces (Table 4.2).

Table 4.2 Human and cow marker concentration in point source samples per gram of wet weight. Detection limits were 9 copies/g for the HF183/HcytB markers and 1 copy/g for all other markers.

Markers	**Human**		**Bovine**		
Sample	**HcytB (copies/g)**	**HF813 (copies/g)**	**BacR (copies/g)**	**CowM2 (copies/g)**	**AcytB (copies/g)**
Human feces	1.28 x 10^{7}[a]	8.28 x 10^{10}	ND[b]	ND	ND
Septic	9.66 x 10^{6}	9.20 x 10^{10}	ND	ND	ND
CP raw sewage	2.44 x 10^{8}	3.52 x 10^{11}	ND	ND	ND
CP treated WW effluent	ND	8.86 x 10^{6}	ND	ND	ND
MM cow feces	ND	ND	2.03 x 10^{10}	1.71 x 10^{8}	5.73 x 10^{6}

[a] - Geomean
[b] - ND= None detected (no amplification or Ct value)

4.4 HF183 AND HCYTB LIMIT OF DETECTION AND RECOVERY STUDY

A limit of detection (LOD) is defined as the lowest concentration of the marker that can be accurately measured under assay conditions and was determined for the HF183 and HcytB markers. Linear relationships for sample Ct values/ log concentrations of human markers in the ten-fold dilution series were found in both assays (R^2=1). The LOD for the assays were 925 and 917 copies/100 mL for HcytB and HF183, respectively. An LOD of 100 copies/100 mL was previously determined by the Yost lab following the method outlined in Fremaux et al. (2010) and used for the other *Bacteroidales* markers in this study. Spiking experiments showed that the overall marker recovery of 53-60% from the filtration step to DNA extraction was similar in both assays.

4.5 MARKER DETECTION IN MM AND CP WATER SAMPLES

4.5.1 AllBac (Universal) Marker Detection

AllBac, the universal marker for presence of *Bacteroidales* in water samples, was detected in 97.9% of the 333 water samples. Concentrations of AllBac ranged from the LOD to 10^5 copies/100 mL (Table 4.3). CP6 showed an overall lower maximum concentration of AllBac (6.00 x 10^4 copies/100 mL). Seven samples out of the 333 samples contained less than 100 copies per 100 mL due to the absence of Ct values.

Table 4.3 Source-specific *Bacteroidales* and mtDNA marker concentrations in MM and CP water samples. Weak positive signals (WPS) are followed by frequency of occurrence.

Site **Min** ***Max***	**AllBac** **(copies/** **100 mL)**	**BacR** **(copies/** **100 mL)**	**CowM2** **(copies/** **100 mL)**	**HF183** **(copies/** **100 mL**	**HcytB** **(copies/** **100 mL)**	**AcytB** **(copies/** **100 mL)**
MM1	1.26×10^2 ***6.02×10^5***	1.28×10^4 WPS[a] (1)	WPS (2)	6.51×10^3 ***2.77×10^5***	3.29×10^3 ***2.79×10^5***	WPS (1)
MM2	6.28×10^2 ***2.89×10^5***	7.68×10^3 WPS (3)	WPS (2)	6.72×10^3 ***1.58×10^5***		WPS (1)
MM3	4.08×10^2 ***3.49×10^5***	5.04×10^3 ***7.97×10^3***	WPS (4)	1.86×10^4 ***6.27×10^5***	2.41×10^3 ***6.31×10^5***	WPS (2)
MM4	4.86×10^2 ***2.91×10^5***	7.91×10^4	WPS (2)	3.78×10^4 ***6.43×10^4***	1.15×10^3 ***6.99×10^5***	WPS (1)
MM5	1.51×10^2 ***4.18×10^5***	WPS (1)		8.74×10^4 ***2.35×10^5***	1.66×10^5	WPS (1)
CP1	1.37×10^2 ***7.80×10^5***			1.67×10^4 ***1.92×10^5***		
CP3	1.10×10^2 ***1.05×10^5***			1.37×10^4 ***1.54×10^5***		
CP4	1.31×10^2 ***5.93×10^5***					
CP5	4.99×10^2 ***3.70×10^5***			1.26×10^4 ***6.39×10^5***	1.40×10^3	
CP6	1.05×10^2 ***6.00×10^4***			1.37×10^4 ***2.11×10^5***	4.58×10^5	
CP7[b]	2.51×10^3 ***4.81×10^3***					

[a]Samples with Weak Positive Signals (WPS) contained target concentrations below the detection limit due to Ct values being higher than 34. These samples were, however, still considered borderline positives.
[b]CP7 was sampled 4 times only

4.5.2 Human (HF183 and HcytB) Marker Detection

Human markers were detected in both watersheds throughout the sampling season. All sites, with the exception of CP4 and CP7, were positive for human markers during at least one sampling event. At sites positive for human markers, both HF183 and HcytB were detected occasionally throughout 2012-2013, following no seasonal trend with detection during the spring, summer, fall and winter. HF183 was detected 3-6 times at positive sites. HF183 was detected 6 times in total at MM3 and CP6. HcytB was detected 1-5 times at positive sites. HcytB was detected at MM3 five times in total, while it was only detected once at CP5 and CP6 (Table 4.3). When detected, both HF183 and HcytB were found in larger quantities in the order of 10^3 to 10^5 copies/100 mL (Table 4.3). In water samples positive for human markers, HF183 and HcytB were detected together in only 14.5% of the samples.

4.5.3 Ruminant and Bovine (BacR, CowM2 and AcytB) Marker Detection

BacR was detected in the Middle Musquodoboit watershed only. Weak Positive Signals (WPS) of CowM2 and AcytB, below the detection limit of 100 copies/100 mL, were also found in the Middle Musquodoboit watershed (Table 4.3). Concentrations of BacR varied from 5.0×10^3 to 8.0×10^4 copies/100 mL at sites MM1-MM4 (Table 3.4). BacR was detected 1-4 times at each positive site (including WPS). CowM2 was detected 1-4 times as WPS only at positive sites, and AcytB was found less frequently (1-2 times) at positive sites in MM (Table 4.3). BacR, CowM2 and AcytB were detected randomly throughout the winter, spring and summer of 2012-2013, showing no seasonal patterns.

In water samples positive for cow markers, AcytB and CowM2 co-occurred in 15.4% of the samples.

4.6 Occurrence of Bacterial Pathogens in the MM and CP Watersheds

4.6.1 Pathogen Detection in MM and CP Water Samples

As a consequence of the selective enrichment process, pathogens were detected as being present or absent in 500 mL (or 1 g for fecal samples) of filtered water and point source samples. Each site, other than CP6 and CP7, was sampled 33 times in total (Table 4.4). CP6 was ice-covered in January 2012, resulting in 32 sampling runs for this site only. CP7 was sampled four times in total during the summer of 2013.

Analyses by qPCR revealed high detection rates of *Campylobacter* spp. and *Listeria monocytogenes* in both the MM and CP watersheds (Table 4.4). MM1 and CP5 showed the highest rates of *Campylobacter* detection (90% and 87.8%, respectively), while MM2 showed the highest rate of *L. monocytogenes* detection (81.8%) (Table 4.4). Higher rates of *L. monocytogenes* were detected in the Middle Musquodoboit watershed. Rates of *Salmonella* detection were consistent across sites, with MM1 and CP5 showing the highest percentages (30.3% and 33.3%) of positive *Salmonella* samples. *E. coli* O157:H7 was more prevalent in the MM watershed with the exception of MM5, where *E. coli* O157:H7 was not detected in any of the 500 mL sample volumes filtered. The highest percentage of *E. coli* O157:H7 was detected at MM4 (Table 3.5). *E. coli* O157:H7 was not detected at CP4, CP6 and CP7.

Table 4.4 Pathogen detection by Taqman qPCR methods in MM and CP water sampling events

Site	*Campylobacter* spp. (%)	*Salmonella* (%)	*E.coli* O157:H7 (%)	*Listeria monocytogenes* (%)
MM1	90.0 (30/33)[a]	30.3 (10/33)	9.1 (3/33)	69.7 (23/33)
MM2	78.8 (26/33)	27.3 (9/33)	12.1 (4/33)	81.8 (27/33)
MM3	75.8 (25/33)	27.3 (9/33)	15.2 (5/33)	66.7 (22/33)
MM4	81.8 (27/33)	27.3 (9/33)	21.2 (7/33)	72.7 (24/33)
MM5	60.6 (20/33)	27.3 (9/33)	0.0	72.7 (24/33)
CP1	69.7 (23/33)	12.1 (4/33)	6.1 (2/33)	51.5 (17/33)
CP3	75.8 (25/33)	27.3 (9/33)	3.0 (1/33)	54.5 (18/33)
CP4	57.5 (19/33)	18.1 (6/33)	0.0	69.7 (23/33)
CP5	87.8 (29/33)	33.3 (11/33)	6.1 (2/33)	69.7 (23/33)
CP6	53.1 (17/32)	25.0 (8/32)	0.0	68.7 (22/32)
CP7[b]	100 (4/4)	25.0 (1/4)	0.0	25.0 (1/4)

[a] - % Positive samples (No. of positive samples /No. of samples tested, 500 mL filtered)
[b] - CP7 was only sampled 4 times.

4.6.2 *Campylobacter* Species Determination in MM and CP Water Samples

Samples that tested positive for *Campylobacter* spp. by qPCR were further analyzed by triplex PCR to determine if the important human pathogenic species of *C. jejuni, C. lari* and/or *C. coli* were present in the samples. The triplex PCR showed that samples did not contain a mixture of species. For example, *C. jejuni* was never found

together with *C. lari* and/or *C. coli* in one sample. High percentages of *C. jejuni* were found at MM1, MM2, and MM4 (Table 4.5). Occasional hits of *C. lari* were found in both watersheds. *C. coli* was rarely detected with positive samples at MM4 and CP6 only (Table 4.5). Other species of *Campylobacter* (i.e., *C. upsaliensis, C. helveticus,* and/or *C. hyointestinalis)* were prevalent at MM5 and all of the Collin's Park/Lake Fletcher sites (Table 4.5).

Table 4.5 Triplex PCR *Campylobacter* species determination in MM and CP samples

Site	***Campylobacter jejuni*** **(%)**	***Campylobacter lari*** **(%)**	***Campylobacter coli*** **(%)**	***Other -*** *C. upsaliensis, C. helveticus,* and/or *C. hyointestinalis* **(%)**
MM1	50.0 (15/30)[a]	10.0 (3/30)	0.0	40.0 (12/30)
MM2	61.5 (16/26)	15.4 (4/26)	0.0	23.1 (6/26)
MM3	24.0 (6/25)	12.0 (3/25)	0.0	64.0 (16/25)
MM4	55.6 (15/27)	7.4 (2/27)	3.7 (1/27)	33.3 (9/27)
MM5	15.0 (3/20)	0.0	0.0	85.0 (17/20)
CP1	4.3 (1/23)	0.0	0.0	95.7 (22/23)
CP3	28.0 (7/25)	4.0 (1/25)	0.0	68.0 (17/25)
CP4	21.0 (4/19)	5.3 (1/19)	0.0	73.7 (14/19)
CP5	27.6 (8/29)	0.0	0.0	72.4 (21/29)
CP6	5.9 (1/17)	11.8 (2/17)	5.9 (1/17)	76.4 (13/17)
CP7[b]	0.0	0.0	0.0	100.0 (4/4)

[a] - % Positive samples (No. of positive samples /No. of samples tested, 500 mL filtered)
[b] - CP7 was sampled 4 times.

4.6.3 Seasonal Trends in Pathogen Detection

Campylobacter spp. and *L. monocytogenes* were detected year-round in both watersheds in 2012-2013 while *Salmonella* and *E. coli* O157:H7 were detected primarily during the summer months (Figs. 4.1 and 4.2). Higher average water temperature seemed to coincide with a higher number of *Campylobacter* and *L. monocytogenes* positive samples in the summer months.

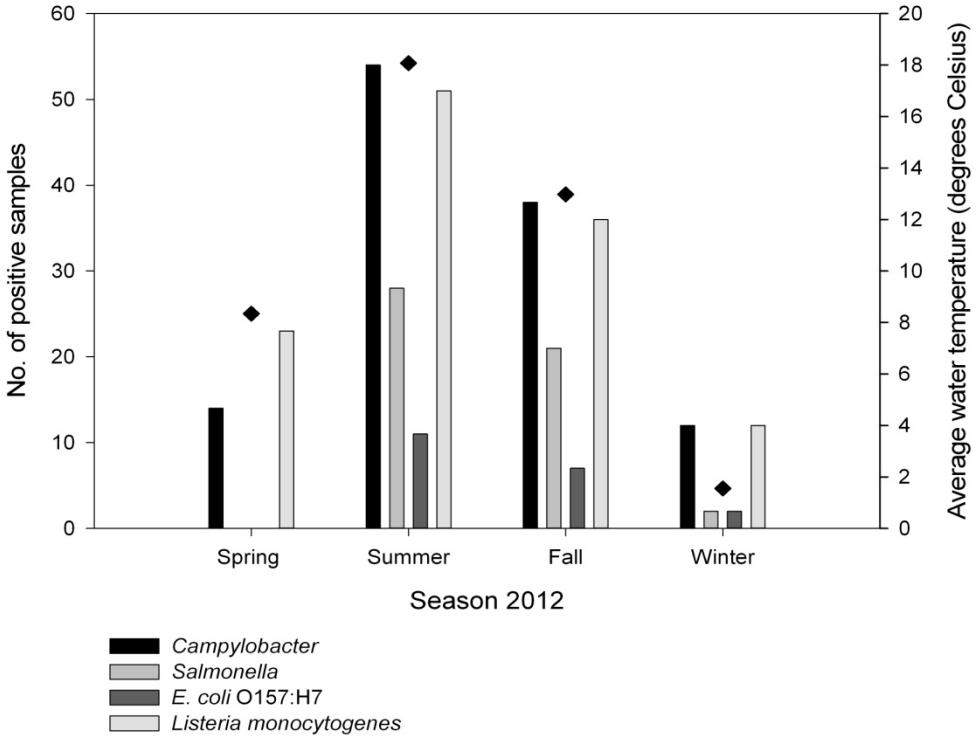

Figure 4.1 Seasonal trends during 2012 in pathogen detection in water samples from both watersheds. Average water temperature is also shown. Numbers of positive samples were tallied for each pathogen for each season. Winter: December - February; Spring: March - May; Summer: June - August; Fall: September - November.

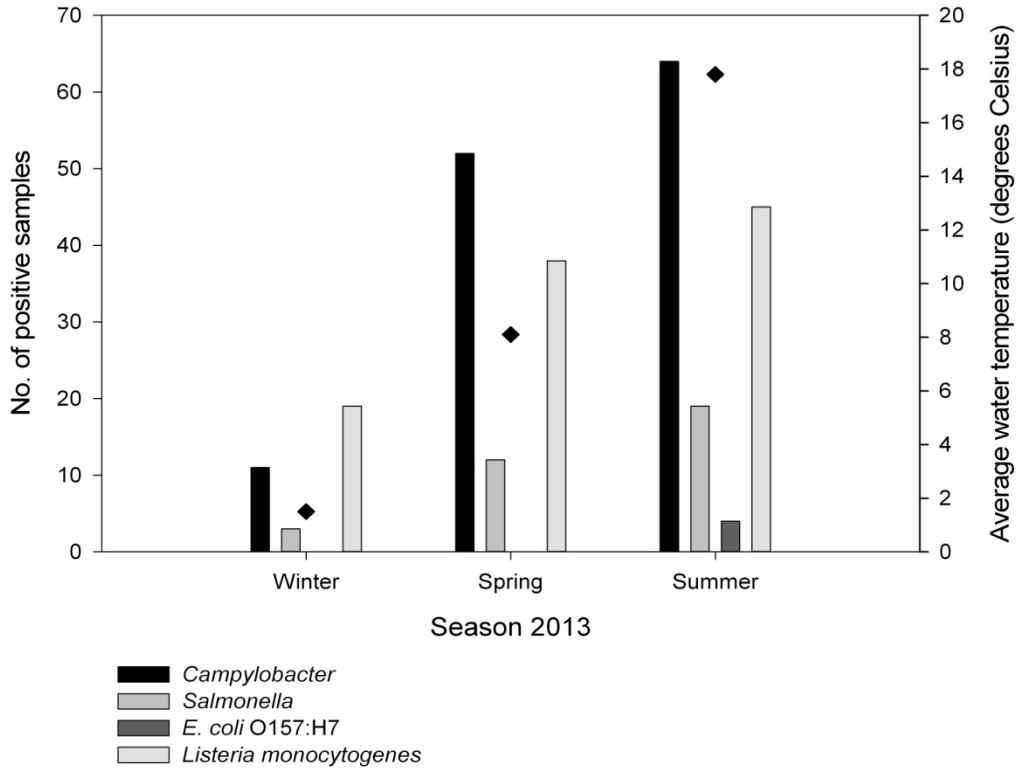

Figure 4.2 Seasonal trends during 2013 in pathogen detection in water samples from both watersheds. Average water temperature is also shown. Numbers of positive samples were tallied for each pathogen for each season. Winter: December - February; Spring: March - May; Summer: June – August.

4.6.4 *Listeria* Diversity in MM and CP

All water samples collected during the 33 sampling runs were subject to a culture-based *Listeria* spp. enrichment and detection protocol using 500 mL sample waters. More *Listeria* isolates were collected from MM water samples (866) than CP water samples (456) (total=1322 isolates). Screening on the Rapid Lmono agar showed that 125 (27%) of the CP isolates were *L. monocytogenes* while 169 (20%) of the MM isolates

were identified as *L. monocytogenes.* The *L. monocytogenes* isolates were divided into serogroups by a multiplex PCR method (total=294 isolates).

The number of positive *Listeria* sampling events as determined by the culture-based method were higher in the Middle Musquodoboit watershed than in the Collin's Park/Fletcher Lake watershed (Fig. 4.3), which is in agreement with the qPCR results showing higher detection rates of *L. monocytogenes* in MM (Table 4.4). However, a closer look revealed that detection rates of *L. monocytogenes* were generally higher when samples were analyzed by qPCR (Table 4.6). The culture-based method demonstrated that the pathogenic *L. monocytogenes* did not dominate in the *Listeria* populations found in the two watersheds (Fig. 4.3). The non-pathogenic group consisting of *L. seeligeri*, *L. grayi* and *L. innocua* was prevalent in MM water samples and less common in CP water samples, where more water samples contained the non-pathogenic *L. welshimeri* (Fig. 4.3). *L. ivanovii*, which is also pathogenic, was a less common occurrence at either watershed (Fig 4.3). It should be noted that the discrimination power of Rapid Lmono© medium is such that *L. seegligeri, L. grayi* and *L. innocua* are placed in the same category of white colonies on the red agar (Bio-Rad, 2005).

Further division of the *L. monocytogenes* isolates into the four serogroups showed that serogroup IIa was predominant at the majority of sites, with the exception of MM1 and CP3 where serogroup IIb dominated, and CP6 where serogroup IVb dominated (Fig. 4.4).

Table 4.6 Detection of *L. monocytogenes* in separate sampling events using qPCR and culture based methods

Watershed	**Rapid *L. mono* Testing (Culture-based Method) (% Detection)**	**qPCR Testing (DNA-based Method) (% Detection)**
Middle Musquodoboit	20.6 (34/165)[a]	66.7 (110/165)
Collin's Park/Lake Fletcher	12.2 (20/164)	57.3 (94/164)

[a] - No. *L. monocytogenes* detected/total no. of samples tested

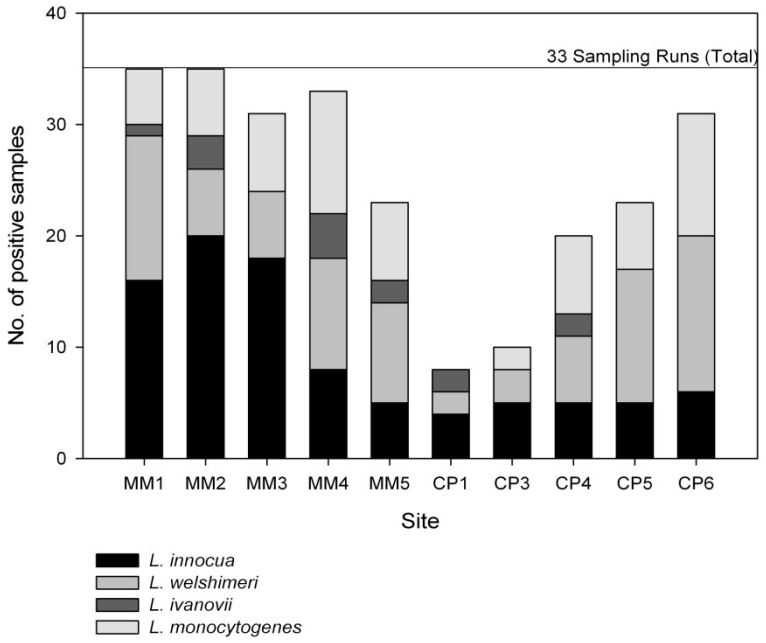

Figure 4.3 *Listeria* species diversity in MM and CP water samples (CP7 was not included in the results). Characterization of the 5 to 8 isolates obtained for each *Listeria* spp. positive sampling event showed that isolates always belonged to the same *Listeria* species. Therefore results shown in this graph show the species diversity per positive sampling event.

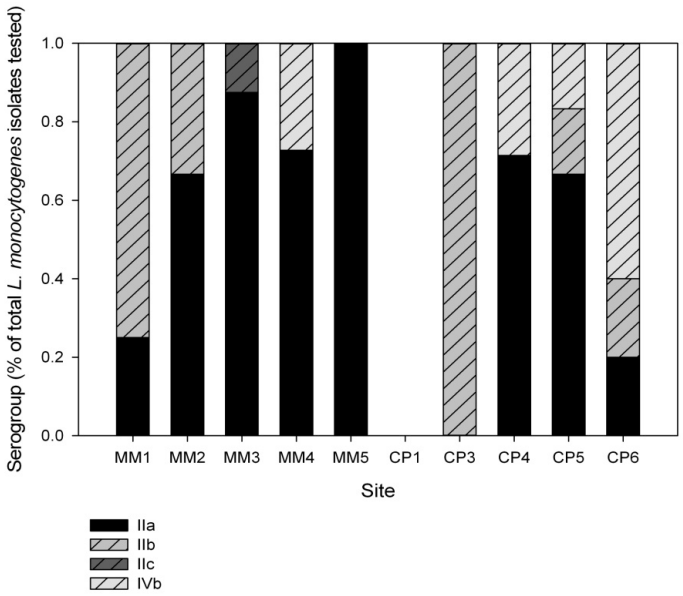

Figure 4.4 *L. monocytogenes* serogroup diversity in MM and CP water samples. Serogroups are shown as percentages of each site's total *L. monocytogenes* population. CP7 was not included in the analysis.

4.6.5 Pathogen Detection in Point Source Samples

Point source samples were also analyzed for the presence of the four pathogens by qPCR. Samples that tested positive for *Campylobacter* were also analyzed by triplex PCR, following the same protocol as were used for the water samples. *Campylobacter* spp. and *L. monocytogenes* were prevalent in cow and raw sewage fecal samples (Table 4.7). In 2012, the majority of dairy cows (randomly selected across all age groups) tested positive for *L. monocytogenes*, while the opposite occurred in 2013 where only 13% of the feces from the cows contained *L. monocytogenes* (Table 4.7). *C. upsaliensis, C. helveticus,* and/or *C. hyointestinalis* dominated in the *Campylobacter* spp. positive cow, raw sewage, septic, and raw water plant samples (Table 4.8). Interestingly, the Collin's Park drinking water treatment plant raw intake water tested positive for *C. jejuni* and *Salmonella* in the 500 ml sample volumes (Tables 4.7 and 4.8).

Table 4.7 Pathogen detection in point source samples

Site	***Campylobacter* spp. (%)**	***Salmonella* (%)**	***E.coli* O157:H7 (%)**	***Listeria monocytogenes* (%)**
Cow	100 (24/24)[a]	0.0	0.0	88.8 (8/9 – 2012 samples) 13.3 (2/15 – 2013 samples)
Raw Sewage	100 (3/3)	33.3 (1/3)	33.3 (1/3)	100 (3/3)
Septic	100 (3/3)	0.0	0.0	0.0
Water Plant, Raw, MM	66.7 (2/3)	0.0	0.0	66.7 (2/3)
Water Plant, Raw, CP	100 (3/3)	33.3 (1/3)	0.0	66.7 (2/3)

[a] - No. of positive samples /No. of samples tested (500 ml was analyzed)

Table 4.8 *Campylobacter* species determination in point source samples

Site	***Campylobacter jejuni* (%)**	***Campylobacter lari* (%)**	***Campylobacter coli* (%)**	***Other -*** *C. upsaliensis, C. helveticus,* and/or *C. hyointestinalis* **(%)**
Cow	0.0	0.0	0.0	100 (24/24)
Raw Sewage	0.0	0.0	0.0	100 (3/3)
Septic	0.0	0.0	0.0	100 (3/3)
Water Plant, Raw, MM	0.0	0.0	0.0	100 (2/2)
Water Plant, Raw, CP	33.3 (1/3)[a]	0.0	0.0	66.7 (2/3)

[a] - No. of positive samples /No. of samples tested (500 ml was analyzed)

4.7 *E. COLI*, AN INDICATOR ORGANISM, IN MM AND CP

4.7.1 *E. coli* Concentrations in MM and CP Water Samples

Mean concentrations of *E. coli* were significantly ($p<0.05$) higher at sites MM1, MM2 and MM3 when compared to CP sites (Figure 4.5). Maximum *E. coli* values of 2080 CFU/100 mL, 1680 CFU/100 mL, and 970 CFU/100 mL for these sites, respectively, were recorded in August 2012. In contrast, the maximum concentration of *E. coli* in the Collin's Park watershed (450 CFU/100 mL) occurred in July 2012 at CP5.

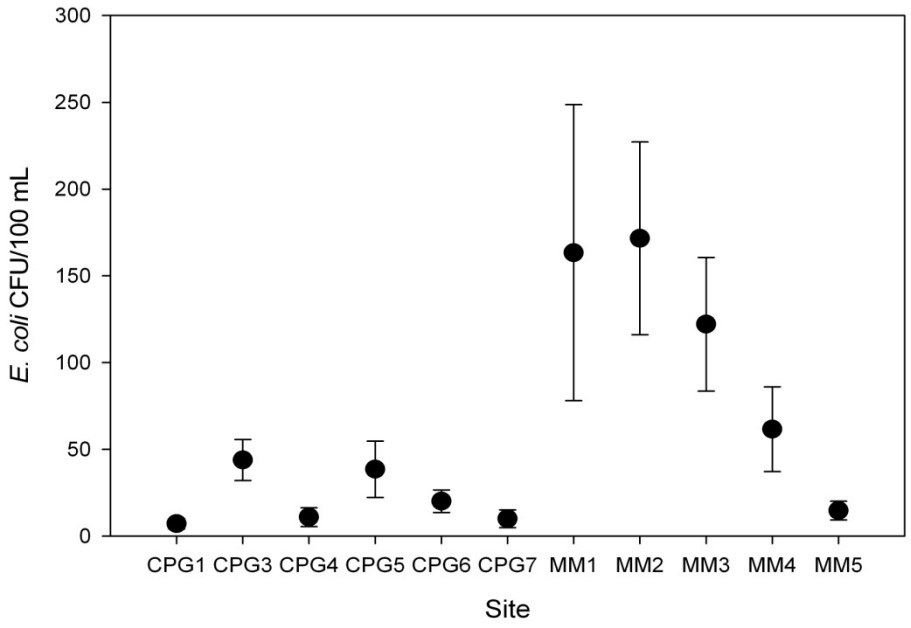

Figure 4.5 Mean concentrations, with associated 95% confidence intervals, of *E. coli* in the MM and CP watershed during the 18-month monitoring period in 2012-13.

4.7.2 *E. coli* and AllBac Concentrations in MM and CP

As discussed in section 4.4.1, AllBac concentrations were similar across sampling sites. *E. coli* and AllBac concentrations were positively correlated ($r^2 = 0.254$, p-value = <0.001, n=333) when analyzed using the Spearman ranking method (Figure 4.6).

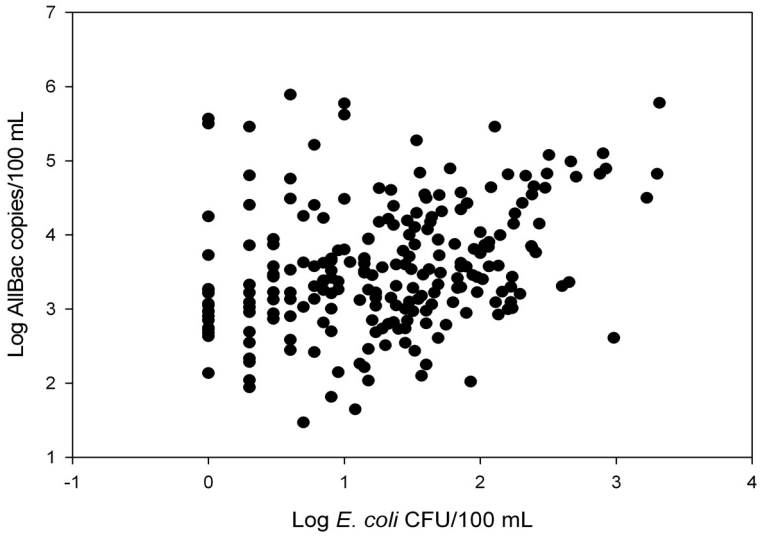

Figure 4.6 Relationship between the concentrations of *E. coli* and AllBac in MM and CP water samples ($r^2 = 0.254$, p-value = <0.001, n=333)

4.7.3 *E. coli* Phylogenetic Grouping by Triplex PCR

A total of 727 *E. coli* isolates were collected from the watersheds between July 2012 and July 2013, representing 21 sampling runs. For every sampling site, up to 5 colonies were randomly selected from the m-Coliblue plate for further analysis. Using this sampling plan, the maximum amount of *E. coli* isolates that could be collected from each site in the sampling time frame was 105 colonies. Point source sample isolates (11) were also obtained and included in the phylogenetic analysis by triplex PCR.

In general, *E. coli* loads were higher in the MM watershed, with more isolates (close to the maximum of 105 at sites MM1-MM3) collected during the July 2012-July 2013 sampling period (Fig. 4.7). The triplex PCR gels revealed that MM sites contained higher number of B1 isolates when compared with Collin's Park. Group D isolates were common at sites CP3, CP4, and CP5, while Group B2 isolates were prevalent at sites MM1 and MM4 (Fig. 4.7). Point source samples tested included septic tank and raw sewage samples. All raw sewage samples tested positive for the B2 phylogenetic group, while all septic tank samples tested positive for phylogenetic group D.

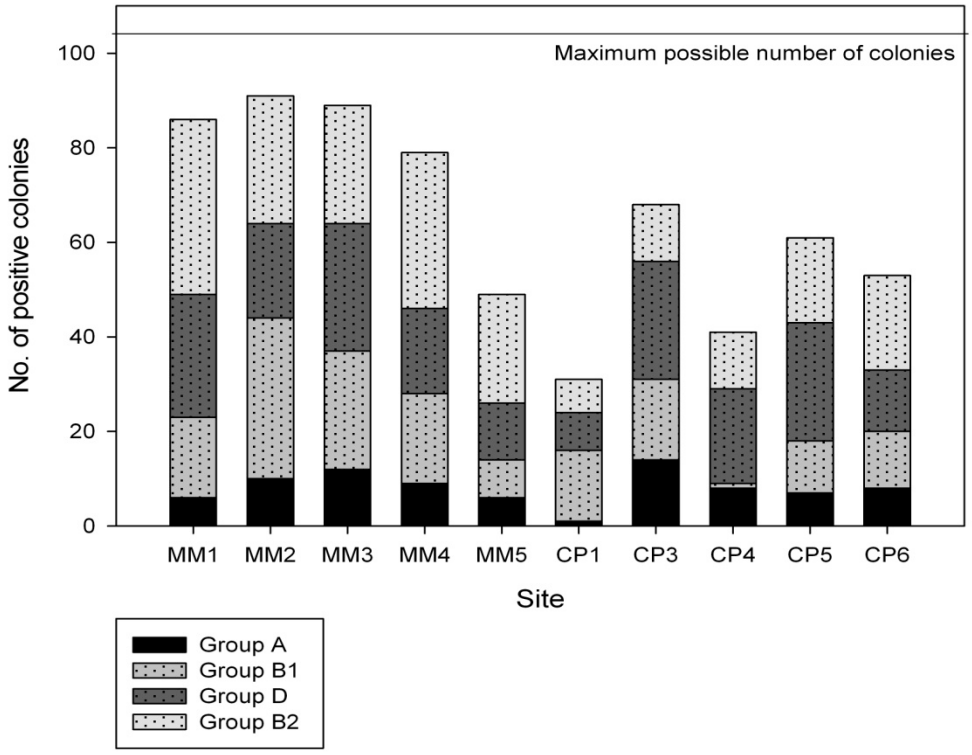

Figure 4.7 Phylogenetic group diversity among *E. coli* isolates obtained from the MM and CP sites

4.7.4 *E. coli* and Total Coliforms as Predictors of Pathogen Presence

Logistic regression models were used to examine the ability of *E. coli* and/or total coliforms to predict pathogen presence/absence in both watersheds. Residual plots were analyzed in SYSTAT during statistical testing. Assumptions of error independence and constant variance were met. Other potential predictors such as AllBac, turbidity, precipitation, and total suspended solids were also examined. Predictor variables such as turbidity are regularly measured by Halifax Water and were included in the analysis to determine if they would work better than *E. coli* as an indicator of pathogen presence. Due to co-linearity of predictors (discussed in further detail in section 4.8), stepwise logistic regression using various predictors was not performed.

The logistic regression analyses showed that detection of *E. coli* weakly predicted *Campylobacter* spp. and *Salmonella* spp. presence in both CP and MM watersheds. The odds of detecting *Salmonella* and *Campylobacter* in a water sample when *E. coli* was present were 1-1.01 and 1-1.03 times greater, respectively, than the odds of detection when *E. coli* was absent (Tables 4.9 & 4.10). These models showed high values for goodness-of-fit terms suggesting better model fits (Tables 4.9 & 4.10). The models for *Listeria* and *E. coli* O157:H7 were not significant ($p>0.05$) indicating that *E. coli* detection cannot be used to indicate presence of these organisms. Total coliforms were significant ($p<0.05$) as a predictor of *Campylobacter* spp. and *Salmonella* spp. presence in the CP watershed only (Table 4.9). Models using other predictors, including AllBac, turbidity, HF183, precipitation, and total suspended solids were not significant ($p>0.05$).

Table 4.9 Prediction of the presence of bacteria pathogens in water samples from CP based on enumeration of *E. coli* (CFU/100 mL) and total coliforms (CFU/100 mL) as analyzed by logistic regression (n=168)

Model Parameters (y,x)	Odds Ratio, Standard Error	95% CI for Odds Ratio	McFadden Rho^2	Area under ROC Curve	p-value
Campylobacter, E. coli	1.026, 0.011	1.004, 1.048	0.064	0.697	<0.001
Salmonella, E. coli	1.014, 0.004	1.006, 1.021	0.100	0.788	<0.001
E. coli O157:H7, *E. coli*	0.996, 0.004	0.988, 1.005	0.004	0.569	0.389
Listeria monocytogenes, E. coli	1.002, 0.003	0.996, 1.008	0.002	0.549	0.520
Campylobacter, Tot. Coliforms	1.012, 0.004	1.004, 1.021	0.097	0.712	<0.001
Salmonella, Tot. Coliforms	1.004, 0.002	1.000, 1.007	0.068	0.737	0.001
E. coli O157:H7, *Tot. Coliforms*	0.996, 0.003	0.990, 1.002	0.021	0.650	0.098
Listeria monocytogenes, Tot. Coliforms	0.999, 0.001	0.998, 1.001	0.009	0.500	0.206

Table 4.10 Prediction of the presence of bacteria pathogens in water samples from MM based on enumeration of *E. coli* (CFU/100 mL) and total coliforms (CFU/100 mL) as analyzed by logistic regression (n=165)

Model Parameters (y,x)	**Odds Ratio, Standard Error**	**95% CI for Odds Ratio**	**McFadden Rho^2**	**Area under ROC Curve**	**p-value**
Campylobacter, E. coli	1.008, 0.004	1.000, 1.016	0.057	0.815	0.001
Salmonella, E. coli	1.002, 0.001	1.000, 1.004	0.051	0.732	0.002
E. coli O157:H7, *E. coli*	1.001, 0.001	0.999, 1.002	0.005	0.513	0.343
Listeria monocytogenes, E. coli	1.002, 0.002	0.999, 1.005	0.019	0.624	0.055
Campylobacter, Tot. Coliforms	1.000, 0.000	1.000, 1.001	0.010	0.699	0.220
Salmonella, Tot. Coliforms	1.000, 0.000	1.000, 1.001	0.023	0.670	0.050
E. coli O157:H7, *Tot. Coliforms*	1.000, 0.000	1.000, 1.001	0.006	0.563	0.335
Listeria monocytogenes, Tot. Coliforms	1.000, 0.000	1.000, 1.001	0.007	0.615	0.265

4.8 Environmental Variables and Storm Events in MM and CP

4.8.1 Relationships between *E. coli* and Environmental Variables

Environmental variables such as total suspended solids, turbidity, pH, dissolved oxygen, precipitation, electrical conductivity and temperature were tested for their relationships with the detection of the fecal indicator organism, *E. coli*, and AllBac/HF183 markers.

Several variables were positively correlated with each other in the scatter plot matrix (Appendix A), leading to the creation of a correlation matrix. The important assumption of multivariate normality was tested in R prior to analysis of the correlation matrix (Shapiro-Wilk test, W = 0.8825, p-value = 0.077, fail to reject the null hypothesis that the data came from a normal distribution).

During PCA analysis, the first two components of PCA were found to account for most of the variation in the dataset (Appendix A). The loadings were examined and the scores were plotted on the bi-plot diagram (Figure 4.8), where a strong association between *E. coli*, total suspended solids and turbidity can be observed. A slightly weaker association between *E. coli* and precipitation was observed. An even weaker association was observed between *E. coli* and AllBac (also described in section 4.7.2) and *E. coli* and HF183 (Fig. 4.8).

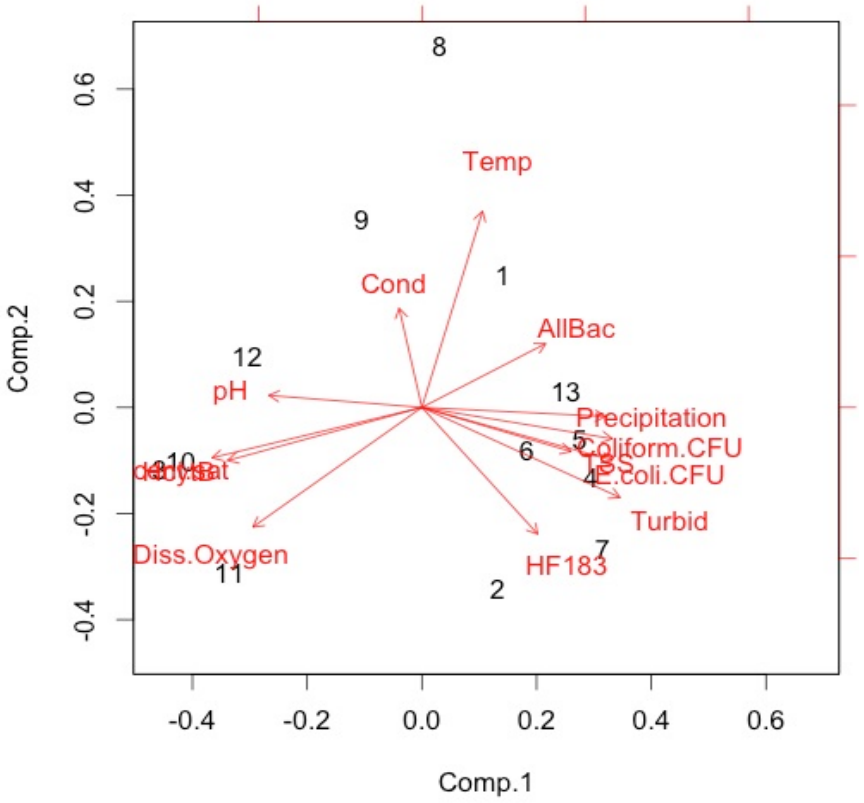

Figure 4.8 Principle components analysis of the multivariate dataset using the correlation matrix. Component 1 accounts for 46.4% of the variance, Component 2 accounts for 16.5% of the variance.

When these important variables were analyzed using the Spearman rank-order method, significant positive correlations were found (Table 4.11). Turbidity and *E. coli* showed the strongest correlation (r^2 =0.586) (Table 4.11).

Table 4.11 Spearman rank-order correlation results for important environmental variables in relation to *E. coli* concentration

Correlation	Coefficient	p-value	N
E. coli vs Turbidity	0.586	<0.001	293
E. coli vs TSS	0.301	<0.001	297
E. coli vs Precipitation	0.309	<0.001	330

4.8.2 Effects of Storm Events in MM and CP

Among the sampling events in MM were six storm events (i.e., sampling runs # 12, 14, 15, 19, 26 and 30), which were captured during the 2012-2013 sampling season. Three storm events (i.e., sampling runs # 14, 15, and 29) were also captured in CP during the experimental monitoring period.

When pathogen occurrence was analyzed with respect to storm events, *E. coli* O157:H7 showed an interesting pattern in both watersheds. Approximately 40% of *E. coli* O157:H7 positive samples were detected during storm events in both watersheds (Table 4.12). In MM only, the majority of *Salmonella* positives occurred during storm events. In contrast storm event associated detection rates of *Campylobacter* spp. and *L. monocytogenes* were low in both watersheds.

Increased precipitation and flow during storm events were associated with elevated *E. coli* concentrations and occasional marker detection, especially during the hurricane season of late August/early September 2012. These trends are outlined in time series diagrams (Figs. 4.9-4.11). In the MM watershed a strong association was detected between increased flow and elevated *E. coli* concentrations (Figs. 4.9-4.10). At the CP3 and MM3 sampling sites, HF183 and BacR (MM3 only) hits occurred during sampling events when *E. coli* concentrations were elevated (Figs. 4.10-4.11). The majority of HcytB positive samples in MM3 occurred when *E. coli* concentrations were not elevated (Figure 4.10).

Table 4.12 Detection rates of pathogens during storm events in MM and CP watersheds

Watershed	***Campylobacter* spp. (%)**	***Salmonella* (%)**	***E. coli* O157:H7 (%)**	***Listeria monocytogenes* (%)**
Middle Musquodoboit	21.9 (28/128)[a]	54.3 (25/46)	42.1 (8/19)	22.5 (27/120)
CP/Lake Fletcher	10.6 (12/113)	18.4 (7/38)	40.0 (2/5)	9.7 (10/103)

[a] - No. positive samples occurring during storm event/total no. positive samples

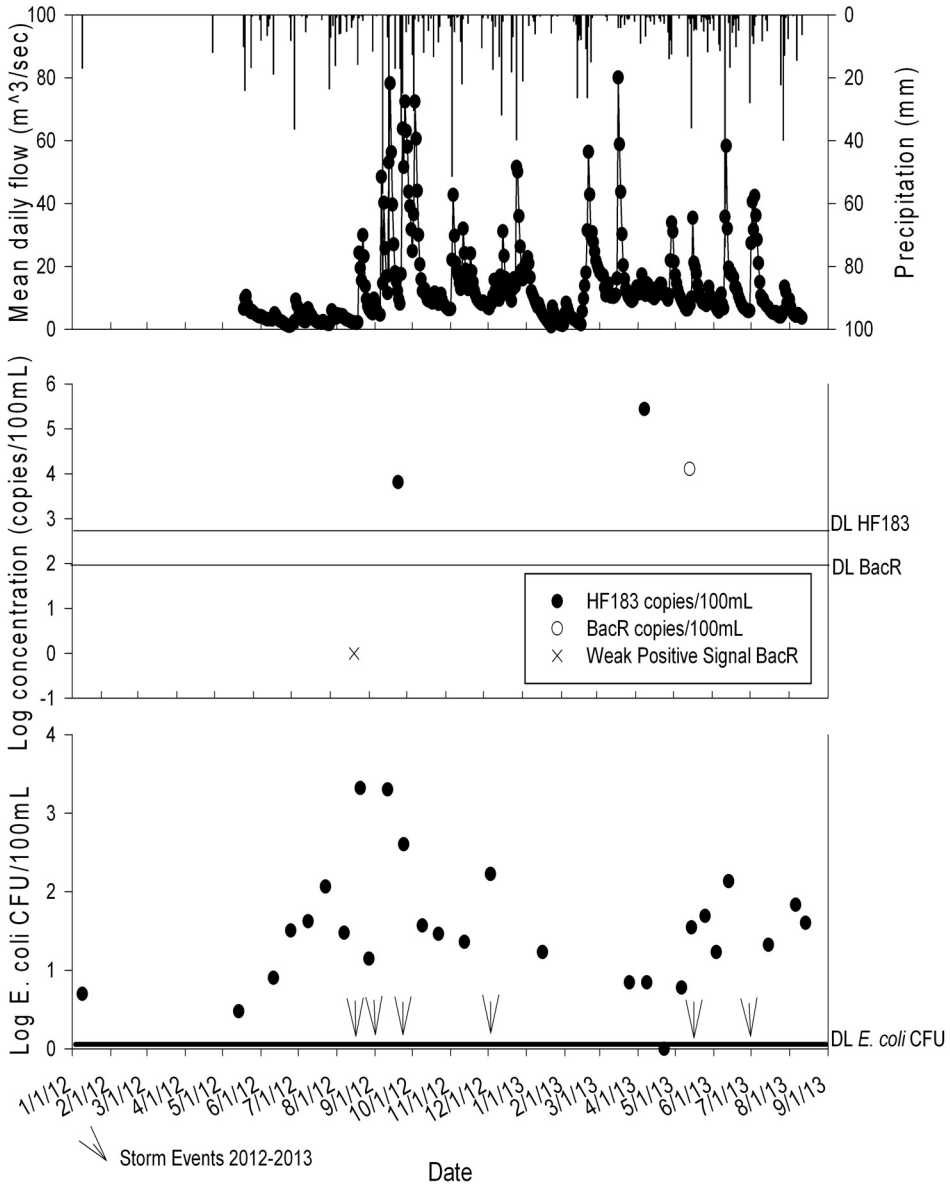

Figure 4.9 MM1 time series of precipitation, flow, *Bacteroidales* marker concentration, and *E. coli*. Note: Continuous flow data measurements did not begin until May 2012. Sampling occurred during all months with the exception of Feb. 2012.

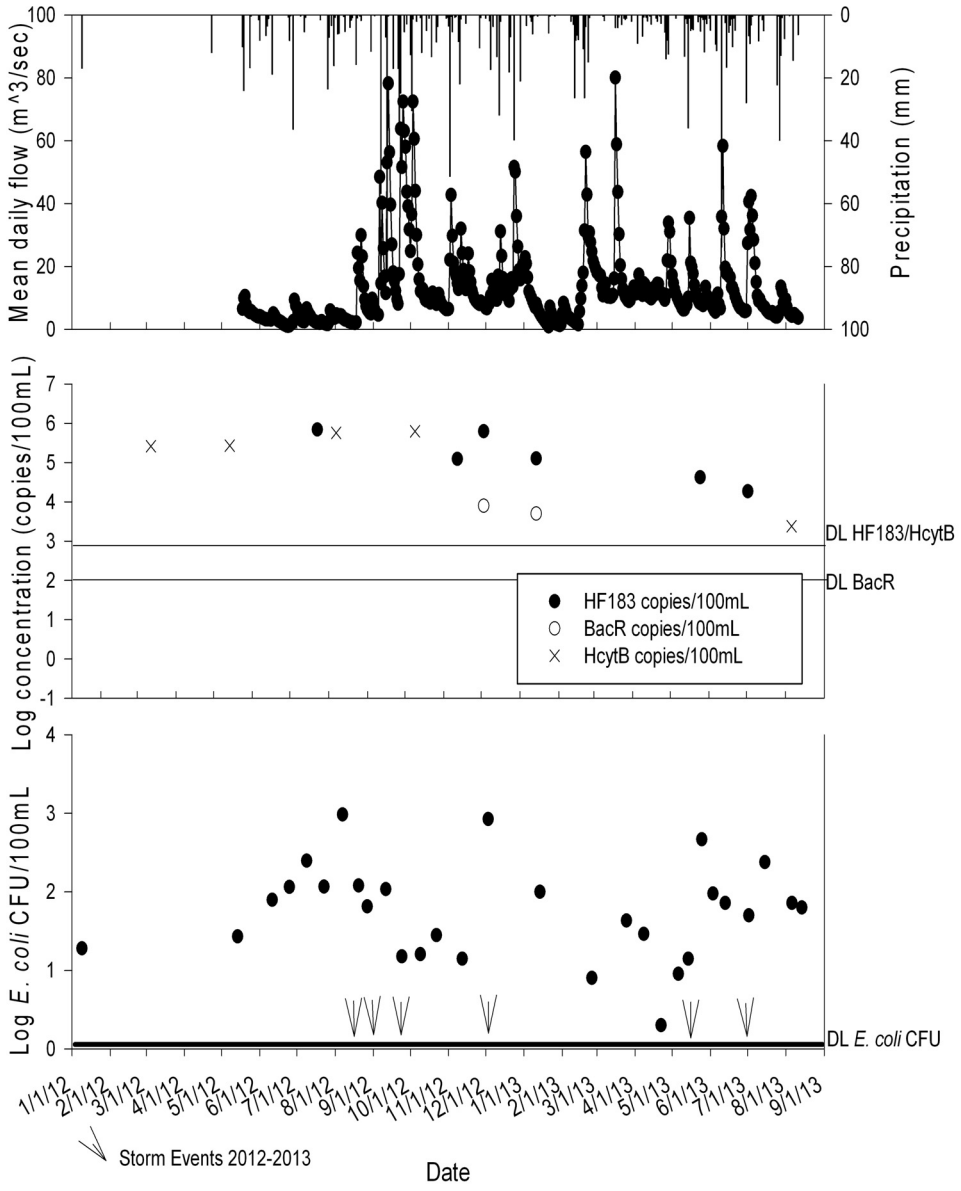

Figure 4.10 MM3 time series of precipitation, flow, marker concentration, and *E. coli*. Note: Continuous flow data measurements did not begin until May 2012. Flow measurements described above are for the outlet, MM1.Sampling occurred during all months with the exception of Feb. 2012.

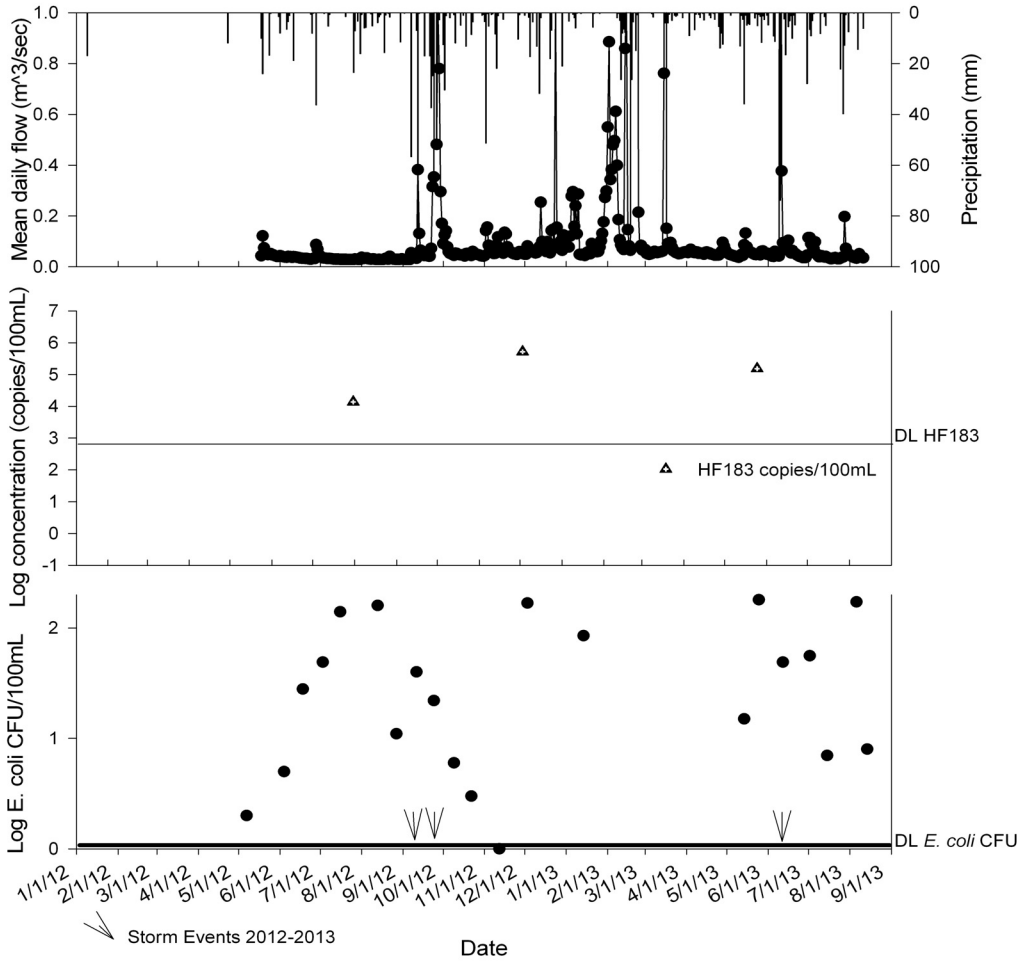

Figure 4.11 CP3 time series of precipitation, flow, *Bacteroidales* marker concentration, and *E. coli*. Note: Continuous flow data measurements did not begin until May 2012. Sampling occurred during all months with the exception of Feb. 2012.

CHAPTER 5 DISCUSSION

5.1 MARKER VALIDATION STUDY

Sensitivity and specificity of the HF183, HcytB, and AcytB markers was determined in order to validate these markers for use in Nova Scotia. Both the human *Bacteroidales* based HF183 and human mtDNA HcytB marker showed excellent sensitivity and specificity (100%, 100%, 86.7% and 96.4%, respectively) suggesting these markers can be used reliably in Nova Scotia to detect human fecal pollution. Other studies have demonstrated excellent sensitivity and specificity of HF183 including in Australia where sensitivity of 100% and specificity of 98-100% were reported (Ahmed et al., 2008, 2009). A false positive HcytB result was obtained from one cow manure sample while several small mammal fecal samples were false positives for HF183. This result is similar to observations by Ridley (2011) where the BacH human marker cross-reacted with small mammal fecal samples. The small mammals used in both studies live at the Shubanacadie Wildlife Park and have regular contact with humans, and this close contact could have affected specificity results. The cow manure sample that tested positive for HcytB could have easily been contaminated with human DNA from the sample collector (although precautions were taken to prevent contamination) or from farm staff working in close proximity to cows. Another study showed quantification of HF183 in 100% of sewage, 67% of human fecal samples, and 8% of cat samples (Werfhorst et al., 2011). Cats could have been in close contact to humans, like the small mammals tested in this study.

Both the HcytB and HF183 markers were present in large quantities (10^7 and 10^{10} copies/g, respectively) in feces. Other studies also found high copy numbers (10^9) of HF183 per gram wet weight in sewage (Seurinck et al., 2005; Werfhorstet al., 2011). HcytB average concentration in this study per g of feces (1.2 x 10^7 copies) was very similar to results from other studies (1.1 x 10^7 copies and 1.4 x 10^7 copies) (Caldwell et al., 2007; Schill & Mathes, 2008).

The HcytB mtDNA marker was able to correctly identify fecal material in 85% of effluents (Baker-Austin et al., 2010). A small percentage of samples in the study by Baker-Austin et al. (2010) showed incorrect results with a weak human signal, most likely due to contamination by lab staff. Using ND5 gene-based mtDNA markers, 85% of effluents were correctly identified with no false positives (Caldwell et al., 2007), which is similar to reports for other CytB mtDNA markers showing sensitivity of 85% and specificity of 99% (Schill & Mathes, 2008). Sensitivity of HcytB, in contrast, was 100% in this study, while the specificity was 86.7%. Differences in sample sizes and sampling handling procedures could account for variation in sensitivity and specificity studies.

In contrast, the AcytB mtDNA-based cow marker may be harder to detect in water samples because of lower abundance in feces (5.73 x 10^6) when compared with markers such as BacR (2.03 x 10^{10}). Although the probe used in the AcytB assay was specific for cow DNA, one false negative and several false positives were detected during marker validation studies. This marker may not be useful for use in future studies because of low starting quantity in cow feces. Schill & Mathes (2008) reported average cow mtDNA concentrations per g feces of (3.45 x 10^7 copies/g) which were an order of magnitude higher than AcytB concentration per g feces found in this study. This may be

due to variability in concentrations of cells shed in the feces between cow breeds used in both studies (Angus versus Holstein breeds) and differences in DNA extraction procedures.

A high LOD was found for ND5-gene based mtDNA markers (2.0×10^6 copies/100 mL) (Caldwell et al., 2007), much higher than the LOD found for the cytochrome B gene based HcytB marker in this study (925 copies/100 mL).

HF183 could be the most promising marker for future studies due to its presence in large quantities in feces and its high sensitivity for detecting human feces. HcytB also showed high sensitivity but was less prevalent in human feces, and for that reason, HF183 appears to be the superior marker for human fecal contamination.

5.2 Marker Detection in MM and CP Water Samples

AllBac, the universal marker, was detected in almost all water samples, showing a ubiquitous occurrence in natural environments. However, AllBac's suitability as an indicator of total fecal pollution has been questioned. The AllBac primer sets are capable of detecting *Bacteroidales* from the hindgut of insects and primer sets from the 16S rRNA genes of *Bacteroidales* showed 100% similarity with *Bacteroidales* species isolated from ecosystems, not limited to gastrointestinal tracts (van der Wielen & Medema, 2010). A recent study found large concentrations of AllBac in pristine alpine soil samples, suggesting a lack of specificity for feces (Vierheilig et al., 2012). Furthermore, AllBac decay curves showed a more persistent population when compared with other targets such as HF183 and *E. coli* (Dick et al., 2010). The AllBac marker was detected until the last day during a persistence study (day 14 and day 15 in the first and

second trials, respectively), similar to the persistence of *E. coli* (Tambalo et al., 2012). In our study it would appear that AllBac, although correlated with *E. coli*, was not a good replacement for *E. coli* as an indicator of pathogen presence in water samples (section 4.4). This is similar to other reports where levels of *E. coli* have been found to be correlated with AllBac (Ridley et al., 2011, Mauffret et al., 2012).

Analyses of host-associated marker detection in both MM and CP watersheds showed surprising results. Based on marker detection rates human contamination was found to be the primary source of fecal pollution in both the urban (CP) and agricultural (MM) watersheds, with several positive hits of the HF183 and HcytB markers obtained throughout the sampling period at all sites except CP4. This could be explained by the fact that the CP4 sampling site is located approximately 1 km into the woods and is also the site located furthest away from human households.

Several studies have found evidence of human fecal contamination in both urban and mixed watersheds from unidentified sources (Schriewer et al., 2010), urban storm water runoff (Sauer et al., 2011; Sidhu et al., 2013), leaky sewers and overflow (Bower et al., 2005; McQuaig et al., 2012) and failing septic tanks (Peed et al., 2011; Drozd et al., 2012; Wilkes et al., 2013).

In both the CP and MM watersheds, the majority of households use onsite wastewater systems. In Lake Fletcher, the CP wastewater treatment plant is located on Lockview Road and serves customers on this road and a high school. New subdivisions on the hill just above Lake Fletcher and homes across the lake were built with onsite wastewater systems. It is unclear if human markers found at CP5 were the result of failing septic systems or leaky sewer system pipes/cross connections present nearby on

the Lockview Road. In MM, homes in general sit on large lots with onsite wastewater systems. Overloading of septic fields or inadequate maintenance (lack of pumping) of tanks could cause wastewater to enter nearby streams. During periods of heavy rainfall, the area around the septic field could become saturated causing the system to fail. Poor design and construction of onsite wastewater systems particularly in older homes in MM could be a contributing factor to septic system failure in this area. Soil and geologic conditions also influence the performance of onsite wastewater systems. Nova Scotia soil type 6 (Queens) is prevalent in MM, while NS soil type 2 (Halifax) is prevalent in the Lake Fletcher area (Quigley & Keys, 2007). The Halifax soil type generally has a sandy loam texture and is also more porous than the Queens soils type, with more potential for water percolation through the unsaturated zone (Quigley & Keys, 2007).

HF183 and HcytB were only found together in 14% of the samples that were positive for human markers. This is in contrast to a recent study which found the BacHum 16S rRNA based *Bacteroidales* marker correlated well with a human mtDNA marker (Kapoor et al., 2013). This discrepancy may be due to HF183 possibly persisting longer than HcytB in environmental waters, especially at lower temperatures. As such, HcytB was detected for up to 2 weeks at 4 ºC (Baker-Austin et al., 2010) while HF183 was detected for up to 24 days at 4 ºC and up to 8 days at 28 ºC (Seurinck et al., 2005). This result, and smaller quantities of HcytB found in feces, could explain why HcytB and HF183 were not found together consistently. This result also suggests the need for a toolbox of microbial water quality approaches when evaluating fecal contamination sources in a particular watershed.

BacR was occasionally found in the MM watershed, but was much less common than human marker detection. Human markers were also found much more frequently than cow *Bacteroidales* markers in another agricultural watershed (Drozd et al., 2012). Other cow markers were detected in low concentrations (weak positive signal, WPS) outside the range of accurate quantification in this study. A potential problem with both the CowM2 *Bacteroidales* marker and AcytB cow mtDNA marker is low initial starting quantity in feces, with quantities of 6.1×10^4 copies/g CowM2 found in manure (Piorkowski, 2013). In another study, cow *Bacteroidales* markers were only detected in highly contaminated samples near the river outflow in an agricultural area close to Lake Michigan (Bower et al., 2005).

Ruminant *Bacteroidales* markers were prevalent in local streams during storm events in a study in southwest Wales (Wyer et al., 2010). A recent study showed strong interactions between *Bacteroidales* markers of fecal pollution, season, stream flow, and farm practices (Wilkes et al., 2013). At low flows, the unrestricted cattle access sites had significantly higher detections (33%) of the ruminant *Bacteroidales* marker than the restricted cattle access site (7%), outlining the impact of cattle access to the stream (Wilkes et al., 2013). Agricultural activities in MM, such as dairy and beef farms, could contribute to detection of BacR at MM sites. However, other ruminants such as deer, sheep and goats, could have contributed to detection of this marker in the watershed (Reischer et al., 2006).

Bayes' theorem (Kildare et al., 2007) was used to determine the conditional probability that any detection of the host specific markers in this study was the result of a true positive for that marker. The HF183 and HcytB markers showed 44.5% and 53%

conditional probability respectively (Table A-1, Appendix A). These values were relatively high but suggest detection of other *Bacteroidales* or mtDNA sequences is possible, and therefore some uncertainty is associated with human marker results. The BacR and AcytB conditional probabilities were low at 32% and 15.2%, respectively, suggesting other incorrect sequences are being detected for the majority of positives (Table A-1, Appendix A). AcytB and BacR may not be completely reliable markers. The CowM2 marker, however, showed 100% conditional probability, due to 100% specificity (Table A-1, Appendix A).

5.3 Pathogen Detection in MM and CP

5.3.1 Pathogen Detection in MM and CP Water Samples

Overall pathogen presence was high in both MM and CP watersheds. Higher detection rates of *L. monocytogenes, C. jejuni* and *E. coli* O157:H7 were found in the Middle Musquodoboit watershed. High pathogen prevalence was also found in surface waters in Australia (Ahmed et al., 2009). In the 32 water samples collected, 25%, 3%, 28%, 44%, and 47% were positive for *C. jejuni mapA*, *Salmonella invA*, enterohaemorrhagic O157 and enterohaemorrhagic VT1 and VT2 genes, respectively (Ahmed et al., 2009). High rates of *Campylobacter* detection were also found in a mainly agricultural Nova Scotian watershed (Ridley et al., 2011). A larger number and variety of animal hosts in the MM watershed could cause differences in pathogen presence. A wide variety of animals such as birds, sheep, cattle, dogs, cats and pigs, as well as humans, could contribute to high rates of *Campylobacter* in the watersheds (Ogden et al., 2009). *C. coli* (3.7% in MM and 5.9% in CP) was rare in both watersheds, while *C. upsaliensis,*

C. helveticus, and/or *C. hyointestinalis* were more common in the urban CP watershed. In a rural study in the United Kingdom, *C. jejuni* was the most prevalent species in all animal samples and was isolated from 15% of water samples. *Campylobacter coli* was commonly found in water (17%) and sheep (21%) samples, but rarely in other samples. *Campylobacter lari* was found in small numbers in birds (7%) and water (5%). *Campylobacter hyointestinalis* was only recovered from cattle (7%) and birds (1%) (Brown et al., 2004). Human raw sewage entering the CP wastewater treatment plant tested positive for all four pathogenic organisms. *C. upsaliensis, C. helveticus,* and/or *C. hyointestinalis* was found in raw sewage samples. Human fecal contamination containing pathogens could also contribute to high prevalence of *L. monocytogenes* and *Campylobacter* in both watersheds.

Cattle can carry *E. coli* O157:H7 in the gastrointestinal tract and shed this bacterium into the environment (Petridis et al., 2002). Several cattle were tested in MM in 2012 and 2013; none were positive for *E. coli* O157:H7 or *C. jejuni*. Interestingly, almost all cattle tested in 2012 were positive for *L. monocytogenes*, while no cows tested positive for the bacterium in 2013. The cattle were healthy at the time the fecal samples were collected, but *L. monocytogenes* has also been isolated from animal feeds (Nightingale et al., 2004; Atil et al., 2011). *L. monocytogenes* presence in the feed in 2012 could have led to its detection in cows that year. *Listeria* spp. are known to be distributed ubiquitously in the natural environment; *Listeria* spp. have been found in many environments including soil, water, food processing plants and farms (Sauders & Wiedmann, 2007; Sauders et al., 2012).

The detection of *Salmonella* spp., which are associated with a wide host range including birds, pigs, cattle and humans (Orji et al., 2005; Dorr et al., 2009), and *E. coli* O157:H7 followed a summer seasonal pattern. The summer seasonal pattern of *E. coli* O157:H7 and *Salmonella* should be a concern for ready-to-eat vegetable producers using untreated irrigation water from local rivers and children/adults using the MM River and Lake Fletcher in the summer for swimming. Increased occurrence of all four pathogens tested during the warmer months could be the result of higher loadings from animal and human sources or increased survival in warmer waters. Almost half of all *E. coli* O157:H7 positive samples were collected during storm events. This result suggests a need for improved runoff and manure management to prevent this harmful pathogen from entering water systems. It is unclear why MM4 showed the highest detection rates of *E. coli* O157:H7, with no farms being in close proximity to the site. Preliminary SWAT modeling, however, showed areas of increased erosion in the upper watershed upstream from MM4 (Appendix B-Erosion Map of MM).

E. coli O157:H7 poses the most risk in waters for recreational use and waters used as source water for drinking. As few as 10-100 cells can cause infection (Petridis et al., 2002). The presence of *C. jejuni*, which causes 80% of human bacterial intestinal disease cases (OIE Terrestrial Manual, 2008), and *Salmonella* spp. in CP drinking water treatment plant raw intake water also indicates the need for adequate water treatment and measures to disinfect the water in the event of system malfunction or high turbidity events. Protecting source water from *E. coli* O157:H7 contamination is extremely important especially in the event of water treatment plant malfunction.

5.3.2 *Listeria* Diversity in MM and CP

Higher detection rates of *L. monocytogenes* occurred when samples were analyzed by qPCR (%) compared to the culture-based method (%). This discrepancy may be due to the fact that *L. innocua,* and possibly other *Listeria* spp., have the ability to outcompete *L. monocytogenes* during the selective enrichment process (MacDonald & Sutherland, 1994). The findings in the present study differ from those reported by Sauders et al., (2012) because high rates of *L. monocytogenes* and *L. innocua* were found in an agricultural watershed. Sauders et al., (2012), in contrast, found a significant association between *L. monocytogenes* and *L. innocua* and urban watersheds ($p<0.05$).

When the culture-based method was used, *L. innocua/grayi/seegligeri* dominated in the MM watershed. *L. monocytogenes* accounted for 20% of the isolates and did not dominate in the *Listeria* spp. population found at any of the sampling sites as determined by the culture-based method. Serogrouping of presumptive *L. monocytogenes* results showed the IIa (1/2a, 3a serovars) and IIb (1/2b, 3b serovars) serogroups were prevalent in both the CP and MM watersheds. These results are similar to another Canadian study, where serogroup IIa was reported to dominate in a mixed-used watershed (Lyautey et al., 2007). Serogroups IIa and IVb were isolated from CP4, the site furthest from any anthropogenic activity, suggesting that the *L. monocytogenes* population is diverse in the natural environment. Four serovars, 1/2a, 1/2c, 1/2b, and 4b (part of serogroup IVb), cause over 98% of human listeriosis cases (Jacquet et al., 2002), and hence include serovars/serogroups found in this study.

One IIc serogroup positive sample was isolated during summer 2013 at MM3. A beaver dam appeared at this location during summer 2013, potentially impacting

microbial water quality. Further characterization of the *L. monocytogenes* isolates obtained during this study using pulsed field gel electrophoresis (PFGE) subtyping methods could provide better understanding of the ecology of this organism and its transmission between different reservoirs.

5.4 *E. COLI*, AN INDICATOR ORGANISM, IN MM AND CP

In general, the *E. coli* concentrations were higher in the MM watershed in comparison to the CP watershed. Again, higher numbers and varieties of animal carriers in MM could be the cause of elevated *E. coli* concentrations in MM. MM sites showed higher number of B1 isolates when compared with Collin's Park. This may point to animal sources of *E. coli* as B1 *E. coli* isolates were found to be more prevalent in sheep, cows, and goats (Carlos et al., 2010). However, all four phylogenetic groups (A, BI, B2 and D) can be found in humans as well as other hosts (Duriez et al., 2001). It should be noted that B1 isolates have demonstrated increased ability to persist in the environment (Walk et al., 2007). Group D isolates were common at sites CP3, CP4, and CP5, while Group B2 isolates were prevalent at sites MM1 and MM4. Group B2 and D isolates have been associated with human and pig fecal contamination, and isolates from these groups also show more virulence potential (Bingen et al., 1998; Carlos et al., 2010). The fact that Group B2 and D isolates were found in both MM and CP watersheds and in human sewage samples supports the general finding of our study that human fecal contamination is the main source of fecal pollution in both watersheds. Group B2 *E. coli* isolates were also found in water samples collected in close proximity to a residential area in the

Thomas Brook watershed located in the Annapolis Valley of Nova Scotia (Piorkowski, 2013, Phd Thesis).

Logistic regression analyses showed that detection of *E. coli* weakly predicted *Campylobacter* spp. and *Salmonella* spp. presence in both CP and MM watersheds. The odds of detecting *Salmonella* and *Campylobacter* in a water sample when *E. coli* was present were 1-1.01 and 1-1.03 times greater, respectively, than the odds of detection when *E. coli* was absent. On the contrary, levels of total coliforms were only meaningful as a predictor of *Campylobacter* spp. and *Salmonella* spp. presence in the CP watershed. Results from this study agree with others that have demonstrated the ability of *E. coli* to predict pathogens, especially *Salmonella* spp. (Walters et al., 2011; McEgan et al., 2013). In the present study, models using other predictors often used in water quality analyses, including AllBac, turbidity, HF183, precipitation, and total suspended solids were not significant, meaning that there parameters could not be used to predict the presence of pathogenic bacteria in the water samples. Host specific *Bacteroidales* markers also did not correlate well with fecal indicator bacteria or pathogens in another study (Walters et al., 2011). *E. coli* was found to be a weak, but significant indicator for pathogens in this particular study, although many issues remain. For example, *E. coli* showed no ability to predict *L. monocytogenes* or *E. coli* O157:H7. Although the latter is an interesting result, low frequencies of *E. coli* O157:H7 could have caused this result during statistical testing. Differences in *E. coli* and *L. monocytogenes* physiology and ecology in the environment could have contributed to this important result.

Although pathogens were frequently found in both watersheds, *E. coli* concentrations rose above the CCME guideline for recreational water (200 CFU/100 mL)

in only 6.3% of all samples tested, indicating that bacterial pathogens may be present in waters with low *E. coli* numbers. *E. coli*'s usefulness as a indicator of pathogen presence in water samples should be evaluated further in the future.

5.5 Environmental Variables and Storm Events in MM and CP

Increased precipitation and flow during storm events were associated with elevated *E. coli* concentrations, especially during the hurricane season of late August/early September 2012. The MM watershed showed a stronger trend between increased flow, precipitation and elevated *E. coli* concentrations. The strong correlations between *E. coli*, total suspended solids and turbidity found in this study had also previously been reported in other studies (Sinclair et al., 2009; Huey & Meyer, 2010; Stumpf et al., 2010).

Although human markers were occasionally found during storm events, results from PCA analysis revealed an association between precipitation and *E. coli*, while a much weaker association existed between precipitation and the human marker HF183 or AllBac. Due to watershed characteristics, flow responses to storm events are flashier in CP, while MM shows a dampened and delayed response to storms and longer storm event timeframes. It is possible that contaminants in the CP watershed could have been transported through the stream systems before storm samples were obtained; therefore elevated levels of *E. coli* and/or markers and bacterial pathogens were not captured. In a North Carolina study, *E. coli* was weakly correlated with total suspended solids but strongly correlated with different stages (base, rising, peak, falling) of the hydrograph (Stumpf et al., 2010). Also, a large amount of intra-storm variation of fecal indicator

bacteria levels was observed (Stumpf et al., 2010). It is recommended that more information about interactions between markers, FIB, and flow be obtained in the future through more intensive storm sampling, especially in urban watersheds such as CP.

CHAPTER 6 CONCLUSION

6.1 PROJECT SUMMARY

Water samples were collected from both an urban and agricultural watershed over an eighteen-month period in Nova Scotia. The HF183 *Bacteroidales* marker and HcytB/AcytB mitochondrial markers were validated for use in Nova Scotia. Both human markers showed excellent sensitivity and specificity suggesting these markers can be used reliably in Nova Scotia to detect human fecal pollution The AcytB marker combined with the bovine specific Taqman® probe may not be useful for use in future studies because of low starting quantity of mtDNA in cow feces.

Detection of *E. coli* weakly predicted *Campylobacter* spp. and *Salmonella* spp. presence in both CP and MM watersheds. The odds of detecting *Salmonella* and *Campylobacter* in a water sample when *E. coli* was present were 1-1.01 and 1-1.03 times greater, respectively, than the odds of detection when *E. coli* was absent. Total coliforms were significant as a predictor of *Campylobacter* spp. and *Salmonella* spp. in the CP watershed only. *E. coli* was found to be a weak, but significant, indicator for the presence of bacterial pathogens when compared with other predictors, such as TSS, turbidity, AllBac or HF183/HcytB markers, for which no significant correlation existed. *E. coli* showed no ability in this study, however, to predict *L. monocytogenes* or *E. coli* O157:H7 presence. Low frequencies of *E. coli* O157:H7 could have affected this result during statistical testing. Results from this study suggest the need for a toolbox approach to water quality monitoring that uses both traditional methods such as FIB and targeted methods to determine sources of fecal contamination.

Overall pathogen detection rates were high in both MM and CP watersheds. Higher rates of *L. monocytogenes, C. jejuni* and *E. coli* O157:H7 and higher *E. coli* concentrations were detected in the Middle Musquodoboit watershed. A larger number and variety of animal hosts in the MM watershed could cause differences in pathogen presence. Serogrouping of presumptive *L. monocytogenes* isolates showed the IIa (1/2a, 3a serovars) and IIb (1/2b, 3b serovars) serogroups were prevalent in both the CP and MM watersheds. *E. coli* phylogenic Group B2 and D isolates found in both MM and CP watersheds and in human sewage samples supported the finding that human fecal contamination is the main source of fecal pollution in both watersheds.

Storm events, elevated *E. coli* concentrations, and bacterial pathogen occurrence should be a particular concern in source water protection. Human contamination was found to be the main source of fecal pollution in both the urban (CP) and agricultural (MM) watersheds. Homeowners with onsite wastewater systems should be aware of the problems that can occur with their systems. Furthermore, adequate manure and runoff management on farms is extremely important to prevent entry of *E. coli* O157:H7 into rivers and lakes. Relying completely on *E. coli* and total coliforms to demonstrate safety of source water could lead to potential microbial water quality and safety problems. Targeted approaches based on known pathogen occurrence data in a particular watershed will result in better water quality management (Field & Samadpour, 2007). In order to fully understand and manage sources of microbial pollution in mixed watersheds, MST techniques should be used along with rigorous pathogen testing to target sources of fecal pollution in a watershed.

6.2 Future Directions

1. Conduct an erosion assessment in the upper watershed area of Middle Musquodoboit

2. Prioritize areas for runoff management to prevent entry of *E. coli* O157:H7 into the river system in MM

3. Inform the public about septic tank maintenance and pollution caused by failing septic systems and determine if the straight-piping of raw sewage into the MM river is an issue

4. Assess groundwater microbial water quality in the future to determine if this important source of drinking water is being impacted by fecal pollution sources

5. Attempt to calibrate and validate a watershed compute simulation model, such as SWAT, to predict *E. coli* dynamics in MM. The model could then be used to predict how water quality is affected by climate and watershed management strategies

6. Analyze *Listeria monocytogenes* isolates using PFGE subtyping methods to provide better understanding of the ecology of this organism and its transmission between different reservoirs

7. Design future rigorous storm event sampling schemes for an urban watershed such as CP to better understand the interactions between flow, precipitation, FIB and marker data during storms

8. Quantify the amount of pathogens in water samples in rivers and lakes to provide data for a more detailed exposure and health risk assessment.

BIBLIOGRAPHY

Ahmed, W., Goonetilleke, A., Powell, D., Chauhan, K., & Gardner, T. (2009). Comparison of molecular markers to detect fresh sewage in environmental waters. *Water Research*, *43*(19), 4908–17. doi:10.1016/j.watres.2009.09.047

Ahmed, W., Sawant, S., Huygens, F., Goonetilleke, A., & Gardner, T. (2009). Prevalence and occurrence of zoonotic bacterial pathogens in surface waters determined by quantitative PCR. *Water Research*, *43*(19), 4918–28. doi:10.1016/j.watres.2009.03.041

Ahmed, W., Stewart, J., Powell, D., & Gardner, T. (2008). Evaluation of *Bacteroides* markers for the detection of human faecal pollution, *Water Research, 46*, 237–242. doi:10.1111/j.1472-765X.2007.02287.x

Anderson, K. L., Whitlock, J. E., Valerie, J., & Harwood, V. J. (2005). Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology*, *71*(6), 3041-3048. doi:10.1128/AEM.71.6.3041

Angulo, F. J., Tippen, S., Sharp, D. J., Payne, B. J., Collier, C., Hill, J. E., & Swerdlow, D. L. (1997). A community waterborne outbreak of salmonellosis and the effectiveness of a boil water order. *American Journal of Public Health*, *87*(4), 580–584.

Agriculture Canada. (2013). *Watershed Evaluation of BMPs.* Retrieved November 1, 2013 from http://www.agr.gc.ca/eng/?id=1285354752471

Atil, E., Ertas, H. B., & Ozbey, G. (2011). Isolation and molecular characterization of *Listeria* spp. from animals, food and environmental samples, *Veterinarni Medicina*. *56*(8), 386–394.

Baker-Austin, C., Morris, J., Lowther, J. A, Rangdale, R., & Lees, D. N. (2009). Rapid identification and differentiation of agricultural faecal contamination sources using multiplex PCR. *Letters in Applied Microbiology*, *49*(4), 529–32. doi:10.1111/j.1472-765X.2009.02679.x

Baker-Austin, C., Rangdale, R., Lowther, J., & Lees, D. N. (2010). Application of mitochondrial DNA analysis for microbial source tracking purposes in shellfish harvesting waters. *Water Science and Technology, 61*(1), 1–7. doi:10.2166/wst.2010.767

Ballesté, E., & Blanch, A. R. (2010). Persistence of *Bacteroides* species populations in a river as measured by molecular and culture techniques. *Applied and Environmental Microbiology*, *76*(22), 7608–16. doi:10.1128/AEM.00883-10

Bell, A., Layton, A. C., McKay, L., Williams, D., Gentry, R., & Sayler, G. S. (2007). Factors influencing the persistence of fecal *Bacteroides* in stream water, *Journal of Environmental Quality, 38*, 1224–1232. doi:10.2134/jeq2008.0258

Bernhard, A. E., & Field, K. G. (2000a). Identification of nonpoint sources of fecal pollution in coastal waters by using host-associated 16S ribosomal DNA genetic markers from fecal anaerobes. *Applied and Environmental Microbiology*, *66*(4), 1587–94.

Bernhard, A. E., & Field, K. G. (2000b) A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Applied and Environmental Microbiology*, *66*(10), 4571-4574

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J. & Denamur, E. (1998). Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *The Journal of Infectious Diseases*, *177*(3), 642–50.

Bio-Rad. (2005) *Rapid L'mono Chromogenic Media*. Retrieved September 20, 2013 from http://www.biorad.com/webroot/web/pdf/fsd/literature/Bulletin_3046.pdf.

Blackburn, B. G., Craun, G. F., Yoder, J. S., Hill, V., Calderon, R. L., Chen, N., Lee, S. H., Levy, D. A. & Beach, M. J. (2004). Surveillance for waterborne disease outbreaks associated with drinking water: United States, 2001–2002. *Mortality Weekly Report, 53(8),* 23-45.

Bower, P. A., Scopel, C. O., Jensen, E. T., Depas, M. M., Mclellan, S. L., & Icrobiol, A. (2005). Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods, *Applied and Environmental Microbiology, 71*(12), 8305–8313. doi:10.1128/AEM.71.12.8305

Brown, P. E., Christensen, O. F., Clough, H. E., Diggle, P. J., Hart, C. A., Hazel, S. & French, N. P. (2004). Frequency and spatial distribution of environmental *Campylobacter* spp. *Applied and Environmental Microbiology, 70*(11), 6501–6511. doi:10.1128/AEM.70.11.6501

Byappanahalli, M.N. & Ishii, S. (2011). Environmental sources of fecal bacteria. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.93-110). Washington DC: American Society for Microbiology.

Cabral, J. P. S. (2010). Water microbiology: bacterial pathogens and water. *International Journal of Environmental Research and Public Health*, *7*(10), 3657–703. doi:10.3390/ijerph7103657

Caldwell, J. M., Raley, M. E., Levine, J. F. & Carolina, N. (2007). Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents, *Environmental Science Technology, 41,* 3277–3283.

Carraro-Colon, M., Wickhard, G. & Turco, R. (2011). Taxonomy, phylogeny, and physiology of the fecal indicator bacteria. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.23-28). Washington DC: American Society for Microbiology.

Carlos, C., Pires, M. M., Stoppe, N. C., Hachich, E. M., Sato, M. Z., Gomes, T.T. & Ottoboni, L. M. (2010). *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiology*, *10*, 161. doi:10.1186/1471-2180-10-161

CCME. (1999). *Canadian Environmental Quality Guidelines*. Retrieved August 12, 2013 from http://ceqg-rcqe.ccme.ca

Center for Disease Control. (2005). *Epidemiology of E. coli O157:H7 United States Outbreaks from 1982-2002.* Retrieved November 1, 2013 from http://wwwnc.cdc.gov/eid/article/11/4/04-0739_article.htm

Center for Disease Control (2011). *Salmonella Surveillance Overview*. Retrieved November 1, 2013 from http://www.cdc.gov/nationalsurveillance/PDFs/NationalSalmSurveillOverview_508.pdf

Cheng, C. M., Lin, W., Van, K. T., Phan, L., Tran, N. N., & Farmer, D. (2008). Rapid detection of *Salmonella* in foods using real-time PCR. *Journal of Food Protection*, *71*(12), 2436–41.

Clark C. G., Price L., Ahmed R., Woodward D. L., Melito P. L., Rodgers F. G., Jamieson F., Ciebin B., Li A. & Ellis A. (2003) Characterization of waterborne outbreak–associated *Campylobacter jejuni*, Walkerton, Ontario. *Emerging Infectious Diseases,* 9, 1232.

Clermont, O., Bonacorsi, S., Bingen, E., & Bonacorsi, P. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Applied and Environmental Microbiology, 66*(10). doi:10.1128/AEM.66.10.4555-4558.2000.Updated

Converse, R. R., Blackwood, A. D., Kirs, M., Griffith, J. F., & Noble, R. T. (2009). Rapid QPCR-based assay for fecal *Bacteroides* spp. as a tool for assessing fecal contamination in recreational waters. *Water Research*, *43*(19), 4828–4837. doi:10.1016/j.watres.2009.06.036

D'Agostino, M., Wagner, M., Vasquez-Boland, J.A., Kutcha, T., Karpiskova, R., Hoofar, J., Novella, S., Scortti, M., Ellison, J., Murray, A., Fernandes, I., Kuhn, M., Pazlarova, J., Heuvlink A., & Cook, N. (2004). A validated PCR-based method to detect *Listeria monocytogenes* using raw milk as a food model: towards an international standard. *Journal Food Protection, 67,* 1646–1655

Dick, L.K., Bernhard, A.E., Brodeur, T.J., Santo Domingo, J.W., Simpson, J.M., Walters, S.P., & Field, K.G., (2005a). Host distributions of uncultivated fecal *Bacteroidales* reveal genetic markers for fecal source identification. *Applied Environmental Microbiology, 71*, 3184–3191.

Dick, L.K., Simonich, M.T., Field, K.G., (2005b). Microplate subtractive hybridization to enrich for source-specific *Bacteroidales* fecal pollution indicators. *Applied Environmental Microbiology, 71*, 3179–3183.

Dick, L. K., Stelzer, E. A, Bertke, E. E., Fong, D. L., & Stoeckel, D. M. (2010). Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. *Applied and Environmental Microbiology*, *76*(10), 3255–62. doi:10.1128/AEM.02636-09

Dorr, P. M., Tadesse, D. A, Zewde, B. M., Fry, P., Thakur, S., & Gebreyes, W. A. (2009). Longitudinal study of *Salmonella* dispersion and the role of environmental contamination in commercial swine production systems. *Applied and Environmental Microbiology*, *75*(6), 1478–86. doi:10.1128/AEM.01632-08

Doumith, M., Buchrieser, C., Glaser, P., Jacquet, C., & Martin, P. (2004). Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. *Journal of Clinical Microbiology, 42*(8), 3819-3822. doi:10.1128/JCM.42.8.3819

Drozd, M., Merrick, N. N., Sanad, Y. M., Dick, L. K., Dick, W. A, & Rajashekara, G. (2012). Evaluating the occurrence of host-associated, general fecal indicators, and bacterial pathogens in a mixed-use watershed. *Journal of Environmental Quality*, *42*(3), 713–25. doi:10.2134/jeq2012.0359

Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chaventré, A, Elion, J., & Denamur, E. (2001). Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology (Reading, England)*, *147*(6), 1671–6.

Edberg, S.C., Rice E.W., Karlin, R.J., & Allen, M.J. (2000). *Escherichia coli:* the best biological drinking water indicator for public health protection*. Journal of Applied Microbiology, 88,* 106-116.

Edge. T. & Boehm, A. (2011). Classical and molecular methods to measure fecal bacteria. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.241-255). Washington DC: American Society for Microbiology.

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

Edge, T.A. & Schaefer, K.A. (2006). *Microbial Source Tracking in Aquatic Ecosystems: The State of the Science and an Assessment of Needs.* NWRI Scientific Assessment Report Series No. 7 and Linking Water Science to Policy Workshop Series (pp. 1-17). Burlington, Ontario: National Water Research Institute.

Eichmiller, J. J., Hicks, R. E., & Sadowsky, M. J. (2013). Distribution of genetic markers of fecal pollution on a freshwater sandy shoreline in proximity to wastewater effluent. *Environmental Science and Technology*, *47*(7), 3395–402. doi:10.1021/es305116c

Environment Canada. (2008). *Microbial Source Tracking, New Forensic Approaches.* Retrieved March 7, 2013 from http://www.ec.gc.ca/scitech/default.asp?lang=En&n=4B40916E-1&xsl=privateArticles2,viewfull&po=B2DDBB17

Field, K. G., & Samadpour, M. (2007). Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research*, *41*(16), 3517–38. doi:10.1016/j.watres.2007.06.056

Fremaux, B., Boa, T., & Yost, C. K. (2010). Quantitative real-time PCR assays for sensitive detection of Canada goose-specific fecal pollution in water sources. *Applied and Environmental Microbiology*, *76*(14), 4886–9. doi:10.1128/AEM.00110-10

Fremaux, B., Gritzfeld, J., Boa, T., & Yost, C. K. (2009). Evaluation of host-associated *Bacteroidales* 16S rRNA gene markers as a complementary tool for detecting fecal pollution in a prairie watershed. *Water Research*, *43*(19), 4838–4849. doi:10.1016/j.watres.2009.06.045

Fremaux, B., Prigent-Combaret, C., & Vernozy-Rozand, C. (2008). Long-term survival of Shiga toxin-producing *Escherichia coli* in cattle effluents and environment: an updated review. *Veterinary Microbiology*, *132*(1-2), 1–18. doi:10.1016/j.vetmic.2008.05.015

Gassman, P. W., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). SWAT modeling. *SWAT Review Series, 50*(4), 1211–1250.

Government of Canada (2012). *Listeria monocytogenes –Pathogen Safety Data Sheet*. Retrieved November 4, 2013 from http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/listeria- monocytogenes-eng.php

Green, H.C., Dick, L.K., Gilpin, B., Samadpour, M., & Field, K.G. (2012). Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. *Applied and Environmental Microbiology, 78(2),* 503-510.

Guillet, C., Join-Lambert, O., Le Monnier, A., Leclercq, A., Mechaï, F., Mamzer-Bruneel, M. F., & Lecuit, M. (2010). Human listeriosis caused by *Listeria ivanovii. Emerging Infectious Diseases*, *16*(1), 136–8. doi:10.3201/eid1601.091155

Halifax Regional Municipality (2013). *Source Water Protection*. Retrieved October 17, 2013 from http://www.halifax.ca/hrwc/

Halifax Regional Municpality (2013). *Water Treatment Facilities*. Retrieved October 17, 2013 from http://www.halifax.ca/hrwc/

Harwood, V. J., Levine, A. D., Scott, T. M., Chivukula, V., Lukasik, J., Farrah, S. R., & Rose, J. B. (2005). Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology, 71(6)*, 3163-3170. doi:10.1128/AEM.71.6.3163

Harwood, V.J., Ryu, H., & Santo Domingo, J. (2011). Microbial source tracking. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.190-202). Washington DC: American Society for Microbiology.

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K., & Korajkic, A. (2013). Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 1-40. doi:10.1111/1574-6976.12031

Health Canada. (2006). *Guidelines for Canadian Drinking Water Quality*: Guideline Technical Document — Bacterial Waterborne Pathogens — Current and Emerging Organisms of Concern. Water Quality and Health Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Health Canada. (2011). *Guidelines for Canadian Drinking Water Quality*: Guideline Technical Document — Enteric Viruses. Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario

Health Canada. (2012). *Guidelines for Canadian Drinking Water Quality*: Guideline Technical Document — Enteric Protozoa: *Giardia* and *Cryptosporidium*. Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Ho, A. J., Her, Z., Grohn, Y. T., McDonough, P. L., Wiedmann, M., Nightingale, K. K., & Fortes, E. D. (2004). Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. *Applied and Environmental Microbiology, 70*(8), 4458-4467 doi:10.1128/AEM.70.8.4458

Huang, D. B., & White, A. C. (2006). An updated review on *Cryptosporidium* and *Giardia*. *Gastroenterology Clinics of North America*, *35*(2), 291–314, viii. doi:10.1016/j.gtc.2006.03.006

Huey, G. M., & Meyer, M. L. (2010). Turbidity as an indicator of water quality in diverse watersheds of the Upper Pecos River Basin. *Water*, *2*(2), 273–284. doi:10.3390/w2020273

Ibekwe, A. M., Watt, P. M., Grieve, C. M., Sharma, V. K., & Lyons, S. R. (2002). Multiplex fluorogenic real-time PCR for detection and quantification of *Escherichia coli* O157:H7 in dairy wastewater wetlands. *Applied and Environmental Microbiology, 68*(10), 4853–4862. doi:10.1128/AEM.68.10.4853

Jacquet, C., Gouin, E., Jeannel, D., Cossart, P., & Rocourt, J. (2002). Expression of ActA, Ami, InlB, and Listeriolysin O in *Listeria monocytogenes* of human and food origin. *Applied and Environmental Microbiology, 68*(2), 616–622. doi:10.1128/AEM.68.2.616

Jiang, S. C., Chu, W., Olson, B. H., He, J.W., Choi, S., Zhang, J., & 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. *Applied Microbiology and Biotechnology*, *76*(4), 927–934. doi:10.1007/s00253-007-1047-0

Kapoor, V., Smith, C., Santo Domingo, J. W., Lu, T., & Wendell, D. (2013). Correlative assessment of fecal indicators using human mitochondrial DNA as a direct marker. *Environmental Science and Technology*, *47*(18), 10485–93. doi:10.1021/es4020458

Kérouanton, A., Marault, M., Petit, L., Grout, J., Dao, T. T., & Brisabois, A. (2010). Evaluation of a multiplex PCR assay as an alternative method for *Listeria monocytogenes* serotyping. *Journal of Microbiological Methods*, *80*(2), 134–137. doi:10.1016/j.mimet.2009.11.008

Khan, I. H., & Edge, T. A. (2007). Development of a novel triplex PCR assay for the detection and differentiation of thermophilic species of *Campylobacter* using 16S-23S rDNA internal transcribed spacer (ITS) region. *Journal of Applied Microbiology*, *103*(6), 2561–2569. doi:10.1111/j.1365-2672.2007.03511.x

Kildare, B. J., Leutenegger, C. M., McSwain, B. S., Bambic, D. G., Rajal, V. B., & Wuertz, S. (2007). 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: a Bayesian approach. *Water Research*, *41*(16), 3701–3715. doi:10.1016/j.watres.2007.06.037

Kortbaoui, R., Locas, A., Imbeau, M., Payment, P., & Villemur, R. (2010). Universal mitochondrial PCR combined with species-specific dot-blot assay as a source-tracking method of human, bovine, chicken, ovine, and porcine in fecal-contaminated surface water. *Water Research*, *43*(7), 2002–2010. doi:10.1016/j.watres.2009.01.030

Kreader, C.A., (1995). Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. *Applied and Environmental Microbiology, 61*, 1171–1179.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., & 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. *Applied and Environmental Microbiology, 72*(6), 4214–4224. doi:10.1128/AEM.01036-05

Lee, J.E., Lee, S., Sung, J., & Ko, G. (2011). Analysis of human and animal fecal microbiota for microbial source tracking. *ISME Journal, 5,* 362-365.

Litton, R. M., Ahn, J. H., Sercu, B., Holden, P. A, Sedlak, D. L., & Grant, S. B. (2010). Evaluation of chemical, molecular, and traditional markers of fecal contamination in an effluent dominated urban stream. *Environmental Science and Technology*, *44*(19), 7369–7375. doi:10.1021/es101092g

Lu, J., Santo Domingo, J., & Shanks, O.C. (2007) Identification of chicken- specific fecal microbial sequences using a metagenomic approach. *Water Research, 41,* 3561–3574

Lund, M., Nordentoft, S., Pedersen, K., & Madsen, M. (2004). Detection of *Campylobacter* spp. in chicken fecal samples by real-time qPCR. *Journal of Clinical Microbiology, 42*(11), 5125–5132. doi:10.1128/JCM.42.11.5125

Lyautey, E, Hartmann, a, Pagotto, F., Tyler, K., Lapen, D. R., Wilkes, G., & Topp, E. (2007). Characteristics and frequency of detection of fecal *Listeria monocytogenes* shed by livestock, wildlife, and humans. *Canadian Journal of Microbiology*, *53*(10), 1158–1167. doi:10.1139/W07-084

Lyautey, E., Lapen, D. R., Wilkes, G., McCleary, K., Pagotto, F., Tyler, K., & Topp, E. (2007). Distribution and characteristics of *Listeria monocytogenes* isolates from surface waters of the South Nation River watershed, Ontario, Canada. *Applied and Environmental Microbiology*, *73*(17), 5401–10. doi:10.1128/AEM.00354-07

MacDonald,F., & Sutherland, A.D (1994). Important differences between the generation times of *Listeria monocytogenes* and *Listeria innocua* in two *Listeria* enrichment broths. *Journal of Dairy Research, 61*, 433-436.

Marion, J. W., Lee, J., Lemeshow, S., & Buckley, T. J. (2010). Association of gastrointestinal illness and recreational water exposure at an inland U.S. beach. *Water Research*, *44*(16), 4796–804. doi:10.1016/j.watres.2010.07.065

Marsalek, J. & Schreier, H. (2008). *Innovation in Stormwater Management in Canada : The Way Forward.* Retrieved March 7, 2013 from http://www.mcgill.ca/files/water2010/Article-Marsalek_and_Schreier.pdf

Marti, R., Zhang, Y., Lapen, D.R., & Topp, E. (2011). Development and validation of a microbial source tracking marker for the detection of fecal pollution by muskrats. *Journal of Microbiological Methods, 87(1)*, 82-88.

Mauffret, A., Caprais, M.P., & Gourmelon, M. (2012). Relevance of *Bacteroidales* and F-specific RNA bacteriophages for efficient fecal contamination tracking at the level of a catchment in France. *Applied and Environmental Microbiology*, *78*(15), 5143–52. doi:10.1128/AEM.00315-12

McElhany, K.G. & Pillai, S.D. (2011). Prevalence and fate of gut associated pathogens in the environment. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.217-240). Washington DC: American Society for Microbiology.

McEgan, R., Mootian, G., Goodridge, L. D., Schaffner, D. W., & Danyluk, M. D. (2013). Predicting *Salmonella* populations from biological, chemical, and physical indicators in Florida surface waters. *Applied and Environmental Microbiology*, *79*(13), 4094–105. doi:10.1128/AEM.00777-13

Moffatt, H., & Struck, S. (2011). *Water-borne Disease Outbreaks in Canadian Small Drinking Water Systems.* Vancouver, British Colombia: National Collaborating Centres for Public Health.

McQuaig, S., Griffith, J., & Harwood, V. J. (2012). Association of fecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. *Applied and Environmental Microbiology*, *78*(18), 6423–32. doi:10.1128/AEM.00024-12

NCBI (2013). *Polymerase Chain Reaction*. Retrieved November 1, 2013 from http://www.ncbi.nlm.nih.gov/projects/genome/probe/doc/TechPCR.shtml

NWRI Scientific Assessment Report (2006). *Microbial Source Tracking: State of the Science and Assessment of Needs*. Retrieved March 7, 2013 from http://www.ec.gc.ca/inre-nwri/default.asp?lang=En&n=D575CDF5-1

NS Wastewater Society (2009). *A Homeowner's Guide to Septic Systems.* Retrieved November 10, 2013 from http://www.novascotia.ca/nse/wastewater/docs/Homeowners.Guide.to.Septic.Systems.pdf

O'Connor, D. (2002). *Report of the Walkerton Commission of Inquiry.* Retrieved November 10, 2013 from http://www.attorneygeneral.jus.gov.on.ca/english/about/pubs/walkerton/part1/

OIE Terrestrial Manual. (2008). *Campylobacter jejuni* and *Campylobacter coli.* Retrieved November 1, 2013 from http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.09.03_CAMPYLO.pdf

Ogden, I. D., Dallas, J. F., MacRae, M., Rotariu, O., Reay, K. W., Leitch, M., & Strachan, N. C. (2009). *Campylobacter* excreted into the environment by animal sources: prevalence, concentration shed, and host association. *Foodborne Pathogens and Disease*, *6*(10), 1161–70. doi:10.1089/fpd.2009.0327

Okabe, S., Okayama, N., Savichtcheva, O., & Ito, T. (2007). Quantification of host-associated *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology*, *74*(4), 890–901. doi:10.1007/s00253-006-0714-x

Oliver, J.D. (2005). The viable but nonculturable state in bacteria. *Journal of Microbiology, 43*, 93-100

Orsi R. H., Bakker H. C. & Wiedmann M. (2011) *Listeria monocytogenes* lineages: Genomics, evolution, ecology, and phenotypic characteristics. *International Journal of Medical Microbiology, 301*, 79-96. .

Orji, M. U., Onuigbo, H. C., & Mbata, T. I. (2005). Isolation of *Salmonella* from poultry droppings and other environmental sources in Awka, Nigeria. *International Journal of Infectious Diseases, 9*(2), 86–9. doi:10.1016/j.ijid.2004.04.016

Park, H., Lee, K., Hwang, J., & Ka, J. (2010). Quantitative analysis of human and cow-specific 16S rRNA gene markers for assessment of fecal pollution in river waters by real-time PCR. *Journal of Microbiology and Biotechnology, 20*(2), 245–253. doi:10.4014/jmb.0908.08022

Peed, L. A, Nietch, C. T., Kelty, C. A, Meckes, M., Mooney, T., Sivaganesan, M., & Shanks, O. C. (2011). Combining land use information and small stream sampling with PCR-based methods for better characterization of diffuse sources of human fecal pollution. *Environmental Science & Technology*, *45*(13), 5652–9. doi:10.1021/es2003167

Petridis, H., Kidder, G., & Ogram, A. (2002). *E. coli O157 : H7 A Potential Health Concern.* Retrieved October 24, 2013 from http://edis.ifas.ufl.edu/pdffiles/SS/SS19700.pdf

Piorkowski, G. (2013). Genetic diversity of *E. coli* in solids and sediments of an agricultural watershed and their spatiotemporal influences on water quality. Phd Thesis-Dalhousie University, Nova Scotia, Canada.

Piorkowski, G. S., Jamieson, R. C., Hansen, L. T., Bezanson, G. S., & Yost, C. K. (2013). Characterizing spatial structure of sediment *E. coli* populations to inform sampling design. *Environmental Monitoring and Assessment*. doi:10.1007/s10661-013-3373-2

Quigley, E., & Keys, K. (2007). *Forest Soil Types of Nova Scotia*. Retrieved November 10, 2013 from http://www.novascotia.ca/NATR/forestry/reports/NS-Soils.pdf

Reischer, G H, Kasper, D. C., Steinborn, R., Farnleitner, A. H., & Mach, R. L. (2007). A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. *Letters in Applied Microbiology, 44*, 351–356. doi:10.1111/j.1472-765X.2006.02094.x

Reischer, Georg H, Kasper, D. C., Steinborn, R., Mach, R. L., & Farnleitner, A. H. (2006). Quantitative PCR Method for Sensitive Detection of Ruminant Fecal Pollution in Freshwater and Evaluation of This Method in Alpine Karstic Regions. *Applied and Environmental Microbiology, 72*(8), 5610–5614. doi:10.1128/AEM.00364-06

Reischl, U. & Kochanowski, B. (1999). *Methods in Molecular Medicine: Quantitative PCR Protocols*: Totowa, NJ: Humana Press.

Ridley, C., Yost, C., Truelstrup Hansen, L., & Jamieson, R.J. (2011, February 28). Microbial source tracking within agricultural watersheds. *Canadian Water Network Annual Meeting.* Ottawa, ON.

Roach, P. D., Olson, M. E., Whitley, G., & Wallis, P. M. (1993). Waterborne *Giardia* cysts and *Cryptosporidium* oocysts in the Yukon, Canada. *Applied and Environmental Microbiology, 59*(1), 67-73.

Rodríguez-Lázaro, D., Hernández, M., Scortti, M., Esteve, T., Vázquez-Boland, J. A., Pla, M., & Herna, M. (2004). Quantitative detection of *Listeria monocytogenes* and *Listeria innocua* by real-time PCR: Assessment of *hly, iap*, and *lin02483* targets and ampliflour technology. *Applied and Environmental Microbiology, 70*(3), 1366-1377. doi:10.1128/AEM.70.3.1366

Rosen. B. (2000). *Waterborne Pathogens in Agricultural Watersheds.* Retrieved March 7, 2013 from ftp://ftpfc.sc.egov.usda.gov/WSI/pdffiles/Pathogens_in_Agricultural_Watersheds.pdf

Roslev, P., & Bukh, A. S. (2011). State of the art molecular markers for fecal pollution source tracking in water. *Applied Microbiology and Biotechnology, 89,*1341–1355. doi:10.1007/s00253-010-3080-7

Ryser, E.T & Marth E.H. (2007). *Listeria, Listeriosis & Food Safety.* Boca Raton: Taylor and Francis Group

Sauders, B. D., Overdevest, J., Fortes, E., Schukken, Y., Lembo, A., Windham, K., & Wiedmann, M. (2012). Diversity of *Listeria* species in urban and natural environments. *Applied and Environmental Microbiology, 78*(12), 4420-4433. doi:10.1128/AEM.00282-12

Sauders B. D., & Wiedmann M. (2007) Ecology of *Listeria* species and *L. monocytogenes* in the natural environment. In E.T Ryser & E.H Marth (3rd Ed) *Listeria, Listeriosis & Food Safety.* pp. 21-44. Boca Raton: Taylor and Francis Group.

Sauer, E. P., Vandewalle, J. L., Bootsma, M. J., & McLellan, S. L. (2011). Detection of the human specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research*, *45*(14), 4081–91. doi:10.1016/j.watres.2011.04.049

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

Schill, W.B. & Mathes, M.V. (2008). Real-Time PCR detection and quantification of nine potential sources of fecal contamination by analysis of mitochondrial Cytochrome b Targets. *Environmental Science and Technology, 42*(14), 5229–5234.

Schriewer, A., Miller, W. A, Byrne, B. A, Miller, M.A, Oates, S., Conrad, P. A, & Wuertz, S. (2010). Presence of *Bacteroidales* as a predictor of pathogens in surface waters of the central California coast. *Applied and Environmental Microbiology*, *76*(17), 5802–14. doi:10.1128/AEM.00635-10

Seurinck, S., Defoirdt, T., Verstraete, W., & Siciliano, S. D. (2005). Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, *7*(2), 249–59. doi:10.1111/j.1462-2920.2004.00702.x

Shanks, O. C., Atikovic, E., Blackwood, A. D., Lu, J., Noble, R. T., Domingo, J. S., & Seifring, S. (2008). Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology, 74*(3), 745-752. doi:10.1128/AEM.01843-07

Sidhu, J. S., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., & Toze, S. (2013). Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *The Science of the Total Environment*, *463-464*, 488–496. doi:10.1016/j.scitotenv.2013.06.020

Simpson, J. M., Domingo, J. W. S., & Reasoner, D. J. (2002). Critical review microbial source tracking : state of the science. *Environmental Science and Technology, 36*(24), 5279–5288.

Sinclair, A., Hebb, D., Jamieson, R., Gordon, R., Benedict, K., Fuller, K., & Madani, A. (2009). Growing season surface water loading of fecal indicator organisms within a rural watershed. *Water Research*, *43*(5), 1199–206. doi:10.1016/j.watres.2008.12.006

Stewart, J.R., Santo Domingo, J., & Wade, T.J. (2007). Fecal pollution, public health and microbial source tracking. In: J. Santo Domingo & M. Sadowsky (Eds.), *Microbial Source Tracking* (Ch. 1). Washington, DC: American Society for Microbiology Press.

Stumpf, C. H., Piehler, M. F., Thompson, S., & Noble, R. T. (2010). Loading of fecal indicator bacteria in North Carolina tidal creek headwaters: hydrographic patterns and terrestrial runoff relationships. *Water Research*, *44*(16), 4704–15. doi:10.1016/j.watres.2010.07.004

Tambalo, D.D., Fremaux, B., Boa, T., & Yost, C.K. (2012). Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. *Water Research, 46,* 2891-2904.

Taskin, B., Gozen, A. G., & Duran, M. (2011). Selective quantification of viable *Escherichia coli* bacteria in biosolids by quantitative PCR with propidium monoazide modification. *Applied and Environmental Microbiology*, *77*(13), 4329–35. doi:10.1128/AEM.02895-10

Tobe, S. S., & Linacre, A. T. (2008). A multiplex assay to identify 18 European mammal species from mixtures using the mitochondrial cytochrome b gene. *Electrophoresis, 29,* 340–347. doi:10.1002/elps.200700706

U.S. Environmental Protection Agency. (2002). *A Pocket Guide to Protecting Your Drinking Water*. Drinking water pocket guide #3. Retrieved March 7, 2013 from www.epa.gov/safewater

U.S. Environmental Protection Agency. (2005). *Microbial Source Tracking Guide Document*. Cincinnati, OH: National Risk Management Research Laboratory.

Van der Wielen, P.W. & Medema, G. (2010). Unsuitability of quantitative *Bacteroidales* 16S rRNA gene assays for discerning fecal contamination of drinking water. *Applied and Environmental Microbiology, 76*(14), 4876-4881.

Vierheilig, J., Farnleitner, A. H., Kollanur, D., Blöschl, G., & Reischer, G. H. (2012). High abundance of genetic *Bacteroidetes* markers for total fecal pollution in pristine alpine soils suggests lack in specificity for feces. *Journal of Microbiological Methods*, *88*(3), 433–435. doi:10.1016/j.mimet.2012.01.009

Walk, S. T., Alm, E. W., Calhoun, L. M., Mladonicky, J. M., & Whittam, T. S. (2007). Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environmental Microbiology*, *9*(9), 2274–2288. doi:10.1111/j.1462-2920.2007.01341.x

Walters, S.P., Gannon, V.J., & Field, K.G. (2007). Detection of *Bacteroidales* fecal indicators and the zoonotic pathogens *E. coli* 0157:H7, *Salmonella*, and *Campylobacter* in river water. *Environmental Science and Technology, 41*, 1856–1862.

Walters, S. P., González-Escalona, N., Son, I., Melka, D. C., Sassoubre, L. M., & Boehm, A. B. (2013). *Salmonella enterica* diversity in central Californian coastal waterways. *Applied and Environmental Microbiology*, *79*(14), 4199–4209. doi:10.1128/AEM.00930-13

Walters, S. P., Thebo, A. L., & Boehm, A. B. (2011). Impact of urbanization and agriculture on the occurrence of bacterial pathogens and stx genes in coastal waterbodies of central California. *Water Research*, *45*(4), 1752–1762. doi:10.1016/j.watres.2010.11.032

Ward, T. J., Usgaard, T., & Evans, P. (2010). A targeted multilocus genotyping assay for lineage, serogroup, and epidemic clone typing of *Listeria monocytogenes*. *Applied and Environmental Microbiology*, *76*(19), 6680–4. doi:10.1128/AEM.01008-10

Werfhorst, L. C., Van De Sercu, B., & Holden, P. A. (2011). Comparison of the host specificities of two *Bacteroidales* quantitative PCR assays used for tracking human fecal contamination. *Applied and Environmental Microbiology, 77*(17), 6258–6260. doi:10.1128/AEM.00239-11

Whitman, R. L., Shively, D. A., Pawlik, H., Nevers, M. B., & Byappanahalli, M. N. (2003). Occurrence of *Escherichia coli* and Enterococci in *Cladophora* (Chlorophyta) in nearshore water and beach sand of Lake Michigan. *Applied and Environmental Microbiology, 69*(8), 4714-19. doi:10.1128/AEM.69.8.4714

Wilkes, G., Brassard, J., Edge, T. A, Gannon, V., Jokinen, C. C., Jones, T. H., & Lapen, D. R. (2013). Coherence among different microbial source tracking markers in a small agricultural stream with or without livestock exclusion practices. *Applied and Environmental Microbiology*, *79*(20), 6207–19. doi:10.1128/AEM.01626-13

Winter, T. C., Harvey, J. W., Franke, O. L., & Alley, W. M. (1998). Ground water and surface water: a single resource. *U. S. Geological Survey Circular, 1139*. Denver, Colorado.

Wolf, S., Hewitt, J., & Greening, G. E. (2010). Viral multiplex quantitative PCR assays for tracking sources of fecal contamination. *Applied and Environmental Microbiology*, *76*(5), 1388–94. doi:10.1128/AEM.02249-09

World Health Organization. (2004). *Water Treatment and Pathogen Control*. Retrieved November 8, 2013 from http://apps.who.int/iris/bitstream/10665/42796/1/9241562552.pdf

World Health Organization. (2011). *Campylobacter*. Retrieved November 10, 2013 from http://www.who.int/mediacentre/factsheets/fs255/en/

World Health Organization. (2013). *Water Quality and Health Strategy*. Retrieved November 1, 2013 from http://www.who.int/water_sanitation_health/publications/2013/water_quality_strategy/en

Wyer, M. D., Kay, D., Watkins, J., Davies, C., Kay, C., & Moore, H. (2010). Evaluating short-term changes in recreational water quality during a hydrograph event using a combination of microbial tracers, environmental microbiology, microbial source tracking and hydrological techniques. *Water Research*, *44*(16), 4783–4795. doi:10.1016/j.watres.2010.06.047

Yost. C., Diarra, M. & Topp, E. (2011). Animals and humans as sources of fecal indicator bacteria. In M.J. Sadowsky & R.L. Whitman (Eds.) *The Fecal Bacteria.* (pp.67-91). Washington DC: American Society for Microbiology.

APPENDIX A SUPPLEMENTARY RESULT DATA

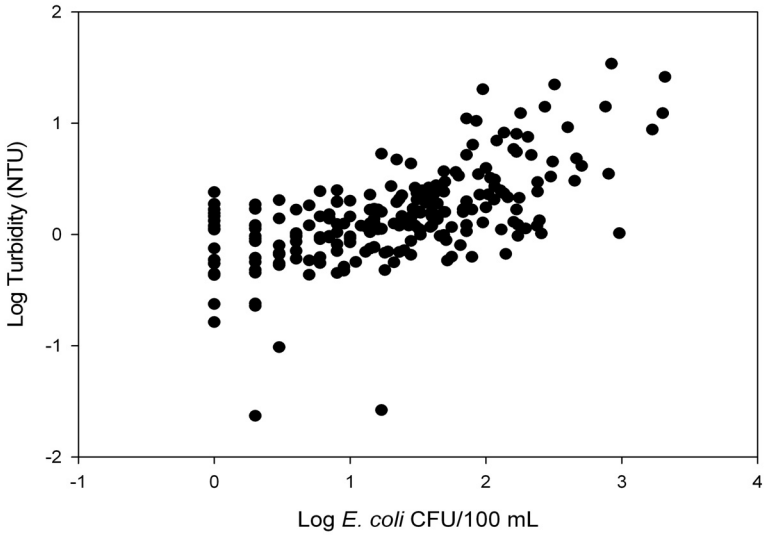

Figure A-1 Relationship between *E. coli* and turbidity

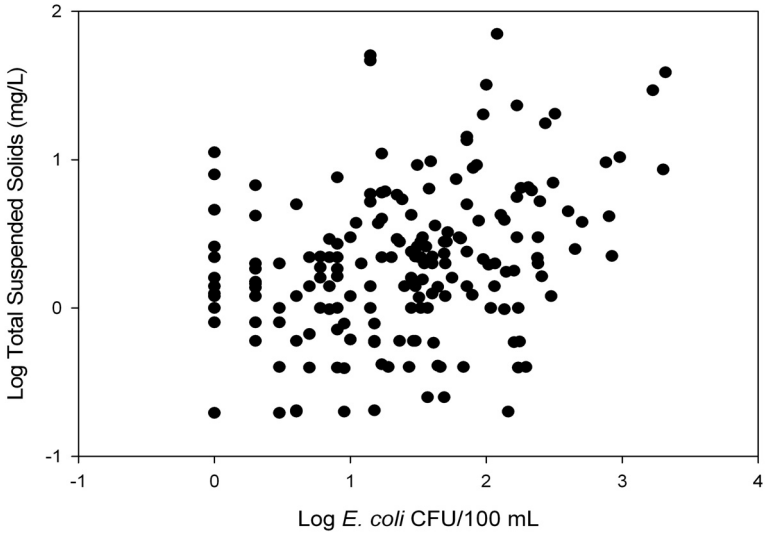

Figure A-2 Relationship between *E. coli* and total suspended solids

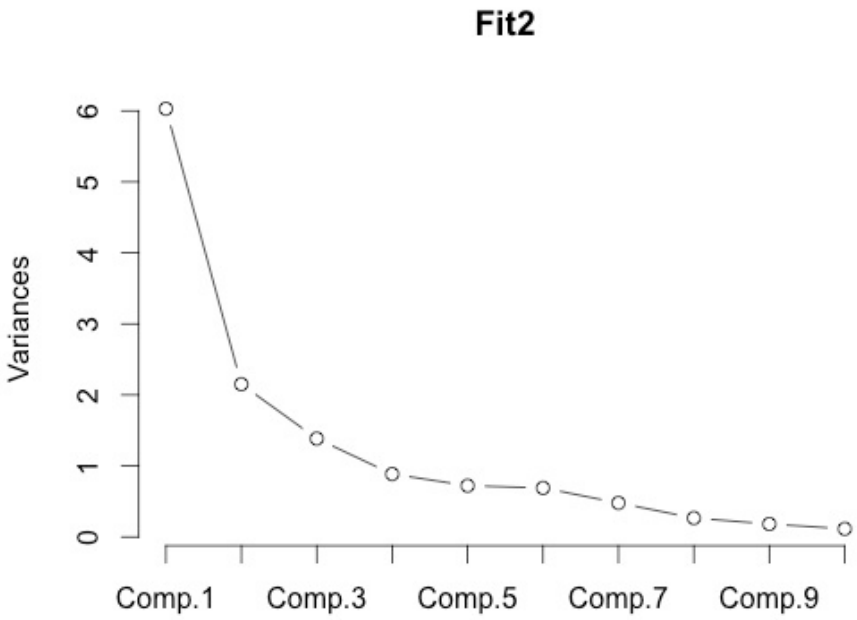

Figure A-3 Scree plot for PCA of environmental variables

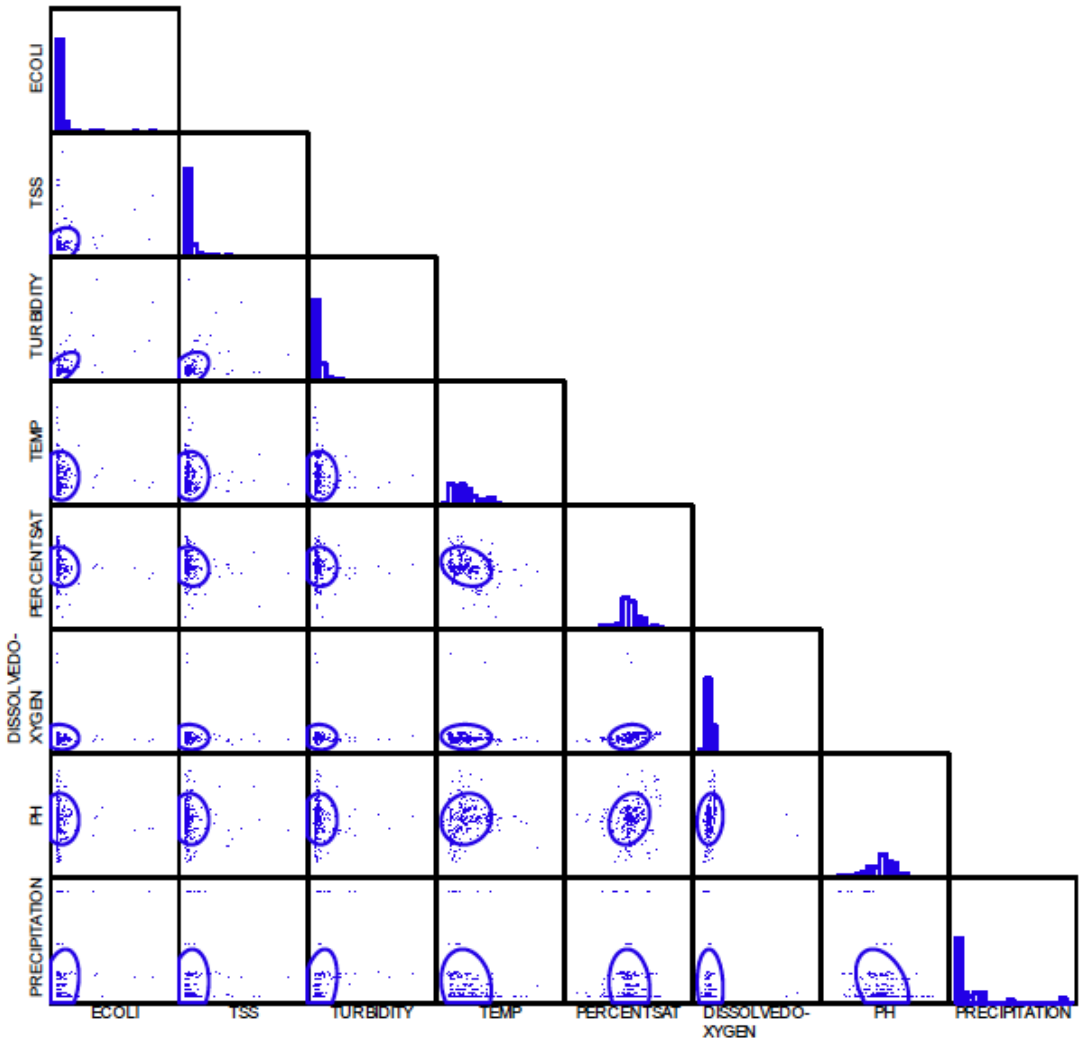

Figure A-4 Scatterplot matrix of environmental variables

Table A-1 Conditional probabilities using Bayes' Theorem (Kildare et al., 2007) for the HF183, HcytB, BacR, AcytB and CowM2 markers detected in this study

Marker	Conditional Probability (%)
HF183	44.5
HcytB	53.0
BacR	32.0
AcytB	15.2
CowM2	100.0

APPENDIX B SOIL AND EROSION MAPS OF MM

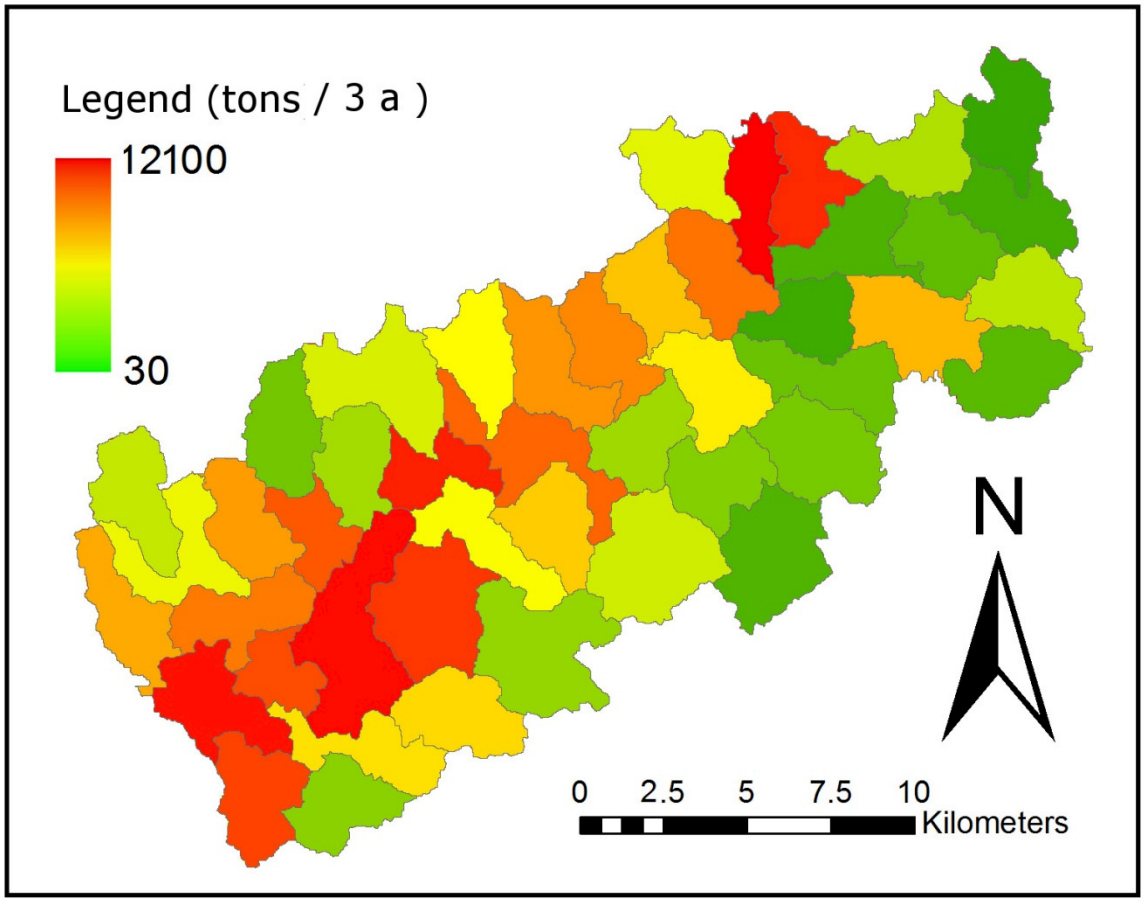

Figure B-1 Erosion map of Middle Musquodoboit (tons/acre over 3 years)

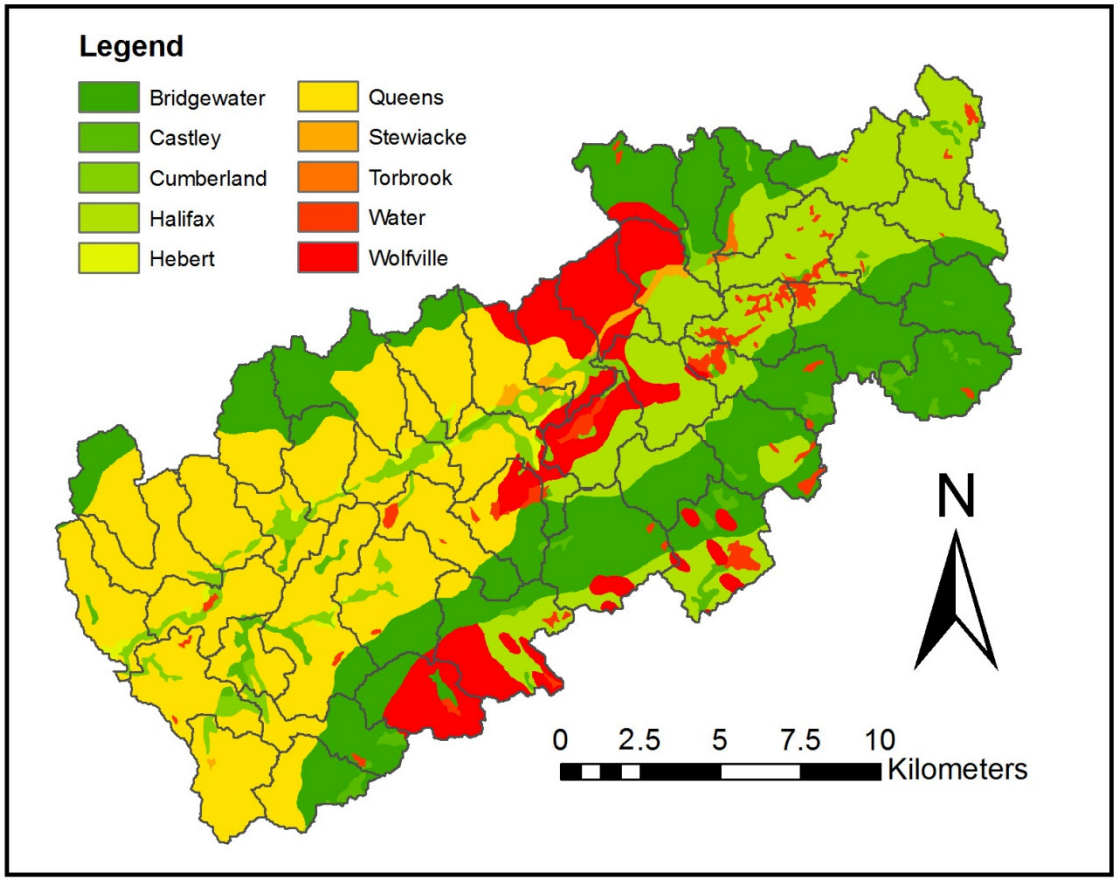

Figure B-3 Soils map of Middle Musquodoboit

APPENDIX C ALIGNMENT/PRIMER BLAST DATA FOR HCYTB AND ACYTB

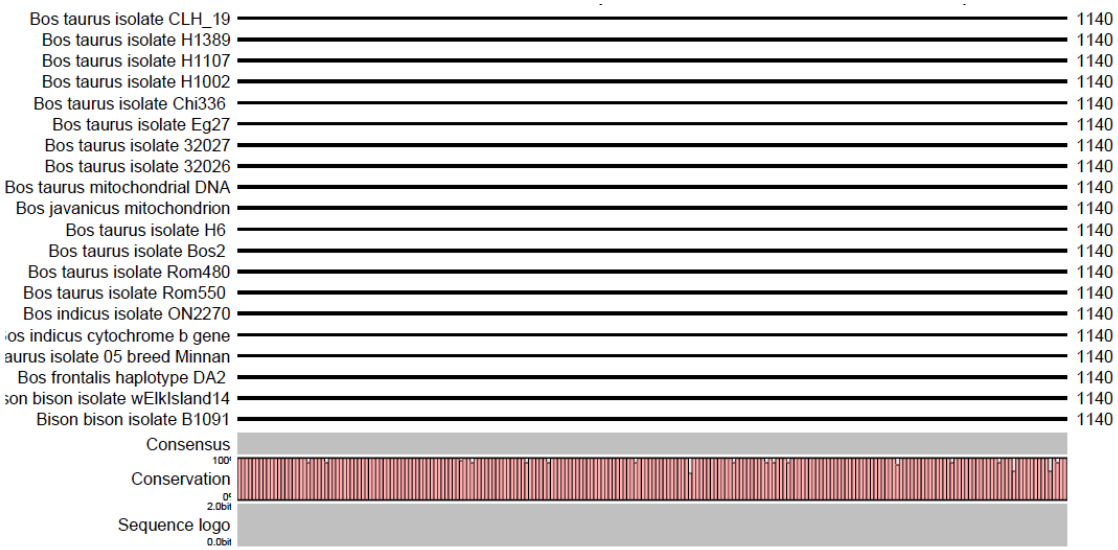

Figure C-1 AcytB cow marker alignment using various cow DNA sequences collected from the NCBI database and aligned using ClustalW within CLC Bio Workbench 6.9.

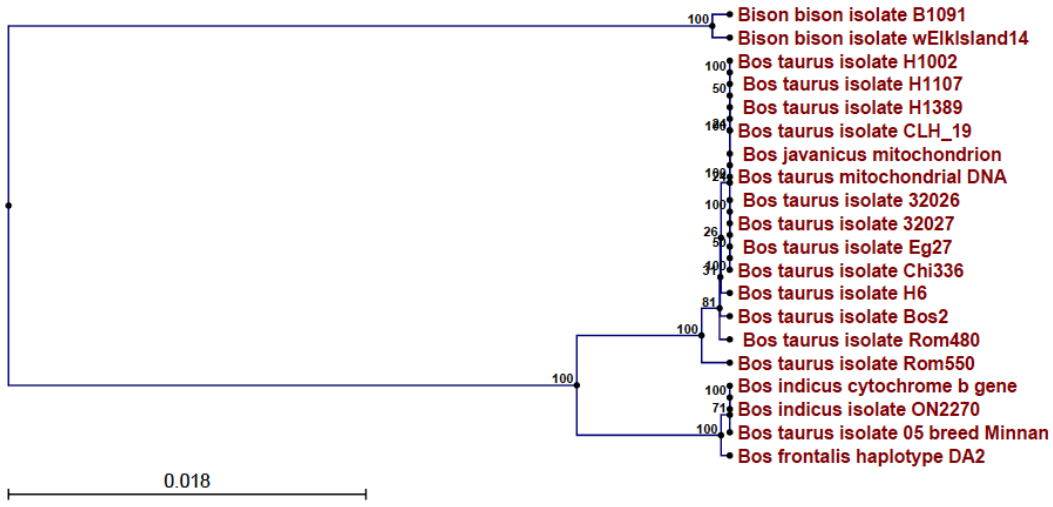

Figure C-2 Similarity tree for the cow cytochrome B sequences

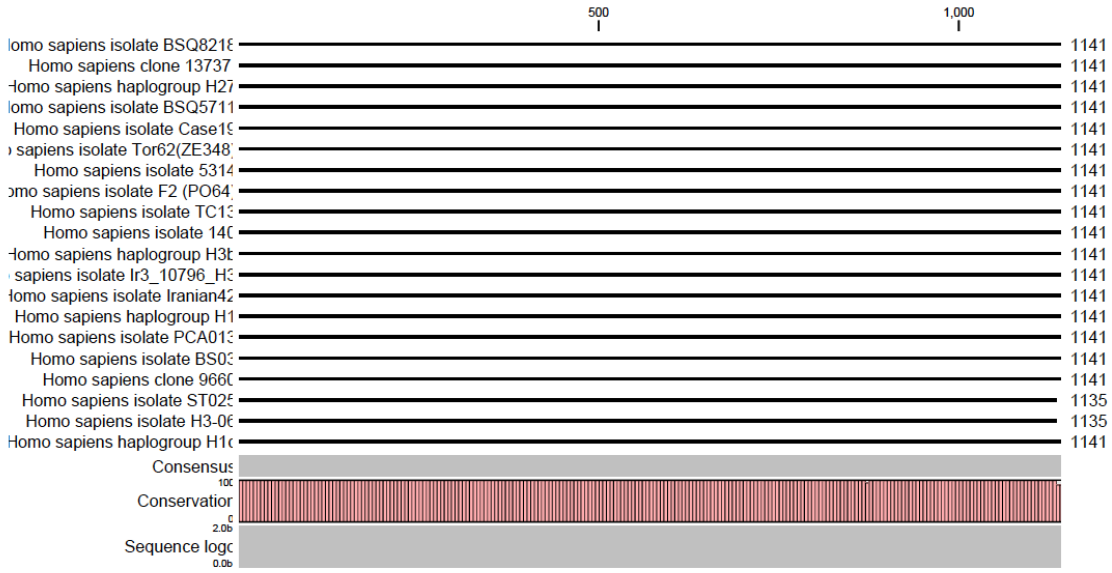

Figure C-3 HcytB human marker alignment using various human sequences collected from the NCBI database and aligned using ClustalW within CLC Bio Workbench 6.9.

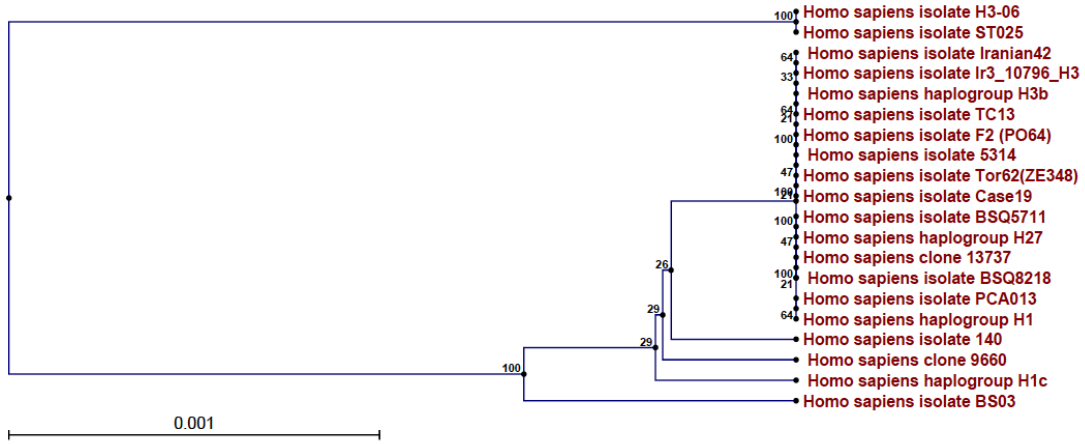

Figure C-4 Similarity tree for human cytochrome B sequences