Development, Validation, and Application of Molecular Microbial Source Tracking Methods to be Used in the Assessment of Environmental Waterways

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Richard Charles Sawdey

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Michael J. Sadowsky

November 2009

Acknowledgements

I would like to thank my advisor, Dr. Michael J. Sadowsky, for his teaching and extended guidance throughout the course of my studies. Dr. Sadowsky afforded me the opportunity to extend my pursuits in higher education and has been a friend along the way. I would also like to thank Dr. James Anderson and Dr. James Cotner for serving on my committee and for what they have taught me.

Several people made the progression of my projects possible. I would first like to acknowledge the teachings and guidance of Matthew Hamilton, who was instrumental in my success as a student, and became a dear friend along the way. I would also like to acknowledge Dan Norat for his numerous consultations, and help with sample collection and processing. Further, I would like to thank Satoshi Ishii for his guidance and support. I would also like to thank my lab members for their suggestions, guidance, and laughs along the way.

Additionally, I would like to thank my wonderful wife Sara, whom without this would not have been possible. I am indebted for her limitless support and love. I would also like to thank my parents, Mary Beth and Stan Sawdey. They provided me every opportunity, and I would not have sought this degree without their support, encouragement, and love. Finally, I would like to thank my brother Joe Sawdey, whose inspiration led me down the path I am on today. Thank you.

Dedication

For Sara. I love you.

Abstract

The fecal loading of aquatic environments by various animal hosts is of concern for public and environmental health. The microbiological contamination of waterways is amongst the most commonly listed water quality impairment in the U. S. The federal Clean Water Act (CWA) requires individual states to provide the U. S. Environmental Protection Agency (USEPA) with an aggregate quality assessment of its waterways every biennium, in effort to identify waters that do not meet state and federal quality standards. Once waterways are deemed impaired, states must conduct a Total Maximum Daily Load (TMDL) assessment in order to mitigate the impairment and restore the water body to acceptable quality. Successful bacterial TMDL implementation strategies require the use of microbial source tracking (MST) technologies that accurately and efficiently characterize the host-specific source of bacterial loading of waterways, and the relative quantities of each bacterium. Here I report that suppression subtraction hybridization (SSH) was found useful to identify gene markers specific to *E. coli* derived from swine fecal sources. The ability of this marker gene set to identify 62.3% of *E. coli* isolated from swine hosts suggests it may be useful in determining their fecal contribution to impaired waterways. I also investigated the influence of cattle grazing operations on the microbiological impairment of a small stream system in Southeastern Minnesota. Impairment by fecal indicator bacteria (FIB) was assessed by using plate count analyses and a quantitative PCR (q-PCR) was developed to estimate the presence of a bovine-specific *Bacteroides* marker gene in the waterway. The q-PCR data were compared to *E. coli* plate count data, revealing a lack of correlation between the two methods. Several

physical and environmental factors likely influenced the level of *E. coli* found in the stream, confirming the hypothesis that other information will be needed to supplement current efforts to monitor fecal indicator bacteria in order to determine accurate source-specific fecal impacts. Lastly, spatial and temporal variation in the population structure of a fecal pathogen (*Salmonella)* in association with an alternate host and habitats, including the green alga *Cladophora* found in stream and lake water, aquatic plants, beach sand, and sediments, was evaluated by use of horizontal fluorophore-enhanced rep-PCR (HFERP) DNA fingerprinting. It was revealed that *Salmonella* populations associated with *Cladophora* varied both spatially and temporally, suggesting potentially different input sources of *Salmonella* over space and time. In addition, differing environmental stressors may play a role in selecting particular *Salmonella* genotypes that are best suited for growth in these environments The use of rapid molecular-based assays to determine the presence and source-specific loading of fecal indicator bacteria and pathogens has the potential to improve the accuracy and efficacy of TMDL studies, and to expedite implementation strategies to remediate impaired waterways.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Fecal bacteria in waterways

The fecal loading of aquatic environments by various animal hosts is of concern for public and environmental health. The microbiological contamination of waterways is amongst the most commonly listed water quality impairment in the U. S. (USEPA, 2005). The federal Clean Water Act (CWA) requires individual states to provide the Environmental Protection Agency (USEPA) with an aggregate quality assessment of its waterways every biennium in effort to identify waters that do not meet state and federal quality standards. Once waterways are deemed impaired, states must conduct a Total Maximum Daily Load (TMDL) assessment in order to mitigate the impairment and restore the water body to acceptable quality. In the case of fecal bacteria impairments, determination of fecal sources to a given body of water is of paramount importance for calculations of TMDL wasteload and load allocations. Techniques to determine sources of fecal bacteria in the environment have been developed and are referred to as microbial source tracking methods (MST). Traditional MST methodologies often do not accurately identify the sources and quantity of fecal bacteria, can be time intensive, require specialized personnel to conduct the studies, and are expensive. Thus, successful TMDL implementation strategies will require the use of MST technologies that accurately and efficiently characterize the host-specific source of bacterial loading of waterways, and the relative quantities of each bacterium. The use of rapid molecular-based assays to determine fecal bacterial sources and source-specific loading quantities has the potential

to improve the accuracy and efficacy of TMDL studies, and to expedite their implementation strategies.

1.2. Water quality regulation

1.2.1 Federal Water Pollution Control Act and the Clean Water Act.

In 1948, Congress passed the Federal Water Pollution Control Act (FWPCA) mandating that the sanitary condition of the nation's surface and groundwater be addressed. The FWPCA marked the first federal mandate to curb distribution of emerging contaminants that threatened public and environmental health. Since its passage, the FWPCA has been amended numerous times, typically to enhance water quality programs, determine discharge allowances, and to insert a framework for state funding to fulfill the mandates of the Federal Code (USFWS, 2009). Notably, the FWPCA amendments of 1972 mandated the restoration and maintenance of the chemical, physical, and biological integrity for the Nation's waters (86 Stat. 47; P.L. 92-500, 33 U.S.C. 1251), and the legislation became known as the Clean Water Act (CWA).

Paramount to this framework, was the mandate that States and Tribal entities establish Water Quality Standards (WQS) as a measure to initiate the quality assessment process of surface waters nationwide (33 U.S.C. 1313). More specifically, section 303 (c) of the WQS program requires States and Tribes to designate uses for the regulated water body. Designated uses that require scale assessments of fecal contamination are recreational use (primary and secondary), public water supplies, aquifer protection, and protection and propagation of fish, shellfish, and wildlife (USEPA, 2001). Once WQS are

set for respective designated uses, States and Tribes are to assess their waters on a biennial basis. Known as the section 305(b) list, the purpose of the biennial water quality assessment is to identify regulated water bodies that meet or exceed WQS per designated use; and those that fail to meet WQS per designated use. Bodies of water that meet or exceed WQS goals are subject to anti-degradation laws within the syntax of section 305 of the CWA. Bodies of water that do not meet designated WQS are subsequently placed on the section 303(d) impaired waters list. Section 303(d) list waters are then subject to Total Maximum Daily Load (TMDL) studies and subsequent TMDL implementation plans to restore water quality. The 1972 amendments also launched the National Pollution Discharge Elimination system (NPDES), which required municipalities and industrial entities to obtain permits when discharging contaminants of concern to regulated water bodies.

1.2.2 Minnesota Clean Water Legacy Act.

Passed in 2006, the Minnesota Clean Water Legacy Act (CWLA) was enacted "to protect, restore, and preserve the quality of Minnesota's surface waters." The act provided authority, direction, and resources to achieve and maintain water quality standards for surface waters as required by section 303(d) of the federal Clean Water Act. An important aspect of the Minnesota CWLA is that it mandates that the public and private entities involved in TMDL determinations are charged with the implementation of restoration. Also, the CWLA required funding for microbial contaminant research. The Minnesota Department of Agriculture (MDA) is currently provided with funding to issue

grants that are used to study and identify agricultural contributions to microbiological impairments that have historically been associated with an increased risk of waterborne pathogen exposure (MPCA, 2009).

1.3. Waterborne pathogens

1.3.1 Introduction to waterborne pathogens.

Waterborne pathogens present a persistent threat to human health (Cabelli et al., 1979; Dufour, 1977; Sharma et al., 2003; Szewzyk et al., 2000). In 1971, the Center for Disease Control (CDC), along with the USEPA and Council of State and Territorial Epidemiologists (CSTE) launched the Waterborne Disease and Outbreak Surveillance System (WBDOSS) to track waterborne disease outbreaks associated with drinking water, and extended the effort to include recreational water in 1978 (CDC, 2008). Pathogenic agents commonly cited by the WBDOSS are; the parasites *Cryptosporidium* and *Giardia*, the bacteria *Campylobacter*, *Shigella*, *Pseudomonas*, *Legionella*, *Salmonella* and *Escherechia*; and the viral agents norovirus and hepatitis B.

While it may be intuitive that waterborne pathogen monitoring efforts test for the presence of infectious agents directly, it is not practical. This is due to the unculturability of some pathogenic organisms, their often sporadic distribution, but potentially low infectious dose, and the cost of simultaneously analyzing samples for a several pathogens (Field and Samadpour, 2007).

1.3.2 Waterborne pathogen ecology.

It has been shown that some pathogenic bacteria propagate in secondary habitats once introduced into aquatic environments. Ishii et al. (2006(a)) reported that the alga *Cladophora* (Cladophoraceae), harbors enteric bacterial pathogens including *Shigella*, *Campylobacter*, *Salmonella*, and Shiga toxin-producing *E. coli* (STEC). Similarly, *Camplyobacter jejuni* has been isolated from soil, surface water, beach sand, and waterborne protozoans (Bolton et al., 1999; Colles et al., 2003; Schallenberg et al., 2005). Further, it has been shown that *Salmonella, Shigella,* and *E. coli* have the ability to survive in soils, sediments and surface water (Byappanahalli and Fujioa, 2004; Ishii et al., 2006(b); Whitman et al., 2001; Winfield and Groisman, 2003). While these studies identified the relationship of select pathogens to various environmental hosts, work is still needed in this area, as it is likely that additional aquatic habitats exist for these organisms. Future efforts to reduce the risk of potential waterborne pathogen exposure will need to consider the ecological dynamics of pathogen introduction and propagation in aquatic environments, as failure to do so will potentially result in consistent exposure risk even after source inputs have been reduced. Further, genetic analysis over spatial and temporal scales of pathogen communities existing in secondary aquatic habitats will provide valuable information regarding their potential input source, and may be coupled with existing monitoring efforts.

1.4. Fecal contaminant monitoring

It is widely held that the major source of waterborne pathogens is fecal loading to aquatic environments. Due to the impracticality of testing for most pathogenic bacteria directly, fecal indicator bacteria (FIB), non-pathogenic bacteria that are putatively derived from fecal material, are currently used to indirectly gauge the potential presence of pathogens. The epidemiological basis correlating FIB presence and the risk potential of related disease incidence has been well documented (Holmes, 1989; Pruss, 1998; Wade et al., 2003; Wiedenmann et al., 2006; USEPA, 1986). Such epidemiological risk assessments have served as the primary basis by which microbiological water quality standards are established. Subsequently, target FIB are monitored in regulated water bodies to identify impairments.

The USEPA currently recommends the use of the FIB *Escherichia coli* and *Enterococci spp.* for the assessment of freshwater WQS compliance (USEPA, 2001). In 2008, Minnesota Rules chapter 7050, "Water Quality" was revised to reflect this recommendation, and *E. coli* are now the WQS target organism used in Minnesota with a target standard of 126 CFU/100 milliliters. However, it has been shown that numbers of *E. coli* and entercocci do not often correlate well with the presence of several pathogenic agents, such as *Salmonella spp.*, *Campylobacter spp.*, protozoa *Giardia spp*. and *Cryptosporidium*, and some human enteroviruses, (Lemarchand and Lebaron, 2003; Bonadonna et al., 2002; Horman et al., 2004; Lund, 1996; Harwood et al., 2005; Geldenhuys and Pretorius, 1989; Pusch et al., 2005). Also, *E. coli* and *Enterococcus spp.* have been shown to propagate and establish communities once excreted from the host

organism in lakes and streams, sand sediments, soil, plants, and algal biomass (Anderson et al., 2005; Byappanahalli et al., 2003, 2006(a), 2006(b); Byappanahalli and Fujioka, 2004; Fujioka et al., 1999; Ishii et al., 2006; Olapade et al., 2006; Power et al., 2005; Solo-Gabriele et al., 2000; Whitman et al., 2003, 2005). This is of particular concern, as both *E. coli* and *Entercoccus spp*. may be present in bodies of water that are not receiving significant fecal loading. As a result, naturalized FIB can skew monitoring results leading to unnecessary regulatory action (e.g. beach closings) and impairment designations. Further, FIB do not identify the source of fecal loading, a significant shortcoming upon determining TMDL load calculations. This has led to federal TMDL recommendations that suggest the sources of fecal loading in a given watershed be determined by "reasonably accurate estimates," or "gross allotments". To rectify the limitations of FIB utilization, techniques that identify the source (MST studies) and scale of fecal loading are a current focal point of much research.

1.5. Sources of fecal pollution to aquatic systems

Bacteria are amongst the most common cause of river and stream impairments (USEPA, 2000). Required TMDL studies will be successful only if the sources of fecal loading contributing to the impairment are identified. Fecal loading of waterwys are typically due to (1) Agricultural contributions; (2) Urban runoff; and (3) Wildlife contributions (Santo Domingo et al., 2002).

From an agricultural perspective, there are several potential input sources, including delivery from adjacent feedlot operations, runoff from fields that have received manure

application, and runoff from manure deposited by livestock grazing operations (Edwards et al., 2000; Heinonen-Tanski et al., 2001; Trevisan et al., 2002). It has been shown that the grazing of cattle significantly increases FIB levels in adjacent water bodies. One study reported that 80% of water samples taken from areas proximate to grazing operations exceeded designated water quality standards (Edwards et al., 1997; Howell et al., 1995).

With the continuing expansion of urban and suburban areas in the U.S., watershed basins are experiencing less infiltration capacity due to increases in impervious surfaces. As urban and sub-urban communities continue to expand, the task of identifying fecal source inputs to local water bodies will become increasingly complicated. This is of concern, as FIB levels in tested stormwater discharges consistently exceed acceptable levels (Novotny et al., 1985). Due to the collective stormwater output of municipal separate storm sewer systems (MS4's), it is becoming increasingly difficult to assign fecal contributions originating from commercial and residential entities. Further, there are temporal phenomena that affect observed FIB levels, such as the time elapsed between rain events (Santo Domingo et al., 2002). Wildlife has also been shown to influence FIB levels in various bodies of water in urban, sub-urban, and rural watersheds (Blankenship, 1996; Simmons et al., 2000). While it is important to identify fecal contributions from wildlife to aquatic systems, is often extremely difficult to remediate these input sources. A general knowledge of the potential fecal sources in a given watershed is the first step in determining TMDL load and wasteload allocations for the impaired water body. While it is important to distinguish between point and non-point source fecal contributions,

efficient TMDL studies will require the specific identification of animal non-point source over a short period of time.

1.6. Microbial Source Tracking (MST)

The presence of fecal input sources in watersheds has led to the development of microbial source tracking (MST) technologies that aim to identify source-specific fecal loading to aquatic environments. MST methods rely on the basic premise that a particular trait unique to a given type of fecal material is directly associated with a specific host, and can be identified using various techniques. Once this trait has been putatively identified, it must undergo a series of tests to confirm its use in monitoring and abatement strategies.

MST techniques can be separated into two categories: (1) molecular genotypic based; and (2) phenotypic, non-molecular, based methods. These two categories are further divided based on the requirement for culturing of the targeted microorganism. Non-molecular methods, primarily antibiotic resistance analysis (ARA), were the first MST methods developed. ARA is a library dependent method, based on the assumption that humans are subjected to antibiotic agents at a greater frequency and of different type than those used for domestic animals, livestock, and wildlife (Kaspar et al., 1990). Consequently, it is inferred that the intestinal flora of humans will possess distinguishable antibiotic resistance characteristics, and these characteristics can be identified and used to track inputs of human fecal bacteria in waterways. Antibiotic resistance studies test fecally or environmentally obtained *E. coli* and *Enterococci spp.*

against a series of antibiotics to determine resistance patterns, that in turn aid in the identification of human versus non-human fecal pollution (Carroll et al., 2005; Edge and Hill, 2005; Hagedorn et al., 1999; Harwood et al., 2000, 2003; Parveen et al., 1997; Wiggins et al., 1999). While the premise of ARA initially seemed logical, antibiotic resistance traits have been shown to be geographically and temporally unstable due to various stressors on the gene, such as degree of antibiotic exposure, and selective pressure (Field and Samadpour, 2007; Samadpour et al., 2005). As a result, it has been shown that antibiotic resistance methods are often ineffective in identifying host species contributing to fecal loading of waterways (Griffith et al., 2003; Moore et al., 2005; Samadpour et al., 2005).

In an attempt to alleviate shortcomings of ARA and other early MST methods, significant advances have been made in the development of molecular based techniques. Several DNA fingerprinting methods, such as ribotyping, repetitive extragenic palindromic polymerase chain reaction (Rep-PCR), and amplified fragment length polymorphism (AFLP), use known-source isolate libraries to identify the host origin of environmental isolates. As with any library dependent method, the size of the known-source library is paramount to the efficacy of DNA fingerprinting methods (Albert et al., 2003; Hassan et al., 2005; Indest et al., 2005; Johnson et al., 2004; Leung et al., 2004; Stoeckel et al., 2004). The average rate of correct classification (ARCC) is a statistical test applied to library dependent MST methods to determine the potential of the library to identify environmental isolates (Field and Samadpour, 2007; Moore et al., 2005; Stoeckel and Harwood, 2007). In general, it has been shown that smaller libraries have higher

ARCC's, but are often less effective at identifying environmental isolates at a watershed scale as are larger libraries (Moore et al., 2005; Wiggins et al., 2003). However, large libraries of fecal bacteria from know animal sources is very labor intensive and expensive to construct, and often are also unable to correctly match environmental isolates to known-source library fecal isolates (Hassan et al., 2005; Stoeckel et al., 2004). Further, known-source libraries are limited by the temporal and geographical variation of target bacterial genotypes, variation in the diets of host animal groups, and the naturalization of target organisms in aquatic environments (Byappanahalli et al., 2003, 2006; Gordon, 2001; Hartel et al., 2002, 2003; Ishii et al., 2006; Olapade et al., 2006; Power et al., 2005; Scott et al., 2003; Whitman et al., 2003, 2005).

More recently, library independent bacterial gene target assays have been, and are currently being, developed to circumvent the use of large known-source libraries. Host-specific bacterial gene markers have been identified by using terminal restriction length fragment polymorphism (T-RLFP) analyses of the 16S rDNA gene, and by subtraction hybridization (Argon et al., 2001; Bernhard and Field, 2000(a); Dick et al., 2005; Hamilton et al. 2006; Harakava and Gabriel, 2003; Hsieh and Pan, 2004; Janssen et al., 2001, Liu et al., 2003; Shanks et al., 2006, 2007). Once marker genes are identified, polymerase chain reaction (PCR) primers are developed and used to amplify the marker directly from environmental samples, bypassing the culture step required in DNA fingerprinting or RFLP methods. As a result, gene targeted assays have the potential to expedite the source tracking process, a key element in future TMDL determinations. Such assays also eliminate culture bias as they are not limited to the use of only easily

culturable organisms.

To date, several PCR primer sets have been developed to determine fecal contributions from various animal hosts to aquatic systems (Table 1.1). Currently, fecal anaerobes are typically used as target organisms for genetic marker development. While Hamilton and coworkers (2006) identified a gene marker with the ability to detect goose-and-duck-specific *E. coli*, fecal anaerobes represent a larger portion of total bacteria in feces than do coliforms and enterococci, and they are unable to propagate in oxygen rich environments once excreted from the host organism (Hamilton et al., 2006; Savage, 2001). Further, it has been shown that bacteria of the genera *Bifidobacterium* and *Bacteroides* have unique distributions depending on their animal host (Allsop and Stickler, 1985; Fiksdal et al., 1985; Kreader, 1995; Resnick and Levin, 1981). Host-specific *Bacteroidales* gene markers have previously been shown to identify feces from cattle, dogs, elk, horses, humans, pigs, and ruminants (Bernhard and Field, 2000; Dick et al., 2005(a); Layton et al., 2006; Okabe et al., 2007; Shanks et al, 2006, 2008).

While promising, work still needs to be done in the development of host-specific bacterial gene target assays. Effective gene markers identified need to be useful over broad geographically areas, as method validation will be required when applied to new locations (USEPA 2005). Also, the correlation between markers developed for organisms other than FIB and risk of pathogen exposure has yet to be established (Field and Samadpour, 2007). Lastly, there has yet to be markers developed for several animal host species that may potentially contribute fecal material to regulated water bodies.

1.7. Summary of Thesis

Due to the magnitude of regulated water bodies that exceed microbiological water quality standards, TMDL determinations will require rapid and effective methods to identify fecal source contributions within a given watershed. The current state of MST methodology lies in the development and validation of animal host-specific gene markers that can be identified directly from environmental samples. The body of research presented in the second and third chapters of this thesis concentrate on the identification and evaluation of animal host-specific gene markers. In Chapter 2, I used suppression subtraction hybridization (SSH) to identify gene markers specific to *E. coli* derived from swine fecal sources. The subtraction yielded ten gene targets that were potentially swine specific. Subsequently, DNA sequence analysis identified 4 unique genetic markers that were subsequently used to develop primer sets for PCR amplification. The ability of this marker set to potentially identify a high percentage of *E. coli* isolated from swine hosts suggests it may be useful in determining their fecal contribution to impaired waterways.

In Chapter 3, I examined the influence of cattle grazing operations on a small creek system in Southeastern Minnesota using a PCR-based MST method. Environmental DNA was extracted bi-monthly over the summer months of 2008 and 2009 at 5 separate sampling locations each year. In addition, membrane filtration was conducted on acquired samples to obtain *E. coli* plate count data per EPA Method 1603 protocol (USEPA, 2002). DNA's were first subjected to conventional PCR analyses using two separate primer sets that targeted: (1) the 16s rDNA gene belonging to all bacteria in the genera *Bacteroides* (AllBac); and (2) the human specific HF 183 *Bacteroides* marker

(Bernhard and Field, 2000(b); Layton et al., 2006). A quantitative PCR (q-PCR) assay was then developed and evaluated to estimate the total presence of bovine-specific *Bacteroides* by use of the Bac3 gene marker (Shanks et al., 2006). Subsequently, the q-PCR data were compared to the *E. coli* plate count data, revealing a lack of consistent correlation between the two methods. Several physical and environmental factors likely influenced the level of *E. coli* found in the stream, confirming the hypothesis that information will be needed to supplement current FIB monitoring efforts in order to determine accurate source-specific fecal impacts. In Chapter 4, I examined whether secondary habitats for fecal bacteria also harbor human pathogens. Specifically, I examined spatial and temporal variation in the population structure of *Salmonella* isolated from *Cladophora* found in stream and lake water, aquatic plants, beach sand, and sediments by use of horizontal fluorophore-enhanced rep-PCR (HFERP). Analysis of HFERP fingerprints obtained from 2005-2007 at various Lake Michigan locations revealed that *Salmonella* populations associated with *Cladophora* varied both spatially and temporally, suggesting potentially different input sources of *Salmonella* over space and time. In addition, differing environmental stressors may play a role in selecting particular *Salmonella* genotypes that are best suited for growth in this environment, indicating the potential for genetic drift of the introduced *Salmonella.* Chapter 5 is a summary of my conclusions and future directions regarding the research presented in this document.

1.8. Tables and Figures

Table 1.1 Genetic markers identifying source-specific bacteria derived from animal fecal material.

Primer	Fecal Source	Sequence	PCR Type	Target Organism	Source
HF183F Bac708R	Human	5'-ATCATGAGTTCACATGTCCG-3' 5'-CAATCGGAGTTCTTCGTG-3'	Conventional	*Bacteroides-Prevotella*	Bernhard and Field, 2000
HuBac566F HuBac629R	Human	5'-GGGTTTAAAGGGAGCGTAGG-3' 5'-CTACACCACGAATTCCGCCT-3'	Conventional	*Bacteroides*	Layton et al., 2006
Bac3F Bac3R	Cattle	5'-CTAATGGAAAATGGATGGTATCT-3' 5'-GCCGCCCAGCTCAAATAG-3'	Conventional	*Bacteroidales*	Shanks et al., 2006
CowM3F CowM3R	Cattle	5'-CCTCTAATGGAAAATGGATGGTATCT-3' 5'-CCATACTTCGCCTGCTAATACCTT-3'	qPCR	*Bacteroides*	Shanks et al., 2008
PF163F Bac708R	Swine	5'-GCGGATTAATACCGTATGA-3' 5'-CAATCGGAGTTCTTCGTG-3'	Conventional	*Bacteroidales*	Dick et al., 2005
HoF597F Bac708R	Horse	5'-CCAGCCGTAAAATAGTCGG-3' 5'-CAATCGGAGTTCTTCGTG-3'	Conventional	*Bacteroidales*	Dick et al., 2005

1.9. References

Agron, P. G.; Walker, R. L.; Kinde, H.; Sawyer, S. J.; Hayes, D. C.; Wollard, J.; Andersen, G. L. 2001. Identification by subtractive hybridization of sequences specific for *Salmonella enterica* serovar Enteritidis. Appl. Environ. Microbiol. 67:4984–4991.

Albert, J.M.; Munakata-Marr, J.; Tenorio, L.; Siegrist, R.L. 2003. Statistical evaluation of bacterial source tracking data obtained by rep-PCR DNA fingerprinting of *Escherichia coli*. Environ. Sci. Technol. 37 (20), 4554–4560.

Alexander, L. M.; Heaven, A.; Tennant, A.; Morris, R. 1992. Symptomatology of children in contact with sea water contaminated with sewage. J. Epidemiol. Community Health 46:340–344.

Allsop, K.; Stickler, J.D. 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. J. Appl. Bacteriol. 58, 95–99.

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

Bernhard, A. E., and Field, K. G. 2000(a). Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66:1587–1594.

Bernard, A.E., and Field, K.G. 2000(b). 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, v. 66, no. 10, p. 4571-4574.

Blankenship, K. 1996. DNA library would give investigators inside poop on pollution sources. Bay Journ. 6(6).

Bolton, F. J.; Surman, S. B.; Martin, K.; Wareing, D. R.; Humphrey, T. J. 1999. Presence of *Campylobacter* and *Salmonella* in sand from bathing beaches. Epidemiol. Infect. 122:7–13.

Byappanahalli, M.; Fowler, M.; Shively, D.; Whitman, R. 2003. Ubiquity and persistence of *Escherichia coli* in a Midwestern coastal stream. Appl. Environ. Microbiol. 69 (8), 4549–4555.

Byappanahalli, M., and Fujioka, R. 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water. Sci. Technol. 50 (1), 27–32.

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

Byappanahalli, M.; Whitman, R.L.; Shively, D.A.; Ting, W.T.; Tseng, C.C.; Nevers, M.B. 2006(b). Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. J. Water Health 4 (3), 313–320.

Cabelli, V.J.; Dufour, A.P.; Levin, M.A.; McCabe, L.J.; Haberman, P.W. 1979. Relationship of Microbial Indicators to Health Effects at Marine Bathing Beaches. Am. J. Public Health. 69: 690-696.

Carroll, S., Hargreaves, M., Goonetilleke, A. 2005. Sourcing fecal pollution from onsite wastewater treatment systems in surface waters using antibiotic resistance analysis. J. Appl. Microbiol. 99, 471–482.

Centers for Disease Control and Prevention. 1998. Outbreak of *Vibrio parahaemolyticus* infections associated with eating raw oysters, Pacific Northwest, 1997. Morb. Mortal. Wkly. Rep. 147:45762.

Centers for Disease Control and Prevention. 2008. Surveillance Summaries. Surveillance for Waterborne Disease and Outbreaks Associated with Recreational Water Use and Other Aquatic Facility-Associated Health Events — United States, 2005–2006 and Surveillance for Waterborne Disease and Outbreaks Associated with Drinking Water and Water not Intended for Drinking — United States, 2005–2006 . MMWR 2008;57(No. SS-9).

Cheung,W.H.S. (Ed.), 1988. Health Effects of BeachWater Pollution in Hong Kong. Vincent Blue Copy Co., Hong Kong.

Colles, F. M.; Jones, K.; Harding, R. M.; Maiden, M. C. J. 2003. Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. Appl. Environ. Microbiol. 69:7409–7413.

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

Dick, L.K.; Simonich, M.T.; Field, K.G. 2005(b). Microplate subtractive hybridization to enrich for source-specific Bacteroidales fecal pollution indicators. Appl. Environ. Microbiol. 71, 3179–3183.

Dufour, A.P. 1977 *Escherichia coli*: The faecal coliform. In Bacterial Indicators/Health Hazards Associated with Water ed. Hoadley, A.W. & Dutka, B.J. ASTM STP 635, pp. 48-58. Philadelphia: American Society for Testing and Materials.

Edge, T.A.; Hill, S. 2005. Occurrence of antibiotic resistance in Escherichia coli from surface waters and fecal pollution sources near Hamilton, Ontario. Can. J. Microbiol. 51 (6), 501–505.

Edwards, D. R.; Coyne, M. S.; Vendrell, P. F.; Daniel, T. C.; Moore, P. A., Jr.; Murdoch, J. F. 1997. *Escherichia coli* survival in mantled karst springs and streams, Northwest Arkansas Ozarks, USA. JAWRA, *33*, 413-422.

Edwards, D.R.; Larson, B.T.; Lim, T.T. 2000. Runoff nutrient and fecal coliform content from cattle manure application to fescue plots. J. Am. Water Resour. Assoc. 36:711–721

Field, Katharine G., Samadpour, Mansour. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research. 41. 3517–3538.

Fiksdal, L.; Make, J.S.; LaCroix, S.J.; Staley, J.T. 1985. Survival and detection of *Bacteroides spp*., prospective indicator bacteria. Appl. Environ. Microbiol. 49 (1), 148–150.

Fujioka, R.; Dian-Denton, C.; Castro, J.; Morphew, K. 1999. Soil: the environmental source of *Escherichia coli* and *Enterococci* in Guam's streams. J. Appl. Microbiol. 85, 83S–89S.

Gordon, D. M. 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology 147:1079–1085.

Griffith, J.F.; Weisburg, S.B.; McGee, C.D. 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. J. Water Health 1, 141–152.

Hagedorn, C.; Robinson, S.L.; Filtz, J.R.; Grubbs, S.M.; Angier, T.A.; Reneau Jr., R.B. 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. 65 (12), 5522–5531.

Hamilton, M.J.; Yan, T.; Sadowsky, M.J.; 2006. Development of goose- and duck-specific DNA markers to determine sources of *Escherichia coli* in waterways. Appl. Environ. Microbiol. 72 (6),4012–4019.

Harakava, R.; and Gabriel, D. W. 2003. Genetic differences between two strains of *Xylella fastidiosa* revealed by suppression subtractive hybridization. Appl. Environ. Microbiol. 69:1315–1319.

Hartel, P. G.; Summer, J. D.; Hill, J. L.; Collins, J.; Entry, J. A.; Segars, W.I. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. J. Environ. Qual. 31:1273–1278.

Hartel, P. G.; Summer, J. D.; Segars, W.I. 2003. Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking. Water Res. 37:3263–3268.

Harwood, V.J.; Whitlock, J.; Withington, V. 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. 66 (9), 3698–3704.

Harwood, V.J.; Wiggins, B.A.; Hagedorn, C.; Ellender, R.D.; Gooch, J.; Kern, J.; Samadpour, M.; Chapman, A.C.H.; Robinson, B.J. 2003. Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Water Health 1, 153–166.

Hassan, W.M.; Wang, S.Y.; Ellender, R.D. 2005. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Appl. Environ. Microbiol. 71 (1), 512–518.

Heinonen-Tanski, H.; Uusi-Kamppa, J. 2001. Runoff of fecal microorganisms and nutrients from perennial grass ley after application of slurry and mineral fertilizer. Water Sci. Technol. 43:143–146

Hsieh, W. J., and M. J. Pan. 2004. Identification *Leptospira santarosai* serovar shermani specific sequences by suppression subtractive hybridization. FEMS Microbiol. Lett. 235:117–124.

Holmes, P.R. 1989. Research into health risks at bathing beaches in Hong Kong. J. Inst. Water Environ. Manage. 3,488–495.

Howell, J. M.; Coyne, M. S.; Cornelius, P. 1995. Fecal bacteria in agricultural waters of the bluegrass region of Kentucky. J. Environ. Qual, 24, 411-419.

Indest, K.J.; Betts, K.; Furey, J.S. 2005. Application of oligonucleotide microarrays for bacterial source tracking of environmental *Enterococcus sp.* isolates. Int. J. Environ. Res. Public Health 2 (1), 175–185.

Ishii, S.; Yan, T.; Shively, D.A.; Byappanahalli, M.N.; Whitman, R.L.; Sadowsky, M.J. 2006(a). *Cladophora* (Chlorophyta) spp. harbor human bacterial pathogens in nearshore water of Lake Michigan. Appl. Environ. Microbiol. 72 (7), 4545–4553.

Ishii, S.; Ksoll, W.B.; Hicks, R.E.; Sadowsky, M.J. 2006(b). Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), 612–621.

Janssen, P. J; Audit, B.; Ouzounis, C. A. 2001. Strain-specific genes of *Helicobacter pylori*: distribution, function and dynamics. Nucleic Acids Res. 29:4395–4404.

Jenkins, M. B.; Hartel, P. G.; Olexa, T. J.; Stuedemann, J.A. 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. J. Environ. Qual. 32:305–309.

Johnson, L.K.; Brown, M.B.; Carruthers, E.A.; Ferguson, J.A.; Dombek, P.E.; Sadowsky, M.J. 2004. Sample size, library composition, and genotypic diversity among natural populations of Escherichia coli from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70 (8), 4478–4485.

Kaspar, C.; Burgess, J.; Knight, I.; Colwell, R. R.1990. Can. J. Microbiol, 36, 891-894.

Kreader, C.A. 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. 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. Appl. Environ. Microbiol. 72(6):4214-4224.

Lemarchand, K., and Lebaron, P. 2003. Occurrence of *Salmonella spp*. and *Cryptosporidium spp.* in a French coastal watershed: relationship with fecal indicators. FEMS Microbiol. Lett. 218,203–209.

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

Liu, L.; Spilker, T.; Coenye, T.; LiPuma, J. J. 2003. Identification by subtractive hybridization of a novel insertion element specific for two widespread *Burkholderia cepacia* genomovar III strains. J. Clin. Microbiol. 41:2471–2476.

Minnesota Pollution Control Agency. 2009. Bacteria TMDL protocols and submittal requirements. wq-iw1-08 revised 03-09.

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

Novotny, V.; Sung, H.-M.; Bannerman, R.; Baum, K. 1985. Estimating nonpoint pollution from small urban watersheds. Journal WPCF 57, 339-348.

Olapade, O.A.; Depas, M.M.; Jensen, E.T.; McLellan, S.L. 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Appl. Environ. Microbiol. 72 (3), 1932–1938.

Parveen, S.; Murphree, R.L.; Edminston, L.; Kaspar, C.W.; Portier, K.M.; Tamplin, M.L. 1997. Association of multiple-antibioticresistance profiles with point and nonpoint sources of *Escherichia coli* in Apalachicola Bay. Appl. Environ. Microbiol. 63, 2607–2612.

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

Pruss, A. 1998. Review of epidemiological studies on health effects from exposure to recreational water. Int. J. Epidemiol.27 (1), 1–9.

Resnick, I.G.; Levin, M.A. 1981. Assessment of *Bifidobacteria* as indicators of human fecal pollution. Appl. Environ. Microbiol.42 (3), 433–438.

Samadpour, M.; Roberts, M.C.; Kitts, C.L.; Mulugeta, W.; Alfi, D. 2005. The use of ribotyping and antibiotic resistance patterns for identification of host sources of *Escherichia coli* strains. Lett. Appl. Microbiol. 40, 63–68.

Schallenberg, M.; Bremer, P. J.; Henkel, S.; Launhardt, A.; Burns, C. W. 2005. Survival of *Campylobacter jejuni* in water: effect of grazing by the freshwater crustacean *Daphnia carinata* (Cladocera). Appl. Environ. Microbiol. 71:5085–5088.

Scott, T. M.; Parveen, S.; Portier, K. M.; Rose, J. B.; Tamplin, M. L.; Farrah, S. R.; Koo, A.; Lukasik, J. 2003. Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. Appl. Environ. Microbiol. 69:1089–1092.

Shanks, O.C.; Santo Domingo, J.W.; Lamendella, R.; Kelty, C.A.; Graham, J.E. 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72 (6), 4054–4060.

Shanks, O.C.; Domingo, J.W.; Lu, J.; Kelty, C.A.; Graham, J.E. 2007. Identification of bacterial DNA markers for the detection of human fecal pollution in water. Appl. Environ. Microbiol. 73 (8), 2416–2422.

Shanks, O. C.; Atikovic, Emina; Blackwood, A.; Denene, Lu; Jingrang, Noble; Rachel T.; Domingo, J.S.; Seifring, Shawn; Sivaganesan, Mano; Haugland, Richard A. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 74: 745-752

Sharma, S.; Sachdeva, P.; Virdi, J. S. 2003. Emerging water-borne pathogens. Appl. Microbiol. Biotechnol. 61:424–428.

Simmons, G.; Waye, D.; Herbein, S.; Myers, S.; Walker, E. 2000. Presented at Virginia Water Research Symposium. Blacksburg, VA, Younos, T., Poff, J., Eds.; VWRRC Special Report SR-19-2000.

Solo-Gabriele, H.M.; Wolfert, M.A.; Desmarais, T.R.; Palmer, C.J. 2000. Sources of *Escherichia coli* in a coastal subtropical environment. Appl. Environ. Microbiol. 66 (1), 230–237.

Souza, V.; Rocha, M.; Valera, A.; Eguiarte, L.E. 1999. Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. Appl. Environ. Microbiol. 65 (8), 3373–3385.

Stoeckel, D.M.; Mathes, M.V.; Hyer, K.E.; Hagedorn, C.; Kator, H.; Lukasik, J.; O'Brien, T.L.; Fenger, T.W.; Samadpour, M.; Strickler, K.M.; Wiggins, B.A. 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. 38 (22), 6109–6117.

Stoeckel, D.M., and Harwood, V.J. 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. 73 (8), 2405–2415.

Szewzyk, U.; Szewzyk, R.; Manz, W.; Schleifer, K. H. 2000. Microbiological safety of drinking water. Annu. Rev. Microbiol. 54:81–127.

Trevisan, D.; Vansteelant, J. Y.; Dorioz, J. M. 2002. Survival and leaching of fecal bacteria after slurry spreading on mountain hay meadows: consequences for the management of water contamination risk. Water Res., 36, 275-283.

U.S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA-A440/5-84-002. U.S.Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2000. The quality of our Nation's waters; EPA-841-S-00-001. U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2001. Protocol for Developing Pathogen TMDLs. EPA 841-J-00-002. 132 pp. Office of Water (4503F). U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2002. Method 1603—*Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar. EPA821-R-02-23, 9. U. S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2005. Microbial Source Tracking Guide Document. EPA-600/R-05/064, 131pp. Office of Research and Development, Washington, DC.

U.S. Fish and Wildlife Service. 2009 (copyright). The Federal Water Pollution Control Act (Clean Water Act). www.fws.gov/laws/lawsdigest/FWATRPO.html.

Wade, T.J.; Pai, N.; Eisenberg, J.N.S.; Colford, J.M.J. 2003. Do US EPA water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environ. Health Perspect. 111, 1102–1109.

Whitman, R. L.; Horvath, T. G.; Goodrich, M. L.; Nevers, M. B.; Wolcott, M. J.; Haack, S. K. 2001. Characterization of E. coli levels at 63rd Street Beach. Report to the City of Chicago Department of the Environment and the Chicago Park District. [Online.] U.S. Geological Survey, Porter, IN.
http://www.glsc.usgs.gov/_files/reports/63rdStreetBeachReport.pdf.

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. Appl. Environ. Microbiol. 69, 4714–4719.

Whitman, R.L.; Byers, S.E.; Shively, D.A.; Ferguson, D.M.; Byappanahalli, M. 2005. Occurrence and growth characteristics of *Escherichia coli* and enterococci within the accumulated fluid of the northern pitcher plant (Sarracenia purpurea L.). Can. J. Microbiol. 51 (12), 1027–1037.

Wiedenmann, A.; Kruger, P.; Dietz, K.; Lopez-Pila, J.M.; Szewzyk, R.; Botzenhart, K. 2006. A randomized controlled trial assessing infectious disease risks from bathing in fresh recreational waters in relation to the concentration of *Escherichia coli*, intestinal enterococci, *Clostridium perfringens*, and somatic coliphages. Environ. Health Perspect. 114 (2), 228–236.

Wiggins, B.A.; Andrews, R.W.; Conway, R.A.; Corr, C.L.; Dobratz, E.J.; Dougherty, D.P.; Eppard, J.R.; Knupp, S.R.; Limjoco, M.C.; Mettenburg, J.M.; Rinehardt, J.M.; Sonsino, J.; Torrijos, R.L.; Zimmerman, M.E. 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. 65, 3483–3486.

Wiggins, B.A.; Cash, P.W.; Creamer, W.S.; Dart, S.E.; Garcia, P.P.; Gerecke, T.M.; Han, J.; Henry, B.L.; Hoover, K.B.; Johnson, E.L.; Jones, K.C.; McCarthy, J.G.; McDonough, J.A.; Mercer, S.A.; Noto, M.J.; Park, H.; Phillips, M.S.; Purner, S.M.; Smith, B.M.; Stevens, E.N.; Varner, A.K. 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. 69, 3399–3405.

Winfield, M. D., and Groisman, E. A. 2003. Role of non-host environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. 69:3687–3694.

Chapter 2. Development of swine-specific DNA markers to determine sources of *Escherichia coli* in waterways.

2.1. Overview

Fecal contamination of waterways represents a consistent threat to public health. Swine and turkey production is abundant in Minnesota's agricultural regions, often leading to manure run-off and contamination of waterways. In this study I developed DNA markers to detect *E. coli* originating from swine feces. Pooled genomic DNAs representing 21 unique *E. coli* strains isolated from swine and 40 non-swine strains (5 strains each from chickens, cows, goats, geese, horses, humans, sheep, and turkeys) were chosen as tester and driver DNAs, respectively, for suppression subtractive hybridization. Tester and driver DNA's were selected from dendograms generated using horizontal, fluorophore-enhanced rep-PCR (HFERP) DNA fingerprinting. Subtraction products were ligated into a vector, 192 clones were picked for the subtraction library, and insert DNAs were screened by dot-blot hybridization to the tester and driver DNAs. Twenty three probes were found to hybridize specifically to the swine DNAs. Southern hybridization analyses indicated that 14 of the 23 clones were potentially swine-specific. Colony hybridizations to 1,316 *E. coli* strains in a library containing16 animal species and humans indicated that 4 hybridization probes had the ability to detect 62.3% of all tested swine *E. coli* strains. Interestingly, the probes cross-reacted with 39.7% of the tested turkey *E. coli*. DNA sequence analysis of clonal insert DNA's identified 4 unique genetic markers that may subsequently used to develop primer sets for PCR amplification. The ability of this marker set to potentially identify a high percentage of *E. coli* isolated from

swine and turkey hosts suggests it may be useful in determining their fecal contribution to impaired waterways.

2.2. Introduction

Fecal loading to aquatic environments by various animal hosts is of concern for public and environmental health. Such loading is the main vehicle for the introduction of human pathogen into waterways. Microbiological contamination of waterways is amongst the most commonly listed water quality impairment in the U. S. (USEPA, 2005). The federal Clean Water Act (CWA) requires individual states to provide the United States Environmental Protection Agency (USEPA) with an aggregate quality assessment of its waterways every biennium (section 305(b) list) in effort to identify waters that do not meet state and federal quality standards (section 303(d) list). Once deemed impaired, States must conduct a total maximum daily load (TMDL) assessment in order to mitigate the impairment and restore the water body to acceptable quality. Successful fecal bacteria TMDL implementation strategies require microbial source tracking (MST) technologies that accurately and efficiently characterize the host-specific source of bacterial loading.

Non-molecular methods, primarily antibiotic resistance analysis (ARA), were the first MST methods developed. ARA is a library dependent method, based on the assumption that humans are subjected to antibiotic agents at a greater frequency and of different type than those used for domestic animals, livestock, and wildlife (Kaspar et al. 1990). Consequently, it is inferred that the intestinal flora of humans will possess distinguishable antibiotic resistance characteristics, and these characteristics can be

identified and used to track inputs of human fecal bacteria in waterways. While the premise of ARA initially seemed logical, antibiotic resistance traits have been shown to be geographically and temporally unstable due to various stressors on the gene, such as degree of antibiotic exposure, and selective pressure (Field and Samadpour, 2007; Samadpour et al., 2005). As a result, it has been shown that antibiotic resistance methods are often ineffective in identifying host species contributing to fecal loading of waterways (Griffith et al., 2003; Moore et al.,2005; Samadpour et al., 2005).

In an attempt to alleviate shortcomings of ARA and other early MST methods, significant advances have been made in the development of molecular based techniques. Several DNA fingerprinting methods, such as ribotyping, repetitive extragenic palindromic polymerase chain reaction (Rep-PCR), and amplified fragment length polymorphism (AFLP), use known-source isolate libraries to identify the host origin of environmental isolates. As with any library dependent method, the size of the known-source library is paramount to the efficacy of DNA fingerprinting methods (Albert et al., 2003; Hassan et al., 2005; Indest et al., 2005; Johnson et al., 2004; Leung et al., 2004; Stoeckel et al., 2004). The average rate of correct classification (ARCC) is a statistical test applied to library dependent MST methods to determine the potential of the library to identify environmental isolates (Field and Samadpour, 2007; Moore et al., 2005; Stoeckel and Harwood, 2007). In general, it has been shown that smaller libraries have higher ARCC's, but are often less effective at identifying environmental isolates at a watershed scale as are larger libraries (Moore et al., 2005; Wiggins et al., 2003; Yan et al., 2007). However, large libraries of fecal bacteria from know animal sources are very labor

intensive and expensive to construct, and often are also unable to correctly match environmental isolates to known-source library fecal isolates (Hassan et al., 2005; Stoeckel et al., 2004). Furthermore, known-source libraries are limited by the temporal and geographical variation of target bacterial genotypes, variation in the diets of host animal groups, and the naturalization of target organisms in aquatic environments (Byappanahalli et al., 2003, 2006; Gordon, 2001; Hartel et al., 2002, 2003; Ishii et al., 2006; Olapade et al., 2006; Power et al., 2005; Scott et al., 2003; Whitman et al., 2003, 2005).

To rectify the shortcomings of library-based methodologies, most researchers now focus in the use of culture- and library-independent methods. To date, several genetic markers have been developed that are reported to identify several host-specific fecal inputs into aquatic environments. Bernhard and Field (2000(a), 2000(b)) investigated the use of the 16S rRNA gene from *Bifidobacterium* and members of the *Bacteroides-Prevotella* group to discriminate between human and ruminant fecal source inputs to waterways. Dick and coworkers (2005(a) and 2005(b)) subsequently used suppressive subtraction hybridization (SSH) to identify host specific *Bacteroides* and *Bacteroidales* 16S rRNA gene sequences that identify fecal loading originating from swine and horses. The SSH technique has been shown to effectively identify a variety of host-specific genetic marker genes, many of which are used in MST studies (Argon et al., 2001; Bernhard and Field, 2000(a); Dick et al., 2005(b); Hamilton et al. 2006; Harakava and Gabriel, 2003; Hsieh and Pan, 2004; Janssen et al., 2001, Liu et al., 2003; Shanks et al., 2006, 2007). Similarly, Shanks et al. (2006) used a gene fragment enrichment (GFE)

technique to target cow-specific *Bacteroides* markers. While these methods are promising for future use in bacterial source tracking studies, they do not identify fecal indicator bacteria (FIB), such as *E. coli* and enterococci, which are currently used by governmental entities to determine potential impairments. As a result, Hamilton and coworkers (2006) used SSH of known-source library isolates to identify duck and goose-specific marker genes originating from *E. coli*.

In the study reported here, I developed DNA-based marker genes to detect *E. coli* originating from swine using similar methodology as that described by Hamilton et al. (2006). Identification of marker genes and subsequent analyses was facilitated by using, a robot-based colony arraying and screening process as described by Yan et al., 2007. Identification of swine-specific fecal contributions to waterways will be essential in the establishment of TMDL's in agricultural regions, and the subsequent implementation of agricultural best management practices (BMP's) to restore water quality.

2.3. Materials and Methods

2.3.1 E. coli strains.

The unique *E. coli* strains used in SSH and specificity analyses were selected from a previously constructed library containing unique isolates obtained from feces of cats, chickens, cows, deer, dogs, ducks, Canada geese, goats, horses, pigs, sheep, and turkeys and humans (Dombek et al., 2000; Johnson et al., 2004). Isolates were obtained from 1998 to 2005 in Minnesota and Wisconsin. Unique strains were defined as isolates originating from a single host animal that had DNA similarity coefficients $\geq 92\%$.

Horizontal fluorophore enhanced repetitive extragenic palindromic PCR (HFERP) DNA fingerprints of *E. coli* strains obtained from swine were analyzed for genetic relatedness using BoxA1R primers. Fingerprints were analyzed using Pearson's product-moment correlation coefficient, with 1% optimization, and dendrograms were generated using the unweighted pair group method with arithmetic means (Hamilton et al., 2006) Based on these analyses, 21 *E. coli* strains from pigs (P002, P008, P010, P071, P073, P084, P115, P152, P188, P193, P219, P220, P236, P246, P250, P252, P266, P267, P273, P309, and P318) that showed maximum difference in genetic relatedness were selected as tester strains for SSH and subsequent probe development. Forty non-pig strains from the same library (5 strains each from chickens, cows, goats, geese, horses, humans, sheep, and turkeys) were chosen as driver strains.

2.3.2 Suppressive Subtraction Hybridization.

SSH was done using the CLONTECH PCR-Select bacterial genome subtraction kit (BD Biosciences CLONTECH, Mountain View, CA) according to the manufacturer's instructions. Genomic DNA was extracted from the 21 pig *E. coli* strains and 49 non-pig strains using a cesium chloride density gradient centrifugation method as previously described (Sadowsky et al., 1987; Hamilton et al., 2006). Aliquots (3 μg) of genomic DNA from the 21 pig and 49 non-pig *E. coli* strains were separately pooled and used as tester and driver DNAs, respectively. Aliquots of each pooled sample (2 μg) were digested with *RsaI*, as previously described (Hamilton et al., 2006). To further enrich for tester-specific fragments, the subtraction process was repeated using PCR-amplified

secondary subtraction products as tester DNAs (Hamilton et al., 2006). A library of potential DNA inserts that were specific for pig *E. coli* was created by cloning the final subtraction products into the pGEM-T vector using a T/A cloning procedure (Promega, Madison, WI). A total of 192 clones were selected for further analysis and stored in HMFM medium at -80°C until used.

*2.3.3 Identification of DNA sequences specific for E. coli from swine***.**

The library of potentially pig-specific cloned DNA fragments was screened for hybridization specificity as described by Schleicher & Schuell, Keene, NH (http://www.schleicher-schuell.com/bioscience) and as previously described (Hamilton et al., 2006). Cloned inserts were amplified by PCR using nested primers 1 (5'-TCGAGCGGCCGCCC GGGCAGGT-3') and 2R (5'- AGCGTGGTCGCGGCCGAGGT-3') provided in the CLONTECH SSH kit. PCR of the cloned inserts was carried out as previously described (Hamilton et al., 2006). PCR products (50 ng) were spotted onto duplicate Nytran SuPerCharge nylon membranes (Schleicher & Schuell, Keene, NH) using a dot blot vacuum manifold (Gibco-BRL, Gaithersburg, MD) and the Minifold spotting protocol (Schleicher & Schuell, Keene, N.H.) established by Hamilton et al. (2006). Membranes were baked and prehybridized as previously described (Sambrook, 1989, Hamilton et al., 2006). Aliquots (50 ng) of the *RsaI* digested pooled genomic DNA from the twenty-one pig *E. coli* strains were labeled with ^{32}P dCTP using a random primer labeling kit (Invitrogen, Carlsbad, Calif.) and following the manufacturer's protocol (Hamilton et al., 2006). Probes were hybridized to prepared membranes for 18 h

at 46°C and washed as previously described (Sambrook et al., 1989, Hamilton et al., 2006). Images were captured using a STORM 840 densitometer (GE Biosciences, Piscataway, NJ). Putative pig-specific DNA inserts were identified by hybridization intensity.

Plasmids for Southern Blot analysis were isolated from putative pig-specific clones using a QIAprep Spin miniprep kit (QIAGEN, Valencia, CA) according to the manufacturer's protocol. Insert DNA was amplified by PCR as previously described (Hamilton et al., 2006). DNAs were transferred to Nytran SuPerCharge nylon membranes (Sambrook et al., 1989), and the membranes were probed with the *RsaI*-digested, pooled, genomic DNAs from the 21 selected pig *E. coli* strains, and 49 non-pig strains. Confirmed pig-specific inserts were identified as discussed above.

2.3.4 Colony hybridization for probe evaluation.

The specificity of subtracted DNA inserts was evaluated by using a colony hybridization method as previously described (Hamilton et al., 2006; Yan et al., 2007). Briefly, *E. coli* isolates from 48 cats, 144 chickens, 189 cows, 96 deer, 106 dogs, 81 ducks, 135 geese, 42 goats, 78 horses, 210 humans, 215 swine, 61 sheep, and 126 turkeys were spot inoculated onto positively charged Performa II nylon membranes (22x22cm, Genetix, Boston, MA) using a Genetix QBot robot (Boston, MA) (Johnson et al., 2004; Yan et al., 2007). Membranes were prepared for arraying as described by Yan et al. (2007), except that each membrane subunit consisted of four spots of each isolate instead of nine, and membranes were placed on the surface of LB agar plates and incubated at

37°C for ~8 hr. Colonies on membranes were lysed and prehybridized as previously described (Sambrook et al., 1989; Hamilton et al., 2006).

Putative pig-specific DNA probes were labeled with ^{32}P dCTP, using the Random Primer DNA labeling system (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol, and membranes were hybridized as previously described (Hamilton et al., 2006; Yan, 2007).

2.4. Results

2.4.1 Isolation of DNA fragments specific for swine E. coli.

Upon completion of the SSH protocol, 192 subtraction clones were selected to construct a DNA subtraction library. Of the 192 library clones analyzed, 23 proved to hybridize specifically to DNA from swine *E. coli* by dot-blot analysis. A representative group of 96 clones analyzed by dot-blot analysis are shown in Figure 2.1. Hybridization characteristics of the 23 clones were further assessed by Southern blot analysis using the probes that were used in the dot blot procedure. Southern hybridization analyses indicated that 14 of the 23 clones were specific to swine *E. coli.*

2.4.2 Assessment of insert specificity.

While the Southern hybridization analysis confirmed that 14 DNA inserts from the SSH library were specific to *E. coli* originating from swine, this assessment was limited to the *E.* coli strains used in the original SSH procedure. The hybridization specificity of the 14 DNA inserts were further assessed by colony hybridization

experiments to DNAs from *E. coli* isolates originating from cats (48 isolates 48 cats, 144 chicken, 189 cow, 96 deer, 106 dog, 81 duck, 135 goose, 42 goat, 78 horse), humans (210), swine (215), sheep (61) and turkeys (126). These know-source *E. coli* isolates (Dombek et al., 2000; Johnson et al., 2004) were spot inoculated onto 14 - 22-by-22 cm membranes and individually probed with the 14 ^{32}P- labeled PCR-amplified insert DNA from the confirmed "swine-specific" clones. A representative image of a colony hybridization membrane for probe 2E11 is shown in Figure 2.2. Colony hybridization membranes were analyzed visually to determine positive hybridization at individual colony locations. Based on theses analyses, 10 of the 14 (71%) hybridization probes were found to detect swine *E. coli* isolates from our library.

Sequence analysis revealed that 6 of the 10 probes had identical nucleotide sequences and that four unique DNA fragments (1A10, 1H5, 2E11, and 2G2) showed hybridization specificity to the swine *E. coli* (Table 2.1). The probes identified 10.7 to 39.1% of 215 unique *E. coli* swine strains tested. Probe 1H5 hybridized to the greatest number of swine derived *E. coli* isolates (39.1%), and together the four marker probes hybridized with 62.3% of swine *E. coli* isolates. Interestingly, three of the four marker probes (1H5, 2E11, and 2G2) cross-hybridized 18.3 to 26.2% of turkey *E. coli* strains included in the colony hybridization analyses (Table 2.1), and together, the three marker probes hybridized with 39.7% of *E. coli* isolated from turkeys. In contrast, probe 1A10 hybridized to only 0.01% of the tested *E. coli* strains from turkeys.

2.4.3 Determination of host specificity.

Since the DNA marker genes specific to swine *E. coli* will ultimately be used to identify swine-derived fecal loading to aquatic environments, it is important to characterize the cross reactivity of marker DNA's with a large number of *E. coli* isolates from different animal hosts that may be present in affected watersheds. To do this, I hybridized each of the four DNA marker probes to 1,316 unique *E. coli* strains isolated from cats, chicken, cows, deer, dogs, ducks, goats, geese, humans, sheep, and turkeys. Results of this analysis (Figure 2.3) showed that the DNA marker probes hybridized to 62.3% of swine isolates tested (*n=215*). Interestingly, the probes cross-hybridized with 39.7% of turkey isolates tested. Further, moderate cross-hybridization percentages were observed with chickens, deer, and sheep at 18.75, 19.79, and 19.67%, respectively. The mean frequency at which the marker DNA probes cross-hybridized with *E. coli* isolates from animal hosts other than swine and turkeys was about 10.7%.

2.4.4 Marker DNA sequencing and BLAST analysis.

The four confirmed swine and turkey specific DNA marker probes were sequenced in both directions, and were subsequently subjected to BLASTX analyses using the *E. coli* protein databases at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The size of insert DNA marker probes ranged from 168 to 532 bp. Results of BLASTX homology searches are shown in Table 2.2. The translated 1A10 DNA marker probe sequence showed significant amino acid homology (96%) to a colicin K immunity protein found in *E. coli* (GenBank accession AAX40978). The translated 1H5 DNA

marker gene exhibited 89% amino acid identity to a putative tail protein derived from a bacteriophage in *E. coli* strain UMN026 (accession YP_002411736), the 2E11marker was found to be 100% identical, at the amino acid level, to transposase IS10 (accession YP_002405951), an IS4 family protein found in the same *E. coli* strain (UMN026), and the translated 2G2 marker gene exhibited 97% amino acid identity to transposase ORFB IS629 in *E. coli* (accession ACQ42080).

2.5. Discussion

In this study, I successfully used SSH to identify four DNA marker gene probes that showed high hybridization specificity to *E. coli* isolates originating from swine fecal material. Together, the probes identified 62.3% of swine *E. coli* isolates tested by colony hybridization (*n=215*). On average, the probe set cross-hybridized with 10.7% of *E. coli* isolates originating from other animal hosts and turkeys. In 2006, Hamilton and coworkers adapted the standard SSH protocol by developing a pooled genomic DNA strategy to use as tester and driver DNA's. As opposed to DNA extracted from a single strain, the pooled genomic DNA protocol aimed to increase subtraction products that were goose specific by "pooling" genomic DNA extracts to use as driver and tester DNAs (Hamilton et al., 2006). As reported here, a pool of genomic DNAs from 21 unique *E. coli* isolates derived from swine was used as driver DNA, and pooled genomic DNA from 40 unique *E. coli* isolates originating from 8 non-swine animal hosts (5 each of: chickens, cows, goats, geese, horses, humans, sheep, and turkeys) was used as tester DNA. By increasing the amount of unique strains isolated from animal hosts other than

swine, it was my hope to select for swine-specific subtraction products at a higher stringency than is normally used in SSH. Results of my analyses indicated that, increasing the amount of strains used in the driver genomic DNA pool did not inhibit the selection for swine specific subtraction products, as two swine probes from a previous subtraction were found to identify the same gene target as several probes in the second subtraction (data not shown). This demonstrates that the results were reproducible and suggests that the subtraction procedures were highly efficient. High suppressive subtraction hybridization (SSH) efficiency has been reported in previous studies. Dick and coworkers (2005(b)) used SSH to successfully develop DNA markers able to distinguish between ruminant and human fecal sources. In a similar study, Shanks et al., (2006) used a genome fragment enrichment (GFE) analogous to SSH, to identify a *Bacteroidales*-like DNA marker capable of determining presence of cattle feces.

While the four DNA marker probes identified a maximum of 39.1% of tested swine *E. coli* isolates individually (illustrating the genetic diversity amongst swine derived strains), combined they identified 62.3% swine *E. coli* strains. As a result, field studies will require the use of pooled hybridization probes to determine swine fecal contributions to waterways. This is similar to results reported by Hamilton et al (2006) who showed that the maximum individual probe detection of *E. coli* strains isolated from geese was 48.1%. However, combination of probes used in colony hybridization analysis identified 76% of targeted *E. coli* isolated from geese, and 73% of non-targeted isolates from ducks, suggesting genetic similarities in *E. coli* originating from both animal hosts.

After translation using BLASTX, the marker genes 2E11 and 2G2 exhibited 100 and 97% amino acid identity, respectively, to transposases found in *E. coli*, suggesting that 2E11 and 2G2 markers may be fragments of transposase genes. Transposition is an important process in adapting to changing living conditions, and it is possible this gene contributes to the survival of *E. coli* in the intestinal tract of swine (Reznikoff, 2003). BLASTX analysis also revealed that marker gene 1A10 was 96% identical to the colicin K immunity protein, a colicin is a bacteriocin (antibiotic) produced by some *E. coli* that is active against other *E. coli* and related species (Feldgarden and Riley, 1999). Colicins are plasmid-encoded proteins, and *E. coli* carrying these plasmids likely possess an ecological advantage over those that lack it.

Since the combined four marker genes collectively identified 62.3% of tested swine *E. coli* isolates, I examined whether the probes cross-hybridized with *E. coli* isolates from cats, chicken, cows, deer, dogs, ducks, goats, geese, humans, sheep, and turkeys. Interestingly, the combined marker probes cross-hybridized to 39.7% of turkey isolates tested (n=126). The nature of this cross-reactivity is largely unknown. However, it could be due to a lack of host specificity in this subgroup of *E. coli*, or the transient population structure in the intestine of some animal hosts (Hamilton et al., 2006; Hartl, 1984). In any event, the cross reaction may limit the use of these probes to environments lacking turkeys. In 2006, Hamilton and coworkers reported similar cross-reactivity of hybridization probes to non-targeted animal hosts (Hamilton et al., 2006). It was found that probes reported in the study combined to identify 76% and 73% of *E. coli* isolated from geese and ducks, respectively. Further, the combined probe set cross-reacted with

approximately 10% of *E. coli* strains from other animal hosts. Similar to the results observed in this study and the Hamilton study Soule et al. (2006), used a microarray approach to identify host-specific *Enterococcus* sp. DNA markers. The identified markers detected from 27 to 45% of *Enterococcus* sp. from targeted host animals, while 1 to 7% of *Enterococcus* sp. from nontargeted host animals were detected. In contrast to the results reported in this study, Bernhard and Field (2000(a)), and Dick et al., (2005(b)) reported that PCR primers designed to identify *Bacteroidales* gene targets from given host animals did not detect non-targeted hosts. However, the results of these studies are difficult to compare to ours, as dilute fecal samples were used as opposed to individual colonies (Hamilton et al., 2006).

Traditional MST methods require the use of large known source libraries, and these are labor intensive to construct, and often fail to identify environmental isolates (Hassan, 2005; Stoeckel, 2004). To rectify the shortcomings, investigations on the use of non-culture based, library-independent, methods represent the current focus in MST technology development. Marker genes targeting *the Bifidobacterium* and *Bacteroides-Prevotella* group of fecal bacteria currently exist to discriminate between a variety of fecal source inputs to waterways (Bernhard and Field, 2000(a), 2000(b); Dick, 2005(a), 2005(b); Shanks 2006). However, because *E. coli* is commonly used for the assessment of water quality, *E. coli* based gene markers can be easily coupled with current regulatory abatement efforts (USEPA, 2001) to provide better information to beach managers and regulatory agencies.

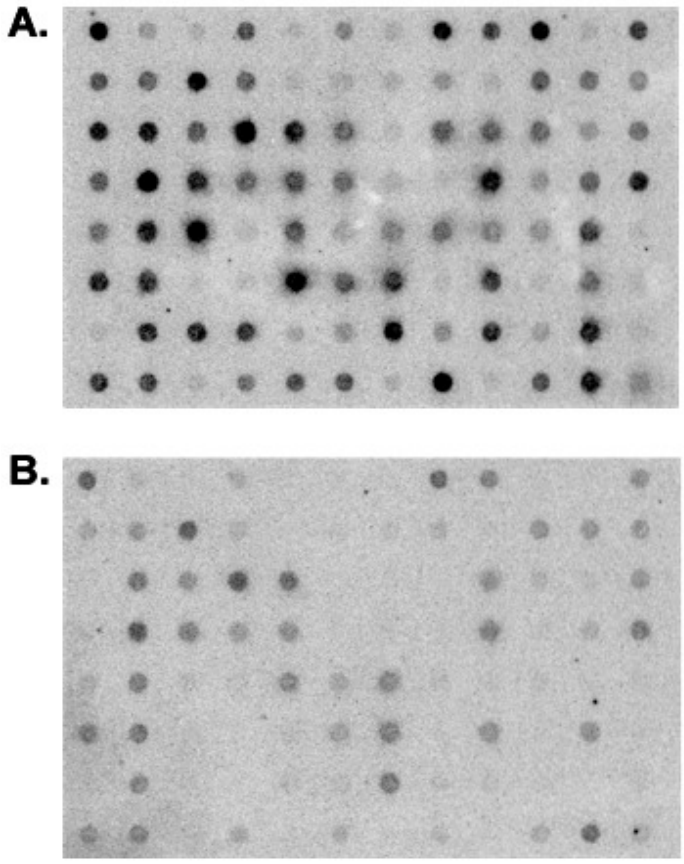
Results of my studies confirm that SSH provides a useful tool for identifying source-specific DNA marker genes from *E. coli.* This study identified a DNA marker gene set capable of identifying a large amount of *E. coli* originating from swine and turkeys, suggesting this probe set may be useful for the assessment of fecal loading in watersheds where only one of these animals is present. While the current study indicates that this marker set may be used as a hybridization probe to identify environmental *E. coli* isolates, more extensive work will be required to determine its geographical specificity and field applicability.

2.6. Acknowledgments

I would like to thank Matthew Hamilton, Daniel Norat, and John Ferguson for their extensive consultation and guidance throughout the course of this study, and for maintenance of the *E. coli* library. I would also like to thank the University of Minnesota Water Resources Center for administration of this grant.

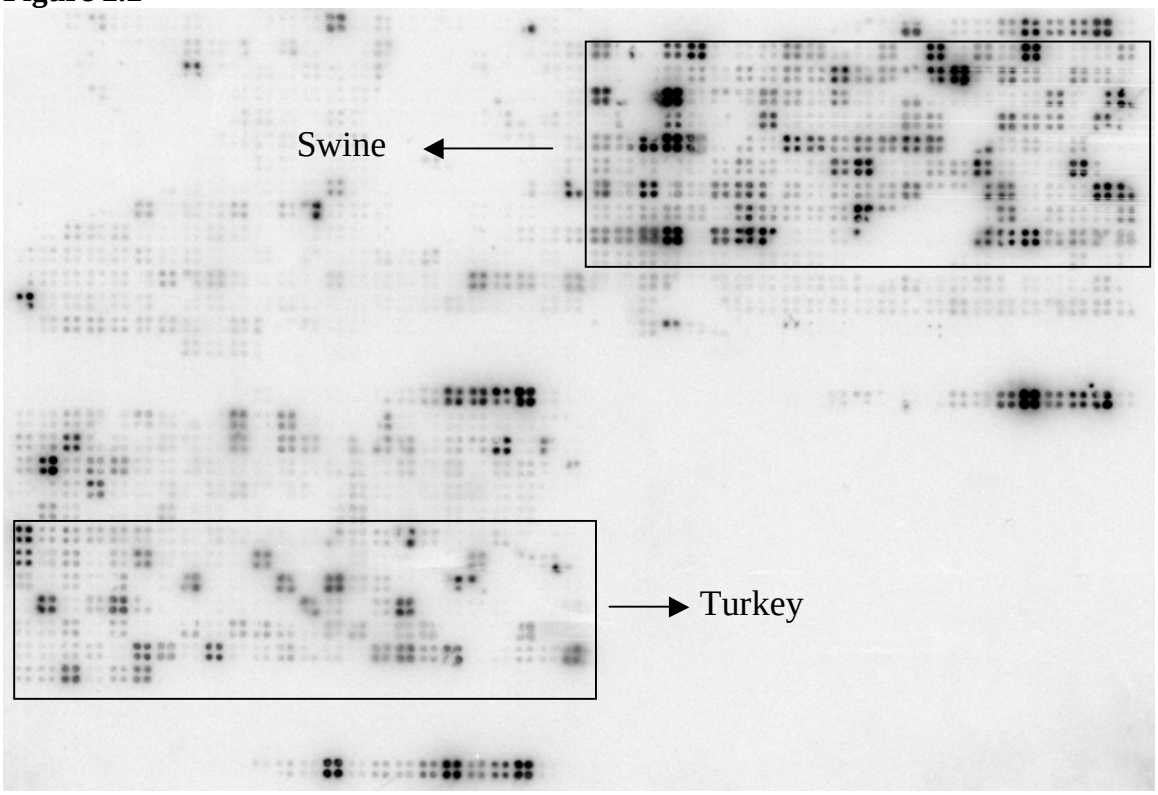
2.7. Tables and Figures

Figure 2.1



Dot blot hybridization of 96 randomly selected subtraction library clones. Panel A. represents PCR-amplified insert DNA's that were hybridized with Rsa1 digested, pooled genomic DNA from *E. coli* isolated from swine. Panel B. represents the amplified inserts subjected to hybridization with Rsa1 digested, pooled genomic DNA's from *E. coli* originating from cats, chickens, cows, deer, dogs, ducks, Canada geese, goats, horses, sheep, turkeys and humans.

Figure 2.2

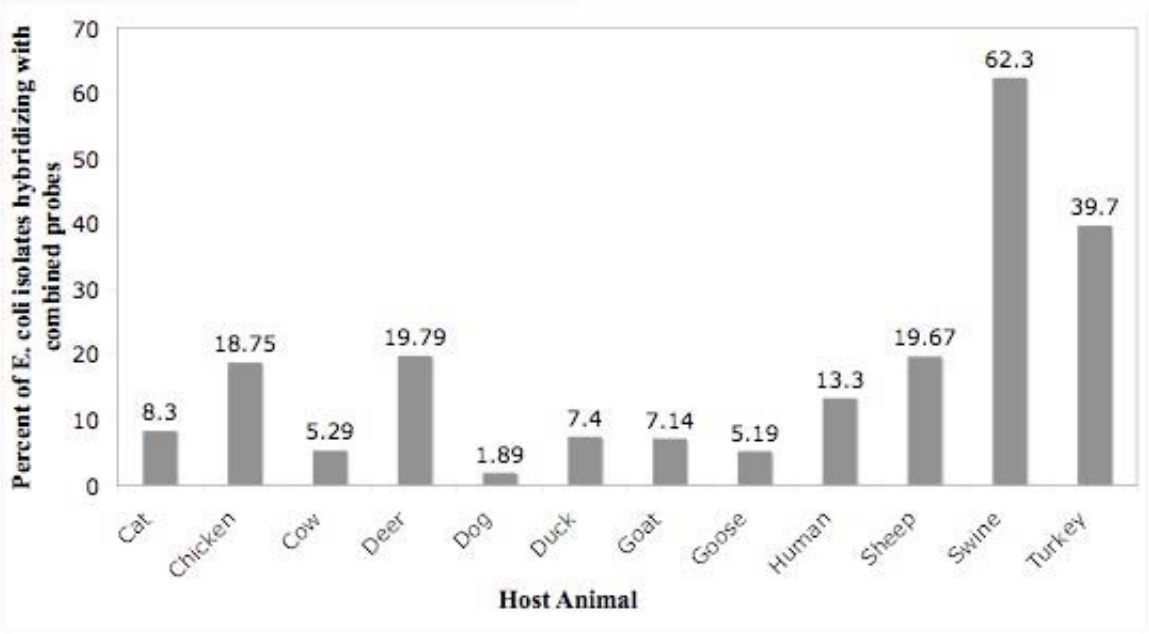


Colony hybridization conducted using probe 2E11. Outlined are membrane regions where swine and turkey isolates were arrayed and subsequently probed. This probe identified 37.% of swine *E. coli* isolates, and 18.3% of turkey isolates.

Table 2.1 Swine-specific DNA marker genes isolated using SSH.

Marker gene	% *E. coli* isolates hybridizing with marker DNA as probes	
	Swine Isolates (*n=215*)	Turkey Isolates (*n=126*)
1H5	39.1	26.2
2E11	37.2	18.3
2G2	28.4	22.2
1A10	10.7	0.01
Total	62.3	39.7

Figure 2.3



Percentages of *E. coli* isolates hybridizing with ^{32}P-labeled marker DNA's 1A10, 1H5, 2E11, and 2G2 combined (to two significant figures, zeros not shown). Values above bars were obtained by visual analysis of colony hybridization intensity.

Table 2.2 DNA probes exhibiting hybridization specificity to swine *E. coli* isolates.

Insert DNA	Length (bp)	Protein Homolog in database	GenBank accession no.	% identical	E value
1A10	460	Colicin K immunity protein (*E. coli*)	AAX40978	96	8.00E-27
1H5	532	Putative tail protein from bacteriophage (*E. coli* UMN026)	YP_002411736	89	1.00E-49
2E11	251	Transposase IS10, a IS4 family protein, (*E. coli* UMN026)	YP_002405951	100	4.00E-43
2G2	168	Transposase ORFB IS629 (*E. coli*)	ACQ42080	97	2.00E-24

2.8. References

Albert, J.M.; Munakata-Marr, J.; Tenorio, L.; Siegrist, R.L. 2003. Statistical evaluation of bacterial source tracking data obtained by rep-PCR DNA fingerprinting of *Escherichia coli*. Environ. Sci. Technol. 37 (20), 4554–4560.

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

Bernhard, A. E., and Field, K.G. 2000(b). Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66:1587–1594.

Byappanahalli, M.; Fowler, M.; Shively, D.; Whitman, R. 2003. Ubiquity and persistence of *Escherichia coli* in a Midwestern coastal stream. Appl. Environ. Microbiol. 69 (8), 4549–4555.

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

Dick, L. K; Bernhard, A. E.; Brodeur, T. J.; Santo Domingo, J. W.; Simpson, J. M.; Walters, S. P.; Field, K.G. 2005(a). Host distributions of uncultivated fecal *Bacteroidales* reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. 71:3184–3191.10.

Dick, L. K.; Simonich, M. T.; Field, K. G. 2005(b). Microplate subtractive hybridization to enrich for *Bacteroidales* genetic markers for fecal source identification. Appl. Environ. Microbiol. 71:3179–3183.

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

Feldgarden, M.; Riley, M. A. 1999. The phenotypic and fitness effects of colicin resistance in Escherichia coli K-12. Evolution. 53(4), 1019-27.

Field, Katharine G.; Samadpour, Mansour. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research. 41. 3517–3538.

Gordon, D. M. 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology 147:1079–1085.

Griffith, J.F.; Weisburg, S.B.; McGee, C.D. 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. J. Water Health 1, 141–152.

Hamilton, M.J.; Yan, T.; Sadowsky, M.J. 2006. Development of goose- and duck-specific DNA markers to determine sources of *Escherichia coli* in waterways. Appl. Environ. Microbiol. 72 (6),4012–4019.

Hartel, P. G., J. D. Summer, J. L. Hill, J. Collins, J. A. Entry, Segars W.I. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. J. Environ. Qual. 31:1273–1278.

Hartel, P. G.; Summer, J. D.; Segars, W.I. 2003. Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking. Water Res. 37:3263–3268.

Hartl, D. L., and D. E. Dykhuizen. 1984. The population genetics of Escherichia coli. Annu. Rev. Genet. 18:31–68.

Hassan, W.M.; Wang, S.Y.; Ellender, R.D. 2005. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Appl. Environ. Microbiol. 71 (1), 512–518.

Indest, K.J.; Betts, K.; Furey, J.S. 2005. Application of oligonucleotide microarrays for bacterial source tracking of environmental *Enterococcus sp.* isolates. Int. J. Environ. Res. Public Health 2 (1), 175–185.

Ishii, S.; Ksoll, W.B.; Hicks, R.E.; Sadowsky, M.J. 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), 612–621.

Jenkins, M. B.; Hartel, P. G.; Olexa, T. J.; Stuedemann, J.A. 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. J. Environ. Qual. 32:305–309.

Johnson, L. K.; Brown, M. B.; Carruthers, E. A.; Ferguson, J. A.; Dombek, P. E.; Sadowsky, M. J. 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70:4478–4485.

Kaspar, C.; Burgess, J.; Knight, I.; Colwell, R. R. 1990. Can. J. Microbiol., 36, 891-894.

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

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

Olapade, O.A.; Depas, M.M.; Jensen, E.T.; McLellan, S.L. 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Appl. Environ. Microbiol. 72 (3), 1932–1938.

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

Reznikoff, William S. 2003. Tn5 as a model for understanding DNA transposition. Molecular Microbiology 47 (5): 1199-1206.

Sadowsky, M. J.; Tully, R. E.; Cregan, P. B.; Keyser, H. H. 1987. Genetic diversity in *Bradyrhizobium japonicum* serogroup 123 and its relation to genotype-specific nodulation of soybeans. Appl. Environ. Microbiol. 53:2624–2630.

Samadpour, M. ; Roberts, M.C. ; Kitts, C.L. ; Mulugeta, W.; Alfi, D. 2005. The use of ribotyping and antibiotic resistance patterns for identification of host sources of *Escherichia coli* strains. Appl. Microbiol. 40, 63–68.

Sambrook, J.; Fritsch, E. F.; Maniatis, T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Scott, T. M.; Parveen, S.; Portier, K. M.; Rose, J. B.; Tamplin, M. L.; Farrah, S. R.; Koo, A.; Lukasik, J. 2003. Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. Appl. Environ. Microbiol. 69:1089–1092.

Shanks, O.C.; Santo Domingo, J.W.; Lamendella, R.; Kelty, C.A.; Graham, J.E. 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72 (6), 4054–4060.

Soule, M.; Kuhn, E.; Loge, F.; Gay, J.; Call, D. R. 2006. Using DNA microarrays to identify library-independent markers for bacterial source tracking. Appl. Environ. Microbiol. 72:1843–1851.

Stoeckel, D.M.; Mathes, M.V.; Hyer, K.E.; Hagedorn, C.; Kator, H.; Lukasik, J.; O'Brien, T.L.; Fenger, T.W.; Samadpour, M.; Strickler, K.M.; Wiggins, B.A. 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. 38 (22), 6109–6117.

Stoeckel, D.M., and Harwood, V.J., 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. 73 (8), 2405–2415.

U.S. Environmental Protection Agency. 2005. Microbial Source Tracking Guide Document. Office of Research and Development, Washington, DC, EPA-600/R-05/ 064, 131pp.

Whitman, R.L.; Shively, D.A.; Pawlik, H.; Nevers, M.B.; Byappanahalli, M. 2003. Occurrence of *Escherichia coli* and enterococci in *Cladophora* (*Chlorophyta*) in nearshore water and beach sand of Lake Michigan. Appl. Environ. Microbiol. 69, 4714–4719.

Whitman, R.L.; Byers, S.E.; Shively, D.A.; Ferguson, D.M.; Byappanahalli, M. 2005. Occurrence and growth characteristics of *Escherichia coli* and enterococci within the accumulated fluid of the northern pitcher plant (Sarracenia purpurea L.). Can. J. Microbiol. 51 (12), 1027–1037.

Wiggins, B.A.; Cash, P.W.; Creamer, W.S.; Dart, S.E.; Garcia, P.P.; Gerecke, T.M.; Han, J.; Henry, B.L.; Hoover, K.B.; Johnson, E.L.; Jones, K.C.; McCarthy, J.G.; McDonough, J.A.; Mercer, S.A.; Noto, M.J.; Park, H.; Phillips, M.S.; Purner, S.M.; Smith, B.M.; Stevens, E.N.; Varner, A.K. 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. 69, 3399–3405.

Yan, T; Hamilton, M.J.; Sadowsky, M.J. 2007. High-Throughput and quantitative procedure for determining sources of *Escherichia coli* in waterways by using host-specific DNA marker genes. Appl Environ Microbiol. 73(3): 890–896.

Chapter 3. Examination and mitigation of bovine fecal loading to the Little Jordan Creek near Chatfield, Minnesota.

3.1. Overview

Fecal contamination of waterways represents a consistent threat to public health. In Minnesota's agricultural regions, cattle grazing operations are abundant, often leading to direct and indirect contamination of waterways. In this study, I assessed the spatial and temporal dynamics of bovine fecal loading to a small creek system in Southeastern Minnesota, The Little Jordan Creek (LJC). The purpose of this study was to (1) determine the scale and source of fecal loading to the creek in 2008; and (2) subsequently employ a best management practice (BMP) at the beginning of and throughout the 2009 sampling season to mitigate fecal loading to the system. The USEPA method 1603 was used to determine the scale of *E. coli* presence at five locations in 2008 and 2009, and conventional polymerase chain reaction (PCR) and quantitative PCR assay were used to determine fecal inputs into the LJC. My analyses revealed that the total *E. coli* counts across all sampling locations in 2008 exceeded state and federal water quality standards, with the highest levels observed where the creek leaves the property. Three major rain events were sampled in 2008, and *E. coli* numbers spiked considerably on these dates. A bovine-specific, *Bacteroidales*-like, marker gene (Bac3) was detected on all sampling dates in 2008, with an overall mean across sites of 1,433 copies for the season. Best Management practice (BMP) remediation plans were initiated in 2009 to reduce direct cattle exposure to the creek, rehabilitate upland areas, and curb over-grazing. The mean

E. coli level in 2009 across sampling locations and dates was 859 CFU *E. coli*/100ml. Unlike 2008, the Bac3 marker gene was not detected on each sampling date, and had an overall mean detection of 998 copies across dates and sites in 2009. The human marker was not detected by conventional PCR in 2008 or 2009. While the *E. coli* levels in 2009 were lower than observed in 2008, it is difficult to surmise the BMP's role, as fewer rain events were sampled in 2009 and the BMP was not completed before the culmination of the sampling season. It is likely that during significant rain events naturalized *E. coli* present in creek sediments are liberated due to increased flow, and observed CFU counts raise accordingly. As a result, USEPA method 1603 may not accurately characterize the scale of fecal loading to stream systems during significant rain events. Further, rain events may skew long-term *E. coli* count data used to determine impairments. Observed CFU *E. coli* did not correlate well with detection of the bovine marker gene, with the exception of sampling locations that had direct cow exposure. The copy number of the bovine marker gene decreased considerably in 2009 at the sampling location exiting the property, suggesting that initiation of the BMP may have helped to achieve lower amounts of fecal loading to the system at this location, and may be successful on a watershed scale.

3.2. Introduction

Fecal loading of aquatic environments by animals and humans is of concern for public and environmental health. The deposition of fecal bacteria into waterways is thought to be one of the main routes by which humans are exposed to pathogens. The

microbiological contamination of waterways is amongst the most commonly listed water quality impairment in the U. S. (USEPA, 2005). The federal Clean Water Act (CWA) requires that individual states provide the U. S. Environmental Protection Agency (USEPA) with an aggregate quality assessment of its waterways every biennium (section 305(b) list) in effort to identify waters that do not meet state and federal quality standards (the section 303(d) list). Once deemed impaired, states must conduct a Total Maximum Daily Load (TMDL) assessment in order to mitigate the impairment and restore the water body to acceptable quality. Successful fecal coliform TMDL implementation strategies require microbial source tracking (MST) technologies that accurately and efficiently characterize the host-specific source(s) of bacterial loading.

Traditional MST methods require the use of large known source libraries. While it has been shown that smaller libraries have higher average rates of correct classification (ARCC), they are not as representative at a watershed scale as larger libraries (Moore et al., 2005; Wiggins et al., 2003). The construction of large libraries is very labor intensive, and despite their size, they often fail to identify the host origins of environmental isolates (Hassan et al., 2005; Stoeckel et al., 2004). Furthermore, known-source libraries are limited by the temporal and geographical variation of target bacterial genotypes, variation in the diets of host animal groups, and the naturalization of target organisms to aquatic environments (Byappanahalli et al., 2003; Byappanahalli et al., 2006; Gordon, 2001; Hartel et al., 2002, 2003; Ishii et al., 2006; Olapade et al., 2006; Power et al., 2005; Scott et al., 2003; Whitman et al., 2003, 2005).

To rectify the shortcomings of library- based MST methodologies, investigation of the use of non-culture based, library independent methods represents the current focus of MST technology development. Such methods aim to circumvent the use of large known source libraries by identifying host-specific bacterial gene targets (markers) that are found in the fecal material of animals. To date, genetic markers have been developed and reported to identify several host-specific fecal inputs into aquatic environments. Bernhard and Field have investigated the use of 16S rRNA-based genes from *Bifidobacterium* and *Bacteroides-Prevotella* that can discriminate between human and ruminant fecal input sources to waterways (Bernhard and Field, 2000(a), 2000(b)). Further, Dick and coworkers have successfully used subtraction hybridization to identify host specific *Bacteroides* 16S rRNA markers, and in a subsequent study, investigated *Bacteroidales* 16S rRNA gene sequences to identify swine and fecal loading due to horses (Dick et al., 2005(a), 2005(b)). These marker genes have been subsequently used by others to examine fecal loading to waterways due to human activities (Bower et al., 2005). Further, Hamilton and coworkers used subtraction hybridization to identify duck and goose specific *E. coli* marker genes (Hamilton et al., 2006). Similarly, Shanks and coworkers used a gene fragment enrichment (GFE) technique to identify cow specific *Bacteroidales*-like marker genes (Shanks et al., 2006, 2008).

In this study, I examined the spatial and temporal dynamics of bovine fecal loading to a small creek system (the Little Jordan) in Southeastern Minnesota. The watershed is home to a host of agricultural practices, and cattle grazing is prevalent in direct upland areas of the creek. Fecal coliform and *E. coli* plate count data was obtained

using USEPA method 1603 to determine potential regulatory impairments (USEPA, 2002). A quantitative PCR protocol was developed for the bovine *Bacteroidales*-like 16S rDNA marker Bac3 (Shanks et al., 2006) to aide in determining the scale of bovine fecal contributions to the Little Jordan Creek system. Results from the Bac3 qPCR assay and plate count analysis were compared, and minimal correlation was observed during rain events. My results suggest that USEPA method 1603 may not be an appropriate tool for classifying the scale of fecal pollution during wet periods.

3.3. Materials and Methods

3.3.1 Sampling location description.

Field investigations were performed on the Little Jordan Creek, near Chatfield, Minnesota (Figure 3.1). Cattle grazing is prevalent within the Little Jordan watershed, and animals are frequently allowed unrestricted access to the creek for the majority of the grazing season. In 2008, five sampling sites were selected in cooperation with a local landowner to assess bovine fecal loading to the waterway. The strategy of location selection was to include areas of the creek that were thought to be significantly impacted by grazing practices, and to those thought to have less cattle impact. In 2008, sites 1 and 3 were selected as areas that would not be significantly influenced by the landowners grazing practices. Sites 2, 4, and 5 all were expected to have high exposure to cattle. In 2009, three new sites were selected upstream to determine potential sources of fecal inputs based on observations from the 2008 sampling season. In 2009, the main spring

source of the creek, site 8, was sampled, along with two additional sites (7 and 9). These were selected based on geographical proximity to the creeks source.

While the creek is relatively small, it is actively managed by the Minnesota Department of Natural Resources (MNDNR) as a designated trout stream. Although there is an abundance of agricultural practices within the watershed area, there is also considerable arboreal growth and very few impervious surfaces. As a result, only large scale rain events had noticeable affects on stream level and turbidity.

3.3.2 Environmental sample collection.

In 2008, samples were collected from April to September. For initial samples, through July 14, 2008, three 700 ml samples were collected per sampling location along the Little Jordan Creek, near Chatfield Minnesota. However, beginning on July 31, 2008, six 1000 ml samples were collected to ensure adequate availability of water for analyses. Samples were collected in either sterile 710 mL capacity Whirl-Pak bags (eNasco, Fort Atkinson, WI) or sterile polypropylene copolymer Nalgene bottles (Thermo Fisher Scientific Inc., Waltham, MA). Time was given to allow stream sediments to settle after entry, and high flow areas at each location were chosen to retrieve sample. Samples were placed on ice during return to laboratory, and held at 4ºC until processed the next morning. Temperature readings were obtained at each location with a digital thermometer during each sampling occasion. Transparency tube (T-tube) measurements were also taken to determine the scale of sediment suspension in the water column (MPCA, 2005).

3.3.3 Filtering strategy and generation of plate count data.

Samples were processed according to EPA Method 1603, using modified membrane-thermotolerant *Escherichia coli* agar medium (modified mTEC) using Millipore S-Pak 0.45 µm sterile gridded filters (Millipore, Billerica, MA) (USEPA, 2002). In addition to using mTEC medium to obtain *E. coli* plate count data, mFC medium (Difco, Detroit, MI) was also used to determine total fecal coliform counts. Three replicate 100, 10, and 1 ml volumes of each sample from each site were filtered and each dilution was performed in duplicate per sample replicate. However, due to occasional sample loss, the replicate numbers obtained from each sampling site sometimes varied. The 10 and 1 ml volumes were placed in 10 and 9 ml of phosphate buffered saline (Sambrook, Fritsch, and Maniatis, 1989), respectively to achieve uniform cell distribution on filters. Filters were placed onto the surface of mTEC and mFC media and incubated at 37 ºC for 2 hours to recover cells, and then at 44.5ºC for 24 hours (Ishii et al., 2006). After incubation, blue colonies on plates were counted. Filtration volumes that yielded colony numbers between 30 and 300 were selected and used to report plate count data. The duplicate counts per volume filtered were averaged for each replicate. Subsequently the average was taken of the three duplicate outputs, and this number was reported as the total count per respective site.

3.3.4 Molecular methods.

Water samples (700 to 1000 ml) for DNA analysis were filtered through 0.45µm sterile filters (Millipore, Billerica, MA). Filters were stored at -80ºC until further

processing. Filters were cut into small pieces using sterile razor blades and DNA was extracted from entrapped microorganism using the MO-BIO Power Soil DNA Extraction Kit (MO-BIO Laboratories, Carlsbad, CA). The concentration of extracted DNA in each sample was determined using an Eppendorf (nanodrop)Bio-Photometer (Eppendorf AG, Hamburg, Germany). DNA concentrations of samples were normalized to 3 ng/µl, such that 5 µl of sample could be added to each PCR tube to obtain 15 ng of total DNA per reaction.

Marker gene presence in environmental samples was determined by using conventional PCR assays targeting total Bacteroides (AllBac) and the human-specific 16S rRNA HF8 gene cluster (HF183) (Bernhard and Field, 2000(b); Layton et al., 2006). A qPCR method was also developed to target a bovine-specific, *Bacteroidales*-like, 16SrRNA gene marker, Bac3 (Table 3.1) (Shanks et al., 2006). The thermal and chemical conditions used for PCR of environmental DNA extracts for the universal (AllBac) and human-specific (HF183) PCR assays were conducted as previously described (Bernhard and Field, 2000(b); Layton et al., 2006) with the exception of the HF183 assay, which was done using an annealing temperature of 63˚C.

Optimization of the Bac3 qPCR assay was done by initially using the thermal protocol as described by Shanks et al. (2006), using 2x iTaq SYBR Green Supermix With ROX (Bio-Rad Laboratories, Hercules, CA), and 200 nM of each primer. Thermal and chemical optimization was necessary to improve qPCR reactions using the Applied Biosystems ABI PRISM 7000 Sequence Detection System (Foster City, CA). The cycling protocol was adjusted to include a 2 min step at 50˚C before initial denaturation

and enzyme activation, and excluded a final 72˚C extension step. Primer concentrations were tested between 200nM and 300nM and annealing temperatures were tested between 60˚C and 61˚C. The optimized Bac3 qPCR assay consisted of the following protocol: 2 min at 50˚C, followed by a 10 min cycle at 95˚C for enzyme activation and denaturation, followed by 40 cycles of denaturing at 95˚C for 15 sec, and annealing and primer extension steps of 1 min at 61˚C. Following amplification, the dissociation of samples was measured between 65ºC and 95ºC. The optimized reaction mixture (25 μl) contained 2x iTaq SYBR Green Supermix With ROX (Bio-Rad Laboratories, Hercules, CA), and 300 nM of each primer. For standard curve generation, 3×10^5, 1×10^5, 3×10^4, 1×10^4, 3×10^3, 1×100^3, 3×10^2, 1×10^2, 30, 10, or 3 copies of the 166 bp Bac3 gene fragment (which was previously cloned into the StrataClone PCR cloning vector pSC-A-amp/kan (Stratagene, La Jolla, CA), were obtained from stock tubes frozen at -80˚C) was added to individual reaction mixtures. Each environmental sample run included a standard curve run to ensure calibration accuracy. For environmental sample runs, 15 ng of extracted DNA was added to the reaction mixture. The qPCR assays were performed in an Applied Biosystems ABI PRISM 7000 Sequence Detection System (Foster City, CA), using optically clear 96 well microplates. Samples were run in triplicate to determine qPCR variability. The mean of the qPCR output of triplicate samples were averaged.

3.3.5 Determination of qPCR detection.

Simple detection was defined as the circumstance when at least one of the triplicate samples gave a copy number average above the limit of detection. Numerical

detection was defined as the circumstance when ≥ 2 of 3 copy number averages from each DNA sample were above limit of detection, and the subsequent average of those replicates was above limit of detection.

3.3.6 Statistical analysis.

Mean CFU counts and mean total copy number were calculated from sample replicates for plate count analysis and the qPCR output, respectively. The standard errors of the means were also calculated and the ANOVA program (www.physics.csbsju.edu/stats/anova.html), at $\alpha = 0.05$, was used to assess the statistical significance of spatiotemporal variation in CFU counts and total Bac3 copy number in the Little Jordan Creek. Pair-wise comparisons were subsequently done by determining Fisher's Least Significant Difference (LSD) value between sample means.

3.4 Results

3.4.1 Plate count data.

In the Little Jordan watershed, there are three landowners actively grazing cattle, providing the animals with unrestricted access to the creek. At least one landowner employs no rotational grazing strategy, allowing significant degradation of riparian vegetation. Recent rain events, proximity of cows to the creek, and water temperature were major factors influencing the *E. coli* and fecal coliform plate count data observed. In 2008, three significant rain events occurred within a 48 h period prior to sampling (June 9, July 13, and July 31). Cattle or fecal material were often observed in and adjacent to

the creek. Early in the 2008 sampling season, cows were distributed amongst the entire property where sampling occurred. As the season progressed, grass based foodstuffs became limited, and the landowner began gathering cattle near site number 4 to feed. In 2009, cattle appeared to be grazing in a similar manner on the property sampled in 2008. A rotational grazing strategy was designed in the summer of 2009 in cooperation with one of the three landowners. While cows were still observed in and adjacent to the creek in 2009, it was with less frequency than in 2008. Water temperature averages during both sampling seasons gradually increased until August, and fell off slightly as September progressed.

While grazing patterns were difficult to assess in 2008, the predictability of either an increase or decrease in observed plate counts at sites 4 and 5 was high, as all fecal inputs to the creek have the potential to influence counts at these locations. In 2009, grazing patterns at designated locations were easier to predict, and the same phenomena were observed. Site number 1 in 2008 was located on the landowner's upstream property line, and received the least exposure to his cattle throughout the season. In 2009, site number 7 was located on the property of a third landowner actively grazing cattle, but not participating in the study.

Results in Figure 3.2 shows count data of *E. coli* and fecal coliforms (as CFU per ml) at each date and site during the 2008 sampling season. The data shown in Figure 3.2 indicate that recent rain events, stockpiled manure, and the proximity of cows to the creek were likely major factors influencing the relatively high counts observed. The Minnesota State standard of 126 CFU *E. coli* per 100 ml water was exceeded at each sampling

location throughout the sampling season. Site 1 consistently had the lowest counts on both media, with the exception of rain-event sampling dates (June 9, July 13, and July 31), where counts averaged 1,007.5 and 880.7 CFU/100ml for fecal coliforms and *E. coli*, respectively. The site 1 exceeded the state standard 55.6% of the dates sampled. Surprisingly, site 2 averaged the second highest counts at 2,401.4 and 2,549.7 CFU/100ml for fecal coliforms and *E. coli*, respectively, and exceeded the 126 CFU *E. coli*/100ml standard 77.7% of sampling visits. Site 3 exceeded the standard 87.5% of sampling visits, averaging 1,764.9 and 1,337.8 CFU/100ml for fecal coliform and *E. coli*, respectively; and sites 4 and 5 exceeded the standard each sampling visit. Average fecal coliform and *E. coli* counts at site 5 were, 3,902.7 and 4038.4 CFU/100ml for fecal coliform and *E. coli*, respectively, were the greatest measured at any site during the 2008 sampling season. Site 4 averages were the third highest with 2,178.6 and 2,393.3 CFU/100ml fecal coliform and *E. coli*, respectively. Interestingly, average *E. coli* counts exceeded fecal coliform averages at sites 2, 4 and 5. Same site sample replicates had minimal variation, and significant differences determined by Fisher's LSD are indicated by alphabetical association (Tables 3.2-3.3).

While precipitation during the 2009 sampling season was comparable to what was observed in 2008 (Table 3.4), it was largely distributed throughout the summer, and fewer significant rain events occurred. As a result, one rain event sampling occurred in 2009 (August 17). Results in shown in Figure 3.3 represent total *E. coli* and fecal coliform counts per date and site during the 2009 sampling season. In 2009, proximity of cows to the stream and the rain event proved to be the largest factors contributing to

fluctuations of numbers of fecal coliforms and *E. coli.* The Minnesota State standard was exceeded at each sampling location at least once in 2009. Site 8, the creeks main spring source, averaged the lowest numbers of both fecal coliform and *E. coli* by a large margin at 116 and 63.7 CFU/100ml, respectively. Moreover, site 8 exceeded the state standard 22.2% of sampling dates. Interestingly, site 7 averaged the highest counts for both fecal coliform and *E. coli*, at 2627.6 and 1607.1 CFU/100ml, respectively, and exceeded the standard 88.9% of sampling date. Site 5 followed averaging 2021.1 and 1235 CFU/100ml for fecal coliform and *E. coli* respectively, and each exceeded the standard on nine sampling dates in 2009. Observed counts at site 4-2 averaged 1846.9 and 920.2 CFU/100ml for fecal coliforms and *E. coli*, respectively, and this site exceeded the state standard on 77.8% of the dates sampled. Site had the second lowest counts at 649.7 and 467.4 CFU/100ml for fecal coliforms and *E. coli*, respectively. Fecal coliform counts averaged higher at each sampling location than *E. coli* counts in 2009. Statistically significant differences in *E. coli* and coliform count averages at the sites and across dates were tested by using Fisher's LSD, and are shown in Tables 3.5 and 3.6.

3.4.2 Conventional PCR to determine presence of all Bacteroides and human derived Bacteroides.

Two conventional PCR primer sets were used in 2008 and 2009 to examine DNA's extracted from water at the Little Jordan Creek study location. The general probe AllBac was used to determine the presence of all *Bacteroides sp.* strains (Layton et al., 2006), while the human-specific probe HF183 was used to identify potential inputs of

human derived *Bacteroides* sp. (Bernhard and Field, 2000(b)). Each replicate DNA extraction tube obtained in 2008 and 2009 was subjected to both of these assays. The AllBac PCR assay identified *Bacteroides* targets in all samples processed, indicated that the DNA extracted was capable of being used for PCR. In contrast, the HF183 marker genes was not observed in any of the environmental samples obtained during the study, indicating that human inputs to the LJC were likely not contributing to elevated fecal loading of this waterway.

3.4.3 Generation of qPCR calibration curves for the Bac3 primer set.

Before environmental samples could be analyzed for concentration of the Bac3, Bovine-specific target gene, optimization of the chemical and thermal protocols for qPCR was necessary. Initially, the thermal protocol as described by Shanks et al. (2006) was used. The primer concentration initially tested was 200 nM as per manufacturers recommendations (Bio-Rad Laboratories, Hercules, CA). The standard run at these conditions yielded an indeterminable detection limit and low amplification efficiency of 71.8% (E=-1+10$^{(-1/slope)}$) (Figure 3.4, Panel A). The Shanks (2006) protocol included a final 72˚C extension step, which was used during the quantification stage in this reaction. Because the majority of target product had likely been extended prior to this step, amplicon quantification was likely inaccurate using this step. Accordingly, a second qPCR run was done excluding the 72˚C extension step. This resulted in a detection limit of 1,000 copies (R^2=0.9983), and a slight increase in amplification efficiency to74.2%. Analysis of the dissociation curve for this experiment revealed that a secondary product

was amplified. This was likely due to primer- dimerization, and amplification of this false target. To alleviate this problem, the annealing temperature was increased to 61˚C. Increasing the annealing temperature to 61˚C improved the detection limit, to 300 copies (R^2= 0.9945), and decreased the amount of secondary product observed in the dissociation curve. However, the amplification efficiency decreased to 69.4%. To improve the amplification efficiency, the primer concentration in subsequent reactions was increased to 300nM. These conditions proved to be optimal, as amplification efficiency improved to 89.9% and the detection limit dropped to 100 copies (R^2=0.9947) (Figure 3.4, Panel B).

3.4.4 Presence of the Bac3 bovine-specific marker gene in environmental samples.

The optimized Bac3 qPCR assay was used to determine the amount of bovine-specific marker genes in environmental DNA extracts obtained from the aforementioned sampling locations in 2008 and 2009. The Bac3 qPCR assay identified the presence of bovine-specific *Bacteroidales*-like target genes in every sample obtained during the 2008 season (Figure 3.5, Panel A). The variability in copy numbers among replicates was high using this assay (data not shown). As a result, site copy number averages spanning the 2008 sampling season were difficult to interpret. The highest observed copy number was at site 2, where the assay quantified 25,058 copies of the *Bacteroidales*-like target. Due to this, site 2 averaged the highest copy number during the 2008 sampling season. Surprisingly, site 1 had the second highest copy number average in 2008 (1,397 copies), despite the fact that the site was initially thought to have limited impact due to grazing

cattle. This was followed by copy number averages of 1,226, 1,131, and 922 at site 5, 3, and 4, respectively. Interestingly, site 5 had the greatest copy number of the bovine marker gene on 66.7% (6/9) of the sampling dates, and site 4 had higher copy numbers than sites 1, 2, and 3 on 55.6% (5/9) of sampling dates. Thus, sites 4 and 5 generally had greater numbers of the Bac3 marker gene than sites 1, 2, and 3 on the majority of sampling dates. This correlated well with the presence of a cattle feeding area located adjacent to site 4, and upstream of site 5.

In contrast to what was found in 2008, the Bac3 qPCR assay did not identify the bovine-specific *Bacteroidales*-like target gene at each sampling date in 2009 (Figure 3.5, Panel B). However, like 2008, replicate copy number variability was high. The greatest copy number, 7,911, was observed at site 9 on September 2, 2009. This resulted in site 9 having the highest overall copy number average in 2009. Site 4-2 averaged the second highest copy number, 1,315, in 2009. This was followed by averages of 1,241, and 734 for sites 7 and 5, respectively. As expected, site 8 averaged the lowest number of bovine-specific marker genes, 189, observed in 2009. Relatively speaking, site 7 had the highest copy number on 37.5% (3/8) of sampling dates, while site 5 did not yield the highest observed copy number at any sampling date during the 2009 season.

3.5. Discussion

3.5.1 Plate count data.

Total fecal coliform counts on the April 2 and April 28, 2008 sampling dates (the first two of the season) were comparatively low, but variation between sites was seen as

expected (Figure 3.2). Low numbers were likely due to several factors, including low water temperatures and low initial influence of cow manure in immediate upland areas throughout the property. In addition, during this early part of the season animals were not given direct access to the stream corridor. Cows were largely confined to the area near site 5, likely explaining the highest counts observed at this site on both dates. Counts from the April 28 sampling date were on average higher than the previous samples, likely due to an increase in water temperature, and the fact that cows were allowed to graze at a closer proximity to the stream near site 5.

On May 12, 2008 the cows were distributed amongst the entire property with no apparent grazing strategy. This was the first observed occasion that all cows were given unrestricted access to the stream. The *E. coli* plate count data for May 12, 2008 likely reflects this observation, as counts were significantly higher at sites 2 to 5, with sites 4 and 5 exceeding the state standard of 126 CFU/ 100mL for the first time. By May 12, 2008 the water temperature had risen since, averaging 11º C with the lower range observed in upstream locations, and the higher range observed downstream. This is true for all temperature data gathered during the 2008 sampling season.

On May 28, 2008, the second sampling date, the cows were again largely distributed amongst the whole property, with the exception of site 5. The land parcel proximate to the road had been sold, and cows that were grazing in the immediate upland areas adjacent to site 5 had been moved upstream of the property border. This did not appear to affect the numbers of fecal coliform and *E. coli* present at this location, however, as the plate count data were higher than on the May 12 sampling date. This was

true for all sites, and likely due to an increase in the average water temperature to 13º C. It should be noted, that for both the May 12 and May 28 sampling dates, high counts were observed at site 3. It was thought that this was likely due to significant amounts of sediment obtained in water samples at this location due to low flow rates. Based on this, it was decided that future samples would be taken from just over the fence-line, where the flow rate was significantly higher.

The June 9 sample marked the first date that sampling occurred within 72 h of a significant rain event (8.9 centimeters of rain in the 48 h prior to sampling (http://www.climate.umn.edu/). The water level had risen 25.4 centimeters since May 28, and high turbidity measurements were observed. The water level was determined by drop-down style measurement from two culverts exiting the property downstream of site 5. The effect of the rain event was manifested in elevated plate count data. Interestingly, counts were similarly high for all sites, including site 1, which had previously been consistently low. This is perhaps due to runoff carrying additional fecal material to sections of the stream that would regularly not be influenced, combined with increased flow disturbing sediment and liberating bacteria that would normally not be identified by the water sampling technique used. By this date that average water temperature had risen to approximately 14.5º C.

By the June 25 sampling date the average water temperature had risen significantly to 17º C. Despite this, *E. coli* counts were less, on average, from those seen on the June 9th rain event sampling, and from the May sampling events at sites 4 and 5. This trend may be explained by the absence of cows from the immediate upland areas of

these sites sometime after the May 12 sampling date, coupled with a reduction in stream turbidity, which had decreased considerably from the June 9, 2008 sampling date.

With the exception of site 2, *E. coli* counts for each of the other four sites exceeded the state water quality standard on the July 14 sampling date. Count data at sites 1 and 3 were high, potentially due to the 2.6 cm rainfall that fell during the previous 72 h. Coliform and *E. coli* counts continued to be greater than expected in the higher flow area near site 3. During this time period the cows were largely confined to an area near Site 4, where the landowner had began to supply feed. Water temperature averages were the same as on June 25, and water level and turbidity were average (average turbidity is ~55 cm of clarity when using a Transparency Tube. The average water level was ~165 cm below the culvert).

On July 31, 2008 a significant rain event (~2.5 cm) began upon arrival to the sampling location. Plate counts for this sampling event were the greatest recorded to this date. Turbidity and water levels exceeded observed averages, and water temperatures were the greatest recorded, averaging over 18ºC. As before, cows were largely confined to a pen area near site 4. Interestingly, on this date plate counts of fecal coliform bacteria and *E. coli* were the second lowest recorded at site 4. This may be in part due to the riparian vegetation that had grown in the upland areas adjacent to site 5, as the new landowner had rotated cows out of this area since mid-June. However, plate count results were similar at all sites, likely due to sediment re-suspension that distributed attached bacteria into sampling areas.

Fewer cows were observed on the property on the August 15 sampling than was previously documented, and all cattle were confined to pen area near site 4. Discussion with the landowner revealed that they were being moved to another property to be sold. Because the average water temperature was approximately 18ºC on this date, the decrease in fecal coliform and *E. coli* plate counts were likely due to the reduction in herd size, and a lack of recent precipitation. There was no *E. coli* plate count data for sites 3 and 4 on this sampling date, due to significant sample loss upon transport to laboratory.

On September 11, 2008, cattle confined to pen area near site 4 were fewer in number than observed on the August 15 sampling date. Water temperatures had dropped to an average of 16 ºC, and the decrease in water temperature, coupled with a lack of recent precipitation likely explains the decrease in fecal coliform counts as compared to the August 15 sampling date. Interestingly, site 3 continued to exceed the state *E. coli* standard, ruling out fecal coliform and *E. coli* contributions from sediment, as the sampling location had been moved to a high flow area. Hence, there are likely unknown fecal inputs to this small tributary that contributed to the higher than expected plate count values observed here. The variation in plate count values observed at each sampling site on September 11 was in line with what was expected, with the exception of site 3.

The last sampling event for the 2008 season occurred on September 28, and at this time less than ten cows were being reared in the pen area near site 4. Average water temperatures were 15.5 ºC and the plate count data for this date showed very similar values at each sampling location. This is likely due to a small rain event (~0.95 cm) that

began upon arrival to the site. Peak rain event plate count values averaged lower on September 28 than on the rain event sampling dates in July and on June 9.

Plate count data was also obtained for the 2009 sampling season. The first sampling of the 2009 season occurred on April 30, and three new sampling locations were chosen in 2009 upstream of the 2008 locations (Figure 3.1). This was in part due to a new property owner and loss of access to some of the old sites. The average water temperature on April 30 was 10.8 ºC, and cows were observed in the creek and in immediate upland area adjacent to site 5. Fecal material that had been recently excreted was found near site 7. Total fecal coliform and *E. coli* counts reflected these observations, as these two sites yielded the highest counts for both respective analyses. *E. coli* counts at both locations exceeded 126 CFU/100ml.

While fecal material was not observed near site 7 on May 12, approximately 60 head of cattle were counted upstream of site 9. Cattle were also present in upland areas adjacent to site 5, yet fewer were observed than on the April 30 sampling date. Water temperature averaged 11.0ºC, and fecal coliform and *E. coli* counts were highest at sites 9 and 5, both exceeded 126 CFU/100ml where cows were observed in the creek. Counts were lower at site 7, as it appeared cows had been vacant from the area for some time. Expectedly, counts at site 8 were low.

Water temperature had risen, averaging 13.6ºC, by the June 1 sampling date. Very little rain had fallen in the watershed, as water levels were identical for each of the first three sampling dates. Cattle were observed in, and adjacent to, the creek upstream of site 9, and were widely distributed upstream of site 4-2. Plate counts for both fecal coliforms

and *E. coli* were the highest observed at sites 7 and 4-2 to that date in the 2009 sampling season, and were likely elevated due to an increase in water temperature. While no cattle or fecal material was observed near site 7, grazing occurred in that general area, which may explain the higher counts observed on June 1. Cattle were absent from site 5, which may explain the slight drop in average CFU's observed here for *E. coli*.

On June 15 the water temperature averaged 13.1ºC, and the watershed had received 0.53 cm of rain 48 h prior to sampling. Also, 4.1 cm of rain fell on the watershed on June 8. This precipitation had residual affects on stream level, as it had risen ~ 4 cm since the May 12 sampling date. While cattle were not observed in proximity to any of the five sampling locations, cow fecal material was still abundant in areas upstream of site 9, so it is likely cows had been in that area recently. As a consequence of this, fecal coliform and *E. coli* counts were the highest at site 9 on June 15. While the increase in precipitation may have had a moderate affect on overall counts, as they averaged higher than May 12, counts were not considerably higher, likely due to minimal sediment re-suspension as determined by turbidity measurements.

Interestingly, by June 29 the water temperature average had fallen to 12.6ºC and site 8 continued to have low numbers of both fecal coliform and *E. coli*. Fecal coliform and *E. coli* counts at site 9 also decreased on this date, and this was likely due to the absence of cattle from this area and lack of fecal material observed in upland areas.

By the July 13 sampling date the watershed had received only trace amounts of precipitation in the weeks previous to sampling. Water temperature had risen on average to 15ºC and fecal material from cattle was observed adjacent to site 7, which likely

contributed to the high fecal coliform and *E. coli* counts observed. At site 9, however, counts continued to fall, as cattle remain absent from this area. There were few cows observed in the upland area near site 5, and counts increased there in comparison to the June 29 sampling date. In contrast, counts at site 4-2 decreased slightly, potentially due to that the lack of cattle at site 2.

The watershed received 3.8 cm of precipitation on July 25, and this was expected to increase counts of coliforms and *E. coli*. However, due to a lack of precipitation to this date, it was likely that the infiltration capacity of soil in the watershed was high. Thus, this precipitation had little affect on stream level and flow on July 29, as the water level was 8.9 cm lower than that measured on June 15. Plate counts for both fecal coliform and *E. coli* were, on average, lower than that observed on July 13. While the water temperature average remained relatively high, 16ºC, no cows were observed in areas adjacent to any of the sampling locations. Interestingly, *E. coli* counts at site 8, which is the spring source, exceeded 126 CFU/100ml. Fecal coliform and *E. coli* counts at the four remaining sites decreased slightly relative to other dates, likely due to the absence of cattle in adjacent upland areas.

Similarly, the area received 1.7 cm of precipitation on August 16. But unlike the other sampling dates, the August 17 sampling date was preceded by rain within a 24 h period. For the first time during the 2009 sampling season, stream turbidity measurements were observed below 60 cm, indicating significant sediment suspension in the water column. Water temperature averaged 16.3ºC, the highest in 2009. Cattle were observed near site 7 and upstream of site 5. Consistent with this, plate counts at sites 7, 4-

2, and 5 averaged the highest recorded during the 2009 season. Fecal coliform and *E. coli* counts at site 9 increased for the first time since June 15, likely due to increased sediment re-suspension, and counts were still relatively high at site 8.

The cattle were largely vacated from the field site by the September 2 sampling date. The watershed received no significant precipitation in the week prior to sampling, and water temperature averaged 14ºC. Accordingly, fecal coliform and *E. coli* plate counts were down significantly from the August 17 sampling date, however, counts were still significantly above the 126 CFU/100ml mark at all sites with the exception of site 8.

In the Little Jordan watershed, there are three landowners actively grazing cattle, and providing the animals with unrestricted access to the creek. At least one landowner employs no rotational grazing strategy, allowing significant degradation of riparian vegetation. Water temperature, recent rain events, and the proximity of cattle to the creek were major factors influencing the *E. coli* and fecal coliform plate count data observed. *E. coli* counts often exceeded the regulatory standard of 126 CFU per 100 ml in both 2008 and 2009. While water temperature did appear to influence levels of observed fecal coliform and *E. coli*, increased streamflow resulting in sediment re-suspension and the proximity of cows to the stream were likely the largest determinants of increased plate count numbers.

While unexpected at first, the relatively uniform plate count data observed during rain events is likely due to the re-suspension of sediment into the stream channel, which causes adherent populations of fecal coliform bacteria and *E. coli* to be liberated from sediments. The consistency at which site 3 showed elevated levels of both indicators was

surprising. Upland land use practices are unknown, but are likely contributing to elevated indicator bacteria levels here. On average, sites 4 and 5 yielded had the greatest numbers of fecal coliforms and *E. coli*. This was not surprising, as all inputs to the creek system have the potential to influence levels of indicator bacteria at these locations. Also, cattle were observed near both of these locations, with the greatest frequency throughout the sampling season. In the beginning of the sampling season, cows were largely distributed over the property. However, as grass became limited in various areas of the property, the landowner began to feed the cows in a rearing pen directly adjacent to site 4. As a result, counts began to decrease at sites 1 and 2, and remained consistently high at sites 4 and 5.

In 2009, three new sampling locations were selected to determine counts at locations upstream of the 2008 sites. Site 5 remained the same as sampled in 2008 and site 4-2 was slightly downstream of the 2008 site 4. However, spatial similarities existed in plate count data observations. Sites 7, 4-2 and 5 averaged the highest counts for both fecal coliforms and *E. coli*, as their exposure to cattle was consistently the highest although significantly lower than observed in 2008. Surprisingly, site 7 averaged higher counts than sites 4-2 and 5, likely due to higher levels of cattle exposure in this area. It was known that cattle were grazing in upland areas adjacent to site 7, but the extent of their presence was previously unknown. Low flow rate was also a factor also potentially contributing to the high counts observed at site 7. This site is upstream of the creeks main source of flow, and just below the spring where the creek begins. Low flow may allow bacteria to accumulate in the stream channel, causing more organisms to be obtained during sampling. Further, fecal material deposited in the creek or on its edge was likely to

remain for a longer period of time. Cows were grazing near site 9 early in the season (May through mid June), considerably affecting the levels of indicator bacteria observed there. As expected, once cattle were removed from this area, plate counts decreased at each sampling date at site 9, with the exception of the rain event sampling obtained on August 17. Increased sediment in the stream channel likely caused fecal coliform and *E. coli* counts to spike as a result of adherent organisms being liberated due to increased flow. Interestingly, site 8 exceeded the state water quality standard twice during the 2009 season. The cause for this is largely unknown. Lower than average plate counts per site in 2009 were likely due to a combination of factors. The new sampling locations, with the exception of site 7 had lower overall exposure to cattle than sites 1 and 2 sampled in 2008. Perhaps the largest contributor to lower overall averages in plate count numbers in 2009 was the sporadic nature of storm events and lack of sediment derived turbidity during sampling, coupled with BMP installment prior to and during the sampling season. While precipitation totals were similar between the two seasons (Table 4), rain events in 2009 were more spread out and difficult to predict. Consequently, only one sampling date was within a 24-hour period of a rain event as opposed to three during the 2008 season.

3.5.2 Bovine-specific (Bac3) qPCR.

In this study, I optimized a qPCR protocol for detecting the bovine-specific *Bacteroidales*-like 16SrRNA gene marker Bac 3 (Figure 3.4). As previously mentioned, numerical detection was defined as the circumstance when ≥ 2 of 3 copy number averages from each DNA sample were above limit of detection, and the subsequent

average copy number of those replicates was above limit of detection. The limit of detection for this assay was 100 copies, that is, this assay had the ability to accurately quantify the Bac3 DNA target down to 100 copies. However, due to the potential of *Bacteroidales* to carry multiple genomes based on growth stage various environmental stressors, copy number is not analogous to cell number. Henceforth, we can not use the qPCR assay developed in this study to determine bovine specific *Bacteroidales* cell number. However, it is reasonable to infer that higher observed copy number translates to greater presence of bovine derived *Bacteroidales,* lending to the conclusion that bovine fecal loading is more substantial. The Bac3 qPCR assay identified the bovine-specific *Bacteroidales* target on every sampling date during the 2008 season (Figure 3.5). This indicates that cattle manure was likely impacting all sites in this study and were contributing to the total fecal loading of the LJC. Interestingly, the distribution of the bovine-specific marker gene varied by sample at each location. This is likely due to the large variability inherent in sampling stream systems, coupled with the sensitivity of the Bac3 qPCR assay.

It appears there was not a strong correlation between numbers of fecal coliform or *E. coli* and the Bac3 assay output on rain event sampling dates. For example, while the highest plate count numbers observed were during the July 31, 2008 sampling date, during a rain event (~2.5 cm upon arrival), this date had very low qPCR copy numbers per site, and were below the limit of detection at sites 1, 3, and 4. Two major factors likely contribute to this phenomenon. First, during the rain event, flow in the stream increases, increasing volume, effectively diluting the amount of Bac3 target present in

water samples. This is in contrast to what findings of a previous study may suggest, as the detection of the Bac3 marker was reported in stream sediments (Lee et al., 2008). Further, Bower and coworkers (2005) found elevated levels of bovine-specific *Bacteroides* markers in a Lake Michigan harbor following a rain event. Secondly, due to increased flow there is a significant increase in turbidity, or sediment re-suspension. Because populations of fecal coliform bacteria and *E. coli* are bound to sediments and released due to increased turbidity, plate count numbers spike to very high levels. There was a general correlation however, between the highest average CFU and Bac3 copy numbers observed at sites 4 and 5 during the majority of sampling dates (see discussion on site 2 below).

Additionally, because *Bacteroides* does not propagate in the environment, direct or adjacent loading to the stream are likely the major influences on copy number values. For example, the highest copy number observed was at site 2 on the May 12 sampling date. During this time, several cows were standing in the water, and fecal material was abundant on the stream bank, and was observed floating in the river. Total copy numbers generally decreased as cows were being moved from the property late in the sampling season. While the number of Bac3 copies averaged the highest at site 2 over the 2008 sampling season, it was primarily due to the May 12 sampling date alone. Aside from this, however, sites 4 and 5 generally had a higher copy number of the Bac3 marker gene over the 2008 sampling season than all other sites.

During the 2009 season the he Bac3 qPCR assay did not identify the bovine-specific *Bacteroides* marker gene t until May 12 at site 9 (Figure 3.5). Similarly to 2008,

the distribution of the marker across sites and dates was not always consistent. In general, copy number values were lower in 2009 than 2008. This could be due relatively low flow conditions that reduced the distribution of the marker in the stream channel. More importantly, however, fewer cows were present in the watershed in 2009 compared to 2008 (sites 1-3). As anticipated, the copy number of the Bac3 marker gene at site 8, over the sampling season, were the lowest recorded in 2009. This location was 5 m s from the spring seepage. The relatively high numbers observed here on August 17 and September 2 were likely due to the presence of cattle in this area before sampling was conducted. Site 8 is approximately 100 m downstream of site 7, where cattle or fecal material had been observed on several occasions. It is likely that the season average at site 7 was so high as a result of this deposition. Interestingly, the high copy numbers of the Bac3 marker gene observed at site 5 were the second lowest during the 2009 sampling season, perhaps in part due to initiation of a grazing best management practice (BMP) that reduced creek exposure to cattle at sites 9, 4-2 and 5.

3.5.3 Installment of land use best management practice (BMP).

Prior to the 2009 sampling season, a best management practice was designed with cooperation of the landowner of upland areas surrounding site 4-2 and 5. The process began with removal of manure from direct upland areas of the creek. This landowner also has corn plots within the LJC watershed, so manure application was addressed and conducted according to proper agronomic rates for each field. Further, a feedlot “fix” implementation plan was initiated to relieve the creek of direct fecal loading by building

a concrete work area where cattle could receive water and feed. To do so, a waterline was constructed to bring water to the concrete work area where cattle were gathered. Further, feedlot buffer areas were established, and it was ensured that a clean water diversion way was determined for springs that had originally carried over the area. Finally, a grazing plan was designed to rehabilitate pasture areas that had been grazed significantly in prior years. A waterline was extended to a 300 gallon tank to supply the cattle with water away from the creek, and out of the feedlot area. Further, pasture areas were seeded with mixture of red clover 2.5 lbs/acre, white clover 0.5 lbs/acre, orchard grass 1 lb/acre, tall fescue 5 lbs/acre, and timothy 0.5 lbs/acre.

Plate count data and Bac3 qPCR copy number detection in 2009 was on average lower than in 2008. While the *E. coli* levels in 2009 were lower than observed in 2008, it is difficult to surmise the BMP's role, as fewer rain events were sampled in 2009. However, site number 7 averaged the highest amount of total *E. coli* in 2009, followed by sites 4-2 and 5. The BMP did not include the area grazed adjacent to site number 7, perhaps explaining why averages were higher here. Based on visual estimates, there were similar numbers of cattle grazing this area to those grazing near sites 4-2 and 5. While averages were still relatively high at sites 4-2 and 5, they were considerably lower than those observed in 2008. For plate count data, this was likely due to fewer significant rain events sampled in 2009, coupled with BMP installment. For Bac3 qPCR copy number output, it is possible the initial BMP implementation contributed to the reduction observed at site 5 in 2009.

3.5.4 Summary of bovine fecal loading to the Little Jordan Creek.

In this study I assessed the spatial and temporal dynamics of bovine fecal loading to a creek system in Southeastern Minnesota. Ample evidence existed that the grazing operations in the Little Jordan Creek watershed were negatively affecting microbiological quality of the creek water, as determined by *E. coli* plate count analysis. *E. coli* numbers often exceeded the Minnesota State Standard of 126 CFU/100ml at several sampling locations spanning the entire creek. The Bac3 qPCR assay, optimized in this study, revealed abundant presence of the bovine-specific *Bacteroidales* gene marker, confirming fecal loading from grazing cattle. Recent rain events and cow proximity to creek sampling location appeared to be the two most significant factors influencing plate counts and observed Bac3 copy number. These results corroborate well with previous findings that bovine fecal loading to aquatic systems decreases as cattle proximity to the water body increases (Larsen, 1996; Bailey and Welling, 1999; Tate et al., 2003). It is likely that during significant rain events, naturalized *E. coli* present in creek sediments are liberated due to increased flow, and observed CFU counts raise accordingly. As a result, USEPA method 1603 may not accurately characterize the scale of fecal loading to stream systems during significant rain events. Further, rain events may skew long-term *E. coli* count data used to determine impairments.

Observed CFU *E. coli* did not correlate well with detection of the bovine marker gene, with the exception of sampling locations that had direct cow exposure. The copy number of the bovine marker gene decreased considerably in 2009 at the sampling location exiting the property, suggesting that the BMP achieved lower amounts of fecal

loading to the system at this location, and would likely be successful on a watershed scale. While the Bac3 qPCR assay effectively identified the scale of the bovine-specific *Bacteroidales* marker gene, work is still needed in accurately characterizing the numerical genome possession of *Bacteroidales* so cell numbers can be determined and scale of bovine fecal loading assessed. Further, fractional characterization of *Bacteroidales* in bovine intestinal flora will be necessary to normalize observed cell numbers in water samples with amount of feces loaded to waterways. Future studies will also be required to establish epidemiological risk relationships to gene targets derived from fecal anaerobe targets.

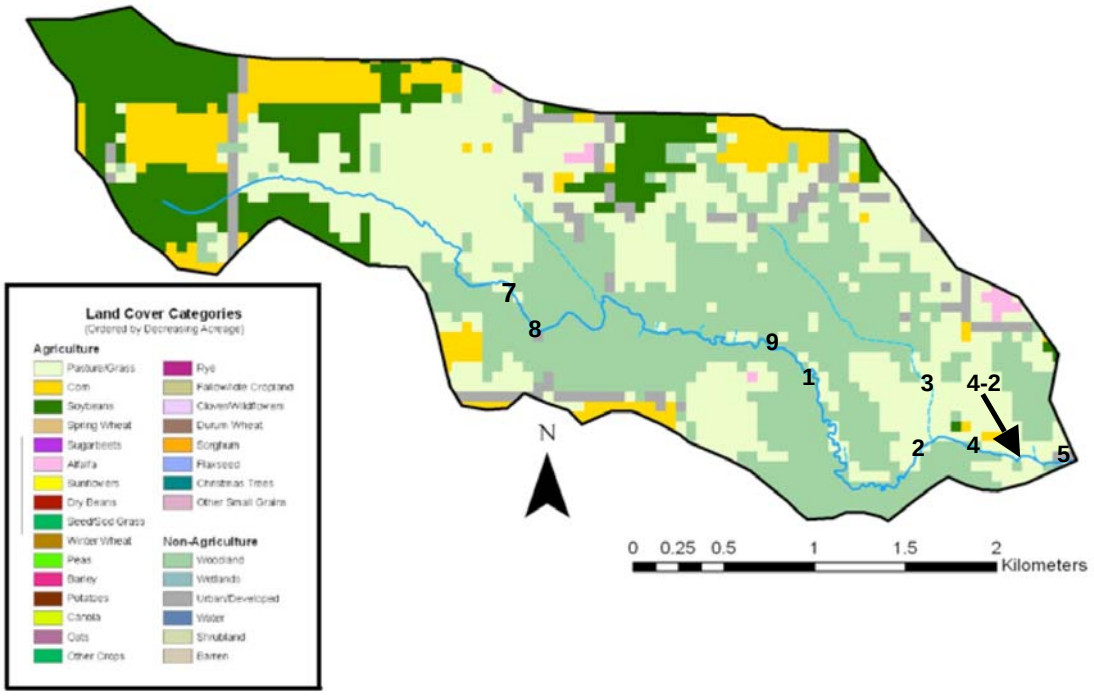
While the BMP installed prior to the 2009 sampling season appeared to have desirable affects, cows were on occasion still granted access to the stream corridor. Future mitigation efforts must include all landowners within the watershed, and enforcement mechanisms must be administered to ensure compliance of BMP design.

3.6. Acknowledgments

I would like to thank Dan Norat, Matthew Hamilton, and Jessica Eichmiller for their extensive consultation, and guidance throughout the course of this study, and for assistance with sample acquisition and processing. I would also like to thank Dr. Adam Birr of the Minnesota Department of Agriculture for his extensive work and dedication to this project.

3.7. Tables and Figures

Figure 3.1



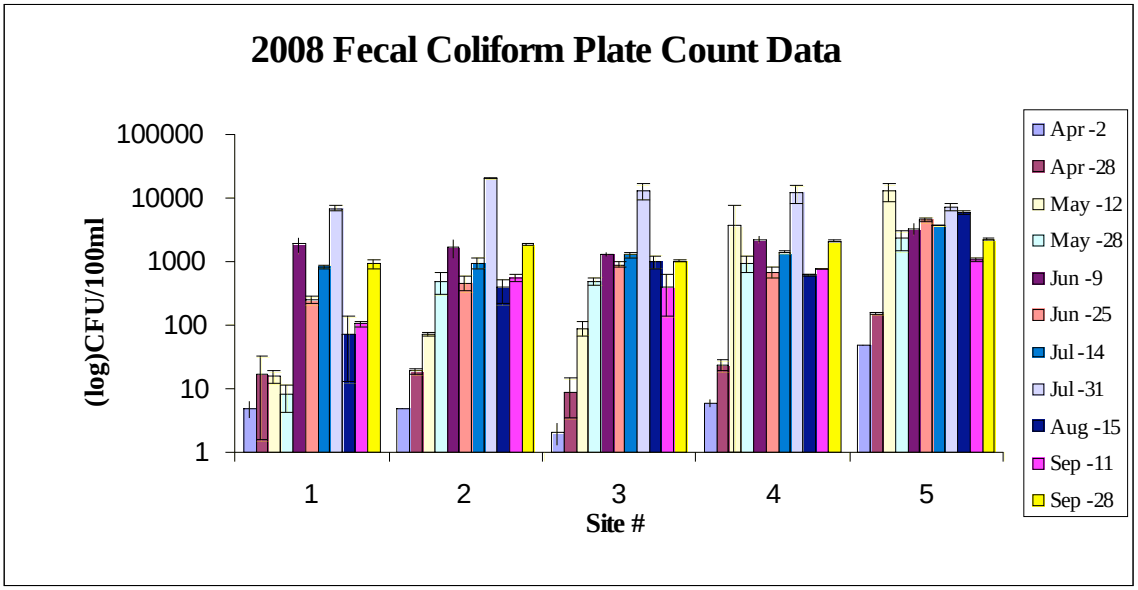
Land-use map of the Little Jordan Creek watershed. Sampling locations are indicated by number. In 2008, sites 1 thorugh 5 were sampled. In 2009, sites 7, 8, 9, 4-2, and 5 were sampled

Table 3.1 Conventional and real-time PCR primers used in this study.

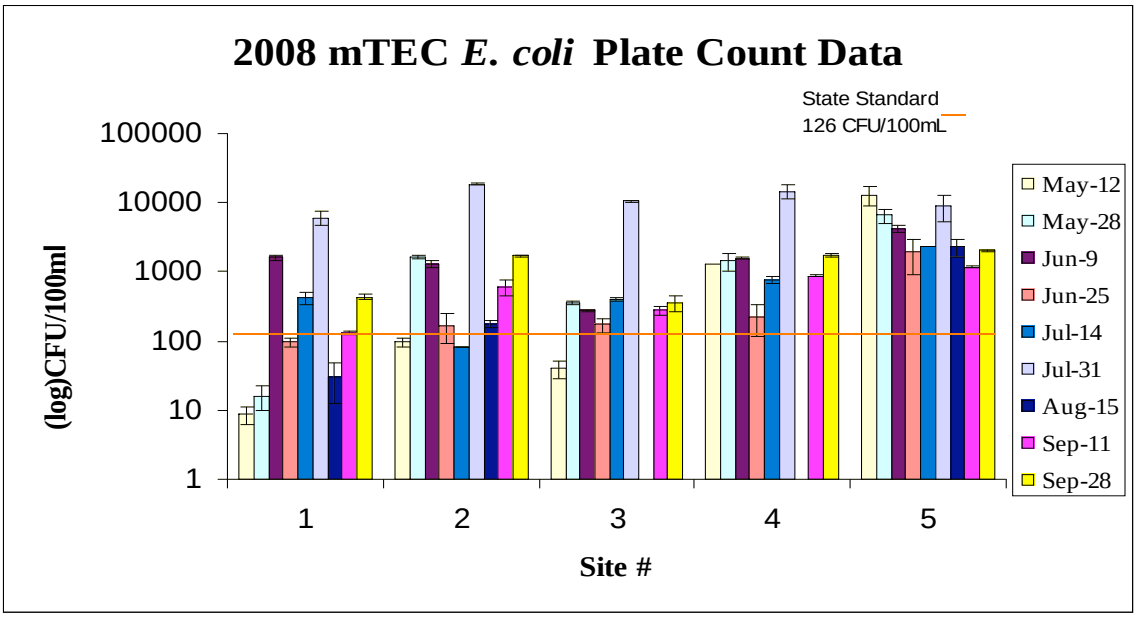
Primer Set	Host	Sequence (5'-3')	Amplicon Lengh (bp)	Annealing Temp. (ºC)	Cycles
AllBac296F AllBac412R	All Bacteroides	5'-GAGAGGAAGGTCCCCCAC-3' 5'-CGCTACTTGGCTGGTTCAG-3'	106	60	35
HF183F Bac708R	Humans	5'-ATCATGAGTTCACATGTCCG-3' 5'-CAATCGGAGTTCTTCGTG-3'	525	63	35
Bac3F Bac3R	Cattle	5'-CTAATGGAAAATGGATGGTATCT-3' 5'-GCCGCCCAGCTCAAATAG-3'	166	61	40

Figure 3.2

A.



B.



Fecal coliform (A) and *E.* coli (B) plate count data from the 2008 sampling season. Significant rain events had occurred within a 72 hour period previous to sampling on June 9th, July 14th, and July 31st. No *E. coli* counts were performed for site 3 and 4 on August 15th, due to significant sample loss upon transportation to laboratory.

Table 3.2 Pair-wise comparison of 2008 *E. coli* plate count data by site on each date using ANOVA and Fisher's LSD. Sites across each sampling date followed by the same letter are not statistically different as determined by ANOVA and Fisher's LSD at $\alpha = 0.05$. No letters are associated with sites 3 and 4 on August 15th due to sample loss (reads across).

	Sites				
Sample Date	1	2	3	4	5
May 12	A	A	A	A	B
May 28	A	A	A	A	B
June 9	A	A	B	A	C
June 25	A	AB	AB	B	C
July 14	A	B	A	C	D
July 31	A	B	C	C	A
August 15	A	A	N/A	N/A	B
September 11	A	B	C	D	E
September 28	A	B	A	B	C

Table 3.3 Pair-wise comparison by date of 2008 *E. coli* plate count data at each site using ANOVA output to determine Fisher's LSD. Dates across each sampling site that share the same letter were not significantly different at $\alpha = 0.05$ (reads across).

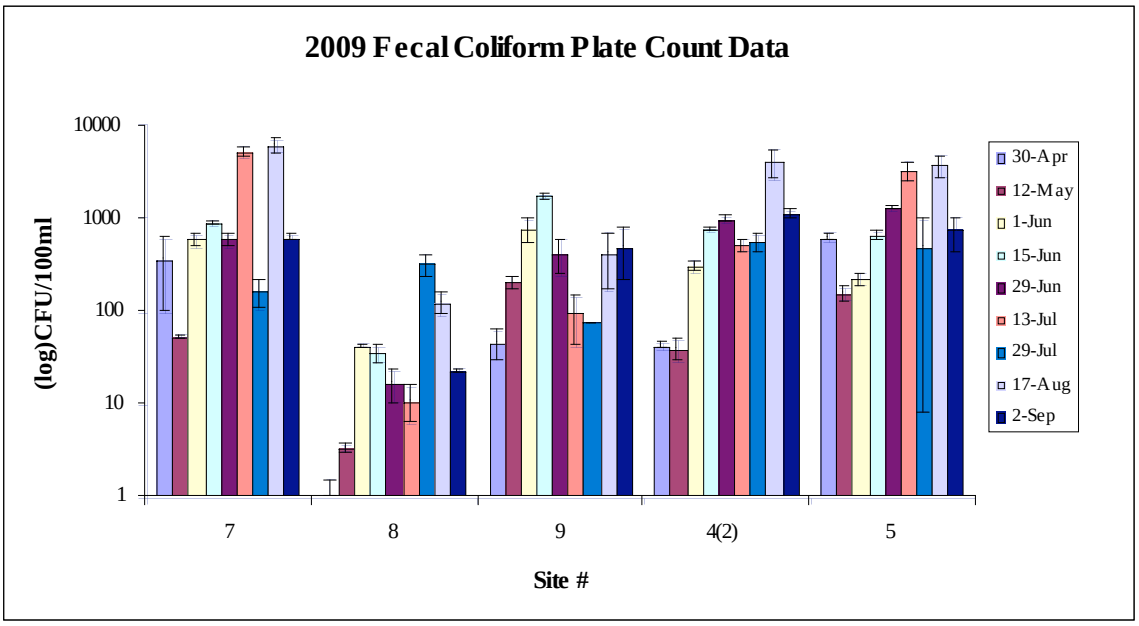
	Sampling Date								
Sites	May 12	May 28	June 9	June 25	July 14	July 31	August 15	September 11	September 28
Site #1	A	A	B	AB	AB	C	AB	AB	AB
Site #2	A	B	B	A	A	C	A	B	D
Site #3	A	A	A	A	A	B	N/A`	A	A
Site #4	A	A	A	A	A	B	N/A`	A	A
Site #5	A	B	B	B	B	C	B	B	B

Table 3.4 Precipitation comparisons by month for the 2008 and 2009 sampling seasons.

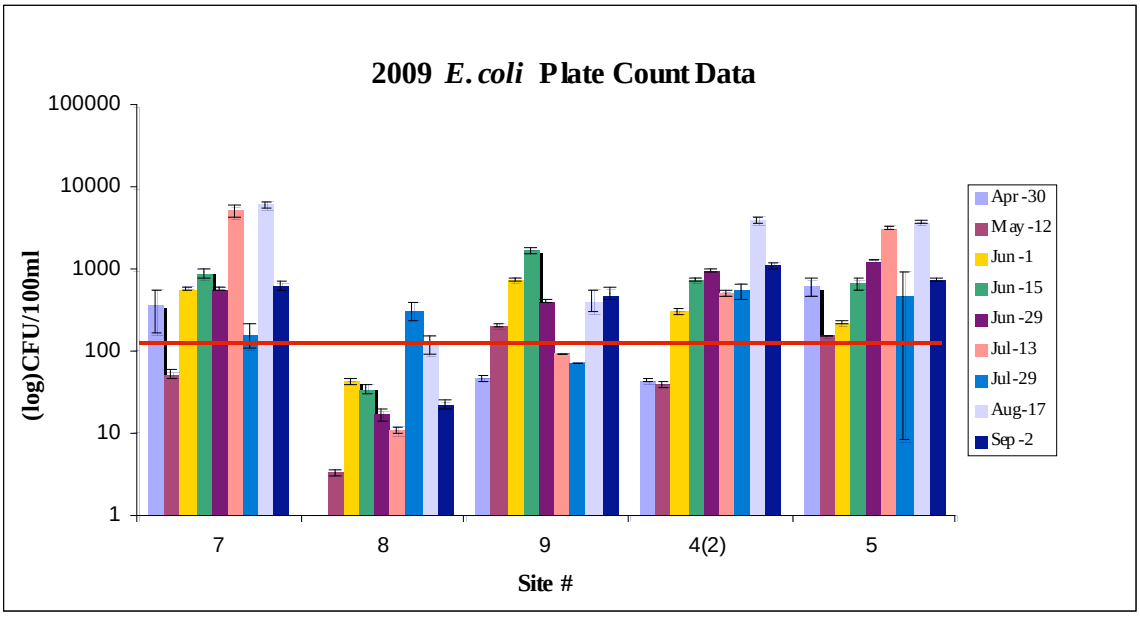
	2008	2009
Month	Precip(cm)	Precip(cm)
April	16.8	7.2
May	10.1	13.2
June	15.9	9.3
July	7.1	9.4
August	7.3	12.4
September	3.2	3.8
Total	60.4	55.3

Figure 3.3

A.



B.



Fecal coliform (A) and *E.* coli (B) plate count data from the 2009 sampling season. August 17th represents the only sampling date when precipitation had fallen in the watershed within a 24-h period prior to arrival.

Table 3.5 Pair-wise comparison of 2009 *E. coli* plate count data by site on each date using ANOVA and Fisher's LSD. Sites across each sampling date followed by the same letter are not statistically different as determined by ANOVA and Fisher's LSD at α = 0.05 (reads across)

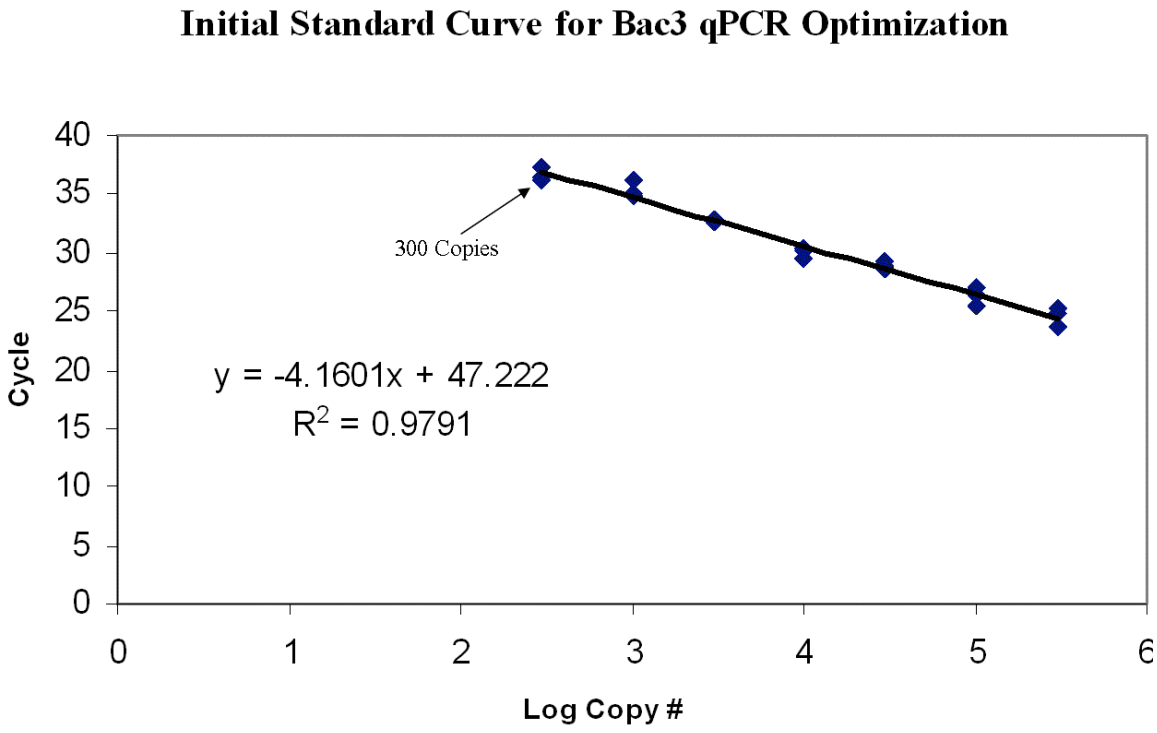
	Sites				
Date	7	8	9	4(2)	5
April 30	A	B	B	B	C
May 12	A	B	C	D	E
June 1	A	B	C	D	E
June 15	A	B	C	AD	D
June 29	A	B	C	D	E
July 13	A	B	B	B	C
July 29	A	B	C	D	E
August 17	A	B	B	C	C
September 2	D	B	AD	C	D

Table 3.6 Pair-wise comparison by date of 2009 *E. coli* plate count data at each site by use of ANOVA output to determine Fisher's LSD. Dates across each sampling site that share the same letter were not significantly different at α = 0.05 (reads across)

	Dates								
Site	April 30	May 12	June 1	June 15	June 29	July 13	July 29	August 17	September 2
7	AB	A	AB	B	AB	C	A	D	AB
8	A	A	B	BC	C	A	D	E	C
9	A	B	C	D	E	A	A	E	E
4(2)	A	A	B	C	D	B	BC	E	F
5	AE	B	B	A	C	D	E	F	A

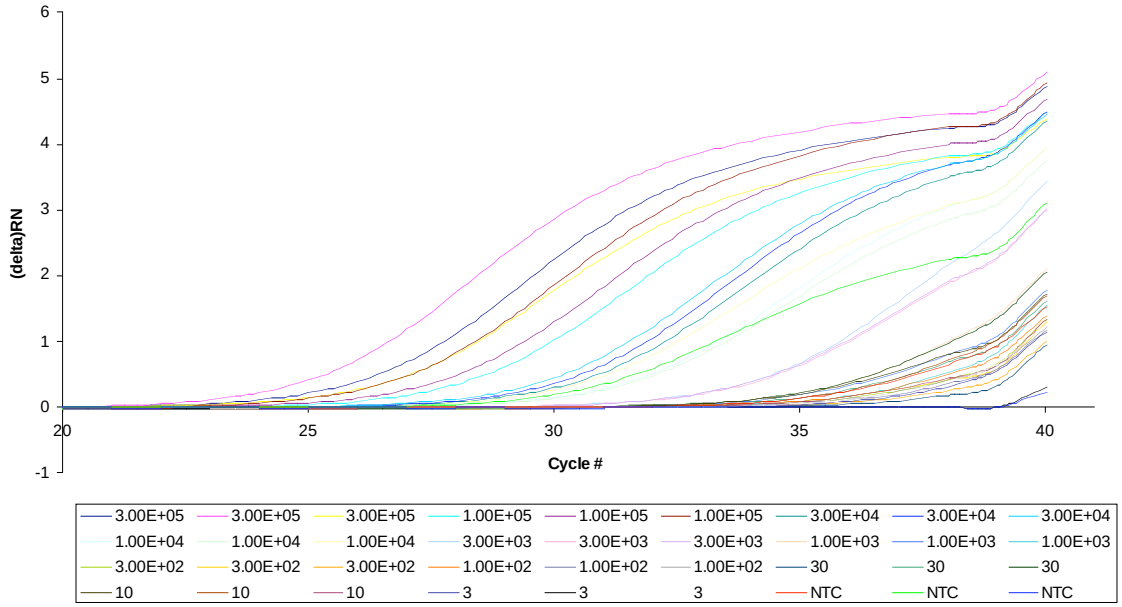
Figure 3.4 Standard curve, amplification plot, and melting curve for the initial optimization run (A) and final optimization run (B)[1]. The initial optimization run (A) was conducted using conventional thermalcycling conditions as previously described, and 200nM of each primer (Shanks, 2006). The final optimization run (B) was conducted using the following thermalcyling conditions: 2 min at 50˚C followed by a 10 min cycle at 95˚C for enzyme activation and denaturation and then 40 cycles of a denaturing step at 95˚C for 15 seconds, combined with an annealing and primer extension step at 61˚C for 1 min. 300nM of each primer was used in the optimized reaction mixture. Dissociation was measured as described above.

A.

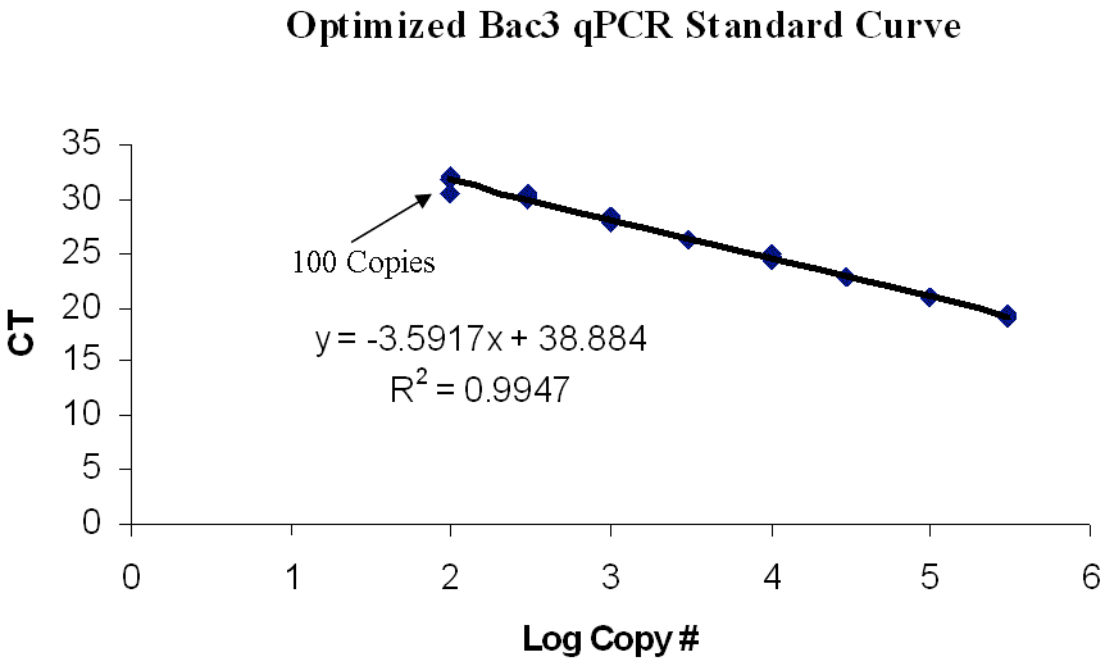


[1] The initial optimization run did not include a dissociation experiment. A melting curve is shown for the final optimization run.

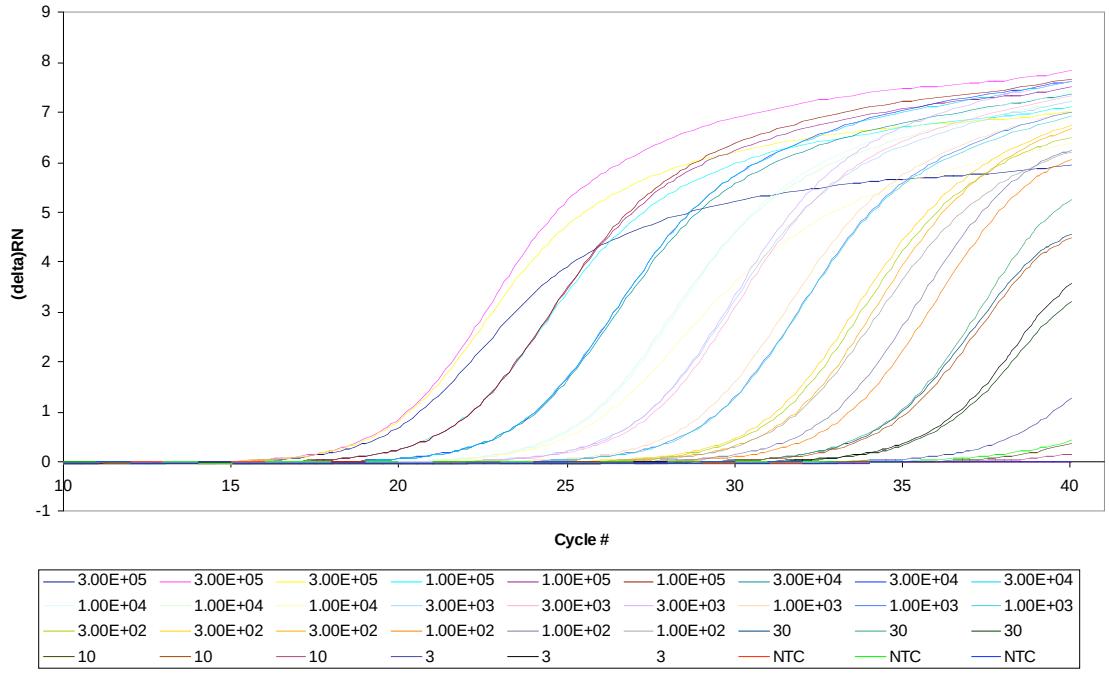
Initial Amplification Plot for Bac3 qPCR Optimization



B.



Optimized Bac3 qPCR Amplification Plot



Optimized Bac3 qPCR Melting Curve

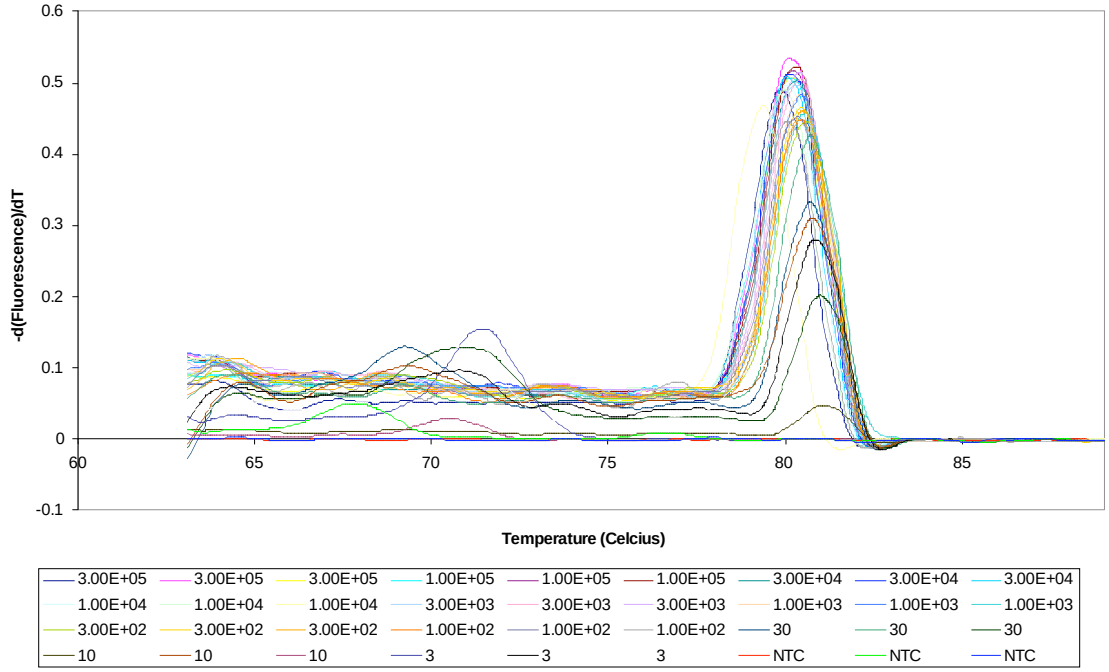
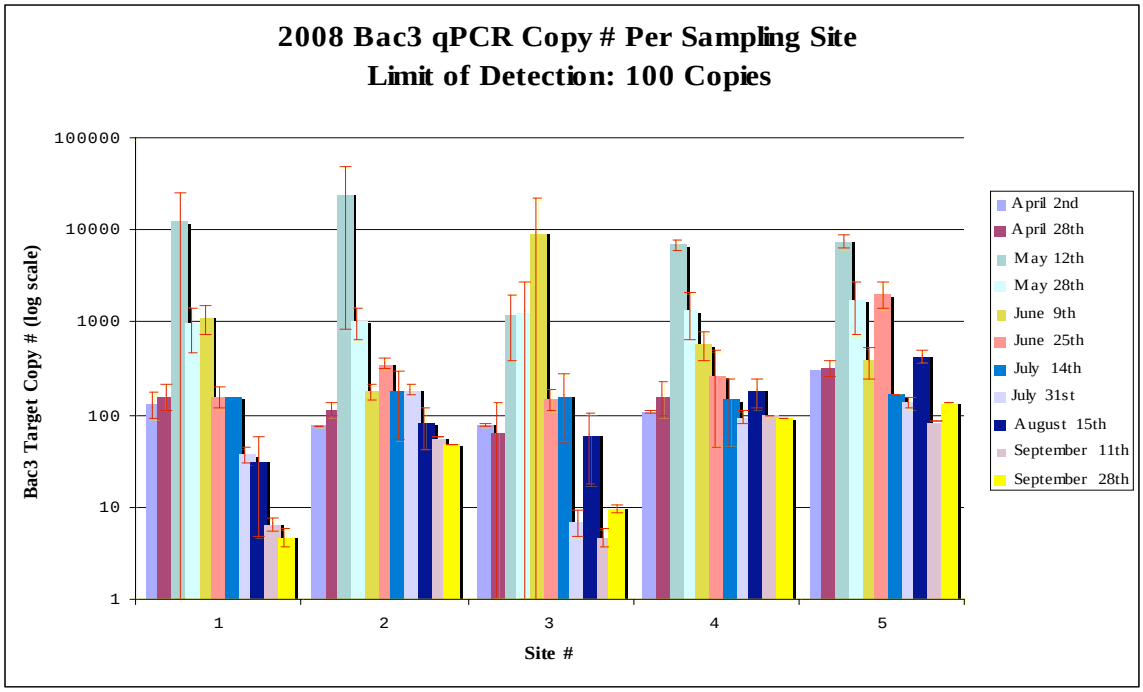
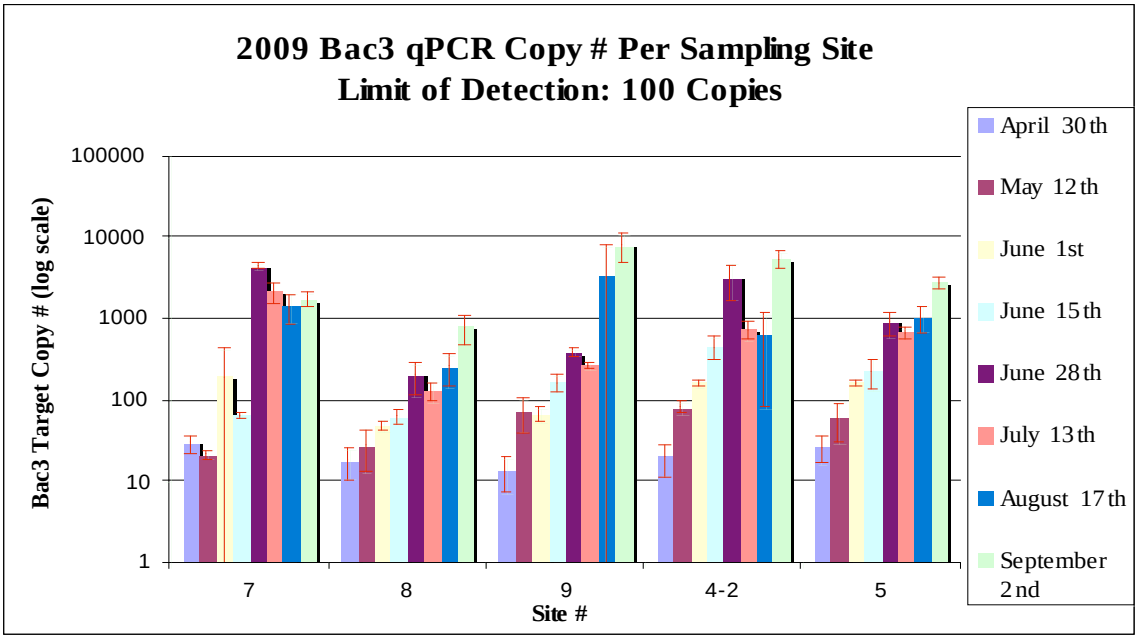


Figure 3.5

A.



B.



Total Bac3 qPCR target copy number at each sampling site during the 2008 (A) and 2009 (B) sampling seasons. The limit of detection was determined to be 100 copies.

3.8 References

Bailey, D.W., and G. R. Welling. 1999. Modification of cattle grazing distribution with dehydrated molasses supplement. J. Range Manage. 52:575-582.

Bernhard, A. E., and K. G. Field. 2000(a). Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66:1587–1594.

Bernard, A.E., and Field, K.G. 2000(b). 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, v. 66, no. 10, p. 4571-4574.

Bower, Patricia A., Scopel, Caitlin O.; Jensen, Erika T.; Depas, Morgan M.; McLellan, Sandra L. 2005. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. Appl. Environ. Microbiol. 71: 8305-8313

Byappanahalli, M., Fowler, M.; Shively, D.; Whitman, R.L. 2003. Ubiquity and persistence of *Escherichia coli* in a Midwestern coastal stream. Appl. Environ. Microbiol. 69 (8), 4549–4555.

Byappanahalli, M.; Whitman, R.L.; Shively, D.A.; Ting, W.T.; Tseng, C.C.; Nevers, M.B. 2006. Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. J. Water Health 4 (3), 313–320.

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

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

Gordon, D. M. 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology 147:1079–1085.

Hartel, P. G.; Summer, J. D.; Hill, J. L.; Collins, J.; Entry, J. A.; Segars, W.I. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. J. Environ. Qual. 31:1273–1278.

Hartel, P. G.; Summer, J. D.; Segars, W.I. 2003. Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking. Water Res. 37:3263–3268.

Hassan, W.M.; Wang, S.Y.; Ellender, R.D. 2005. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Appl. Environ. Microbiol. 71 (1), 512–518.

Ishii, S.; Ksoll, W.B.; Hicks, R.E.; Sadowsky, M.J. 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), 612–621.

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. Appl. Environ. Microbiol. 72(6):4214-4224.

Larsen, R.E. 1996. Survival and transport of fecal pathogens in grazed watersheds, p.59-71. In: M.R. George (ed.) Livestock management in grazed watersheds. Univ. Calif. Div. Agric. and Nat. Resourc. Pub. 3381.

Lee, Yong-Jin; Molina, Marirosa; Santo Domingo, Jorge W.; Willis, Jonathan D.; Cyterski, Michael; Endale, Dinku M.; Shanks, Orin C. 2008. Temporal assessment of the impact of exposure to cow feces in two watersheds by multiple host-specific PCR assays. Appl. Environ. Microbiol. 74: 6839-6847

Minnesota Pollution Control Agency. 2005. Using Transparency Tube and total suspended solids data to assess stream turbidity. wq-csm1-01. http://www.pca.state.mn.us/publications/wq-csm1-01.pdf.

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

Olapade, O.A.; Depas, M.M.; Jensen, E.T.; McLellan, S.L. 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Appl. Environ. Microbiol. 72 (3), 1932–1938.

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

Sambrook, Fritsch, and Maniatis (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, volume 3, appendix B.12

Scott, T. M.; Parveen, S.; Portier, K. M.; Rose, J. B.; Tamplin, M. L.; Farrah, S. R.; Koo, A.; Lukasik, J. 2003. Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. Appl. Environ. Microbiol. 69:1089–1092.

Shanks, O.C.; Santo Domingo, J.W.; Lamendella, R.; Kelty, C.A.; Graham, J.E. 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72 (6), 4054–4060.

Shanks, O. C.; Atikovic, Emina; Blackwood, A. Denene; Lu, Jingrang; Noble, Rachel T.; Domingo, Jorge Santo; Seifring, Shawn; Sivaganesan, Mano; Haugland, Richard A. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 74: 745-752

Stoeckel, D.M.; Mathes, M.V.; Hyer, K.E.; Hagedorn, C.; Kator, H.; Lukasik, J.; O'Brien, T.L.; Fenger, T.W.; Samadpour, M.; Strickler, K.M.; Wiggins, B.A. 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. 38 (22), 6109–6117.

Tate, K.W.; Atwill, E.R.; McDougald, N.K.; George, M.R. 2003. Spatial and Temporal Patterns of Cattle Feces Deposition on Rangeland. J. Range Manage. Vol. 56, No. 5 pp. 432-438

U.S. Environmental Protection Agency. 2002. Method 1603—*Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar. EPA821-R-02-23, 9. U. S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2005. Microbial Source Tracking Guide Document. EPA-600/R-05/064, 131pp. Office of Research and Development, Washington, DC.

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. Appl. Environ. Microbiol. 69, 4714–4719.

Whitman, R.L.; Byers, S.E.; Shively, D.A.; Ferguson, D.M.; Byappanahalli, M. 2005. Occurrence and growth characteristics of *Escherichia coli* and enterococci within the accumulated fluid of the northern pitcher plant (Sarracenia purpurea L.). Can. J. Microbiol. 51 (12), 1027–1037.

Wiggins, B.A.; Cash, P.W.; Creamer, W.S.; Dart, S.E.; Garcia, P.P.; Gerecke, T.M.; Han, J.; Henry, B.L.; Hoover, K.B.; Johnson, E.L.; Jones, K.C.; McCarthy, J.G.; McDonough, J.A.; Mercer, S.A.; Noto, M.J.; Park, H.; Phillips, M.S.; Purner, S.M.; Smith, B.M.; Stevens, E.N.; Varner, A.K. 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. 69, 3399–3405.

Seasonal stability of *Cladophora*-associated *Salmonella* in Lake Michigan watersheds.

*Byappanahalli, M.; Sawdey, R.C.; Ishii, S.; Shively, D.A.; Ferguson, J.A.; Whitman, R.L.; Sadowsky, M.J. 2009. Seasonal stability of Cladophora-associated Salmonella in Lake Michigan watersheds. Water Res. 43(3):806-14.

**This project was conducted in collaboration with the United States Geological Survey at the Lake Michigan Ecological Research Station located in Porter, Indiana.

Chapter 4. Seasonal stability of *Cladophora*-associated *Salmonella* in Lake Michigan watersheds.

4.1. Overview

The bacterial pathogens *Shigella, Salmonella, Campylobacter*, and shiga toxin-producing *E. coli* (STEC) were recently found to be associated with *Cladophora* growing in southern Lake Michigan. Preliminary results indicated that the *Salmonella* strains associated with *Cladophora* were genetically identical to each other. However, because of the small sample size (n=37 isolates), and a lack of information on spatial-temporal relationships, the nature of the association between *Cladophora* and *Salmonella* remained speculative. In this study, I investigated the population structure and genetic relatedness of a large number of *Cladophora*-borne *Salmonella* isolates from Lake Michigan (n=133), as well as those isolated from stream and lake water (n=31), aquatic plants (n=8), and beach sands and sediments (n=8) from adjacent watersheds. *Salmonella* isolates were collected during 2005-2007 between May and August from Lake Michigan beachsheds in Wisconsin, Illinois, and Indiana. The genetic relatedness of *Salmonella* isolates was examined by using the horizontal, fluorophore-enhanced rep-PCR DNA fingerprinting technique. While the *Salmonella* isolates associated with *Cladophora* exhibited a high degree of genetic relatedness ($\geq$ 92% similarity), the isolates were not all genetically identical. Spatial and temporal relationships were evident in the populations examined, with tight clustering of the isolates both by year and location. These findings suggest that the relationship between *Salmonella* and *Cladophora* is likely casual and is

related to input sources (e.g. wastewater, runoff, birds) and the predominant *Salmonella* genotype surviving in the environment during a given season. Our studies also indicate that *Cladophora* is likely an important reservoir for *Salmonella* and other enteric bacterial pathogens in Lake Michigan beachsheds, which in turn may influence nearshore water quality.

4.2. Introduction

Cladophora (Cladophoraceae) are filamentous green alga found in fresh and marine waters. The accumulation of *Cladophora* along the shorelines of the Great Lakes has become a serious beach management problem, affecting the aesthetic and recreational experiences, and potential health of visitors. Whitman et al. (2003) reported high densities of the fecal indicator bacteria *Escherichia coli* (E. coli) and enterococci in *Cladophora* mats, with counts often exceeding 1×10^5 colony-forming units (CFU) per g algal tissue. Preliminary studies suggested that such high bacterial densities are likely attributed to *in situ* growth under certain environmental conditions (Byappanahalli et al., 2003(b)). In addition to *E. coli*, sulfate-reducing bacteria and other phylogenetically diverse groups of bacteria have also been detected in *Cladophora* mats (Olapade et al., 2006).

Recently, Ishii et al. (2006(b)) reported that enteric bacterial pathogens, including Shiga toxin-producing *E. coli* (STEC), *Shigella*, *Salmonella,* and *Campylobacter* can also be recovered from rock-attached *Cladophora* in Lake Michigan. DNA fingerprint analyses of a small number of strains (n = 37), done using the horizontal, fluorophore-

enhanced rep-PCR (HFERP) DNA fingerprinting technique (Johnson et al., 2004; Ishii et al., 2006(a)) and the BOX A1R primer (Versalovic et al., 1991), showed that almost all Salmonella strains recovered from Cladophora consisted of a single genotype (≥95% similarity). Serotype analyses of two representative strains indicated that the isolated bacteria belonged to *Salmonella enterica* subsp. *enterica* serovar Newport, suggesting that the *Salmonella* strains associating with *Cladophora* were likely human pathogens. It has also been reported that genetically diverse *S. enterica* subsp. *enterica* serovar Typhimurium strains are present in southern Lake Michigan beach water (Whitman et al., 2001), suggesting that either *Cladophora* acquires specific genotypes of this pathogen directly from the environment or that only a limited number of Salmonella genotypes have the ability to survive in lake water. In contrast, Ishii et al. (2006(b)) reported that *Cladophor*a-borne *Campylobacter* strains were genotypically diverse, with genetic similarity values ranging from 2 to 98%, suggesting that lake water contains highly diverse populations of *Campylobacter* and that *Cladophora* does not display any preference for a given bacterial genotype.

While our initial results showed that *Cladophora* is a reservoir for *Salmonella* and other potential human pathogens in Lake Michigan, the data were not robust enough to make definitive conclusions concerning genotypic diversity of the *Salmonella* for two reasons: (a) only a small sample size was used (n = 37 isolates), and (b) there was a lack of information about the geographical, spatial, and temporal factors that might influence the relationship between *Cladophora* and *Salmonella*.

To overcome these limitations, in the study presented here I investigated the temporal and spatial variation in population structure of a large number of *Salmonella* isolates obtained from *Cladophora* (n = 180) found in stream and lake water, from aquatic plants, and from beach sand and sediments. The objectives of this study were to determine whether populations of *Cladophora*-associated *Salmonella* varied with time and location, and whether there was a relationship between *Salmonella* populations associated with *Cladophora* and those obtained from several potential input sources (e.g. sand, water) where these bacteria have been previously recovered. My underlying hypothesis was that the prevailing *Salmonella* genotypes present in lake water and sediments become associated with *Cladophora* and that these genotypes may vary each year due to changing input sources. To test this hypothesis, I compared the DNA fingerprint patterns of *Salmonella* isolated from *Cladophora* and other sources from several Lake Michigan watersheds in 2006 and 2007 to each other, and to the DNA fingerprint patterns of isolates obtained during the 2005 season (Ishii et al., 2006(b)).

4.3. Materials and Methods

4.3.1 Study location and sampling strategy.

Cladophora-associated *Salmonella* strains isolated in 2005 were obtained in August from the lakeside of the Ogden Dunes breakwater at the Indiana Dunes National Lakeshore in northwest Indiana, as previously described (Ishii et al., 2006(b)). In 2006, *Cladophora* and water samples were collected between June and August from several Lake Michigan beaches where *Cladophora* accumulations are common, including

Wisconsin (the Door Peninsula and Washington Island, Door County, and North Beach and Wind Point Beach, Racine), Illinois (63rd Street Beach, Chicago), Michigan (Sleeping Bear Dunes National Lakeshore, Empire), and Indiana (the lake- and ditch-sides of the breakwater near Ogden Dunes beach, west branch of the Little Calumet River near Portage Marina, and Dunes Creek near Indiana Dunes State Park, Porter) (Figure 4.1, Table 4.1).

In addition to water and *Cladophora*, other potential sources of *Salmonella* were also examined, including beach sand (n = 3) and sediment (n = 5) from the 63rd St. Beach (June 2006), and an aquatic plant, *Potamogeton* sp. (*Stukenia* sp.) (n = 8), from North Beach (August 2006). The characteristics of 63rd St. Beach and the Wisconsin locations are described in greater detail elsewhere (Whitman and Nevers, 2003; Ishii et al., 2006(b); Kleinheinz et al., 2006).

In 2007, intensive sampling in August focused on selected watersheds in northwest Indiana; Dunes Creek, the Little Calumet River (both east and west branches), and Salt Creek, where *Salmonella* was frequently detected (Fig. 4.1, Table 4.1). Characteristics of these coastal watersheds are described in detail elsewhere (Whitman, 1983; Byappanahalli et al., 2003(a); Ishii et al., 2006(b); Byappanahalli et al., 2007). In brief, the Little Calumet River (LCR) feeds into Lake Michigan via Burns Ditch at Ogden Dunes, Indiana. Salt Creek drains into the east branch of the Little Calumet River in Portage, Indiana, and Dunes Creek feeds into Lake Michigan within the Indiana Dunes State Park in Porter, Indiana. In general, *E. coli* densities in Burns Ditch are usually higher than the adjacent lake. Here I refer to the ditch- and lakeside locations as polluted

and less polluted areas, respectively. The majority of the *Cladophora* collections took place within the Little Calumet River watershed, in an area surrounding the breakwater adjacent to the Ogden Dunes beach.

Cladophora samples were collected by hand using latex gloves and immediately transferred to sterile Whirl-Pak® bags. Water samples were collected in sterile bottles or Whirl-Pak® bags. Between sampling, hands were sanitized with 70% alcohol, washed several times with sterile water, and fresh gloves were used at each sampling location to prevent cross-contamination. The samples were placed on ice in a cooler during transportation and analyzed within 24 h of collection.

4.3.2 Bacterial elutriation.

Bacterial isolates were obtained from *Cladophora*, sand, and sediment samples as previously described (Whitman et al., 2003; Ishii et al., 2006(b)). In brief, homogenized sub-samples (usually 75 g wet *Cladophora*, sand, or sediment) were shaken in 300 ml of phosphate-buffered water (pH 6.8) containing 0.01% hydrolyzed gelatin as described by Ishii et al., 2006(b). Bottles were shaken for 30 min on a horizontal shaker, allowed to stand for 20 min, and the upper phase served as the initial dilution (0.2 g wet algae/ml) for analysis.

4.3.3 Most Probable Number analysis.

Aliquots (0.5 ml or 5 ml) of the initial dilution were added to test tubes (n = 5) containing 4.5 ml or 45 ml of tetrathionate broth (TTB) (Difco Laboratories, Sparks,

MD). In addition, 50-ml aliquots of the initial dilution were used to reconstitute 2.3 g of the dehydrated TTB medium, followed by the addition of 1 ml iodine solution (6 g iodine crystals and 5 g potassium iodide in 20 ml water). This resulted in three dilutions (10, 1, and 0.1 g wet algae/ml), which were subsequently used for five-tube MPN analyses. Culture tubes were incubated at 37 °C for 48 h.

Enumeration of *Salmonella* in water, sand, sediment, and *Cladophora* samples was done using three-dilution, five-tube MPN analysis. Aliquots (1, 10, and 100 ml) of water samples (n = 5) were filtered through 0.45-µm membranes (Millipore, Billerica, MA). The membranes were transferred to tubes containing TTB medium and incubated at 37 °C for 48 h. During 2007, an additional 400 ml of each water sample was filtered and processed as described above.

The presence/absence of *Salmonella* in each MPN tube was determined by *Salmonella*-specific PCR as described previously (Ishii et al., 2006(b)). Positive (*S. enterica* subsp. *enterica* ser. Typhimurium ATCC 14028) and negative (uninoculated TTB) controls were used for PCR. MPN counts were calculated based on published tables (Alexander, 1982). Bacterial counts were expressed as log MPN g^{-1} dry weight (*Cladophora*, sand, sediment) or ml^{-1} (water). Ranges of MPN counts of three replicate samples are reported on a dry-weight basis.

4.3.4 Salmonella isolation and confirmation.

Presumptive *Salmonella* in MPN tubes were isolated by using selective and differential agar media as described by Ishii et al., 2006(b). *Salmonella* isolates were

stored in 50% glycerol at −80 °C until used. Three isolates per MPN tube were obtained in this manner and subsequently confirmed as *Salmonella* sp. by PCR using genus-specific primers (Ishii et al., 2006(b)). *S. enterica* subsp. *enterica* ser. Typhimurium strain ATCC 14028 was used as a positive control strain for cultural and PCR confirmations. Using this approach, 37, 79, and 64 *Salmonella* isolates were collected in 2005, 2006, and 2007, respectively, for a total of 180 isolates.

4.3.5 DNA fingerprint analysis.

Salmonella isolates were genotyped by using the HFERP DNA fingerprinting and the BOX1AR primer as previously described (Johnson et al., 2004; Ishii et al., 2006(b)). Gel images were analyzed by using BioNumerics version 3.0 software (Applied-Maths, Sint-Martens-Latem, Belgium). DNA fingerprint similarities were calculated by using Pearson's product-moment correlation coefficient, with 1% optimization, and dendrograms were generated by using the unweighted pair-group method with arithmetic means (UPGMA) as previously described (Johnson et al., 2004; Ishii et al., 2006(b)). A binary band-matching character table was generated by using the BOX-derived HFERP DNA fingerprint data, and this table was analyzed by using multivariate analysis of variance (MANOVA), a form of discriminant analysis. MANOVA, which was done accounting for the covariance structure, can be used to determine even small differentiating features in user-specified groups. Clustering of isolates was performed by using Jackknife analysis (Dombek et al., 2000), which is useful in determining whether entries in user defined groups are correctly classified.

4.4. Results

4.4.1 Salmonella detection and quantification by MPN-PCR from Cladophora.

In 2006, *Salmonella* was detected in 23% (7 of 31) of *Cladophora* samples from only the southwest side of Lake Michigan bordering Illinois and Indiana, with densities ranging from 0.16 to 1.21 MPN/g. In 2007, *Salmonella* was detected in 72% (13 of 18) of *Cladophora* samples, with densities ranging from 0.16 to 51.98. In contrast, while the frequency of *Salmonella* in *Cladophora* in 2005 was much lower (i.e. 24%, 8 of 33) than in 2007, bacterial densities were much higher ranging from 3.79 to 89.46 (Ishii et al., 2006(b)). For the years 2005–2007, *Cladophora* samples were collected from the lake- and ditch-sides of the Ogden Dunes breakwater in Indiana. In 2005, *Salmonella* was not detected by the MPN method in *Cladophora* samples collected from the ditch side, but was detected in 40% (8 of 20) of the lakeside samples, with 3.79 to 89.46 MPN/g algal tissue. In contrast, *Salmonella* was not detected in *Cladophora* samples collected in 2006 from the lakeside, but was detected in 75% (3 of 4) of ditch-side samples, with densities ranging from 0.16 to 1.21. In 2007, however, *Salmonella* was detected in 67% of ditch-side samples (2/3) and in all three of lakeside samples of the Ogden Dunes breakwater, with densities ranging from 2.94 to 13.76 and 0.16 to 41.26, respectively. In Dunes Creek, *Salmonella* was detected in 2 of 6 (33%) *Cladophora* samples in 2006 and in 2 of 3 (67%) samples in 2007, with densities ranging from 0.3 to 0.51 and 6.89 to 51.98 MPN/g algal tissue, respectively. Various cultural, biochemical, and molecular tests

confirmed that the presumptive bacterial isolates recovered from *Cladophora* mats were indeed *Salmonella*.

These results support our previous observations that *Cladophora* is likely an environmental reservoir of *Salmonella* in Lake Michigan watersheds (Ishii et al., 2006(b)). In 2006, *Salmonella* was detected in 57% (4 of 7) of the water samples, albeit in much lower densities (0.002 to 0.008 MPN/ml) relative to that seen with *Cladophora*. In 2007, *Salmonella* was detected in 33% (4 of 12) water samples, with densities ranging from 0.002 to 0.017 MPN/ml. No water samples were analyzed for *Salmonella* in 2005 (Ishii et al., 2006(b)). The high frequency of *Salmonella* in *Cladophora* relative to water suggested that either *Salmonella* multiplied on *Cladophora*, or that the algae may simply have acted as an entrapping matrix for the water-borne *Salmonella* cells. Since it was previously shown that algal washings provide nutrients to support the in vitro growth of enteric bacteria, such as *E. coli* (Byappanahalli et al., 2003(b)), our data support the hypothesis that *Salmonella* likely multiplies on *Cladophora* in the water column. Interestingly, no *Salmonella* was detected in *Cladophora* samples obtained from Door Peninsula and Washington Island (Door County, WI) or from Sleeping Bear Dunes National Lakeshore in Michigan (Figure 4.1), despite the fact that massive *Cladophora* accumulations are very common at these locations. These results suggest that the occurrence of *Salmonella* on *Cladophora* may be related to input sources, especially those coming from anthropogenic activities (e.g.wastewater, agricultural waste). Taken together, these results indicate that both the detection of *Salmonella* and their corresponding densities in *Cladophora* and water were highly variable over the study

period. While the variability is not completely unexpected, given the difference in the number and location of samples between years, it was surprising that *Salmonella* was detected in samples collected on both sides of the breakwater only in 2007. Variability in the association of *Salmonella* with *Cladophora* has been previously reported (Ishii et al., 2006(b)). Such variations may be explained by changes in input sources (Winfield and Groisman, 2003), perhaps due to runoff from rain events (Gaertner et al., 2008), to predation (Rhodes and Kator, 1988; Stevik et al., 2004), or other environmental factors such as dilution, temperature (Rhodes and Kator, 1988), solar radiation (McCambrdge and McMeekin, 1981), and the viable but non-culturable (VBNC) state of this microorganism (Roszaket al., 1984).

4.4.2 Population structure of Salmonella associated with Cladophora.

DNA fingerprint analyses indicated that the *Salmonella* strains isolated from *Cladophora* displayed a high degree of genetic similarity (>92%). Previously, it was shown that strains having HFERP DNA fingerprints with ≥92 similarity could be considered to be genetically identical (Ishii et al., 2006(b)). Due to the large number of isolates analyzed in my studies presented here, which resulted in a large correlation tree, a full-length dendrogram is not presented. However, when compressed for presentation purposes using a 92% cutoff value, the 133 *Salmonella* isolates clustered into 15 distinct groups, with 2 to 37 clonal isolates per group (Figure 4.2). Strikingly, all isolates from 2005 were in a single group (n ¼ 37), whereas those in 2006 and 2007 were split between several groups containing from 2 to 24 isolates. Only 7 (5%) of the isolates comprised

individual lineages, indicating that the strains were highly related to each other. In general, the 2005 and 2006 isolates were more closely related to each other than to the 2007 isolates. Four isolates from 2006 were very closely related (at the 85% similarity level) to those in 2005. Likewise, 4 isolates from 2006 were distantly related (48% similarity level) to those in 2005. The high degree of genetic relatedness among *Salmonella* isolates was unexpected, given that the bacterial isolates were independently obtained from *Cladophora* samples collected over time and space to ensure that both numerically dominant and less prevalent strains would be isolated from *Cladophora* mats. Moreover, some of the genotypically-identical *Salmonella* isolates were obtained from *Cladophora* samples attached to different rocks, as much as 25 m apart on either side of the embayment near Ogden Dunes beach (Ishii et al., 2006(b)). These results suggest that either nearly identical *Salmonella* cells colonized *Cladophora* or that certain strains of *Salmonella* have a selective advantage in associating with *Cladophora* (strain-specific associations) that enable their growth or enhanced survival in the environment (Englebert et al., 2008).

4.4.3 Spatial–temporal relationships.

MANOVA analysis of isolates obtained in 2005, 2006, and 2007 (Figure 4.3) indicated that the *Salmonella* isolates strongly clustered by year. The strong clustering of isolates was also seen in Jackknife analysis (Table 4.2); percent correct assignment of isolates by year was 100, 74, and 100% for 2005, 2006, and 2007, respectively. Jackknife analyses of geographically distinct strains collected in 2007 (Table 4.3) showed that the

HFERP DNA fingerprints of *Salmonella* isolates from *Cladophora* from Burns ditch near the steel plant (DNP), Wind Point (WP), and Portage Marina (PM) sites were correctly assigned to their corresponding sampling sites 100% of the time, and the 63rd Street Beach (63rd) isolates were correctly assigned 89% of the time, again indicating that there was a strong association of isolates by location. In contrast, the percent correct assignment of *Cladophora*-borne *Salmonella* from Dunes Creek (DC), the ditch-(IDD) and lake- (IDL) sides of Indiana Dunes, and Salt Creek-700 N (700 N) locations was only 0, 13, 66, and 50%, respectively. While some of the genotypes were shared among certain locations (e.g., DC, IDD, and IDL) that were likely impacted by similar contaminant sources, the majority of *Salmonella* genotypes were both spatially and temporally different, suggesting that these genotypes were likely derived from different sources (Ishii et al., 2006(b)).

4.4.4 Genetic relatedness of Salmonella collected from other sources.

Besides *Cladophora*, *Salmonella* was also recovered from a variety of samples, including beach sand, sediment, creek and lake water, and an aquatic plant, *Potamogeton sp.* Overall, the *Salmonella* isolates obtained from these sources (n = 47) displayed a high degree of genetic relatedness (>92% similarity), but the isolates were not all identical. In general, the *Salmonella* isolates clustered by sample type and location (data not shown). Despite genotypic differences within the populations examined, nearly identical *Salmonella* isolates were found in certain samples. For example, 100, 75, and 82% of the isolates obtained from the Little Calumet River (west branch), *Potamogeton sp.* (North

Beach), and river water from the Burns Ditch side of breakwater were >92% similar, respectively, to other isolates obtained from the same locations. *Salmonella* isolates from 63rd Street Beach sand were closely related, at the 88% similarity level, to those from water at the Little Calumet River (west branch).

Some of the *Cladophora*-borne *Salmonella* isolates collected during 2006 and 2007 at each sampling location were nearly identical to those obtained from the other sources discussed above. For example, 75% (9 of 12 isolates) isolated from Burns Ditch water samples were identical to 69% (9 of 13) of the *Cladophora*-derived isolates, and in 2006, 3 of 3 and 2 of 5 *Salmonella* isolates from sand and sediment samples from the 63rd Street Beach, respectively, were identical to 55% of the *Cladophora*-borne *Salmonella* isolates obtained from the same location. The recovery of *Salmonella* from the aquatic plant *Potamogeton sp*., which is not phylogenetically related to *Cladophora*, indicates that the *Salmonella–Cladophora* association is less likely to be inter-dependent, instead it may reflect a survival strategy of *Salmonella* in non-host environments (Winfield and Groisman, 2003). Taken together, these findings suggest that *Cladophora* likely acquires *Salmonella* from its adjacent environment (e.g. sand, sediment, and water) and there may be an exchange of genotypes between sources.

In the 2007 multi-watershed sampling, the presence of a particular *Salmonella* genotype at given location in a waterway had limited influence on those found further downstream, and the genotypes were most often different. Whitman et al. (2001) also reported the recovery of genetically diverse *S. enterica* subsp. *enterica* ser. Typhimurium isolates from Lake Michigan water and sediments. Similarly, diverse serovars of

Salmonella have been previously isolated from marine beaches (Aulicino et al., 2001; Martinez-Urtaza et al., 2004) and beach sand (Sanchez et al., 1986; Bolton et al., 1999; Shatti and Abdullah, 1999). Based on the results of our findings, I suggest that *Salmonella* population in water, sediment, and sand may genetically be more diverse than that associated with *Cladophora*. Overall, the *Salmonella* strains isolated in this study were relatively diverse. The Shannon diversity index for the entire (all sources included) population was 2.71, with a species richness value of 29. Interestingly, when the five dominant groups (i.e. groups with more than 10 members) were removed from the analysis, the diversity index did not change substantially (from 2.71 to 2.84). Moreover, when isolates derived from the same MPN tube were removed to reduce potential enrichment bias, only two groups remained that possessed more than 10 members, with the resulting diversity index of 2.88. Taken together, these results suggest that the diversity was spread fairly evenly across all the groups. The Shannon diversity index reported here is very similar to that reported for *Salmonella enteritidis* (2.81) obtained from human and animal sources (Cho et al., 2007). The current findings support previous results obtained by Ishii et al. (2006(b)) that showed that the nuisance green alga *Cladophora* can serve an environmental reservoir for the enteric bacterial pathogen *Salmonella* in Lake Michigan beachsheds.

At this time, the source of these bacteria remains speculative. However, the association of this pathogen with *Cladophora* raises serious implications for beach water quality and public health, especially in the Great Lakes where *Cladophora* accumulations are common during the recreational season.

4.5. Discussion

In this study, I used cultural, biochemical, and molecular approaches to determine if *Salmonella* was commonly associated with the green alga *Cladophora* and whether specific *Salmonella* strains (genotypes) have a propensity to associate with *Cladophora* at the same and different sites over time. Taken as a whole, my data indicates that: (1) *Cladophora* in Lake Michigan serves as an environmental reservoir for *Salmonella* and potentially other pathogens, such as shiga toxin-producing *E. coli*, *Shigella*, and *Campylobacter*, as previously described; (2) Genotypically identical *Salmonella* strains associate with geographically separated *Cladophora* and other matrices; (3) *Salmonella* strains associated with *Cladophora* varied extensively by year that may be explained by the *Salmonella* genotype that numerically dominates at a particular site each year; and (4) The association of potentially pathogenic strains of *Salmonella* with *Cladophora* warrants additional studies to assess risks to public health and impact on regulatory issues.

4.6. Acknowledgements

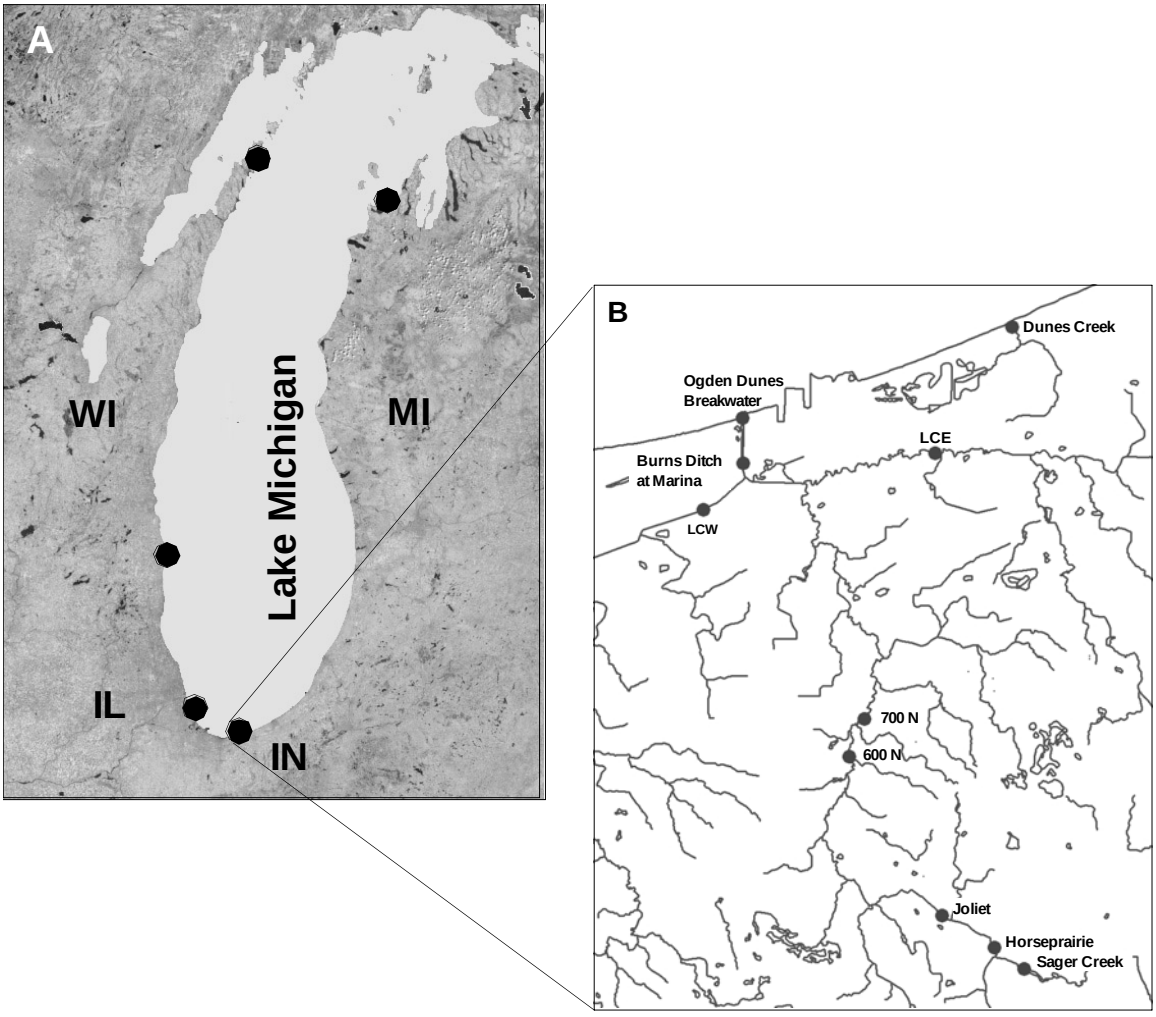
I would like to thank Dr. Muruleedhara N. Byappanahalli for his extended work on this project, and for permission to use this publication. I would also like to thank Dr. Satoshi Ishii for his extended direction and guidance on this project. I would also like to thank John Ferguson for his assistance in using Bioinformatics for HFERP analysis. Finally, I would like to extend thanks to the entire Sadowsky lab for support and guidance lended throughout the course of this project.

4.7. Permission of use.

This study was previously published in the journal *Water Research*. It is used here by permission of Elsevier Ltd., under license number 2260270481183.

4.8. Tables and Figures

Figure 4.1



(A) Map showing general study area with sampling locations (●). *Cladophora* samples were collected between May and August during 2005-2007 along the shorelines of Lake Michigan bordering Wisconsin, Illinois, Indiana, and north Michigan. During 2007, sampling was focused in Dunes Creek, the Little Calumet River, and the Salt Creek watersheds (B), where *Salmonella* was frequently detected *Cladophora* in previous years (Ishii et al., 2006).

Table 4.1 Sampling location, sample type, and number of samples collected during 2006 and 2007.

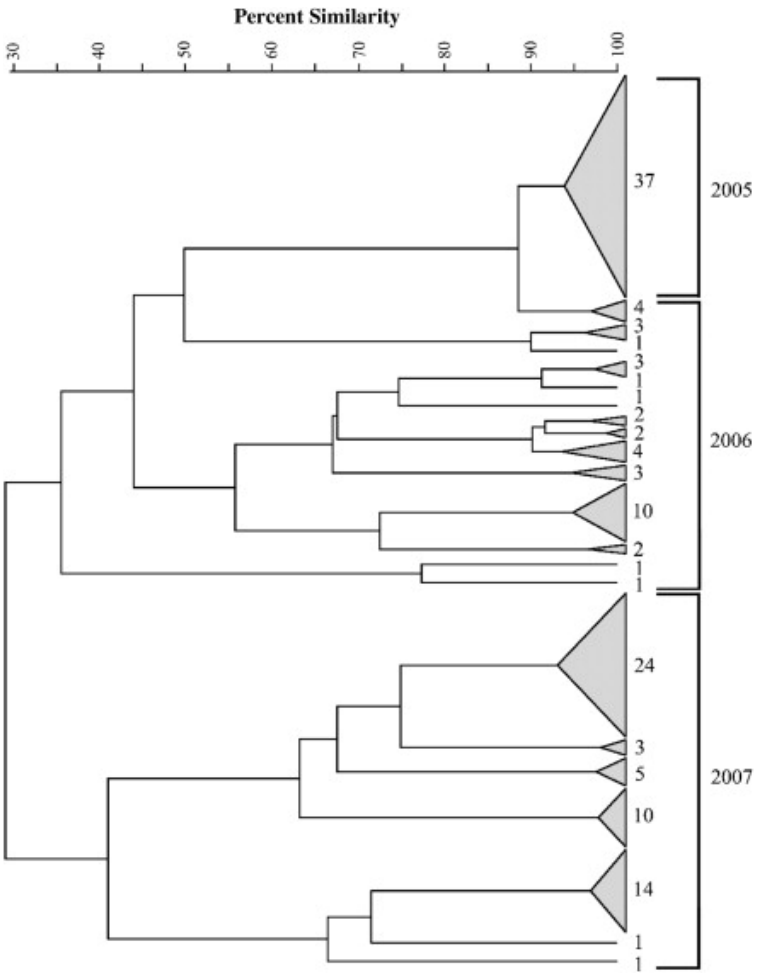
Location	State	Year	Substrate (N)
Burns Ditch, breakwater	IN	2006	*Cladophora* (5), Water (2)
Lake Michigan, breakwater	IN	2006	*Cladophora* (5), Water (1)
Dunes Creek	IN	2006	*Cladophora* (6)
Little Calumet River, west	IN	2006	Water (1)
63rd St. Beach	IL	2006	*Cladophora* (1), Water (3)
Wind Point	WI	2006	*Cladophora* (5)
North Beach	WI	2006	Non-*Cladophora* plant (1)
Murphy[a]	WI	2006	*Cladophora* (1)
Ephraim[a]	WI	2006	*Cladophora* (1)
Lakeside[a]	WI	2006	*Cladophora* (1)
Jackson Port[a]	WI	2006	*Cladophora* (1)
Whitefish Dunes[a]	WI	2006	*Cladophora* (1)
County Rd. 651[a]	MI	2006	*Cladophora* (1)
Glen Haven Cannery[a]	MI	2006	*Cladophora* (1)
Platte River Point[a]	MI	2006	*Cladophora* (1)
Esch Rd.[a]	MI	2006	*Cladophora* (1)
Burns Ditch, breakwater	IN	2007	*Cladophora* (3), Water (1)
Lake Michigan, breakwater	IN	2007	*Cladophora* (3), Water (1)

Location	State	Year	Substrate (N)
Dunes Creek	IN	2007	*Cladophora* (3), Water (1)
Little Calumet River, west[a]	IN	2007	Water (1)
Burns Ditch, marina	IN	2007	*Cladophora* (3), Water (1)
Burns Ditch, near plant	IN	2007	*Cladophora* (3), Water (1)
Little Calumet River, east[a]	IN	2007	Water (1)
Salt Creek, 700 N	IN	2007	*Cladophora* (3), Water (1)
Salt Creek, 600 N[a]	IN	2007	Water (1)
Salt Creek, Joliet Rd.[a]	IN	2007	Water (1)
Salt Creek, Horse prairie Rd.[a]	IN	2007	Water (1)
Sager Creek, Sager Rd.[a]	IN	2007	Water (1)

All sampling locations in Indiana (IN) are located in Porter County; Wind Point and North Beach sites are located in Racine County, Wisconsin (WI); Murphy, Ephraim, Lakeside, Jackson Port, and Whitefish sites are located in Door County, WI; County Road 651, Glen Haven cannery, Platte River Point, and Esch Road sites are located Leelanau County, Michigan (MI).

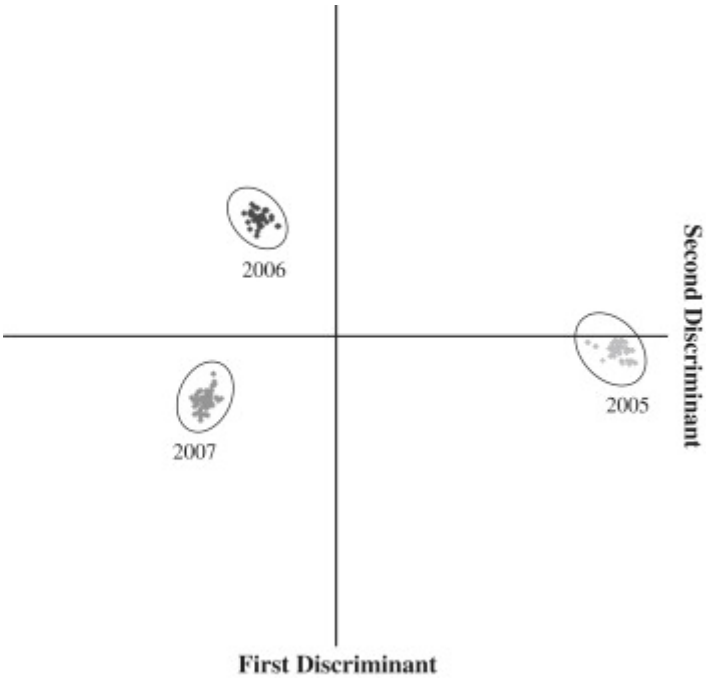
[a] Denotes sampling locations where isolates were not obtained.

Figure 4.2



Dendrogram showing the relatedness of *Salmonella* strains isolated from *Cladophora* as determined by HFERP DNA fingerprint analysis using the Box A1R primer. DNA fingerprint similarities were calculated by using the curve-based, cosine coefficient and dendrograms were generated by the unweighted pair-group method using arithmetic averages (UPGMA). Triangles represent genetically unique (i.e. ≥92% similarity) *Salmonella* isolates clustering into distinct groups; numbers to the right of the each triangle represent the number of isolates in that group.

Figure 4.3



MANOVA analysis of *Salmonella* isolates obtained from *Cladophora* collected in 2005, 2006, and 2007.

Table 4.2 Jackknife analysis of relative average similarities of *Salmonella* isolates by year.

Assigned Group	**Average similarities (%)**		
	2005	**2006**	**2007**
2005	100	10	0
2006	0	*87*	0
2007	0	3	100

Table 4.3 Jackknife analysis of relative, average similarities of *Salmonella* isolates obtained in 2007 by site.

Assigned Group	Average Similarities (%)									
	63rd	**700 N**	**BD, near plant**	**Dunes Creek**	**BD, breakwater**	**Lake, Breakwater**	**LCR, east**	**North Beach**	**BD, marina**	**Wind Point**
63rd	*24*	0	0	0	0	0	0	0	0	0
700 N	0	0	0	0	0	0	0	0	0	0
BD, near plant	0	0	*50*	0	0	0	0	0	0	0
Dunes Creek	0	14	0	63	0	0	0	0	0	0
BD, breakwater	71	0	3	5	100	0	0	0	0	0
Lake, breakwater	0	0	6	0	*0*	100	0	0	0	0
LCR, east	0	14	28	0	0	0	100	0	0	0
North Beach	0	52	8	27	0	0	0	100	0	50
BD, marina	6	10	0	5	0	0	0	0	100	0
Wind Point	0	10	6	0	0	0	0	0	0	*50*

Abbreviations for locations: 63rd, 63rd Street Beach, Chicago, IL; 700 N, 700th Street North, Salt Creek IN; BD near plant, Burns Ditch near steel plant, Indiana Dunes National Lakeshore, IN; Dunes Creek, Dunes Creek, IN; BD breakwater, Burns Ditch side of breakwater, Indiana Dunes National Lakeshore, IN; Lake breakwater, Lakeside of breakwater, Indiana Dunes National Lakeshore, IN; LCR East, Little Calumet River East, IN; North Beach, North Beach, Racine, WI; BD marina, Burns Ditch near marina, Portage, IN; Wind Point, Wind Point beach, Racine, WI.

4.9. References

Alexander, M. 1982. Most probable number method for microbial populations. In: A.L. Page, R.H. Miller and D.R. Keeney, Editors, Methods of Soil Analysis, Part 2 (second ed.), Agron. Monogr vol. 9, American Society for Agronomy and Soil Science Society of America, Madison, WI, pp. 815–820.

Aulicino, F.A.; Orsini, P.; Carere, M.; Mastrantonio, A. 2001. Bacteriological and virological quality of seawater bathing areas along the Tyrrhenian coast. Int. J. Environ. Health Res. 11 (1), pp. 5–11.

Bolton, F.J.; Surman, S.B.; Martin, K.; Wareing, D.R.; Humphrey, T.J. 1999. Presence of *Campylobacter* and *Salmonella* in sand from bathing beaches. Epidemiol. Infect. 122, pp. 7–13.

Byappanahalli, M.; Fowler, M.; Shively, D.; Whitman, R. 2003(a). Ubiquity and persistence of *Escherichia coli* in a midwestern coastal stream. Appl. Environ. Microbiol. 69 (8), pp. 4549–4555.

Byappanahalli, M.; Shively, D.A.; Nevers, M.B.; Sadowsky, M.J.; Whitman, R.L. 2003(b). Growth and survival of *Escherichia coli* and enterococci populations in the macro-alga *Cladophora* (Chlorophyta). FEMS Microbiol. Ecol. 46, pp. 203–211.

Byappanahalli, M.; Whitman, R.L.; Shively, D.A.; Ferguson, J.; Ishii, S.; Sadowsky, M. 2007. Population structure of *Cladophora*-borne *Escherichia coli* in nearshore water of Lake Michigan. Water Res. 41, pp. 3649–3654.

Cho, S.; Boxrud, D.J.; Bartkus, J.M.; Whittam, T.S.; Saeed, M. 2007. Multiple-locus variable-number tandem repeat analysis of *Salmonella enteritidis* isolates from human and non-human sources using a single multiplex PCR. FEMS Microbiol. Lett. 275 (1), pp. 16–23.

Dombek, P.E.; Johnson, L.K.; Zimmerley, S.T.; Sadowsky, M.J. 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl. Environ. Microbiol. 66 (6), pp. 2572–2577.

Englebert, E.T.; McDermott, C.; Kleinheinz, G.T. 2008. Impact of the alga *Cladophora* on the survival of *E. coli*, *Salmonella*, and *Shigella* in laboratory microcosm. J. Great Lakes Res. 34 (2), pp. 377–382.

Gaertner, J.P.; Garres, T.; Becker, J.C.; Jimenez, M.L.; Forstner, M.R.J.; Hahn, D. 2008. Temporal analyses of *Salmonella* in a headwater spring ecosystem reveals the effects of precipitation and runoff events. J. Water Health 7 (1), pp. 115–121.

Ishii, S.; Ksoll, W.B.; Hicks, R.E.; Sadowsky, M.J. 2006(a). Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), pp. 612–621.

Ishii, S.; Yan, T.; Shively, D.A.; Byappanahalli, M.; Whitman, R.L.; Sadowsky, M.J. 2006(b). *Cladophora* (Chlorophyta) spp. harbor human bacterial pathogens in nearshore water of Lake Michigan. Appl. Environ. Microbiol. 72 (7), pp. 4545–4553.

Johnson, L.-A.K.; Brown, M.B.; Carruthers, E.A.; Ferguson, J.A.; Dombek, P.E.; Sadowsky, M.J. 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70 (8), pp. 4478–4485.

Kleinheinz, G.T.; McDermott, C.M.; Leewis, M.-C.; Englebert, E. 2006. Influence of sampling depth on *Escherichia coli* concentrations in beach monitoring. Water Res. 40, pp. 3831–3837.

Martinez-Urtaza, J.; Saco, M.; de Novoa, J.; Perez-Piñeiro, P.; Peiteado, J.; Lozano-Leon, A.; Garcia-Martin, O. 2004. Influence of environmental factors and human activity on the presence of *Salmonella* serovars in a marine environment. Appl. Environ. Microbiol. 70 (4), pp. 2089–2097.

McCambrdge, J.; McMeekin, T.A. 1981. Effect of solar radiation and predacious microorganisms on survival of fecal and other bacteria. Appl. Environ. Microbiol. 41 (5), pp. 1083–1087.

Olapade, O.A.; Depas, M.M.; Jensen, E.T.; McLellan, S.L. 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Appl. Environ. Microbiol. 72 (3), pp. 1932–1938.

Rhodes, M.W.; Kator, H. 1988. Survival of *Escherichia coli* and *Salmonella spp*. in estuarine environments. Appl. Environ. Microbiol. 54 (12), pp. 2902–2907.

Roszak, D.B.; Grimes, D.J.; Colwell, R.R. 1984. Viable but nonrecoverable stage of *Salmonella enteritidis* in aquatic systems. Can. J. Microbiol. 30 (3), pp. 334–338.

Sanchez, P.S.; Agudo, E.G.; Castro, F.G.; Alves, M.N.; Martins, M.T. 1986 Evaluation of the sanitary quality of marine recreational waters and sands from beaches of the Sao Paulo State, Brazil. Water Sci. Technol. 18, pp. 61–72.

Shatti, J.A.; Abdullah, T.H.A. 1999. Marine pollution due to wastewater discharge in Kuwait. Water Sci. Technol. 40 (7), pp. 33–39.

Stevik, T.K.; Aa, K.; Ausland, G.; Hanssen, J.F. 2004. Retention and removal of pathogenic bacteria in wastewater percolating through porous media: a review. Water Res. 38 (6), pp. 1355–1367.

Versalovic, J.; Koeuth, T.; Lupski, R. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19 (24) (1991), pp. 6823–6831.

Whitman, R.L. 1983. The influence of a combined sewage outfall on the chemical and biological quality of Salt Creek, Porter County, Indiana. Proc. Indiana Acad. Sci. 92, pp. 275–282.

Whitman, R.L.; Horvath, T.G.; Goodrich, M.L.; Nevers, M.B.; Wolcott, M.J.; Haack, S.K. 2001. Characterization of *E. coli* levels at 63rd Street Beach. Report to the City of Chicago, Department of the Environment and the Chicago Park District, Chicago, Illinois. U.S. Geological Survey, Porter, IN. Available from: http://www.glsc.usgs.gov/_files/reports/63rdStreetBeachReport.pdf (Online).

Whitman, R.L., and Nevers, M.B. 2003. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. Appl. Environ. Microbiol. 69 (9), pp. 5555–5562.

Whitman, R.L.; Shively, D.A.; Pawlik, H.; Nevers, M.B.; Byappanahalli, M. 2003. Occurrence of *Escherichia coli* and enterococci in *Cladophora* (Chlorophyta) in nearshore water and beach sand of Lake Michigan. Appl. Environ. Microbiol. 69 (8), pp. 4714–4719.

Winfield, M.D.; Groisman, E.A. 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. 69 (7), pp. 3687–3694.

Chapter 5. Conclusions and Future Directions

The fecal loading of aquatic environments by various animal hosts is of concern for public and environmental health. The microbiological contamination of waterways is amongst the most commonly listed water quality impairment in the U. S. Once waterways are deemed impaired, states must conduct a Total Maximum Daily Load (TMDL) assessment in order to mitigate the impairment and restore the water body to acceptable quality. The use of rapid molecular-based assays to determine fecal bacterial sources and source-specific loading quantities has the potential to improve the accuracy and efficacy of TMDL studies, and to expedite their implementation strategies. The current state of MST methodology lies in the development and validation of animal host-specific gene markers that can be identified directly from environmental samples. The body of research presented in the second and third chapters of this thesis concentrate on the identification and evaluation of animal host-specific gene markers.

In Chapter 2, I used suppression subtraction hybridization (SSH) to identify gene markers specific to *E. coli* derived from swine fecal sources. Results of my studies confirm that SSH provides a useful tool for identifying source-specific DNA marker genes from *E. coli.* This study identified a DNA marker gene set capable of identifying a large amount of *E. coli* originating from swine and turkeys, suggesting that this probe set may be useful for the assessment of fecal loading in watersheds where only one of these animals is present. While the current study indicates that this marker set may be used as

a hybridization probe to identify environmental *E. coli* isolates, more extensive work will be required to determine its geographical specificity and field applicability.

In Chapter 3, I examined the influence of cattle grazing operations on a small stream system in Southeastern Minnesota using plate count analysis and an optimized qPCR method. Environmental DNA's were extracted bi-monthly over the summer months of 2008 and 2009 at 5 separate sampling locations each year. In addition, membrane filtration was conducted on acquired samples to obtain *E. coli* plate count data per EPA Method 1603 protocol (USEPA, 2002). DNA's were first subjected to conventional PCR analyses using two separate primer sets that targeted: (1) the 16s rDNA gene belonging to all bacteria in the genera *Bacteroides* (AllBac); and (2) the human specific HF 183 *Bacteroides* marker (Bernhard and Field, 2000; Layton et al., 2006). A quantitative PCR (q-PCR) assay was then developed and evaluated to estimate the total presence of bovine-specific *Bacteroides* by use of the Bac3 gene marker (Shanks et al., 2006). Subsequently, the q-PCR data were compared to the *E. coli* plate count data, revealing a lack of consistent correlation between the two methods. Several physical and environmental factors likely influenced the level of *E. coli* found in the stream, confirming the hypothesis that information will be needed to supplement current FIB monitoring efforts in order to determine accurate source-specific fecal impacts. While the Bac3 qPCR assay effectively identified the scale of the bovine-specific *Bacteroidales* marker gene, work is still needed in accurately characterizing the numerical genome possession of *Bacteroidales* so cell numbers can be determined and scale of bovine fecal loading assessed. Further, fractional characterization of

Bacteroidales in bovine intestinal flora will be necessary to normalize observed cell numbers in water samples with amount of feces loaded to waterways. Future studies will also be required to establish epidemiological risk relationships to gene targets derived from fecal anaerobe targets.

In 2009, an agricultural BMP was facilitated prior to the sampling season. While plate count outputs in 2009 could not be correlated to facilitation of the BMP due to a lack of rain event samplings, the Bac3 qPCR copy data was more encouraging. Decreased overall counts at sampling locations involved in the BMP suggested the desired affect was achieved. However, E. coli counts often still exceeded the state standard, so employment of a watershed scale BMP may be necessary to ultimately mitigate the fecal contamination of the Little Jordan Creek.

In Chapter 4, I examined whether secondary habitats for fecal bacteria also harbor human pathogens. Specifically, I examined spatial and temporal variation in the population structure of *Salmonella* isolated from *Cladophora* found in stream and lake water, aquatic plants, beach sand, and sediments by use of horizontal fluorophore-enhanced rep-PCR (HFERP). Analysis of HFERP fingerprints obtained from 2005-2007 at various Lake Michigan locations revealed that *Salmonella* populations associated with *Cladophora* varied both spatially and temporally, suggesting potentially different input sources of *Salmonella* over space and time. In addition, differing environmental stressors may play a role in selecting particular *Salmonella* genotypes that are best suited for growth in this environment, indicating the potential for genetic drift of the introduced *Salmonella.* The association of potentially pathogenic strains of *Salmonella* with

Cladophora warrants additional studies to assess risks to public health and impact on regulatory issues.

5.1. References

Bernard, A.E., and Field, K.G. 2000(b). 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, v. 66, no. 10, p. 4571-4574.

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. Appl. Environ. Microbiol. 72(6):4214-4224.

Shanks, O.C.; Santo Domingo, J.W.; Lamendella, R.; Kelty, C.A.; Graham, J.E. 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72 (6), 4054–4060.

U.S. Environmental Protection Agency. 2002. Method 1603—*Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar. EPA821-R-02-23, 9. U. S. Environmental Protection Agency, Washington, DC.

References

Agron, P. G.; Walker, R. L.; Kinde, H.; Sawyer, S. J.; Hayes, D. C.; Wollard, J.; Andersen, G. L. 2001. Identification by subtractive hybridization of sequences specific for *Salmonella enterica* serovar Enteritidis. Appl. Environ. Microbiol. 67:4984–4991.

Albert, J.M.; Munakata-Marr, J.; Tenorio, L.; Siegrist, R.L. 2003. Statistical evaluation of bacterial source tracking data obtained by rep-PCR DNA fingerprinting of *Escherichia coli*. Environ. Sci. Technol. 37 (20), 4554–4560.

Alexander, M. 1982. Most probable number method for microbial populations. In: A.L. Page, R.H. Miller and D.R. Keeney, Editors, Methods of Soil Analysis, Part 2 (second ed.), Agron. Monogr vol. 9, American Society for Agronomy and Soil Science Society of America, Madison, WI, pp. 815–820.

Alexander, L. M.; Heaven, A.; Tennant, A.; Morris, R. 1992. Symptomatology of children in contact with sea water contaminated with sewage. J. Epidemiol. Community Health 46:340–344.

Allsop, K.; Stickler, J.D. 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. J. Appl. Bacteriol. 58, 95–99.

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

Aulicino, F.A.; Orsini, P.; Carere, M.; Mastrantonio, A. 2001. Bacteriological and virological quality of seawater bathing areas along the Tyrrhenian coast. Int. J. Environ. Health Res. 11 (1), pp. 5–11.

Bailey, D.W., and G. R. Welling. 1999. Modification of cattle grazing distribution with dehydrated molasses supplement. J. Range Manage. 52:575-582.

Bernhard, A. E., and Field, K. G. 2000(a). Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. 66:1587–1594.

Bernard, A.E., and Field, K.G. 2000(b). 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, v. 66, no. 10, p. 4571-4574.

Blankenship, K. 1996. DNA library would give investigators inside poop on pollution sources. Bay Journ. 6(6).

Bolton, F. J.; Surman, S. B.; Martin, K.; Wareing, D. R.; Humphrey, T. J. 1999. Presence of *Campylobacter* and *Salmonella* in sand from bathing beaches. Epidemiol. Infect. 122:7–13.

Bower, Patricia A., Scopel, Caitlin O.; Jensen, Erika T.; Depas, Morgan M.; McLellan, Sandra L. 2005. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. Appl. Environ. Microbiol. 71: 8305-8313

Byappanahalli, M.; Fowler, M.; Shively, D.; Whitman, R. 2003(a). Ubiquity and persistence of *Escherichia coli* in a Midwestern coastal stream. Appl. Environ. Microbiol. 69 (8), 4549–4555.

Byappanahalli, M.; Shively, D.A.; Nevers, M.B.; Sadowsky, M.J.; Whitman, R.L. 2003(b). Growth and survival of *Escherichia coli* and enterococci populations in the macro-alga *Cladophora* (Chlorophyta). FEMS Microbiol. Ecol. 46, pp. 203–211.

Byappanahalli, M., and Fujioka, R. 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water. Sci. Technol. 50 (1), 27–32.

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

Byappanahalli, M.; Whitman, R.L.; Shively, D.A.; Ting, W.T.; Tseng, C.C.; Nevers, M.B. 2006(b). Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. J. Water Health 4 (3), 313–320.

Byappanahalli, M.; Whitman, R.L.; Shively, D.A.; Ferguson, J.; Ishii, S.; Sadowsky, M. 2007. Population structure of *Cladophora*-borne *Escherichia coli* in nearshore water of Lake Michigan. Water Res. 41, pp. 3649–3654.

Cabelli, V.J.; Dufour, A.P.; Levin, M.A.; McCabe, L.J.; Haberman, P.W. 1979. Relationship of Microbial Indicators to Health Effects at Marine Bathing Beaches. Am. J. Public Health. 69: 690-696.

Carroll, S., Hargreaves, M., Goonetilleke, A. 2005. Sourcing fecal pollution from onsite wastewater treatment systems in surface waters using antibiotic resistance analysis. J. Appl. Microbiol. 99, 471–482.

Centers for Disease Control and Prevention. 1998. Outbreak of *Vibrio parahaemolyticus* infections associated with eating raw oysters, Pacific Northwest, 1997. Morb. Mortal. Wkly. Rep. 147:45762.

Centers for Disease Control and Prevention. 2008. Surveillance Summaries. Surveillance for Waterborne Disease and Outbreaks Associated with Recreational Water Use and Other Aquatic Facility-Associated Health Events — United States, 2005–2006 and Surveillance for Waterborne Disease and Outbreaks Associated with Drinking Water and Water not Intended for Drinking — United States, 2005–2006 . MMWR 2008;57(No. SS-9).

Cheung,W.H.S. (Ed.), 1988. Health Effects of BeachWater Pollution in Hong Kong. Vincent Blue Copy Co., Hong Kong.

Cho, S.; Boxrud, D.J.; Bartkus, J.M.; Whittam, T.S.; Saeed, M. 2007. Multiple-locus variable-number tandem repeat analysis of *Salmonella enteritidis* isolates from human and non-human sources using a single multiplex PCR. FEMS Microbiol. Lett. 275 (1), pp. 16–23.

Colles, F. M.; Jones, K.; Harding, R. M.; Maiden, M. C. J. 2003. Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment. Appl. Environ. Microbiol. 69:7409–7413.

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

Dick, L.K.; Simonich, M.T.; Field, K.G. 2005(b). Microplate subtractive hybridization to enrich for source-specific Bacteroidales fecal pollution indicators. Appl. Environ. Microbiol. 71, 3179–3183.

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

Dufour, A.P. 1977 *Escherichia coli*: The faecal coliform. In Bacterial Indicators/Health Hazards Associated with Water ed. Hoadley, A.W. & Dutka, B.J. ASTM STP 635, pp. 48-58. Philadelphia: American Society for Testing and Materials.

Edge, T.A.; Hill, S. 2005. Occurrence of antibiotic resistance in *Escherichia coli* from surface waters and fecal pollution sources near Hamilton, Ontario. Can. J. Microbiol. 51 (6), 501–505.

Edwards, D. R.; Coyne, M. S.; Vendrell, P. F.; Daniel, T. C.; Moore, P. A., Jr.; Murdoch, J. F. 1997. *Escherichia coli* survival in mantled karst springs and streams, Northwest Arkansas Ozarks, USA. JAWRA, *33*, 413-422.

Edwards, D.R.; Larson, B.T.; Lim, T.T. 2000. Runoff nutrient and fecal coliform content from cattle manure application to fescue plots. J. Am. Water Resour. Assoc. 36:711–721.

Englebert, E.T.; McDermott, C.; Kleinheinz, G.T. 2008. Impact of the alga *Cladophora* on the survival of *E. coli*, *Salmonella*, and *Shigella* in laboratory microcosm. J. Great Lakes Res. 34 (2), pp. 377–382.

Feldgarden, M.; Riley, M. A. 1999. The phenotypic and fitness effects of colicin resistance in Escherichia coli K-12. Evolution. 53(4), 1019-27.

Field, Katharine G., Samadpour, Mansour. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research. 41. 3517–3538.

Fiksdal, L.; Make, J.S.; LaCroix, S.J.; Staley, J.T. 1985. Survival and detection of *Bacteroides spp.*, prospective indicator bacteria. Appl. Environ. Microbiol. 49 (1), 148–150.

Fujioka, R.; Dian-Denton, C.; Castro, J.; Morphew, K. 1999. Soil: the environmental source of *Escherichia coli* and *Enterococci* in Guam's streams. J. Appl. Microbiol. 85, 83S–89S.

Gaertner, J.P.; Garres, T.; Becker, J.C.; Jimenez, M.L.; Forstner, M.R.J.; Hahn, D. 2008. Temporal analyses of *Salmonella* in a headwater spring ecosystem reveals the effects of precipitation and runoff events. J. Water Health 7 (1), pp. 115–121.

Gordon, D. M. 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology 147:1079–1085.

Griffith, J.F.; Weisburg, S.B.; McGee, C.D. 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. J. Water Health 1, 141–152.

Hagedorn, C.; Robinson, S.L.; Filtz, J.R.; Grubbs, S.M.; Angier, T.A.; Reneau Jr., R.B. 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. 65 (12), 5522–5531.

Hamilton, M.J.; Yan, T.; Sadowsky, M.J.; 2006. Development of goose- and duck-specific DNA markers to determine sources of *Escherichia coli* in waterways. Appl. Environ. Microbiol. 72 (6),4012–4019.

Harakava, R.; and Gabriel, D. W. 2003. Genetic differences between two strains of *Xylella fastidiosa* revealed by suppression subtractive hybridization. Appl. Environ. Microbiol. 69:1315–1319.

Hartel, P. G.; Summer, J. D.; Hill, J. L.; Collins, J.; Entry, J. A.; Segars, W.I. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. J. Environ. Qual. 31:1273–1278.

Hartel, P. G.; Summer, J. D.; Segars, W.I. 2003. Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking. Water Res. 37:3263–3268.

Harwood, V.J.; Whitlock, J.; Withington, V. 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. 66 (9), 3698–3704.

Harwood, V.J.; Wiggins, B.A.; Hagedorn, C.; Ellender, R.D.; Gooch, J.; Kern, J.; Samadpour, M.; Chapman, A.C.H.; Robinson, B.J. 2003. Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Water Health 1, 153–166.

Hassan, W.M.; Wang, S.Y.; Ellender, R.D. 2005. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Appl. Environ. Microbiol. 71 (1), 512–518.

Hartl, D. L., and D. E. Dykhuizen. 1984. The population genetics of Escherichia coli. Annu. Rev. Genet. 18:31–68.

Heinonen-Tanski, H.; Uusi-Kamppa, J. 2001. Runoff of fecal microorganisms and nutrients from perennial grass ley after application of slurry and mineral fertilizer. Water Sci. Technol. 43:143–146

Hsieh, W. J., and M. J. Pan. 2004. Identification *Leptospira santarosai* serovar shermani specific sequences by suppression subtractive hybridization. FEMS Microbiol. Lett. 235:117–124.

Holmes, P.R. 1989. Research into health risks at bathing beaches in Hong Kong. J. Inst. Water Environ. Manage. 3,488–495.

Howell, J. M.; Coyne, M. S.; Cornelius, P. 1995. Fecal bacteria in agricultural waters of the bluegrass region of Kentucky. J. Environ. Qual, 24, 411-419.

Indest, K.J.; Betts, K.; Furey, J.S. 2005. Application of oligonucleotide microarrays for bacterial source tracking of environmental *Enterococcus sp.* isolates. Int. J. Environ. Res. Public Health 2 (1), 175–185.

Ishii, S.; Yan, T.; Shively, D.A.; Byappanahalli, M.N.; Whitman, R.L.; Sadowsky, M.J. 2006(a). *Cladophora* (Chlorophyta) spp. harbor human bacterial pathogens in nearshore water of Lake Michigan. Appl. Environ. Microbiol. 72 (7), 4545–4553.

Ishii, S.; Ksoll, W.B.; Hicks, R.E.; Sadowsky, M.J. 2006(b). Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), 612–621.

Janssen, P. J; Audit, B.; Ouzounis, C. A. 2001. Strain-specific genes of *Helicobacter pylori*: distribution, function and dynamics. Nucleic Acids Res. 29:4395–4404.

Jenkins, M. B.; Hartel, P. G.; Olexa, T. J.; Stuedemann, J.A. 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. J. Environ. Qual. 32:305–309.

Johnson, L.K.; Brown, M.B.; Carruthers, E.A.; Ferguson, J.A.; Dombek, P.E.; Sadowsky, M.J. 2004. Sample size, library composition, and genotypic diversity among natural populations of Escherichia coli from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70 (8), 4478–4485.

Kaspar, C.; Burgess, J.; Knight, I.; Colwell, R. R.1990. Can. J. Microbiol, 36, 891-894.

Kleinheinz, G.T.; McDermott, C.M.; Leewis, M.-C.; Englebert, E. 2006. Influence of sampling depth on *Escherichia coli* concentrations in beach monitoring. Water Res. 40, pp. 3831–3837.

Kreader, C.A. 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. 61, 1171–1179.

Larsen, R.E. 1996. Survival and transport of fecal pathogens in grazed watersheds, p.59-71. In: M.R. George (ed.) Livestock management in grazed watersheds. Univ. Calif. Div. Agric. and Nat. Resourc. Pub. 3381.

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. Appl. Environ. Microbiol. 72(6):4214-4224.

Lee, Yong-Jin; Molina, Marirosa; Santo Domingo, Jorge W.; Willis, Jonathan D.; Cyterski, Michael; Endale, Dinku M.; Shanks, Orin C. 2008. Temporal assessment of the impact of exposure to cow feces in two watersheds by multiple host-specific PCR assays. Appl. Environ. Microbiol. 74: 6839-6847

Lemarchand, K., and Lebaron, P. 2003. Occurrence of *Salmonella spp*. and *Cryptosporidium spp.* in a French coastal watershed: relationship with fecal indicators. FEMS Microbiol. Lett. 218,203–209.

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

Liu, L.; Spilker, T.; Coenye, T.; LiPuma, J. J. 2003. Identification by subtractive hybridization of a novel insertion element specific for two widespread *Burkholderia cepacia* genomovar III strains. J. Clin. Microbiol. 41:2471–2476.

Martinez-Urtaza, J.; Saco, M.; de Novoa, J.; Perez-Piñeiro, P.; Peiteado, J.; Lozano-Leon, A.; Garcia-Martin, O. 2004. Influence of environmental factors and human activity on the presence of *Salmonella* serovars in a marine environment. Appl. Environ. Microbiol. 70 (4), pp. 2089–2097.

McCambrdge, J.; McMeekin, T.A. 1981. Effect of solar radiation and predacious microorganisms on survival of fecal and other bacteria. Appl. Environ. Microbiol. 41 (5), pp. 1083–1087.

Minnesota Pollution Control Agency. 2005. Using Transparency Tube and total suspended solids data to assess stream turbidity. wq-csm1-01. http://www.pca.state.mn.us/publications/wq-csm1-01.pdf.

Minnesota Pollution Control Agency. 2009. Bacteria TMDL protocols and submittal requirements. wq-iw1-08 revised 03-09.

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

Novotny, V.; Sung, H.-M.; Bannerman, R.; Baum, K. 1985. Estimating nonpoint pollution from small urban watersheds. Journal WPCF 57, 339-348.

Olapade, O.A.; Depas, M.M.; Jensen, E.T.; McLellan, S.L. 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Appl. Environ. Microbiol. 72 (3), 1932–1938.

Parveen, S.; Murphree, R.L.; Edminston, L.; Kaspar, C.W.; Portier, K.M.; Tamplin, M.L. 1997. Association of multiple-antibioticresistance profiles with point and nonpoint sources of *Escherichia coli* in Apalachicola Bay. Appl. Environ. Microbiol. 63, 2607–2612.

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

Pruss, A. 1998. Review of epidemiological studies on health effects from exposure to recreational water. Int. J. Epidemiol.27 (1), 1–9.

Resnick, I.G.; Levin, M.A. 1981. Assessment of *Bifidobacteria* as indicators of human fecal pollution. Appl. Environ. Microbiol.42 (3), 433–438.

Reznikoff, William S. 2003. Tn5 as a model for understanding DNA transposition. Molecular Microbiology 47 (5): 1199-1206.

Rhodes, M.W.; Kator, H. 1988. Survival of *Escherichia coli* and *Salmonella spp*. in estuarine environments. Appl. Environ. Microbiol. 54 (12), pp. 2902–2907.

Roszak, D.B.; Grimes, D.J.; Colwell, R.R. 1984. Viable but nonrecoverable stage of *Salmonella enteritidis* in aquatic systems. Can. J. Microbiol. 30 (3), pp. 334–338.

Sadowsky, M. J.; Tully, R. E.; Cregan, P. B.; Keyser, H. H. 1987. Genetic diversity in *Bradyrhizobium japonicum* serogroup 123 and its relation to genotype-specific nodulation of soybeans. Appl. Environ. Microbiol. 53:2624–2630.

Samadpour, M.; Roberts, M.C.; Kitts, C.L.; Mulugeta, W.; Alfi, D. 2005. The use of ribotyping and antibiotic resistance patterns for identification of host sources of *Escherichia coli* strains. Lett. Appl. Microbiol. 40, 63–68.

Sambrook, J.; Fritsch, E. F.; Maniatis, T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sanchez, P.S.; Agudo, E.G.; Castro, F.G.; Alves, M.N.; Martins, M.T. 1986 Evaluation of the sanitary quality of marine recreational waters and sands from beaches of the Sao Paulo State, Brazil. Water Sci. Technol. 18, pp. 61–72.

Schallenberg, M.; Bremer, P. J.; Henkel, S.; Launhardt, A.; Burns, C. W. 2005. Survival of *Campylobacter jejuni* in water: effect of grazing by the freshwater crustacean *Daphnia carinata* (Cladocera). Appl. Environ. Microbiol. 71:5085–5088.

Scott, T. M.; Parveen, S.; Portier, K. M.; Rose, J. B.; Tamplin, M. L.; Farrah, S. R.; Koo, A.; Lukasik, J. 2003. Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. Appl. Environ. Microbiol. 69:1089–1092.

Shanks, O.C.; Santo Domingo, J.W.; Lamendella, R.; Kelty, C.A.; Graham, J.E. 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72 (6), 4054–4060.

Shanks, O.C.; Domingo, J.W.; Lu, J.; Kelty, C.A.; Graham, J.E. 2007. Identification of bacterial DNA markers for the detection of human fecal pollution in water. Appl. Environ. Microbiol. 73 (8), 2416–2422.

Shanks, O. C.; Atikovic, Emina; Blackwood, A.; Denene, Lu; Jingrang, Noble; Rachel T.; Domingo, J.S.; Seifring, Shawn; Sivaganesan, Mano; Haugland, Richard A. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 74: 745-752

Sharma, S.; Sachdeva, P.; Virdi, J. S. 2003. Emerging water-borne pathogens. Appl. Microbiol. Biotechnol. 61:424–428.

Shatti, J.A.; Abdullah, T.H.A. 1999. Marine pollution due to wastewater discharge in Kuwait. Water Sci. Technol. 40 (7), pp. 33–39.

Simmons, G.; Waye, D.; Herbein, S.; Myers, S.; Walker, E. 2000. Presented at Virginia Water Research Symposium. Blacksburg, VA, Younos, T., Poff, J., Eds.; VWRRC Special Report SR-19-2000.

Solo-Gabriele, H.M.; Wolfert, M.A.; Desmarais, T.R.; Palmer, C.J. 2000. Sources of *Escherichia coli* in a coastal subtropical environment. Appl. Environ. Microbiol. 66 (1), 230–237.

Soule, M.; Kuhn, E.; Loge, F.; Gay, J.; Call, D. R. 2006. Using DNA microarrays to identify library-independent markers for bacterial source tracking. Appl. Environ. Microbiol. 72:1843–1851.

Souza, V.; Rocha, M.; Valera, A.; Eguiarte, L.E. 1999. Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. Appl. Environ. Microbiol. 65 (8), 3373–3385.

Stevik, T.K.; Aa, K.; Ausland, G.; Hanssen, J.F. 2004. Retention and removal of pathogenic bacteria in wastewater percolating through porous media: a review. Water Res. 38 (6), pp. 1355–1367.

Stoeckel, D.M.; Mathes, M.V.; Hyer, K.E.; Hagedorn, C.; Kator, H.; Lukasik, J.; O'Brien, T.L.; Fenger, T.W.; Samadpour, M.; Strickler, K.M.; Wiggins, B.A. 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. 38 (22), 6109–6117.

Stoeckel, D.M., and Harwood, V.J. 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. 73 (8), 2405–2415.

Szewzyk, U.; Szewzyk, R.; Manz, W.; Schleifer, K. H. 2000. Microbiological safety of drinking water. Annu. Rev. Microbiol. 54:81–127.

Tate, K.W.; Atwill, E.R.; McDougald, N.K.; George, M.R. 2003. Spatial and Temporal Patterns of Cattle Feces Deposition on Rangeland. J. Range Manage. Vol. 56, No. 5 pp. 432-438

Trevisan, D.; Vansteelant, J. Y.; Dorioz, J. M. 2002. Survival and leaching of fecal bacteria after slurry spreading on mountain hay meadows: consequences for the management of water contamination risk. Water Res., 36, 275-283.

U.S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA-A440/5-84-002. U.S.Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2000. The quality of our Nation's waters; EPA-841-S-00-001. U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2001. Protocol for Developing Pathogen TMDLs. EPA 841-J-00-002. 132 pp. Office of Water (4503F). U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2002. Method 1603—*Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar. EPA821-R-02-23, 9. U. S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency. 2005. Microbial Source Tracking Guide Document. EPA-600/R-05/064, 131pp. Office of Research and Development, Washington, DC.

U.S. Fish and Wildlife Service. 2009 (copyright). The Federal Water Pollution Control Act (Clean Water Act). www.fws.gov/laws/lawsdigest/FWATRPO.html.

Versalovic, J.; Koeuth, T.; Lupski, R. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19 (24) (1991), pp. 6823–6831.

Wade, T.J.; Pai, N.; Eisenberg, J.N.S.; Colford, J.M.J. 2003. Do US EPA water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environ. Health Perspect. 111, 1102–1109.

Whitman, R.L. 1983. The influence of a combined sewage outfall on the chemical and biological quality of Salt Creek, Porter County, Indiana. Proc. Indiana Acad. Sci. 92, pp. 275–282.

Whitman, R. L.; Horvath, T. G.; Goodrich, M. L.; Nevers, M. B.; Wolcott, M. J.; Haack, S. K. 2001. Characterization of E. coli levels at 63rd Street Beach. Report to the City of Chicago Department of the Environment and the Chicago Park District. [Online.] U.S. Geological Survey, Porter, IN. http://www.glsc.usgs.gov/_files/reports/63rdStreetBeachReport.pdf.

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. Appl. Environ. Microbiol. 69, 4714–4719.

Whitman, R.L., and Nevers, M.B. 2003. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. Appl. Environ. Microbiol. 69 (9), pp. 5555–5562.

Whitman, R.L.; Byers, S.E.; Shively, D.A.; Ferguson, D.M.; Byappanahalli, M. 2005. Occurrence and growth characteristics of *Escherichia coli* and enterococci within the accumulated fluid of the northern pitcher plant (Sarracenia purpurea L.). Can. J. Microbiol. 51 (12), 1027–1037.

Wiedenmann, A.; Kruger, P.; Dietz, K.; Lopez-Pila, J.M.; Szewzyk, R.; Botzenhart, K. 2006. A randomized controlled trial assessing infectious disease risks from bathing in fresh recreational waters in relation to the concentration of *Escherichia coli*, intestinal enterococci, *Clostridium perfringens*, and somatic coliphages. Environ. Health Perspect. 114 (2), 228–236.

Wiggins, B.A.; Andrews, R.W.; Conway, R.A.; Corr, C.L.; Dobratz, E.J.; Dougherty, D.P.; Eppard, J.R.; Knupp, S.R.; Limjoco, M.C.; Mettenburg, J.M.; Rinehardt, J.M.; Sonsino, J.; Torrijos, R.L.; Zimmerman, M.E. 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. 65, 3483–3486.

Wiggins, B.A.; Cash, P.W.; Creamer, W.S.; Dart, S.E.; Garcia, P.P.; Gerecke, T.M.; Han, J.; Henry, B.L.; Hoover, K.B.; Johnson, E.L.; Jones, K.C.; McCarthy, J.G.; McDonough, J.A.; Mercer, S.A.; Noto, M.J.; Park, H.; Phillips, M.S.; Purner, S.M.; Smith, B.M.; Stevens, E.N.; Varner, A.K. 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. 69, 3399–3405.

Winfield, M. D., and Groisman, E. A. 2003. Role of non-host environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl. Environ. Microbiol. 69:3687–3694.

Yan, T; Hamilton, M.J.; Sadowsky, M.J. 2007. High-Throughput and quantitative procedure for determining sources of *Escherichia coli* in waterways by using host-specific DNA marker genes. Appl Environ Microbiol. 73(3): 890–896.