Project Report to Minnesota Department of Agriculture

Growth, Survival, and Genetic Structure of E. coli found in Ditch Sediments and Water at the Seven Mile Creek Watershed

2008-2010

Michael Sadowsky[1] and Scott Matteson[2]

Matt Hamilton[3] and Ramyavardhanee Chandrasekaran[1]

[1]BioTechnology Institute, and [3]Microbiology Department, University of Minnesota, St. Paul, MN; and [2]Water Resource Center, Minnesota State University, Mankato MN

TABLE OF CONTENTS

OVERVIEW ABSTRACT

Fecal contamination of waterways is a widespread public health problem. *Escherichia coli* is currently used as an indicator of fecal contamination in freshwater systems. The presence of *E. coli* in water is thought to indicate recent fecal contamination and in turn signals the possible presence of pathogens. Several studies have reported the presence of naturalized *E. coli* strains that persist and potentially grow in the environment. This confounds the use of this bacterium as an indicator organism. In the current study, we examined the fecal inputs and spatial and temporal distribution of *E. coli* in water and sediments of the Seven Mile Creek (SMC), a small man-made waterway in Nicollet County, MN. Results of this study indicated that *E. coli* counts varied considerably across sites and by dates and were likely affected by seasonal parameters, such as temperature and rainfall. Host specific PCR assays indicated that cattle were likely major contributors to the fecal loading of the SMC, although swine and poultry fecal markers were also detected sporadically. HFERP DNA fingerprint analysis indicated that the *E. coli* populations present in SMC were very diverse, but consisted of both transient and persistent strains. Some of the persistent strains appeared to be naturalized to the environment, particularly in the sediments. Multivariate analysis of variance (MANOVA) showed that water and sediment isolates from a given year clustered together suggesting mixing of *E. coli* strains in the sediment and water column. *E. coli* populations, however, shifted from year to year. Isolates obtained during times of flow conditions clustered together, suggesting that mixing and transport between sites occurs during conditions of flow. In addition, isolates from no flow conditions clustered into distinct groups. Furthermore, isolates from late spring to early fall clustered together at each site suggesting that temperature and growth of naturalized strains are likely

factors affecting the dynamics of population in SMC. Taken together, results of this study suggest that both newly acquired and indigenous *E. coli* strains are present in the SMC and this has obvious implications for water quality monitoring programs and for TMDL determinations.

INTRODUCTION

A large number of water bodies in the United States are considered to be impaired because of contamination by pathogenic microbes. As most enteric pathogens are transmitted through the fecal-oral route, fecal pollution is generally regarded as the major contributor of pathogens to waterways. Sources of fecal contamination include sewage and septic systems, livestock feedlots, wildlife, and run off from urban and agricultural land (17,21). There are some technical difficulties involved in directly monitoring pathogens. Each type of waterborne pathogen requires a specific test method, and some pathogens are difficult to identify in water systems. Therefore, water monitoring programs assess the level of contamination at a given site by determining numbers of fecal indicator bacteria (FIB), such as *Escherichia coli* and *Enterococcus* sp. The presence of these microorganisms in waterways is hypothesized to be due to contamination from human feces and indicate the likely presence of fecal pathogens, such as *Salmonella*, *Shigella*, and gastrointestinal viruses. Several studies have shown that elevated counts of FIB correlate with an increased risk of gastrointestinal disease after contact with contaminated water (6, 11, 14, 37). Thus, in the state of Minnesota and in most freshwater systems, *E. coli* are frequently used as the indicator organism for fecal contamination.

Monitoring of FIB can only determine the degree of contamination and not the host species responsible for it. Enteric pathogens are derived from both human and animal sources (*Salmonella*, *E. coli* O157:H7). However, risk from human feces is considered to be significantly higher due to presence of human viruses, which are highly host-specific. Hence, identifying the host source can help estimate human health risks associated with exposure to contaminated water. Furthermore, according to current regulatory requirements, waste loads need be allocated

to the potential contributing sources to establish total maximum daily loads (TMDL) for impaired water bodies.

Historically, *E. coli* have generally been thought of as a harmless commensal organisms found in the lower gut of warm blooded animals, although several pathogenic *E. coli* strains, such as EPEC or STEC, have been isolated (27). The use of *E. coli* as an indicator organism relies partly on the notion that it does not persist or grow in the environment, thus signaling the presence of recent contamination (15). However, our initial studies have shown that some soil-borne *E. coli* are genetically-distinct from those isolated from known- source animals and identical soil-borne *E. coli* genotypes are present in soils over a period of two years. Several studies, have reported on the persistence and potential growth of *E. coli* in water, sediment, and soil ecosystems, and in association with macrophytic algae (6-8, 16, 17, 18, 19, 23, 24, 31, 36). Such strains are termed as," naturalized *E. coli*". Results from several studies have shown that specific strains of *E. coli* have become naturalized to or indigenous to these environments, and these strains may contribute to elevated *E. coli* counts through the inoculation of water by runoff of *E. coli* from soils and sands, and through sediments mixing into the water column (17, 20, 36). The presence of indigenous *E. coli* in sediments and water confounds the use of this bacterium as an indicator of fecal pollution and may misdirect abatement strategies developed for TMDLs. This further necessitates the need for a better understanding of the sources of fecal contamination and their fate and ecology in the environment. The focus of our current project is to help determine the sources of bacterial contamination in ditches, and their survival and growth in the environment.

Both seasonal and weather-related factors likely play large roles in the population dynamics of naturalized *E. coli,* as the ability to persist and grow in natural environments is affected by changes in temperature, moisture, salinity, organic matter content and predation (4, 5, 32, 33, and 34). Several studies have shown that naturalized *E. coli* strains can survive winter temperature extremes and freeze thaw cycles, as well as the summer months in northern temperate climate soils (6, 18). The transport of naturalized strains from soils and to water and the re-suspension of stream sediment-borne *E. coli* during high flow periods has also been documented (20).

In light of the studies described above, in the study reported here we examined the spatial and temporal changes in *E. coli* populations present in sediments and water in the Seven Mile Creek (SMC), a small constructed waterway in Nicollet County, MN. The purposes of this study was to: i) examine the spatial and temporal distribution of *E. coli* in water and sediments of the SMC, ii) examine the persistence and transport of these bacteria in the SMC, iii) use DNA fingerprinting analyses to examine the genetic structure of *E. coli* populations obtained from water and sediment samples to determine if these bacteria are likely growing in this environment or if they arise as a result of new inputs, and iv) determine potential sources of fecal contamination in the SMC by using molecular biological tools.

MATERIALS AND METHODS

Study Site Description

The site chosen for this study was a subwatershed of the Seven Mile Creek (SMC) in Nicollet County, MN (Figure 1). The seven Mile creek watershed covers nearly 36.8 sq miles in Nicollet County about 14 miles from Mankato. The 24,551 acre watershed consists mainly of 86% agricultural land. About 20% of the watershed receives manure fertilizer each year. Other potential sources of fecal bacteria within the SMC watershed include 24 animal feedlots for cattle and swine. Most of the drainage entering the creek itself comes from three constructed ditches and two tile systems. The drainage ditches within watershed have intermittent flow and often do not contribute water to the creek after July. Water and sediment samples were collected for analysis from four sites within the SMC watershed, designated as SM1 - SM4 (Figure 1).

Sample Collection and Processing:

Samples were collected at all four sites (SM1-SM4). The sampling period is shown in the table below.

SAMPLING YEAR	SAMPLING PERIOD
2008	July through October
2009	April through October
2010	April through June

Water samples were collected as described in the USEPA microbiology methods manual (Bordner *et al.,* 1978, Section II, A). Sediment samples were collected 2 cm below the water column. Samples were shipped on ice to the laboratory after collection for processing. Bacteria present in the samples were concentrated by membrane filtration prior to enumeration and DNA extraction. Approximately 200 to 600 ml of water from each sample was filtered through a 0.45 µm (47mm) cellulose ester membrane (Millipore, Billerica, MA). Sediment samples (10 g) were suspended in 95 ml of 0.1 M sodium phosphate buffer amended with 0.1% hydrolyzed gelatin (22) in a milk dilution bottles containing 10 g of 3 mm glass beads. Samples were vigorously agitated using a wrist action shaker for 30 minutes to release bacteria bound to sediment particles. Samples were allowed to settle for 30 min and 25-80 ml of the upper phase was filtered through 0.45 µm membranes as described above. Water and sediment filtrations were done in triplicate, or greater. Membrane filters for DNA extraction were stored frozen at -80C until further use. Filters for colony isolation were placed into a conical centrifuge tube containing 10 ml of phosphate buffered saline and 5 g of 3 mm glass beads. Bacteria were removed from the filters by gentle agitation for 30 min using a wrist action shaker. Glycerol was added to a final concentration of 10%, and bacterial preparations were stored frozen at -80C until use.

Enumeration of *E. coli*

Counts of *E. coli* in water and sediment samples were obtained by using the Colilert® Quanti-tray 2000®, a most probable number (MPN) based analysis system, according to the manufacturer's instructions. Count data are expressed as MPN / 100 ml or MPN / g sediment (dry weight) for water and sediment samples, respectively.

E. coli isolation

E. coli were isolated from concentrated water and sediment filter wash samples using modified mTEC medium as previously described (35, 39). Approximately 2-3 ml of each filter wash sample was spread-plated onto the surface of mTEC medium in 20 x 20cm Q-tray bioassay plates (Genetix, Boston, MA). Each plate contained 250 ml of modified mTEC medium. Plates were incubated at 37^{o} C for 2 hr, then at 42^{o} C for 16 hr. After incubation, plates were stored at 4^{o} C overnight to facilitate the development of blue pigment in colonies, allowing for the differentiation of *E. coli* from other fecal coliform bacteria. Twenty-four well isolated blue colonies were hand-picked into 96 well microtiter plates containing 150 µl of Hogness modified freezing medium (HMFM). In some cases, less than 24 isolates were collected due to low numbers of *E. coli* present. *E. coli* were not collected from sediment samples obtained in April and May 2009 because *E. coli* was not detected in these samples. Microtiter plates were incubated at 37^{o} C overnight then stored frozen at -80C before use.

DNA Extraction

Total DNA was extracted from frozen membrane filters using a PowerSoil DNA kit (MO BIO, Carlsbad, CA) with slight modifications. Filters were finely chopped using a sterile razor blade and used as the input material for DNA extraction. Total DNA was eluted using 100 µl of sterile, nuclease free water. DNA concentrations were determined using an Eppendorf BioPhotometer (Eppendorf, New York, NY). DNA samples were diluted to 3 ng/µl in sterile, nuclease free water and were used as templates in PCR reactions.

PCR analyses of environmental DNA samples

PCR analyses were done to detect the presence of DNA from *Bacteriodes* sp. and *Brevibacterium* sp. strains. The primer pairs HF183F/Bac708R, PF163F/Bac708R, LA35F/LA35R, and CowM3F/CowM3R were used to detect the presence of *Bacteroides* associated with humans, swine, poultry, and cattle, respectively (3, 12, 30). The primer pair LA35F/LA35R were used to detect the presence of poultry-specific *Brevibacterium* sp. (V. J. Harwood, unpublished data). All PCR reactions were performed as previously described. Each reaction used 15 ng of environmental DNA as template. Positive control reactions were performed using template DNA isolated from human sewage or the feces of swine, chickens, or cattle. Negative controls consisted of PCR reactions without added template DNA. All samples were also tested using a universal Bacteroides primer pair, AllBac296F/Bac708R, to test for PCR inhibitors that may be present in the extracted DNA sample (25).

HFERP DNA fingerprinting

Horizontal, fluorophore enhanced, repetitive element PCR was performed using BOXA1R primers as previously described (21). Gel images were captured using a Typhoon 8600 Variable Mode Imager (GE Healthcare, Chalfont St. Giles, UK) and analyzed using Bionumerics (version 2.1) software (Applied Maths, Kortrijk, Belgium) as previously described (13, 21).

Statistical Analysis

Dendrograms were constructed from the HFERP fingerprint data using the curve-based, Pearson's product-moment correlation coefficient and the unweighted pair group method with arithmetic means (UPGMA) clustering method. Isolates sharing 92% similarity or greater were defined as clonal (21). The Shannon index of diversity was calculated as previously described (9). The HFERP fingerprint data was also used to generate a binary band-matching character table which was analyzed using multivariate analysis of variance (MANOVA) to cluster isolates (9, 13, 19) and to evaluate the significance of the discriminant analysis. There is a P value and an L value associated with each MANOVA analysis. The P value represents the probability of obtaining same level of separation by random grouping of isolates and L value determines if the group assignment was significant or not. Further, Jackknife analysis was performed on these isolates to determine the rates of correct assignment. ID bootstrap analysis (at P = 0.9), done using a Bionumerics script (http://www.applied-maths.com/bn/scripts/bnscripts.htm), was performed to identify the potential sources of *E. coli* isolates in SMC (13, 19).

RESULTS

E. coli counts

The number of *E. coli* in water and sediments samples at the Seven Mile Creek site was evaluated from mid-summer to fall in 2008, from spring to fall in 2009 and from spring to mid-summer in 2010. The *E. coli* counts for individual water and sediment samples varied greatly by site and by date (Figure 2)**.** Relative to current numbers, in 2008, *E. coli* counts were consistently elevated. Of the 52 water samples collected in 2008, 33 (63.5%) exceeded the state standard of 126 CFU/100 ml. Sediment samples from site SM1 had higher counts than SM2, SM3, and SM4 for 8 of the 12 (75%) sampling dates in 2008. While *E. coli* were detected in all sediment samples at varying levels through the summer months, several samples from fall 2008 did not contain culturable *E. coli*. Moreover, samples from spring 2009 generally had lower counts than those from summer 2008 and in 2009, the number of *E. coli* in water and sediment samples began to increase in late spring and early summer. Forty two of the 116 (36.2%) water samples from 2009 and only 8 out of 64 (12.5%) water samples from 2010 exceeded the state standard and all of the samples that exceeded were obtained from May or later. *E. coli* was not detected in several of the sediment samples, especially those from earlier dates in 2009 and 2010. Further, in 2009, the site SM3 was almost dry during the month of August and sites SM3 and SM4 were dry towards the end of September. Hence, water and sediment samples from SM3 during August and water samples from SM3 and SM4 during the end of September for the year 2009 could not be obtained for PCR- based 16S rRNA and HFERP fingerprinting analysis.

Host Source-Specific PCR

PCR-based 16S rRNA analyses were used to examine water and sediment DNA samples from Seven Mile Creek for the presence host specific *Bacteroides* strains originating from human, bovine, swine, and poultry sources. This was done using host-source-specific DNA primers targeting these animal sources. Results of this analysis indicated that most samples (63%) from mid to late summer 2008 were positive for the bovine specific *Bacteriodes* marker. About 76% of the samples collected in 2009 and 96% of samples collected in 2010 were positive for the bovine specific Bacteriodes marker. In contrast, the human fecal marker gene was not detected in any sample tested from 2008-2010, and the markers indicative of poultry and swine fecal material were detected infrequently in the 2008, 2009 and 2010 samples. All samples were positive for a universal *Bacteroides* marker indicating that PCR inhibitors were not responsible for negative results. Host specific PCR results are summarized in Table 1.

HFERP DNA Fingerprinting to Determine Diversity of *E. coli* in the Watershed

DNA fingerprint analysis of *E. coli* strains collected from each sample was used to examine genetic diversity, population dynamics and persistence of *E. coli* strains at Seven Mile Creek. The analysis will also help in determining if growth of *E. coli* in water and sediments was involved in exceedance of state and national *E. coli* limits. Analysis of the dendrogram generated from fingerprint data obtained during 2008-2009 indicated that the *E. coli* populations present at the SMC were quite diverse (the dendrogram is not shown due to size constraints). Overall, fingerprint similarity ranged from 1.10 to 100%, and the isolates were grouped into three groups-

two of which were large and diverse containing 55 and 1512 isolates. The 1568 isolates examined were comprised of 452 different strains. Of these, 287(63.5%) were represented by a single isolate. The remaining 165 strains were represented by 2 to 66 isolates, and several of these strains were found repeatedly in samples from both 2008 and 2009 and across different sampling sites and types, suggesting that these strains persist at the sites. The Shannon diversity index for all isolates from 2008-2009 was 5.04. The Shannon diversity indices for water and sediment samples isolates from 2008 and 2009 were 4.43, 4.06, 4.59 and 4.04 respectively. This suggests that the number of genotypes in most samples are evenly distributed among all the genotypes present.

A dendrogram showing the genetic relatedness of one of these strains is shown in Figure 5. DNA fingerprint analysis of the SMC isolates were compared to those obtained from a library of known source *E. coli* from Minnesota using ID bootstrap analysis. None of the SMC isolates were identified as originating from a specific source using this method. MANOVA analysis of fingerprint data obtained from water and sediment isolates from 2008 to April 2010 was used to determine if isolates found at one site and during one year could be found at other sites during the same year and into the next season. Results in Figure 6 showed that isolates clustered well by year into overlapping groups of water and sediment isolates. The MANOVA explained 78% of the variation in two dimensions, showing that the isolates were fairly tightly clustered by dates and sites. The P value of the analysis was 0.001% suggesting that the probability of obtaining same level of separation by random grouping of isolates is negligible.

A MANOVA analysis was also used to compare water isolates obtained from conditions were flow was occurring to those found at each site under no flow condition (isolated sites)

found in 2008. This analysis showed that isolates obtained during flow conditions clustered together regardless of site, while no flow isolates clustered into distinct groups containing isolates from a single site, with the exception of no flow isolates from sites 3 and 4, which clustered together (Figure 7). The P and the L value of this analysis were low suggesting that the specified groups were valid. Further, Jacknife analysis of the HFERP fingerprint data obtained from low and no flow isolates provided a similar result, where low flow isolates were often assigned to other low flow groups (Table 2).

The water isolates from April to August 2009 were analyzed individually for each site (SM1-SM4) using MANOVA. Analysis indicates that the water *E. coli* isolates during early spring and late fall clustered into distinct groups whereas the isolates from late spring to early fall clustered together with the exception of site SM4 where they were all clustered into distinct groups (Figure 8). Due to extremely low *E. coli* counts, only a few isolates could be obtained from April and May for analyses. However, although the project ended in June 2010, we are continuing the analyses of fecal counts, source tracking data, and DNA fingerprints. This information will be included in our final publications. No additional data will be collected or analyzed after September 30, 2010.

DISCUSSION

In this study, we examined the population structure of *E. coli* strains present in water and sediment samples from the SMC. Count data indicated that the water at the SMC is highly contaminated with *E. coli*, especially during the summer and fall months (Figure 3). Sediment samples had the highest *E. coli* counts in the summer months, although count data indicated significant variability across sites and dates. Low temperatures may explain the failure to detect *E. coli* in several sediment samples from late fall 2008 and during early spring 2009 and 2010. In fall 2008, the average daily temperatures were around 15^{o} C near the SMC, whereas summer 2008 average temperature were consistently above 20^{o} F. Average daily temperatures were less than 4^{o} C for the earliest 2009 and 2010 sample dates and increased to around 20^{o} C for the summer months. These results are similar to those previously reported, where soil *E. coli* counts from several sites in northern Minnesota were highest in the summer months and decreased over the winter, presumably due to temperature effects (18). The inherent variability of environmental samples and the patchy distribution of *E. coli* in the environment (2, 9, 18, 28, 38) were also a likely contributors to count variability. It is also important to note that the *E. coli* enumeration method used in this study requires the use of a culture technique and, therefore, isolates in a viable but not culturable (VBNC) state would not be detected. Previous studies have suggested that the VBNC state is important for bacteria to survive extreme or unfavorable conditions (29) and may account for the survival of *E. coli* during low temperature conditions (18).

Host specific PCR assays for fecal bacteria showed that cattle were major contributors to the fecal loading of SMC (Table 1). More than 60% of samples from 2008 and more than 75% of samples from 2009 and 2010 were positive for a bovine-specific marker gene. The presence of several large scale cattle feedlots and manure amended agricultural fields throughout the watershed are likely sources for bovine fecal pollution which would reach the creek through run-off. The bovine marker was detected infrequently in the fall of 2008, which may be related to low water levels and a lack of run-off during this time.

About 13% of the samples obtained during 2008-2010 were positive for swine and only few samples from 2009 and 2010 had the poultry specific markers. Out of the 218 samples tested from 2008-2010, only 5 (<3%) were positive for poultry specific marker (Figure 4). Humans likely do not contribute to fecal loading at the SMC as evidenced by the lack of any positive samples. Since the vast majority the SMC subwatershed is agricultural land or otherwise undeveloped, the lack of the human-specific marker gene is not surprising. It is probable that wildlife and other sources also impact the fecal loading of SMC, although due to a lack of validated source tracking tools, we did not test for fecal bacteria from other sources. Thus, the percentages of isolates positives for the host sources tested do not add up to 100% (Table 3.)

Despite the prevalence of the bovine-specific fecal marker gene in the samples, none of the 1,568 *E. coli* isolated from SMC water and sediment samples was identified as originating from bovine using DNA fingerprint data and ID bootstrap analysis (at P = 0.9). In fact, none of the tested *E. coli* isolates matched with any of the known source *E. coli* isolates in a library obtained from cats, dogs, chickens, cows, deer, ducks, geese, goats, humans, pigs, sheep, and

turkeys. This suggests that the library used is likely not representative of *E. coli* strains found at this site (39). Failure to assign any of the *E. coli* collected in this study may be a result of temporal or geographic limitations of the known source library used for this analysis which was largely collected from 2000 to 2005 from across Minnesota (13, 21). A reference library containing *E. coli* isolates from livestock and wildlife obtained within the SMC watershed would be better than the current library, and may serve to identify some of the isolates as originating from sources within the watershed.

In addition to direct input of fecal-derived strains, *E. coli* inputs into SMC may also include soil and other environmental strains that enter the system via run-off from rain events. It has been previously shown that run-off from soil may contribute up to 19% of *E. coli* detected in some waterways (18). Unfortunately, rain events occurred sporadically through the sampling seasons and did not track well to our sampling dates; therefore, it is extremely difficult to gauge inputs from these sources. Obtaining additional samples collected immediately after run-off reaches the creek would be helpful to determine if the *E. coli* isolates collected after rain events are different than those present during periods without rain.

Results from HFERP DNA fingerprinting revealed the presence of very diverse *E. coli* populations at the SMC study site. The Shannon diversity indices for water and sediment isolates from 2008 and 2009 were much higher than those reported from the analysis of the ribotypes of *E. coli* isolated from cattle, horses, and humans (1). Nearly 63.5% of the strains were only detected once, suggesting that transient strains make up a large proportion of the *E. coli* found in SMC water and sediments. However, several strains were detected in multiple samples, across months and years, and at different sites, suggesting that some strains in SMC can

persist for extended periods of time or grow at the SMC sites. In some cases, especially in sediment samples, clonal isolates from a single strain were found many times in the same sample (Figure 5). Some sediment samples consisted almost entirely of isolates belonging to a single strain. This result suggests that these isolates were likely growing or colonizing in the sediment or water, as the isolation method does not have an enrichment step that would allow for single isolates to multiply. Growth in the sediments and water of stream and lake ecosystems has been reported in the past (9, 17), and the results reported here support the idea that *E. coli* can become naturalized to these environments.

The diversity amongst *E. coli* isolates from SMC was also seen using a MANOVA analysis comparing isolates from water and sediment from June 2008 to April 2010. Results of this analysis (Figure 6) showed that isolates clustered well with their respective groups and water and sediment isolates from a given year were more closely related to each other than to isolates from different years. The apparent similarity between sediment and water *E. coli* and the detection of identical strains from both sediment and water sources further supports the idea that mixing of sediments with the water column can contribute to elevated *E. coli* counts (20). Furthermore, these results indicate that while the total population of *E. coli* shifts from year to year, overlapping clusters provide further evidence that some strains can persist for extended periods.

An additional analysis of fingerprint data from 2008 using MANOVA showed that isolates obtained during flow conditions clustered together regardless of sampling site and those from no flow condition clustered into distinct groups (Figure 7). Jackknife analysis of the fingerprint data yielded similar results, where isolates from flow conditions were assigned to the

correct group at lower frequencies than those from no flow condition. These results suggests that mixing of the water column and transport of these isolates downstream during flow conditions were likely influencing the population structure of *E. coli* in the SMC. This result is similar to that of Jamieson et al. which also showed the resuspension of tracer *E. coli* strains from sediments and downstream transport (20). During no flow condition, water became stagnant at the sampling sites and evaporation caused the sites to no longer be physically connected to one another, which could have allowed for differentiation of *E. coli* populations at the respective sites. Interestingly, isolates from sites SM3 and SM4 clustered together during flow conditions. The reason for this is not clear, but this result may be related to similar inputs, as these sites are relatively close to each other in the watershed.

Further, MANOVA analysis of water isolates from April to August 2009 for each site (SM1-SM4) indicated that the water *E. coli* isolates during spring and late fall clustered into distinct groups whereas the isolates from late spring to early fall clustered together(with the exception of site SM4 where they were all clustered into distinct groups) (Figure 8). This result could be due to environmental conditions such as temperature, flow, additional inputs such as rainfall or variation in the area of sampling within each site. The weather history for Mankato obtained from Weather Underground shows that the temperature increased steadily from late spring to early fall, which favored the replication of *E. coli* within each site. This period was also marked by poor mixing capacity and accumulation of the water at each site due to no flow leading to the isolation of closely related *E. coli* population, irrespective of the location of sampling within the site. Whereas, the months of April, September and October were relatively cold and April is marked by transport downstream due to snow melt and run off causing the *E.*

coli population to be transient. Hence, *E. coli* population isolated from these months clustered into separate groups in the analysis. Further, the dendrogram constructed from all the isolates showed that some strains persisted at the same site and were detected predominantly during June to August (Figure 5). This suggests that, in the absence of new input, growth of these naturalized *E. coli* strains might be responsible for the closely related population found at a particular site during these months.

Table 1. Results of host species specific PCR assays for fecal bacteria originating from bovine, human, poultry and swine sources.

	Number of samples positive for host-specific markers[1]							
	Sediment				**Water**			
	Bovine Marker	**Human Marker**	**Poultry Marker**	**Swine Marker**	**Bovine Marker**	**Human Marker**	**Poultry Marker**	**Swine Marker**
7/23/2008	4 (100)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)
8/13/2008	3 (75)	0 (0)	0 (0)	0 (0)	3 (75)	0 (0)	0 (0)	0 (0)
8/27/2008	1 (25)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
9/5/2008	0 (0)	0 (0)	0 (0)	1 (25)	0 (0)	0 (0)	0 (0)	0 (0)
9/22/2008	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
10/7/2008	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)	0 (0)	0 (0)	1 (25)
4/1/2009	4 (100)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)
4/15/2009	4 (100)	0 (0)	2 (50)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)
4/29/2009	4 (100)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)
5/12/2009	4 (100)	0 (0)	0 (0)	1 (25)	4 (100)	0 (0)	0 (0)	1 (25)

5/26/2009	4 (100)	0 (0)	0 (0)	1 (25)	4 (100)	0 (0)	0 (0)	1 (25)
6/8/2009	4 (100)	0 (0)	0 (0)	1 (25)	4 (100)	0 (0)	0 (0)	1 (25)
6/23/2009	4 (100)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	0 (0)
7/7/2009	4 (100)	0 (0)	0 (0)	1 (25)	4 (100)	0 (0)	0 (0)	1 (25)
7/21/2009	4 (100)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	1 (25)
8/4/2009	2(67)	0(0)	0(0)	0(0)	2(67)	0(0)	0(0)	0(0)
8/18/2009	3(100)	0(0)	0(0)	0(0)	3(100)	0(0)	0(0)	1(33)
9/2/2009	1(25)	0(0)	0(0)	2(50)	3(75)	0(0)	0(0)	1(25)
9/14/2009	4(100)	0(0)	0(0)	0(0)	3(75)	0(0)	0(0)	0(0)
9/30/2009	4(100)	0(0)	0(0)	0(0)	2(100)	0(0)	0(0)	0(0)
10/14/2009	0(0)	0(0)	0(0)	0(0)	4(100)	0(0)	0(0)	1(25)
10/29/2009	2(50)	0(0)	0(0)	0(0)	3(75)	0(0)	0(0)	0(0)
4/5/2010	4(100)	0(0)	0(0)	0(0)	4(100)	0(0)	0(0)	2(50)
4/21/2010	4(100)	0(0)	1(25)	2(50)	4(100)	0(0)	1(25)	2(50)
5/6/2010	4(100)	0(0)	0(0)	2(50)	4(100)	0(0)	0(0)	0(0)
5/17/2010	4(100)	0(0)	0(0)	1(25)	2(50)	0(0)	0(0)	0(0)

6/2/2010	4(100)	0(0)	0(0)	0(0)	4(100)	0(0)	0(0)	1(25)
6/16/2010	4(100)	0(0)	0(0)	1(25)	4(100)	0(0)	0(0)	1(25)

[1] Values in parentheses are percentages.

Table 2. Jackknife analysis of water isolates from low flow and no flow conditions at the SMC sites. Values are expressed as percentages.

		Maximum similarities (%)							
		Low Flow				**No Flow**			
Group Assignment		**SM1**	**SM2**	**SM3**	**SM4**	**SM1**	**SM2**	**SM3**	**SM4**
low Flow	**SM1**	79.2	20.8	8.3	4.3	0	0	4.2	4.2
	SM2	12.5	33.3	29.2	8.7	0	0	0	0
	SM3	4.2	12.5	41.7	13	0	0	0	0
	SM4	0	8.3	8.3	56.5	8.3	4.2	0	0
No Flow	**SM1**	0	12.5	8.3	4.3	91.7	0	4.2	4.2
	SM2	0	0	0	8.7	0	95.8	4.2	4.2
	SM3	0	4.2	0	4.3	0	0	83.3	4.2
	SM4	4.2	8.3	4.2	0	0	0	4.2	83.3

Table 3. Summary of PCR results from 2008-2010.

A.

SAMPLE TYPE	TOTAL SAMPLES TESTED	SAMPLES NEGATIVE FOR ALL HOST SOURCES *	SAMPLES POSITIVE FOR ATLEAST ONE HOST SOURCE*
WATER	108	23 (21.3%)	85(78.7%)
SEDIMENT	110	24(21.8%)	86(78.2%)

Note: *- Host sources include only Human, Bovine, Swine and Poultry.

B.

SAMPLE TYPE	TOTAL SAMPLES TESTED	BOVINE	HUMAN	POULTRY	SWINE
WATER	108	84(77.7%)	0	1 (0.9%)	15 (13.9%)
SEDIMENT	110	84 (76.4%)	0	4 (3.6%)	13(11.8%)

Note: Some samples may be positive for more than one host source.

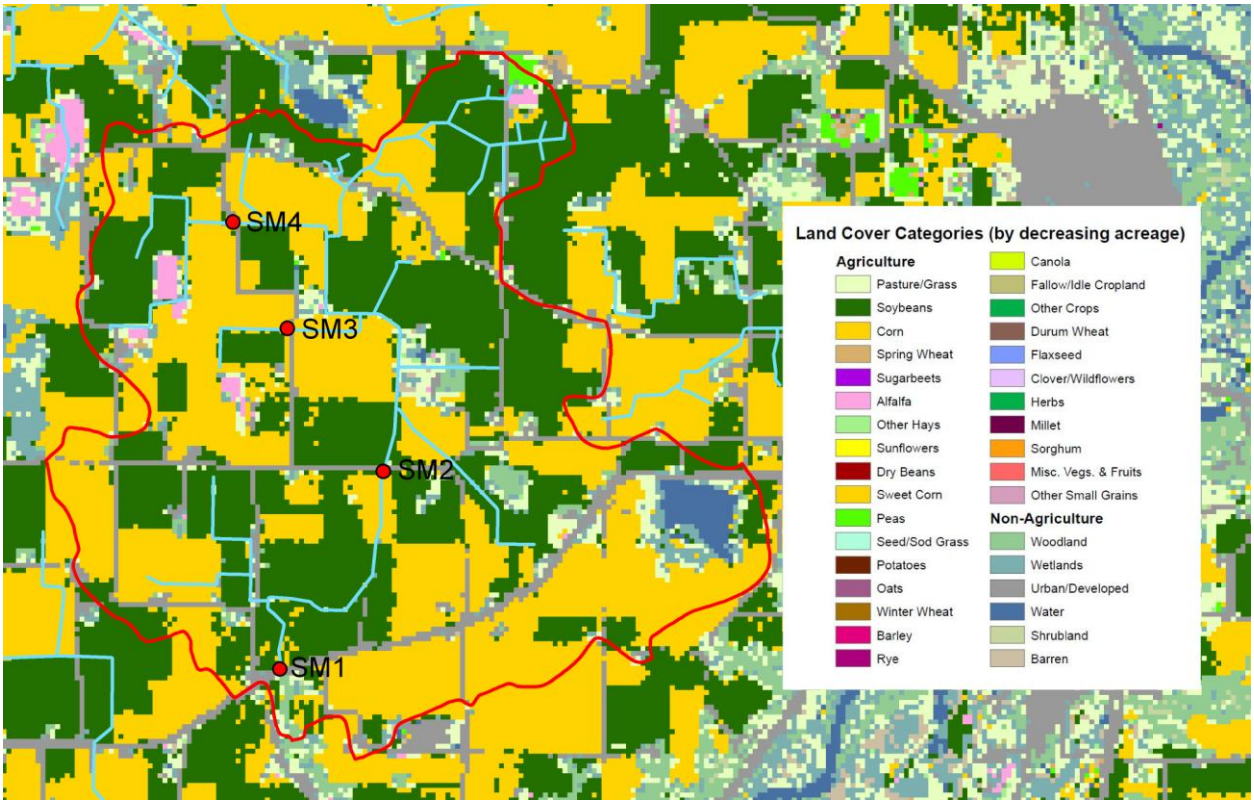
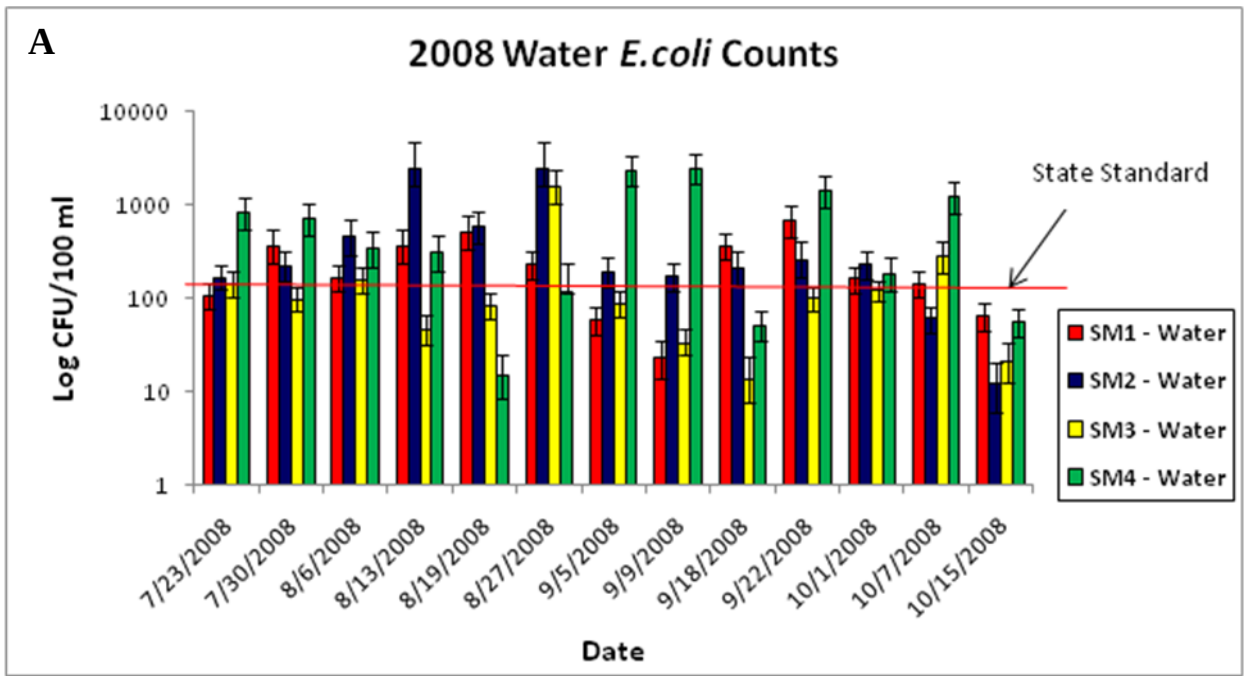


Figure 1. GIS land use map of the Seven Mile Creek subwatershed examined in this study. Samples sites, SM1-4 are denoted by yellow circles. The creek flows from SM4 towards SM1.



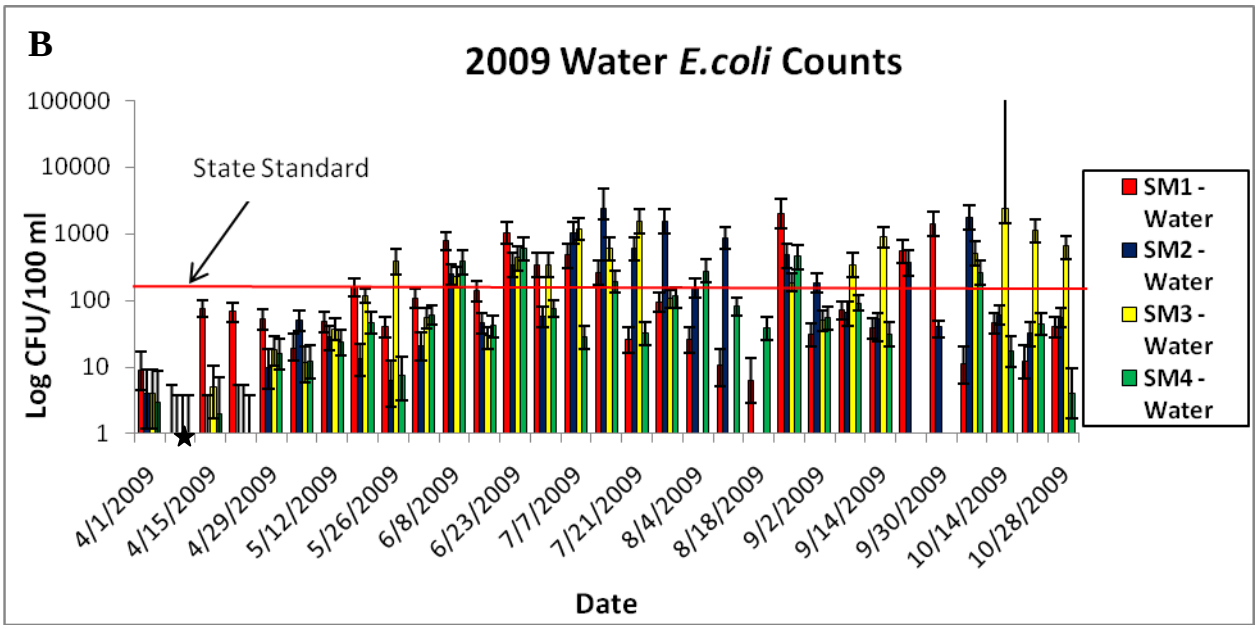
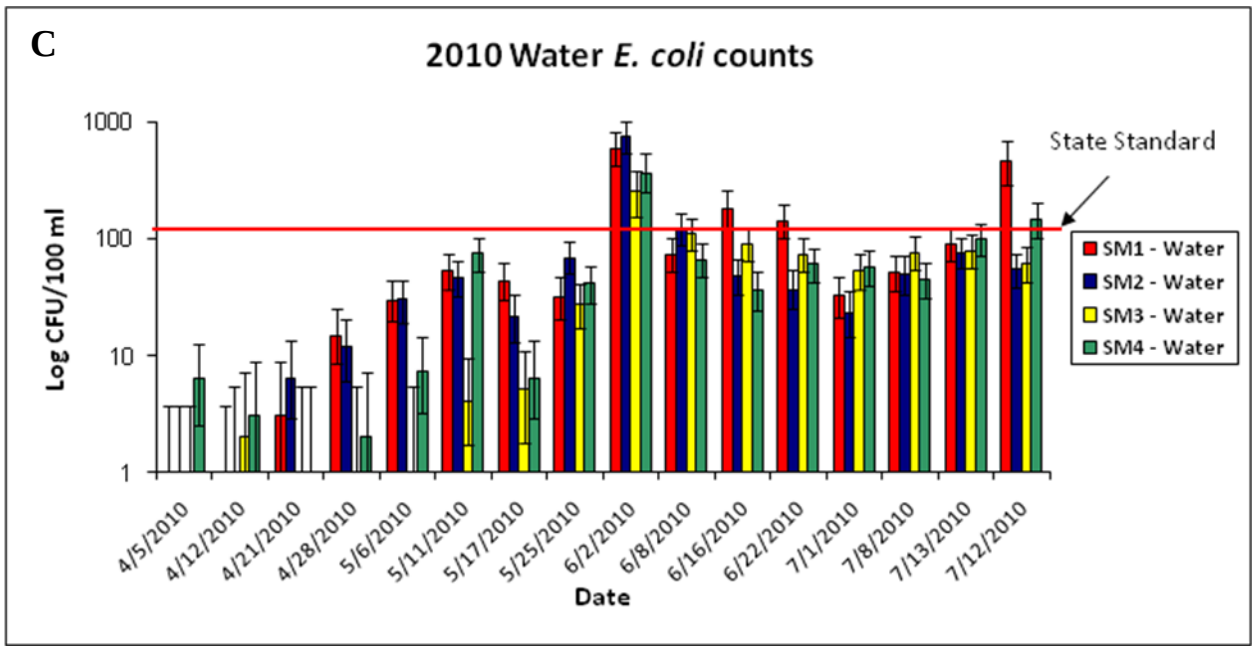


Figure 2 – continued



Note: The ★ denotes *E. coli* counts that were zero in samples.

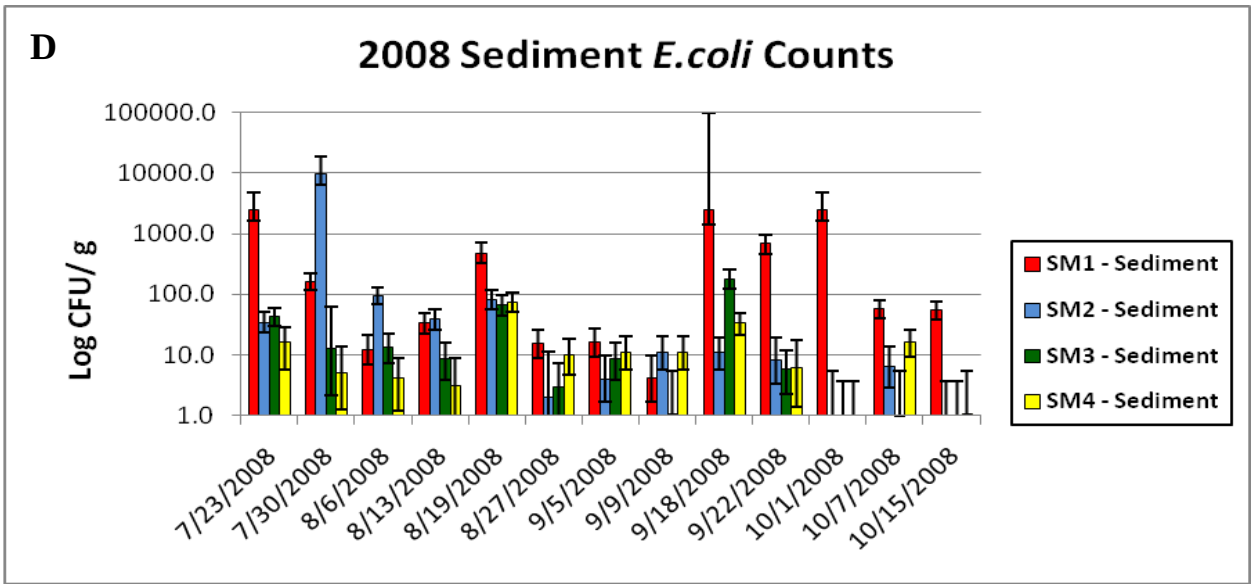
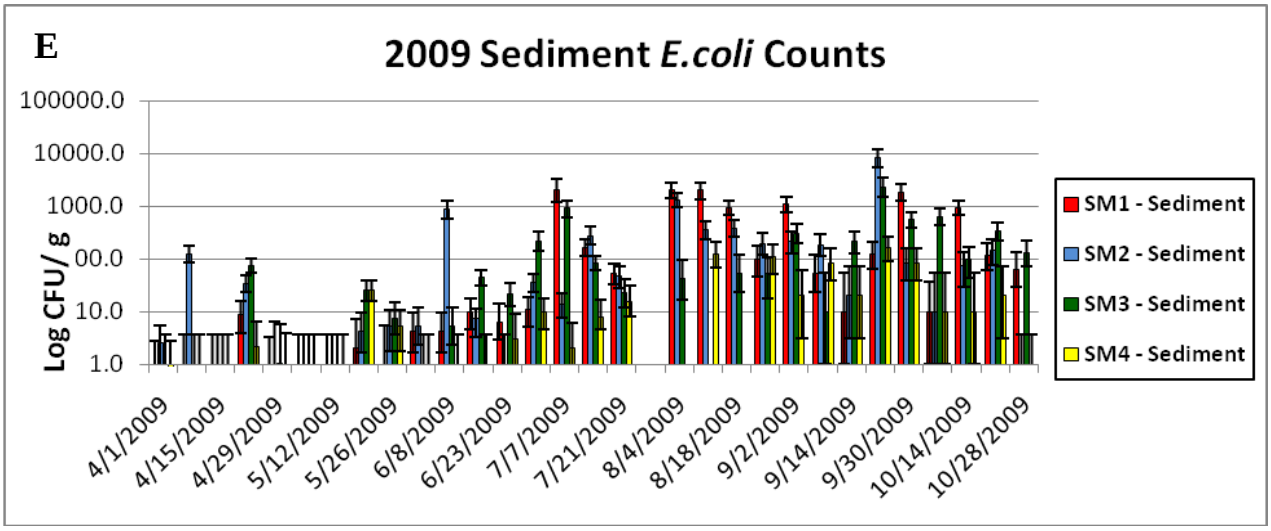
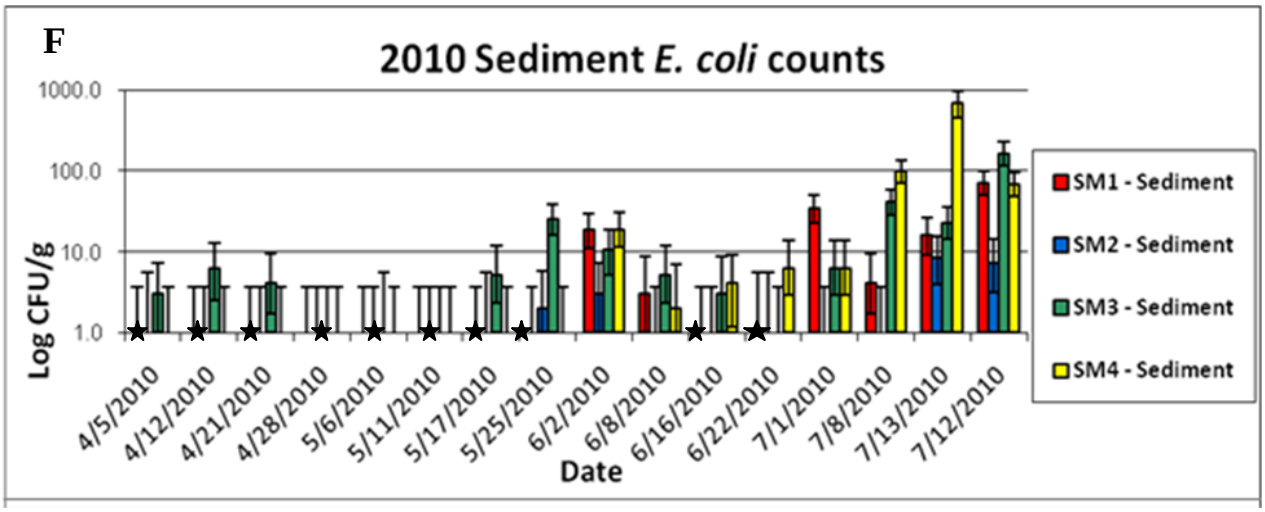


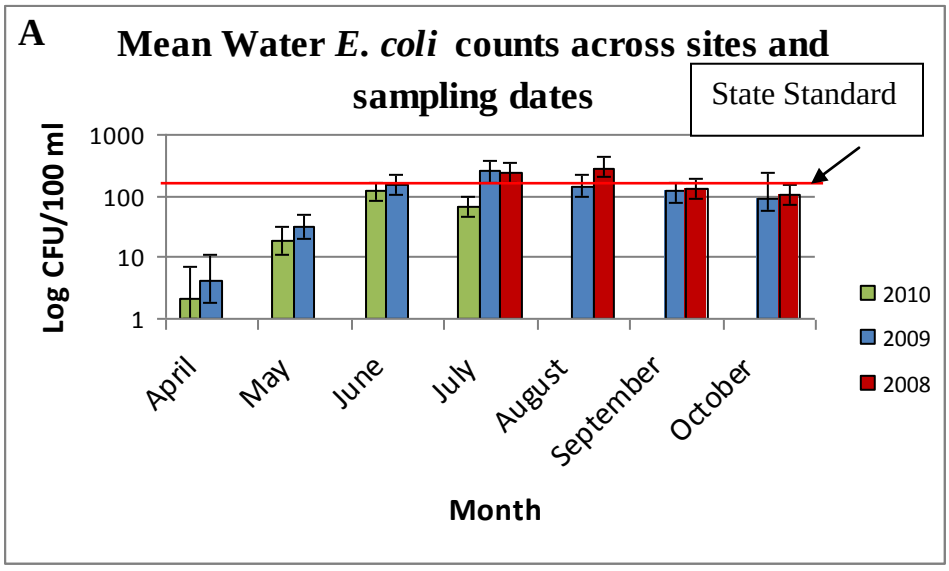
Figure 2 continued





Note: The ★ denotes *E. coli* counts that were zero in samples.

Figure 2. *E. coli* counts in water and sediment samples collected from July to October 2008, April to October 2009 and April to July 2010. Water samples from 2008, 2009 and 2010 are shown separately in panels A, B and C and sample sites SM1, SM2, SM3, and SM4 are represented as (■), (■), (■), (■), respectively. Sediment samples are shown in panels D, E and F and sample sites SM1, SM2, SM3, and SM4 are represented as (■), (■), (■), (■), respectively. The red line (——) represents the state standard of 126 CFU / 100 ml.



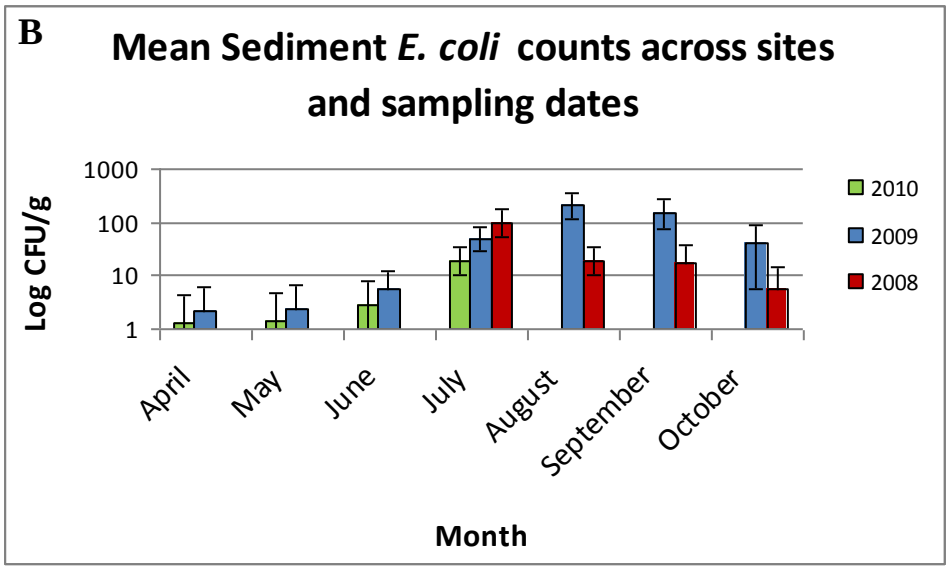


Figure 3. *E. coli* count data was averaged across sites and sampling dates for each month to emphasize the effect of temperature. 2008, 2009 and 2010 counts are shown as ( ), ( ) and ( ) respectively

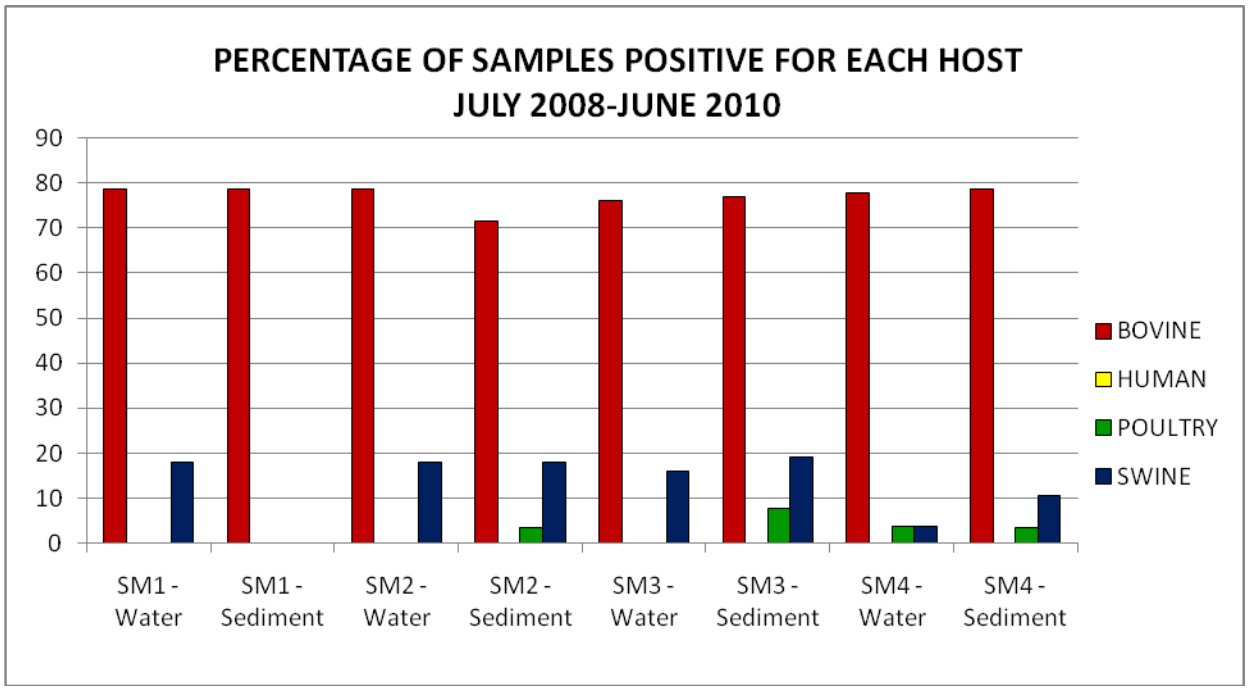


Figure 4. Percentages of samples positive for each host are represented across site (SM1, SM2, SM3 and SM4) and sample type (Water or Sediment). Color code is specific to the host.

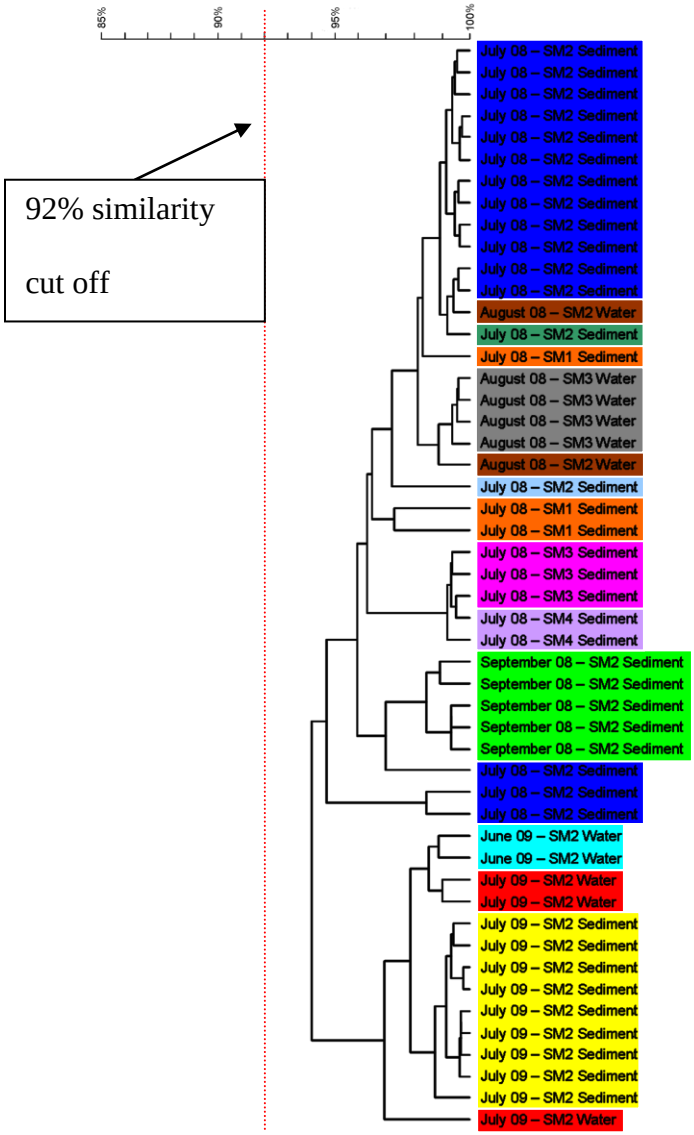


Figure 5. Dendrogram generated using HFERP DNA fingerprint data from isolates identified as a single strain. Similarity ranged from 95.67 to 99.83%. Sample date, location and type for individual isolate is shown on the right. Isolates from a single sample are grouped by color.

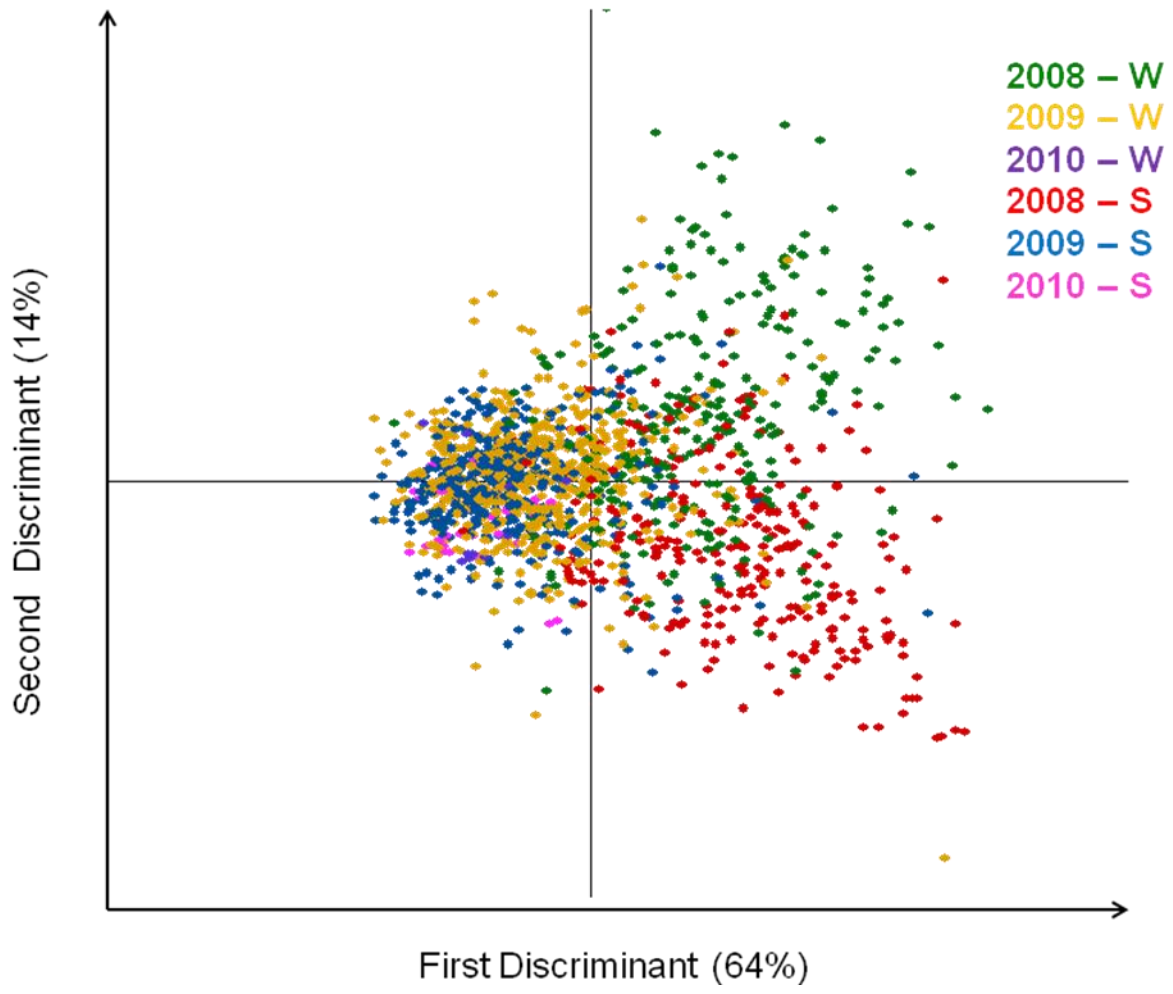


Figure 6. MANOVA analysis of all SMC *E. coli* isolates, grouped by year and sample type. 2008 water and sediment isolates are shown as (●) and (●), respectively. Water and sediment isolates from 2009 are represented as (●) and (●), respectively. Water and sediment isolates from 2009 are represented as (●) and (●), respectively.

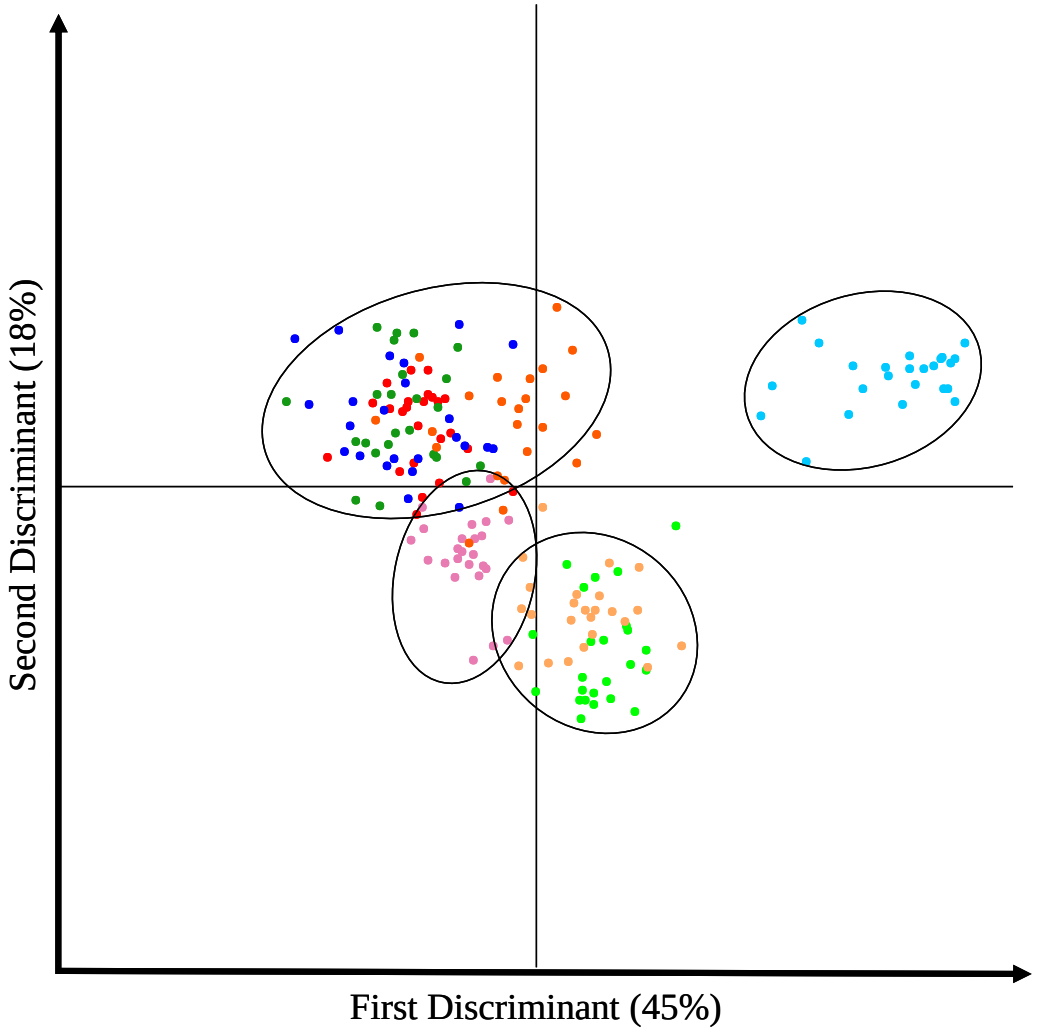


Figure 7. MANOVA analysis of low and no flow SMC *E. coli* isolates from water in 2008, grouped by flow condition and by site. High flow isolates from sites SM1, SM2, SM3, and SM4 are shown as (●), (●), (●), and (●), respectively. Low flow isolates from sites SM1, SM2, SM3, and SM4 are shown as (●), (●), (●), and (●), respectively.

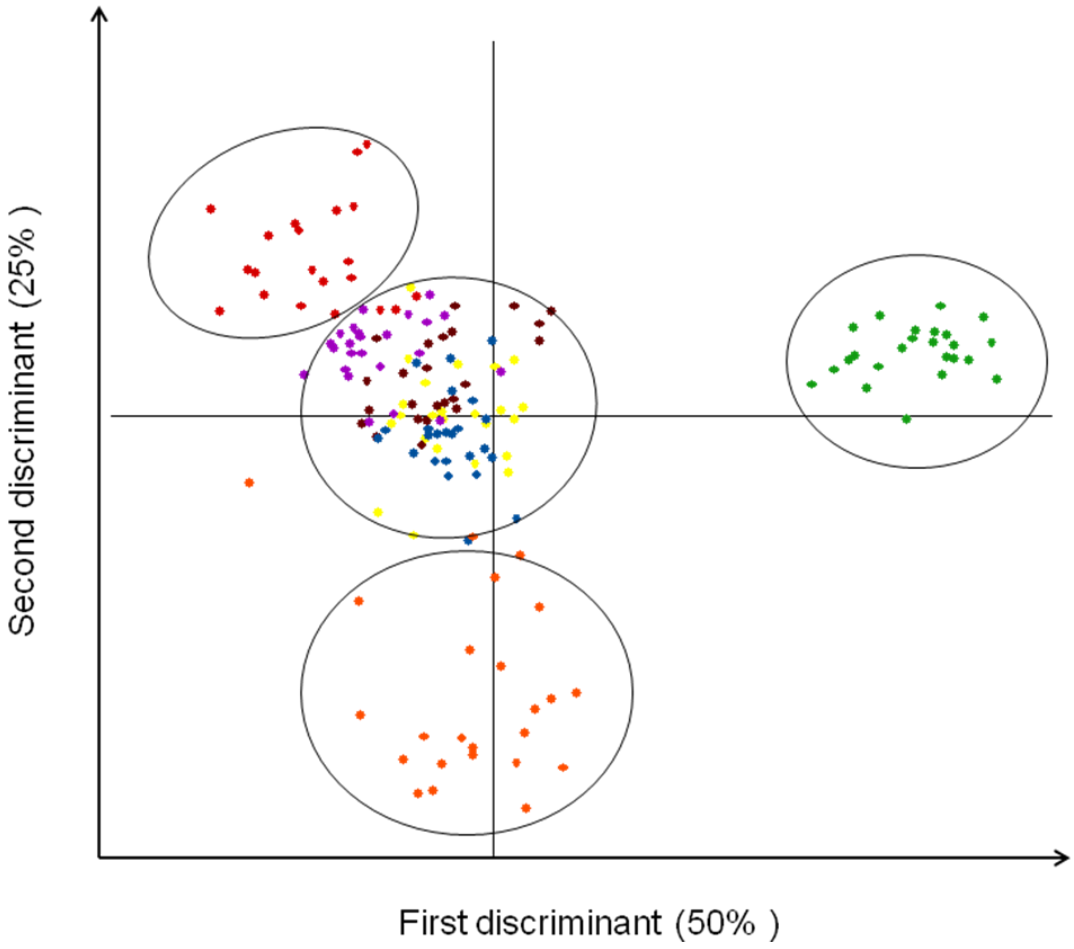


Figure 8. MANOVA analysis of the water isolates from SM1 obtained in 2009. *E. coli* isolates from April, May, June, July, August, September & October are represented as (●), (●), (●), (●), (●), (●), (●) respectively.

CONCLUSIONS

In this study we showed that the *E. coli* counts varied considerably over time and this is likely due to seasonal effects, including temperature and rainfall amounts. Sediment and water isolates grouped together suggesting that mixing of the sediments with the water column has effects on the *E. coli* populations. Also, populations appeared to shift from year to year and were influenced by flow conditions in the creek. The study also showed that *E. coli* populations present in the SMC are dynamic and are made up of both transient and persistent strains which are likely growing in the environment, particularly in the sediments. Persistent strains are likely indigenous to the sites and may grow in the sediments and/or water. Dispersal of these strains in the water may increase *E. coli* counts even in the absence of new fecal inputs, which has obvious implications for water quality monitoring and TMDL determinations. Seasonal effects such as temperature and flow conditions likely play important roles in the survival and transport of *E. coli* from site to site. Futures studies and continued sampling are necessary to conclusively determine the inputs and examine the long term dynamics of *E. coli* at the Seven Mile Creek watershed.

Overall Summary

- *E. coli* counts were highest during summer than the rest of the sampling period.
- Cattle are likely major contributors to fecal pollution at the Seven Mile Creek sediments.
- Water and sediment isolates are closely related in each year, suggesting that there is mixing of strains between the sediment and water column.
- Mixing and transport of *E. coli* isolates occur between sites during flow conditions.
- Flow rate, temperature effects, rainfall and run-off events are likely factors influencing the distribution of E. coli populations in SMC.
- Results of this study indicate that indigenous *E. coli* strains are likely present in Seven Mile Creek sediments and that some strains appear to grow in the sediments. Thus, sites likely contain a mixture of newly acquired and resident strains
- These results confound the use of *E. coli* in water quality assessment.

ACKNOWLEDGEMENTS

This work was supported by the Minnesota Department of Agriculture. We thank Scott Matteson and the field crew at Minnesota State University Mankato for sample collection and the *E. coli* count data, Jack Bovee at the MN Soil and Water Conservation Board for the collection of physical data from the SMC, Adam Birr for his help in organizing this project, and John Ferguson for his assistance with statistical analyses.

REFERENCES

1. **Anderson, M. A., J. E. Whitlock, and V. J. Harwood.** 2006. Diversity and distribution of *Escherichia coli* genotypes and antibiotic-resistant phenotypes in feces of humans, cattle, and horses. Appl. Environ. Microbiol. **72:**6914-6922.
2. **Atlas, A. M., and R. Bartha.** 1998. Microbial ecology: Fundamentals and applications, 4th ed. Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.
3. **Bernhard, A. E., and K. G. Field.** 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. **66:**4571-4574.
4. **Berry, C., B. J. Lloyd, and J. S. Colbourne.** 1991. Effect of heat shock on recovery of *Escherichia coli* from drinking water. Water Sci. Technol. **24:**85-88

5. **Byappanahalli, M. N., and R. Fujioka.** 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:**27-32
6. **Byappanahalli, M. N., and R. S. Fujioka.** 1998. Evidence that tropical soil environment can support the growth of *Escherichia coli*. Water Sci. Technol. **38:**171-174.
7. **Byappanahalli, M. N., D. A. Shively, M. B. Nevers, M. J. Sadowsky, and R. L. Whitman.** 2003. Growth and survival of *E. coli* and enterococci populations in the macro-alga *Cladophora* (Chlorophyta). FEMS Microbiol. Ecol. **46:**203-211.
8. **Byappanahalli, M. N., R. L. Whitman, D. A. Shively, J. Ferguson, S. Ishii, and M. J. Sadowsky.** 2007. Population structure of *Cladophora*-borne *Escherichia coli* in nearshore water of Lake Michigan. Water Res. **41:**3649-3654.
9. **Byappanahalli, M. N., R. L. Whitman, D. A. Shively, M. J. Sadowsky, and S. Ishii.** 2006. Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. Environ. Microbiol. **8:**504-513.
10. **Cabelli, V. J., A. P. Dufour, L. J. McCabe, and M. A. Levin.** 1982. Swimming-associated gastroenteritis and water quality. Am. J. Epidemiology **115:**606-616.
11. **Corbett, S. J., G. L. Rubin, G. K. Curry, and D. G. Kleinbaum.** 1993. The health effects of swimming at Sydney beaches. Am. J. Public Health. **83:**1701-1706
12. **Dick, L. K., A. E. Bernhard, T. J. Brodeur, J. W. S. Domingo, J. M. Simpson, S. P. Walters, and K. G. Field.** 2005. Host distributions of uncultivated fecal *Bacteroidales* reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. **71:**3184-3191.

13. **Dombek, P. E., L. K. Johnson, S. T. Zimmerley, and M. J. Sadowsky.** 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.
14. **Dufour, A. P.** 1984. Health effects criteria for fresh recreational waters. Report No. EPA-600/1-84-004. United States Environmental Protection Agency. Washington, D. C.
15. **Field, K. G., and M. Samadpour.** 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. **41:**3517-3538.
16. **Fujioka, R., C. Sian-Denton, M. Borja, J. Castro, and K. Morphew.** 1999. Soil: The environmental source of *Escherichia coli* and enterococci in Guam's streams. *J. Appl. Microbiol.* **85:**83S-89S
17. **Ishii, S., D. L. Hansen, R. E. Hicks, and M. J. Sadowsky.** 2007. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. Environ. Sci. Technol. **41:**2203-2209.
18. **Ishii, S., W. B. Ksoll, R. E. Hicks, and M. J. Sadowsky.** 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. App. Environ. Microbiol. **72:**612-621.
19. **Ishii, S., T. Yan, D. A. Shively, M. N. Byappanahalli, R. L. Whitman, and M. J. Sadowsky.** 2006. *Cladophora* (Chlorophyta) spp. harbor human bacterial pathogens in nearshore water of Lake Michigan. Appl. Environ. Microbiol. **72:**4545-4553.
20. **Jamieson, R. C., D. M. Joy, H. Lee, R. Kostaschuk, and R. J. Gordon.** 2005. Resuspension of sediment-associated *Escherichia coli* in a natural stream. J. Environ. Qual. **34:**581-589.

21. **Johnson, L. K., M. B. Brown, E. A. Carruthers, J. A. Ferguson, P. E. Dombek, and M. J. Sadowsky.** 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.

22. **Kingsley, M. T., and B. B. Bohlool.** 1981. Release of *Rhizobium* spp. from tropical toils and recovery for immunofluorescence enumeration. Appl. Environ. Microbiol. **42:**241-248.

23. **Kinzelman, J., S. L. McLellan, A. D. Daniels, S. Cashin, A. Singh, S. Gradus, and R. Bagley.** 2004. Non-point source pollution: determination of replication versus persistence of *Escherichia coli* in surface water and sediments with correlation of levels to readily measurable environmental parameters. J. Water Health. **2:**103-114.

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

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

26. **Mokady, D., U. Gophna, and E. Z. Ron.** 2005. Extensive gene diversity in septicemic *Escherichia coli* strains. J. Clin. Microbiol. **43:**66-73.

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

28. **Nunan, N., K. Ritz, D. Crabb, K. Harris, K. Wu, J. W. Crawford, and I. M. Young.** 2001. Quantification of the *in situ* distribution of soil bacteria by large-scale imaging sections of undisturbed soil. FEMS Microbiol. Ecol. **37:**67-77.

29. **Roszak, D. B., and R. R. Colwell.** 1987. Survival strategies of bacteria in the natural environment. Microbiol. Rev. **51:**365-379.

30. **Shanks, O. C., E. Atikovic, A. D. Blackwood, J. Lu, R. T. Noble, J. Santo Domingo, S. Seifring, M. Sivaganesan, and R. A. Haugland.** 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol **74:**745-752.

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

32. **Tassoula, E. A.** 1997. Growth possibilities of *E. coli* in natural waters. Int. J. Environ. Stud. **52:**67-73.

33. **Tate III, R. L.** 1978. Cultural and environmental factors affecting the longevity of *Escherichia coli* in histosols. Appl. Environ. Microbiol. **35:**925-929.

34. **Terzieva, S. I., and G. A. McFeters.** 1991. Survival and injury of *Escherichia coli*, *Campylobacter jejuni*, and *Yersinia enterocolitica* in stream water. Can. J. Microbiol. **37:**785-790.

35. **United States Environmental Protection Agency.** 2002. Method 1603: *Escherichia coli* (*E. coli*) in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar. U. S. Environmental Protection Agency. Washington, D. C.

36. **Whitman, R. L., and M. B. Nevers.** 2003. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. Appl. Environ. Microbiol. **69:**5555-5562.

37. **Wiedenmann, A., P. Kruger, K. Dietz, J. M. Lopez-Pila, R. Szewzyk, and K. Botzenhart.** 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:**228-236.

38. **Wollum, I. A. G., and D. K. Cassel.** 1984. Spatial variability in *Rhizobium japonicum* in two North Carolina (USA) soils. Soil Sci. Soc. Am. J. **48:**1082-1086.

39. **Yan, T., and M. J. Sadowsky.** 2007. Determining sources of fecal bacteria in waterways. Environ. Monit. Assess. **129:**97-106