

ABSTRACT

Molecular- and physiological-based characterization of fecal coliform reservoirs in Maumee Bay.

Abstract of the final report to the Ohio Lake Erie Commission (November, 2007)

Prepared by: Von Sigler, Ph.D.

The University of Toledo Department of Environmental Sciences

Fecal bacteria contamination is a primary cause of swimming advisories at Lake Erie Beach (LEB) (Oregon, OH). While previous research implicated Berger Ditch as a contributor of fecal pollution, we hypothesized that other nearby drainage ditches also play a role in the pollution of LEB. Throughout one year, biweekly water samples collected from LEB and three nearby drainage ditches were (i) analyzed for *E. coli* density and (ii) assessed for community structure using a novel community genetic fingerprinting method, denaturing gradient gel electrophoresis (DGGE) analysis, as well as an isolate based method, BOX-PCR fingerprinting. The results showed that *E. coli* density in LEB and the three ditches were dynamic, seasonally variable and mostly dependant on rainfall. Fingerprinting analyses indicated that although Berger Ditch remains a major contributing source of *E. coli* to LEB throughout the year, other nearby ditches (Tobias- and McHenry Ditches) appeared to contribute pollution during at least a portion of the swim season. BOX-PCR analyses of *E. coli* libraries collected during the swim season suggested that temporal variation in pollution sources is as important as geographic location and must be considered when developing pollution mitigation plans.

**Molecular- and physiological-based characterization of fecal coliform
reservoirs in Maumee Bay.**

Final Report to the Ohio Lake Erie Commission (November, 2007)

Prepared by: **Von Sigler, Ph.D.**

The University of Toledo Department of Environmental Sciences

Table of Contents

	<u>Page</u>
List of Figures	iii
List of Tables	iii
I. Introduction	1
II. Materials and Methods	5
III. Results	11
IV. Discussion and Conclusion	15
V. References	22

List of Figures

Figure 1	Map of locations	5
Figure 2	Rapid <i>E. coli</i> 2 agar	6
Figure 3	Example of rarefaction analysis	10
Figure 4	<i>E. coli</i> densities among the three beach sites	11
Figure 5	Combination of rainfall and <i>E. coli</i> composite densities	11
Figure 6	DGGE analysis of <i>E. coli</i> community fingerprints	13
Figure 7	MDS analysis of <i>E. coli</i> isolates based on sampling location and time	13
Figure 8	Rarefaction analysis of <i>E. coli</i> isolate library	14

List of Tables

Table 1	Seasonal correlation tests of <i>E. coli</i> density	12
Table 2	Jackknife analysis of <i>E. coli</i> isolates based on location and time	14

INTRODUCTION

Pathogens originating from the fecal material of warm blooded animals, are frequently detected in recreational waters and pose a variety of health hazards, including gastrointestinal, respiratory, dermatologic, and ear, nose, and throat infections (Henrickson, et al, 2001). To protect beachgoers from exposure to these pathogens, the monitoring of the microbiological quality of beach water in coastal and Great Lakes states is required (USEPA, 2000b). Indicator bacteria, such as *E. coli* or Enterococci, are used to signal the presence of pathogens without intensive exercise of assaying a water sample for the numerous pathogens that might be present. If the indicator bacteria level exceeds EPA standards, an advisory is posted to warn the public of the health risk associated with primary contact with the water. While the role of advisory posting is to protect public health, it can also create an extensive economic impact. Lake Erie Beach (LEB) at Maumee Bay State Park (MBSP) is a popular destination in northwest Ohio for water recreation including swimming, fishing, boating, windsurfing, as well as non-water-contact activities, such as sunbathing, and activities in the sand. MBSP, along with LEB, is a vital economic resource for northwest Ohio. LEB is partially responsible for tourism that contributes \$6.2 million to the local economy annually. Of the 1.4 million visitors annually to MBSP, 17% (238,000) go to specifically swim at the beach (Sohnngen et al., 1999). In 2003, LEB was closed for a total of 42 days as a result of bacterial pollution (ODH, 2006), resulting in an estimated loss of \$347,800 to the local economy (assuming a 120 day swimming season).

According to the Ohio Department of Health (ODH, 2006), the total advisory days for LEB were 42, 20 and 8 days during the swimming season (June through August) for 2003, 2004 and 2005, respectively. Before 2005 an advisory was issued whenever the geometric mean density of *E. coli* for any five samples collected within a 30-day period exceeded 126 *E. coli* CFU/100 mL of water. In 2006, the Ohio EPA initiated the use of the single sample maximum standard. This new standard dictated that a beach be closed or posted under an advisory if the *E. coli* density exceeds 235 CFU/100 mL in a single water sample (ODH, 2006). Under the new standard, posted advisories at LEB increased to 17 days in 2006, from 8 days in 2005. While common causes for beach advisories can result from various inputs including stormwater runoff, sewage spills and overflows, algal blooms, dredging, and chemical contamination (NRDC, 2005), all LEB advisories from 2003 to 2005 were attributed to bacteria levels that exceeded established beach water quality standards (ODH, 2006).

Ingestion of pathogens associated with poor water quality is a significant risk to the health of people visiting a beach (Kueh et al., 1995). It was reported that on average, children and adults recreating in water swallow 37 and 16 ml of water, respectively, per every 45 minutes of activity (Dufour et al., 2006). Since the *E. coli* density in beach water can be as high as 2×10^4 CFU/100 mL (McLellan and Jensen, 2005), 7400 and 3200 *E. coli* cells could theoretically be consumed by children and adults, respectively during less than one hour of activity. Ingested bacteria loads in this range can pose a significant health risk. For example, it was established that as few as ten cells of *E. coli* O157:H7 can cause gastrointestinal illness in humans (Mead and Griffin, 1998).

Therefore, even small volumes of ingested water might pose a risk to beachgoers, in particular young children, the elderly, pregnant women, and immunocompromized people (NRDC, 2006).

Although it has been established that bacteria greatly impact water quality, most municipalities have failed to identify sources of bacterial pollution to recreational waters (NRDC, 2006). There are two classes of bacterial pollution sources: host sources (animals or humans), which can directly contaminate a water body, and geographic sources that indirectly contaminate via farmland leaching, septic tank leaking or ditch water transport. Geographic source identification involves identifying the specific location, such as streams, ditches, or agricultural land as the source of fecal pollution. To track the source of fecal pollution, two broad categories of bacterial source tracking (BST) are established including library-dependent - and library-independent methods. Library-dependent methods compare phenotypic or genotypic patterns in bacteria collected from polluted sinks with those from potential sources. These methods present considerable limitations, as they require the generation of isolate libraries containing potentially thousands of isolates collected from the pollution sink and each potential source (Wiggins et al., 2003). This large number is mandated by the high diversity of environmental bacterial populations and broad geographic locations of potential sources. Because of the vast size of appropriate isolate libraries, library-dependent methods can result in excessive expenditure of time, labor and money (Johnson et al., 2004; Wiggins et al, 2003).

In contrast, library-independent methods do not require a reference isolate library. Library independent methods are based on the presence or absence of genes associated with a target organism or function in an environmental sample. For example, specific human and animal DNA markers can be used to identify the host origin (animal vs. human) of fecal pollution (Bernhard and Field, 2000 a, b; Hamilton, et al., 2006) from among the entire microbial community present at the site. Another example of emerging technology is denaturing gradient gel electrophoresis (DGGE) analysis, which can provide fingerprints of microbial communities by separating PCR generated fragments that vary in nucleotide sequence by as little as one base pair (Muyzer et al., 1993). Environmental DGGE analysis is generally used to analyze the structure of microbial community in the soil and water environment (Sigler and Zeyer, 2004; Sekiguchi et al., 2002). While library-dependent methods might require hundreds to thousands of isolates to construct a reference library, DGGE analysis does not require a reference library, yet can still provide a diagnostic fingerprint useful for differentiating complex bacterial assemblages. The effective use of DGGE analysis for microbial source tracking depends on identifying genes that can differentiate bacterial communities collected from pollution sinks and potential sources. Furthermore, a gene associated with an established fecal indicator would further enhance the use of DGGE analysis for source tracking exercises. One such gene is *uidA*. *uidA* is expressed in approximately 95% of environmental *E. coli* (Martins et al., 1993) and codes for the enzyme, β -glucuronidase, which has been shown to be an appropriate molecular target for *E. coli* detection (Bej et al., 1991). It has been reported that nucleotide sequence analysis of β -glucuronidase (*uidA*) from 182 human and animal fecal *E. coli* isolates

resulted in 75% correct assignment to the primary host, suggesting the utility of *uidA* as a microbial source tracking target (Ram et al., 2004). DGGE analysis of *uidA* is also an effective tool to differentiate unique *E. coli* populations in natural waters (Sigler and Pasutti, 2006; Farnleitner et al., 2000a).

In 1996, a study commissioned by MBSP and the City of Toledo found that *E. coli* densities at LEB were highest in water samples collected from the east end of the beach and declined in samples collected from the west end (Anonymous, 2003). In addition, research conducted by our laboratory in the summer of 2004 reached the same conclusion (data not shown). These studies led to the hypothesis that Berger Ditch, a drainage ditch located adjacent to the east end of the beach, might be a significant source of *E. coli* to LEB. Prior to the development of this hypothesis, several studies attempted to identify the pollution source. Lauber (2001) used FAME- (fatty acid methyl ester) and plasmid analyses of *E. coli* isolates from water collected from LEB, the Inland Lake Beach, and nearby ditches to ascertain the contribution of *E. coli* from Berger Ditch. The study did not conclusively define the source of fecal pollution at MBSP, as plasmid analysis was limited in tracing diffuse pollution sources. Further efforts by Francy et al. (2005) determined that the spatial distribution of *E. coli* could be explained by several environmental factors including rainfall, wind direction and wave height. However, no definitive identification of fecal contamination sources could be made. One variable that could have complicated these studies is the presence of other nearby ditches. While Berger ditch has been historically implicated as the primary source of bacterial pollution, additional sources such as McHenry- and Tobias Ditches, are located to the west of the beach and drain significant areas of the surrounding watershed into Lake Erie. Therefore, these sites could also have contributed to the bacterial pollution at the beach.

Most previous attempts to identify geographic sources of bacterial pollution have relied on point-in-time samplings. However, because seasonal patterns in climate, vegetation, and animal- and human activity are not consistent, it is likely that the sources, and therefore bacterial pollution volume and composition, will change throughout the year. The collection of microbiological data throughout the year can provide several advantages over single season-, or point-in-time samplings. First, a long-term sampling effort can integrate many of the short-term, periodic issues that have been shown to affect *E. coli* densities in natural waters (Whitman and Nevers, 2004). These issues include environmental factors, such as UV radiation, wind speed and direction, wave height and relative lake level, temperature, rainfall and stream flow volume (Whitman et al., 2004; Lipp et al., 2001; Kinzelman et al., 2004). Additional factors include fecal inputs from wild and domestic animals, such as gulls (Lévesque et al., 1993), swan (Khatib et al., 2003) and cattle (Doran and Linn, 1979). If one or more of the aforementioned factors is strong, but short lived, a one-time sampling might result in a conclusion that inflates the true contribution of the source. Long-term (year-long) sampling might avoid this bias by allowing the integration of multiple point-in-time analyses. Finally, although summer is the primary season for water recreation, it is not the only season for water-enthusiasts. Sports such as windsurfing and fishing are performed at MBSP year-round (personal observation). Therefore, integrated, yearlong

sampling can account for water quality conditions during secondary recreation seasons as well as the primary recreation season.

The overarching goal of this study was to identify the geographic sources of bacterial pollution at LEB. This was addressed by performing DGGE analysis of *E. coli* communities from water collected from LEB, Berger-, McHenry- and Tobias Ditches, from October 2005 to September 2006.

The specific objectives of this study were to:

1. Use *E. coli* density correlations and DGGE analysis of *E. coli* communities to identify the geographic sources of fecal pollution impacting water quality at LEB.
2. Monitor the year-long dynamics of *E. coli* densities and community structure in water collected from LEB and nearby ditches, including Berger-, McHenry- and Tobias Ditches to determine if the source ditch(es) is changing over time.
3. Confirm the results of DGGE analysis by using a high-resolution library-dependent technique, BOX-PCR on a library of *E. coli* isolates collected from the beach and three ditches during the swimming season (June to September 2006).

MATERIALS AND METHODS

Site descriptions. Water samples were collected from LEB as well as three ditches used for storm water management, including Berger-, McHenry-, and Tobias Ditches (Fig. 1). Each of the sites is described in detail below.

Lake Erie Beach is located in Maumee Bay State Park in northwest Ohio. It is approximately 750 m in length and defined by six embayments to limit wave energy, each approximately 100 m in length.

Berger Ditch. Berger Ditch is approximately 5,500 m long and flows from south to north. It connects with Wolf Creek (about 15,000 m long) approximately 2,400 m south of MBSP. Berger Ditch flows through varying land management systems, including residential areas, farmland, wetlands, forest, and ultimately through the MBSP's marina before emptying into Lake Erie directly east of LEB. Berger Ditch is approximately one meter wide at its headwaters, and it gradually becomes wider (~6 m) as it approaches the outfall to the lake. Berger Ditch continually flows throughout the year.

McHenry- and Tobias Ditches.

McHenry Ditch is approximately 4,900 m long, flows south to north and eventually empties into Lake Erie approximately 900 m west of LEB. Tobias Ditch flows for 3,500 m through residential areas and farmland areas, and adjacent to a horse farm (about 25 acres in area and containing 32 horses) before emptying into Lake Erie approximately 1,750 m from the west end of the beach. Compared with Berger Ditch, McHenry- and Tobias Ditches are shorter, shallower and narrower. Neither ditch flows continually throughout the year, but both flow after rainfall. Therefore, it is not always possible to sample water from these two ditches. All three ditches are open to the air, except when they cross a road, at which point the water travels through an underground culvert.



Fig.1. Aerial view of Lake Erie Beach and Berger-, McHenry- and Tobias Ditches. Sampling locations are denoted with a star.

Water collection. Water samples were collected along LEB and the ditches at the following specific locations (Fig.1):

- (i) Four sampling sites along Berger Ditch (the point where Wolf Creek and Berger Ditch merge, MBSP entrance, pre-marina and the mouth of Berger Ditch as it enters Lake Erie),

- (ii) Three sites along LEB (west-, middle- and east-most embayments),
- (iii) Two sites, representing the terminal (north-most) section of sewer pipe and mouth of McHenry Ditch and Tobias Ditch.

Samples were collected monthly from October 2005 to September 2006 from all sites and twice per month from the beach sites from April 2006 to September 2006. Surface water (10-20 cm depth) from each site was collected in triplicate, sterile, 1 L plastic bottles and stored on ice until laboratory analysis was performed (within 6 hours after collection). Samples were usually collected between 11:00 am and 2:00 pm within 36 hours following a rainfall event (0.13-8 cm) to ensure that water in each ditch was positively flowing (south-to-north) into Lake Erie. Depending on rainfall and wind conditions, backflow would occur when northerly currents forced lake water southward, into the ditches. Under backflow conditions, water samples were not collected. Rainfall data were collected from NOAA, (<http://www.weather.gov/climate/index.php?wfo=cle>).

The data representing each of the twelve months during which samplings were performed were expressed within the context of three arbitrary “seasons”: (i) Season One (nonswimming season, from October, 2005 to January, 2006); (ii) Season Two (nonswimming season, from February, 2006 to May, 2006), and (iii) Season Three (swimming season, from June, 2006 to September, 2006).

***E. coli* enumeration.** *E. coli* densities in the water samples were determined following standard membrane filtration procedures (USEPA, 2000c). Specifically, three volumes (1 mL, 25 mL, 100 mL) of each water sample were passed through a mixed cellulose ester membrane (0.45 μ m pore size). For the 1 ml volume, 1 ml of sample and 9 ml of 10 mM sodium phosphate buffer (pH 7.0) were filtered together to ensure adequate dispersal of the bacteria over the membrane. Membranes were placed onto Rapid *E. coli* 2 medium (BioRad Laboratories, Inc), and incubated upside down at 37° C for 24 h. Following incubation, purple colonies were counted in the plates that contained between 10 and 100 colonies. Rapid *E. coli* 2 medium is a medium for the enumeration of *E. coli* and other coliforms that takes advantage of the activity of two enzymes: β -D-glucuronidase and β -D-galactosidase. To detect the activity of these two enzymes, the medium contains two chromogenic substrates; 4-methylumbelliferyl- β -D-glucuronide (MUG) and 5-bromo-4-chloro-3-indolyl- β -d-galactosidase (X-Gal). When β -D-glucuronidase cleaves MUG, colonies will turn violet, while β -D-galactosidase cleavage of X-Gal results in blue-green colonies. *E. coli* strains express both enzymes simultaneously, resulting in violet colonies. Coliform bacteria that are not *E. coli* only express β -D-galactosidase, and will remain blue-green (Fig. 2). *E. coli* density was expressed as colony forming



Fig.2 *E. coli* and coliform bacteria growing on Rapid *E. coli* 2 agar. *E. coli* in this agar exhibits purple coloration, while other coliforms are light blue.

unit per 100 mL water (CUF 100 mL⁻¹). The mean *E. coli* density (CFU 100 mL⁻¹) for each time point (month) was obtained by calculating the mean *E. coli* density at each site of the beach and in each ditch. The temporal correlation of *E. coli* densities between ditch and beach samples was calculated using Statistical Analysis Software (v 9.1, SAS Institute, Inc).

For eventual DNA extraction (described below), large volumes of water (75 - 800 ml) were filtered and incubated (as described above) to generate lawns of *E. coli*. Incubation was performed at 44.5° C for 24 h. This incubation temperature was higher than that of the enumeration protocol to limit the growth of bacteria exhibiting lower optimal growth temperatures, thereby maximizing *E. coli* enrichment.

***E. coli* community DNA isolation.** DNA isolation from the enrichment membranes was performed according to the protocol of Sigler and Zeyer (2002). Specifically, each membrane containing an enriched *E. coli* community was cut into small pieces with sterile scissors and then transferred to a sterile, 2 ml microcentrifuge tube containing 0.4 - 0.5 ml of glass beads (0.10 mm diameter) and 1.0 ml of DNA extraction buffer (5% SDS, 50 mM NaCl; 50 mM Tris-HCl, pH 7.6; 50 mM EDTA). Tubes were shaken with a Fast-Prep (Thermo Electron Corporation) instrument for 30 seconds at 5.5 ms⁻¹, followed by centrifugation for three minutes at 14,000 x g. The supernatant (~ 0.4 - 0.6 ml) was transferred to a new 1.5 ml microcentrifuge tube and purified by adding 0.5 volumes of both phenol (Tris-saturated, pH 8.0) and chloroform/isoamyl alcohol (24:1) (final volume ~ 1 ml). The solution was vortexed on a high setting until milky (5 - 10 s) followed by centrifugation for three minutes at 14,000 x g. The aqueous phase containing the DNA was transferred to a new 1.5 ml tube and an equal volume of chloroform was added, vortexed (5 - 10 s), followed by centrifugation for three minutes at 14,000 x g. The upper, aqueous phase was transferred to a new 1.5 ml tube. DNA was precipitated by adding 10% and 70% of the DNA extract volume of 3 M sodium acetate solution and isopropanol, respectively. Tubes were mixed by inverting several times by hand, and then centrifuged at 14,000 x g for 30 minutes at 10° C. The supernatant was removed by aspiration and the remaining DNA pellet was washed with 0.5 ml of ice-cold 70% ethanol by gently inverting the tube several times. The ethanol was aspirated and the DNA pellet was suspended in 50 µL of DNase/RNase-free water. Following quantification by spectrophotometry (A₂₆₀), each DNA sample was diluted to a concentration of approximately 100 µg mL⁻¹ and stored at -20° C.

***uidA* polymerase chain reaction.** A portion of the *uidA* gene of *E. coli* was amplified by PCR analysis using primers UAL-1939 and UAR-2105, which target the carboxyl-coding region of the *uidA* gene (Bej et al., 1991). Each reaction contained the desired DNA at a concentration of 1 ng µl⁻¹, PCR buffer (1X), 2.5mM MgCl₂, BSA (bovine serum albumin) (1 mg mL⁻¹), 0.2 mM each dNTP, 0.3 µM each primer), 0.02 U µL⁻¹ *Taq* DNA polymerase (Promega Corporation), and DNase/RNase-free water to a final volume of 50 µL. A positive and negative control reaction contained DNA isolated from *E. coli* strain DH5α, and water (no added DNA) was also included in each PCR assay. The PCR program consisted of an initial denaturing cycle (94° C for 5 min), 30 cycles of 94° C for 45 s, 50° C for 45 s, and 72° C for 60 s, and a final extension cycle of 72° C for 30 min.

PCR products were screened by electrophoresis in 1x TAE buffer at 100 v for 15 min and archived by image analysis with a Gel Logic 200 system (Eastman Kodak Company).

DNA community analysis by DGGE. DGGE analysis (Sigler et al., 2004) was performed in a DCODE DGGE system (BioRad, Laboratories, Inc) for 1000 V•h (200 V for 5 h) at 60° C using a denaturing gradient of 40-55% (a 100% denaturing gel contains 7M urea and 40% formamide). Approximately 30 µL of PCR product was loaded into each lane, while a standard marker was loaded such that a maximum of five samples separated each marker well. The standard marker was generated by mixing the PCR products of the DNA from six bacterial strains (each amplified by using 16S rRNA primers [Muyzer et al., 1993]). These strains included *Pseudomonas aeruginosa*, *Ralstonia pickettii*, and four environmental *E. coli* strains. Following electrophoresis, the gel was stained for 15 min in 50 mL of a 1:10,000 dilution of GelStar nucleic acid stain (BioWhittaker, Inc), and band patterns were archived by an image analysis.

DGGE fingerprint patterns were analyzed with GelCompar II software (version 4.6, Applied Maths, Inc). DGGE fingerprints were normalized using the abovementioned standard marker. The similarity of *E. coli* community fingerprints was calculated using the band-based, Dice coefficient (Dice, 1945) with 0% optimization and 3% band position tolerance settings in the GelCompar II software. Bands in each lane were identified by the software and then checked manually to avoid missassignment. Similarity coefficients were calculated for each fingerprint pair possibility, from which a similarity matrix was constructed. A dendrogram displaying the inferred relationship between all fingerprints was generated using the default parameters of the GelCompar II software.

***E. coli* isolate libraries and BOX-PCR analysis.** This study represented the first use of DGGE to identify geographic sources of fecal pollution. Therefore, a confirmation of the method's validity was necessary. BOX-PCR is a library dependent method that is an effective tool for differentiating human and animal sources of fecal pollution (Dombek et al., 2000). Therefore, in an effort to validate the results of the DGGE analyses, BOX-PCR analysis of *E. coli* libraries generated during a portion of the study was performed.

A library of *E. coli* colonies was generated during Season Three (June to September of 2006). Specifically, during each month, approximately 25 isolates were collected from the beach and each ditch, resulting in a total of 100 isolates from the four locations each month. The *E. coli* were collected from the same filter membranes that were generated during the enumeration analysis. Isolates were picked with streaking needles, transferred twice to fresh Rapid *E. coli* 2 agar and incubated (37° C overnight). Pure, putative *E. coli* isolates were confirmed for identity by determining morphology on eosin-methylene blue agar (EMB) (OXOID, Inc). EMB agar is a differential medium used in the identification and isolation of Gram-negative enteric rods. *E. coli* colonies on EMB medium appear black, with a metallic green sheen (due to the fermentation-driven precipitation of methylene blue from the medium). The identity of the putative *E. coli* was further confirmed by PCR using 16S rRNA gene *E. coli* primers specific for *E. coli* (Sabat et al., 2000). DNA was isolated from each confirmed isolate

using a Wizard SV 96 Genomic DNA Purification System (Promega Corporation). Following isolation, the DNA was quantified as described above and stored at -20° C.

BOX PCR was performed according to the protocol of Rademaker et al. (1998). To facilitate the electrophoresis of the numerous BOX-PCR samples, three electrophoresis systems were used in this study; (i) Owl Model A1 Gator Large Gel Electrophoresis System (Owl Separation Systems), (ii) Fisher Biotech Horizontal Electrophoresis Midigel System (Fisher Scientific, Inc.), and (iii) a BioRad Sub-cell GT (BioRad, Laboratories, Inc.). Migration distance in these three systems was approximately 13 cm, however to achieve comparable fingerprint separation, it was necessary to vary the running time used for each system (90 min, 120 min and 180 min for BioRad, Fisher and Owl systems, respectively). To assess the impact of different electrophoresis running times, a 100 bp DNA marker was loaded in several gel lanes and electrophoresed in each system for the duration of time mentioned above. The marker fingerprints were analyzed as described above, which revealed 99.99% similarity among the 68 marker test lanes (27 for Owl system, 23 for BioRad, and 18 for FisherBiotech), indicating that the selection of running times for each machine were appropriate and would not alter the analysis of fingerprint patterns.

BOX-PCR fingerprints were generated from 30 µl of each reaction mixture loaded onto a 1.5% agarose gel containing ethidium bromide (0.005 mg L⁻¹). A 100 bp DNA size ladder (Fisher Scientific, Inc.) was loaded into the two outside- and middle wells of the gel. Electrophoresis was performed as described above. Isolate fingerprints were archived and analyzed by GelCompar II software as described above, except that normalization were performed using the 100 bp DNA ladder instead of the DGGE marker. In an effort to express the BOX-PCR fingerprint data in a manner that could validate the DGGE results, multidimensional scaling (MDS) was performed to provide a convenient visual interpretation of the similarity among the fingerprints of total *E. coli* isolates. MDS produces a 'map' in three dimensions that describes the distance (relative similarity) between isolates, allowing researcher to visualize proximity of isolates both within and among source groups (Ritter et al., 2003). In this study MDS analysis identified relationships (e.g. separation or clustering) between the *E. coli* isolates based either on temporal or geographic patterns.

Jackknife analysis (maximum similarity) was performed to determine how accurately the BOX-PCR fingerprint similarity coefficients could predict the environment from which any given isolate was obtained. Briefly, the isolates were first manually assigned to the correct source groups (geographic areas and month of sampling), following which each isolate was individually removed from the database. Similarity between the removed isolate and each of the remaining isolates in each source group was calculated and a maximum group similarity coefficient was determined. The removed isolate was assigned to the source group that possessed the greatest maximum group similarity coefficient, and the percentage of accurately assigned isolates for each source group was then calculated, resulting in a rate of correct classification (maximum similarity) of the isolates collected from each location and month of sampling. If the rate of correct classification was high for each location, it meant that BOX-PCR fingerprints of the *E.*

coli isolates were discriminative, and vice-versa.

BOX- PCR reproducibility. To test whether the fingerprints of the same strain analyzed by BOX PCR in one gel were reproducible in other gels, DNA from five different *E. coli* isolates (e.g. strain A-E) was BOX-PCR amplified in five, separately mixed reactions followed by electrophoresis and image analysis as described above. The reproducibility of BOX-PCR was expressed as the fingerprint similarity (as calculated by GelCompar II) among the five repeated fingerprints generated from each of the five *E. coli* strains tested (e.g. A1-A5). The cut off value was calculated by averaging the similarity of five repeated fingerprints of each strain as shown in the calculation $\{[(\text{similarity of A1-A5}) + (\text{similarity of B1-B5}) + (\text{similarity of C1-C5}) + (\text{similarity of D1-D5}) + (\text{similarity of E1-E5})] / 5\}$. This resulting average similarity was used as a threshold similarity, above which isolates can be considered identical. Using this threshold value, the number of *E. coli* phlotypes in the libraries could be determined and used for rarefaction analysis (below).

Since BOX-PCR analysis could require from hundreds to thousands of isolates to correctly discriminate the fecal sources from human to animals (Dombek et al., 2000; McLellan et al., 2003; Johnson et al., 2004). It was necessary to determine if our library size was appropriate. Rarefaction analysis (Analytic Rarefaction software v1.3, University of Georgia) is a technique to compare species richness among samples of different sizes and was performed by in an effort to determine whether the sample size is enough for representing the richness of sample species. Each data point in a rarefaction curve represents the number of expected novel individuals resulting from a certain sampling effort. A high slope indicates a high likelihood that a novel member of the community will be sampled if sampling was to continue, and therefore suggests that a large proportion of the bacterial diversity has not been sampled. In contrast, a low slope (or zero slope in the most extreme case) indicates that repeated sampling will yield only a small number of additional species, suggesting that an appropriate sampling effort has occurred. For example, the straight line in Fig. 3 indicates a high likelihood that new phlotypes of *E. coli* will be sampled with continued sampling, which means the sampling size is inadequate, and further sampling is required. In contrast, the curved line shows that more intensive sampling will result in low likelihood of collecting new *E. coli* phlotypes.

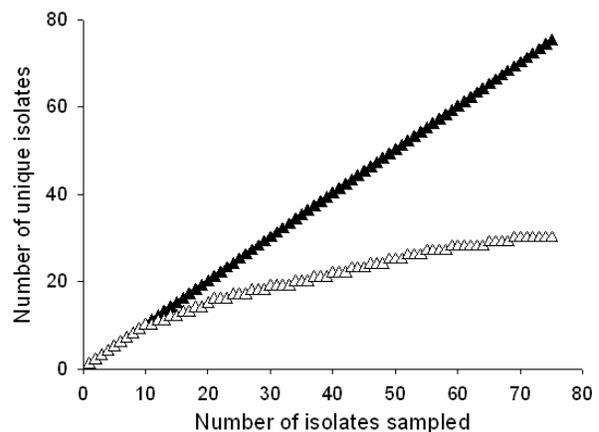


Fig. 3 Example of the results of rarefaction analysis. Closed triangles represent a non-representative sampling (no saturation of the number of unique phlotypes occurs). The open triangles represent an appropriate, representative sampling in which saturation of the number of unique phlotypes occurs.

RESULTS

***E. coli* enumeration.** *E. coli* densities in LEB (including east, middle and west) were dynamic, ranging from as low as 10 CFU 100 mL⁻¹ in November 2005 to as high as 1,700 100 mL⁻¹ in July 2006 (Fig. 4). In general, *E. coli* densities at three sites of the beach were more consistent in Seasons One and Two than in Season Three. Specifically, in Seasons One and -Two, *E. coli* densities in the three beach sites were low, less than 100 CFU 100 mL⁻¹, except in January, when density was greater than 500 CFU 100 mL⁻¹. In Season Three, *E. coli* densities in the three beach sites increased, beginning in mid-June, when *E. coli* density was approximately 1,600 CFU 100 mL⁻¹ in the east, and 1,200 CFU 100 mL⁻¹ in the west. While *E. coli* density in the three beach sites decreased during the next three months, the density on the east side of the beach was always significantly higher than that of the west side ($p < 0.05$). No significant difference in *E. coli* density between east and west sides was observed in Seasons One and Two ($p = 0.44$).

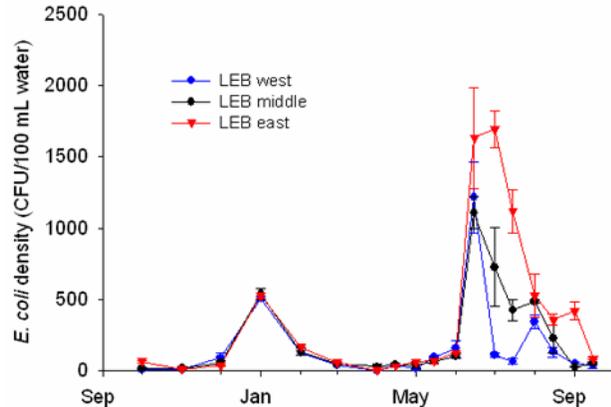


Fig. 4. *E. coli* densities at the west, middle and east sites of LEB from October, 2005 to September, 2006. Error bars indicate the standard error of the mean density ($n=3$).

Average *E. coli* densities in each ditch- and beach water sample were variable throughout the year, ranging from 10¹ to 10⁵ CFU 100 mL⁻¹ (Fig. 5). *E. coli* densities in the ditches and the beach were consistent during Seasons One and -Three. During Season One, *E. coli* densities in Berger Ditch and the beach increased over time. In Season Three, densities in all ditches and the beach increased from June to August, and decreased in September. However, in Season Two, the trend of *E. coli* densities in ditches and the beach was inconsistent, as densities in Berger Ditch decreased in March, and then increased in April, while densities in the beach decreased from February to April. As compared with *E. coli* densities in ditches, the densities at the beach were consistently lower, ranging from 10¹ to 10³ CFU 100 mL⁻¹.

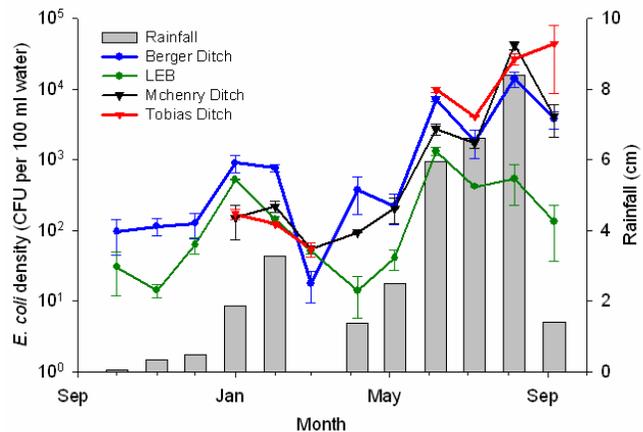


Fig. 5. Impact of rainfall on monthly *E. coli* densities in water collected from each ditch and the beach. The missing density data for McHenry and Tobias Ditches (March) was the result of low water flow. Each rainfall bar includes the amount of rainfall during the 96 h prior to sampling. Error bars represent the standard error of the mean ($n=3$).

The average rainfall depth during the 96 h prior to sampling was highest during Season Three (5.6 cm) and lowest during Season One (0.7 cm), with intermediate rainfall occurring during Season Two (1.8 cm). Rainfall depth was significantly correlated with *E. coli* densities in three of the four sites; LEB ($r=0.68$), Berger Ditch ($r=0.80$) and McHenry Ditch ($r=0.67$). *E. coli* density in Tobias Ditch was not correlated with rainfall volume ($r=0.12$).

Correlations between *E. coli* densities in the ditches and beach. The correlation of *E. coli* densities between each ditch and the beach were variable (Table 1). For example, during Season One, the *E. coli* density in the beach water was significantly correlated with densities in Berger Ditch ($r=0.98$, $p<0.05$). In contrast, *E. coli* densities in McHenry and Tobias Ditches were negatively correlated with those at the beach. During Season Two, the *E. coli* densities in all ditches correlated with densities at the beach. Specifically, correlation coefficients between Berger-, McHenry-, and Tobias Ditches and LEB were 0.63, 0.59 and 0.94, respectively ($p < 0.05$) (Table 1). During Season Three, the *E. coli* densities in each ditch were uncorrelated ($r < 0.2$) with those at the beach.

Table 1. Correlation of *E. coli* densities in water samples from each ditch and the beach. All correlations were significant ($P \leq 0.05$) unless noted otherwise (ns).

Site compared with LEB	Pearson correlation coefficient	Number of samples (n)
Season One		
Berger Ditch	0.98	12
McHenry Ditch	-0.73 (ns)	3
Tobias Ditch	-0.70 (ns)	3
Season Two		
Berger Ditch	0.63	12
McHenry Ditch	0.59	12
Tobias Ditch	0.94	6
Season Three		
Berger Ditch	0.13 (ns)	12
McHenry Ditch	-0.14 (ns)	12
Tobias Ditch	-0.56 (ns)	12

***E. coli* community structure.** In an effort to characterize the dynamics of the bacterial pollution at LEB, fingerprints were generated that represented the *E. coli* community in each ditch and the three beach sites (Fig. 6). These fingerprints identified temporal dynamics of *E. coli* and were compared to determine community similarity among the sites, under the assumption that high similarity of *E. coli* communities of the ditch and beach samples would imply a source-sink relationship between the sites. Dendrogram analysis of community fingerprints showed that in Season One, *E. coli* communities in Berger Ditch, Tobias Ditch and the beach (west, middle and east sites) were 85% similar, forming a distinct cluster, to which the Tobias Ditch communities were 44% similar.

Analysis of fingerprints generated from *E. coli* communities during Season Two showed that all communities were more similar to each other (62% overall similarity) than those from Season One (44%). Distinct clustering (85% similarity) was observed among *E. coli* community fingerprints generated from the beach sites. Of the three ditches, Berger Ditch harbored *E. coli* communities most similar to those of the beach sites (69%), followed by McHenry- and Tobias Ditches (62%).

E. coli community fingerprints from Season Three were more similar to each other than those from Seasons One and Two, exhibiting 88% overall similarity. Within this tight clustering, *E. coli* communities from the different sites were highly similar indicating that the communities from the three ditches were all similar to those at the beach.

Confirming the results of the *E. coli* community structure analyses. Since this study represented the initial use of DGGE fingerprinting to identify geographic sources of bacterial pollution, BOX PCR, an established, library dependent method of source identification, was used to validate the DGGE results. To limit the resource expenditure associated with library dependent methodology, BOX-PCR validation of DGGE was limited to samples collected during Season Three, since this duration of time represented the swimming season and the most sensitive season for source tracking. Cluster analysis of the 358 BOX-PCR fingerprints revealed that 83 distinct *E. coli* phylotypes (using the 85% similarity threshold as described in Materials and Methods) were collected during the Season Three sampling (See attachment; BOX-PCR Dendrogram). No geographic-specific clustering of the *E. coli* phylotypes were observed and similar *E. coli* phylotypes were present in all three ditches and the beach sites. For example, multidimensional scaling (MDS) failed to separate the geographic origin of the phylotypes (Fig. 7A). Similar results were obtained following Jackknife analysis, as the rates of correct geographic classification of the *E. coli* isolates were low to moderate, ranging from 53% to 71% (Table 2A).

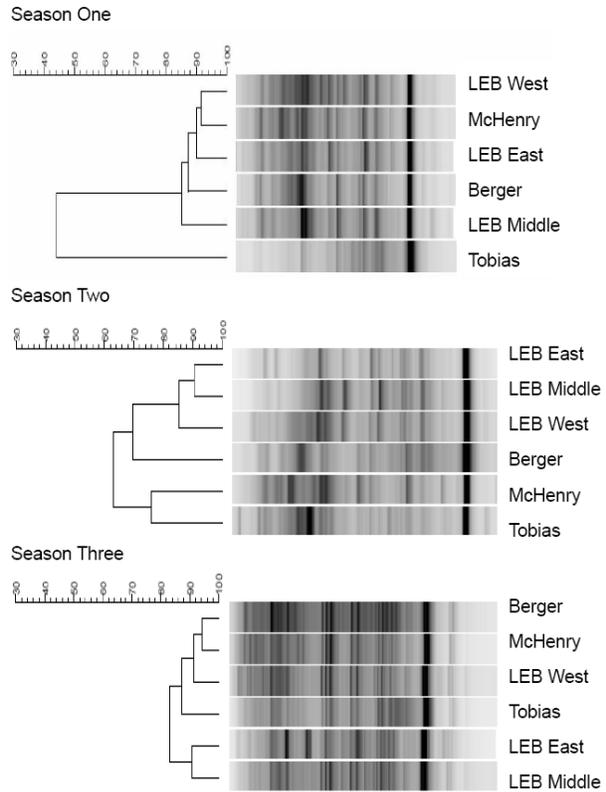
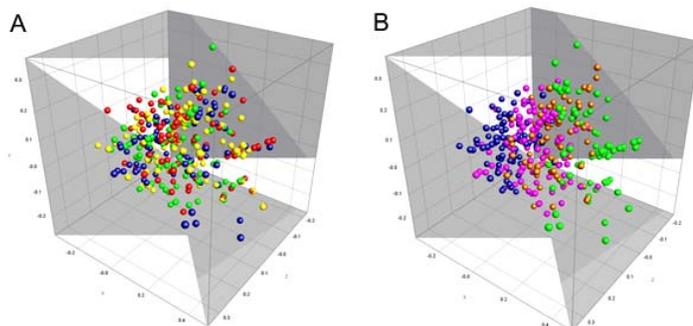


Fig. 6. Dendrograms showing relationships between the *E. coli* community in each ditch and the beach (LEB west, LEB middle and LEB east) throughout the three sampling seasons. The scale indicates the percent similarity of *E. coli* community DGGE fingerprints based on Pearson's correlation coefficient.

Fig. 7. MDS analysis of *E. coli* isolate fingerprints based on (A) location and (B) month of sampling. Panel A legend: yellow, LEB; red, Berger Ditch; blue, McHenry Ditch; green, Tobias Ditch. Panel B legend: blue, June; orange, July; green, August; purple, September.



For example, a temporal pattern of *E. coli* phylotypes was detected by MDS analysis of the BOX-PCR fingerprints (Fig. 7B) suggesting that BOX-PCR analysis was able to classify *E. coli* isolates on a temporal (month of collection) scale. Furthermore, Jackknife analysis showed that the *E. coli* isolates were correctly classified to the month during which they were collected at rates ranging from 69% to 86% (Table 2B).

A

	Tobias Ditch	Berger Ditch	McHenry Ditch	LEB
Tobias Ditch	70.93	26.53	16.46	26.32
Berger Ditch	17.44	57.14	15.19	13.68
McHenry Ditch	4.65	4.08	60.76	53.68
LEB	6.98	12.24	7.59	6.32

B

	June	July	Aug.	Sept.
June	78.65	7.78	0.00	12.77
July	12.36	73.33	12.94	12.77
Aug.	0.00	10.00	85.88	5.32
Sept.	8.99	8.89	1.18	69.15

Table 2. Jackknife analysis of *E. coli* isolates based on (A) the location from which they were collected and (B) month during which they were collected.

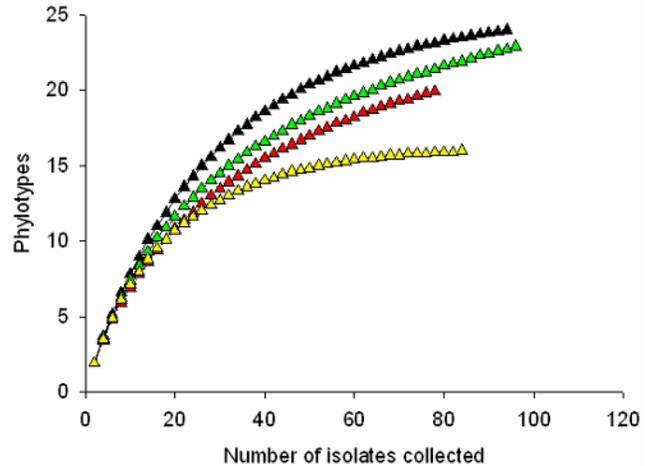


Fig. 8. Rarefaction analyses of *E. coli* isolates collected from each ditch and the beach during Season Three. Legend: black, LEB; green, Berger Ditch; red, McHenry Ditch; yellow, Tobias Ditch.

As with all library dependent analyses, it was necessary to determine if our *E. coli* library was of appropriate size. To test if the number of *E. coli* isolates selected from each site was large enough to draw meaningful conclusions about the BOX-PCR analysis, and therefore the validity of the DGGE analysis, rarefaction analysis was used. As shown in Fig. 8, rarefaction curves for all *E. coli* collections showed a strong trend towards saturation (slope of zero). This suggested that the number of *E. coli* in the four isolate libraries was reasonable and more sampling would not have appreciably increased the number of *E. coli* phylotypes.

DISCUSSION AND CONCLUSION

The purpose of this study was to determine the geographic source of bacterial pollution impacting LEB in MBSP over the course of one year. This was performed by characterizing *E. coli* density and community structure in Lake Erie and several drainage ditches thought to contribute bacterial pollution. While all sites exhibited varying densities throughout the year, the *E. coli* density in LEB water (10^1 to 10^3 CFU 100 mL^{-1}) was consistently one to two orders of magnitude lower than densities in the three ditches (10^1 to 10^5 CFU 100 mL^{-1})(Fig. 5). This observation was not surprising, as it is likely that *E. coli* exiting the ditch onto the foreshore beach was either diluted by lake water or was associated with sediments that dropped out of the water column close to the east side of the beach. Those populations that remained in deep water were likely less detectable, as less sediment resuspension will occur in deeper water. For example, Whitman and Nevers (2003) observed that the *E. coli* density of water collected at 90 cm depth of water was about half of that at 45 cm depth. The variation observed in both *E. coli* density and community structure was expected, given the position of LEB on Lake Erie. LEB is located in the Western Basin of Lake Erie, which is shallower than the central and eastern portions of the lake. The average depth of Lake Erie is 19 meters, however, the average depth in the Western Basin is only 7.4 meters (USEPA, 2006). This discrepancy in water level will result in greater impact of currents with regard to water movement and sediment resuspension and movement. This dynamic will have a profound impact on water quality since sediment can harbor fecal bacteria and be resuspended in turbulent water (Obiri-Danso and Jones, 1999; Jamieson et al., 2005).

Previous work showed that the *E. coli* density was consistently higher at the east end of LEB than at the west end (Sigler, unpublished data), which suggested that Berger Ditch, located at the east end of the beach, was a principal contributor to fecal pollution in LEB. During Season Three, we observed a similar phenomenon (Fig. 4), as *E. coli* densities were significantly higher at the east end of the beach than at the west end ($p=0.047$). While this result suggested a role of Berger Ditch in the elevated *E. coli* densities in LEB during Season Three, no significant difference in *E. coli* densities was observed across the beach during Seasons One and Two ($p=0.48, 0.36$, respectively). Several possible reasons exist for this seasonal contribution. Berger Ditch is the longest ditch surveyed in this study, and originates in Wolf Creek, which has been shown to harbor high densities of *E. coli* ($\sim 46,000$ CFU/100 mL) (Anonymous, 2003). Therefore, it is a logical assumption that Berger Ditch might experience a greater degree of potential fecal inputs than the other ditches. For example, agricultural inputs, in particular those associated with biosolids application, can dramatically increase the *E. coli* density in ditches that receive drain water from amended fields. Land application of biosolids is a popular practice in northwest Ohio, and is currently performed on several fields adjacent to Berger Ditch. Scarbro (2006) found that densities of *E. coli* in water draining into Berger Ditch from a biosolids-amended field increased from 100 to approximately 8,500 CFU 100 mL^{-1} . These *E. coli* were detected as long as three months after the biosolids application. Since Berger Ditch is not only the longest of the ditches, but also carries the highest volume of water to Lake Erie compared with the other two ditches, increased animal activity (waste inputs from muskrats, possum, groundhogs, deer, wild ducks,

birds and geese) (Struffolino, personal communication) during Season Three (with increasing temperature) would further separate its contribution from the other ditches. For example, animals are more active in the summer than in winter, as food is more plentiful than in cooler months. Further separation of Berger Ditch during Season Three likely resulted from increased rainfall, which combined with the large size of Berger Ditch, led to greater overall discharge into the lake when compared of McHenry- and Tobias Ditches. Under these conditions, the output from Berger Ditch during Season Three could have resulted in higher *E. coli* densities at the east end of the beach, while simple dilution of the bacteria as they were transported to the west end resulted in lower densities at the other beach sites.

E. coli densities in both beach and ditch samples during Season Three was one to two orders of magnitude greater than densities observed in Seasons One and -Two (Fig. 5). This can be explained primarily by the frequency and amount of rainfall that occurred during Season Three. NOAA rainfall data indicated that 49 rainfall events occurred during Season Three, resulting in 45 cm of precipitation, which was higher than the rain frequency and volume during Season One (42 events, 26.4 cm) and Two (45 events, 31.2 cm). Our data revealed a significant correlation between rainfall amount and *E. coli* densities at the beach and in Berger- and McHenry Ditches ($r=0.68, 0.8, 0.67$, respectively). It has been shown that storm water runoff can increase water turbidity by disturbing the underlying sediment/sand in ditches and beach water (Atherholt et al., 1998) and release and resuspend previously trapped *E. coli* and other potential pathogenic bacteria into the water column (Grimes, 1975). Furthermore, stormwater/rainfall can also transport contaminants, such as agriculture manure, and wildlife fecal material into ditches and streams that eventually empty into receiving waters (Mallin, et al., 2000).

It was previously assumed that a significant correlation between *E. coli* densities at the beach and a specific ditch might indicate a source-sink relationship. During Season One, only the *E. coli* density in Berger Ditch was correlated ($r=0.981, p<0.05$) with that of the beach, while densities in all three ditches correlated with those from the beach during Season Two. However, during Season Three, no correlation was observed between any of the three ditches and the beach. The inconsistency of correlation among the three seasons suggested that inputs other than the three ditches played a role in the *E. coli* densities observed in this study. The possible inputs are summarized below:

Lévesque et al (1993, 2000) studied the influence of **gull droppings** on the composition and number of potentially pathogenic bacteria in beach water. They found that *Staphylococcus aureus*, *Aeromonas spp.*, *Campylobacter spp.*, *Salmonella spp.*, and fecal coliforms were major microflora in gull droppings, while total bacteria densities were variable, ranging from 10^4 to 10^8 CFU g^{-1} (Lévesque et al., 2000). LEB is a habitat for both resident- and migratory birds, including gulls and roosting waterfowl, and their droppings are readily visible on the beach. More importantly however, gulls are rarely found visiting or inhabiting the ditches, which suggests that the presence of gulls might reduce the agreement between ditch- and beach *E. coli* densities and community structure.

Another potential factor driving differences between ditch and beach *E. coli* communities is the role of **beach sand** as a pathogen reservoir. Microorganisms are a significant component of beach sand. *E. coli* can sustain population densities of over 2,000 CFU 100 mL⁻¹ in beach sand during summer months without continual lake, human or animal input (Whitman and Nevers, 2003). The study also found that *E. coli* densities in sand were, on average, 5 to 10 times higher than densities in adjacent swimming waters. In addition, a significant correlation was observed between *E. coli* in beach sand and the *E. coli* concentration in the water on the following day, suggesting that *E. coli* in beach sands influence water densities. Beyond simply inhabiting beach sands, *E. coli* can grow in sand (Alm et al., 2006). In the current study, *E. coli* density in the beach water was similar to or higher than densities in the ditches in March when there was no rain, indicating other sources such as sand may play a role in the elevated *E. coli* in the beach water.

Algae were shown to provide protection to pathogens, thereby serving as a reservoir that can enhance *E. coli* densities in beach water (McLellan and Jensen, 2005). In the current study, during the August and September samplings, algal mats were observed along the shore of Lake Erie Beach (personal observation). These algae were in dark blue in color, filamentous, and formed an extensive mat that covered the foreshore beach. Algal mats maintain high moisture content, protect bacteria from sunlight and are rich in nutrients, thus providing an appropriate habitat for bacterial survival and growth (Whitman et al., 2003; Byappanahalli et al., 2003). The contribution of pathogens harbored in algal mats is likely limited to the beach environment, as mats were never observed in or around any of the ditches, thereby providing another potential bacteria source exclusive to the beach.

Finally, **humans and their related activities** can impact the bacteriological quality of recreational water. Since human activity is concentrated on the beach at MBSP and not in the ditches, they represent an exclusive source of *E. coli* for the beach. Human inputs result from poor hygiene (improperly disposed of baby diapers or poorly toilet trained children) (Keene, et al., 1994) and motorboats that disturb and resuspend the sediment near the beach (Hilton and Phillips, 1982), resulting in the liberation of trapped bacteria in the sediments into the beach water (Grimes, 1975). Bathing activities can also elevate *E. coli* in beach water via shedding of indicator bacteria or resuspension of contaminated sediments (Obiri-Danso and Jones, 1999). Furthermore, people who spend time both on the beach and in the water can also play a role in transporting bacteria from the sand to the water (Sherry, 1986).

The results of *E. coli* community fingerprinting analyses by DGGE showed that the *E. coli* communities in three ditches and the beach were similar (with the exception of Tobias Ditch during Season One) but seasonally dynamic (Fig. 6). Because these results were based on *E. coli* community composition and not simply on densities, the relationships

between the *E. coli* communities observed in beach- and ditch water were likely more reliable than those based on *E. coli* densities. This is because densities cannot provide a reliable assessment of from where the bacteria originated. During Seasons One and Three, the fingerprints of *E. coli* communities from all ditches were at least 85% similar to those from Lake Erie Beach, with the exception of Tobias Ditch in Season One (discussed below). The similarity of *E. coli* community fingerprints among the ditches and the beach was expected because the three ditches were all located in close approximate to Maumee Bay State Park. Berger Ditch (the easternmost of the ditches) and Tobias Ditch (the westernmost) are separated by approximately 2200 m (McHenry Ditch lies between Berger- and Tobias Ditches). Along this ca. 2 km of shoreline is found a mixture of impervious surfaces, residential areas and farmland, and no single land use type defines the drainage area of any one of the ditches. Therefore, it is possible that common sources of *E. coli* could have influenced each of the ditches. Common sources in this region might include residential septic tank leaking, wild animal input, paved surface runoff, and agricultural runoff and leaching. For example, a survey by the Lucas County Public Health Department suggested 40% of the residential septic systems in the drainage area of MBSP might not function properly (Sinsabaugh and Glatzer, 1999). It was also reported that 79% of all agricultural lands in the Maumee River basin were cropland (Hess, 1995) and a considerable proportion of these are subject to biosolids application. Inputs from each of these sources are likely enhanced by significant rainfall. This is evident particularly during Season Three (Fig. 6), as the increased rainfall appears to contribute not only the bacterial load into Lake Erie, but also the similarity of the *E. coli* communities contributed by each ditch.

During Season Two, the similarity of the beach *E. coli* community fingerprints to those from the ditches decreased, likely due to any of a number of environmental factors. The average air temperature in February and March was -1°C and 3°C , respectively, with at least one snow precipitation event (18 cm). The low temperature might have lowered the average *E. coli* densities in the three ditches by lowering the inputs into the ditches (e.g. lowered activities of animals), and ultimately resulting in little contribution to the beach. Because snowfall does not infiltrate to the soil or provide immediate runoff, a lag in the ditch discharge of snow precipitation will occur until snowmelt (Yang et al., 2003). When temperatures increase, with a concomitant increase in rainfall (March and April), accumulated sediments with attached *E. coli* are likely released from the ditch bottoms and discharged into the lake. This build-up and eventual release might drive differences in the *E. coli* communities among the ditches and the beach, especially if the built up sediments harbor *E. coli* from numerous, temporally driven sources.

Compared to community fingerprints from McHenry- and Tobias Ditches, those from Berger Ditch were usually higher in similarity to fingerprints from the beach, suggesting the consistent contribution of Berger ditch to the pollution at the beach. The similarities of the *E. coli* community fingerprints between Berger Ditch and the beach were 87%, 70% and 93% in Seasons One, -Two and -Three respectively (Fig. 6), while McHenry- (88%, 65%, and 83%) and Tobias Ditches (46%, 62%, and 83%) were less similar. The high similarity of *E. coli* communities from Berger Ditch to those from the beach was expected, primarily because of factors characterizing Berger Ditch that were mentioned

above. Overall, the data suggested that a yearlong contribution from Berger Ditch was occurring. However, the contribution of each potential point source within the ditch has not been studied and requires further investigation. Since primary interest concerning bacterial pollution has surrounded Berger Ditch, the role of McHenry Ditch in pollution at the beach has not been investigated. It is likely however, that McHenry Ditch is at least a modest contributor of bacterial pollution, as fingerprints from Seasons One and Three were over 88% similar to those from the beach.

The similarity of *E. coli* fingerprints in *E. coli* collected from Tobias Ditch to those from the beach increased from 46% to 83% throughout the three sampling seasons. The primary driver of this difference is likely the small relative size of Tobias Ditch. Because of the comparatively low volume of water flowing through Tobias Ditch, it is subject to occasional periods of little or no flow. This was especially evident during Season One, as little to no water was flowing during three out of the four sampling events. It is clear that this low level of water activity limited the contribution of Tobias Ditch to the pollution at Lake Erie Beach during Season One. In contrast, more frequent and heavier rainfall (Fig. 5), probably led to more consistent water flow in the ditch during Seasons Two and Three. As was described above, it is likely that periods of little or no water flow allowed for the accumulation of *E. coli* in ditch sediments that were released en-masse during periods of heavy rainfall (e.g. Season Three), resulting in contribution to the bacterial pollution at Lake Erie Beach.

Since this study represents the first application of DGGE to identify potential geographic sources of bacterial pollution, it was prudent to validate the results by analyzing a subset of the samples with an established, library-dependant method. Therefore, BOX PCR (Dombek et al., 2000) was used to confirm the results of the DGGE analysis of *E. coli* communities. The 358 isolates collected during Season Three represented a sampling of the isolates present throughout the entire swim season. Three analyses associated with BOX-PCR, including cluster analysis, multidimensional scaling analysis and jackknife analysis, were used to analyze the BOX-PCR results and validate the results of DGGE. Overall, the analysis of BOX-PCR fingerprints confirmed the results of the Season Three DGGE analysis, revealing that the *E. coli* communities from each site were similar, and geographically non-discernable (see attachment). The results of MDS and jackknife analyses showed that the *E. coli* communities could not be classified based on location of sampling (Fig. 7A) but were easily classified based on season of sampling (Fig. 7B). We also used jackknife analysis tests the ability of BOX-PCR fingerprinting to classify the *E. coli* isolates into their environment of origin (Dombek et al., 2000). Both analyses revealed high similarity of the *E. coli* communities sampled during Season Three, regardless of geographic origin or month of sampling, which confirmed the results of the DGGE analysis.

To accurately validate the DGGE analysis, MDS and jackknife analyses were dependent on an *E. coli* isolate library of appropriate size. The number of *E. coli* phylotypes represented in our library was determined by calculating the number of fingerprint clusters that formed above an empirically derived similarity threshold of 85% (described in Materials and Methods). In other words, if the similarity of *E. coli* isolate fingerprints

was above 85%, they were considered identical phylotypes. Using a similar method of threshold determination, Kinzelman et al. (2004) found 87% similarity to be an appropriate threshold value for computer recognition of identical fingerprint patterns. Rarefaction analysis was used to ultimately test the validity of the library size. For each library (based on the location of origin), the rarefaction curves became asymptotic, indicating that more intensive sampling would not have added additional phylotypes to the collection (Fig. 8). While the sample size of *E. coli* isolates necessary for geographic source tracking is not standardized, reports suggest that host origin library sizes can vary from one hundred to more than two thousand *E. coli* isolates from human and various animals (Dombek et al., 2000; Hartel et al., 2002; McLellan et al., 2003; Johnson et al., 2004). The number of isolates necessary is dependant on the variation in genetic structure (Gordon, 1997; Souza et al., 1999), host specificity (Gordon and Lee, 1999), and temporal changes in alleles (DNA sequences) (Pupo and Richardson, 1995; Gordon, 1997). Despite reports of large library sizes, increasing the size of libraries does not always improve classification and can even result in decreased discriminating capability (Johnson et al., 2004), possibly as a result of the presence of clonal strains (multiple strains collected from a single host) (McLellan, 2003), or errors due to gel-to-gel variation (Johnson et al., 2004).

Although numerous studies have attempted to identify sources of bacterial pollution, the majority have drawn conclusions based on a limited sampling period, varying from a one time sampling to sampling collected over several months (Dombek et al., 2000; Kinzelman et al., 2004; Ram et al., 2004). The current study and others highlight the importance of variability in environmental factors as controls over *E. coli* densities. Factors such as temperature, rainfall, light, water level, available nutrients, salinity, and environmental pollutants all influence the density and strain composition of fecal contamination over time and affect the results of source identification (Lipp et al., 2001; Kinzelman et al., 2004; Whitman et al., 2004; Tassoula, 1997; Korhonen and Martikainen, 1991; Pathak and Bhattacharjee, 1994). For example, during Season Two, DGGE analysis revealed low similarity between communities from the beach and each ditch, indicating that the ditches might not represent the principal contributors to the bacterial pollution. In contrast, analysis of Season Three samples showed that all ditch communities were similar to those from the beach, and likely contributing to the pollution. Therefore, understanding the true dynamics and ecology of bacterial pollution requires long-term sampling of fecal indicator bacteria.

Based on the data in this study, we found both the *E. coli* density and community structure in Lake Erie Beach and the three ditches were dynamic and seasonally variable during the year-long sampling period. The weak correlation between *E. coli* density in the ditch and beach water demonstrated that monitoring bacterial densities is limited in its utility for bacterial source tracking and is best combined with complimentary methods. One such method highlighted in this study is DGGE, which effectively described the relationship of *E. coli* communities in the ditches and the beach. While results suggested that all three ditches sampled were involved in the bacterial pollution at some time, the contribution was seasonally-based. The contribution from Tobias Ditch, in particular, was highly rainfall dependent, while Berger Ditch was the most consistent

contributor of the three ditches, regardless of rainfall. BOX-PCR of *E. coli* libraries collected during Season Three validated the results of DGGE analysis, indicating that DGGE was a reliable method for rapidly screening the putative geographic sources of bacterial pollution in Lake Erie Beach.

REFERENCES

Alm, E.W., J. Burke, and E. Hagan. 2006. Persistence and potential growth of the fecal indicator bacteria, *Escherichia coli*, in shoreline sand at Lake Huron. J. Great Lakes Res. 32:401-405.

Anonymous, 2003. Wolf creek bacterial impact on Maumee Bay state park beach. Summary report, Accessed on November 25, 2006 at http://www.tmacog.org/Environment/maumee_bay_RFP/Wolf_Creek_Summary.pdf

Atherholt, T.B., M.W. LeChevallier, W.D. Norton, and J.S. Rosen. 1998. Effect of rainfall on Giardia and Cryptosporidium. J. Am. Water Works Assoc. 90:66-80.

Bej, A.K., J.L. Dicesare, L. Haff and R.M. Atlas. 1991. Detection of *Escherichia coli* and *shigella spp.* in water by using the polymerase chain reaction for identifying fecal pollution sources in a multi-use estuary. Wat. Res 37:0=909-913

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

Dice, L.R. 1945. Measure of the amount of ecological association between species. Ecol. 26:297-302

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.

Dufour, A.P., O. Evans, T.D. Behymer and R. Cantú. 2006. Water ingestion during swimming activities in a pool: A pilot study. J. Water Health 04: 425-430.

Farnleitner, A.H., N. Kreuzinger, G.G. Kavka, S. Grillenberger, J. Rath and R.L. Mach. 2000a. Simultaneous detection and differentiation of *Escherichia coli* populations from environmental freshwaters by means of sequence variations in a fragment of the β -D-glucuronidase gene. Appl. Environ. Microbiol. 66:1340-1346

Francy, D.S., P. Struffolino, A.M.G. Braday and D.F. Dwyer. A spatial multivariable approach for identifying proximate sources of *Escherichia coli* to Maumee Bay, Lake Erie, Ohio. Open-File Report 2005–1386. Accessed on July 11, 2006 at <http://pubs.usgs.gov/of/2005/1386/pdf/OFR%202005-1386.pdf>

Gordon, D.M. 1997. The genetic structure of *Escherichia coli* populations in feral house mice. Microbiology 143:2039-2046.

Gordon, D. M., and J. Lee. 1999. The genetic structure of enteric bacteria from Australian mammals. Microbiology 145:2673-2682.

- Grimes, D.J. 1975. Release of sediment-bound fecal coliforms by dredging. *Appl. Microbiol.* 29:109-111.
- Hagedorn, C., S.L. Robinson, J.R. Filtz, S.M. Grubbs, T.A. Angier, and R.B. Reneau, Jr. 1999. Using antibiotic resistance patterns in the fecal streptococci to determine sources of fecal pollution in a rural Virginia watershed. *Appl. Environ. Microbiol.* 65:5522–5531.
- Hancock, D.D., T.E. Besser, D.H. Rice, D.E. Herriott, P.I. Tarr. 1997. A Longitudinal Study of *Escherichia coli* O157 in Fourteen Cattle Herds. *Epidemiol. Infect.* 118:193-195.
- Hartel, P. G., J. D. Summer, J. L. Hill, J. Collins, J. A. Entry, and W. I. Segars. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. *J. Environ. Qual.* 31:1273–1278.
- Hess, P. 1995. Transect; Software for managing cropland survey data, 1996 user guide: West Lafayette, Ind., Purdue University.
- Hilton, J., and G.L. Phillips. 1982. The effect of boat activity on turbidity in a shallow broadland river. *J. Appl. Ecology.* 19:143-150.
- 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.
- 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.
- Keene, W.E., J.M. McAnulty, F.C. Hoesly, L.P.Jr. Williams, K. Hedberg, G.L. Oxman, T.J. Barrett, M.A. Pfaller and D.W. Fleming. 1994. A swimming-associated outbreak of hemorrhagic colitis caused by *Escherichia coli* O157:H7 and *Shigella sonnei*. *N. Engl. J. Med.* 331:579-84.
- Khatib, L.A., Y.L. Tsai, and B.H. Olson. 2003. A biomarker for the identification of swine fecal pollution in water, using the STII toxin gene from enterotoxigenic *Escherichia coli*. *Appl. Microbiol. Biotech.* 63:231-238.
- 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 and health.* 2 :103-114
- Korhonen, L.K., and P.J. Martikainen. 1991. Survival of *Escherichia coli* and *Campylobacter jejuni* in untreated and filtered lake water. *J. Appl. Bacteriol.* 71:379-382

Kueh, C.S.W., T.Y. Tam, T. Lee, S.L. Wong, O.L. Lloyd, I.T.S. Yu, T.W. Wong, J.S. Tam and D.C.J. Bassett. 1995. Epidemiological study of swimming-associated illnesses relating to bathing-beach water quality. *Wat. Sci. Tech.* 31:1-4

Lauber, C.L. 2001. Characterization of diffuse sources of fecal pollution and pathogenic constituents of *Escherichia coli* populations of Maumee Bay state park. Master thesis, University of Toledo.

Lévesque, B., P. Brousseau, P. Simard, E. Dewailly, M. Meisels, D. Ramsay, and J. Joly. 1993. Impact of the ring-billed gull (*Larus delawarensis*) on the microbiological quality of recreational water. *Appl. Environ. Microbiol.* 59:1228-1230.

Lévesque, B., P. Brousseau, F. Bernier, E. Dewailly, and J. Joly. 2000. Study of the bacterial content of ring-billed cull droppings in relation to recreational water quality. *Wat. Res.* 34:1089-1096.

Lipp, E.K., R. Kurz, R. Vincent, C. Rodriguez-Palacios, S. R. Farrah, and J. B. Rose. 2001. The effects of seasonal variability and weather on microbial fecal pollution and enteric pathogens in a subtropical estuary. *Estuaries.* 24:266-276.

Louws, F.J., D.W. Fulbright, C.T. Stephens, F.J. de Bruijn. 1994. Specific genomic fingerprints of phytopathogenic *Xanthomonas* and *Pseudomonas* pathogens and strains generated with repetitive sequences and PCR. *Appl. Environ. Microbiol.* 60:2286 -2295.

Mallin, M.A. K.E. Williams, E. C. Escham, and R. P. Lowe. 2000. Effect of human development on bacteriological water quality in coastal watersheds. *Ecol. Appl.* 10:1047-1056.

McLellan, S.L., A.D. Daniels, and A.K. Salmore. 2003. Genetic characterization of *Escherichia coli* populations from host sources of fecal pollution using DNA fingerprinting. *Appl. Environ. Microbiol.* 69:2587-2594.

McLellan, S.L. and E.T. Jensen. 2005. Identification and quantification of bacterial pollution at Milwaukee county beaches, Great Lakes WATER Institute Technical Report

Mead, P.S., and P.M. Griffin. 1998. *Escherichia coli* O157:H7. *Lancet.* 352:1207-1212.

Muyzer, G., E.C. De Waal and A.G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* 59:695-700

Muyzer, G. and K. Smalla. 1998. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek* 73:127-141

Obiri-Danso, K., and K. Jones. 1999. The effect of a new sewage treatment plant on faecal indicator numbers, campylobacters and bathing water compliance in Morecambe

Bay. J. Appl. Microbiol. 86:603-614.

ODH (Ohio Department of Health), 2006. Beach sampling results from 2001 to 2005. Accessed on July 6, 2006 at <http://www.odh.ohio.gov/ASSETS/8E877F6BD85F4BF988C02F92A21927FF/maumeeri e.PDF>

Pathak, S.P., and J.W. Bhattacharjee. 1994. Effect of pollutants on survival of *Escherichia coli* in microcosms of river water. Bull. Environ. Contam. Toxicol. 53:198-203

Pupo, G.M., and B.J. Richardson. 1995. Biochemical genetics of a natural population of *Escherichia coli*: seasonal changes in alleles and haplotypes. Microbiology 141:1037-1044.

Rademaker, J.L.W., F.J. Louws and F.J. de Bruijn. 1998. Characterization of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting. p. 1-27. In A. D.L. Akkermans, J.D. van Elsas, and F.J. de Bruijn (ed.), Molecular microbial ecology manual, supplement 3. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Ram, J.L., R.P. Ritchie, J. Fang, F.S. Gonzales, and J.P. Seleguean. 2004. Sequence-based source tracking of *Escherichia coli* based on genetic diversity of β -Glucuronidase. J. Environ. Qual. 33:1024-1032.

Ritter, K.J., E. Carruthers, C.A. Carson, R.D. Ellender, V.J. Harwood, K. Kingsley et al. 2003. Assessment of statistical methods used in library-based approaches to microbial source tracking. J. Water Health. 1:209-223.

Sabat, G., P. Rose, W.J. Hickey, and J.M. Harkin. 2000. Selective and sensitive method for PCR amplification of *Escherichia coli* 16S rRNA genes in soil. Appl. Environ. Microbiol. 66:844-849.

Scarbro, L.S. 2006. Master thesis. Microbial Transport through the Vadose Zone of a Biosolid Amended Cropland: Implications for land application of Class B biosolids and surface water quality. University of Toledo.

Sherry, J.P. 1986. Temporal distribution of faecal pollution indicators and opportunistic pathogens at a Lake Ontario bathing beach. J. Great Lakes Res. 12:154-160

Sigler, W.V., and J. Zeyer, 2002. Microbial diversity and activity along the forefields of two receding glaciers. Microb. Ecol. 43:397-407.

Sigler, W.V., C. Miniaci and J. Zeyer. 2004. Electrophoresis time impacts the denaturing gradient gel electrophoresis-based assessment of bacterial community structure. J. Microbiol. Methods 57:17-22

Sigler, W.V. and L. Pasutti. 2006. Evaluation of denaturing gradient gel electrophoresis to differentiate *E. coli* communities in secondary environments. *Environ. Microbiol.* 8:1703-1711.

Sinsabaugh, R.L., and L. Glatzer. 1999. Tracing diffuse sources of fecal contamination through phenotypic and genotypic characterization of *E. coli* isolates. Final project report to Lake Erie protection fund.

Sohngen, B., F. Lichtkoppler, and M. Bielen, 1999, The value of Lake Erie beaches, The Ohio State University and Sea Grant publication FS-078

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

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

USEPA, 2000b. Beaches environmental assessment and coastal health act of 2000. PUBLIC LAW 106-284—OCT. 10, 2000.

USEPA. 2000c. Improved enumeration methods for the recreational water quality indicators-Enterococci and *Escherichia coli*, Washington, D.C., EPA/821/R-97/004 March 2000.

USEPA, 2006. Accessed on Jan. 17th, 2007 at <http://www.epa.gov/glnpo/basicinfo.html#erie>

Van Donsel, D.J., E.E. Geldreich, and N.A. Clarke. 1967. Seasonal variations in survival of indicator bacteria in soil and their contribution to storm-water pollution. *Appl. Microbiol.* 15:1362-1370.

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.

Whitman, R.L., D.A. Shively, H. Pawlik, M.B. Nevers, and M.N. Byappanahalli. 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 M.B. Nevers. 2004. *Escherichia coli* sampling reliability at a frequently closed Chicago Beach: monitoring and management implications. *Environ. Sci. Technol.* 38:4241-4246.

Whitman, R.L. M.B. Nevers, G.C. Korinek, and M.N. Byappanahalli. 2004. Solar and temporal effects on *Escherichia coli* concentration at a Lake Michigan swimming beach. *Appl. Environ. Microbiol.* 70:4276-4285.

Wiggins, B.A., R.W. Andrews, R.A. Conway, C.L. Corr, E.J. Dobratz, D.P. Dougherty, et al. 1999. Identification of sources of fecal pollution using discriminant analysis: Supporting evidence from large datasets. *Appl. Environ. Microbiol.* 65:3483-3486.

Wiggins, B.A., P.W. Cash, W.S. Creamer, S.E. Dart, P.P. Garcia, T.M. Gerecke, et al. 2003. Use of Antibiotic Resistance Analysis for Representativeness Testing of Multiwatershed Libraries. *Appl. Environ. Microbiol.* 69: 3399-3405.

Yang, D., D. Robinson, Y. Zhao, T. Estilow, and B. Ye. 2003. Streamflow response to seasonal snow cover extent changes in large Siberian watersheds. *J. Geophys. Res.* VOL. 108, NO. D18, 4578, doi:10.1029/2002JD003149

LAKE ERIE PROTECTION FUND

SMALL GRANT - FINAL ACCOUNTING

Grant Number: LEPF 04-21

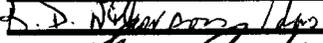
v2008

Budget Categories	Original Budget	Funds Spent	Current Balance	Matching Funds
A. Salaries & Wages				
	29255	28991	264	11114
B. Fringe Benefits				
	2634	2629	5	3705
Graduate Student Tuition				24458
C. Total Salaries & Benefits (A+B)	\$31,889.00	\$31,620.00	\$269.00	\$39,277.00
D. Non-expendable Equipment				
E. Expendable Materials & Supplies				
	15052	16357	-1305	
F. Travel				
	1500	997	503	
G. Services or Consultants				
H. Computer Costs				
I. Publications/Presentations				
	600	67	533	
J. All other direct costs				
K. Total Direct Costs (C thru J)	\$49,041.00	\$49,041.00	\$0.00	\$39,277.00
L. Indirect Costs				
	4904	4904	0	3928
Indirect Costs Under Recovery				28261
Total Costs (K + L)	\$53,945.00	\$53,945.00	\$0.00	\$71,466.00

Ohio Lake Erie Commission
 One Maritime Plaza, 4th Floor
 Toledo, OH 43604
 p 419-245-2514
 f. 419-245-2519
www.epa.state.oh.us/oleo

I certify that the grant expenditures listed and descriptions of the charges are true and accurate to the best of my knowledge. These expenditures represent approved grant costs that have been previously paid for and for which complete documentation is on file.

Project Director
 Authorizing Agent
 Fiscal Agent

	Date
	11/21/02
	11/28/07
	10/21/07