Repetitive element (REP)-polymerase chain reaction (PCR) analysis of *Escherichia coli* isolates from recreational waters of southeastern Lake Huron

Tanya Kon, Susan C. Weir, E. Todd Howell, Hung Lee, and Jack T. Trevors

Abstract: Repetitive element-polymerase chain reaction (REP-PCR) DNA fingerprinting and library-based microbial source tracking (MST) methods were utilized to investigate the potential sources of *Escherichia coli* pollution in recreational waters of southeastern Lake Huron. In addition to traditional sources such as humans, agriculture, and wildlife, environmentally persistent *E. coli* isolates were included in the identification library as a separate library unit consisting of the *E. coli* strains isolated from interstitial water on the beach itself. Our results demonstrated that the dominant source of *E. coli* pollution of the lake was agriculture, followed by environmentally adapted *E. coli* strains, wildlife, and then humans. A similar ratio of contributing sources was observed in all samples collected from various locations including the river discharging to the beach in both 2005 and 2006. The high similarity between the compositions of *E. coli* communities collected simultaneously in the river and in the lake suggests that tributaries were the major overall sources of *E. coli* to the lake. Our findings also suggest that environmentally adapted strains (EAS) of *E. coli* should be included as one of the potential sources in future microbial source tracking efforts.

Key words: beach, environment, *Escherichia coli*, microbial pollution, REP-PCR, surface water, survival, tracking, watershed.

Résumé : L'empreinte ADN par REP-PCR (*Repetitive element-polymerase chain reaction*) et l'identification des sources de contamination par MST (*Microbial source tracking*) à partir de banques ont été utilisées pour investiguer les sources potentielles de pollution par *E. coli* dans les eaux de baignade sud-est du Lac Huron (Canada). En plus des sources traditionnelles que sont l'humain, l'agriculture et la faune, des isolats persistants d'*E. coli* consistant en souches d'*E. coli* isolées des eaux interstitielles de la plage elle-même ont été inclus dans la banque d'identification comme unités indépendantes. Nos résultats ont démontré que la source dominante de pollution par *E. coli* du lac était l'agriculture, suivie par les souches d'*E. coli* adaptées à l'environnement, la faune et finalement, l'humain. Un ratio similaire de sources contribuant à la pollution a été observé dans tous les échantillons recueillis à différents endroits, y compris à la décharge de la rivière en 2005 et 2006. Le haut niveau de similarité dans la composition des communautés d'*E. coli* recueillies simultanément dans la rivière et le lac suggère que les affluents sont les sources générales majeures d'*E. coli* du lac. Nos résultats suggèrent aussi que les souches d'*E. coli* adaptées à l'environnement devraient être incluses parmi les sources potentielles de contamination microbienne dans les protocoles de MST futurs.

Mots-clés : plage, environnement, *Escherichia coli*, pollution microbienne, REP-PCR, eaux de surface, survie, repérage, ligne de partage des eaux.

[Traduit par la Rédaction]

Introduction

The microbial pollution of recreational water is a serious environmental problem that is of considerable public health concern. Human and other activities occurring on or adjacent to a beach can be responsible for lake water pollution. Identifying the sources of this pollution is important for assessing public health risks and deciding what management strategies could be used in a region that is susceptible to such risks. Microbial source tracking (MST) studies have been developed to address these issues (Simpson et al. 2002). Traditionally, the studies have assumed a direct link between the presence of *Escherichia coli* in recreational waters and the originating source(s). Many studies have limited their focus to well-known sources such as agriculture, sewage treatment plants, combined sewer overflows, shorebirds, wildlife, and pet droppings on a beach (George et al. 2004; Fogarty et al. 2003; Saini et al. 2003), and have assumed limited bacterial survival between the sources and surface waters. However, more recent studies have demonstrated that high bacterial counts in surface waters along

Received 21 October 2008. Accepted 24 October 2008. Published on the NRC Research Press Web site at cjm.nrc.ca on 27 March 2009.

T. Kon, H. Lee,[1] and J.T. Trevors.[2] University of Guelph, Guelph, ON N1G 2W1, Canada.
S.C. Weir and E.T. Howell. Ontario Ministry of the Environment, 125 Resources Rd., Toronto, ON M9P 3V6 Canada.

[1]Corresponding author (e-mail: hlee@uoguelph.ca).
[2]Corresponding author (e-mail: jtrevors@uoguelph.ca).

shorelines may be a result of bacterial survival in beach sand, which can contribute to high indicator bacterial counts in the absence of fecal input (Alm et al. 2003; McLellan and Salmore 2003). Environmental survival of *E. coli* strains outside of animal hosts has been reported in subtropical waters (Anderson et al. 2005), tropical soils (Byappanahalli and Fujioka 1998), temperate soils (Ishii et al. 2006), and beach sand (Alm et al. 2006; Ishii et al. 2007; Kon et al. 2007*a*, 2007*b*; Whitman and Nevers 2003). While little is known about the mechanism(s) by which *E. coli* may adapt to such an environment, it is now established that some *E. coli* strains are capable of persisting in the secondary environment (Beversdorf et al. 2007). In the literature they are called naturalized (Ishii et al. 2006) or environmentally adapted strains (EAS) (Kon et al. 2007*a*).

We hypothesize that the EAS of *E. coli* represent a significant source of water pollution. It is not known if sources of increased *E. coli* counts represent recent inputs or survival in the interstitial environment and periodic release from sand as suggested by some authors (Ishii et al. 2007). The characterization of continuous, localized sources of microbial indicators is essential to complement current water-monitoring strategies. Differentiation between freshly introduced and resident *E. coli* strains at the shore could assist in understanding the microbial ecology of the beach environment and water quality.

The objective of this research was to determine the sources of lake water pollution at a beach in southeastern Lake Huron and to look at EAS as one of the potential microbial sources. To investigate this possibility we utilized a library-based microbial source tracking method known as repetitive element-polymerase chain reaction (REP-PCR). In this approach the sources of pollution are determined by comparing DNA profiles of *E. coli* isolates from contaminated waters with profiles of *E. coli* isolated from known suspected sources collected within the same geographic area or watershed. A database of known isolates, referred to as a "library", is required for this method (United States Environmental Protection Agency (US EPA) 2005). REP-PCR DNA fingerprinting is a widely accepted technique for distinguishing between different sources of water contamination using a library-based approach because it is reproducible, rapid, and highly discriminatory (Dombek et al. 2000; Olive and Bean 1999). The limitations of this method are its dependency on the library and geographical variability from 1 watershed to another (Seurinck et al. 2005). To address this issue we generated a REP-PCR library based on *E. coli* isolates obtained locally from various agricultural, human, wildlife, and environmental sources within the same watershed to determine the sources of recreational water pollution at the adjacent shores of Lake Huron. Along with the traditional human, agricultural, and wildlife source units, we generated an environmental source library unit that included *E. coli* isolates from the interstitial water of the study beach.

Materials and methods

Study site

The water samples were collected between May 2005 and November 2006 in the watershed of Eighteen Mile River and at the Ashfield Township Park beach and adjoining shoreline on the southeastern shore of Lake Huron (Fig. 1). The study area of the beach consisted of 2 parts: privately owned land and rural-type public beach. The beach is dry with sand and small gravel deposits as a substrate. It is backed by clay cliffs followed by the agriculturally dominated Eighteen Mile River watershed with small tributaries discharging into the lake. The Eighteen Mile River is the largest tributary within the study area. It discharges directly into the centre of the study area. The sampling area encompassed 5 km of the shoreline. Each sampling station had its own unique identifier and coordinates defined by a global positioning system (GPS).

Sample collection

The sampling strategy consisted of 3 parts: (*i*) a full survey of the lake covering the entire study area both shoreline and nearshore up to 4 km offshore about once every 2 months; (*ii*) roughly biweekly sampling of 5 nearshore lake stations near the mouth of the river; and (*iii*) biweekly sampling at the intensive-monitoring station in the river (simultaneously with the 5 lake stations) at a site not affected by the lake water that might occasionally come into the river.

Surface water (lake and river) samples were collected in sterile 300 mL bottles, leaving at least a 2.5 cm air space in each bottle according to previously described procedures (Ontario Ministry of the Environment 2004*a*). The collection of lake samples was performed as follows: (*i*) for lake surveys, at different depths by 3 water-monitoring crews, walking waist-deep at the shoreline of the lake, nearshore sampling from a vessel at the depth of 3 m starting from 200 to 1100 m, and up to a 4 km distance from the shore, and sampling from a small boat over an area between the shoreline and a 3 m depth of the lake; (*ii*) for regular biweekly sampling, only the waist-deep samples were collected from the lake concurrently with the sample from the intensive river monitoring site.

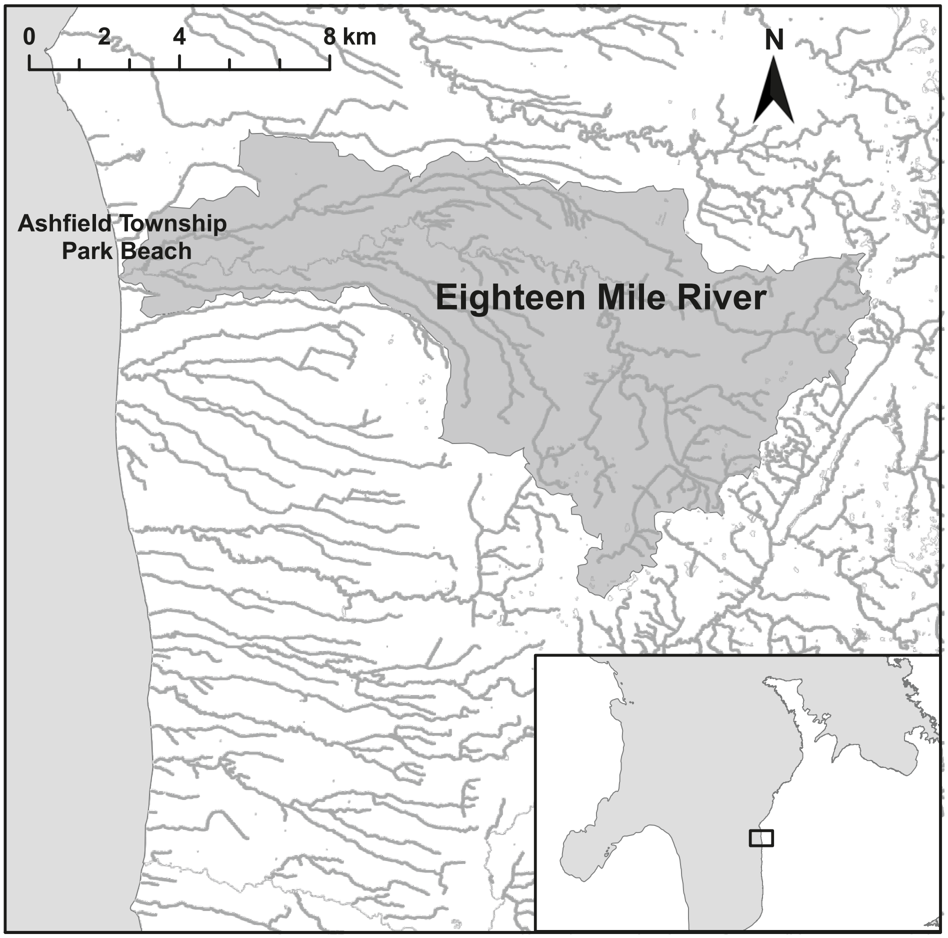
Environmental sources were isolates from interstitial (pore) water collected from the beach over the study area. Our previous studies showed high concentrations of *E. coli* in interstitial waters (Kon et al. 2007*a*), and this water is easily transportable to the swimming area in the lake. Sampling locations for interstitial water were excavated on the Ashfield Township Park beach with an alcohol-disinfected shovel to just below the water table at each station 25 cm from the observed swash zone. Interstitial water from these sampling locations was collected in the same type of sterile bottles as for surface water and analyzed for *E. coli* within 48 h. The analysis for *E. coli* was performed as described (Kon et al. 2007*a*).

Fecal material samples were collected using sterile plastic scoops and placed into sterile Whirl-Pak bags (Ontario Ministry of the Environment 2004*b*). Some samples were individual and some were composites of 5–10 individual samples. Agricultural samples were considered as composites from many animals because they came from manure storages. All samples were transported to the laboratory on ice at a temperature <10 °C and analyzed within 24 h.

Isolation of *E. coli* from water and fecal material samples

Water samples were subjected to membrane filtration (MF) within 24 h of collection. The water sample was passed

Fig. 1. Map of Eighteen Mile River watershed.



through a sterile 47 mm diameter cellulose ester disk filter (0.45 μm average pore size; PALL Life Sciences, Mississauga, Ontario). Filters were placed on mFC-BCIG agar (Difco, Sparks, Maryland; consisting of 10.0 g of tryptose, 5.0 g of proteose peptone, 3.0 g of yeast extract, 1.5 g of bile salts, 5.0 g of sodium chloride, and 15.0 g of agar/L) plates and incubated at 44.5 ± 0.5 °C for 24 ± 2 h. For each fecal sample, 11 g of wet mass were added to 99 mL of a sterile 0.85% (*m*/*v*) NaCl dilution blank contained in a flask and manually shaken for 2 min. The resulting slurries were serially diluted and subjected to membrane filtration as described above. mFC-BCIG media allowed the selection of colonies that have β-galactosidase and β-glucuronidase activities. β-Glucuronidase activity, which is specific for *E. coli* among the thermotolerant coliform group, was assessed by the conversion of BCIG (5-bromo-4-chloro-3-indolyl-β-D-glucuronide) and the production of a blue colour. Blue colonies (putative *E. coli*) were picked and restreaked on BHI agar (EMD Chemicals, Gibbstown, New Jersey) for isolated colonies. Individual isolates were confirmed as *E. coli* on ChromCult agar (Merk, Darmstandt, Germany), which, in addition to confirmation of β-galactosidase and β-glucuronidase activity, contains tryptophan to improve the indole reaction, and frozen at –20 °C in Microbank bead (Pro-Lab Diagnostics, Richmond Hill, Ontario) cryovials containing preservatives as per the manufacturer's instructions. Five colonies from each water sample (≤5 if 5 were not available) were used for DNA fingerprinting.

REP-PCR DNA fingerprinting

Genomic DNA from individual pure cultures of *E. coli* isolates was extracted as described (Kon et al. 2007*a*). Cells were suspended in 200 μL of Tris–EDTA lysis buffer with proteinase K (0.5 mg/mL) and lysed for 1 h at 37 °C, followed by incubation for 10 min at 80 °C. Cell debris was pelleted by centrifugation for 10 min at 10 000*g*, and 1 μL of supernatant was used for PCR amplification with the BOX1AR primer, 5′-CTACGGCAAGGCGACGCTGACG-3′ (Dombek et al. 2000). Amplification was performed in a thermal cycler (Barloworld Scientific) using the following program: 35 cycles of 94 °C for 20 s, 60 °C for 20 s, and 65 °C for 5 min, with initial denaturation at 94 °C for 2 min and a final extension at 65 °C for 5 min (Edge and Hill 2007). PCR products were separated on 1% (*m*/*v*) agarose gel in Tris–acetate–EDTA buffer (40 mmol/L Tris, 20 mmol/L acetic acid, 1 mmol/L EDTA, pH 8.3) and visualized under UV transillumination after staining with ethidium bromide (Sambrook and Russell 2001). A 100 bp (100–3000 bp) DNA ladder (Fermentas, Burlington, Ontario) was used as the standard. Gel images were captured and stored electronically using GeneSnap software (Syngene, Cambridge, United Kingdom).

MST library

To build the MST library, we collected samples from manure storage tanks, septic tanks, and wildlife in the Eighteen Mile River watershed. The land use within the Eighteen Mile River watershed is predominantly agricultural with a focus on livestock farming (Statistics Canada 2001) and has very limited urban development. *Escherichia coli* from the samples for the library were isolated and frozen at –20 °C for DNA fingerprinting as described. *Escherichia coli* isolates were taken from frozen stock, grown on BHI agar (EMD Chemicals), and their REP-PCR DNA fingerprints were generated as described above. The fingerprints were grouped into library units based on their known animal source. Our MST library consists of the following library units: agriculture, wildlife, human, and environmental.

Computer-assisted data analysis

REP-PCR fingerprint analysis was performed with Bionumerics version 4.0 software (Applied Maths, Sint-Martens-Latem, Belgium). The positions of the PCR fragments on each gel were normalized with respect to the 100 bp DNA ladder as an external reference standard. The normalization allowed a comparison of multiple gels (Dombek et al. 2000). Identifications were carried out using *k*-nearest neighbour (*k*-NN) analysis with $k = 10$. In *k*-NN, source assignment is based on the unknown's proximity to *k* of the most similar fingerprints from the library of known sources. The unit of the identification library that has the largest number of entries belonging to *k*-NN is the best matching unit (Yao and Ruzzo 2006). *k*-NN is reported to be the best option for disproportional libraries such as our MST library (Robinson et al. 2007). If the fingerprint to be identified matched 2 library units equally, then it was assigned as unidentified. Utilizing the position tolerance function of Bionumerics, we determined the optimal position tolerance and performed our analysis with an optimization of 1.14% and a position tolerance of 1.90%. The performance of the MST library was assessed using the Jackknife analysis feature in Bionumerics in which library isolates were removed from the library one by one and treated as unknowns. Their correct or incorrect assignments were used to calculate the rate of correct assignment (RCA) (Wiggins et al. 2003). The library was decloned; clonal isolates (over 90% similarity) were removed from the library. This threshold value of 90% was established by comparison of the same DNA sample that was run on all gels used in this study. The similarities between the same DNA samples varied from 90.2% to 100% owing to gel-to-gel variability.

Results

The MST library

The size and representation of the library are important factors that determine the accuracy of its predictive ability. The MST library described in this study was constructed in proportion to the relative contribution of fecal material from each source within the Eighteen Mile River watershed, based on data from the Agricultural Census Report of Canada (Statistics Canada 2001) and calculated based on the Fleming and Ford (2001) report. The samples for the human source library unit were collected from septic tanks, since there is no sewage treatment plant (STP) or combined sewer outflow (SCO) within the Eighteen Mile River watershed. The samples included the septic tank of the public washroom of the study beach. The samples from manure storage tanks were used as a source for building the agricultural source library unit because they represent microbial population that might be released into the environment through different agricultural practices such as manure spreading. The environmental source library unit consisted of *E. coli* from interstitial water on the beach because they represent EAS (Kon et al. 2007*b*).

A total of 1432 isolates were used to construct the MST library (Table 1). The wildlife library contained 301 DNA fingerprints from *E. coli* isolated from seagull, goose, deer, duck, and raccoon droppings collected within the Eighteen Mile River watershed. One hundred and five colonies of *E. coli* from septic tanks from the watershed were isolated and their DNA fingerprints were used to build the human source library unit. Manure storage tank samples included 799 isolates from dairy, beef, horse, swine, sheep, and poultry farms in the Eighteen Mile River watershed. This representation is comparable with contribution from fecal material by different animal species in the Eighteen Mile River

Table 1. The composition of the microbial source tracking library.

Library unit	Animal source	No. of samples	No. of *Escherichia coli* isolates
Wildlife	Seagull	13 (6)	157
	Goose	8 (4)	76
	Duck	1	28
	Deer	1	18
	Raccoon	4 (2)	36
Agriculture	Cow	26	428
	Pig	20	242
	Horse	9 (2)	82
	Sheep	4	40
	Chicken	5	39
Human	Human	7	105
Environmental	Environmentally adapted strains (EAS)	31	250

Note: Samples were composites of 5–10 individual samples or represented manure storages. The numbers in brackets indicate additional, not composite, samples.

Table 2. Contribution from potential sources within the Eighteen Mile River watershed based on an agricultural census report from Statistics Canada (2001).

Fecal source	Total fecal material within the Eighteen Mile River watershed (%)*	No. of isolates
Cow	66	428
Pig	24	242
Poultry	7	39
Other (sheep, horse)	2	122
Human population	1	105

*Based on kg/day of fecal production, as calculated according to the Fleming and Ford (2001) report.

Table 3. Jackknife analysis: rates of correct classifications by the repetitive element (REP)-PCR library.

Known source of isolates	Predicted source category (%) Agriculture	Wildlife	Human	Environmental
Agriculture	**88.1**	6.7	1.6	3.6
Wildlife	29.6	**65.7**	2.6	2.2
Human	36.1	9.3	**50**	4.7
Environmental	30	5.6	0.6	**63.9**

Table 4. Identification of *Escherichia coli* isolates from Lake Huron at the Ashfield Township Park beach and from the Eighteen Mile River.

				Source category (%)				
Year	Location	Range of *E. coli* (CFU/100 mL of water)	Total No. of colonies	Agriculture	Wild	Human	EAS	Unidentified
2005	Five lake stations	2–1500	170	60.0	13.6	2.9	15.9	7.9
	Lake surveys	1–4800	388	62.4	8.7	3.3	17.8	8.5
	River*	11–4900	341	60.4	12.6	3.2	16.1	7.6
2006	Five lake stations	1–2800	155	60.0	12.9	2.6	15.5	9.0
	Lake surveys	1–210	132	59.3	4.7	2.1	17.5	16.4
	River*	22–6500	142	59.2	7.7	1.4	22.5	9.2

*River sampled at an intensive-monitoring site.

watershed (Table 2). The environmental source library unit was constructed with 227 DNA fingerprints from *E. coli* isolated from interstitial water on the beach.

The performance of the MST library was assessed by Jackknife analysis (Wiggins et al. 2003). The average rate of correct assignment (ARCA) was 66.9% (Table 3). The highest rate of misassignments (36.0%) was observed for the human isolates that were assigned as being of agricultural origin. The rates of misassignment to agricultural sources of samples originating from wildlife, human, and environmental sources were 29.6%, 36.1%, and 30%, respectively. Assignment of unknown samples to agricultural sources is likely biased high.

Identification of *E. coli* isolates from Lake Huron water samples

A total of 845 *E. coli* isolates from water samples collected at the shoreline and in the nearshore lake at the Ashfield Township Park beach of Lake Huron were subjected to REP-PCR DNA fingerprinting and their sources were identified using the MST library that we constructed. Out of 845 lake isolates, 558 were collected in 2005 and 287 in 2006. The results demonstrated that the dominant source of *E. coli* in lake water samples was agriculture, ranging from 59% to 62% (Table 4). The next prevalent source was EAS, ranging from 16% to 18%, followed by wildlife, which varied from 5% to 14%. The isolates assigned to the human source library unit were the least frequent among all fingerprints analyzed and ranged from 2% to 3%. An unidentified component was also present and it varied from 8% to 16%.

The results demonstrated very negligible differences between sampling locations (surveys of the study area of the lake versus the 5 nearshore lake monitoring stations) and between the 2 years (2005 and 2006).

Identification of *E. coli* isolates from the Eighteen Mile River samples

Escherichia coli isolates from water samples collected in the Eighteen Mile River intensive-monitoring station were subjected to the same REP-PCR analysis as the lake water isolates. Out of 483 *E. coli* isolates examined, 341 were collected in 2005 and 142 in 2006. The results revealed that the dominant source of *E. coli* in the river is agriculture (59% and 60% in 2005 and 2006, respectively), followed by the EAS (23% and 16% in 2005 and 2006, respectively), and then the wildlife (13% and 8% in 2005 and 2006, respectively) (Table 4). The ratio of different contributing sources in the river was similar to those observed for lake water isolates.

Discussion

Sources of *E. coli* contribution to the lake water

This study was undertaken to investigate the major contributing sources of *E. coli* pollution at the shoreline of southeastern Lake Huron over recreationally developed shoreline receiving discharge from small tributaries using the Ashfield Township Park beach as the study site. The shoreline is typical of the area and has features that are characteristic for southeastern Lake Huron such as sandy beaches backed by clay cliffs (Huron Fringe) followed by gentle slope plains in the direction of the lake (Huron Slopes) and abundance of small tributaries discharging into the lake (Howell et al. 2005). It is influenced by the Eighteen Mile River, which drains the 106 km^2 watershed and discharges into Lake Huron at the beach site.

The Huron Slope is parallel to the Huron Fringe. This unique geographic region is characterized by a narrow strip of sand and by the twin beaches of glacial Lake Warren that flank Wyoming Moraine. It is covered by a 1 m thick layer of clay above till deposits (Singer et al. 2003). The study beach is located in this area and has both small gravel bars and sand dunes backed by clay cliffs. The cliffs are forested for the most part and have random cottages and dwellings embedded into the narrow wooded area followed by heavily developed agricultural lands. Residences along the shoreline of the study site and within the watershed of Eighteen Mile River rely on septic systems for disposal of sanitary waste. A specialization in the region is livestock farming. There is also extensive pasture and crop farming over land that is extensively tile-drained (Howell et al. 2005). The dominant crops are soybean, wheat, and corn. Cattle and swine manure is abundant in the region and routinely applied to the fields as fertilizer.

The study site is located in southern Ontario, which is situated in a temperate climate zone with 4 seasons and precipitation spread evenly throughout the year as either snow or rain (Singer et al. 2003). The southern shores of Lake Huron are characterized by the heaviest snowfalls of the entire region of southern Ontario owing to local topography, wind, and proximity to Lake Huron. Prevailing winds are from southwest to northeast (onshore on the study beach).

The site consists of 2 parts: private and public beach (rural type). The beach does not have any standing water. The public beach has been periodically posted as unsafe for swimming, along with several beaches in southeastern Lake Huron, because of elevated *E. coli* numbers in the water. During long-term, beach-water-quality monitoring by the Huron County Health Unit (HCHU) the frequency of sample sets exceeding the Ontario Provincial Water Quality Objective (PWQO) of 100 CFU / 100 mL of water was variable from 1993 to 2003 (Howell et al. 2005). In 4 years (1994, 1998, 2000, and 2001), >50% of the sample sets exceeded 100 CFU / 100 mL; only in 2002 and 2003 did <30% of the sample sets exceed 100 CFU / 100 mL. The concentrations of *E. coli* varied greatly throughout the year in the samples used for this study; however, the overall ranges in the numbers were similar between the years and sampling locations, ranging from 1 to 4800 CFU /1 00 mL of water in the lake and from 11 to 6500 CFU / 100 mL of water in the river in samples used for source apportioning (Table 4).

There is a great necessity to understand the sources of microbial pollution at the shores of the Great Lakes to improve recreational water quality and perhaps stabilize it in the long term. There have been efforts by research groups to characterize *E. coli* sources at the beaches of Lakes Michigan, Superior, and Ontario. They employed different genotypic and phenotypic methods of MST such as REP-PCR, antibiotic resistance analysis, and genetic markers. To date there is no consensus in the scientific community as to which method is the best and the most applicable to a broad range of situations. These studies revealed different dominant sources in different areas of study. In an urban beach on Lake Ontario (Hamilton Harbor), the dominant source of fecal pollution was found to be bird droppings (Edge and Hill 2005). In another study, the dominant source in Lake St. Clair at the Clinton River watershed was human (Ram et al. 2004). However, it is never a single source that contributes to lake pollution but rather multiple sources in different ratios and combinations. The Great Lakes recreational shoreline is often in close proximity to heavily developed urban or agricultural lands. Depending on the type of land use, the sources contributing to lake pollution vary in different locations. In addition, recent studies in Lake Superior suggest that beach sand may be a temporal sink and a source of human- and waterfowl-derived *E. coli* (Ishii et al. 2007).

Utilizing our MST library of *E. coli* isolates from the known contributing sources, we determined that the major source of pollution was agriculture, followed by EAS, wildlife, and human sources. This result is not surprising, as the Eighteen Mile River watershed is predominantly agricultural by its land use and has a small human population. Wildlife is an inevitable source of fecal pollution anywhere and, therefore, is expected to be found in the lake.

Environmentally adapted strains of *E. coli* are also recognized to be widespread in the environment. A certain portion of those strains exists in many different secondary environments such as beach sand (Alm et al. 2006; Ishii et al. 2007; Kon et al. 2007*b*; McLellan et al. 2003), tropical soils (Byappanahalli and Fujioka 1998), and temperate soils from the Lake Superior watersheds (Ishii et al. 2006). We believe they are widespread in the environment, which is consistent with the published literature (Byappanahalli et al. 2003). This explains the significant contribution of EAS to both lake and river water, especially taking into consideration the possibility of their replication outside of animal hosts.

The unidentified component in our study varied from 7.6% to 16.4%, which is within the average range reported in other microbial source tracking studies. For example, Edge and Hill (2007) reported variations of between 8% and 27%. Vogel et al. (2007) reported a 6% variation. Ksoll et al. (2007) noted that the unidentified component could range from 2% up to 44% and stated that such observation is not uncommon in microbial source tracking studies (Ksoll et al. 2007). It can be explained by the fact that it is unlikely that the library can have matches for all lake or river isolates, therefore, the unidentified component will always exist in this type of a study. It can also be explained by the exchange of *E. coli* strains between different animals that occurs in the environment. The examples of such exchange are seagulls feeding on the fields where manure was spread (we have noticed flocks following the manure spreader), geese on the pastures, and raccoons on garbage dumps and cans. All these birds and animals can serve as the vectors delivering *E. coli* to the beach either directly or via tributaries. A source-tracking study by REP-PCR in an agriculturally dominated watershed was performed in the Finger Lakes region of western New York, and the results showed that wildlife was the major contributing source of fecal pollution (Somarelli et al. 2007).

Consistency of sources throughout the study area

Interestingly, when we examined the entire data set by year and sampling locations, we found no temporal or spatial differences between the subsets. The ratio of contributing sources did not change between the 2 years and from location to location, suggesting homogenous distribution of the *E. coli* community in the study area. Moreover, when we identified *E. coli* isolated from the samples collected in the river concurrently with samples collected in the lake, the results revealed that the river isolates matched the sources of the identification library just the same way as the lake isolates did. There were no differences in the ratio between contributing sources.

A possible explanation for the stability of this distribution is that the river is the major contributing source. Everything that is collected by the tributaries is mostly discharged into the lake, diluted, and distributed by the currents. That is possibly why there is such a high level of similarity between the composition of *E. coli* communities collected simultaneously in the river and in the lake. There are several small ephemeral creeks that are seasonally active and discharge into the lake within our study site. They can also contribute *E. coli*; however, the Eighteen Mile River is the largest tributary and, therefore, is the major contributor of discharge from the watershed. We focused our sampling on the intensive-monitoring site where water was present all the time during the study period, but with widely varying discharge to the lake. Our study beach is typical of areas of Lake Huron recreational shoreline adjacent to agricultural watersheds and affected by small tributaries discharging directly into the lake. It is possible that our findings can help us better understand this particular type of Great Lakes shoreline with respect to microbial pollution of the rural beaches.

Similar source-tracking studies were performed in a rural Virginia watershed dominated by livestock farming, and the results revealed that cattle was the dominant source of water pollution in the stream and neither seasonality nor sampling locations had an effect on the outcome as determined by library-based antibiotic resistance analysis (Graves et al. 2007).

In conclusion, we observed a striking consistency in the proportion of different *E. coli* populations spread uniformly throughout the study area and which were not affected by variations in weather conditions or total numbers of *E. coli* in the water. This suggests a stability of *E. coli* input from all contributing sources during the 2 years of study in the Eighteen Mile River watershed on southeastern shore of Lake Huron and raises interesting questions for microbial ecologists. The microbial source tracking methodology in general is a developing field and needs some refinement. We suggest that one refinement should be the inclusion of EAS of *E. coli* into the libraries for library-based approaches in future source-tracking efforts.

Acknowledgements

Our thanks are extended to H. House for providing agricultural samples and to P. Scharfe, A. Crowe. and T. Edge for the samples from septic tanks. We are indebted to the EMRB field crew for sample collection and the Microbiology Laboratory of the Ontario Ministry of the Environment for membrane filtrations and enumeration of *E. coli* in the samples. This work was financed by the Best in Science Program of the Ontario Ministry of the Environment. J.T.T. and H.L. acknowledge infrastructure and equipment support from the Canadian Foundation for Innovation and the Ontario Innovation Trust.

References

Alm, E.W., Burke, J., and Spain, A. 2003. Fecal indicator bacteria are abundant in wet sand at freshwater beaches. Water Res. **37**: 2978–2982.

Alm, E.W., Burke, J., and Hagan, E. 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. doi:10.3394/0380-1330(2006)32[401:PAPGOT]2.0.CO;2.

Anderson, K.L., Whitlock, J.E., and Hardwood, V.J. 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. **71**: 3041–3048. doi:10.1128/AEM.71.6.3041-3048.2005. PMID:15933000.

Beversdorf, L.J., Bornstein-Forst, S.M., and McLellan, S.L. 2007. The potential for beach sand to serve as a reservoir for *Escherichia coli* and the physical influences on cell die-off. J. Appl. Microbiol. **102**: 1372–1381. doi:10.1111/j.1365-2672.2006.03177.x. PMID:17448172.

Byappanahalli, M.N., and Fujioka, R.S. 1998. Evidence that tropical soil environment can support the growth of *Escherichia coli*. Water Sci. Technol. **38**: 171–174. doi:10.1016/S0273-1223(98)00820-8.

Byappanahalli, M., Fowler, M., Shively, D., and Whitman, R. 2003. Ubiquity and persistence of *E. coli* in a Midwestern coastal stream. Appl. Environ. Microbiol. **69**: 4549–4555. doi:10.1128/AEM.69.8.4549-4555.2003. PMID:12902241.

Dombek, P.E., Johnson, L.K., Zimmerley, S.T., and Sadowsky, M.J. 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl. Environ. Microbiol. **66**: 2572–2577. doi:10.1128/AEM.66.6.2572-2577.2000. PMID:10831440.

Edge, T.A., and Hill, S. 2005. Occurrence of antibiotic resistance in *Escherichia coli* from surface waters and fecal pollution sources near Hamilton, Ontario. Can. J. Microbiol. **51**: 501–505. doi:10.1139/w05-028. PMID:16121229.

Edge, T.A., and Hill, S. 2007. Multiple lines of evidence to identify the sources of fecal pollution at a freshwater beach in Hamilton Harbour, Lake Ontario. Water Res. **41**: 3585–3594. doi:10.1016/j.watres.2007.05.012. PMID:17575998.

Fleming, R., and Ford, M. 2001. Human versus animals-comparison of waste properties. Ridgetown College, University of Guelph.

Fogarty, L.R., Haack, S.K., Wolcott, M.J., and Whitman, R.L. 2003. Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull feces. J. Appl. Microbiol. **94**: 865–878. doi:10.1046/j.1365-2672.2003.01910.x. PMID:12694452.

George, I., Anzil, A., and Servais, P. 2004. Quantification of fecal coliform inputs to aquatic systems through soil leaching. Water Res. **38**: 611–618. doi:10.1016/j.watres.2003.10.022. PMID:14723930.

Graves, A.K., Hagedorn, C., Brooks, A., Hagedorn, R.L., and Martin, E. 2007. Microbial source tracking in a rural watershed dominated by cattle. Water Res. **41**: 3729–3739. doi:10.1016/j.watres.2007.04.020. PMID:17582454.

Howell, T., Abernathy, S., Charleton, M., Crowe, A., Edge, T., House, H., et al. 2005. Sources and mechanisms of delivery of *E. coli* (bacteria) pollution to the Lake Huron shoreline of Huron County. A report from the Lake Huron Science Committee to the Ontario Ministry of the Environment.

Ishii, S., Ksoll, W.B., Hicks, R.E., and Sadowsky, M.J. 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. **72**: 612–621. doi:10.1128/AEM.72.1.612-621.2006. PMID:16391098.

Ishii, S., Hansen, D.L., Hicks, R.E., and Sadowsky, M.J. 2007. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. Environ. Sci. Technol. **41**: 2203–2209. doi:10.1021/es0623156. PMID:17438764.

Kon, T., Weir, S.C., Howell, E.T., Lee, H., and Trevors, J.T. 2007*a*. Genetic relatedness of *Escherichia coli* isolates in interstitial water from a Lake Huron (Canada) beach. Appl. Environ. Microbiol. **73**: 1961–1967. doi:10.1128/AEM.02437-06. PMID: 17261522.

Kon, T., Weir, S.C., Trevors, J.T., Lee, H., Champagne, J., Meunier, L., et al. 2007*b*. Microarray analysis of *Escherichia coli* strains from interstitial beach waters of Lake Huron (Canada). Appl. Environ. Microbiol. **73**: 7757–7758. doi:10.1128/AEM. 01333-07. PMID:17890330.

Ksoll, W.B., Ishii, S., Sadowsky, M.J., and Hicks, R.E. 2007. Presence and sources of fecal coliform bacteria in epilithic periphyton communities of Lake Superior. Appl. Environ. Microbiol. **73**: 3771–3778. doi:10.1128/AEM.02654-06. PMID:17468280.

McLellan, S.L., and Salmore, A.K. 2003. Evidence for localized bacterial loading as the cause of chronic beach closings in a freshwater marina. Water Res. **37**: 2700–2708. doi:10.1016/ S0043-1354(03)00068-X. PMID:12753847.

McLellan, S.L., Daniels, L.D., and Salmore, A.K. 2003. Genetic characterization of *Escherichia coli* populations from host sources of fecal pollution by using DNA fingerprinting. Appl. Environ. Microbiol. **69**: 2587–2594. doi:10.1128/AEM.69.5.2587-2594.2003. PMID:12732525.

Olive, D.M., and Bean, P. 1999. Principles and applications of methods for DNA-based typing of microbial organisms. J. Clin. Microbiol. **37**: 1661–1669. PMID:10325304.

Ontario Ministry of the Environment. 2004*a*. A membrane filtration method for the detection and enumeration of total coliform, *Escherichia coli*, *Pseudomonas aeruginosa*, and fecal streptococci in environmental samples. Laboratory Services Branch, Quality Management Unit, Etobicoke, Ontario. Report No. MFMICRO-E3371.

Ontario Ministry of the Environment. 2004*b*. Isolation, detection, and enumeration of *Escherichia coli* in biosolids. Laboratory Services Branch, Quality Management Unit, Etobicoke, Ontario. Report No. MICROBIO-E3433.

Ram, J.L., Ritchie, R.P., Fang, J., Gonzales, F.S., and Selegean, J.P. 2004. Sequence-based source tracking of *Escherichia coli* based on genetic diversity of β-glucuronidase. J. Environ. Qual. **33**: 1024–1032. PMID:15224940.

Robinson, B.J., Ritter, K.J., and Ellender, R.D. 2007. A statistical appraisal of disproportional versus proportional microbial source tracking libraries. J. Water Health, **5**: 503–509. doi:10.2166/wh. 2007.044.

Saini, R., Halverson, L.J., and Lorimor, J.C. 2003. Rainfall timing and frequency influence on leaching of *Escherichia coli* RS2G through soil following manure application. J. Environ. Qual. **32**: 1865–1872. PMID:14535331.

Sambrook, J., and Russell, D. 2001. Molecular cloning. A laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Seurinck, S., Verstraete, W., and Siciliano, D. 2005. Microbial source tracking for identification of fecal pollution. Reviews in Environ. Sci. Technol. **4**: 19–37.

Simpson, J.M., Santo Domingo, J.W., and Reasoner, D.J. 2002. Microbial source tracking: state of the science. Environ. Sci. Technol. **36**: 5279–5288. doi:10.1021/es026000b. PMID:12521151.

Singer, S.N., Cheng, C.K., and Scafe, M.G. 2003. The hydrogeology of Southern Ontario. 2nd ed. Ontario Ministry of the Environment, Toronto.

Somarelli, J.A., Makarewicz, J.C., Sia, R., and Simon, R. 2007. Wildlife identified as major source of *Escherichia coli* in agriculturally dominated watersheds by BOX A1R-derived genetic fingerprints. J. Environ. Manage. **82**: 60–65. doi:10.1016/j. jenvman.2005.12.013. PMID:16551490.

Statistics Canada. 2001. Agricultural Census Report of Canada.

United States Environmental Protection Agency (US EPA). 2005. Microbial source tracking guide. United States Environmental Protection Agency, Washington, D.C. Document No. EPA/600/ R-05/064.

Vogel, J.R., Stoeckel, D.M., Lamendella, R., Zelt, R.B., Santo Domingo, J.W., Walker, S.R., and Oerther, D.B. 2007. Identifying fecal sources in a selected catchment reach using multiple source-tracking tools. J. Environ. Qual. **36**: 718–729. doi:10. 2134/jeq2006.0246. PMID:17412907.

Whitman, R.L., and Nevers, M.B. 2003. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. Appl. Environ. Microbiol. **69**: 5555–5562. doi:10.1128/ AEM.69.9.5555-5562.2003. PMID:12957945.

Wiggins, B.A., Cash, P.W., Creamer, W.S., Dart, S.E., Garcia, P.P., Gerecke, T.M., et al. 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69**: 3399–3405. doi:10.1128/AEM.69.6. 3399-3405.2003. PMID:12788742.

Yao, Z., and Ruzzo, W.L. 2006. A regression-based *K* nearest neighbor algorithm for gene function prediction from heterogeneous data. BMC Bioinformatics, **7**(Suppl. 1): S11. doi:10.1186/ 1471-2105-7-S1-S11. PMID:16723004.