IDENTIFICATION OF FECAL POLLUTION SOURCES OF KARST WATERS BY A COMBINATION OF *BACTEROIDALES* GENETIC MARKER, BACTERIAL INDICATORS, AND ENVIRONMENTAL VARIABLES

BY

YA ZHANG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Environmental Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Adviser:

Professor Wen-Tso Liu

ABSTRACT

Fecal contamination in Midwestern karst regions was evaluated by a combination of *Bacteroidales*-based microbial source tracking (MST), traditional bacterial indicators, and environmental variables. Water samples from springs and wells were collected from Illinois, Wisconsin, Kentucky and Missouri quarterly. Quantitative PCR with seven primer sets targeting different members of *Bacteroidales* was used to discriminate human and livestock fecal pollution. The results suggested that contamination by both human and animal waste was the dominant type of pollution among all samples, and only a few samples showed pollution solely by human or animal. Furthermore, spring water tended to be more contaminated than well water, and urban spring systems were detected with a higher number of fecal biomarkers than rural spring systems. However, correlation between traditional bacterial indicators and fecal contamination determined by *Bacteroidales*-based MST was weak. Among all the environmental parameters examined, *Escherichia coli*, Cl^-, and ORP were significantly correlated ($p < 0.05$) with level of *Bacteroidales* fecal biomarkers. Well samples were affected by ORP and Cl^- the most, and samples from springs were influenced by all three parameters. The results illustrated that a large portion of karst groundwater systems in Midwestern regions was co-contaminated with human and livestock feces. The inclusion of traditional bacterial indicators and a variety of environmental variables in addition to *Bacteroidales*-based MST provided an effective toolbox for identifying fecal contamination in karst regions.

To Father and Mother

ACKNOWLEDGMENTS

This work was funded by the Midwest Technology Assistance Center (MTAC), which supports Environmental Protection Agency Strategic Plan Goal 2 of Clean and Safe Water, Objective 2.1: Protect Human and Sub-Health Objective 2.1.1: Water Safe to Drink. The authors thank Walton R. Kelly from Illinois State Water Survey, and Samuel V. Panno from Illinois State Geological Survey for field work and all the assistance and guidance throughout the study. We appreciate the help from Wisconsin Geological and Natural History Survey, Kentucky Geological Survey, and Perry County Soil and Water Conservation District. We also thank local farmers' cooperation with field study. Reference DNAs were obtained from Peiying Hong, a former student of our group. We thank two undergraduate students, Jinwei Hu and Bonnie Coats for their help with experiments. Finally, special thanks to my adviser, Professor Wen-Tso Liu, for his constant guidance throughout this research.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1
1.1 Traditional microbial indicators of fecal pollution 1
1.2 Microbial source tracking (MST) 3
1.2.1 Concept of MST 3
1.2.2 Chemical-based MST 3
1.2.3 Microbe-based MST 4
1.3 Karst region 11
1.4 Studies on tracking contaminations in karst region 13
1.5 Objective of the study 14
CHAPTER 2: MATERIALS AND METHODS 16
2.1 Study sites 16
2.2 Sample collection and analysis 17
2.3 Genomic DNA extraction 19
2.4 PCR and gel electrophoresis analysis 20
2.5 Q-PCR 21
2.6 Statistical analysis 22
CHAPTER 3: RESULTS 24
3.1 Composition of water samples 24
3.2 Environmental parameters 24
3.3 PCR screening 29
3.4 Standard curves for Q-PCR 29
3.5 Using Q-PCR technique to identify sources of fecal contamination 31
3.6 Contamination identified with concentrations of chloride and nitrogen 36
3.7 Correlation between environmental variables and fecal contamination 38
CHAPTER 4: DISCUSSION 40
4.1 Sensitivity and specificity of Q-PCR-based MST 40
4.2 Multi-metric approaches: relationship among traditional indicators (viruses), environmental variables and *Bacteroidales*-MST 43
4.3 Implication for management practice in karst region 45
CHAPTER 5: CONCLUSION 46
REFERENCES 48

CHAPTER 1
INTRODUCTION

1.1 Traditional microbial indicators of fecal pollution

Indicators are selected to signify the presence of fecal pollution in water bodies and forecast the associated risk for human health. These indicators are abundant in human and animal feces with distinct characteristics that can be easily differentiated from other microorganisms. Due to technical limitations, those characteristics are typically based on cultivation methods. It is suggested that an ideal indicator should have the following characteristics: i) specific to fecal origin; ii) no replication or regrowth outside of the host; iii) similar survival characteristics with pathogens in natural environments; iv) strongly associated with presence of pathogens; and v) simple to be tested and easy to recover from environments (68).

Several indicators such as total coliform, *Escherichia coli*, enterococci, *Clostridium perfringens* and their spores, and coliphages are commonly used in monitoring ambient water. Enumeration of total coliform, *E. coli*, and/or enterococci in freshwater and marine water is recommended to indicate the fecal pollution. Although coliforms were thought to be primarily of fecal origin, they are widely distributed in the environment, making the representativeness of their presence problematic (39). In comparison, *E. coli* is considered to be a better indicator of fecal contamination because it is more restricted to feces of human and warm-blooded animals (15). Enterococci are often used as bacterial indicators for recreational surface waters due to their salt-resistant characteristics (3). However, they are approximately 100- to 1000-fold less numerous in aquatic environments than *E. coli*, making them a less effective indicator when a low level of fecal contamination occurs (15). *C. perfringens* is an anaerobic bacterium and will turn

into spores under unfavorable conditions in water environment. Its long survival time under the spore state makes it suitable to indicate groundwater pollution (15). However, detection of *C. perfringens* and its spores requires the experiments to be performed under strict anaerobic conditions, which can be complicated and expensive. Coliphages are viruses that infect members of the coliform group, and thus are not specific only to *E. coli*. Coliphages are a useful indicator for the presence of enteric viruses in water because of their similar characteristics with human viruses (81). However, coliphages can only be found in less than 50% of human feces (15) and detecting coliphages by cultivation is time-consuming and requires good technical skills. As the results, the use of coliphages as an indicator of fecal contamination has not been widely used in field-testing.

Until now, total coliform, *E. coli*, and enterococci are still the most widely used microbial indicators for presence of fecal contamination and their detection methods are much easier and less expensive. However, it is hard to differentiate the sources of fecal pollution with these traditional indicators. First, coliforms and *E. coli* are present not only in human feces but also in animal sources. Second, these indicators can survive out of hosts for a long time and even replicate in the environment. The correlation between the presence of indicators and human health risks is also not strong in most cases (19, 24, 25, 47, 68, 69, 74, 80). Knowing only the level of contamination without fecal origins provides little information for management of land use and implementation of remediation. Therefore, effective microbial source tracking (MST) methods are needed for discriminating sources of fecal pollution.

1.2 Microbial source tracking (MST)

1.2.1 Concept of MST

MST aims to trace the origin of fecal pollution by using chemical, microbiological, genotypic, and phenotypic methods (68). It works on the assumptions that there are host-specific characteristics showing spatial and temporal stability. MST methods can be divided according to target types into two large categories: chemical-based methods and microbe-based methods. Recent studies use a combination of methods from these two categories rather than adopting them separately (21, 34, 47, 53, 58, 60, 65, 79). In addition to identify and/or quantify sources of contamination in the area of interest, MST can support the development and implementation of total maximum daily loads (TMDLs) and best management practice (BMP), thus lowering health risks associated with fecal pollution (64, 68, 72).

1.2.2 Chemical-based MST

Chemical-based fecal MST targets chemicals that are specific to human wastewater to indicate pollution of anthropogenic activities. A variety of human-specific chemicals have been proposed from compounds identified in wastewater effluent. A few chemicals showed promising application, including caffeine, stanols/sterols, which were metabolic by-products of cholesterols (coprostanol), pharmaceuticals (carbamazepine and diphenhydramine) and fluorescent whitening agents (22). For examples, Furtula et al. (21) combined chemical (sterols) and microbial MST methods as well as nitrogen to explore the potential contamination sources of surface water in Nathan Creek watershed, British Columbia, Canada. Sterol-based MST was shown to be effective in discriminating

pollution from human and nonhuman. Another study (53) utilized caffeine, nutrients and indicator bacteria to track human and non-human sources of pollution in rural freshwater and urban marine system. In rural area, concentration of caffeine was sometimes below detection limit. Caffeine and nitrogen did not correlate well with traditional fecal indicators. In the urban system, caffeine was linked to traditional fecal indicator and human activity. Chemical-based MST is fast to perform and avoids changes in consideration due to the regrowth and decay of microbes in the environment. However, analyzing those chemical compounds generally need specialized and expensive equipment. It is also hard to detect these compounds in natural water systems when their concentrations are usually too low to be detected in original pollution sources after dilution. The application of these methods in field studies also needs information of their persistence under ambient environment conditions and relationship with traditional microbial indicators. Lastly, the method might not be able to distinguish among different animal sources (27). With all the limitations, chemical-based MST might not be qualified as the primary MST method.

1.2.3 Microbe-based MST

MST based on microbes can be library-independent and library-dependent. Library-independent techniques determine sources of pollution based on known host-specific characteristics of microorganisms without referring to a library when water samples are tested. These techniques need the establishment of a library of bacterial isolates to determine fecal pollution sources. Generally, certain characteristics of targeting bacteria isolated from various feces (e.g. human and animal) are recorded to

generate a library. A database of known sources is extracted from the library using statistical analysis (discriminant analysis or clustering analysis). Then bacterial isolates from unknown sources are compared with the database of known sources to determine the possible sources of pollution (26).

Library-dependent methods can be further divided into two approaches based on phenotypes and genotypes. The most commonly used phenotypic method is antibiotic resistance analysis (ARA). ARA relies on the detection of antibiotic resistance patterns of fecal streptococci, enterococci, or *E. coli* (26, 60, 72, 84). Many field studies demonstrated that ARA is a simple and reliable approach to discriminate sources of fecal contamination. Hagedorn *et al.* (26) established a large database of ARA for fecal streptococci with more than 80% accuracy in source identification. The database was applied to a rural watershed in Virginia and identified cattle as the dominant source of pollution. Based on these results, management practice was done to limit cattle access to the watershed, and improved water quality was observed. Another study (60) examined the resistance patterns with *E. coli* and fecal streptococci to small watersheds in Kentucky in order to discriminate human from non-human sources of pollution. Using their database, higher than 60% of correct classification was achieved. The study concluded that ARA detection in combination with information on the watershed and its land use was helpful to the remediation process of impaired water environments.

Although ARA is easy to be utilized from technical aspects and can discriminate multiple fecal sources, its accuracy (e.g., correct classification) is still low (generally more than 50%) and dependent on the size and representative of reference database (23, 26, 28, 60, 84). Furthermore, it is difficult to apply this method to complex systems with

multiple sources of pollution. Another concern of the method is that horizontal gene transfer via plasmids could carry antibiotic resistance genes from target organisms to other non-fecal contamination related organisms. This will change the antibiotic resistance profile of known isolates in the database (72).

Genotypic-based library-dependent methods, including ribotyping, pulse field gel electrophoresis (PFGE), and repetitive element PCR, adopt similar principles as DNA fingerprinting. Bacteria from known sources need to be isolated, cultured, and identified first. Ribotyping approach identifies pieces of genomic DNA cut with restriction enzymes by oligonucleotide probes targeting rRNA genes. The process is highly reproducible but labor-intensive. PFGE utilizes gel electrophoretic analysis to diagnose genomic DNA fingerprints after treatment with rare-cutting restriction enzymes. Correlation between PFGE profile and isolate source is weak due to the high variation in genomic DNA among closely related bacterial isolates at strain level. Repetitive element PCR uses PCR targeting interspersed repetitive DNA sequences in prokaryotic genomes followed by gel electrophoresis (68, 72). This method is easier to perform comparing with ribotyping and PFGE. Application of repetitive element PCR for MST has seldom been performed. In comparison with library-dependent phenotypic methods, library-independent genotypic methods are more expensive and sometimes more labor-intensive and time-consuming. Their usage in field study of MST is still limited compared with ARA. Due to the disadvantages of library-dependent methods and the advance of molecular techniques, library-independent methods based on molecular markers start to gain the popularity. These methods include host-specific PCR/quantitative PCR (Q-PCR), gene-specific PCR,

phage typing, and microbial community analysis [denaturing gradient gel electrophoresis (DGGE) and terminal restriction fragment length polymorphism (T-RFLP)] (72).

Among all the possible host-specific bacteria, most analyses chose *Bacteroidales* as the target. *Bacteroidales* is an order in the phylum *Bacteroidetes*. *Bacteroidetes* species are obligate anaerobes, which account for about half of all bacterial populations in human intestine and are at much larger percentage than total coliform and *E. coli* (42). The dominant members of the genus *Bacteroides* include *B. vulgatus* (31% of all *Bacteroidetes* and 1.4-15% of all bacteria), *B. thetaiotaomicron* (12% of all *Bacteroidetes* and 0.2-6.2% of all bacteria), *B. fragilis* (5% of all *Bacteroidetes* and 2.5-13.3% of all bacteria), *B. distasonis* (0.8% of all *Bacteroidetes* and 0.4% of all bacteria), and *B. caccae* (0.2-1.1% of all bacteria) (14, 30, 42, 76). Because of their prevalence in feces, *Bacteroides* were considered as a good candidate target for MST. Early studies focused on assessing the use of *B. fragilis* as an indicator for human fecal contamination. Cultivation methods were used to examine their prevalence in various animal feces and ambient environments and their degradation rate in aerobic environment (1, 2, 19). These studies demonstrated that *Bacteroides* are host specific and have limited survival in the environment. With its high abundance, it can provide the ability to discriminate human fecal contamination from other different fecal contamination.

Methods such as Q-PCR and T-RFLP targeting the 16S rRNA genes of *Bacteroides* were developed and applied to circumvent the difficulty to cultivate anaerobic *Bacterioides* species. T-RFLP is a community analysis technique that can identify all possible major microbial populations in one electropherogram (see reference 43 for detailed explanation on the fundamentals of the technique). Using primer sets

targeting the order *Bacteroidales*, specific terminal restriction fragments (tRFs) representing different fecal origins in an electropherogram are selected. Fecal sources of unknown samples can be easily identified with those selected tRFs. This method is a semi-quantitative technique that can provide relative abundance information for different contamination sources (6, 17, 43, 66). T-RFLP technique is reported to be high throughput and reproducible. Multiplex reactions are not needed since representative T-RF peaks associated with different hosts could be determined in one reaction. And the analysis can be finished within a few hours. Several representative fragments are reported, including 117-119 bp human-specific marker and 222-227 bp cow-specific marker (6, 66). However, T-RFLP is likely to have many false-positive results, possibly because each tRF can represent more than one target. At the same time, this feature might be beneficial in the cases that false negatives are intolerant (66). In addition, analysis of T-RFLP needs sequencing machines, which can limit its application to labs without access to those expensive machines.

Q-PCR can quantify the amplification of PCR amplicons in real time. Assuming a constant amplification efficiency (E), equation (1) describing a normal PCR reaction:

$$P_C = P_0 \times (E+1)^n \qquad (1)$$

where n is the number of thermocycles, P_n is the copy number of product and P_0 is the starting number of template. To predict P_0, we need to know E (equation 2):

$$P_0 = {}^{P_n}/_{(E+1)^n} \qquad (2)$$

Equation (2) can be converted into equation (3) using logarithm transformation.

$$Log\ (P_0) = -n \times Log\ (E+1) + Log\ (P_n) \qquad (3)$$

If $Log\ (P_n)$ is constant, $Log\ (P_0)$ and n have a linear relationship with $Log\ (E+1)$ as the slope and $Log\ (P_n)$ as the intercept. If multiple couples of n and $Log\ (P_0)$ are obtained, then the slope $Log\ (E+1)$ can be calculated using linear regression. To satisfy all the requirements, a series of DNAs with known concentrations are amplified in individual tubes and compared at a point, at which same amounts of DNA products was obtained. A series of n corresponding to the point are recorded and $Log\ (E+1)$ and $Log\ (P_n)$ can be calculated. For unknown DNA samples, record n corresponding to the same point chosen with DNA templates of known quantity. With all the parameters known on the right side of equation (3), $Log\ (P_0)$ is easy to be calculated (29, 56, 62).

A breakthrough for Q-PCR is the ability to quantify the accumulation of amplicons using fluorophores such as SYBR. The point where each reaction contains identical amounts of DNA products is determined arbitrarily at which the fluorescent signal intensity increases above the background threshold. That fluorescent intensity is named as the threshold and the corresponding n is called threshold cycle (Ct). The regression line between Ct and $Log\ (P_0)$ is the standard curve. To minimize non-target amplification, a melting curve can be carried out after the normal amplification, which would reduce the occurrence of false-positives (29, 31, 48, 56, 62)

Q-PCR specifically targeting *Bacteroidales* is widely used for tracking fecal contamination from most human, cow, and pig feces in different environments (38, 47, 48, 65, 70). Other possible contamination sources include dog, cat, elk, deer, and gull (13, 18)(12). The targeted DNA biomarkers of *Bacteroidales* include 16S rRNA genes, functional genes, and genes encoding hypothetical proteins. However, it can be difficult to design specific primers to target *Bacteroidetes* species from different sources due to

the close homology of the DNA biomarkers among *Bacteroides* species. Dick *et al.* (12) analyzed 16S rRNA gene sequences from feces of eight different hosts including human, bovine, pig, horse, dog, cat, gull, and elk. They observed that tight and high-similarity clusters were formed only for *Bacteroides* species detected in cow, pig, and horse feces. 16S rRNA gene sequences of *Bacteroides* species from human, gull, dog, and cat fecal samples were clustered together, making it difficult to design specific primers for Q-PCR to distinguish *Bacteroides* species from different hosts. A study tested 10 different Q-PCR assays for human-associated *Bacteroidales*, only three out of 10 primer sets satisfied all the six criteria used. This suggests that it is important to validate the primer sets before applying them in MST (71).

Q-PCR detection is not only quantitative but also sensitive and specific. It was theoretically calculated that Q-PCR could detect as low as three copies of targeted DNA per reaction (8), in comparison to the range of detection adopted in most MST studies, which is approximately 25 - 1×10^6 copies of targeted gene per reaction (38, 48, 69, 71). Generally, Q-PCR for *Bacteroidales*-based MST methods can correctly identify fecal sources with an accuracy of more than 80% when samples from known sources are tested, which is much higher than the specificity of ARA and T-RFLP (38, 59, 71). Considering only the Q-PCR step using a 96-well reaction plate, it can be finished in 3 hours. Comparing with T-RFLP, there is no need to use a DNA sequencing machines. Because of the advantages and successes in field studies, Q-PCR has become a more reliable method than methods like T-RFLP in MST studies.

1.3 Karst region

MST has been applied to study fecal contamination in various environments, such as freshwater in general (21, 38, 48, 60, 66, 79), rural freshwater (26, 47, 53), urban marine water (53), urban stormwater systems (65), and groundwater from karst regions (33, 34, 59, 82) . Among all the different environments, groundwater from karst regions is the most vulnerable one to fecal contamination.

Karst topography is referred to terrain shaped by dissolution of carbonate rocks and characterized by crevices, sometimes large conduits, sinkholes, caves, and springs and sinking streams (Figure 1) (11, 35, 55). Groundwater from karst regions is a main source of water for local people, and supports roughly a quarter of the world's population (11). This valuable water resource, however, is very vulnerable to microbial contamination coming from surface water through sinkholes, fissures, and conduits and bypassing the natural filtration through soil filtration or sorption. Transport of microbial contaminants in karst is further facilitated by high groundwater flow velocities, especially in the case of heavy rainfalls or melt-water happening in agricultural areas, in which high loadings of nutrient elements, pesticides, fertilizers, as well as fecal materials from human and animal are flushed into groundwater system (9, 35, 37, 50, 52, 63). Human wastes enter karst aquifers via on-site disposal systems or community treated wastewater effluents. Contamination from animals can occur either from discharge from waste lagoons and pits at livestock facilities or from the leaching of manure (7, 32, 44, 57). Human and animal wastes carry not only high concentrations of nutrient elements, but also elevated level of microorganisms and possible pathogens. A close link between the occurrences of

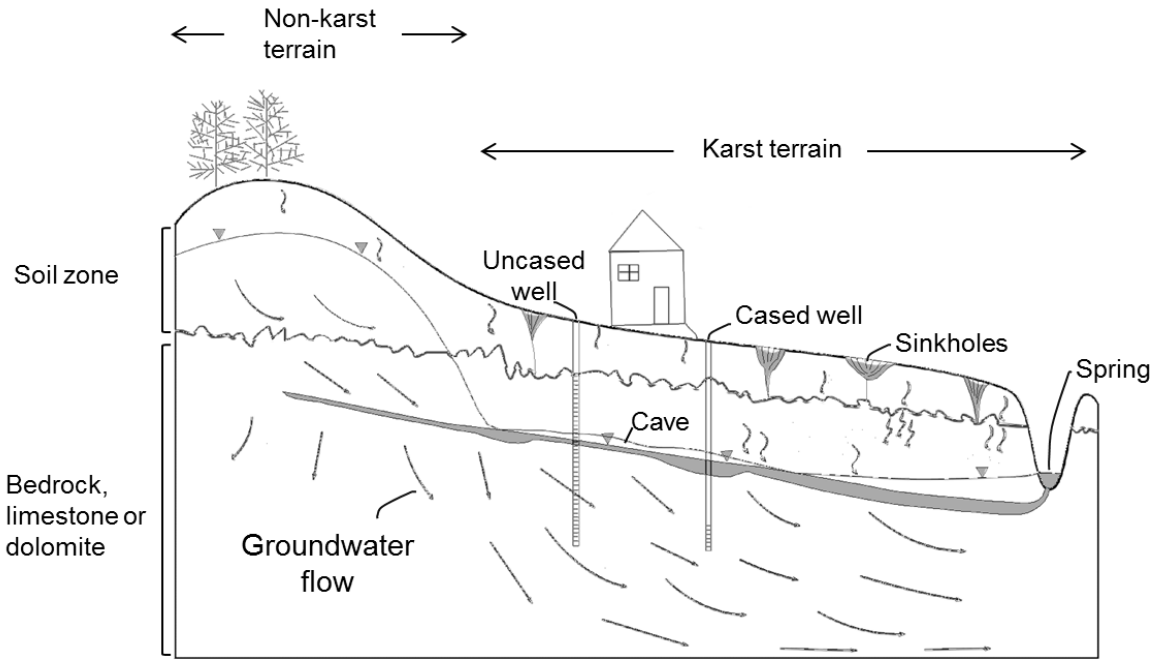

Figure 1. Karst topography (35).

impaired water quality and the presence of karst topography of small distribution systems in France has been reported (5). In addition, a few outbreaks have been traced back with drinking water originated from karst systems (4, 10, 75, 78). Therefore, there is increasing need to track the origin of microbial contamination, so that BMP can be applied to guide agricultural activities in karst regions.

Springs and wells are the two main water bodies in karst regions. There are generic differences of the possible contamination sources between wells and springs. Compared with wells, springs tend to receive recharge from a larger area, which increases the likelihood of occurrence of fecal pollution. Concentrations of total coliform were found generally higher in springs than in wells (35). However, water movement in springs is much faster than that of wells. The contamination level is determined by the relative effect of loading rate and water movement. If a source of pollution has entered into

springs, the contamination level can be lowered much faster if the effect of water movement is stronger than rate of loading.

1.4 Studies on tracking contaminations in karst region

Direct monitoring inorganic and organic compounds, bacterial indicators and viruses associated with contamination of karst regions, can give a clue on the state of water system, and is often the first step to recognize the presence of non-point source pollution. In most cases, elevated-level of chloride, nitrate and traditional fecal indicators are first observed in the karst area studied. But the correlation between the presence of enteric bacteria and elevated-level of chloride and/or nitrate was not significant (34, 35, 45). In addition to routine monitoring of traditional indicator bacteria, enteric viruses were also monitored in some studies (24, 33, 34, 40, 41, 46, 47). Based on the monitoring results and local land use information, the possible source of pollution, such as septic discharge, livestock waste, or agricultural runoff can be inferred.

The next step is to carry out MST studies to identify the possible pollution sources. Due to the relative high concentrations of pesticides and pharmaceuticals in vulnerable karst regions, chemical-based MST is suitable in addition to the widely-used microbe-based MST. Detections of previously-banned and currently-used pesticides and their degradation products, such as DEET (N, N-diethyl-meta-toluamide) were used (20, 34). In some situations, the concentrations of these chemicals could exceed the standard (20, 34). Another study (82) investigated concentrations of natural estrogen, a type of steroid hormones in Missouri Ozark Plateau Aquifer. The concentrations of estrogen changed seasonally and were higher than the level leading to adverse health effect in fish at more

than 60% of the time. The sources of estrogen were identified to be derived from sewage-treated wastewater and agricultural runoff (82).

Although many studies have been conducted to trace sources of pollution from karst region in European countries such as Austria (59) and Switzerland (55), and in US such as Florida (34), Tennessee (33), Arkansas (54), Kentucky (57), and Missouri (20, 82), direct quantification of human and animal fecal pollution in karst region is lacking except one study (59). The comparison of different studies from various regions is difficult because of different methods used by these studies, thus a systematic investigation of karst regions covering a large area is needed to reflect the spatial differences.

1.5 Objective of the study

The objective of this study was to distinguish between human and livestock sources of fecal contamination in karst aquifers of four Midwestern states (Wisconsin, Illinois, Kentucky, and Missouri) by using a combination of *Bacteroidales*-based MST, traditional bacterial indicators, and environmental parameters. This thesis examined 73 water samples collected over a two-year period from more than 40 wells and springs. A Q-PCR approach including seven primer sets, one for overall contamination, four for human, and one each for swine and bovine was used. In addition, sources of contamination identified by *Bacteroidales*-based MST were compared with results derived from nitrate and chloride concentrations. Traditional bacterial indicators were monitored to explore their relationship with fecal contamination determined by *Bacteroidales*-based MST. Another task was to determine the potential environmental

variables that associate with fecal contamination in springs and wells. Standard water quality parameters including water temperature, pH, specific conductivity (SpC), oxidation-reduction potential (ORP) as well as key inorganic chemicals and metal were recorded for each sample. Ordination analysis was used to extrapolate the correlation. This was the first study to examine and compare fecal contamination in Midwestern karst aquifers by a multi-metric approach. Results of the study could provide useful information for BMP implementation and better land use management. Also, identified environmental parameters associated with fecal contaminations are important for monitoring practices in the future.

CHAPTER 2
MATERIALS AND METHODS

2.1 Study sites

Study sites are located in the karst regions of four Midwestern states, including Illinois, Wisconsin, Missouri, and Kentucky. Well-known karst areas within these regions include the Bluegrass Region and the Western Pennyroyal area of Kentucky, and the Illinois Sinkhole Plain. Samples taken from Illinois and Wisconsin were collected from wells, whereas samples from Missouri and Kentucky were collected from springs.

Sampling sites in Illinois were located primarily within the counties of St. Clair, Monroe, and Randolph, an area known as the Illinois Sinkhole Plain. This area is known for having more than 10,000 large (typically 30 to over 100 m in diameter) sinkholes, large karst springs, and numerous branch work caves with actively flowing streams. About 65% of the land use in this area is row-crop agriculture, with some livestock including hogs and cattle. Hogs are generally confined indoors while cattle are allowed to graze in pastures most of the year. Hog manure is used as fertilizers for crop lands, but represents only a small fraction of the fertilizer applied in this region. Most of the crop lands are fertilized with anhydrous ammonia during late fall or early spring. These fertilizers and on-site wastewater treatment systems all contribute to the local groundwater contamination as suggested from previous studies (35).

Much of Wisconsin's karst area locates in the south-central and eastern parts of the state, which is masked by thick layers of glacial deposits. The bedrock is the Silurian dolomite or the Sinnipee and Prairie du Chien dolomite. Samples from wells were taken from WI. Six sites were sampled quarterly to investigate seasonal effect. These sampling

sites located in counties with noticeable karst features, including Fond du Lac, Calumet, Kewaunee, and Door.

Samples taken from Missouri were located in Perry County within the Perryville Karst. This is also a sinkhole plain and contains thousands of sinkholes and large caves with actively flowing streams (large-river caves). The size of sinkholes in this area ranges from several meters to several kilometers in diameters and can reach up to 30 meters in depth. The unique large-river caves are relatively shallow and are distributed primarily in the central part of the county. The climate within these caves is cool (55-60 °F) and moist. The sampling sites in Missouri included Mystery Entrance, Apple Creek Shrine, Scholl Creek, Dry Fork, Thunder Hole, Mystery Resurgence, Huber Brance, Blue Spring, Keyhole Spring, Briar Spring, and Ball Mill Spring. The streams investigated in this work support a variety of species, including some rare species listed in Missouri Species of Concern (20) .

Kentucky is famous for its karst area in the world with approximately 55% of area underlaid by limestone or dolomite. Lots of caves longer than 30 miles can be found in KY. Both urban and rural karst regions were sampled. Urban samples were taken from Louisville city, whereas rural samples were taken from LaRue, Grayson, and Hardin coutnties, which located in the Western Pennyroyal karst area. The spring samples taken from Kentucky were not for drinking water supply.

2.2 Sample collection and analysis

Water samples were collected on a quarterly basis from wells and springs in the karst regions of the four states. A total of one to five liters per sample was collected

during each sampling point. Wells were purged for at least 15 min until field parameters (temperature, pH, dissolved oxygen, specific conductance, and ORP) were stabilized. Samples for chemistry were collected in acid-washed plastic containers using appropriate preservation techniques (see below). Samples used for molecular method analysis and bacterial indicators were taken using sterile techniques. For samples collected from private wells, samples were collected from spigots. Prior to sampling, the spigot was flame-sterilized. Autoclaved bottles were used for sampling. Samples were preserved as appropriate and stored in ice-filled coolers before transporting back to the analytical laboratories. Spring samples were collected by dropping the sampling container to the spring discharge point as close as possible. Deionized water blanks were collected for each sampling trip.

Standard water quality parameters including water temperature, pH, specific conductivity, and ORP were measured in the field with specific probes (Hydrolab). Samples used for chemical analyses were filtered through a 0.45 μm filter and stored at 4^{o}C prior to analysis. Analyses of key inorganic chemicals were conducted according to the standard methods (3). For the measurements of NO_3-N, PO_4, alkalinity, and Cl^-, water samples were collected in 400 mL acid-washed plastic bottles, and analyzed within 24 h. Water samples for metal concentration analysis were collected in 400 mL acid-washed bottle, preserved with HCl, and analyzed within two weeks of collection. The metals analyzed included Al, As, B, Ba, Be, Ca, Cd, Co, Cr, Cu, Fe, K, Li, Mg, Mn, Mo, Na, Ni, Pb, S, Sb, Se, Si, Sn, Sr, Ti, Tl, V, and Zn. Other inorganic parameters measured included F, Br, and SO_4^{2-}. Based on these data, total dissolved solids (TDS) and hardness were calculated. Bacterial indicators determined included total coliform, *Escherichia coli*, and

Enterococcus by IDEXX method (45). The sample collection and analysis for chemicals and bacterial indicators were conducted in collaboration with the Illinois State Water Survey and the Illinois State Geological Survey.

For molecular-based MST analysis, water samples were processed according to the procedure shown in Figure 2. Water samples were filtered through 0.22 µm filters to concentrate bacterial cells. The concentrated bacterial biomass was used for molecular analysis.

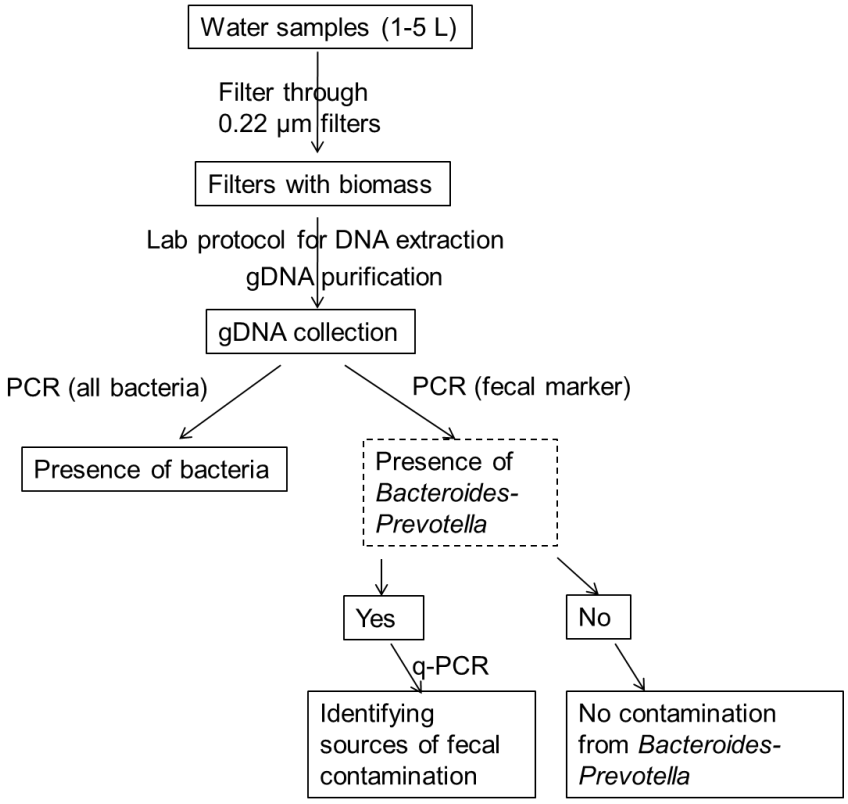

Figure 2. Experimental procedure.

2.3 Genomic DNA extraction

Genomic DNA was extracted from biomass of water samples according to the protocol described by Schmidt *et al.* (67) with some modification. In brief, filters with

biomass were cut into pieces and put into 2 ml centrifuge tubes. Then, 600µl of extraction buffer (0.1 M Tris-HCl, 0.1 M EDTA, 0.75 M Sucrose) was added to 2 g of biomass with 0.3 g glass beads, and homogenized using a bead beater (maker and location) for 90 s. The mixture wad added with 6 µl of lysozyme (100 mg/ml), briefly vortexed, and incubated at 37 °C for 30 min. To fully lyse the cell, 3 µl of 20 mg/ml proteinase and 60 µl of 10 % SDS were added to the tube, and incubated at 37 ℃ for another 2 hrs. CTAB (60 µl of 10 % CTAB with 84 µl of 5M NaCl) was used to remove polysaccharides. The supernatants were purified twice with phenol-chloroform extraction followed by isopropanol precipitation and ethanol wash. After DNA extraction, the purity of gDNA was evaluated by measuring the absorbance at wavelengths of 230 nm, 260 nm, and 280 nm with a UV-spectrophotometer. Ratios of 2.0-2.2 for OD260/OD230 and 1.7-2.0 for OD260/OD280 suggested that good quality of gDNA was extracted. Otherwise, DNA extraction process should be repeated. The gDNA was also analyzed through gel electrophoresis using 1% agarose gel.

2.4 PCR and gel electrophoresis analysis

All the gDNA were purified before PCR to ensure successful amplification using Wizard® Genomic DNA Purification Kit (Promega). Genomic DNA from individual water samples was PCR-amplified separately using a domain *Bacteria*-specific primer set, 11F (5'-GTT TGA TCC TGG CTC AG-3') and 1492R (5'-GGY TAC CTT GTT ACG ACT T-3'), and a *Bacteroides-Prevotella*-specific primer set, Bac32F (5'-AAC GCT AGC TAC AGG CTT-3') and 1492R (5'-GGY TAC CTT GTT ACG ACT T-3') (Hong *et al.*, 2009). Each PCR mixture (25 µl in volume) contained 20 to 30 ng of gDNA in 1X

DNA polymerase buffer, 25 nM of one set of forward and reverse primer, 100 mM of dNTP, and 0.5 U of Ex-Taq DNA polymerase. The reaction mixture was subjected to 30-35 cycles of thermal amplification, consisting of denaturation (95 ^{o}C for 30 s), annealing (55 ^{o}C for 45 s), and extension (72 ^{o}C for 60 s), to generate amplicons. PCR products were visualized with gel electrophoresis.

2.5 Q-PCR

Q-PCR was used to identify the sources of fecal pollution. A total of seven primer sets were tested for each sample (Table 1). PCR amplicons from four bacterial reference strains (*B. fragilis* BCRC10619, *B. uniformis* JCM5828, *B. caccae* JCM9498, and *B. vulgatus* BCRC12903), one uncultivated pig-specific *Bacteroidales* clone, and one uncultivated cow-specific *Bacteroidales* clone) were used as reference 16S rRNA genes. The PCR condition was 96 ℃ for 5 min, followed by 25 cycles of denaturation at 96 ℃ for 10 s, annealing at 55 ℃ for 30 s, extension at 72 ℃ for 45 s, and after the cycles a further extension at 72 ℃ for 10min. PCR was performed with a C1000 Thermal Cycler (Bio-Rad). The PCR products were purified using Wizard® SV Gel and PCR Clean-Up System (Promega). The Q-PCR experiments were performed as described previously (Hong *et al*., 2009). Standard curves were established in duplicates with 400 nM of forward and reverse primers, 1×iQ SybrGreen mastermix (Bio-Rad, CA), and different concentrations of the six reference DNA (0.0001-100 pg per 25ul of reaction mixture). Forty cycles of thermal program [denaturation at 96 ℃ for 10 s, annealing at 60 ℃ (57 ℃ for cow-specific primer set) for 30 s, extension at 72 ℃ for 30 s] were performed with CFX96 Real-Time System (Bio-Rad). Melting curve was carried out following the

normal Q-PCR program to check the specificity of annealing. Ct values were plotted against logarithmic concentration of DNA template to generate standard curves using the software by CFX96 Real-Time System (Bio-Rad). Quantification of bacterial targets from water samples was performed in duplicates with 10 ng of gDNA. The molecular weight of each reference strains was calculated to be 9.2×10^5 g/mol. The mass per copy of reference strains was estimated to be 1.53×10^{-18} g-DNA.

Table 1. Primers used in Q-PCR for determining sources of fecal contamination (31, 38)

Target	**Primer**	**Primer sequence (5'-3')**
Bacteroides-Prevotella	Bac32F	AAC GCT AGC TAC AGG CTT
	Bac303R	CCA ATG TGG GGG ACC TT
Pig-specific uncultivated *Bacteroidales*	Bac32F	AAC GCT AGC TAC AGG CTT
	P163R	TCA TAC GGT ATT AAT CCG C
Cow-specific uncultivated *Bacteroidales*	C367F	GGA AGA CTG AAC CAG CCA AGT A
	C467R	GCT TAT TCA TAC GGT ACA TAC AAG
B. fragilis	927F	GGG CCC GCA CAA GCG G
	Bfrg1024R	TCA CAG CGG TGA TTG CTC A
B. uniformis	927F	GGG CCC GCA CAA GCG G
	Bufm1018R	CTG CCT TGC GGC TGA CA
B. caccae	927F	GGG CCC GCA CAA GCG G
	Bcc1066R	CGT ATG GGT TTC CCC ATA A
B. vulgatus	927F	GGG CCC GCA CAA GCG G
	Bvg1016R	ATG CCT TGC GGC TTA CGG C

2.6 Statistical analysis

Cluster analysis (CA) was performed using selected important environmental parameters to evaluate differences among sampling sites. CA was done with Primer 6 (http://www.primer-e.com/primer.htm). Gradient analysis was created with the software Canoco (http://www.pri.wur.nl/uk/products/canoco/). These analyses were used to identify important environmental factors, and facilitate the interpretation of findings from

bacterial indicators and fecal markers (21, 73, 77). These parameters included pH, SpC, ORP, B, Ba, Ca, Cu, Fe, K, Mg, Mn, Na, Si, Sr, Zn, alkalinity, F, Cl, NO_3-N, SO_4^{2-}, o-PO_4-P, NVOC, TKN, NH_3-N, and calc TDS.

CHAPTER 3
RESULTS

3.1 Composition of water samples

A total of 73 water samples were collected within the two-year period, in which five blank samples were included as the negative control. These sampling sites were located in east Wisconsin, at the border of Illinois and Missouri, and border of Indiana and Kentucky (Figure 3). Forty-six samples were collected in the spring season compared with nine for summer, seven for fall, and 11 for winter as shown in Figure 4. Six sites in Wisconsin were sampled more than once at different seasons. Samples from Illinois and Wisconsin (47 samples) were all obtained from wells, while samples from Missouri and Kentucky (26 samples) were from springs.

3.2 Environmental parameters

Major environmental parameters of water samples were analyzed. They included, temperature, pH, dissolved oxygen (DO), ORP, alkalinity, calcium, iron, phosphorus, chloride, nitrate, sulfate, ammonia, and metals. Temperature for all water samples was between 8 and 22 ℃. pH of all samples ranged from 6.56 to 7.54 with the exception of three samples having pH higher than 8. DO was under 10 mg/L with an average of 5.5 mg/L. Samples from KY had lower alkalinity (95-227 mg $CaCO_3$/L) than all the other samples (200-360 mg $CaCO_3$/L). Concentrations of TKN, phosphorus, and sulfate were low for most water samples (Figure 5). Chloride concentrations were averaged 28.80 mg/l. When 15 mg Cl^-/l was considered as the background level (49), 40 out of 47 well

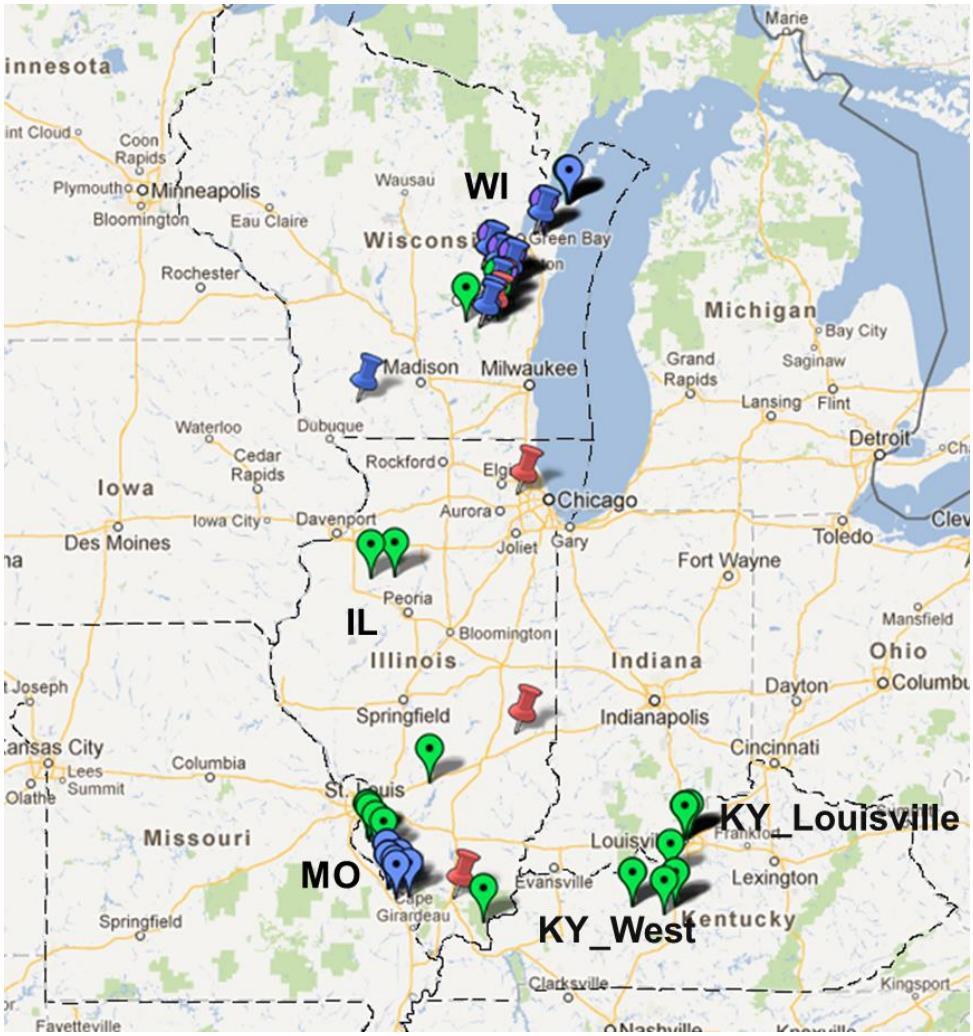

Figure 3. Locations of sampling sites. Samples were taken from WI, MO, IL, KY_West, and KY_Louisville during four seasons and six sites from WI had repeat samples in different seasons. Green, red, purple, and blue colors indicate samples taken in spring (Mar 20-Jun 21), summer (Jun 22-Sep 22), fall (Sep 23-Nov 6) and winter (Nov 7-Mar 19), respectively.

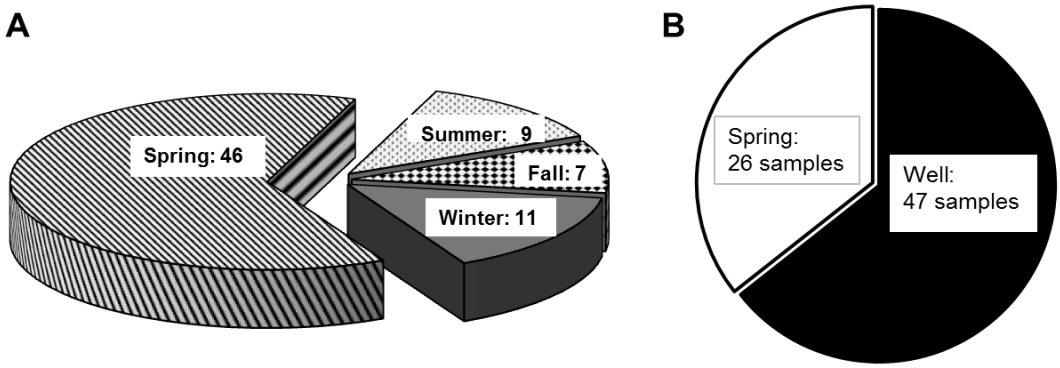

Figure 4. Sample composition categorized by seasons (A) and sample locations (B) (73 water samples in total).

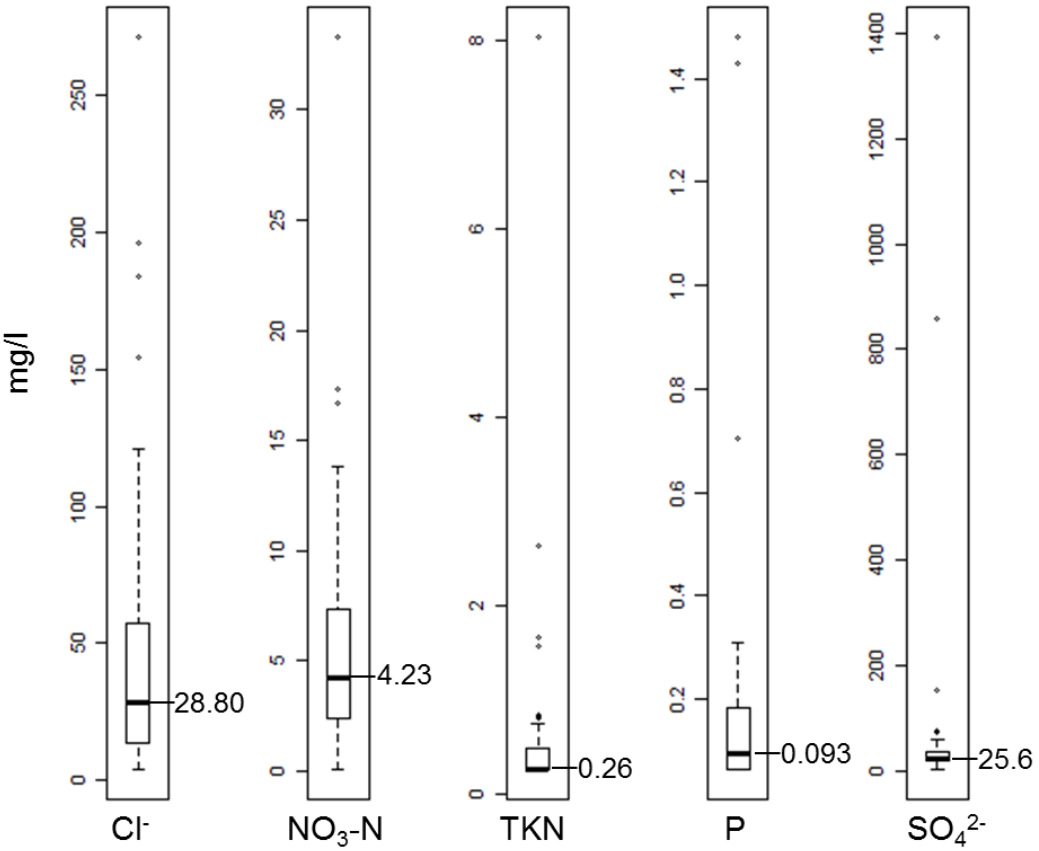

Figure 5. Selected water chemistry data as presented using box-and-whisker plot for all samples. The boundaries of the box indicated 25th and 75th percentile while the middle line within the box represented 50th percentile. Whiskers above and below the box were 10th and 90th percentiles. A "◦" marks outliner measurements. The median for each parameter was shown on the right side.

samples and 11 out of 22 spring samples have exceeded this criterion. When 2.5 mg/l was set as the background concentration of nitrate (51), 33 well samples were higher than background, compared with 16 spring samples. Thirty-two well samples and ten spring samples with high level of nitrate overlapped with samples with elevated level of chloride.

All samples were clustered based on selected environmental parameters (Figure 6). Most samples formed one single cluster in the center of Figure 6A, indicating high similarities among these samples. Four samples from IL were grouped together, due to elevated boron concentration typically originated from detergents containing sodium

perborate (34, 83). Two WI samples (cluster "3") were observed with high level of metals (Cu, Fe, and Mn). Two MO samples (cluster "2") had elevated level for nutrients (P, NVOC, TKN, and NH_3-N). Samples S22, S19, S34, and S47 each formed distinct clusters. S19 was observed to have elevated concentrations of SpC, Ba, Sn, SO_4^{2-} and calc TDS). S34 had elevated concentrations of B, Ca, Fe, Mg, Mn, Ni, alkalinity, SO_4^{2-} and calc TDS. S22 had elevated concentrations of Na^+, and Cl^-. S47 had high concentrations of B, Cu, Fe, K, Mn, P, o-PO_4-P, TKN, and NH_3-N. The center cluster "1" in Figure 6A could be further divided into five subgroups and four singletons (Fig. 6B). The grouping pattern of those samples mostly coincided with the classification of geological regions except for three samples at the intersection of IL and KY_ Louisville, which likely shared similar geological conditions.

In addition to chemical parameters, common bacterial indicators were also tested in those water samples (Table 2). More than half of the water samples (45 samples) were positive for total coliform. Among them, 31 were positive for *E. coli*, and 29 samples were positive for enterococci. When samples from wells and springs were separated, more samples from springs exhibited positive for all the three fecal indicators, and most of them exceeded the detection limit except one well sample. This suggested that these water samples were probably contaminated with human or animal feces, especially samples taken from springs, and could be tested with other MST techniques to identify the contamination source.

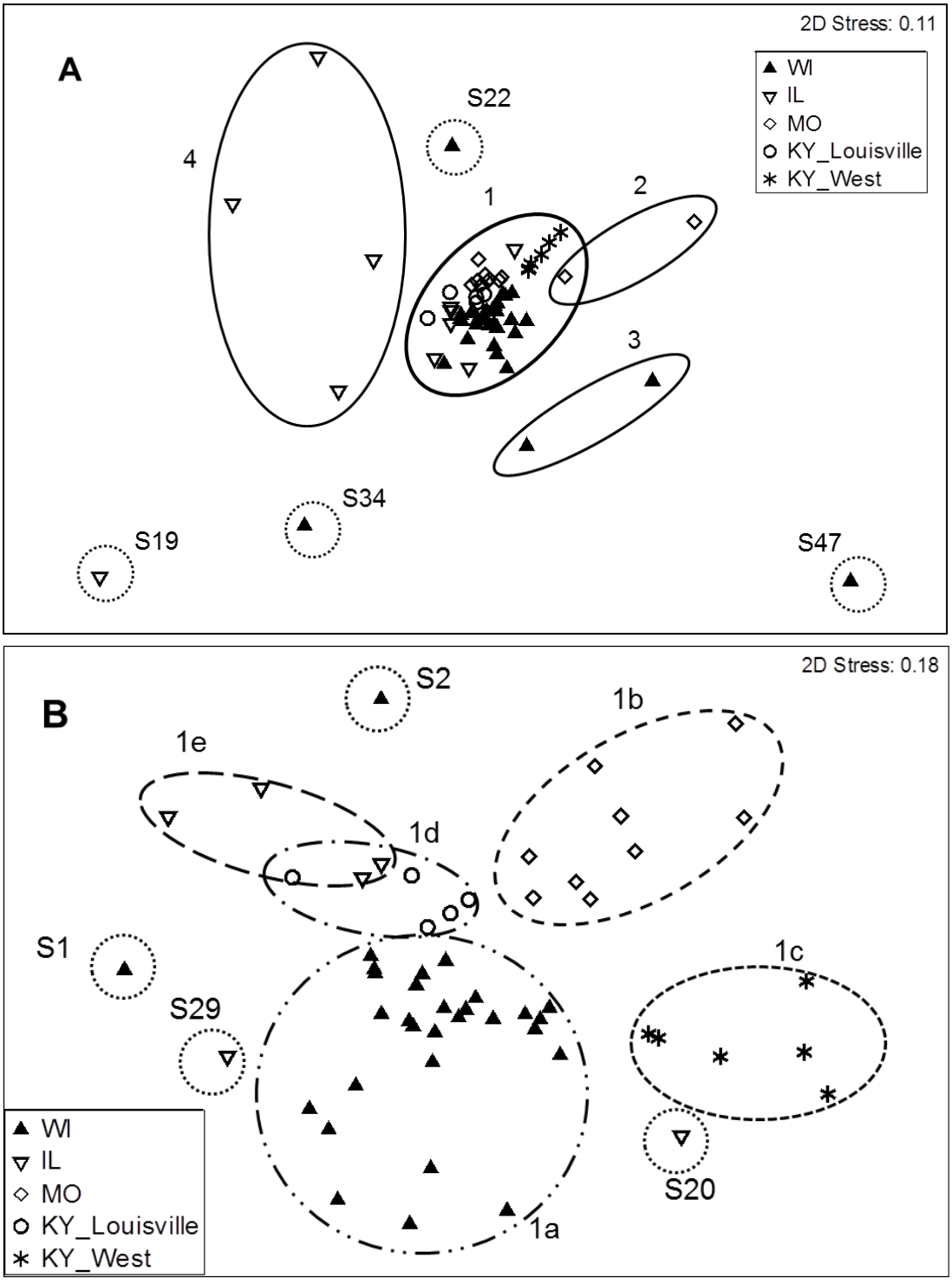

Figure 6. Multidimentional scaling (MDS) analysis plot based on D1 Euclidean distance with log (X+1) transformation of environmental parameters. Environmental parameters used in the analysis included pH, SpC, ORP, B, Ba, Ca, Cu, Fe, K, Mg, Mn, Na, Si, Sr, Zn, alkalinity, F, Cl, NO_3-N, SO_4^{2-}, o-PO_4-P, NVOC, TKN, NH_3-N, and calc TDS. All the samples except five negative controls were included in Panel A. The centered cluster "1" (60 samples) was further clustered into subgroups in Panel B.

Table 2. Testing results for bacterial indicators

	Sample numbers		
	All samples	Well samples	Spring samples
Positive for total coliform	45	22	23
Positive for *E. coli*	31	10	21
Positive for enterococci	29	8	21
⩾1419 CFU for total coliform	14	1	13
⩾3419 CFU for *E. coli*	3	0	3
⩾2419 CFU for enterococci	1	1	0
Total	73	47	26

3.3 PCR screening

To determine the source of fecal contamination, two PCR screening steps were used (Figure 7). First, all samples were tested for primer set targeting all bacteria. Among the 73 samples, 62 were tested positive for the presence of bacterial cells. Among those bacteria-positive samples, 49 were tested positive with the primer set specific for fecal *Bacteroides-Prevotella*, and were furthered analyzed using Q-PCR-based MST.

3.4 Standard curves for Q-PCR

Standard curves were established using the six reference 16S rRNA genes obtained (Table 3). The correlation coefficient (R^2) was always higher than 0.99. The amplification efficiency (E) was between 60 and 75% for six primer sets and 52% for the primer set targeting *B. uniformis*. All these efficiencies were not as high as reported in other studies, which were between 78 and 102%. The target DNA concentration was between 653- 6.53×10^6 copies, and was comparable with other studies, which were from

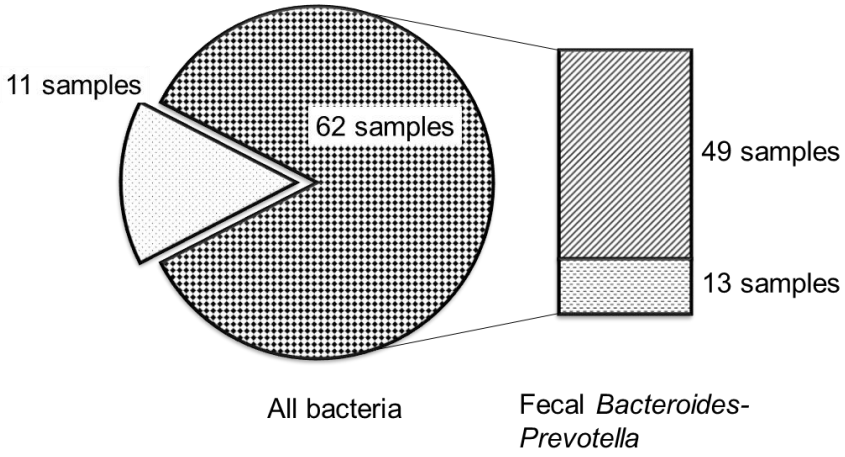

Figure 7. Water sample screening process. All the samples were first tested against primer sets targeting all bacteria. Samples that were positive were further analyzed using primer sets targeting fecal *Bacteroides-Prevotella*.

Table 3. Q-PCR standard curve equations and performance characteristics

Assay	Standard curve	Amplification efficiency (E)	Range of quantification (ROQ) (copies) for target DNA	R^2	%CV across ROQ
Bacteroides-Prevotella	y=8.28-4.33x	70.25%	6-6.53×10^6	0.998	2.27
B. fragilis	y=6.84-4.83x	61.05%	65-6.53×10^6	0.996	4.19
B. caccae	y=6.39-4.30x	70.88%	6-6.53×10^6	0.998	2.48
B. uniformis	y=10.17-5.45x	52.56%	653-6.53×10^6	0.994	0.87
B. vulgatus	y=7.33-4.13x	74.70%	6-6.53×10^6	0.998	2.60
Swine	y=8.19-3.96x	78.96%	65-6.53×10^6	0.999	1.61
Bovine	y=9.77-4.47x	67.46%	65-6.53×10^6	0.999	2.55

"Amplification efficiency", equal to ($10^{-1/slope}$)-1.
"Range of quantification for target DNA", the range of quantifications of reference DNA for each standard curve.
"%CV across ROQ", the mean percent coefficient of variations measured for quantifications of standards.

10 to 2.5×10^6 copies (38, 48, 69, 71). As the standard curve was repeated each time when water samples were tested, statistics for those curves would be slightly different.

3.5 Using Q-PCR technique to identify sources of fecal contamination

Seven Q-PCR were conducted with one primer set targeting fecal *Bacteroides-Prevotella* group and the remaining six primer sets targeting different host-specific contamination (four for human, and one each for swine and bovine). Among those 49 samples tested positive with Q-PCR targeting fecal *Bacteroides-Prevotella*, six samples had concentrations below the detection limit and were excluded from testing with the remaining primer sets. Primer sets targeting *B. fragilis*, *B. caccae*, *B. uniformis*, and *B. vulgatus* were used to indicate human fecal contamination, and primer sets targeting certain *Prevotella* group were used to indicate fecal contamination from animal (swine and bovine). ANOVA analysis ($p<0.05$) confirmed the observation that tests with *B. fragilis*, *B. vulgatus,* and bovine primer sets gave significantly higher concentrations (with median more than 10^2 copies/ng-gDNA) than all the other three primer sets (Figure 8). One sample that was negative for fecal marker in PCR screening process was randomly chosen to check with primer sets targeting *Bacteroides-Prevotella* by Q-PCR. No target could be detected or the target was below detection limit.

Level of fecal biomarkers between spring samples and well samples were separated and compared due to the different mechanisms of pollution associated with each water type (Figure 9). In general, spring samples contained higher level of biomarkers than well samples. Concentrations of biomarkers for *B. vulgatus*, swine, and bovine were significantly different ($p<0.05$) between the two water types. Average concentrations of fecal markers within each sampling area were shown in Table 4. For

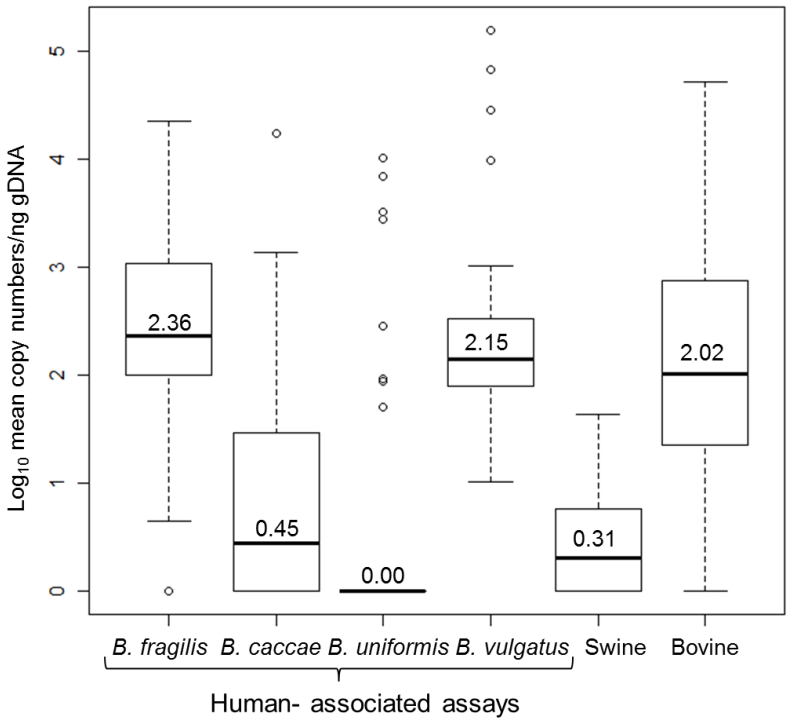

Figure 8. Q-PCR results for determining fecal contamination using box-and-whisker plot. Estimated gene copy numbers are shown in the unit, copies/(ng-gDNA) (log_{10} scale). The boundaries of the box indicated the 25th and 75th percentile while the middle line within the box represented the 50th percentile. Whiskers above and below the box were the 10th and 90th percentiles. A "◦" marks outliner measurements. The total number of samples included is 43.

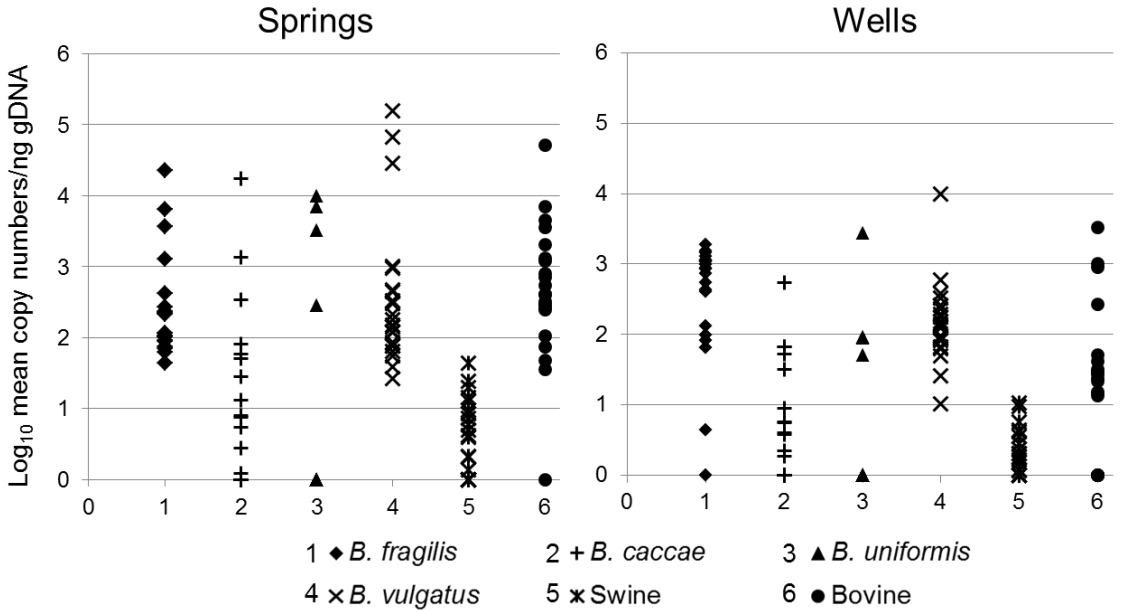

Figure 9. Concentrations of fecal biomarkers from springs (22 samples) and wells (21 samples).

samples taken from Lousiville, KY, the biomarkers were observed to be the highest, and were likely associated with the urban spring systems, which were the most vulnerable system within karst regions. Spring samples tended to have higher level of bovine markers than well samples. Well samples had higher concentrations of biomarkers for human pollution.

Table 4. Average concentration of each biomarker in the unit of copies/ng gDNA among water samples (combing results from PCR screening process and Q-PCR process).

Sample sites	Water Type	Number of samples	Human				Animal	
			B. fragilis	*B. caccae*	*B. uniformis*	*B. vulgatus*	Swine	Bovine
IL	Wells	10	87.0	54.0	284.1	989.8	0.4	332.0
WI	Wells	25	594.5	9.9	7.3	128.0	2.0	99.3
KY_Louisville	Springs	5	6821.8	3760.3	4125.2	50672.1	17.2	12288.6
KY_WEST	Springs	6	107.7	14.6	0.0	142.7	6.1	701.0
MO	Springs	11	143.9	40.3	0.0	258.8	3.8	981.7

Five negative controls and 11 samples were not included. There was not enough gDNA for those 11 water samples.

Based on Q-PCR results, we could discriminate sources of fecal pollution for each water sample (Table 5). Human fecal pollution was determined based on the presence of at least two or three detectable primer sets. When two primer sets were used, almost all spring samples were contaminated with human feces and bovine feces. Contamination by swine feces occurred to a lesser extent, but was still observed with more than half of the spring samples. Only two well samples from IL contained more than two detectable biomarkers for human pollution, compared with more than 70% (18 samples) from WI. When three positive biomarkers were used, samples from WI, west of KY, and MO were affected the most, suggesting that the level of contamination for a few samples were near

the detection limit or less. Although the presence of all four human biomarker was ideal, it could not be applied in this study because one of the human primer set was not working well.

Table 5. Number of samples that were contaminated by human and animal wastes

Sample sites	Water type	Number of samples	Human		Animal	
			+ in ≥ 2 primer sets	+ in ≥ 3 primer sets	Swine	Bovine
IL	Wells	10	2	2	2	1
WI	Wells	25	18	13	13	14
KY_Louisville	Springs	5	5	4	4	5
KY_WEST	Springs	6	6	2	4	5
MO	Springs	11	11	7	6	11

Five negative controls and 11 samples were not included. There was not enough gDNA for those 11 water samples.

Contamination by both human and animal fecal pollution was the dominant type of pollution among all samples (25 out of 57) (Table 6). Only three samples showed pollution solely by human whereas 15 samples for animal. Forty-three samples were contaminated by at least one source. Again, all the spring samples were contaminated by human or animal sources. Six well samples from WI were contaminated only by animals. Table 6 further indicates a correlation between presences of contamination with total coliform. When data from each sample were examined individually, there was discrepancy between those two methods (Figure 10). Five samples could not be detected with *Bacteroidales*-MST, but showed positive for traditional indicators. Another nine samples contained detectable *Bacteroidales*-MST biomarkers, but showed negative for traditional indicators.

Table 6. Number of samples contaminated within each sampling area and comparison with number of samples that was positive for traditional indicators.

Sample sites	Water type	number of samples	Human only	Animal only	Both human and animal	Contaminated by at least one source	Traditional indicators		
							Total coliform	*E. coli*	Enterococci
IL	Wells	10	0	0	2	2	2	0	0
WI	Wells	25	2	6	11	19	17	8	8
KY_Louisville	Springs	5	0	1	4	5	4	4	4
KY_WEST	Springs	6	1	4	1	6	6	6	6
MO	Springs	11	0	4	7	11	11	11	11

Five negative controls and 11 samples were not included. There was not enough gDNA for those 11 water samples.

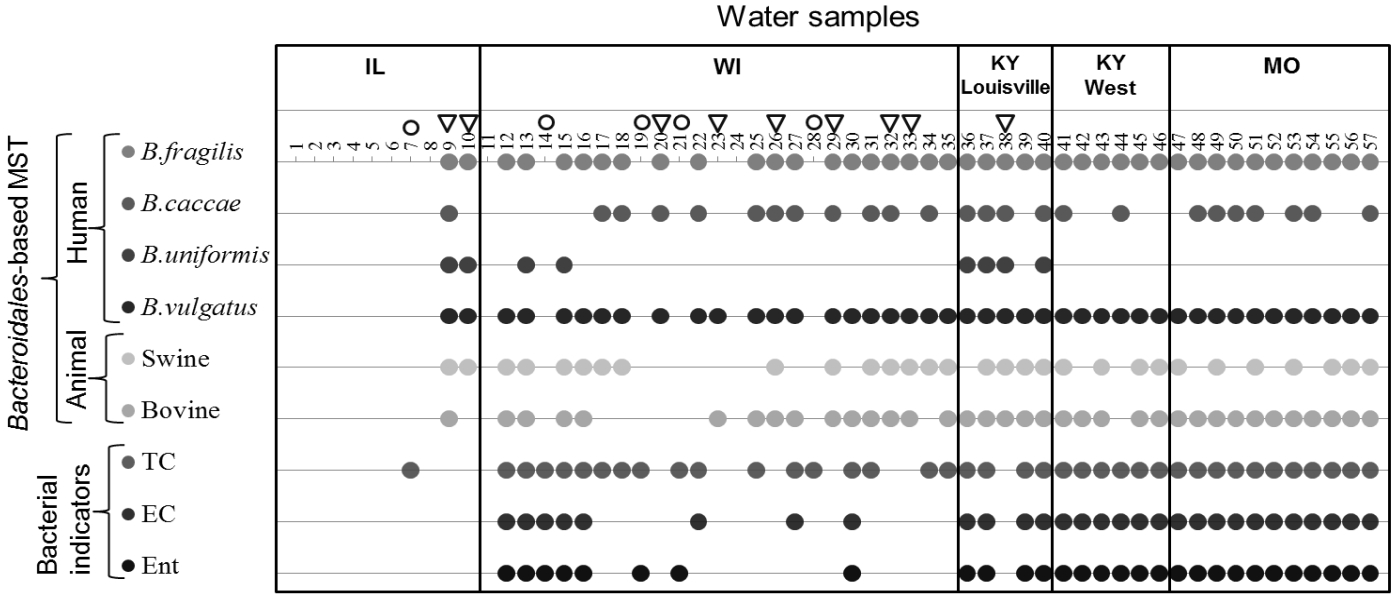

Figure 10. Comparison of contamination identified by *Bacteroidales*-based MST and traditional bacterial indicators. **O** indicated samples that were positive for traditional bacterial indicators but could not be detected with *Bacteroidales*-MST. **▽** highlighted samples contained detectable *Bacteroidales*-MST biomarkers, but showed negative for traditional indicators.

3.6 Contamination identified with concentrations of chloride and nitrogen

Infiltration of fecal materials into wells and springs is usually associated with elevated levels of nitrogen and chloride. Therefore, possible sources of contamination (fertilizer, road salt or human/animal waste) might be identified with known concentrations of nitrogen and chloride (49, 51). Background concentrations of total nitrogen (TN) and Cl^- are usually defined as 2.5 mg/l and 15 mg/l, respectively (Figure 11). When concentration of total nitrogen is more than 15 mg/l, the sample is considered to be contaminated significantly by human/animal waste. When TN and chloride are within 2.5-15 mg/l and 0-15 mg/l, respectively, the elevated concentrations are usually originated from anthropogenic contamination. Samples with elevated TN concentrations and background Cl^- level are usually associated with contamination by fertilizers. High Cl^- level with background TN level is usually associated with road salts application.

However, the aforementioned method showed partial correlation with our *Bacteroidales*-MST method. Identified sources by the two methods did not agree with each other for spring samples. All spring samples were contaminated by human and/or animal feces identified by our results. However, the graph classified most samples into fertilizer category. Even samples fell within background region were determined to be contaminated. On the contrary, well samples classified into background, anthropogenic contamination and significant influence categories were placed correctly, except road salt category. Therefore, the method based on concentrations of chloride and nitrogen was not suitable for identifying fecal sources of spring samples, while it was useful for well samples.

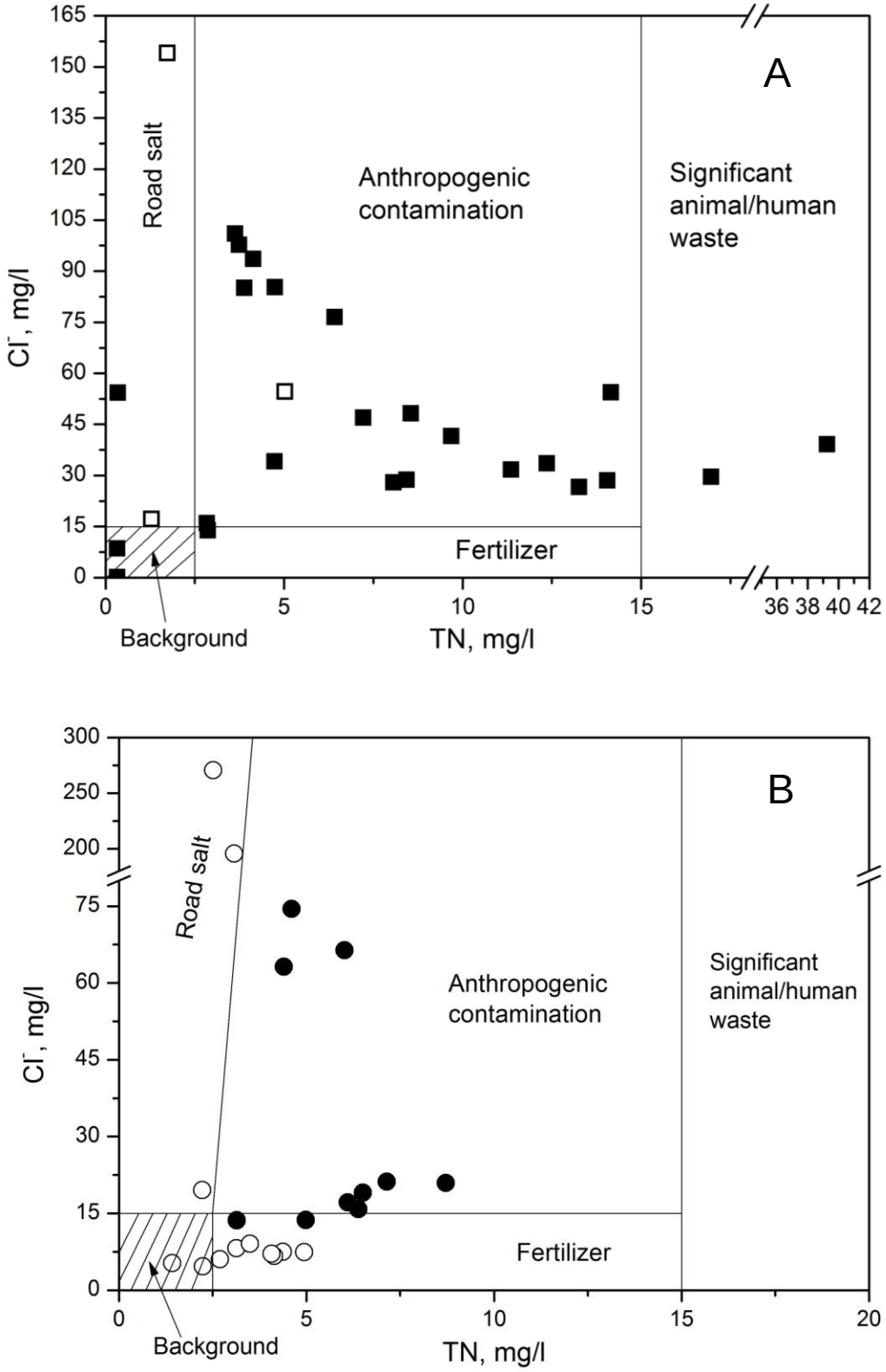

Figure 11. Cl^- vs. TN plots for (a) well samples and (b) spring samples. The domains were simplified according to two previous studies (49, 51). Solid signs indicated agreement with *Bacteroidales*-MST results and open signs were for disagreement.

3.7 Correlation between environmental variables and fecal contamination

To find putative correlations between environmental variables and fecal contamination, ordination analysis was performed (Figure 12). Three environmental parameters were determined to correlate with level of *Bacteroidales* fecal biomarkers. These variables were *E. coli*, Cl^-, and ORP. The two axes shown in Figure 12 could explain almost all variations between samples. *E. coli* was shown to be a very important factor associated with fecal contamination (15, 16, 38). However, total coliform count did not correlate well with those fecal biomarkers. It was reported that the dominant source of Cl^- in karst regions was from row crops followed by septic effluent and animal wastes (49). The Cl^- level varied greatly from 4.63 to 271 mg/l among all samples. Spring and well samples had similar levels of ORP except a few well samples with extreme high concentrations. ORP of almost all water samples were greater than 300 mV, suggesting that all samples were oxygenated.

Well samples seemed to be affected by ORP and Cl^- the most, and to a lesser extent by *E. coli*, whereas samples from springs were influenced by all three parameters. Bacterial indicators were not as frequently detected from well samples as from spring samples, which agreed with this result. Negative controls were separated from well and spring samples. Distribution of samples from Louisville, KY was more dispersed compared with samples from west of KY. Samples from MO and WI were correlated more with *E. coli* and ORP, respectively.

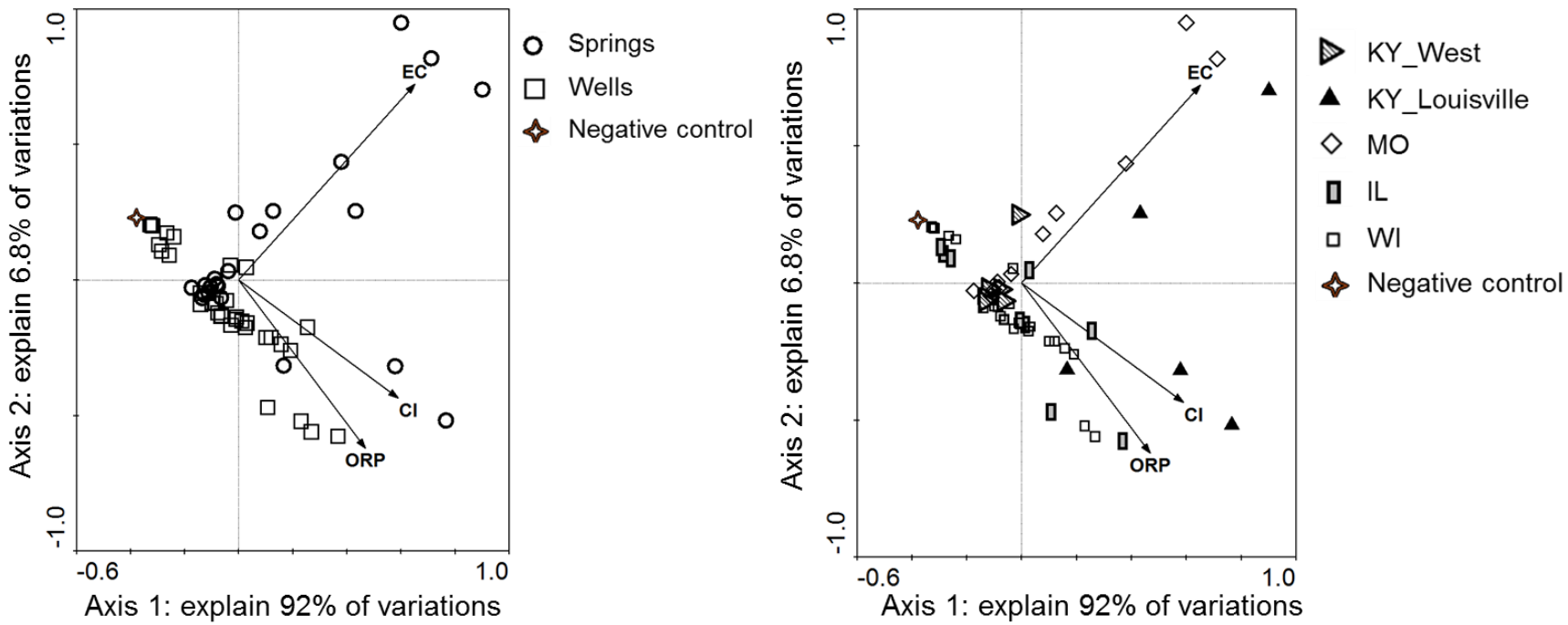

Figure 12. Ordination analysis of environmental variables and *Bacteroidales* fecal biomarkers by redundancy analysis. Six individual biomarkers were included in addition to selected important environmental parameters. *E. coli*, Cl^-, and ORP were identified as the significant factors ($p<0.05$). This diagram accounted for 98.8% of the variance in the environmental variable. Samples in Panel A were categorized by water types (springs and wells), whereas Panel B was grouped by sampling areas. Eleven samples that did not have enough gDNA were excluded. Each arrow represented an environmental variable and the direction of the arrow pointed the steepest increase of values of environmental variable. Distance between sample points indicated the dissimilarity of their fecal biomarker composition, measured by Euclidean distance.

CHAPTER 4
DISCUSSION

4.1 Sensitivity and specificity of Q-PCR-based MST

To obtain useful MST information involves not only the Q-PCR results but also the whole analytical process, including selection of sampling sites, sample collection and filtration, DNA extraction, removal of PCR inhibitors, establishment of reference DNA and standard curves, control of cross contamination, and estimation of gene copy numbers based on raw data. Although Q-PCR technique is reliable and easy to apply to the field study, maintain high quality of control during each step in the measurement is rather difficult. Field et al. (18) reported that during the comparison of methodology among five participated laboratories, on average one laboratory could not obtain sufficient gDNA from samples tested. Another study reported a spike recovery rate of as low as 63% for plasmids (38), indicating the high likelihood of having low DNA recovery. In the beginning of this study, only 1-liter water was collected for 11 samples, which were detected negatively with PCR targeting all bacteria due to likely the inability to extract sufficient good quality gDNA. Even when sufficient concentrations of gDNA were obtained, presence of inhibitors in gDNA could lead to failure of PCR and Q-PCR. To prevent inhibition, gDNA extracts were purified using a commercially available kit and then diluted to 10 ng/µl, which could reduce the concentration of potential inhibitors such as humic acids. Another solution is to use internal amplification control to evaluate the effect of inhibitors in PCR-based reactions (18, 71).

One problem with Q-PCR is the low amplification efficiency. Amplification efficiency for most primer sets was between 61 and 79%, and was lower than that reported previously between 78 to 102% (38, 48, 69, 71). We further examined the

secondary structure of amplicons obtained from primer sets targeting *B. fragilis*, *B. uniformis*, and bovine, and observed the formation of complex secondary structure. This could also affect the amplification efficiency, gave false-negative results, and underestimated the concentrations of *Bacteroidales* biomarkers in all samples.

Shanks *et al.* tested ten Q-PCR assays for human-associated *Bacteroidales*. They suggested the performance of the assay can be evaluated by six criteria: "1) limit of quantification $\geqslant$ 25 gene copies, 2) %CV across ROQ $\leqslant$ 5.0%, 3) fitted calibration curve R^2 values eria: "1) limit of quantification $\geqslant$ 25 gene copies, 2) %ng total DNA per reaction >99%, 5) mean nontarget abundance of genetic marker <0.50 $\log_{10}$ mean copy number, and 6) target abundance of genetic marker >3.0 log10 mean copy number" (71). In their study, only three out of ten primer sets satisfied all the criteria, suggesting that vigorous validation on the primer sets designed is very important for obtaining convincing MST results. Our primer sets were not included in their study, but was previously tested using another molecular technique, where high level of specificity and sensitivity were demonstrated (31).

A few groundwater samples tested in our study were highly contaminated with human and/or bovine feces. Shank et al. (69) reported a mean of three copies per ng gDNA for bovine fecal *Bacteroidales* marker in ruminant feces. Our results showed a much higher copy number, ranging from 15 to 5.12×10^4 per ng gDNA. This different might come from highly polluted groundwater, the sampling collection efficiency, and DNA extraction process. Another study, which tried to compare different primer sets targeting *Bacteroidales* for human fecal pollution, determined concentrations of human fecal markers in 54 sewage samples to be 10^2-10^5 copies/ng gDNA. Our tests got similar

ranges of concentrations for human markers with primer sets targeting *B. fragilis* (44 - 2.24 $\times 10^4$ copies/ng gDNA), *B. caccae* (1 - 5.37 $\times 10^2$ copies/ng gDNA), *B. uniformis* (51 - 1 $\times 10^4$ copies/ng gDNA), and *B.vulgatus* (10 - 1.58 $\times 10^5$ copies/ng gDNA). The high level of fecal contamination identified in our study is not surprising, as previous studies on Midwestern karst regions have detected elevated level of nutrients, bacterial indicators, natural estrogen, pesticides and presence of male-specific coliphage (20, 35, 57, 82).

Because different units were used to report the final concentrations of fecal biomarkers, the results obtained in this study could be converted to copy number/100ml water samples to enable further comparison. On average, the gDNA concentrations extracted from two liters of water samples were 90 ng/µl with a final volume of 50 µl. Thus, a total of 4500 ng gDNA could be recovered from two liters, and the concentration was equal to 225 ng gDNA/100 ml. If Q-PCR process identified 10^3 copies/ ng gDNA, then the copy number would be 2.25 $\times 10^5$ in 100 ml water sample. Previous study on river water samples showed that the concentration of total *Bacteroides* to be 2.5 $\times 10^2$-1.0 $\times 10^4$ copies/100ml, which was one log lower than our estimated number (66). Another study reported much higher concentrations of total and human-specific *Bacteroides* markers with averages of 9.8 $\times 10^8$ and 4.8 $\times 10^7$ copies/100ml, respectively, for stormwater outfalls. This result could be expected due to their extremely high fecal indicator concentrations (4900-3,410,000 CFU/100ml water samples). To connect this number with cell number, the copy number of 16S rRNA genes in one cell would need to be assumed. In general, each genome of *Bacteroides* contains four to six copies of 16S rRNA genes with an average of 5.57 copies (36).

4.2 Multi-metric approaches: relationship among traditional indicators (viruses), environmental variables and *Bacteroidales*-MST

The complexity and cost of MST are dependent on the composition and available knowledge of water systems. If the composition of landscape, infrastructure, and human and animal populations is simple, the potential sources of contamination can be easily screened and confirmed with methods effectively targeting the specific sources. In this case, dominant fecal sources (contributing >50%) can be detected with high confidence with assays having specificity higher than 50% while sources contributing 10-50% of the pollution would need methods with specificity higher than 90%. Pathogens, if present, can be identified and isolated following a confirmation step. In the situation that the composition of study area is complicated, such as urban systems with significant aging infrastructure and numerous non-point sources of pollution, Q-PCR may serve as an easy and cost-effective method to get an initial knowledge of the sources (65). With the previous statistical analysis, only dominant and mediate sources (contributing more than 10% of pollution) might be effectively detected considering a specificity of Q-PCR at 80-100%. Combining with information of local land-use and human and animal activities, more specific and focused MST methods can be applied to hotspot areas, and this can lead to better understanding of the sources of fecal contamination.

Although traditional bacterial indicators are easy, less expensive, and suitable for routine monitoring, their correlation or occurrence with fecal pollution and presence of pathogens is weak (33, 34, 40, 41, 46-48, 53, 65) except a few studies (13, 16, 38). This weak relationship is also illustrated in our study. Five samples contained traditional indicators could not be detected with *Bacteroidales*-MST and another nine samples

having detectable *Bacteroidales*-MST biomarkers without traditional indicators. At the same time, the presence of *Bacteroidales* markers correlates well with presence of pathogens (47, 79). This discrepancy of indication function between traditional indicator and *Bacteroidales* markers may come from the different survival ability of the two groups. In addition, this ability can also be affected by factors such as temperature, sunlight, nutrient level, availability of oxygen, and presence of predators (48, 61).

In addition to traditional indicators, we used a method based on concentrations of chloride and total nitrogen to discriminate fecal source pollution. However, the method did not correlate with *Bacteroidales*-MST. As all approaches (i.e., traditional indicators, *Bacteroidales*-MST, and nutrient level) did not agree, multi-metric or multi-tiered approaches have been proposed by a number of studies (21, 34, 47, 53, 58, 60, 65, 79). These approaches used more than one of the following categories to validate the sources of contamination: organic and inorganic compounds (N, P, Cl^-, organic wastewater compounds, pharmaceutical compounds), indicator bacteria, viruses, and MST (caffeine, sterols, ARA, *Bacteroidales*). The selection of a proper combination of techniques, which can verify results from different techniques as well as get different perspectives of the problem, depends on resources availability, such as manpower, time, and fund. For example, Furtula et al. (2012) adopted a multi-metric approach to investigate the effects of agricultural activities on surface water quality. The toolbox included inorganic nitrogen measurement, sterol analysis, and *Bacteroidales*-based MST. Sterol ratios showed a combination of human and animal contamination for most sampling events. Compared with results from *Bacteroides*-MST methods, only part of animal

contamination with no human contamination was confirmed. To solve the discrepancy between the two methods, other tiers need to be considered.

4.3 Implication for management practice in karst region

The karst regions are vulnerable to fecal contamination and lack of natural remediation. Many studies have been done to trace sources of pollution from this type of area worldwide (20, 33, 34, 54, 55, 57, 59). Therefore, the findings could be used to implement BMP and eventually reduce the level of pollution. For this purpose, timing of sampling and selection of representative sites are two important factors to consider. For the latter, it is crucial to identify hotspots and map them according to their vulnerabilities. In our study, all samples were likely associated with human sewage contamination, indicating that local septic tank systems need to be improved and well maintained. Spring systems are more vulnerable for animal fecal contamination than well systems. Limiting the access of livestocks, in particular cows to watershed can ease the pressure. In addition, karst region is susceptible to organic chemicals, such as pharmaceutical compounds. Future studies should focus on identifying contamination sources by combining chemical-MST and microbe-MST and evaluation of the effectiveness of BMP to provide a powerful toolbox for karst region management.

CHAPTER 5
CONCLUSION

Our results demonstrated that karst aquifers in Midwest were vulnerable to both human and non-human fecal contamination. Most water samples showed co-contamination by human and animal wastes as determined by *Bacteroidales*-based Q-PCR. This result was in accordance with the observation on the presence of on-site wastewater treatment systems in rural areas and anthropogenic activities in urban areas. Samples from urban spring systems (Lousiville, KY) had the highest level of fecal contamination, suggesting the vulnerability of karst aquifers due to human activities. Generally, spring systems were more vulnerable due to recharge from a much larger area. The overall contamination level from wells was much lower than from springs. Dilution effect was not enough to mitigate contamination in karst systems associated primarily with springs.

Q-PCR method was effective at distinguishing fecal contamination sources and more sensitive than traditional bacterial indicators to evaluate the overall contamination level. The Q-PCR method used in this study involved even reactions (each primer set was tested separately) for each sample, which was not convenient for analyzing a large number of samples. To enhance productivity in the future, studies could apply multiplex Q-PCR or other multiplexing analysis such as hierarchical oligonucleotide primer extension reaction (HOPE) (31) to obtain both qualitative and quantitative information for fecal contamination.

Bacteroidales-based MST, traditional bacterial indicators, and a variety of environmental variables were used in this study as an effective toolbox to identify fecal contamination in karst regions. Important environmental parameters, *E. coli*, Cl^-, and

ORP were correlated well with the level of *Bacteroidales* fecal biomarkers. Another concern of karst region was organic chemical contamination, such as pharmaceutical compounds, pesticides and their degradation products. The presence of these compounds in karst aquifers not only indicated anthropogenic influence but also was associated with human and animal health risks. Previous studies have shown the presence of these chemicals in karst groundwater systems (21, 22, 53). Future studies combining chemical-MST and microbe-MST to differentiate fecal source contamination are recommended to obtain the overall contamination profile in a specific geographic region like karst aquifers. With all the information, appropriate management practice can be designed and applied to reduce overall contamination level in the vast karst area around the world.

REFERENCES

1. **Allsop, K., and D. J. Stickler.** 1984. The enumeration of Bacteroides fragilis group organisms from sewage and natural waters. J. Appl. Bacteriol. **56:**15-24.

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

3. **APHA, AWWA, and WEF.** 1998. **Standard Methods for the Examination of Water and Wastewater, 20th edition**, p. 9-47-9-70. *In* Anonymous.

4. **Beaudeau, P., H. de Valk, V. Vaillant, C. Mannschott, C. Tillier, D. Mouly, and M. Ledrans.** 2008. Lessons learned from ten investigations of waterborne gastroenteritis outbreaks, France, 1998-2006. J. Water. Health. **6:**491-503. doi: 10.2166/wh.2008.051.

5. **Beaudeau, P., D. Valdes, D. Mouly, M. Stempfelet, and R. Seux.** 2010. Natural and technical factors in faecal contamination incidents of drinking water in small distribution networks, France, 2003-2004: a geographical study. J. Water. Health. **8:**20-34. doi: 10.2166/wh.2009.043.

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

7. **Breed, R. S., and J. F. Norton.** 1937. Nomenclature for the Colon Group. Am. J. Public Health Nations Health. **27:**560-563.

8. **Bustin, S. A., V. Benes, J. A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M. W. Pfaffl, G. L. Shipley, J. Vandesompele, and C. T. Wittwer.** 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry. **55:**611-622. doi: 10.1373/clinchem.2008.112797.

9. **Currens, J.** 2002. Changes in groundwater quality in a conduit-flow-dominated karst aquifer, following BMP implementation. Environmental Geology. **42:**525-531. doi: 10.1007/s00254-001-0515-6.

10. **D'Antonio, R. G., R. E. Winn, J. P. Taylor, T. L. Gustafson, W. L. Current, M. M. Rhodes, G. W. Gary Jr, and R. A. Zajac.** 1985. A waterborne outbreak of cryptosporidiosis in normal hosts. Ann. Intern. Med. **103:**886-888.

11. **Derek C. Ford, Paul Williams.** 2007. Karst Hydrogeology and Geomorphology. Wiley.

12. **Dick, L. K., A. E. Bernhard, T. J. Brodeur, J. W. Santo Domingo, J. M. Simpson, S. P. Walters, and K. G. Field.** 2005. Host distributions of uncultivated fecal

Bacteroidales bacteria reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. **71:**3184-3191. doi: 10.1128/AEM.71.6.3184-3191.2005.

13. **Dick, L. K., and K. G. Field.** 2004. Rapid estimation of numbers of fecal Bacteroidetes by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. **70:**5695-5697. doi: 10.1128/AEM.70.9.5695-5697.2004.

14. **Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, and D. A. Relman.** 2005. Diversity of the human intestinal microbial flora. Science. **308:**1635-1638. doi: 10.1126/science.1110591.

15. **Edberg, S. C., E. W. Rice, R. J. Karlin, and M. J. Allen.** 2000. Escherichia coli: the best biological drinking water indicator for public health protection. Symp. Ser. Soc. Appl. Microbiol. **(29):**106S-116S.

16. **Farnleitner, A. H., G. Ryzinska-Paier, G. H. Reischer, M. M. Burtscher, S. Knetsch, A. K. T. Kirschner, T. Dirnböck, G. Kuschnig, R. L. Mach, and R. Sommer.** 2010. Escherichia coli and enterococci are sensitive and reliable indicators for human, livestock and wildlife faecal pollution in alpine mountainous water resources. J. Appl. Microbiol. **109:**1599-1608. doi: 10.1111/j.1365-2672.2010.04788.x.

17. **Field, K. G., E. C. Chern, L. K. Dick, J. Fuhrman, J. Griffith, P. A. Holden, M. G. LaMontagne, J. Le, B. Olson, and M. T. Simonich.** 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. J. Water. Health. **1:**181-194.

18. **Field, K. G., E. C. Chern, L. K. Dick, J. Fuhrman, J. Griffith, P. A. Holden, M. G. LaMontagne, J. Le, B. Olson, and M. T. Simonich.** 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. J. Water. Health. **1:**181-194.

19. **Fiksdal, L., J. S. Maki, S. J. LaCroix, and J. T. Staley.** 1985. Survival and detection of Bacteroides spp., prospective indicator bacteria. Appl. Environ. Microbiol. **49:**148-150.

20. **Fox, J. T., G. Adams, M. Sharum, and K. L. Steelman.** 2010. Passive sampling of bioavailable organic chemicals in Perry County, Missouri cave streams. Environ. Sci. Technol. **44:**8835-8841. doi: 10.1021/es1019367.

21. **Furtula, V., H. Osachoff, G. Derksen, H. Juahir, A. Colodey, and P. Chambers.** 2012. Inorganic nitrogen, sterols and bacterial source tracking as tools to characterize water quality and possible contamination sources in surface water. Water Res. **46:**1079-1092. doi: 10.1016/j.watres.2011.12.002.

22. **Glassmeyer, S. T., E. T. Furlong, D. W. Kolpin, J. D. Cahill, S. D. Zaugg, S. L. Werner, M. T. Meyer, and D. D. Kryak.** 2005. Transport of Chemical and Microbial

Compounds from Known Wastewater Discharges: Potential for Use as Indicators of Human Fecal Contamination. Environ. Sci. Technol. **39:**5157-5169. doi: 10.1021/es048120k.

23. **Graves, A., C. Hagedorn, A. Teetor, M. Mahal, A. Booth, and R. Reneau.** 2002. Antibiotic resistance profiles to determine sources of fecal contamination in a rural Virginia watershed. J. Environ. Qual. **31:**1300-1308.

24. **Griffin, D., K. Donaldson, J. Paul, and J. Rose.** 2003. Pathogenic human viruses in coastal waters. Clin. Microbiol. Rev. **16:**129-+. doi: 10.1128/CMR.16.1.129-143.2003.

25. **Griffin, D., E. Lipp, M. McLaughlin, and J. Rose.** 2001. Marine recreation and public health microbiology: Quest for the ideal indicator. Bioscience. **51:**817-825. doi: 10.1641/0006-3568(2001)051[0817:MRAPHM]2.0.CO;2.

26. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. M. Grubbs, T. A. Angier, and R. B. Reneau.** 1999. Determining Sources of Fecal Pollution in a Rural Virginia Watershed with Antibiotic Resistance Patterns in Fecal Streptococci. Applied and Environmental Microbiology. **65:**5522-5531.

27. **Hagedorn, C., and S. Weisberg.** 2009. Chemical-based fecal source tracking methods: current status and guidelines for evaluation. Reviews in Environmental Science and Biotechnology. **8:**275-287. doi: 10.1007/s11157-009-9162-2.

28. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of Antibiotic Resistance Patterns of Indicator Bacteria by Discriminant Analysis: Use in Predicting the Source of Fecal Contamination in Subtropical Waters. Applied and Environmental Microbiology. **66:**3698-3704. doi: 10.1128/AEM.66.9.3698-3704.2000.

29. **Higuchi, R., C. Fockler, G. Dollinger, and R. Watson.** 1993. Kinetic Pcr Analysis - Real-Time Monitoring of Dna Amplification Reactions. Bio-Technology. **11:**1026-1030. doi: 10.1038/nbt0993-1026.

30. **Hong, P., J. Wu, and W. Liu.** 2008. Relative Abundance of Bacteroides spp. in Stools and Wastewaters as Determined by Hierarchical Oligonucleotide Primer Extension. Appl. Environ. Microbiol. **74:**2882-2893. doi: 10.1128/AEM.02568-07.

31. **Hong, P., J. Wu, and W. Liu.** 2009. A high-throughput and quantitative hierarchical oligonucleotide primer extension (HOPE)-based approach to identify sources of faecal contamination in water bodies. Environ. Microbiol. **11:**1672-1681. doi: 10.1111/j.1462-2920.2009.01892.x.

32. **John, D. E., and J. B. Rose.** 2005. Review of factors affecting microbial survival in groundwater. Environ. Sci. Technol. **39:**7345-7356.

33. **Johnson, T. B., L. D. McKay, A. C. Layton, S. W. Jones, G. C. Johnson, J. L. Cashdollar, D. R. Dahling, L. F. Villegas, G. S. Fout, D. E. Williams, and G. Sayler.** 2011. Viruses and Bacteria in Karst and Fractured Rock Aquifers in East Tennessee, USA. Ground Water. **49:**98-110. doi: 10.1111/j.1745-6584.2010.00698.x.

34. **Katz, B. G., and D. W. Griffin.** 2008. Using chemical and microbiological indicators to track the impacts from the land application of treated municipal wastewater and other sources on groundwater quality in a karstic springs basin. Environ. Geol. **55:**801-821. doi: 10.1007/s00254-007-1033-y.

35. **Kelly, W., S. Panno, K. Hackley, A. Martinsek, I. Krapac, C. Weibel, and E. Storment.** 2009. Bacteria Contamination of Groundwater in a Mixed Land-Use Karst Region. Water Quality, Exposure and Health. **1:**69-78. doi: 10.1007/s12403-009-0006-7.

36. **Klappenbach, J. A., P. R. Saxman, J. R. Cole, and T. M. Schmidt.** 2001. rrndb: the Ribosomal RNA Operon Copy Number Database. Nucleic Acids Res. **29:**181-184.

37. **Laroche, E., F. Petit, M. Fournier, and B. Pawlak.** 2010. Transport of antibiotic-resistant Escherichia coli in a public rural karst water supply. Journal of Hydrology. **392:**12-21. doi: 10.1016/j.jhydrol.2010.07.022.

38. **Layton, A., L. McKay, D. Williams, V. Garrett, R. Gentry, and G. Sayler.** 2006. Development of Bacteroides 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. **72:**4214-4224. doi: 10.1128/AEM.01036-05.

39. **Leclerc, H., D. A. A. Mossel, S. C. Edberg, and C. B. Struijk.** 2001. ADVANCES IN THE BACTERIOLOGY OF THE COLIFORM GROUP: Their Suitability as Markers of Microbial Water Safety. Annu. Rev. Microbiol. **55:**201.

40. **Leguyader, F., V. Apairemarchais, J. Brillet, and S. Billaudel.** 1993. Use of Genomic Probes to Detect Hepatitis-a Virus and Enterovirus Rnas in Wild Shellfish and Relationship of Viral Contamination to Bacterial-Contamination. Appl. Environ. Microbiol. **59:**3963-3968.

41. **Leguyader, F., E. Dubois, D. Menard, and M. Pommepuy.** 1994. Detection of Hepatitis-a Virus, Rotavirus, and Enterovirus in Naturally Contaminated Shellfish and Sediment by Reverse Transcription-Seminested Pcr. Appl. Environ. Microbiol. **60:**3665-3671.

42. **Ley, R. E., D. A. Peterson, and J. I. Gordon.** 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell. **124:**837-848. doi: 10.1016/j.cell.2006.02.017.

43. **Liu, W. T., T. L. Marsh, H. Cheng, and L. J. Forney.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. **63:**4516-4522.

44. **Mahler, B. J., J. -. Personné, G. F. Lods, and C. Drogue.** 2000. Transport of free and particulate-associated bacteria in karst. Journal of Hydrology. **238:**179-193. doi: 10.1016/S0022-1694(00)00324-3.

45. **Neill, H., M. Gutiérrez, and T. Aley.** 2004. Influences of agricultural practices on water quality of Tumbling Creek cave stream in Taney County, Missouri. Environmental Geology. **45:**550-559. doi: 10.1007/s00254-003-0910-2.

46. **Noble, R., and J. Fuhrman.** 2001. Enteroviruses detected by reverse transcriptase polymerase chain reaction from the coastal waters of Santa Monica Bay, California: low correlation to bacterial indicator levels. Hydrobiologia. **460:**175-184. doi: 10.1023/A:1013121416891.

47. **Noble, R., J. Griffith, A. Blackwood, J. Fuhrman, J. Gregory, X. Hernandez, X. Liang, A. Bera, and K. Schiff.** 2006. Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California. Appl. Environ. Microbiol. **72:**1604-1612. doi: 10.1128/AEM.72.2.1604-1612.2006.

48. **Okabe, S., N. Okayama, O. Savichtcheva, and T. Ito.** 2007. Quantification of host-specific Bacteroides-Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. Appl. Microbiol. Biotechnol. **74:**890-901. doi: 10.1007/s00253-006-0714-x.

49. **Panno, S. V., K. C. Hackley, H. H. Hwang, S. E. Greenberg, I. G. Krapac, S. Landsberger, and D. J. O'Kelly.** 2006. Characterization and Identification of Na-Cl Sources in Ground Water. Ground Water. **44:**176-187. doi: 10.1111/j.1745-6584.2005.00127.x.

50. **Panno, S. V., and W. R. Kelly.** 2004. Nitrate and herbicide loading in two groundwater basins of Illinois' sinkhole plain. Journal of Hydrology. **290:**229-242. doi: 10.1016/j.jhydrol.2003.12.017.

51. **Panno, S. V., W. R. Kelly, A. T. Martinsek, and K. C. Hackley.** 2006. Estimating Background and Threshold Nitrate Concentrations Using Probability Graphs. Ground Water. **44:**697-709. doi: 10.1111/j.1745-6584.2006.00240.x.

52. **Pasquarell, G. C., and D. G. Boyer.** Agricultural Impacts on Bacterial Water Quality in Karst Groundwater. J. Environ. Qual. **24:**959-969.

53. **Peeler, K. A., S. P. Opsahl, and J. P. Chanton.** 2006. Tracking anthropogenic inputs using caffeine, indicator bacteria, and nutrients in rural freshwater and urban marine systems. Environ. Sci. Technol. **40:**7616-7622. doi: 10.1021/es061213c.

54. **Peterson, E., R. Davis, and H. Orndorff.** 2000. 17 Beta-Estradiol as an Indicator of Animal Waste Contamination in Mantled Karst Aquifers. J. Environ. Qual. **29:**826-834.

55. **Pronk, M., N. Goldscheider, J. Zopfi, and F. Zwahlen.** 2009. Percolation and particle transport in the unsaturated zone of a karst aquifer. Ground Water. **47:**361-369.

56. **Raeymaekers, L.** 2000. Basic principles of quantitative PCR. Mol. Biotechnol. **15:**115-122; 122. doi: 10.1385/MB:15:2:115.

57. **Reed, T., A. Fryar, G. Brion, and J. Ward.** 2011. Differences in pathogen indicators between proximal urban and rural karst springs, Central Kentucky, USA. Environmental Earth Sciences. **64:**47-55. doi: 10.1007/s12665-010-0816-8.

58. **Reischer, G. H., J. M. Haider, R. Sommer, H. Stadler, K. M. Keiblinger, R. Hornek, W. Zerobin, R. L. Mach, and A. H. Farnleitner.** 2008. Quantitative microbial faecal source tracking with sampling guided by hydrological catchment dynamics. Environ. Microbiol. **10:**2598-2608. doi: 10.1111/j.1462-2920.2008.01682.x.

59. **Reischer, G. H., D. Kollanur, J. Vierheilig, C. Wehrspaun, R. L. Mach, R. Sommer, H. Stadler, and A. H. Farnleitner.** 2011. Hypothesis-Driven Approach for the Identification of Fecal Pollution Sources in Water Resources. Environ. Sci. Technol. **45:**4038-4045. doi: 10.1021/es103659s.

60. **Ritchey, S., and M. Coyne.** 2009. Applying MAR Analysis to Identify Human and Non-Human Fecal Sources in Small Kentucky Watersheds. Water, Air, & Soil Pollution. **196:**115-125. doi: 10.1007/s11270-008-9761-5.

61. **Roslev, P., L. Bjergbaek, and M. Hesselsoe.** 2004. Effect of oxygen on survival of faecal pollution indicators in drinking water. J. Appl. Microbiol. **96:**938-945. doi: 10.1111/j.1365-2672.2004.02209.x.

62. **Rutledge, R., and C. Cote.** 2003. Mathematics of quantitative kinetic PCR and the application of standard curves. Nucleic Acids Res. **31:**e93. doi: 10.1093/nar/gng093.

63. **Ryan, M., and J. Meiman.** 1996. An Examination of Short-Term Variations in Water Quality at a Karst Spring in Kentucky. Ground Water. **34:**23-30. doi: 10.1111/j.1745-6584.1996.tb01861.x.

64. **Santo Domingo, J. W., D. G. Bambic, T. A. Edge, and S. Wuertz.** 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. Water Res. **41:**3539-3552. doi: DOI: 10.1016/j.watres.2007.06.001.

65. **Sauer, E. P., J. L. VandeWalle, M. J. Bootsma, and S. L. McLellan.** 2011. Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water Res. **45:**4081-4091. doi: 10.1016/j.watres.2011.04.049.

66. **Savichtcheva, O., and S. Okabe.** 2009. Qualitative and quantitative estimation of host-specific fecal pollution using Bacteroides-Prevotella 16S rRNA genetic markers by T-RFLP and real-time PCR analyses. Water Sci. Technol. **59:**1831-1840. doi: 10.2166/wst.2009.213.

67. **Schmidt, T. M., E. F. DeLong, and N. R. Pace.** 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. **173:**4371-4378.

68. **Scott, T. M., J. B. Rose, T. M. Jenkins, S. R. Farrah, and J. Lukasik.** 2002. Microbial Source Tracking: Current Methodology and Future Directions. Appl. Environ. Microbiol. **68:**5796-5803. doi: 10.1128/AEM.68.12.5796-5803.2002.

69. **Shanks, O. C., E. Atikovic, A. D. Blackwood, J. Lu, R. T. Noble, J. S. Domingo, S. Seifring, M. Sivaganesan, and R. A. Haugland.** 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. **74:**745-752. doi: 10.1128/AEM.01843-07.

70. **Shanks, O. C., J. W. S. Domingo, J. Lu, C. A. Kelty, and J. E. Graham.** April 15, 2007. Identification of Bacterial DNA Markers for the Detection of Human Fecal Pollution in Water. Applied and Environmental Microbiology. **73:**2416-2422. doi: 10.1128/AEM.02474-06.

71. **Shanks, O. C., K. White, C. A. Kelty, M. Sivaganesan, J. Blannon, M. Meckes, M. Varma, and R. A. Haugland.** 2010. Performance of PCR-Based Assays Targeting Bacteroidales Genetic Markers of Human Fecal Pollution in Sewage and Fecal Samples. Environ. Sci. Technol. **44:**6281-6288. doi: 10.1021/es100311n.

72. **Simpson, J. M., J. W. Santo Domingo, and D. J. Reasoner.** 2002. Microbial source tracking: state of the science. Environ. Sci. Technol. **36:**5279-5288.

73. **Singh, K. P., A. Malik, V. K. Singh, D. Mohan, and S. Sinha.** 2005. Chemometric analysis of groundwater quality data of alluvial aquifer of Gangetic plain, North India. Anal. Chim. Acta. **550:**82-91. doi: 10.1016/j.aca.2005.06.056.

74. **Solo-Gabriele, H. M., M. A. Wolfert, T. R. Desmarais, and C. J. Palmer.** 2000. Sources of Escherichia coli in a Coastal Subtropical Environment. Applied and Environmental Microbiology. **66:**230-237. doi: 10.1128/AEM.66.1.230-237.2000.

75. **Stevanovic, Z. P.** 1988. The regime of the quality of karst ground waters in Eastern Serbia, Yugoslavia., p. Volume 2. *In* International Association of Hydrogeologists. Congress (ed.), Karst Hydrogeology and Karst Environment Protection: Proceedings of the 21st Congress of the International Association of Hydrogeologists, Guilin, China, 10-15 October 1988. IAHS Press. http://books.google.com/books?id=FLEnSAAACAAJ.

76. **Suau, A., R. Bonnet, M. Sutren, J. J. Godon, G. R. Gibson, M. D. Collins, and J. Dore.** 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. **65:**4799-4807.

77. **Terrado, M., D. Barcelo, and R. Tauler.** 2011. Chemometric analysis and mapping of environmental pollution sources in the ebro river basin, p. 331-372. *In* D. Barcelo and A. G. Kostianoy (eds.), The Handbook of Environmental Chemistry. Springer-Verlag, Berlin, Germany.

78. **USEPA.** 2006. Occurrence and monitoring document for Final Ground Water Rule.*In* Anonymous USEPA publication, no. 815-R-06-012 ed., vol. no. 815-R-06-012. Office of Water, U.S. Environmental Protection Agency, Washington DC.

79. **Walters, S. P., V. P. J. Gannon, and K. G. Field.** 2007. Detection of Bacteroidales fecal indicators and the zoonotic pathogens E-Coli O157 : H7, Salmonella, and Campylobacter in river water. Environ. Sci. Technol. **41:**1856-1862. doi: 10.1021/es0620989.

80. **Weiskel, P., B. Howes, and G. Heufelder.** 1996. Coliform contamination of a coastal embayment: Sources and transport pathways. Environ. Sci. Technol. **30:**1872-1881. doi: 10.1021/es950466v.

81. **WHO,** 2006. Guidelines for Drinking Water Quality. World Health Organization, Geneva.

82. **Wicks, C., C. Kelley, and E. Peterson.** 2004. Estrogen in a karstic aquifer. Ground Water. **42:**384-389. doi: 10.1111/j.1745-6584.2004.tb02686.x.

83. **Widory, D., E. Petelet-Giraud, P. Negrel, and B. Ladouche.** 2005. Tracking the sources of nitrate in groundwater using coupled nitrogen and boron isotopes: A synthesis. Environ. Sci. Technol. **39:**539-548. doi: 10.1021/es0493897.

84. **Wiggins, B. A., R. W. Andrews, R. A. Conway, C. L. Corr, E. J. Dobratz, D. P. Dougherty, J. R. Eppard, S. R. Knupp, M. C. Limjoco, J. M. Mettenburg, J. M. Rinehardt, J. Sonsino, R. L. Torrijos, and M. E. Zimmerman.** 1999. Use of Antibiotic Resistance Analysis To Identify Nonpoint Sources of Fecal Pollution. Applied and Environmental Microbiology. **65:**3483-3486.