2-28-2012

Bacterial Source Tracking of E. coli in a Constructed Wetland

Rose M. Martin

University of Connecticut, rose.m.martin.31@gmail.com

Recommended Citation

http://digitalcommons.uconn.edu/gs_theses/224

This work is brought to you for free and open access by the University of Connecticut Graduate School at DigitalCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.
Bacterial Source Tracking of \textit{E. coli} in a Constructed Wetland

Rose M. Martin

B.S., University of Rhode Island, 2009

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the degree of
Master of Science
at the
University of Connecticut
2012
Master of Science Thesis

Bacterial Source Tracking of *E. coli* in a Constructed Wetland

Presented by
Rose M. Martin, B.S.

Major Advisor
John C. Clausen

Associate Advisor
Daniel Gage

Associate Advisor
Donald Les

University of Connecticut

2012
Acknowledgements

Funding for this study came from a USDA NIFA Hatch grant and the Connecticut Association of Wetland Scientists (Michael Lefor Grant). The Department of Natural Resources and the Environment and the Center for Applied Genetics and Technology at the University of Connecticut provided facilities, materials and supplies for this research.

I would like to thank my advisor, Dr. Jack Clausen, for his contribution to this research and for the guidance he has provided me over the past two and a half years. I also sincerely thank Dr. Daniel Gage and Dr. Donald Les, my thesis committee members. Their input has been crucial to the development of this project.

Special thanks go to Ranyelle Craig, Amanda Dupuy, Keith Kozelka, and Emma White for their assistance with laboratory methods and to Dr. Linda Strausbaugh and Lu Li for assistance with and access to sequencing equipment.
# TABLE OF CONTENTS

**LIST OF TABLES**................................................................................................................. vi

**LIST OF FIGURES**................................................................................................................. vi

**ABSTRACT**............................................................................................................................... vii

**LITERATURE REVIEW**............................................................................................................. 1

**INTRODUCTION**....................................................................................................................... 1

**URBANIZATION AND CONTAMINATED RUNOFF**................................................................. 2

**Fecal Contamination of Water**............................................................................................... 2

**Indicator Bacteria**..................................................................................................................... 7

**Conducted Wetlands**............................................................................................................... 8

**Bacterial Source Tracking**..................................................................................................... 9

- Direct Sequence Observation for BST..................................................................................... 14
- Primer Selection......................................................................................................................... 16

**Fate and Transport of Fecal Bacteria**..................................................................................... 17

- Fate and Transport of Bacteria in Soil...................................................................................... 18
- Fate and Transport of Bacteria in Water and Sediments......................................................... 19
- Growth and Naturalization........................................................................................................ 20

**Conclusions**............................................................................................................................ 21

**References**............................................................................................................................... 23

**Bacterial Source Tracking of E. coli in a Conducted Wetland**............................................... 32

**Introduction**............................................................................................................................. 32

**Materials & Methods**.......................................................................................................... 34

- Study Area............................................................................................................................... 34
- Sample Collection.................................................................................................................... 34
- E. coli Isolation.......................................................................................................................... 35
- PCR Amplification..................................................................................................................... 35
- Dye-terminator sequencing...................................................................................................... 37
- Sequence Analysis.................................................................................................................... 37
LIST OF TABLES

Literature Review
Table 1. Summary of advantages and disadvantages of various BST methods............. 12

Bacterial source tracking of \textit{E. coli} in a constructed wetland
Table 1: Pigeon, cattle, inlet water and outlet water isolates corresponding to seven distinct 593 bp sequences of the E. coli \textit{mdh} gene. .......................................................... 43

LIST OF FIGURES

Bacterial source tracking of \textit{E. coli} in a constructed wetland
Figure 1. The sequence of the 940 bp \textit{mdh} gene. .............................................................. 44
Figure 2: Dendrogram showing 30 pigeon, cattle, inlet water and outlet water isolates. 45
Figure 3: Dendrogram showing all 30 pigeon, cow, inlet water and outlet water isolates along with GenBank \textit{mdh} sequences. ............................................................... 46
ABSTRACT

Bacterial source tracking was used to identify sources of fecal contamination in a constructed wetland. Nucleotide sequence differences in the *Escherichia coli* malate dehydrogenase (*mdh*) gene were used to distinguish between strains isolated from pigeon and cattle feces. Fourteen *E. coli* isolates were taken from cattle and pigeon fecal samples and sixteen *E. coli* isolates were taken from wetland water samples. A region of the *E. coli mdh* gene was amplified via PCR and sequenced. Twelve distinct sequences were obtained. Water samples indicated the presence of both pigeon and cattle fecal contamination in the wetland. Six sequences distinct from those isolated from pigeon and cattle feces were also present. Three of these sequences were pigeon-specific and two were cattle-specific. The presence of host-specific sequences indicates that sequence-based source tracking methods show promise for identifying fecal contamination.
LITERATURE REVIEW

INTRODUCTION

Microbial contamination can result from the introduction of fecal matter to ground and surface water (46). Such contamination presents a risk to public health, and is recognized by the U.S. Environmental Protection Agency as a leading cause of surface water impairment (124). For this reason, water must be monitored routinely for compliance with government standards in order to protect public health (45, 129).

Traditionally, the presence of commensal fecal coliforms or *Escherichia coli* has been used to indicate fecal contamination originating from human and animal waste. Although these indicator organisms are not pathogenic, they suggest the possible co-occurrence of pathogenic microbes, which also are part of the normal intestinal flora (4).

The simple presence of indicator bacteria signifies potential contamination but does not provide information about its source, which is crucial to implementing effective control strategies (17). To remedy this problem, a variety of bacterial source tracking (BST) techniques have been developed to associate fecal contamination with specific human and animal sources (124). Consequently, these techniques have the potential to aid in the control of contamination at its source.

The objective of this review is to examine the current state of knowledge pertaining to microbial contamination of water, the use of BST methods to identify contamination sources, and the fate and transport of fecal bacteria in soil and water. This review will identify areas requiring further research, with particular emphasis on direct sequencing BST methods.
URBANIZATION AND CONTAMINATED RUNOFF

A well-established correlation exists between urbanization and an increasing proportion of impervious surfaces such as roads and roofs that prevent water from infiltrating into soil (5). Impervious surfaces increase the velocity and volume of surface water and decrease infiltration (5, 81) and stormwater has been shown to mobilize microbial contaminants (81).

Urban stormwater runoff is understood to contain contaminants hazardous to public health including raw or poorly treated sewage (42, 81, 121, 127). High levels of microorganisms in urban stormwater can create health risks for those using the water as a drinking source (99) or for recreational purposes such as swimming (81). In a study reporting waterborne disease outbreaks in the United States between 1948 and 1994, a positive, significant (P = 0.002) association was found between extreme precipitation and disease outbreaks (30). Extreme precipitation events were defined as storms with an intensity of more than two in/d. The occurrence and extent of microbial contamination in stormwater runoff can be quantified using bacterial counts; however, only BST methods enable the origin, transport, and fate of the microorganisms to be determined.

Fecal contamination of water

Fecal contaminants in watersheds originate both from point and nonpoint sources (43). Nonpoint sources of fecal contamination include livestock operations (27, 37, 112, 119), wildlife (2, 110) and pets (131). Combined sewer overflows (CSO’s) (83), failed septic systems, leaking sewer lines, and cesspools (80) also can introduce fecal contamination into runoff.
Impacts of Fecal Contamination

The contamination of water by fecal bacteria is a widespread, persistent problem that adversely impacts public health as well as local and national economies (24, 47, 99). In 2004, the US Environmental Protection Agency assessed 16% of stream reaches in the United States and found 44% of those to be impaired, with pathogens cited as the leading cause of impairment (125). When 29% of estuaries and bays was assessed, 30% of those were found to be impaired. Pathogens, again, were cited as the cause (125).

The greatest risk from microbial contamination present in water involves the consumption of drinking water contaminated with fecal matter, which can contain a variety of pathogenic microbes (47, 80, 132). These pathogens include Norwalk virus, hepatitis A, hepatitis E, toxin-producing *E. coli*, and Shigella (80). In the United States, it is estimated that 96% of rural Americans rely on groundwater sources for their water supply (99). In developing regions and small communities in particular, where groundwater is the typical source of drinking water, microbial contamination of groundwater can have significant negative implications for public health (7, 99).

Swimming in fecally contaminated recreational water is consistently reported to cause illness as well (130). Settling of microbes associated with particles has been linked to accumulations of viable bacteria in sediment. When re-mobilized, these bacterial reservoirs can create health risks for users of recreational waters (63). In order to protect human health, beaches and shellfish harvesting areas may be shut down when high fecal coliform indicator bacteria counts are recorded for an area (16, 130). Although indicator bacteria are used widely to determine the sanitary quality of water and to determine beach closures, this approach has been criticized. Due to the lack of precision inherent in the
enumeration of coliform indicators, the use of fecal coliforms as indicators requires thoughtful experimental design to indicate the sanitary quality of water adequately (45). In a California study of one location contaminated predominantly by non-point sources, Colford et al. (26) found that traditional fecal indicators were not actually associated with health risks. They concluded that traditional indicators may not adequately assess the risk of illness when point sources of human contamination are not the predominant source of fecal contamination.

**Human Waste**

Human waste in urban areas generally is transported to a treatment plant either by combined sewer systems or by separate storm and sanitary sewers. Combined sewer systems are intended to transport stormwater and sanitary sewage from domestic, commercial, and industrial wastewater sources to a treatment facility via a single pipe. During periods of rainfall or snowmelt, when increased wastewater flows can exceed the capacity of the system or treatment facilities, the system is designed to overflow directly into surface water bodies. These events (known as combined sewer overflows or CSO’s), can be major contributors to water pollution (121). Because microbes in the water column tend to adsorb to particles, they fall out of solution and into sediment (21), where they may persist for longer durations than they would in the water column (14, 29, 68). One study of river bed sediments around a combined sewer outfall (63) indicated that bacteria persisted in sediments and could be resuspended, adversely impacting water quality even when no wastewater discharge was taking place.
In rural and suburban areas, on-site wastewater treatment systems such as septic systems are commonplace. Septic systems are intended to treat domestic wastewater, preventing microbiological and nutrient pollutants from contaminating surface and groundwater (3). However, on-site waste disposal systems sometimes fail due to clogging, overloading, poor separation from the water table, and low soil permeability (120). Arnade (6) found that wells close to septic tanks showed a greater likelihood than others of being contaminated with fecal coliforms during Florida’s wet season. Septic systems located on sand and gravel aquifers are particularly prone to contaminate groundwater used for human consumption (105).

Agriculture

Livestock waste is a major source of pathogens on agricultural land (43, 111, 117), with 133 million tons of manure on a dry weight basis produced annually in the United States (13). Fecal waste from confined animal feeding operations (CAFO’s) typically is stored as semiliquid slurry in lagoons, leading to conditions favoring enteric pathogen persistence (43). Even when cattle are grazed on pastureland, potential pathogens in excreta may survive for up to 56 days (111).

Bacteria present in manure can be mobilized by rainfall (53) and animal fecal waste can enter the environment via leakage from manure lagoons or during major precipitation events that result in either overflowing of the lagoon or runoff of waste applied recently to agricultural fields (13, 27). Although the nutrient content and availability of animal waste makes it valuable as a crop fertilizer, over-application of wastes or application to saturated soils can lead to the movement of contaminants into
receiving waters and aquifers (13). The use of treated wastewater for irrigation also can introduce fecal microbes into the environment (1).

**Wildlife**

In developed areas, rodents, pigeons, and waterfowl congregate in large numbers due to a quantity of available shelters and food (39, 56). The combination of large numbers of animals and a high proportion of impervious surfaces facilitates the introduction of fecal contaminants into water systems. Urban bird populations are implicated as a source of fecal pollution (83). Pigeons live in close proximity to humans and are common in both urban and rural areas where they harbor abundant levels of enteric bacteria (0.5 x 10^6 cells/g feces) (39). Pigeons are a known source of pathogenic *E. coli* O157:H7 (107), as well as *Clostridium perfringens*, *Listeria monocytogens*, *Salmonella enterica*, *Yersinia* spp., *Campylobacter jejuni*, *Campylobacter coli*, *Coxiella burnetti*, and *Chlamydia psittaci* (55). Although attempts to control pigeon populations by culling and avian birth control historically have been unsuccessful (55, 56), public education programs aimed at reducing pigeon feeding have proven effective (54).

**Domestic Animals**

Domestic animals also have been recognized as a source of fecal contamination (24, 83, 120). Dogs and cats living in watersheds with a high proportion of impervious surfaces present a risk to public health because they can carry zoonotic pathogens (97). The USEPA (120) advocates leash laws and regulations mandating pet waste pickup to
reduce the contamination of runoff by fecal material. In the United Kingdom, where animal waste is more seriously recognized as a threat to public health, dog wardens and feces collection bins have been put in place (39).

**Urban Runoff**

Runoff is a significant contributor of non-point source pollution in urban areas. In developed areas, higher densities of impervious surface in the form of roofs, roads, and other paved areas result in faster movement of water over the landscape and decreased infiltration (5). An increase in impervious surface leads proportionally to an increase in runoff (5). As water moves over the landscape, it mobilizes and carries with it microbes and particles associated with microbes. When stormwater enters receiving water bodies, microbial cells existing individually in the water tend to remain more mobile, while those associated with particles tend to settle out (21).

**INDICATOR BACTERIA**

Indicator bacteria should be nonpathogenic, rapidly detected, easily cultured, should possess survival characteristics similar to those of the pathogens of concern, and should be associated strongly with the presence of pathogens (109). Traditionally, the occurrence of fecal indicators such as *E. coli*, enterococci, total coliform, fecal coliform and *Clostridium perfringens* in receiving waters has been used to signify the potential presence of pathogenic microorganisms (4, 19) because these microorganisms are abundant in the guts of warm-blooded animals (70). Problems with using such fecal bacteria as indicators of pathogens can arise due to weak correlations between the
concentration of bacteria and the presence of pathogenic microorganisms in water (61) and the potential for indicator bacteria to replicate outside the host and become naturalized in the environment, particularly in tropical areas (15, 18, 64, 65). In addition, indicator bacteria counts do not specify the source of the contamination, which is needed to effectively direct remediation efforts.

**CONSTRUCTED WETLANDS**

Over the past two decades, the use of constructed wetlands to improve water quality has gained prominence (32, 62, 73) because they provide for low-cost, environmentally friendly wastewater treatment (34). Two main types of constructed wetlands typically are used: surface flow wetlands, in which wastewater flows horizontally over wetland sediments, and subsurface flow wetlands, in which wastewater flows vertically through permeable sediment and collects in drains (72, 129). Constructed wetlands have been used to treat roof runoff (60), treatment facility effluent (87), domestic wastewater (9), and dairy wastewater (62). These wetlands can effectively remove bacteria from wastewater (35, 51, 60). Reported coliform removal efficiencies typically exceed 90%, with significantly higher removal rates reported for vegetated systems (72). One study on the University of Connecticut campus found a greater than 98% reduction in *E. coli* abundance in roof runoff after constructed wetland treatment (60). Several studies support the observation that wetlands incorporating vegetation are better than non-vegetated wetlands at removing bacteria from wastewater (35, 73). Wetlands planted with vegetative polycultures perform consistently year-round, with removal rates for fecal coliforms ranging from 98% in the fall to 82% in the winter (73).
Emergent vegetation provides resistance to flow and thereby slows surface water flow (71). Because bacteria usually adsorb to small sediments that take longer to settle out of the water column (31), an increased retention time induced by emergent vegetation predictably would facilitate sedimentation and thus bacteria removal. Additionally, emergent plants reduce wind velocities near the water’s surface, which also reduces re-suspension (12). However, Vacca et al. (126) and Hier (60) found that the presence of plants had no effect on bacteria removal in one treatment wetland. Thus, the role of wetland vegetation in removal of fecal bacteria in wastewater requires further study.

**BACTERIAL SOURCE TRACKING**

Although direct pathogen monitoring in water provides information pertaining to potential health risks, the hundreds of different pathogens that are found in water contaminated with feces make it infeasible to routinely monitor water for all possible pathogens (4). Moreover, the indicator bacteria typically used to test for fecal contamination (4,109) do not indicate the origin of the contamination. As an alternative, BST approaches use genotypic and phenotypic differences in animal host intestinal bacteria to determine the source of fecal contamination in water (109). This technology assumes that members within a bacterial species have become adapted to a specific host (109). Therefore, if a match to a strain unique to a particular host is found in contaminated water, the source of fecal contamination may be inferred.

BST protocols (Table 1) involve both phenotypic and genotypic techniques. Phenotypic BST techniques are based on the observation of expressed physical traits (such as antibiotic resistance) typical of the strain being tracked. Examples include
F"RNA coliphage typing, antibiotic resistance analysis (ARA) and multiple antibiotic resistance analysis (MAR) (25, 91, 93).

However, most recent BST research employs genotypic techniques. For genotypic BST, a portion of a microbial genome is amplified via polymerase chain reaction (PCR) and then characterized genetically to identify patterns of variation uniquely associated with each strain. Characterization can incorporate various types of fragment analysis such as repetitive element PCR (rep-PCR), amplified fragment length polymorphism (AFLP), pulsed-field gel electrophoresis (PFGE), and ribotyping (110). Methods that rely on direct sequence observation may also be employed (66).

BST techniques can be subdivided further based on whether it is necessary to culture organisms from samples prior to analysis or if a reference library consisting of known isolates is required. Both library preparation and culturing can substantially increase the time and resources needed to complete an assay (44).

Culture independent methods of microbial community analysis examine signature biochemicals taken directly from environmental samples (10). Culture-independent BST methods include terminal restriction fragment length polymorphism (T-RFLP), denaturing gradient gel electrophoresis (DGGE), length heterogeneity PCR (LH-PCR), and the use of 16SrRNA gene clone libraries (88, 110) (Table 1).

Library independent BST methods (e.g. DGGE, LH-PCR, and host-specific marker gene PCR) are rapid and easy to perform (17, 110), and have successfully matched fecal water contamination to host sources (52) (Table 1). However, library independent BST methods require further method development (110, 114), produce non-quantitative results (4), and sometimes fail to differentiate among multiple host groups.
Most BST methods (e.g., T-RFLP, 16SrRNA clone libraries, ARA, MAR, rep-PCR, AFLP, PFGE, ribotyping, and sequencing) require a library (36, 66, 88, 91, 93, 110) (Table 1). Library dependent BST methods require a comparison of microbes isolated from water samples to serve as a reference for fecal samples obtained from the host organisms. These BST methods are more time consuming (44). Regional differences in intestinal host floras may require the isolation of separate libraries for each watershed studied (19, 75) and the clonal composition of isolates can differ temporally, e.g., during the transition from primary habitat (the host) to secondary habitat (the environment) (75).

Although BST techniques have been used successfully to identify sources of fecal bacteria (8, 58, 89, 90, 96), the need for further study and refinement of currently used BST techniques is emphasized (20, 88, 109, 110, 114, 115).
Table 1. Summary of various BST methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Library Dependent?</th>
<th>Culture Dependent?</th>
<th>Phenotypic or Genotypic?</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-RFLP (Terminal Restriction Fragment Length Polymorphism)</td>
<td>Yes</td>
<td>No</td>
<td>Genotypic</td>
<td>uses restriction enzymes coupled with PCR fragments tagged with a fluorophore are detected.</td>
<td>Marsh, 1999; Simpson, 2002; Blackwood et al., 2003; Meays et al., 2004; Hartmann and Widmer, 2008</td>
</tr>
<tr>
<td>DGGE (Denaturing Gradient Gel Electrophoresis)</td>
<td>No</td>
<td>No</td>
<td>Genotypic</td>
<td>electrophoretic analysis of PCR products based on melting properties of amplified DNA sequence; discriminates species</td>
<td>Simpson, 2002; Meays et al., 2004</td>
</tr>
<tr>
<td>LH-PCR (Length heterogeneity PCR)</td>
<td>No</td>
<td>No</td>
<td>Genotypic</td>
<td>separates pcr products generated for host-specific genetic markers by length</td>
<td>Simpson, 2002; Meays et al., 2004</td>
</tr>
<tr>
<td>F’RNA Coliphage Typing</td>
<td>No</td>
<td>Yes</td>
<td>Genotypic</td>
<td>variability of viruses infecting coliform bacteria fertility factors indicate animal or human fecal contamination</td>
<td>Cole et al., 2003</td>
</tr>
<tr>
<td>Host-specific marker gene PCR</td>
<td>No</td>
<td>Yes</td>
<td>Genotypic</td>
<td>discriminates E. coli genes from strains associated with host species</td>
<td>Call et al., 2007</td>
</tr>
<tr>
<td>16SrRNA gene clone libraries</td>
<td>Yes</td>
<td>No</td>
<td>Genotypic</td>
<td>combine LH-PCR and T-RFLP on fecal anaerobes, discriminates humans and cattle</td>
<td>Meays et al., 2004; McGarvey et al., 2004</td>
</tr>
<tr>
<td>Method</td>
<td>Library Dependent?</td>
<td>Culture Dependent?</td>
<td>Phenotypic or Genotypic?</td>
<td>Description</td>
<td>References</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>ARA, MAR (Antibiotic Resistance, Multiple Antibiotic Resistance)</td>
<td>Yes</td>
<td>Yes</td>
<td>Phenotypic</td>
<td>observes variability in resistance to antibiotics</td>
<td>Moore et al., 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Olivas et al., 2008</td>
</tr>
<tr>
<td>Rep-PCR (Repetitive Element PCR)</td>
<td>Yes</td>
<td>Yes</td>
<td>Genotypic</td>
<td>PCR amplification of palindromic DNA sequences with electrophoretic analysis</td>
<td>Dombek et al., 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simpson, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mohapatra et al., 2008a and 2008b</td>
</tr>
<tr>
<td>AFLP (Amplified Fragment Length Polymorphism)</td>
<td>Yes</td>
<td>Yes</td>
<td>Genotypic</td>
<td>DNA fingerprinting using rare and frequent cutting restriction enzymes</td>
<td>Simpson, 2002</td>
</tr>
<tr>
<td>PFGE (Pulsed Field Gel Electrophoresis)</td>
<td>Yes</td>
<td>Yes</td>
<td>Genotypic</td>
<td>DNA fingerprinting using rare cutting restriction enzymes paired with electrophoretic analysis</td>
<td>Scott et al., 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simpson, 2002</td>
</tr>
<tr>
<td>Ribotyping</td>
<td>Yes</td>
<td>Yes</td>
<td>Genotypic</td>
<td>southern hybridization of genomic DNA cut with restriction enzymes and probed with ribosomal sequences</td>
<td>Scott et al., 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simpson, 2002</td>
</tr>
<tr>
<td>Sequencing</td>
<td>Yes</td>
<td>Yes</td>
<td>Genotypic</td>
<td>PCR amplification of sample DNA, dye-terminator sequencing to determine order of nucleotides</td>
<td>Ivanetich et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Olive and Bean, 1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ram et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ram et al., 2007</td>
</tr>
</tbody>
</table>
Direct Sequence Observation for BST

In contrast to genotypic BST methods which use fragment analysis, genotypic BST methods that use DNA sequencing allow for direct sequence observation, which enables better reproducibility and precision, as well as the opportunity to survey a greater extent of the genome in order to detect differences (102). Direct sequencing BST methods involve PCR amplification of a selected DNA region specific to the organism being studied, followed by sequencing of the PCR products. Suitable primers must first be developed to suitably target the region of interest. Like other genotypic BST techniques, direct sequencing methods assume that bacterial strains, after being acted on by the selective pressure of their host environment, are unique to that particular host (109). Therefore, if an exact match can be found in a water sample, the source of contamination can be identified.

Despite the high degree of accuracy and reproducibility in DNA sequencing achieved as a result of recent technological advances, literature pertaining to sequence-based BST methods is scarce. Olive and Bean (1999) concluded that despite their potential utility for BST, sequencing methods have some limitations. They pointed out that few regions selected in previous studies had met criteria essential for strain differentiation: i.e., they comprise a variable region flanked by highly conserved regions; the level of variability detected sufficiently discerned different strains of a species; and the region was not susceptible to horizontal gene transfer. However, subsequent studies have identified and developed effective BST methods by targeting specific regions of the *E. coli* genome. Ram et al. (101) sequenced the β-glucuronidase (*uidA*) gene of *E. coli* at different sites to quantify genetic differences between populations. They also were able
to assign different fecal isolates to possible host sources (including humans, farm animals, birds and pets) using interpopulational allelic variation in *uidA*. By excluding alleles common to all hosts, they achieved a 75% level of correct fecal sample assignment to host organisms. Ram et al. (102) subsequently used the *uidA* gene to evaluate humans, pets and urban wildlife as sources of contamination to storm sewers, this time achieving a 65% level of correct classification.

Use of the *gusA* gene in BST methods is complicated by uncertainty regarding the extent of horizontal gene transfer near the *gusR, A, B, C* operon. In order for a sequence-based BST method to be effective, genetic diversity at the gene locus being sequenced must be consistent with the evolutionary history of the organism being studied (66).

Following criteria established by Olive and Bean (94), Ivanetich et al. (66) evaluated several regions of the *E. coli* genome for the development of a BST assay. They selected the malate dehydrogenase gene (*mdh*), as it satisfies all criteria presented by Olive and Bean (94) and because there is no evidence that more than a single copy of the gene exists in the *E. coli* genome. They also emphasized other advantages of using the *mdh* gene for sequence-based BST methods including their observation that no *E. coli mdh* alleles were found frequently either in the host species or in most environmental samples and also because little variation was observed between strains sampled from hosts in different geographic locations (66). They designed primers to amplify an 825 base pair portion of the 936 base pair *mdh* gene and focused specifically on a 150 base pair region of the gene determined to have the highest level of polymorphism. However, by shortening the target sequence, some polymorphisms were eliminated making differentiation between some hosts' strains impossible. A total of ten polymorphic sites
was identified, and these were used to distinguish between host species, which included dog, deer, seagull, horse and human. A dendrogram of sequence variants clustered horse, seagull and dog isolates into identifiable groups, though human isolates were highly variable. In a blind analysis of environmental fecal isolates, 95% of dog, deer, seagull and horse isolate sequences were matched successfully with host reference sample sequences using this approach (66).

**Primer Selection**

Genomic BST methods require PCR primers for amplification. Primers are short, single strands of nucleic acid that contain sequences complementary to target DNA regions and serve as starting points for DNA synthesis. Pairs of primers are designed to work in the forward and reverse directions on a strand of DNA, allowing for amplification of a specific region (106). Selecting or designing the correct pair of primers is crucial to the success of PCR, as poor primer selection can lead to poor PCR yield (106). Primers must be designed for specificity, and will only work to produce good PCR results when each primer anneals stably to the target sequence of DNA in the desired organism (106). In order to develop effective primers, a target region is selected on the bacterial chromosome and specific sequences are designed to selectively amplify a region between annealing sites of the two oligonucleotides (105). Requirements for successful primers include proper base composition, with guanine and cytosine content between 40% and 60% and an even distribution of adenine, thymine, cytosine and guanine; an 18 to 25 nucleotide long region of the primer complementary to the template; the sizes of the two primers not differing by more than 3 bp’s; the absence of inverted
repeat or self complimentary sequences to avoid secondary structures; and proper melting temperature (106). Many web-based primer design tools have been created to assist with the technical aspects of primer design (23, 78, 100, 128).

Boyd et al. (11) designed primers to amplify an 864 base pair segment of the malate dehydrogenase (mdh) gene from E. coli. These primers were used initially by Ivanetich et al. (66) but later were modified to amplify a region of higher polymorphism. Chen and Griffiths (22) developed primers targeting an 884 bp sequence of E. coli DNA adjacent to the universal stress protein gene (uspA). This region specifically differentiated E. coli from other Gram-negative bacteria, and a test assay was highly specific for E. coli (22). Turner (118) used this primer set in a study employing T-RFLP to distinguish between strains of E. coli derived from feral pigeons and cattle, but was unable to isolate unique T-RF’s from cattle-sourced E. coli. Ram et al. (101, 102) designed primer sets to target the E. coli β-glucuronidase gene for the purpose of sequence-based source tracking.

FATE AND TRANSPORT OF FECAL BACTERIA

Fecal microbes can survive in soil, water, and sediment for varying durations depending on the suitability of local conditions including moisture, temperature, pH, nutrient availability, and competing microorganisms (67). They can even form naturalized populations (64, 65).
Fate and Transport of Bacteria in Soil

Microbial concentrations in soil or percolated water are dependent upon the survival of organisms in the soil and the soil’s retention ability (49). Fecal bacteria can survive for long periods in soil, with one study reporting a duration of 68-80 days (28). Bacteria survive longer in organic soils than in sandy soils (116) because organic matter increases nutrient retention, provides bacteria with a carbon source, and retains moisture (43). A review by England (41) indicated that clays may enhance bacteria survival in soil by providing pore spaces small enough to protect bacterial cells from predation. Generally, fecal bacteria survive longer in soil at temperatures between 4° and 10°C than they do at warmer temperatures (28, 74). Shorter survival times associated with warmer temperatures may be due to a combination of thermal effects and increased activity of potentially competitive native soil flora (28). Freezing is lethal to bacterial populations, which can be reduced by up to 95% after several freeze-thaw cycles (74). Enteric bacteria exhibited lower survival rates in low-pH soils than in those of more moderate pH (49). Van Donsel et al. (127) showed that the time required to remove 90% of bacteria introduced into soil increased from as little as 2.7 days in the summer to as long as 20.1 days during the winter. Bacterial species also display different survivorship in different soil types. In a study of *E. coli* and *Enterococcus* spp. derived from pig slurry, *E. coli* survived best in a sandy soil, whereas *Enterococcus* spp. survived best in loamy soil (28). Gannon et al. (48) observed that transport of different types of soil bacteria was influenced by the electrostatic charge of their cell surface, hydrophobicity, cell size, and the presence of capsules and flagella.
Microbes that enter the soil move downward with infiltrating water and sediment (103). Soil provides filtering and adsorption sites for microorganisms, and microbial movement through a saturated soil depends on soil properties, including sand, silt, clay, and organic matter content (43). The majority of such filtering seems to occur at the soil surface, through straining, sedimentation and adsorption (49). McCoy and Hagedorn (85) found that bacterial populations are reduced rapidly as they enter the soil system. However, once the organisms reach a highly conductive soil zone, long distances are required for further population reduction. Microbial organisms are transported more rapidly when soil is saturated (30, 57). Macropores such as earthworm burrows may further facilitate the movement of microorganisms through the soil and potentially even into groundwater (69).

Fate and Transport of Bacteria in Water and Sediments

The availability of pathogens transported in runoff is influenced by the die-off rate of any enteric bacteria in soil and to the extent of waste applied to it (103). Once in the water column, microbes associate with particles, clump together to form aggregates, or are suspended individually (21, 108). Microbes that exist individually in the water column or are adsorbed to less dense particles will remain more mobile, whereas microbes that associate with dense inorganic particles generally settle out of the water column much earlier (21).

Bacteria survive longer in marine and freshwater sediments than in the overlying water (14, 29, 40, 68, 108). This phenomenon is attributed to the greater availability of nutrients available in sediments (29). Possible re-suspension of microorganisms may
occur during natural turbulence or disturbance by humans (29, 108). Sediment properties also influence the distribution of bacteria. Bacterial concentrations are higher in sediments comprised of fine particles than in those comprised of coarse particles (14, 29, 31). Individual bacterial species and strains exhibit different preferences for sediment size classes (21, 68, 98). Jeng et al (68) found that enterococci attached preferentially to particles with a diameter of 10–30 μm, but that fecal coliforms and *E. coli* were not as selective. Bacterial survival in sediment also increases with increased availability of organic material (31, 76), which is higher in finer sediments.

**Growth and Naturalization**

In addition to surviving in soil, water, and sediments, enteric bacteria are capable of growing and even of forming naturalized populations in these environments (38, 64, 65, 82). Several studies document growth of fecal bacteria in soil and sediment. Fecal coliforms were able to multiply in storm drain sediments (82) as well as in sewage sludge applied to a forest clearcut (38). Ishii et al. (64, 65) reported naturalized populations of *E. coli* in temperate soils.

**Fate of Bacteria in Wastewater Treatment Wetlands**

Many treatment wetlands use the process of bacterial adsorption to sand, silt and clay particles which then undergo sedimentation to remove fecal bacteria from wastewater (32, 113). Though numerous studies detail removal rates of fecal bacteria from wastewater using constructed wetlands, less research details the fate of bacteria in these wetlands.
Microorganisms persist in wetland sediments, acting as reservoirs of living bacteria (32, 113). Because bacteria adsorb preferentially to smaller particles (31), treatment systems must effectively facilitate the settling out of fine particles (32, 113). Besides sedimentation, other processes that remove bacteria in constructed wetlands include filtration through the substrate and associated biofilm, aggregation, oxidation, exposure to biocides, predation, attacks from lytic bacteria and viruses, antibiosis, naturally occurring die-off, and competition for limiting nutrients and trace elements (50, 51). Predation by nematodes, rotifers, and protozoa also is thought to be important to the removal of bacteria from wastewaters treated by constructed wetlands (51).

CONCLUSIONS

Microbial contamination of water threatens public health as well as local and national economies. Pigeons carry a wide array of pathogenic enteric microorganisms, but no BST studies have focused on pigeons as potential sources of fecal contamination. Although bacterial counts can provide information regarding the presence and concentrations of indicator bacteria, they offer little information about the source of the microorganisms. Though numerous studies have indicated various promising BST methods, there is a need to refine currently used techniques for increased reproducibility and accuracy. Sequencing is a promising BST technique due to inherently high levels of precision and reproducibility. However, the need for culturing and library development is a drawback. Culturing increases the amount of time and technical skill needed to perform these assays. Library development can be problematic due to regional differences in host intestinal flora and changes in isolate composition during transition.
between hosts and the environment. Genome regions must meet certain criteria in order to be suitable for sequence-based BST applications and such regions are uncommon in bacteria. The *E. coli* malate dehydrogenase gene (*mdh*) meets all necessary criteria for sequencing BST, but only one study has focused on this approach.
REFERENCES


Carolina: septic tanks vs. storm-water runoff as fecal coliform sources. J. Coastal Res. 22(2):319-327.


BACTERIAL SOURCE TRACKING OF E. COLI IN A CONSTRUCTED WETLAND

INTRODUCTION

Water contamination resulting from fecal inputs is a widespread problem, with pathogen contamination a leading cause of surface water impairment (33) that impacts public health and local and national economies (3, 5, 11, 26). Fecal contamination can result from many sources including human waste, agriculture and livestock operations, indigenous wildlife, domestic animals, and urban runoff (15, 9, 12, 7, 20). Runoff effectively transports fecal contaminants, and a well-known correlation exists between urbanization and increased runoff (22). Consequently, greater contamination of water by pathogens is expected as urbanization progresses, due to the increased volume of runoff and its load of fecal contaminants.

Because their droppings contain high levels of enteric bacteria (8), pigeons have emerged as a public health concern due to their abundance in urban and rural areas and their frequent close proximity to humans. They are a known source of *Clostridium perfringens*, *Listeria monocytogenes*, *Salmonella enterica*, *Yersinia* spp., *Campylobacter jejuni*, *Campylobacter coli*, *Coxiella burnetii*, and *Chlamydia psittaci* (13, 14), and pathogenic *E. coli* O157:H7 (30). Despite the fact that pigeons carry such an array of pathogenic microbes, no studies have focused on them as a potential source of fecal contamination.

Traditionally, indicators such as *E. coli*, enterococci, and *Clostridium perfringens* have been used to signify the potential presence of pathogenic microorganisms (1, 6). Because these microorganisms co-occur abundantly in the guts of warm-blooded animals,
their presence in receiving waters indicates the existence of fecal contamination and the presumed presence of pathogenic microorganisms (19). However, simple quantification of indicator bacteria does not enable the origin of a contaminant to be determined. Source identification is necessary for the implementation of management strategies to curtail contamination and thus lead to improved water quality.

Bacterial source tracking (BST) uses genotypic and phenotypic differences in intestinal bacterial communities to determine the source of fecal contamination in water. BST assumes that strains of bacteria become adapted, and thus unique to, particular animal hosts. Therefore, bacterial isolates from the environment may be matched to animal sources (6, 24).

Techniques that use gene sequence data to directly analyze specific regions of the \textit{E. coli} genome potentially are highly effective in matching environmental samples to their sources (18, 27, 28). Specifically, the malate dehydrogenase (\textit{mdh}) gene is a particularly promising target region for such applications because it satisfies several essential criteria: it is not vulnerable to horizontal gene transfer, it consists of a highly variable region flanked by highly conserved regions, it is relatively short in length but contains sufficient allelic polymorphism to differentiate between strains (25), and occurs as a single copy in the \textit{E. coli} genome (18). Only one previous study has tracked sources of fecal contamination using \textit{E. coli} \textit{mdh} gene polymorphisms (18).

The objective of this study was to determine whether occurrences of \textit{E. coli} detected in roof runoff and a constructed wetland originated from cattle or pigeon sources. Nucleotide sequence differences in the \textit{E. coli} malate dehydrogenase (\textit{mdh}) gene were used to genotype strains derived from these two hosts, and these differences
were used to screen the microbial composition of water samples derived from the runoff and wetland sources.

**MATERIALS & METHODS**

*Study Area*

The study site was a 416 m² constructed wetland that received roof runoff from a 602.6 m² portion of the roof of the Kellog Dairy Center at the University of Connecticut in Storrs, CT. Until 2004, the wetland received milkhouse waste from the Center. A flock of over 100 feral pigeons has been observed roosting on the roof. Water is directed from the roof to a monitoring station and from there to three cells vegetated primarily by emergent wetland species. The water leaves the wetland through an effluent monitoring station.

*Sample Collection*

Fresh fecal reference samples were collected from cattle in the Kellog Dairy Center. Pigeon reference samples were obtained by placing aluminum foil under groups of roosting pigeons at a nearby barn silo. Care was taken to procure fecal samples from individual animals; no mixed samples were taken. Fecal material was collected with sterile polyester swabs and placed into TWIRL’EM® sterile polystyrene sampling bags (Thermo Fisher Scientific, Waltham, MA). Samples were placed on ice for transport to the lab.
Water samples were collected from the inlet and outlet pipes of the constructed wetland using TWIRL’EM® bags during rainstorms when flow into and out of the wetland occurred. Samples were placed on ice for transport to the lab.

E. coli Isolation

*Escherichia coli* was isolated from feces and water samples using two culturing steps. Serial dilutions of fecal material in sterile deionized water were prepared and plated on MacConkey agar to select for gram-negative, lactose-metabolizing bacteria (10). Cultures were incubated overnight at 37ºC. Well-isolated suspected *E. coli* colonies (10) from each plate were re-streaked onto MacConkey agar for isolation and again incubated overnight. Isolate colonies were inoculated into 1 mL of Colilert-18 medium (Idexx Laboratories Inc., Westbrook, ME) according to the manufacturer’s protocol to confirm the presence of *E. coli* (27). Cultures that tested positive (fluorescent) for *E. coli* were stored for downstream use at -20ºC.

Petrifilm™ plates for *E. coli* and coliforms (3M™, St. Paul, MN) were used according to the manufacturer’s protocol to isolate *E. coli* from water samples. Suspected *E. coli* colonies from each plate were streaked onto MacConkey agar and treated as above. If colonies were too numerous on the Petrifilm™, a sterile pipette tip was touched to the plate and streaked onto MacConkey agar as above.

**PCR Amplification**

Two primers first developed by Ivanetich et al. (16) (5’-TGAAAGTCGCAGTCTCGG-3’; 5’-GGTAAAAACGCGGTGGA-3’) (Figure 1)
were used to amplify an 825 base pair region of the malate dehydrogenase gene (mdh) by PCR (18). Each 30 μL reaction contained 1 μL of positive Colilert-18 culture, 0.2 μL Bio-Rad iTaq DNA polymerase, 2 μL 2 mM deoxynucleoside triphosphates (dNTP’s), 1 μL 5 μM forward and reverse primers, 6 μL 5X cresol red, 3 μL 25 mM MgCl₂, 3 μL 10X PCR buffer, and 8.8 μL sterile deionized water (32). A master mix containing all components of the reaction except the template DNA was prepared, vortexed briefly, and aliquoted into 0.2 mL tubes where template DNA was added. Positive controls contained *E. coli* XL1 Blue genomic DNA shown previously to produce a product with this procedure and negative controls contained sterile water in place of culture.

Reactions were carried out in a MyCycler thermal cycler (Bio-Rad Laboratories Inc. Hercules, CA). Samples were denatured for 5 min. at 94°C, and then exposed to 32 amplification cycles consisting of: 2 min at 94°C (denaturation), 1 min at 60°C (annealing), and 1 min at 72°C (extension). After a final extension step of 25 min. at 72°C, reactions were held at 10°C until retrieved.

To confirm successful amplification of the 825 base pair region, PCR products were electrophoresed for 30 minutes at 140 V alongside 1 Kb+ DNA ladder on a 1% agarose gel in Tris/Borate/EDTA buffer containing ethidium bromide. Gels were imaged with a GelDoc XR + UV camera (Bio-Rad Laboratories Inc., Hercules, CA). PCR products were purified for downstream application using a QiaQuick PCR purification kit (Qiagen Inc., Valencia, CA) according to the manufacturer’s protocol and were processed either immediately or stored at -20°C.
Dye-terminator sequencing

An 825 bp region of the \textit{mdh} gene was sequenced using the same primers used for PCR. Each 10 \textmu L reaction contained 1 \textmu L QiaQuick-purified PCR product, 0.5 \textmu L BigDye® Terminator, 1.75 \textmu L BigDye® buffer, 1 \textmu L 5 \textmu M primer, and 5.75 \textmu L sterile deionized water. Master mixes for each primer were prepared separately and aliquots were placed into 0.2 mL tubes to which purified PCR product was added. Positive controls contained purified PCR product previously shown to produce clean sequences. Negative controls contained sterile water in place of DNA. Reactions were initially denatured for 1 minute at 96°C and then exposed to 25 amplification cycles consisting of: 10 seconds at 96°C (denaturation), 5 seconds at 50°C (annealing), and 4 minutes at 60°C (extension). Reactions were held at 4°C until retrieved.

Sequencing reaction products were precipitated using a Qiagen Dye-Ex Kit (Qiagen Inc., Valencia, CA) following the manufacturer’s protocol. Once precipitated, PCR products were resuspended in highly-deionized formamide (Hi-Di\textsuperscript{TM}, Applied Biosystems, Carlsbad, CA), vortexed to mix, and loaded into sequencing plates. Sequencing was performed on an ABI 3130 Genetic Analyzer (Applied Biosystems, Carlsbad, CA).

Sequence Analysis

Sequences were inspected and reverse compliments of reverse primer sequences were generated in FinchTV (Perkin Elmer, Waltham, MA). Sequences were exported in FASTA format and then trimmed and assembled in CLC Main Workbench (CLC bio, Aarhus, Denmark). In the event of poor sequence quality, sequencing was repeated.
Poor sequence quality included: sequence trace chromatograms < approximately 500 bp; low signal strength; or “messy” sequence peaks.

Sequences for all 30 pigeon, cattle, inlet water and outlet water samples were aligned using the CLC Main Workbench multiple alignment function. From this alignment, a dendrogram was constructed using the neighbor-joining algorithm (29). Redundant, or identical, sequences were identified. Bootstrapping, a statistical technique that uses resampling to evaluate dendrogram reliability, was performed with 1000 replicates. The Basic Local Alignment Search Tool (BLAST) algorithm (2) was used to compare unique sequences identified to GenBank *E. coli mdh* sequences.

### RESULTS & DISCUSSION

#### Sample Collection & *E. coli* Isolation

Nine isolates were taken from two cattle fecal samples. Of these, Colilert-18 tests confirmed that seven of the isolates (78%) were *E. coli*. Nineteen isolates were obtained from two pigeon fecal samples, with seven of the isolates (37%) testing positive for *E. coli*. All of the eight inlet water sample isolates and the eight outlet water sample isolates tested positively for *E. coli*.

#### Sequence Comparison and Analysis

Trimming to remove poor quality sequence ends for all 30 pigeon, cattle, inlet water and outlet water samples yielded a 593 bp region when sequences were aligned. A reverse primer closer to the end of the *mdh* gene was designed in CLC Workbench in an attempt to improve gene coverage. However, this primer generated secondary products.
during PCR and so the original published primer was used (18). Twelve distinct sequences (Table 1) were identified and aligned (Appendix I). Among these sequences, 21/593 sites (3.5%) were polymorphic, with 17 substitutions (80.9%) representing transitions. A cluster analysis (Figure 2) grouped isolates with identical sequences. Bootstrap values, shown on dendrograms, indicate how strongly branching structure is supported by the data.

Sequence 1 was found most often (20%) among the 30 isolates evaluated (Figure 2, Table 1). Cattle isolates included four of the 12 sequences, as did pigeon isolates. Two of the sequences were unique to inlet water samples and three were unique to outlet water samples. Three sequences were obtained from pigeon (and not cow) isolates and two sequences from cow (but not pigeon) isolates. Two sequences were retrieved from both cow and pigeon isolates.

In order to be useful for source tracking applications, *E. coli* sequences must be host-specific. In this investigation, sequences detected only either in pigeon (and water) isolates (Sequences 1, 5, and 7) potentially are indicative of pigeon fecal contamination. Similarly, sequences detected only in cattle (and water) isolates (Sequences 2 and 10) are potential indicators of cattle fecal contamination. In both cases, a much greater sampling of host and water samples would be required before the extent of host specificity could be established with certainty. The sequences found in both pigeon and cow isolates (Sequences 6 and 9) would not be useful for source tracking due to their lack of host specificity.

The sequences found exclusively in water (inlet or outlet) samples (Sequences 3, 4, 8, 11 and 12) could represent fecal contamination from different host sources. They
also could be sequences of strains that were present in the hosts sampled but not isolated, or strains from earlier effluent inputs that became naturalized in the constructed wetland (4, 16, 17). The latter scenario seems unlikely due to the occasional cycles of drying of the wetland during the summer months and freezing during the winter (21). No sequences were found in both inlet water and outlet water isolates. Sequences from strains isolated from inlet samples but not outlet samples suggest that those strains could have been removed by the constructed wetland. Sequences from strains isolated from outlet but not inlet samples possibly implicate the wetland as a source of *E. coli*. However, it is also possible that the small number of samples and isolates taken did not fully characterize the extent of diversity for all strains present in both inlet and outlet water samples.

The presence of sequences isolated from pigeon feces in inlet and outlet water samples (Sequences 1 and 5) was expected, as large numbers of feral pigeons defecate on the roof from which runoff is directed into the wetland. Sequence 10 putatively indicates cattle fecal contamination in roof runoff due to its exclusive presence in cattle and inlet water isolates, at least among the samples evaluated. Its presence may be due to several factors. First, the *E. coli* strain from which the sequence was obtained could have become naturalized after its introduction during dairy facility effluent inputs prior to 2004, though summer drying and winter freezing cycles make it unlikely that naturalized *E. coli* populations would survive. Second, *E. coli* from cattle feces may have become airborne (31, 34, 35) and settled on the roof of the dairy facility. From there, rain could carry *E. coli* into the constructed wetland. Third, it is possible that this sequence is also
found in pigeon strains but was not isolated from pigeon feces among the samples processed during this study.

A BLAST search of GenBank using the 12 unique sequences identified in this study mainly matched \textit{mdh} sequence strains isolated from mammalian sources. However, this result could be attributable to the presence of fewer \textit{mdh} sequences in the GenBank database. The BLAST matches represented a variety of sources from different locations, suggesting that sequences isolated in this study represented strains that are widely distributed. A dendrogram, which compares the \textit{mdh} sequences from GenBank strains to sequences obtained in this study (Figure 3) illustrates these relationships. GenBank matches for which a source was not specified are not shown.

The use of more widely separated primer sites could allow for analysis of a larger region of the \textit{mdh} gene. Further sampling of pigeon, cattle, inlet water and outlet water would provide greater assurance that host-specific sequences exist. Expanding host sampling to include other species could disclose the source of the strains observed in water but not in pigeon or cow isolates. Dogs, horses, Canada geese, starlings, sparrows, and groundhogs all have been observed near the study site and could contribute to fecal contamination in the constructed wetland.

\textbf{CONCLUSIONS}

Sequence variation in a 593 bp region of the \textit{mdh} gene indicated different \textit{E. coli} strains in cattle and pigeon which potentially identify host-specific strains within this dataset. The distribution of sequences evaluated in this study pointed to pigeon fecal contamination in wetland inlet and outlet samples. Sequences indicating cattle fecal
contamination matched only inlet samples. Wetland inlet and outlet samples indicated the presence of *E. coli* strains not isolated from pigeons or cattle in this study. No sequences were found in both wetland inlet and outlet water samples.

Though preliminary, this study demonstrates the potential for source tracking methods using nucleotide sequence differences in the *mdh* gene to identify host-specific *E. coli* strain sequences.
Table 1. Pigeon, cattle, inlet water and outlet water isolates corresponding to seven distinct 593 bp sequences of the *E. coli* *mdh* gene.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Isolates</th>
<th>Sequence</th>
<th>Isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pigeon 4</td>
<td>7</td>
<td>Pigeon 7</td>
</tr>
<tr>
<td></td>
<td>Pigeon 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outlet Water 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outlet Water 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outlet Water 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outlet Water 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cattle 6</td>
<td>8</td>
<td>Inlet Water 3</td>
</tr>
<tr>
<td>3</td>
<td>Outlet Water 2</td>
<td>9</td>
<td>Pigeon 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cattle 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inlet Water 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inlet Water 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inlet Water 7</td>
</tr>
<tr>
<td>4</td>
<td>Inlet Water 5</td>
<td>10</td>
<td>Cattle 2</td>
</tr>
<tr>
<td></td>
<td>Inlet Water 6</td>
<td></td>
<td>Cattle 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cattle 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cattle 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inlet Water 8</td>
</tr>
<tr>
<td>5</td>
<td>Pigeon 3</td>
<td>11</td>
<td>Outlet Water 1</td>
</tr>
<tr>
<td></td>
<td>Pigeon 6</td>
<td></td>
<td>Outlet Water 3</td>
</tr>
<tr>
<td></td>
<td>Inlet Water 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Cattle 4</td>
<td>12</td>
<td>Outlet Water 4</td>
</tr>
<tr>
<td></td>
<td>Pigeon 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. The sequence of the 940 bp $mdh$ gene. Forward and reverse primers used for PCR amplification and dye-terminator sequencing are shown.
Figure 2. Dendrogram showing 30 pigeon, cattle, inlet water and outlet water isolates. The Neighbor-joining algorithm was used for dendrogram construction.
Figure 3. Dendrogram showing all 30 pigeon, cow, inlet water and outlet water isolates along with GenBank *mdh* sequences. The Neighbor-joining algorithm was used for dendrogram construction.
REFERENCES


APPENDIX I

Alignment of 593 bp region of the *mdh* gene of unique sequences isolated from pigeon, cow, inlet and outlet water isolates. Polymorphisms are indicated (1 of 3).
Alignment of 593 bp region of the *mdh* gene of unique sequences isolated from pigeon, cow, inlet and outlet water isolates.

Polymorphisms are indicated (2 of 3).
Alignment of 593 bp region of the *mdh* gene of unique sequences isolated from pigeon, cow, inlet and outlet water isolates. Polymorphisms are indicated (3 of 3).

<table>
<thead>
<tr>
<th></th>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
<th>Sequence 5</th>
<th>Sequence 6</th>
<th>Sequence 7</th>
<th>Sequence 8</th>
<th>Sequence 9</th>
<th>Sequence 10</th>
<th>Sequence 11</th>
<th>Sequence 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td>TGGCCAGGCAGCTGCACGTCTTGGTCTGTCCTCTGGTCCG 593</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX II

Colilert-18 *E. coli* Presence / Absence Test

□ Prepare Colilert-18 medium. Dissolve one packet in 100 mL sterile water.

□ Inoculate well-isolated pink colony from MacConkey agar into 1 mL Colilert-18 medium

□ If the sample is not already at 33–38°C, place the sample in a 35°C waterbath for 20 minutes or a 44.5°C waterbath for 7–10 minutes

*(This prewarming time is part of (not in addition to) the 18-hour incubation period)*

□ Incubate 18 hours at 35.5°C

□ Fluorescence indicates presence of *E. coli*

Adapted from kit insert, Colilert-18 medium (Idexx Laboratories Inc., Westbrook, ME)
APPENDIX III

PCR of E. coli DNA extracted from Feces

Samples:_____________________________ Extraction method:______________
Extraction Date:_____________ Pages describing extraction:____________
Taq, MgCl2, and PCR buffer used: ABI □/Bio-Rad □
Notes:

<table>
<thead>
<tr>
<th>30 μl PCR:</th>
<th>X Master Mix (μl):</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA ___ ng/μL (___μl)</td>
<td>NO DNA in Master Mix</td>
</tr>
<tr>
<td>Taq polymerase (___μl)</td>
<td>Taq (___μl)—5 units/μl □</td>
</tr>
<tr>
<td>dNTPs (___μl)</td>
<td>dNTPs (___μl)—2 mM □</td>
</tr>
<tr>
<td>up primer (___μl)</td>
<td>up primer (___μl)—5 μM □</td>
</tr>
<tr>
<td>down primer (___μl)</td>
<td>down primer (___μl)—5 μM □</td>
</tr>
<tr>
<td>cresol red (___μl)</td>
<td>cresol red (___μl)—5X □</td>
</tr>
<tr>
<td>MgCl2 (___μl)</td>
<td>MgCl2 (<em><strong>μl)—</strong></em> mM □</td>
</tr>
<tr>
<td>PCR buffer (___μl)</td>
<td>PCR buffer (___μl)—10X □</td>
</tr>
<tr>
<td>H2O (___μl)</td>
<td>H2O (___μl) □</td>
</tr>
<tr>
<td>=___μl</td>
<td>=___μl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>tube</th>
<th>MM added?</th>
<th>DNA added?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cycling Conditions

94°C for 5 min.

32 cycles of:
94°C for 2 min.
60°C for 1 min.
72°C for 1 min.

72°C for 25 min
10°C for ∞

Start Time: ________________
End Time: ________________

Adapted from Turner (2010)
APPENDIX IV

Cleanup using QiaQuick PCR purification kit

1. □ Add 5x product/digest volume in Buffer PB to product/digest
   -PB volume ___________μl
2. □ Transfer to spin column; let sit 1 min.
3. □ Spin w/ balance 1 min. @ max speed
4. □ Discard flowthrough.
5. □ Add 750μl Buffer PE
6. □ Spin w/ balance 1 min. @ max speed
7. □ Discard flowthrough.
8. □ Spin w/ balance 1 min. more @ max speed
9. □ Transfer column to a new 1.5ml tube
10. □ Add 60μl dH₂O; let sit 1 min.
11. □ Spin w/ balance 1 min. @ max speed

DNA is now ready for downstream application.
Store @ 4°C if not to be used right away.

Adapted from QiaQuick PCR purification kit insert. (Qiagen Inc., Valencia, CA)
APPENDIX V

Sequencing Reaction Setup

Samples:_____________________________

*Reactions for forward and reverse primers set up separately*

<table>
<thead>
<tr>
<th>10 μl PCR:</th>
<th>X Primer 1 Master Mix (μl):</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR Product</td>
<td>NO DNA in Master Mix</td>
</tr>
<tr>
<td>1 μl</td>
<td>Big Dye (μl)</td>
</tr>
<tr>
<td>Big Dye Terminator</td>
<td>Buffer (μl)</td>
</tr>
<tr>
<td>0.5 μl</td>
<td>Forward primer (μl)</td>
</tr>
<tr>
<td>5x Big Dye Buffer</td>
<td>H₂O (μl)</td>
</tr>
<tr>
<td>1.75 μl</td>
<td></td>
</tr>
<tr>
<td>Up primer</td>
<td>= μl</td>
</tr>
<tr>
<td>1 μl</td>
<td>10 μl</td>
</tr>
<tr>
<td>Down primer</td>
<td></td>
</tr>
<tr>
<td>1 μl</td>
<td></td>
</tr>
<tr>
<td>H₂O</td>
<td></td>
</tr>
<tr>
<td>5.75 μl</td>
<td></td>
</tr>
</tbody>
</table>

\[ \text{X Primer 2 Master Mix (μl):} \]

<table>
<thead>
<tr>
<th>NO DNA in Master Mix</th>
<th>Big Dye (μl)</th>
<th>Buffer (μl)</th>
<th>Reverse primer (μl)</th>
<th>H₂O (μl)</th>
<th>= μl</th>
</tr>
</thead>
</table>

Cycling Conditions

96°C for 1 min.

25 cycles of:
96°C for 10 sec.
50°C for 5 sec.
60°C for 4 min.
10°C for ∞

Adapted from Applied Biosystems BigDye® Terminator v1.1 Cycle Sequencing Kit protocol
APPENDIX VI

Precipitation of Sequence PCR Products (Using Qiagen DyeEx Kit)

Sample #:__________________

- Gently vortex the Dye Ex spin column to resuspend the resin and label the caps.
- Loosen the cap of the column a quarter of a turn.
- Snap off the bottom closure of the spin column and place the spin column in the 2 ml collection tube.
- Centrifuge for 3 minutes at the calculated speed (see kit insert).
- Carefully transfer the spin column to a clean microcentrifuge tube.
- Apply the contents (10 uL) of the sequencing reaction product to the gel bead.
- Centrifuge for 3 minutes at the calculated speed.
- Remove the spin column and place the collection tube onto the hot plate for approximately 20 minutes, until all liquid has evaporated.
- After completely dry, resuspend pellet in 20 uL Hi-Di formamide.
- Vortex for 20 seconds.
- Transfer all contents in the collection tube into sequencer plate, load plate into ABI 3130 sequencer, and sequence.

Adapted from Qiagen Dye-Ex kit insert. (Qiagen Inc., Valencia, CA)