Quantitative PCR-based approach for detection of fecal pollution in water

by

Jia Xue

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 6, 2016

Keywords: Microbial source tracking, qPCR, *Bacteroidales*,
Fecal pollution, DNA adsorption, primer evaluation

Approved by

Yucheng Feng, Chair, Professor of Crop, Soil and Environmental Sciences
Eve Brantley, Associate Professor of Crop, Soil and Environmental Sciences
Luxin Wang, Assistant Professor of Animal Sciences
William Deutsch, Visiting Professor of School of Fisheries, Aquaculture and Aquatic Sciences
C. Wesley Wood, Professor of Soil and Water Sciences Department, University of Florida

Abstract

The water quality of many waterways in our state and nation is deteriorating due to point and nonpoint source pollution from human and animal wastes. Accurate identification of contamination sources is essential for developing cost-effective pollution control strategies. Direct detection of host-specific genetic markers by polymerase chain reactions (PCR) has been widely used in identifying sources of fecal contamination in environmental waters. Four studies were conducted and in the first study, experiments were conducted to validate genetic markers associated with deer/elk, goose, dog, and cattle for bacterial source tracking in Alabama. End-point PCR was performed using DNA extracted from 143 fecal samples of target and non-target animal species. The results showed that one of the two cattle markers, the goose markers, as well as the elk/deer-associated markers had acceptable specificity and sensitivity and thus can be used for bacterial source tracking in Alabama. Field validation showed that both humans and Canada geese contributed to fecal pollution in Parkerson Mill Creek. In addition, water samples collected after a significant rainfall event had the highest frequency of host-associated marker detection.

In second study, quantitative PCR was used to determine concentrations of host-associated genetic markers. A more practical and reliable approach was developed to determine the limits of detection and quantification for qPCR assays at both analytical and process levels for two cattle-associated genetic markers. Our results indicated the cattle marker, CowM3, had better performance characteristics overall compared with the CowM2 marker.

The objective of the third study was to determine if humans and cattle contribute to fecal pollution at a municipal beach in the eastern shore of Mobile Bay, which has been included on Alabama's 303(d) list due to elevated enterococci concentrations in coastal waters. DNA extracted from water samples was subjected to quantitative PCR targeting general *Bacteroidales* as well as human- and cattle-associated *Bacteroidales*. Enterococci were found in all water samples ranging from 2 to 8000 CFU/100 ml. High concentrations of enterococci frequently occurred after significant rainfall events. There was a positive correlation between enterococci and the general *Bacteroidales* marker. The human-associated marker was detected in 49 out of 101 samples, but only nine samples had concentrations high enough for quantification.

Adsorption of DNA by sediment increases the persistence of free DNA in the aquatic environment and thus may cause ambiguities in the identification of recent fecal pollution sources when PCR-based methods are used. In the fourth study, the adsorption and desorption of DNA molecules on both freshwater and marine sediments were quantified using quantitative PCR. Both DNA extracted from raw sewage and purified PCR products were used in the experiment, and their sorption kinetics showed different trends. More DNA was adsorbed on both sediments in stream water than in 5 mM NaCl solution. DNA adsorption on both freshwater and marine sediments was increased in the presence of Mg^{2+} and Ca^{2+}. Clay content in the sediments was another important factor influencing DNA adsorption capacity. Adsorption data were fitted with equations of Langmuir and Freundlich. The observed DNA adsorption capacity was higher than the maximal capacity estimated from the Langmuir equation, suggesting the presence of multilayer adsorption. Desorption experiments were performed using various solutions and 5–22% of adsorbed DNA was desorbed. The results indicate that more DNA molecules were adsorbed on sediment through ligand bonding than electrostatic bonding.

Taken together, quantitative PCR-based bacterial source tracking methods hold great promise for accurate identification of fecal contamination sources in surface waters. Future research is needed to better understand the influence of sediments on the outcome of bacterial source tracking studies.

Acknowledgments

I would like to thank my major professor, Dr. Yucheng Feng, for her expert advice and encouragement throughout this challenging project. This project would have been impossible without her support and valuable advice. I would like to give my great thanks to my committee for their advice, support, and encouragement: Dr. Eve Brantley, Dr. Luxin Wang, Dr. William Deutsch, and Dr. C. Wesley Wood. I really appreciate for the learning opportunities provided by my committee. Also, I want to give my thanks to Dr. Tung-Shi Huang for serving as my research dissertation's University Reader and to Dr. Stephen Ditchkoff, Ms. Kathryne Jernigan and Patty Tyler for assisting with fecal sample collection. My completion of this project would not have been accomplished without the support from my labmates: Michele Owsley, Udenika Wijesinghe, Reji Mathew, and Lu Fan. I would also like to give my thanks to all my friends and classmates: Jianping Wang, Rui Yang, Xu Wang, Boxin Chang, Yaru Lin, and Zack Ogles. Finally, I would like to express my deepest appreciation and gratitude to my loving and supportive wife Bowen Zhang, my daughter Olivia Xue, and my family for their encouragement, support and love throughout my study at Auburn University!

Table of Contents

List of Tables

List of Illustrations

Chapter 1

Literature Review

1.1 Introduction

Fecal pollution from humans and animals is one of the major concerns in relation to water bodies used for drinking, recreational activities, and seafood harvesting. Pathogens associated with fecal pollution can lead to human disease and economic losses in industries due to the closures of beaches and seafood harvesting (Fong and Lipp, 2005; Nayak et al., 2015; Scott et al., 2002; Webster et al., 2013). Pollution sources, such as stormwater runoff, broken sewer pipeline, cattle farms, wildlife, and agricultural runoff may contribute to the higher concentration of fecal indicator bacteria. People made great efforts to minimize fecal input into water, but the problem persists, partly due to an inability to reliably identify nonpoint sources. Knowing the source of contamination is the key to solving the problem. Water microbiological quality was assessed by enumerating fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*) and enterococci in recreational waters (USEPA, 2008). However, the majority of these organisms are not limited to humans but also exist in the intestines of many other warm-blooded animals, such as bovine, horse, dog, cat, duck and goose (Johnson et al., 2004; McLellan et al., 2003). Also this method requires at least 24 h to obtain results (Boehm et al., 2002; Leecaster and Weisberg, 2001). Also, *E. coli* and enterococci may survive and reproduce in the sediment after being released into environment (Anderson et al, 2005; Desmarais et al., 2002; Solo-Gabriele et al., 2000). The FIB method, however, does not identify the source of fecal contamination. Due to these disadvantages, the effectiveness of using FIB method to predict the

presence of human or animal waste impact on health and to develop effective pollution control strategies is limited.

Microbial source tracking (MST) has been widely used during the last decade. The concept that the origin of fecal pollution can be traced using microbiological, genotypic, and chemical methods has been termed microbial source tracking. MST includes a group of methodologies that are aimed to determine the origins of fecal pollution in waterbodies (Scott et al., 2002; Stoeckel and Harwood, 2007). Many of them are based on library-independent molecular based techniques, which significantly increase the speed and reliability of identification of fecal sources.

Library-independent methods rely on identification of host-specific microbial species or genotypic traits of microorganisms to identify sources of fecal contamination. One emerging MST method is based on the detection of host-specific 16S rRNA markers that target the order of *Bacteroidales* due to its host specificity, broad geographic stability, and more abundance in the gastrointestinal tract of warm-blooded animals (Bernhard and Field, 2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006). Members of *Bacteroidales* are gram negative, rod shaped, bile-resistant, non-spore forming obligate anaerobes living in the intestinal tract of warm-blooded animals (Okabe et al., 2007; Roslev and Bukh, 2011; Wexler, 2007). It has been reported that *Bacteroidales* are one of the most promising alternative indicator organisms for MST, even though report showed *Bacteroides* may survive for several days at low temperature in surface water (Kreader, 1998). Compared to fecal coliforms (e.g., *E. coli*) and *Enterococcus*, *Bacteroidales* are obligate anaerobes, and have a shorter life span in the secondary habitat. This makes the order of *Bacteroidales* a better candidate for microbial source tracking studies.

This study focuses on Fairhope municipal beach and Parkerson Mill Creek Watershed in Alabama. Recently, a portion of the eastern shore of Mobile Bay has been included on Alabama's 303(d) list due to elevated enterococci concentrations in coastal waters. Fairhope city officials have debated the cause of contaminated water on the beach of the Fairhope Pier, where high counts of bacteria often force the beach to be closed. The possible pollution sources, such as stormwater runoff, leaky sewer line, waterfowl at the duck pond on the north beach, cattle farms, wildlife, and agricultural runoff may contribute to the higher concentration of fecal indicator bacteria.

Parkerson Mill Creek Watershed is part of the Chewacla Watershed in the lower Tallapoosa River Basin located in Lee County, Alabama. Parkerson Mill Creek is impaired because it does not meet water quality criteria to support its designated use as a fish and wildlife stream. In 2007, A 6.85-mile segment of the creek was listed on the Alabama Department of Environmental Management (ADEM) 303(d) list of impaired waters for pathogens from point and non-point pollution sources. (http://adem.alabama.gov/programs/water/303d.cnt). Urban runoff, pet waste, wildlife, and leaky sewer lines have been thought to contribute to the high level of *E. coli* concentration (Parkerson Mill Creek Watershed Management Plan, 2010). But more information is needed for both studying sites to identify the major sources of fecal pollution so that a better strategy to protect humans against health risk posed by polluted water can be made.

Increasing interest is now being directed towards the use of library- and cultivation-independent microbial source tracking (MST) methods based on polymerase chain reaction (PCR) targeting host-specific molecular markers. The MST methods combined with end-point PCR and qPCR will be able to provide results that reflect most recent fecal pollution and identify

the sources, which could provide better strategy to protect humans against health risk posed by polluted water (Wade et al., 2010).

1.2 Objectives

The objectives of these studies were to:

1. Evaluate the specificity and sensitivity of Canada goose-, deer-, cattle-, and dog-specific genetic markers targeting animal species that likely to affect a local watershed (Parkerson Mill Creek) for fecal source identification in Alabama.
2. Examine the applicability of cattle-associated genetic markers (CowM2 and CowM3) in the field and develop an accurate and standardizable method of calculating a lower limit of detection (LLOD) and lower limit of quantification (LLOQ) for each marker.
3. Determine whether and to what extent humans and cattle contribute to fecal pollution at the Fairhope municipal beach.
4. Determine DNA adsorption and desorption by freshwater and marine sediments.

1.3 Literature Review

1.3.1. Fecal pollution in water

Fecal contamination in rivers and coastal waters originate from both human and non-human derived sources. Pathogens associated with fecal pollution would lead to human disease such as eye infections, gastrointestinal (GI) illness and skin complaints (Balarajan et al., 1991; Haile et al., 1999). The United States Environmental protection agency (USEPA) has examined the potential for illness from exposure to non-human fecal contamination compared to the potential for illness from exposure to human fecal contamination. The results indicate both human and animal feces in recreational waters pose similar risks to human health (USEPA

2009). A recent study suggested that careful consideration is needed for the management of recreational waters impacted by cattle sources as well as by human sources (Soller et al., 2010). The United States Department of Agriculture (USDA) estimated that there are 92.0 million head cattle and calves on Jan. 1st 2016, up 3% from 89.1 million head in 2015 (USDA/NASS 2014). Fecal pollution from cattle is usually caused by inappropriate land application of manure (Pell, 1997), direct contact of livestock and water ways (Burt et al., 2013), and livestock grazing (DeRamus, 2004; Roche et al., 2013). Fecal contamination can also cause economic losses in industries due to the closures of beaches and seafood harvesting (Fong and Lipp, 2005; Scott et al., 2002). These problems were common to all countries regardless of economic status, although the condition of fecal pollution as well as the primary agents of disease differed among countries.

The Clean Water Act (CWA) is a series of legislative acts that form the foundation for protection of U.S. water resources. The goal of the CWA was to restore and maintain the chemical, physical, and biological integrity of the nation's waters. One of the most important goals of CWA was to protect and restore waters for swimming. An recommended water quality criteria (RWQC) published by U.S. EPA in 2012 accurately reflect the latest scientific knowledge on the identifiable effects on health and welfare that might be expected from the presence of pollutants in water bodies (U.S. EPA 2012). In the RWQC, EPA recommended using the fecal indicator bacteria (FIB) *E. coli* as indicators of fecal contamination for fresh water and enterococci for marine water. These recommendations are referred to as 304(a) criteria (Table 1.1) for the purpose of protecting human health in coastal recreation waters that used for swimming, bathing, or similar water contact activities. Based on RWQC, EPA recommended that each state needs to make a risk management decision regarding illness rate with will the most appropriate criteria for their waters. The geometric mean (GM) of fecal indicator

concentration of water body should not be greater than the selected GM magnitude in any 30-day interval (Table 1). Also, there should not be greater than a ten percent excursion frequency of the selected statistical threshold value (STV) magnitude in the same 30-day interval.

Section 303(d) of the CWA requires each state in the United States to establish a list of impaired water bodies that currently do not support their designated uses. Based on 2012 Draft Alabama 303(d) list, there were 229 (296 in 2016 draft Alabama 303 (d) list) impaired water bodies on the list, and pathogens was one of the major causes. The state also has to establish the total maximum daily loads (TMDL) that will meet water quality standards for each listed water body. TMDL is the sum of the individual waste-load allocations for point sources and load allocations for nonpoint sources and natural background. It is simple to identify whether the water is contaminated with fecal materials or not. However, identification of sources of contamination was not that easy. Determining the sources of fecal pollution was the key for developing effective pollution control strategies and best management practices (BMPs).

1.3.2 Fecal indicator bacteria (FIB)

Water pollution caused by fecal contamination is a serious environmental problem due to human health risks. World health organization(WHO) reported that 1.5 billion people suffered from a lack of safe drinking water and 3.4 million people died each year due to these water borne diseases (WHO 2001). In order to protect humans from these polluted waters, U.S. EPA made drinking water regulations for pathogens and indicators to monitor the biological water quality. CWA defined pathogen indicator as a substance that indicates the potential for human infectious disease (U.S. EPA 2012a). Inadequately treated water may contain disease-causing pathogens, which include various types of bacteria, viruses, and protozoan parasites (U.S. EPA 2012b). For more than a century, the biological water quality has been assessed by using FIB, like fecal

coliforms *Escherichia coli* (*E. coli*), and enterococci. Most strains of enterococci and *E. coli* do not cause human illness, however they are used as indicators because they often co-occur with other fecal contaminants that are pathogens.

FIB includes three groups of organisms: total coliform bacteria, fecal coliforms, and enterococci. Total coliform are common in ambient water and may be injured by environmental stresses such as lack of nutrients. EPA considered total coliforms to be a useful indicator because the absence of total coliforms in the water system indicates a reduced likelihood of fecal contamination.

1) Fecal coliforms

Fecal coliforms are Gram-negative, non-spore forming, rod-shaped, aerobic or facultative anaerobic bacteria that are able to ferment lactose with gas and acid with 48h at 35°C (Rompre et al., 2002). The fecal coliform definition has also been revised to thermotolerant coliform which produced indole from tryptophane at a temperature of 44±0.5°C, as this is a more accurate description of the group (Tallon et al., 2005). The characterization studies showed *E. coli* represented over 94% of the thermotolerant coliforms isolated from human feces, while the other thermotolerant coliforms identified as KEC members (*Klebsiella*, *Enterobacter,* and *Citrobacter*) ranged from 3.2-7.4%. Fecal coliforms were considered to be present specifically in the gut and feces of warm-blooded animals. It was originally believed that the presence of the selected coliforms was purely fecal origin and that was why these organisms were referred to as fecal coliform. But research showed “fecal coliform” was not a promising method to indicate fecal contamination in water, because fecal coliforms were present in the environment where no fecal pollution had occurred (Byamukama et al., 2000; Edberg et al., 2000; McLellan et al., 2001).

2) *E. coli*

In 1893 *E. coli* was first introduced as an indicator of fecal contamination by Blachstein. The use of *E. coli* as an indicator of microbiological water quality dates from their first isolation from feces at the end of the 19th century. But *E. coli* was not commonly used during the early years simply because of the detection method was not suitable for routine testing. There were two key factors that led to the use of *E. coil* as the indicator for detection of fecal pollution in waters: First, the finding that some "fecal coliform" were of non-fecal origin; second, the development of testing method for *E. coli*. *E. coli* is a rod shape, Gram-negative, and facultative aerobic bacterium belonging to the fecal coliform group (Rompre et al., 2002). It has a rapid growth rate, with a generation time of 20 minutes under optimum conditions. *E. coli* is the major member of fecal coliform that produces indole from tryptophan and had a positive reaction with enzyme β-glucuronidase (Tallon et al., 2005). *E. coli* is present in large numbers among the intestinal flora of healthy humans and other warm-blooded animals, and thus is found in fecal wastes. In 1986, EPA recommended the use of *E. coli* as indicator bacteria for fresh water (USEPA, 1986).

There has been a debate on if *E. coli* should be used as an indicator of the possible presence of enteric pathogens in aquatic systems. Kaper's study found there were at least 700 recognized *E. coli* strains (Kaper et al., 2004), but only about 10% of *E. coli* strains caused disease (Feng, 1995; Kaper et al., 2004). The majority of the strains were the natural inhabitants of the gastrointestinal tract of warm-blooded animals. The disease causing *E. coli* include *E. coli* O157: H7, which was the main cause of hemolytic uremic syndrome in the United States. The "O" and "H" antigens on the bacteria and their flagella distinguish the different serotypes. There were over 160 serogroups recognized based on antigens to specific lipopolysaccharides found either on cell envelope (O antigens) or the flagella (H antigens) and most of them do not cause

disease (Saylers and Whitt 2002). EPA-approved standard methods for *E. coli* did not typically identify the presence of pathogenic *E. coli* strains, which means that an *E. coli* positive result was an indicator of fecal pollution but is not necessarily a measure of waterborne pathogen occurrence (EPA 2008).

3) Enterococci

The enterococci are Gram-positive, non-spore forming, obligate fermentative chemoorganotrophs. They were previously classified into the genus *Streptococcus*. The enterococci were proposed as a division composed of organisms that generally grow at 10 and 45 °C, in 6.5% NaCl, at pH 9.6 and survive at 60°C for 30 min (Martinez-Murcia and Collins 1991). In 1984, *Enterococcus* was separated from *Streptococcus,* and right now there are 36 known *Enterococcus* species, and the most abundant strains include *Ent. faecalis, Ent. faecium, Ent. casseliflavus, Ent. hirae,* and *Ent. mundtii* (Badgley et al., 2010; Ryu et al., 2013). When enterococci were released from gastrointestinal tract of warm-blooded animals into secondary habitats like water environment or sediment, the environmental stressors such as sunlight and salinity would lead to a decline in the population over time (Davies et al., 1995; Davies-Colley et al., 1999). Despite the negative affection by environmental stressors, many studies clearly demonstrated the persistent of some *Enterococcus* spp. in extra-enteric habitats (Table 1.2).

Enterococci are important members of gut communities in many animals, like humans, birds, cattle, swine, and wildlife. They are also opportunistic pathogens that cause millions of infections annually (Robert C. Moellering, 1992). Ostrolenk et al. were among the first to recommend that enterococci might be more appropriate FIB than *E. coli*, and studies have confirmed this recommendation for marine waters (Ostrolenk et al., 1947). Recently, studies have found an association between enterococci densities and illness rates at beaches impacted by

nonpoint sources of contamination (Fleisher et al. 2010). The characteristics of the member enterococci, such as abundance in human and animal feces, easy to cultivate, correlation with human health in marine and freshwaters, and association with the presence of pathogens (USEPA 1986), have led to the widely use as tools for assessing marine and freshwater quality worldwide (Wade et al., 2008).

4) Alternative indicator bacteria: *Bacteroidales*

For decades, FIBs have been used to indicate fecal pollution and potential human health risks in water bodies (USEPA, 1986). However, the FIB method was not specific to any fecal sources and this method requires at least 24 h to obtain results (Boehm et al., 2002; Leecaster and Weisberg, 2001). Besides *E. coli* and enterococci may survive and regrow in the sediment after being released into environment which will adversely affect the identification of the most recent pollution event in water (Anderson et al., 2005; Desmarais et al., 2002; Solo-Gabriele et al., 2000). Accurate identification of contamination sources is essential for developing cost-effective pollution control strategies. The increasing interest is now being directed towards the use of library- and cultivation-independent microbial source tracking (MST) methods based on polymerase chain reaction (PCR) targeting host-specific molecular markers. The MST methods combined with end-point PCR and qPCR will be able to provide results that reflect most recent fecal pollution and identify the sources, which could provide better strategy to protect humans against health risk posed by polluted water (Wade et al., 2010).

Members of the order *Bacteroidales* are Gram negative rods, bile-resistant, non-spore forming and obligate anaerobes. They are the most predominant anaerobes in the gut of warm-blooded animals (Wexler, 2007). *Bacteroidales* have been used as a promising fecal indicator to monitor microbial water quality due to its host specificity, broad geographic stability, and higher

abundance in the gastrointestinal tract of warm-blooded animals (Bernhard and Field, 2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006). Typically, bacteria comprise approximately 1/3 of feces by weight, and about 25% to 40% of the amount of total fecal bacteria belongs to *Bacteroidales*; therefore, it may comprise about 10% of the fecal mass (Stephen and Cummings, 1980). As they have a shorter life span in the secondary habitat, and *Bacteroidales*-based methodologies are designed to target specific diagnostic sequences within the *Bacteroidales* 16S rRNA gene present in feces from different animals, also there is no need to culture individual *Bacteroidales* isolates. So far, host specific *Bacteroidales* genes have been found from human, dogs, cattle, horse, swine, chickens, Canada goose, and elk (Dick et al., 2005; Layton et al., 2006). The size of the genome in *Bacteroidales* is about 4.5 million bp and there are 3.5 to 6 16S rRNA gene copies per genome in this species (Větrovský and Baldrian, 2013; USEPA, 2010). Compared with culturing methods (USEPA, 2000), the qPCR based MST method has the ability of not only provide a reliable and accurate method to estimate contributions of fecal concentrations in water, but also to identify fecal pollution sources in surface water (Scott et al., 2002). All these advantages above make the use of *Bacteroidales* spp. as indicators of the type of host animal significantly better than other fecal coliforms and makes the order of *Bacteroidales* a better candidate for microbial source tracking studies.

1.3.3 Microbial communities in human and animal intestines

The human microbiome is composed of 10^{14} bacterial cells, which is 10 times more than the total number of human cells (Hattori and Taylor, 2009). It was considered that human GI tract may be the most complex ecosystem that consists of bacteria, archaea, yeasts, and filamentous fungi (Miller and Wolin, 1986). The endogenous GI microbial flora played an

important role on human biology including their relationship with health and disease. However, the ecosystem remains incompletely characterized, and our current knowledge of the normal human GI microbial diversity overall was about 30% (Rajilić-Stojanović et al., 2007). There were two major reasons lead to the incomplete characterization: First, due to the fact that only a limited number of individuals have been subjected to the analysis of the human intestinal microbial community; second, variation associated with environment, diet, health status, and host genetics. Besides, the limitations of the approach, such as detection limit, have to be concerned. Our current research on the human GI microbial diversity originates from cultivation-based and molecular studies. Because of the insensitivity of cultivation, investigators have begun to use molecular fingerprinting methods and sequence analysis of microbial small-subunit (SSU) ribosomal RNA genes (16S rDNA).

Over the last decade, the bacterial 16S rRNA gene sequence has been a useful tool for analyzing the microbial diversity. Bacterial community within the human GI tract is exceptionally diverse. Members of eight bacterial phyla were found to inhabit the human GI tract (*Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Fusobacteria, Verrucomicrobia, Euryarchaeota,* and *Cyanobacteria*). Both cultivation-based and molecular approaches showed unequivocally that *Firmicutes* are by far the most diverse group, then followed by *Bacteroidetes* (Table 1.3) (Eckburg, 2005; Hattori and Taylor, 2009; Rajilić-Stojanović et al., 2007). Members of *Proteobacteria* are also common and diverse, but they are usually secondary to the above. Lin et al. (1997) evaluated the microbial community structure of the gastrointestinal tracts of various domestic animals, such as bovine, ovine, caprine, and swine. Bacterial, eukaryotic, and archaeal rRNAs were estimated to account for 60-90%, 3-30%, and 0.5-3%, respectively (Lin et al., 1997). Results showed that the microbial population of lower intestinal bacteria of cattle was

dominated by strict anaerobes like *Bacteroides* spp., and *Bifidobacterium* spp. (Dowd et al., 2008). Kim et al. (2011) studied the composition and distribution of the microbial population in intestinal tracts of swine using culture independent 16S rRNA gene sequencing method. The results showed that bacterial communities of all samples were comprised primarily of *Firmicutes* and *Bacterioidetes*, which accounted for more than 90% of total sequences. At the genus level, 15 genera contained more than 59% of the total sequences, and 14 of the 15 genera belonged to *Firmicutes*. In addition, as the age of pig increased, the proportion of *Firmicutes* increased, but the proportion of *Bacteroidetes* decreased (Kim et al., 2011).

1.3.4 Identification of sources of fecal contamination

Various approaches have been used to identify the origin of fecal pollution in water bodies. They can be divided into two basic groups: chemical method and biological method.

1) Chemical methods

A reliable chemical marker should be constant, regular, and allow the unambiguous elucidation of the sources and the quantification of the magnitude of pollution. Organic wastewater contaminants (OWCs) have been used as chemical markers to identify the sources of fecal contamination in the environment. According to Tyagi et al. (2007), the chemical markers they used (fecal sterols and bile acids) were detected by using gas chromatograph and mass spectrometer (GC-MS). Fecal sterols are C27, C28, and C29 cholestane-based sterols found in fecal material; they are formed as reduction products of cholesterol and the higher molecular weight congeners. These sterols include coprostanol, epicoprostanol, cholesterol, cholestanol, stigmastanol, and stigmasterol. Bile acids are steroidal acids produced in the digestive system of warm blood animals, they include lithocholic acid, deoxycholic acid, cholic acid, chenodeoxycholic acid, ursodeoxycholic acid, and hyodeoxycholic acid (Tyagi et al., 2007).

Based on Kolpin et al. (2002), 95 OWCs were measured in water samples from 139 streams across 30 states during 1999 and 2000, they found the most frequently detected compounds were coprostanol, cholesterol, triclosan, and caffeine, all of these compounds suggest the influence of human activities on the water bodies (Koplin et al. 2002).

Caffeine is a potential chemical marker for domestic wastewater contamination. It is an alkaloid that occurs in more than 60 plant species, such as the seeds of coffee, cacao, tea, and cola tree. The correlation between caffeine concentrations and the population has demonstrated the suitability of caffeine as a quantitative anthropogenic marker for wastewater contamination of surface water (Buerge et al., 2003).

2) Biological methods

Environmental waters were susceptible to fecal pollution from point and nonpoint sources. Water contaminated with human feces is regarded as a risk to human health, especially in developing countries where the drinking water was untreated or insufficiently treated. Traditional biological methods using FIB such as fecal coliform, *E. coli*, and *Enterococcus* spp. to predict the presence of fecal contamination in water, however they cannot differentiate the sources of microbes whether they come from humans or other animals. Recent development of testing methods and analytical techniques make the identification of specific sources of these organisms possible. Microbial source tracking is an increasingly used approach to determine host-specific contributions of fecal contamination to environmental waters by using microbiological, genotypic, and phenotypic methods. An ideal indicator would be non-pathogenic, rapidly detected, easily enumerated and have survival characteristics such as coevolve with their host. Therefore by living in specific host animals for a period of time, the organisms should possess a similar or identical genetic fingerprint, which will differ from those

adapted to a different host (Scott et al., 2002). MST included a group of methodologies that were aimed at identifying, and in some cases quantifying the dominant source of fecal contamination in resource waters, including drinking, ground, recreational, and wildlife habitat waters.

1.3.5 Microbial Source Tracking (MST)

There were two major types of MST, library-dependent method and library-independent method. These two methods can be divided into four categories (Bush et al. 2003): library dependent genotypic methods, library dependent phenotypic methods, library independent phenotypic methods, and library independent genotypic methods.

a) Library dependent genotypic methods: ribotyping, pulse-field gel electrophoresis (PFGE), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) analysis, and repetitive DNA sequences (Rep-PCR).

Ribotyping also referred to as "molecular fingerprinting", is based on the differentiation of genetic differences in the genomic sequences of 16S or 23S ribosomal RNA genes. The ribotyping procedure provides a DNA fingerprint of bacterial genes coding for ribosomal ribonucleic acids (rRNA), which were highly conserved in microorganisms (Field and Samadpour, 2007).

Pulse-field gel electrophoresis (PFGE) is a DNA 'fingerprinting' technique that uses rare cutting restriction enzymes on the entire DNA genome (Hagerdorn et al., 2003). This method involves direct analysis of the microbial genome without performing PCR. Restriction enzymes are used to cut the genomic DNA infrequently, resulting in about 10 to 30 large fragments. The large genomic fragments are then separated by subjecting them to alternately pulsed, perpendicularly oriented electrical fields.

Repetitive DNA sequences (Rep-PCR) is a DNA fingerprinting technique that uses repetitive intergenic DNA sequences to differentiate between sources of fecal pollution (Dombek et al., 2000). This method has been used extensively because it is rapid, simple, and less expensive compared to other genomic methods. The genetic fingerprints generated using rep-PCR contain multiple bands, which can be subsequently analyzed, categorized by host sources, and used to construct a database to identify the source of an unknown isolate.

Randomly amplified polymorphic DNA (RAPD), non-selective primers at high stringency have been used to produce a series of species or strain specific PCR products that depend on both the primer and template used. This method is relatively inexpensive compared to ribotyping and PFGE. But its disadvantage, i.e., poor reproducibility and lab-to-lab variation, limited its use for MST work (USEPA, 2005).

Amplified fragment length polymorphism (AFLP) analysis was developed for plant genome mapping, later on its use was extended to fingerprinting bacterial species. Most of the AFLP analysis published so far have focused on epidemiological studies (USEPA, 2005). Discriminant analysis showed that AFLP gave better isolate separation into host groups than multiple antibiotic analysis and 16S rRNA analysis (Guan et al., 2002).

b) Library dependent phenotypic methods: antibiotic resistant analysis (ARA), carbon source profiling.

Antibiotic resistant analysis (ARA) is used to identify the fecal sources by screening fecal streptococci or *E. coli* isolates against commonly used antibiotics. ARA method was based on patterns of antibiotic resistance of bacteria from human and animal sources. The promise behind this method was that fecal bacteria from human would have greater resistance to specific antibiotics followed by livestock and wildlife, and that livestock will have greater resistance to

other antibiotics (Hager, 2001a). The difference of resistance was because humans are exposed to different antibiotics than cattle, pigs and other wildlife, etc. The resistance pattern of an organism can be used to identify its source. A database of antibiotic resistant patterns from known sources will be needed to compare sample isolate patterns to.

c) Library independent phenotypic methods: fecal bacterial ratio, host specific indicator organisms, toxin biomarkers.

Fecal coliform / fecal streptococci (FC/FS) ratios have been used to assess the general source of nonpoint fecal pollution, with FC/FS > 4 indicating humans, FC/FS between 0.1 and 0.6 indicating domestic animals, and FC/FS < 0.1 indicating wild animals as the source. But later scientists have found this ratio was difficult to use in agricultural settings. That was because The FC/FS ratio is influenced by temperature, the presence of sediment, and sediment particle size (Howell et al., 1996).

d) Library independent genotypic methods: host specific 16S rDNA markers, terminal restriction fragment length polymorphism analysis (T-RFLP), length heterogeneity PCR (LH-PCR).

T-RFLP is a method used to determine the diversity of an entire bacterial community by examining differences in the 16S rRNA gene. This method is considered to be a library independent method, because it doesn't require the isolation of environmental strains.

1.3.6 Host specific markers

Library independent molecular markers (Host specific markers) for MST can target the sequences in host associated microorganisms or sequences derived from the host. The fecal source associated molecular markers can be divided into three groups: molecular markers in prokaryotes, markers in viruses, and markers in eukaryotes (Roslev and Bukh, 2011).

Molecular markers in prokaryotes have been identified in the order *Bacteroidales*. Members of *Bacteroides-Prevotella* group and the genus *Bifidobacterium* are fecal anaerobes and they make up a large portion of the fecal bacteria. The major limitation of *Bacteroides* and *Bifidobacterium* using as indicators of fecal pollution is that they are difficult to grow in culture media. Recently the development of molecular methods has improved the ability for their use in water quality monitoring. In 1995, Kreader developed PCR-based assays to amplify genes from three cultivated strains of *Bacteroides* to monitor fecal pollution from humans (Kreader, 1995). Bernhard and Field further advanced this approach by identifying host-specific *Bacteroidales* 16S rDNA markers for humans and cows (Bernhard and Field, 2000a). Additional host associated *Bacteroidales*-specific polymerase chain reactions have been developed and validated for humans, chickens, elk, dogs, Canada goose, cows, and dogs (Table 1.4).

Over 100 different enteric viruses are excreted in human and animal feces, and they also show good persistence in environmental waters. Even though many of these viruses are not easy to cultivate in environmental samples, the use of real-time quantitative PCR make the detection of molecular markers in human- and animal-associated viruses possible. Quantitative molecular detection assays using qPCR have been developed for many host groups. Human associated viruses include members of adenovirus, enterovirus, norovirus, and polyomavirus. Animal associated viruses like cattle-associated adenoviruses, and enteroviruses, pig-associated adenoviruses, and teschoviruses and so on (Table 1.4).

In 2005, Martellini and colleagues first proposed the idea of using eukaryotic mitochondrial DNA sequence (mtDNA) as direct fecal marker to differentiate human, bovine, porcine, and ovine sources of fecal pollution (Martellini et al., 2005). Feces from humans and other warm-blood animals contain blood and intestinal cells from their host. In that case, to

target the host nucleic acids directly rather than molecular markers will be much more specific, and because of the large amounts of cells from host, this method also have very high sensitivity. More and more studies have developed primers for detection of mtDNA from humans, dogs, cats, Canada goose, and deer in wastewater samples. Right now quantitative detection methods like qPCR have been developed for most of the human and animal mtDNA marker groups (Table 1.4).

1.3.7 Fundamentals of quantitative polymerase chain reaction (PCR)

1) Quantitative PCR (qPCR)

The application of PCR in combination with environmental nucleic acids has been widely developed as culture-independent approaches for detecting the source of fecal contamination in water environment. The PCR process can be summarized in three steps: 1) Double-stranded DNA (dsDNA) separation at temperature >90°C; 2) Primer annealing at 50-75°C; 3) optimal extension at 72-78°C (Mackay, et al., 2002). QPCR methods combined the detection of target template with quantification by recording the amplification of a PCR product via a corresponding increase in the fluorescent signal associated with product formation during each cycle. There are two types of fluorescence detection chemistries used to detect template in qPCR, SYBR green assay and probe-based system (Smith and Osborn, 2009). SYBR green binds to all dsDNA and emits fluorescent signals. During its unbound state, no fluoresce can be detected. So the template amplification can be measured in each cycle by the corresponding increase in fluorescence. For probe-based assays, in annealing process the probe and primers bind to the template. The intact probe includes a fluorescent 5' end (reporter) and one 3' end called quencher. When they are in close proximity on the probe, the quencher hijacks the emissions that resulted in no fluorescence can be detected. As the new strand is synthesized by Taq polymerase,

the 5’ nucleotide will be cleaved by enzyme, once they are no longer in close proximity, the fluorescent signal from the probe is detected and template amplification is recorded by the increase in fluorescence.

The qPCR methods have revolutionized our understanding of the contribution of fecal contamination from different sources. Its rapidity, simple, sensitivity, specificity, reproducibility, culture independency, and the reduced risk of carry-over contamination has made this method the most popular MST method. There are also some disadvantages, such as expensive equipment, technically demanding, and the requirement for prior sequence data of the specific target gene of interest, and there are limited markers for human, cattle, swine, goose, dogs, and so on.

2) Quality control of qPCR experiments

a. Limit of detection (LOD) and limit of quantification (LOQ)

Limit of detection is defined as the lowest concentration that can be reliably detected to be statistically different from a blank (USEPA, 2000). A similar definition is the lowest concentration at which 95% of the positive samples are detected (Bustin et al., 2009). In other words, within a group of replicates containing the target at concentrations at the LOD, no more than 5% failed reactions should occur. This concentration is recommended to be 3 standard deviations above the measured average difference between the sample and blank signals.

Limit of quantification is defined as the level above which quantitative results can be obtained. The corresponding sample/blank difference is recommended to be 10 standard deviations above the blank which corresponds to the 99% confidence level. LOQ is always used to define the lower limit of the useful range of the measurement technology in use. Samples that

do not bear residues at or above the LOQ are often referred to as non-quantifiable (Staley et al., 2012).

b. PCR amplification efficiency

PCR amplification efficiency is used to evaluate how well the target sequence can be amplified. For example, a PCR reaction that amplifies the target sequence with 100% efficiency would theoretically double the amount of PCR products with each cycle. In that case the amount of PCR products (C_n) from C_0 input target molecules after n cycles could be calculated by $C_n = C_0(1+E)^n$. Amplification efficiencies were calculated according to the formula:

$$S = -\frac{1}{log(1+E)}$$

Where E is Amplification efficiency, C_0 is the original amount of target sequence, C_n the amount of target sequence of PCR product, and S is the slope of the standard curve. Therefore,

$$E = 10^{1/-s} - 1.$$

c. Specificity and sensitivity in the context of MST

According to USEPA (2009), specificity is defined as the ability of a particular MST method to discriminate between different animal fecal sources; and sensitivity is the proportion of target organisms that can be detected (Table 1.5). Where TPC is number of samples tested positive correctly, TPI is number of samples tested positive incorrectly, TNC is number of samples tested negative correctly, and TNI is number of samples tested negative incorrectly. The equation expressed as:

$$\text{Specificity} = \frac{TNC}{(TNC+TPI)} \qquad \text{Sensitivity} = \frac{TPC}{(TPC+TNI)}$$

d. Inhibitions

The application of PCR-based methods on DNA extracts from environmental samples has to overcome the inhibition of PCR. Substances that inhibit enzyme activity are present in many materials. Based on Kreader's study in 1996, known inhibitors include EDTA, sodium dodecyl sulfates (SDS), Triton X-100, bile salts (sodium cholate plus deoxycholate), tannic acids, humic acids, bilirubin, and $FeCl_3 \bullet 6H_2O$ and so on (Kreader, 1996). The inhibitory level for each reagent is defined as the lowest concentration of the inhibitor reproducibly reduced the yield of PCR product (Table 1.6).

Several different strategies have been established to overcome the inhibitory effect: 1) inhibitory contaminants can be removed from DNA extracts using cleanup procedures; 2) proteins such as bovine serum albumin (BSA) or phage T4 gene 32 protein can be added to PCR reactions in order to eliminate inhibitors and protect DNA polymerases; 3) DNA samples can be diluted to lower the concentration of co-extracted components and to improve PCR amplification; 4) since co-extracted components from soil especially in strongly acidic forest soil contain many inhibitors, DNA extraction protocols have been optimized to avoid co-extraction of PCR-inhibitors by increasing salt concentration in the lysis buffer (Schneider et al., 2009).

1.4 References

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

ADEM. 2011. Draft total maximum daily load (TMDL) for Parkerson Mill Creek for pathogens (*E. coli*). http://adem.alabama.gov/programs/water/wquality/tmdls/FinalParkersonMillPathogenTMDL.pdf

Badgley, B.D., Nayak, B.S., Harwood, V.J., 2010. The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. Water Research 44, 5857–5866. doi:10.1016/j.watres.2010.07.005

Balarajan, R., Raleigh, V., Yuen, P., Wheeler, D., Machin, D., Cartwright, R., 1991. Health risks associated with bathing in sea-water. Br. Med. J. 303, 1444–1445.

Bernhard, A. E., Field, K. G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66(10), 4571–4574. http://doi.org/10.1128/AEM.66.10.4571-4574.2000

Bernhard, A. E., Field, K. G., 2000b. 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(4), 1587–1594. http://doi.org/10.1128/AEM.66.4.1587-1594.2000

Boehm, A. B., Grant, S. B., Kim, J. H., Mowbray, S. L., McGee, C. D., Clark, C. D., Foley, D. M., Wellman, D. E., 2002. Decadal and shorter period variability of surf zone water

quality at Huntington Beach, California. Environ. Sci. Techno. 36(18), 3885–3892. http://doi.org/10.1021/es020524u

Buerge, I.J., Poiger, T., Müller, M.D., Buser, H.-R., 2003. Caffeine, an anthropogenic marker for wastewater contamination of surface waters. Environ. Sci. Techno. 37, 691–700. doi:10.1021/es020125z

Burt, C., Bachoon, D.S., Manoylov, K., Smith, M., 2013. The impact of cattle farming best management practices on surface water nutrient concentrations, faecal bacteria and algal dominance in the Lake Oconee watershed: Cattle farming best management practices. Water Environ. J. 27, 207–215. doi:10.1111/j.1747-6593.2012.00343.x

Bush, L., W. Kurz, L. Powell, M. Power. G. Putnam, and E. Vandenburgh. 2003. Addressing microbial pollution in costal waters; A reference for local government. Appendix 1: technical source tracking methods. 41-47

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 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

Byamukama, D., Kansiime, F., Mach, R.L., Farnleitner, A.H., 2000. Determination of *Escherichia coli* contamination with chromocult coliform agar showed a high level of discrimination efficiency for differing fecal pollution levels in tropical waters of Kampala, Uganda. Appl. Environ. Microbiol. 66, 864–868. doi:10.1128/AEM.66.2.864-868.2000

Byappanahalli, M., Fujioka, R., 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water Sci. Technol. 50, 27–32.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Ting, W.T.E., Tseng, C.C., Nevers, M.B., 2006. Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. J. Water and Health 4, 313–320. doi:10.2166/wh.2006.018

Davies, C., Long, J., Donald, M., Ashbolt, N., 1995. Survival of fecal microorganisms in marine and fresh-water sediments. Appl. Environ. Microbiol. 61, 1888–1896.

Davies-Colley, R.J., Donnison, A.M., Speed, D.J., Ross, C.M., Nagels, J.W., 1999. Inactivation of fecal indicator microorganisms in waste stabilisation ponds: Interactions of environmental factors with sunlight. Water Res. 33, 1220–1230. doi:10.1016/S0043-1354(98)00321-2

DeRamus, H.A., 2004. Grazing management of ruminant animals in sustainable agriculture. Outlook on Agriculture 33, 93–100. doi:10.5367/000000004773973073

Desmarais, T.R., Solo-Gabriele, H.M., Palmer, C.J., 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Appl. Environ. Microbiol. 68, 1165–1172. doi:10.1128/AEM.68.3.1165-1172.2002

Dick, L.K., Simonich, M.T., Field, K.G., 2005. Microplate subtractive hybridization to enrich for *Bacteroidales* genetic markers for fecal source identification. Appl. Environ. Microbiol. 71, 3179–3183. doi:10.1128/AEM.71.6.3179-3183.2005

Dombek, P.E., Johnson, L.K., Zimmerley, S.T., Sadowsky, M.J., 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal

sources. Appl. Environ. Microbiol. 66, 2572–2577. doi:10.1128/AEM.66.6.2572-2577.2000

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., Edrington, T.S., 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiology 8, 125. doi:10.1186/1471-2180-8-125

Durso, L. M., Harhay, G. P., Smith, T. P. L., Bono, J. L., DeSantis, T. Z., Harhay, D. M., Andersen, G. L., Keen, J. E., Laegreid, W. W., Clawson, M. L., 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Appl. Environ. Microbiol. 76(14), 4858–4862. http://doi.org/10.1128/AEM.00207-10

Eckburg, P.B., 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638. doi:10.1126/science.1110591

Edberg, Sc., Rice, E.W., Karlin, R.J., Allen, M.J., 2000. *Escherichia coli*: the best biological drinking water indicator for public health protection. J.Appl. Microbiol. 88, 106S–116S.

Feng, P., 1995. *Escherichia coli* serotype O157-H7 - novel vehicles of infection and emergence of phenotypic variants. Emerg. Infect. Dis 1, 47–52.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. 41, 3517–3538. doi:10.1016/j.watres.2007.06.056

Fleisher, J.M., Fleming, L.E., Solo-Gabriele, H.M., Kish, J.K., Sinigalliano, C.D., Plano, L., Elmir, S.M., Wang, J.D., Withum, K., Shibata, T., Gidley, M.L., Abdelzaher, A., He, G., Ortega, C., Zhu, X., Wright, M., Hollenbeck, J., Backer, L.C., 2010. The BEACHES Study: health effects and exposures from non-point source microbial contaminants in

subtropical recreational marine waters. Int. J. Epidemiol. 39, 1291–1298. doi:10.1093/ije/dyq084

Fong, T.T., Lipp, E.K., 2005. Enteric viruses of humans and animals in aquatic environments: Health risks, detection, and potential water quality assessment tools. Microbiol. Mol. Biol. Rev. 69, 357–+. doi:10.1128/MMBR.69.2.357-371.2005

Guan, S., Xu, R., Chen, S., Odumeru, J., Gyles, C., 2002. Development of a procedure for discriminating among *Escherichia coli* isolates from animal and human sources. Appl. Environ. Microbiol. 68, 2690–2698. doi:10.1128/AEM.68.6.2690-2698.2002

Haile, R.W., Witte, J.S., Gold, M., Cressey, R., McGee, C., Millikan, R.C., Glasser, A., Harawa, N., Ervin, C., Harmon, P., Harper, J., Dermand, J., Alamillo, J., Barrett, K., Nides, M., Wang, G.Y., 1999. The health effects of swimming in ocean water contaminated by storm drain runoff. Epidemiology 10, 355–363. doi:10.1097/00001648-199907000-00004

Hartz, A., Cuvelier, M., Nowosielski, K., Bonilla, T.D., Green, M., Esiobu, N., McCorquodale, D.S., Rogerson, A., 2008. Survival potential of and enterococci in subtropical beach sand: implications for water quality managers. J. Environ. Qual. 37, 898. doi:10.2134/jeq2007.0312

Hattori, M., Taylor, T.D., 2009. The human intestinal microbiome: a new frontier of human biology. DNA Res. 16, 1–12. doi:10.1093/dnares/dsn033

Howell, J.M., Coyne, M.S., Cornelius, P.L., 1996. Effect of sediment particle size and temperature on fecal bacteria mortality rates and the fecal coliform/fecal streptococci ratio. J. Environ. Qual. 25, 1216–1220.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., Fulton, R.S., Giglio, M.G., Hallsworth-Pepin, K.,

Lobos, E.A., Madupu, R., Magrini, V., Martin, J.C., Mitreva, M., Muzny, D.M., Sodergren, E.J., Versalovic, J., Wollam, A.M., Worley, K.C., Wortman, J.R., Young, S.K., Zeng, Q., Aagaard, K.M., Abolude, O.O., Allen-Vercoe, E., Alm, E.J., Alvarado, L., Andersen, G.L., Anderson, S., Appelbaum, E., Arachchi, H.M., Armitage, G., Arze, C.A., Ayvaz, T., Baker, C.C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M.J., Bloom, T., Bonazzi, V., Brooks, J.P., Buck, G.A., Buhay, C.J., Busam, D.A., Campbell, J.L., Canon, S.R., Cantarel, B.L., Chain, P.S.G., Chen, I.-M.A., Chen, L., Chhibba, S., Chu, K., Ciulla, D.M., Clemente, J.C., Clifton, S.W., Conlan, S., Crabtree, J., Cutting, M.A., Davidovics, N.J., Davis, C.C., DeSantis, T.Z., Deal, C., Delehaunty, K.D., Dewhirst, F.E., Deych, E., Ding, Y., Dooling, D.J., Dugan, S.P., Dunne, W.M., Durkin, A.S., Edgar, R.C., Erlich, R.L., Farmer, C.N., Farrell, R.M., Faust, K., Feldgarden, M., Felix, V.M., Fisher, S., Fodor, A.A., Forney, L.J., Foster, L., Di Francesco, V., Friedman, J., Friedrich, D.C., Fronick, C.C., Fulton, L.L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M.Y., Goldberg, J.M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Haake, S.K., Haas, B.J., Hamilton, H.A., Harris, E.L., Hepburn, T.A., Herter, B., Hoffmann, D.E., Holder, M.E., Howarth, C., Huang, K.H., Huse, S.M., Izard, J., Jansson, J.K., Jiang, H., Jordan, C., Joshi, V., Katancik, J.A., Keitel, W.A., Kelley, S.T., Kells, C., King, N.B., Knights, D., Kong, H.H., Koren, O., Koren, S., Kota, K.C., Kovar, C.L., Kyrpides, N.C., La Rosa, P.S., Lee, S.L., Lemon, K.P., Lennon, N., Lewis, C.M., Lewis, L., Ley, R.E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C.-C., Lozupone, C.A., Lunsford, R.D., Madden, T., Mahurkar, A.A., Mannon, P.J., Mardis, E.R., Markowitz, V.M., Mavromatis, K., McCorrison, J.M., McDonald, D., McEwen, J., McGuire, A.L., McInnes, P., Mehta, T., Mihindukulasuriya, K.A., Miller, J.R., Minx,

P.J., Newsham, I., Nusbaum, C., O'Laughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S.M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K.S., Pop, M., Priest, M.E., Proctor, L.M., Qin, X., Raes, J., Ravel, J., Reid, J.G., Rho, M., Rhodes, R., Riehle, K.P., Rivera, M.C., Rodriguez-Mueller, B., Rogers, Y.-H., Ross, M.C., Russ, C., Sanka, R.K., Sankar, P., Sathirapongsasuti, J.F., Schloss, J.A., Schloss, P.D., Schmidt, T.M., Scholz, M., Schriml, L., Schubert, A.M., Segata, N., Segre, J.A., Shannon, W.D., Sharp, R.R., Sharpton, T.J., Shenoy, N., Sheth, N.U., Simone, G.A., Singh, I., Smillie, C.S., Sobel, J.D., Sommer, D.D., Spicer, P., Sutton, G.G., Sykes, S.M., Tabbaa, D.G., Thiagarajan, M., Tomlinson, C.M., Torralba, M., Treangen, T.J., Truty, R.M., Vishnivetskaya, T.A., Walker, J., Wang, L., Wang, Z., Ward, D.V., Warren, W., Watson, M.A., Wellington, C., Wetterstrand, K.A., White, J.R., Wilczek-Boney, K., Wu, Y., Wylie, K.M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B.P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J.D., Birren, B.W., Gibbs, R.A., Highlander, S.K., Methe, B.A., Nelson, K.E., Petrosino, J.F., Weinstock, G.M., Wilson, R.K., White, O., 2012. Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. doi:10.1038/nature11234

Johnson, L.K., Brown, M.B., Carruthers, E.A., Ferguson, J.A., Dombek, P.E., Sadowsky, M.J., 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70, 4478–4485. doi:10.1128/AEM.70.8.4478-4485.2004

Kaper, J.B., Nataro, J.P., Mobley, H.L., 2004. Pathogenic *Escherichia coli*. Nature Reviews Microbiology 2, 123–140.

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., Isaacson, R.E., 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. Veterinary Microbiol. 153, 124–133. doi:10.1016/j.vetmic.2011.05.021

Kolpin, D.W., Furlong, E.T., Meyer, M.T., Thurman, E.M., Zaugg, S.D., Barber, L.B., Buxton, H.T., 2002. Pharmaceuticals, hormones, and other organic wastewater contaminants in US streams, 1999-2000: A national reconnaissance. Environ. Sci. Technol. 36, 1202–1211.

Kreader, C.A., 1998. Persistence of PCR-detectable *Bacteroides* distasonis from human feces in river water. Appl. Environ. Microbiol. 64, 4103–4105.

Kreader, C.A., 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. Appl. Environ. Microbiol. 62, 1102–1106.

Kreader, C.A., 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. 61, 1171–1179.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., Sayler, G., 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

Lessard, E.J., Sieburth, J.M., 1983. Survival of natural sewage populations of enteric bacteria in diffusion and batch chambers in the marine environment. Appl. Environ. Microbiol. 45, 950–959.

Lin, C.Z., Raskin, L., Stahl, D.A., 1997. Microbial community structure in gastrointestinal tracts of domestic animals: Comparative analyses using rRNA-targeted oligonucleotide probes. FEMS Microbiol. Ecol. 22, 281–294. doi:10.1111/j.1574-6941.1997.tb00380.x

Mackay, I.M., Arden, K.E., Nitsche, A., 2002. Real-time PCR in virology. Nucleic Acids Res. 30, 1292–1305. doi:10.1093/nar/30.6.1292

Martellini, A., Payment, P., Villemur, R., 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Res. 39, 541–548. doi:10.1016/j.watres.2004.11.012

Martinezmurcia, A., Collins, M., 1991. *Enterococcus Sulfureus*, a New Yellow-Pigmented *Enterococcus* Species. FEMS Microbiol. Lett. 80, 69–73.

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

McLellan, S.L., Daniels, A.D., Salmore, A.K., 2001. Clonal populations of thermotolerant *Enterobacteriaceae* in recreational water and their potential interference with fecal *Escherichia coli* counts. Appl. Environ. Microbiol. 67, 4934–4938. doi:10.1128/AEM.67.10.4934-4938.2001

Miller, T., Wolin, M., 1986. Methanogens in human and animal intestinal tracts. Syst. Appl. Microbiol. 7, 223–229.

Moellering, R., 1992. Emergence of Enterococcus as a Significant Pathogen. Clin. Infect. Dis. 14, 1173–1176.

Nayak, B., Weidhaas, J., Harwood, V. J. (2015). LA35 poultry fecal marker persistence is correlated with that of indicators and pathogens in environmental waters. Appl. Environ. Microbiol. 81(14), 4616–4625. http://doi.org/10.1128/AEM.00444-15

Okabe, S., Okayama, N., Savichtcheva, O., Ito, T., 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

Ostrolenk, M., Kramer, N., Cleverdon, R., 1947. Comparative studies of enterococci and *Escherichia coli* as indices of pollution. J. Bacteriol. 53, 197–203.

Pell, A.N., 1997. Manure and microbes: Public and animal health problem? J. Dairy Sci. 80, 2673–2681.

Rajilić-Stojanović, M., Smidt, H., de Vos, W.M., 2007. Diversity of the human gastrointestinal tract microbiota revisited. Environ. Microbiol. 9, 2125–2136. doi:10.1111/j.1462-2920.2007.01369.x

Roche, L.M., Kromschroeder, L., Atwill, E.R., Dahlgren, R.A., Tate, K.W., 2013. Water quality conditions associated with cattle grazing and recreation on national forest lands. PLoS ONE 8, e68127. doi:10.1371/journal.pone.0068127

Rompre, A., Servais, P., Baudart, J., de-Roubin, M.R., Laurent, P., 2002. Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. J. Microbiol. Methods 49, 31–54. doi:10.1016/S0167-7012(01)00351-7

Roslev, P., Bukh, A.S., 2011. State of the art molecular markers for fecal pollution source tracking in water. Appl. Microbiol. Biotechnol. 89, 1341–1355. doi:10.1007/s00253-010-3080-7

Ryu, H., Henson, M., Elk, M., Toledo-Hernandez, C., Griffith, J., Blackwood, D., Noble, R., Gourmelon, M., Glassmeyer, S., Santo Domingo, J.W., 2013. Development of quantitative PCR assays targeting the 16S rRNA genes of *enterococcus* spp. and their

application to the identification of *enterococcus* species in environmental samples. Appl. Environ. Microbiol. 79, 196–204. doi:10.1128/AEM.02802-12

Schneider, S., Enkerli, J., Widmer, F., 2009. A generally applicable assay for the quantification of inhibitory effects on PCR. Journal of Microbiological Methods 78, 351–353. doi:10.1016/j.mimet.2009.06.010

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial source tracking: current methodology and future directions. Appl. Environ. Microbiol. 68, 5796–5803. doi:10.1128/AEM.68.12.5796-5803.2002

Sinton, L.W., Braithwaite, R.R., Hall, C.H., Mackenzie, M.L., 2007. Survival of indicator and pathogenic bacteria in bovine feces on pasture. Appl. Environ. Microbiol. 73, 7917–7925. doi:10.1128/AEM.01620-07

Sinton, L.W., Hall, C.H., Lynch, P.A., Davies-Colley, R.J., 2002. Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Appl. Environ. Microbiol. 68, 1122–1131. doi:10.1128/AEM.68.3.1122-1131.2002

Smith, C.J., Osborn, A.M., 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology: Application of Q-PCR in microbial ecology. FEMS Microbiology Ecology 67, 6–20. doi:10.1111/j.1574-6941.2008.00629.x

Solo-Gabriele, H. M., Wolfert, M. A., Desmarais, T. R., Palmer, C. J., 2000. Sources of *Escherichia coli* in a coastal subtropical environment. Appl. Environ. Microbiol. 66(1), 230–237. http://doi.org/10.1128/AEM.66.1.230-237.2000

Staley, C., Gordon, K.V., Schoen, M.E., Harwood, V.J., 2012. Performance of Two Quantitative PCR Methods for Microbial Source Tracking of Human Sewage and Implications for

Microbial Risk Assessment in Recreational Waters. Appl. Environ. Microbiol. 78, 7317–7326. doi:10.1128/AEM.01430-12

Stephen, A.M., Cummings, J.H., 1980. The microbial contribution to human faecal mass. Journal of Medical Microbiology 13, 45–56.

Stoeckel, D.M., Harwood, V.J., 2007. Performance, Design, and Analysis in Microbial Source Tracking Studies. Appl. Environ. Microbiol. 73, 2405–2415. doi:10.1128/AEM.02473-06

Tallon, P., Magajna, B., Lofranco, C., Leung, K.T., 2005. Microbial indicators of faecal contamination in water: a current perspective. Water Air Soil Pollut 166, 139–166. doi:10.1007/s11270-005-7905-4

Tyagi, P., Edwards, D.R., Coyne, M.S., 2007. Use of selected chemical markers in combination with a multiple regression model to assess the contribution of domesticated animal sources of fecal pollution in the environment. Chemosphere 69, 1617–1624. doi:10.1016/j.chemosphere.2007.05.039

USDA/NASS 2014 State Agriculture Overview for California [WWW Document], n.d. URL http://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=CALIFORNIA (accessed 1.27.16).

U. S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA A440/5-84-002. U. S. Environmental Protection Agency, Washington, D. C.

U. S. Environmental Protection Agency. 2000. Assigning values to nondetected/ non-quantified pesticide residues in human health food exposure assessments.

U. S. Environmental Protection Agency. 2005. Microbial source tracking guide document. U.S. Environmental Protection Agency. EPA/600-R-05-064

U. S. Environmental Protection Agency. 2008. National primary drinking water regulations: revisions to the total coliform rule. EPA-HQ-OW-2008-0878

U. S. Environmental Protection Agency. 2009. Method validation of U. S. Environmental Protection Agency microbiological methods of analysis.

U. S. Environmental Protection Agency. 2010. Method B: *Bacteroidales* in water by quantitative polymerase chain reaction (qPCR) assay.

U. S. Environmental Protection Agency. 2012. Recreational water quality criteria. Office of water 820-F-12-058

U. S. Environmental Protection Agency. 2012b. Basic information about pathogens and indicators in drinking water.

U. S. Environmental Protection Agency. 2009. Review of published studies to characterize relative risks from different sources of fecal contamination in recreational water. EPA 822-R-09-001

Větrovský, T., Baldrian, P., 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. PLoS ONE 8, e57923. doi:10.1371/journal.pone.0057923

Wade, T.J., Calderon, R.L., Brenner, K.P., Sams, E., Beach, M., Haugland, R., Wymer, L., Dufour, A.P., 2008. High sensitivity of children to swim ming-associated gastrointestinal illness - Results using a rapid assay of recreational water quality. Epidemiology 19, 375–383. doi:10.1097/EDE.0b013e318169cc87

Wade, T. J., Sams, E., Brenner, K. P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C. C., Love, D., Li, Q., 2010. Rapidly measured indicators of recreational water quality

and swimming-associated illness at marine beaches: a prospective cohort study. Environ. Health, 9(66), 1–14.

Webster, L. F., Graves, D. A., Eargle, D. A., Chestnut, D. E., Gooch, J. A., Fulton, M. H., 2013. Assessment of animal impacts on bacterial water quality in a South Carolina, USA tidal creek system. Environ. Monitoring and Assessment, 185(9), 7749–7756. http://doi.org/10.1007/s10661-013-3132-4

Wexler, H.M., 2007. *Bacteroides*: the good, the bad, and the nitty-gritty. Clinical Microbiol. Reviews 20, 593–621. doi:10.1128/CMR.00008-07

WHO. 2001. water immediate solution for persistent problems. www.who.int/inf-pr-2001/en/pr2001-12.html

Table 1.1 USEPA recommended water quality criteria (2012)

Criteria Elements	**Estimated illness rate(NGI): 36 per 1,000 primary contact recreators**		**or**	**Estimated illness rate(NGI): 32 per 1,000 primary contact recreators**	
Indicator	**GM (cfu/100 ml)**	**STV (cfu/100 ml)**		**GM (cfu/100 ml)**	**STV (cfu/100 ml)**
Enterococci Marine and Fresh	35	130		30	110
		or			or
E. coli Fresh only	126	410		100	320

Note: GM represents geometric mean value; STV represents statistical threshold value

Table 1.2 The persistence/survival of enterococci in extra-enteric habitats

Habitat	Persistence/Survival	References
Freshwater	Differential survival of enterococci compared to other FIB has been observed	Anderson et al., 2005
Marine water	Survive longer than *E. coli* Decay rates for *E. coli* 25-55h, *Enterococcus* 29-122h	Lessard and Sieburth 1983
Soil	Persists longer than *E. coli*; survive longer than other FIB	Sinton et al., 2007, Byappanahalli and Fujioka, 2004
Sediment	Survive longer in sediments than water	Anderson et al., 2005
Beach sand	Persists longer in moist beach sand and in nearshore and backshore areas, both in fresh and marine beaches	Byappanahalli et al., 2006

Table 1.3 Microbial diversity of the human gut microbiota

SSU rRNA gene sequence based phylogenetic distribution of the human gut microbiota	Reference
Firmicutes (64%), *Bacteroidetes* (23%), *Proteobacteria* (8%), *Actinobacteria* (3%)	Hattori and Taylor, 2009
Firmicutes>*Bacteroidetes*>*Proteobacteria*>*Actinobacteria*>*Fusobacteria*	Rajilic-Stojanovic et al., 2007
Firmicutes and *Bacteroidetes* are the dominate communities	HMP, 2012
(*Firmicutes* & *Bacteroidetes*) > *Proteobacteria* > *Actinobacteria* > *Fusobacteria* > *Verrucomicrobia*	Eckburg et al., 2005

Note: SSU denotes small-subunit

Table 1.4 Examples of human- and other animals-associated molecular markers

(Roslev and Bukh, 2011)

Source association	Marker	Target
Humans	HF134, HF183, BACHum, HuBac, BacH,	*Bacteroidales*
	Human M2, Human M3, Bf, B. theta	
	ADO, DEN	*Bifidobacterium dentium*
	HS-AV, HAdV, HAdV-C, HAdV-F	Adenovirus
	EV, HEV	Enterovirus
	NoVGI, NoVGII	Norovirus
	Humito, Human, HcytB	Mitochondrial DNA
Cattle	CF128, CF193, Cow-Bac, BacR, Cow M2, Cow	*Bacteroidales*
	M3, Rum,	
	BAV, BAdV	Adenovirus
	BEV	Enterovirus
	NoV GIII	Norovirus
	BPyV	Polyomavirus
	Bomito, Bovine, Cow	Mitochondrial DNA
Elk	EF990R, EF447F	*Bacteroidales*
	White tailed deer	Mitochondrial DNA
Dog	DF475, BacCan	*Bacteroidales*
	Dog	Mitochondrial DNA
Goose/Duck	CG-Prev f5, CGOF1-Bac, CGOF2-Bac,	*Bacteroidales*

	Combination of GA9, GB2, GD5, GE3, GE11, GF5, GG11	*E. coli*
	Canada goose	Mitochondrial DNA
Gull	Gull-2	*Catellicoccus marimammalium*
Horse	HoF597	*Bacteroidales*
	Horse	Mitochondrial DNA
Pig	PF163, Pig Bac, Pig-1-Bac, Pig-2-Bac	*Bacteroidales*
	PAV, PAdV, PAdV-3, PAdV-5	Adenovirus
Sheep	CP	*Bacteroidales*
	OAdV	Adenovirus
	NoV GIII	Norovirus
	Ovmito, Sheep	Mitochondrial DNA

Table 1.5 Probabilities of host-specific assays

Test	**Tested Positive**	**Tested Negative**	**Total Probability**
Samples with target feces	TPC	TNI	TPC+TNI
Samples without target feces	TPI	TNC	TPI+TNC
Total Probability	TPC+TPI	TNI+TNC	TPC+TPI+TNI+TNC

TPC is number of samples tested positive correctly, TPI is number of samples tested positive incorrectly, TNC is number of samples tested negative correctly, and TNI is number of samples tested negative incorrectly.

Table 1.6 The inhibitory level for different reagent

Inhibitors	**Concentration**
Bile salts	1-10 ug/ul
EDTA	1 mM
$FeCl_3$	>10 uM
Fulvic acids	0.1 ng/ul
Humic acid	0.1 ng/ul
Hemin and tannic acid	0.1 ng/ul
SDS	0.1 mM

Note: EDTA denotes ethylenediaminetetraacetic acid; SDS denotes sodium dodecyl sulfates.

Chapter 2

Evaluation of host-associated genetic markers for rapid PCR-based identification of fecal contamination sources in water

2.1 Abstract

The water quality of many waterways is deteriorating due to point and nonpoint source pollution from human and animal waste. Accurate identification of contamination sources is essential if we are to develop cost-effective pollution control strategies. The direct detection of host-specific genetic markers by polymerase chain reactions (PCR) has been widely used in identifying sources of fecal contamination in natural waters. In this study, we conducted experiments to validate genetic markers associated with deer/elk, Canada goose, dog, and cattle for bacterial source tracking in Alabama. End-point PCR was performed on 10 raw sewage samples and 133 fecal samples from nine animal species. Our results showed that CowM3, GFD (goose), and deer/elk-associated markers have acceptable specificity and sensitivity, making them suitable for bacterial source tracking studies. However, the dog marker and one of the cattle markers (CowM2) exhibited cross-reactions with other fecal samples. The performance of these host-associated markers in natural waters was evaluated using both end-point and real-time PCR. Human, goose, and dog markers were detected in several water samples by end-point PCR; the human marker and CowM2 marker were also detected by qPCR. Samples collected after a significant rainfall event showed the highest frequency of host-associated marker detection. Both humans and Canada geese contributed to fecal pollution Parkerson Mill Creek.

2.2 Introduction

The water quality of many of the waterways in our state and nation is deteriorating due to contamination by both point and nonpoint source pollution from human and animal wastes. Each year, millions of cases of infectious disease result from swimming and bathing in contaminated water or consumption of shellfish harvested from fecal polluted waters (Shuval, 2003). Parkerson Mill Creek, located in east Alabama, is rated "impaired" because it fails to meet the water quality criteria required to support its designated use as a fish and wildlife stream. In 2007, the creek was included on the Alabama Department of Environmental Management's (ADEM) 303(d) list of impaired waters for pathogens from point and non-point pollution sources. (http://adem.alabama.gov/programs/water/303d.cnt). Urban runoff, pet waste, wildlife, and leaky sewer lines have all been thought to contribute to the high level of *E. coli* concentration (Parkerson Mill Creek Watershed Management Plan, 2010). However, more information is needed to definitively identify the major sources of fecal pollution in order to develop better strategies to protect against the health risks posed by polluted water.

Fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*) and enterococci have been used to indicate fecal pollution and potential human health risks in surface water (USEPA, 1986). However, FIB methods are not specific to any fecal sources and require at least 24 h to obtain results (Boehm et al., 2002; Leecaster and Weisberg, 2001). It is also possible for *E. coli* and enterococci to survive and regrow in sediment after being released into the environment (Anderson et al., 2005; Desmarais et al., 2002; Solo-Gabriele et al., 2000). Given that the accurate identification of contamination sources is essential for developing cost-effective pollution control strategies, increasing interest is now being directed towards the use of library- and cultivation-independent microbial source tracking (MST) methods based on the polymerase

chain reaction (PCR) technique that target host-specific molecular markers. Combining MST methods with end-point PCR and real-time PCR should provide results that reflect the most recent fecal pollution and identify the sources, thus enabling us to develop better pollution control strategies (Wade et al., 2010).

The host-specific genetic markers detected by using PCR assays in MST can be categorized into four major groups: 1) anaerobic bacterial markers (i.e., host-specific *Bacteroidales*) (Bernhard and Field, 2000a, 2000b); 2) bacterial toxin markers (i.e., *E. coli* toxin gene markers) (Scott et al., 2005); 3) viral markers (i. e., adenoviruses and polyomaviruses) (McQuaig et al., 2006; Roslev and Bukh, 2011); and 4) mitochondrial DNA markers (Schill and Mathes, 2008). The organisms targeted by bacterial markers, viral markers and mitochondrial markers consist of prokaryotes, viruses, and eukaryotes, respectively. Members of the order *Bacteroidales* are considered promising fecal indicators with which to monitor microbial water quality due to their host specificity, broad geographic stability, and high abundance in the gastrointestinal tract of warm-blooded animals (Bernhard and Field, 2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006). However, *Bacteroidales* are not present in the feces of every individual member of a species and the concentrations may also vary from one to another. For example, Green et al. (2012) suggested that *Bacteroides* in gulls are scarce and the horizontal transfer of *Bacteroides* from humans to gulls is common. As a result, the goose marker used in the current study targeted another bacterium, namely *Catellicoccus marimammalium.*

The objective of this study was to evaluate host-specific genetic markers targeting animal species of dogs, cattle, geese and deer that are likely to affect the local watershed. Water samples

which were collected from Parkerson Mill Creek were used to determine the performance of these markers in the field.

2.3 Materials and Methods

Sample collection

A total of 133 fecal samples and 10 wastewater samples were collected around three cities in east-central Alabama, Auburn, Opelika, and Montgomery. The fecal specimens represented nine different animal species (cattle, Canada goose, cat, chicken, deer, dog, duck, goat, and horse) likely to affect the watersheds statewide and were collected with sterile wooden spatulas and placed in sterile polyethylene tubes. Ten raw human sewage samples were collected from nearby wastewater treatment plants in sterile 1-liter bottles. All samples were kept on ice and transported to the lab on the day of collection. Sewage samples were centrifuged at 5000 rpm at 4°C for 15 minutes to concentrate the solid materials 10 fold and fecal samples were stored at -80°C until use.

On each of four days during the months of April and May in 2013 (April 12, 19, 26, and May 3rd), environmental water samples were collected from the surface of the water in three different sites in the Parkerson Mill Creek watershed (Fig. 1), a total of 12 samples were collected in sterile 1-liter plastic bottles. In order to extract the bacterial cells from the water, 500 ml of each sample was vacuum filtered through 0.45-µm-pore-size, 47-mm-diameter nitrocellulose membrane filters (Thermo Fisher Scientific, Waltham, MA). The membrane filters were then stored at -20°C prior to DNA extraction.

Enumeration of *E. coli*

The *E. coli* concentrations in the water samples were measured using the membrane filtration method, followed by cultivation on modified mTEC agar (USEPA, 2002). Water

samples were shaken vigorously by hand to distribute the bacteria uniformly before usage. In order to produce 20-80 *E. coli* colonies on the membranes, three sample volumes (1 ml, 3 ml, and 10 ml) were used for each site. In order to spread these small volumes uniformly, phosphate buffered saline (PBS) was used to dilute each water sample and rinse the sides of the funnel. Each water sample was filtered through 0.45-µm-pore-size, 47-mm-diameter nitrocellulose membrane filters (Thermo Fisher Scientific, Waltham, MA), after which sterile forceps were used to aseptically remove the membrane filter from the filter base, and roll it onto the modified mTEC agar. The plates were incubated at 35±0.5°C for the first 2 hours, and then incubated at 44.5±0.2°C for 22-24 h.

Sensitivity and specificity

Presence and absence data generated from end-point PCR assays were used to estimate the sensitivity and specificity of the host-associated genetic markers. Specificity is the ability of a PCR assay to discriminate between the target fecal samples (for example, those from cattle) and those from other animal sources and is expressed as follows:

specificity % = TNC / (TNC + TPI)

where TNC represents the total number of negative samples that tested negative correctly, and TPI is the total number of samples that tested positive incorrectly. The sensitivity of a genetic marker is defined as the ability of a PCR assay to test positive samples correctly from target fecal samples, and is expressed as follows:

sensitivity % = TPC / (TPC + TNI)

where TPC represents the total number of positive samples that tested positive correctly, and TNI is the total number of test negative samples that tested negative incorrectly (Shanks et al., 2010).

DNA extraction

All DNA extractions were performed using the PowerSoil™ DNA Isolation kit (Mo Bio Laboratories, Inc., Carlsbad, CA). Following the manufacturer's instructions, 0.25 g of each fecal sample or 300 µl of concentrated sewage were used for the DNA extraction. DNA from water samples were extracted from membrane filters that had been cut into small pieces prior to extraction. DNA concentrations were quantified using a NanoDrop ND-1000 UV spectrophotometer (Thermo Scientific, Wilmington, USA).

End-point PCR and real-time PCR assays

The primers used in this study are listed in Table 1. End-point PCRs were performed on a $TG_{RADIENT}$ thermal cycler (Whatman Biometra®, Germany). Each 25 µl reaction mixture contained 5 µl of 5x colorless GoTaq® Flexi buffer, 1.5 mM of $MgCl_2$ solution, 0.2 mM of dNTPs, 0.5 µM each of the forward and reverse primers for the genetic markers, 0.4 mg/ml of bovine serum albumin (BSA), 0.08 unit/µl GoTaq® DNA polymerase, 2.0 µl template DNA, and an appropriate volume of PCR grade water. The thermal cycling parameters for each PCR assay were 94°C for 2 min, followed by 30 cycles of 94°C 60 s, 60°C (for the different markers' annealing temperature please see Table 1) 45 s, and 72°C 60 s, then 72°C for 7 min. The end-point PCR products were resolved using 1.5% agarose gel electrophorese and viewed under UV light to verify the absence or presence of the target gene. No template controls (NTC) containing PCR grade water only and positive controls were included in each instrument run for quality control.

Real-time PCR assays (AllBac, HF183, CowM3, and CowM2) were performed using the StepOne real time PCR instrument (Applied Biosystems, NY). The reaction mixture (15 µl) contained 1x SsoAdvanced™ SYBR® Green Supermix (BIO-RAD, CA), 0.7 µg/µl BSA, 0.2

µM of each primer and 5 µl of template DNA. All reactions were performed in duplicate and began with a hold at 95°C for 10 min, followed by 40 cycles of 95°C 15s, 60°C 30 s, and 72°C 30 s. For each set of experiments, a no template control with two replicates was included and a calibration curve with a concentration spanning the range from 10 to 10^6 gene copies per reaction with two replicates was constructed.

Construction of plasmid DNA standards

Plasmid DNA standards were constructed for AllBac, HF183, CowM3, and CowM2 markers. For the plasmid DNA preparation, PCR were performed using each pair of primers according to the conditions described in the previous section. The PCR product was examined by agarose gel electrophoresis and followed by purification using the DNA Clean & Concentrator™ kit (Epigenetics™, CA). Next, the purified PCR products were ligated into the pCR™ 2.1-TOPO® vector cloning system (Invitrogen by Life Technologies, Grand Island, NY). Each clone was sub-cultured on a Luria Broth (LB) plate containing 50 µg/ mL ampicillin and plasmids were extracted from the transformed One Shot® Mach1™-T1^R competent *E. coli* strains. The plasmid DNA was purified with DNA Clean & Concentrator™ prior to sequencing. The cloning products were sequenced using M13 primers on an ABI 3100 DNA Genetic Analyzer (Applied Biosystems, Grand Island, NY). The sequencing results were confirmed by referring to the NCBI website using the nucleotide Basic Local Alignment Search Tool (BLAST). The gene copy numbers for the plasmid were calculated using the equation below. The molecular weight (MW) of TOPO-TA plasmid is 2,486,846 g/mol and the MW of the insert DNA can be calculated via the web site (www.basic.northwestern.edu/biotools/oligocalc.html). The calibration standard curves ranging from 10 to 10^6 gene copies per reaction were prepared using serial dilutions of plasmid DNA extracted from a pure culture of competent *E. coli* strains.

$$Gene\ Copy\ No. = \frac{6.02 \times 10^{23} \times Plasmid\ DNA\ conc.\ (g/\mu l) \times Volume\ used\ in\ qPCR\ (\mu l)}{(MW\ of\ TOPO\ vector + MW\ of\ insert)\ (g/mol)}$$

Data analysis

The amplification efficiencies (AE) were calculated based on the following equation:

$E = 10^{(-1/slope)} - 1$

The related statistical analyses were performed using SAS® 9.3 software. ArcGIS 10.2 software for desk top was used to generate a sampling map for the Parkerson Mill Creek watershed.

2.4 Results

DNA extracts from a total of 133 fecal samples and 10 wastewater samples were analyzed and the results are shown in Table 2. The CowM3 *Bacteroidales* marker exhibited 100% sensitivity and 97.1% specificity, so the assay met the 80% benchmark suggested by the USEPA (2005) for both specificity and sensitivity. However, false positive amplification was also observed in four non-target DNA samples (two sewage and two deer samples). The CowM2 marker was present in 9 of 11 cattle fecal samples, resulting in 81.8% sensitivity. The CowM2 marker cross reacted with 42 non-target fecal DNA samples: 34.6% (n=9) Canada goose, 66.7% (n=6) duck, 50% (n=5) sewage, 71.4% (n=15) dog, and 53.8% (n=7) chicken, resulting in 68.2% specificity.

Although the GFC marker for Canada goose was detected in 84.6% of the goose fecal samples it also exhibited a 100 % cross-reaction with human fecal samples, so in this case, we chose to use the GFD marker in our primer evaluation study instead. The GFD marker was positive in 84.6% and 27.3% of goose and chicken samples, respectively (Fig. 2). The overall specificity and sensitivity of the GFD marker were 97.4% and 84.6%, respectively and although it had some cross reaction with chicken fecal samples (27.3%), it fully distinguished duck

samples. Similarly, the elk marker was positive in 100% of the deer fecal samples, though it also cross reacted with the cattle (3/11) and goat (3/3) samples. The overall specificity and sensitivity of the elk marker were 94.9% and 100%, respectively. The dog marker was detected in 12 out of 22 dog fecal samples, 10 out of 10 sewage samples, and 14 out of 14 horse fecal samples, giving the dog marker the lowest values for specificity and sensitivity, at 80.2% and 54.5%, respectively, of the species tested.

The AllBac genetic marker targeting the general *Bacteroidales* was detected in all 12 environmental water samples (Table 3); the human marker was detected in 6 out of 12 and 7 out of 12 water samples in the end-point and real-time PCR assays, respectively. The Site B samples for all four sampling dates were positive for the human marker, as were the samples collected on April 12 and April 26 at site Q. Similarly, the dog marker was detected in one third (4 out of 12) of the water samples. The marker for Canada goose was detected in 58.3% (7 out of 12) of the water samples, with every sample from site B testing positive for this marker. Two samples from site Q, collected on April 12 and 26, showed positive results for the GFD marker. Neither CowM3 nor elk markers were detected in any of the 12 water samples.

All of the 12 environmental water samples collected from the three locations were positive for *E. coli* (Fig. 3), with concentrations ranging from 225 CFU/100 ml on April 19 at site B to 5200 CFU/100 ml on April 26 at site I. Both Site B and Site Q had significant high concentrations on April 12, probably due to the rainfall on that day. Site I had the largest geometric mean for *E. coli* concentration and site B the lowest. All sites exceeded the USEPA's criterion for recreational water quality (USEPA 2012), which is a geometric mean of 126 CFU/100ml water. Eleven out of the 12 samples also exceeded the USEPA's single sample

maximum for *E. coli* concentration, which is 410 CFU/100ml. The relationship between the *E. coli* concentration and the AllBac marker concentration was weak (results didn't show).

2.5 Discussion

The Canada goose specific genetic markers (GFC and GFD) were chosen for evaluation in our study. The other genetic markers used in our study to identify the sources of fecal pollution were based on *Bacteroidales* and its relatives (Table 1), as the order of *Bacteroidales* is known to be both abundant and common in mammalian feces. However, Lu et al. (2009) characterized the fecal microbial community from Canada goose, suggesting that the majority of the genes sequenced were related to *Clostridia* or *Bacilli* or, to a lesser degree, *Bacteroidetes*. Canada goose *Bacteroidales*-specific genetic markers have also been reported elsewhere (Fremaux et al., 2010), but these genetic markers were not chosen for the current study because although they are relatively temporally stable, they have low sensitivities. In Green et al.'s (2012) study, the GFC and GFD markers targeted were *Catellicoccus marimammalium* and *Helicobater spp*., respectively, with GFC occurring at a higher concentration as more ribosomal operons in *Catellicoccus*. However, the GFC marker failed to distinguish between waste pollution from human and goose samples in the present study, so we discontinued our evaluation of the GFC marker on other fecal samples. This result suggests that genetic markers need to be validated across a range of conditions, even when they appear to be highly specific when initially reported. Here, the GFD marker exhibited a 27.3% cross amplification with chicken samples, which is consistent with Green et al.'s research as GFD was originally developed to detect avian fecal samples. We found that it actually had a higher sensitivity (84.6%) on goose samples than the 68% reported in Green et al.'s original study. Therefore, our results for the GFD marker

support the sensitivity and specificity of PCR assays for identifying Canada goose-associated fecal pollution in freshwater.

CowM2 and CowM3 are both well-developed cattle-associated *Bacteroidales* genetic markers that have been widely used in various MST research studies. Although CowM2 was reported to perform better than CowM3 by Raith et al. (2013), a lower sensitivity for CowM2 (50%) has also been reported elsewhere (Odagiri et al., 2015). However, the applicability of those results to other regions is potentially limited due to factors such as host diet, climate and geographic location. In our study, we found a much lower specificity for CowM2 (68.2%) compared with previous studies, some of which have reported values of over 98% (Ebentier et al., 2013; Raith et al., 2013; Tambalo et al., 2012). There are several possible reasons for this discrepancy. First, geographical differences could affect host-associated *Bacteroidales* markers significantly due to differences in the diet and animal digestive tract physiology. Layton et al. (2006, 2013) found that *Bacteroidales* 16S rRNA gene sequences obtained from pig were more closely related to *Bacteroidales* 16S rRNA gene sequences obtained from humans than to cattle sources, even though pig and cattle are in the same order of Artiodactyla. Second, the evaluation of the same set of samples may produce different results when examined from a presence-absence or quantitative perspective. When there are cross reactions with non-target feces in PCR assays, this is usually at a low level compared with the signal for the target feces and will thus tend to be classified as false positives in end-point PCR evaluations but not in real-time PCR evaluations (Boehm et al., 2013). Third, the decay rates for the host-associated markers in the environment may be different due to their size and function (Rogers et al., 2011). CowM2 targets a 437-bp fragment as encoding an HDIG domain protein involved in energy metabolism and electron transport, while CowM3 targets a 569-bp fragment encoding a sialic acid-specific 9-O-

acetylesterase secretory protein involving cell envelope biosynthesis and the degradation of surface polysaccharides and lipopolysaccharides (Shanks et al., 2006, 2008), so the decay rate of these two proteins in environmental water after release from local animal tracts or the abundance variation of the proteins in fecal samples may explain the discrepancy.

The CowM3 marker, on the other hand, had an overall specificity of 97.1 % and 100% sensitivity, which is consistent with previous studies that reported CowM3 to have both a broader target host distribution and greater stability (Raith et al., 2013; Shanks et al., 2006). The relative abundance of the host-associated genetic markers for CowM3 was 32.6 times greater than the CowM2 marker concentration in the same DNA sample, and this value compares favorably with the results previously reported by Shanks et al. (2010). The amount of target gene in each cell may explain the different target copies detected in the same DNA samples by the different markers. Ridley et al. (2014) pointed out that the CowM2 marker targets a single copy gene involved in energy metabolism. Here we hypothesize that there may be two or more CowM3 target genes involved in cell envelope biosynthesis and the degradation of surface polysaccharides and lipopolysaccharides. This result also indicates that not only is CowM3 more specific, but it also has higher sensitivity and a lower detection limit than CowM2 due to its greater abundance. Thus, it will be necessary to validate the specific genetic marker that will be used in each different geographic location because the performance characteristics may change and will thus affect the evaluation results.

The dog-associated marker DF475F was paired with *Bacteroidales*-specific Bac708R and analyzed in our study against 143 target and non-target DNA samples. Dick et al. (2005), who developed this dog marker, found no cross amplifications with human, cat, cow, pig, chicken, or gull sources. However, they also pointed out that the horizontal transfer of fecal bacteria may

occur among species in close contact, such as humans and their pets, which suggests the potential for cross reactions with the dog marker in human samples. This is probably why our results showed that this primer set amplified 100% of the sewage DNA representing human sources. Since this primer was the first and only dog specific primer that has yet been identified, the similar results for dog primer in Dick et al.'s research suggests the need for future work in this area to optimize the primer and reaction. Elk primer was found to amplify both the cattle and goat samples, which is consistent with previous studies that reported that deer/elk primer could not distinguish between *Bacteroidales* sequences from deer/elk and sheep. Our results suggest that combining the results from CowM2, CowM3, and elk markers should make it possible to distinguish between cattle and deer/elk fecal pollution.

The AllBac genetic marker was designed to target the 16S rRNA genes of *Bacteroides* spp. and provides a rapid direct measurement of fecal contamination in water due to feces from warm-blooded animal sources (Layton et al., 2006). The positive results for the AllBac genetic marker in all the water samples in the present study provides an estimate of the total fecal contamination present in the water samples. The human marker was detected in 6/12 and 7/12 of the water samples using end-point and real-time PCR assays, respectively. All samples from site B has been detected with HF183 marker, which suggests a potential source of human fecal pollution close to this sampling site. However, there was no CowM3 signal detected by either end-point PCR or real-time PCR. The discrepancy results were observed between end-point and real-time PCR assays with CowM2 marker probably due to the cross-reaction of CowM2 marker with non-target feces since the low target concentration was detected. Given the lack of signal detected for the deer/elk marker, the positive signal for CowM2 in the environmental water samples is probably due to the presence of human fecal pollution. The samples that were positive

for human signals also amplified the dog marker; these sites were Site B and Site Q on April 12, and Site B on April 26. Since the dog marker was detected in 100% of the human samples, the positive signal for the dog marker in the water samples was probably due to the presence of human fecal pollution. Similar results for the GFD marker are likely to indicate the presence of fecal pollution from Canada goose; during the sampling season, Canada geese were observed around the sampling site, which is consistent with these results. Interestingly, there was no signal detected due to the CowM3 marker in samples collected from site Q, even though this site is close to the beef teaching center at Auburn University and beef cattle were observed on site. That was probably because site Q was located at the upstream of beef teaching center. Future work may be needed to add more sampling sites locate downstream of beef teaching center. This result suggests the capacity of MST to identify major pollution sources from among many possible sources.

2.6 Conclusions

1. Of all the genetic markers evaluated, the CowM3, GFD and deer/elk genetic markers had acceptable values for both sensitivity and specificity (>80%).
2. The dog marker had the lowest level of sensitivity (54.5%) and 100% false positive signals in horse and human samples. This result indicates that the dog marker is unlikely to be a suitable marker for bacterial source tracking.
3. The CowM2 marker exhibited cross-reactions with human, duck, dog, and Canada goose samples. This may explain the inconsistent results between the CowM3 and CowM2 markers.

4. High concentrations of *E. coli* were found in the environmental water samples from all three sites during the sampling period. Surface runoff after a significant rainfall event was likely an important cause of these high *E. coli* concentrations.
5. Both humans and Canada geese contributed to fecal pollution in Parkerson Mill Creek, especially for site B.

2.7 References

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

Bernhard, A.E., Field, K.G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66, 4571–4574. doi:10.1128/AEM.66.10.4571-4574.2000

Bernhard, A.E., Field, K.G., 2000b. 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. doi:10.1128/AEM.66.4.1587-1594.2000

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E., 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. Environ. Sci. Technol. 36, 3885–3892. doi:10.1021/es020524u

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D., Weisberg, S.B., 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Res. 47, 6812–6828. doi:10.1016/j.watres.2012.12.046

Desmarais, T.R., Solo-Gabriele, H.M., Palmer, C.J., 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Appl. Environ. Microbiol. 68, 1165–1172. doi:10.1128/AEM.68.3.1165-1172.2002

Dick, L.K., Simonich, M.T., Field, K.G., 2005. Microplate subtractive hybridization to enrich for *Bacteroidales* genetic markers for fecal source identification. Appl. Environ. Microbiol. 71, 3179–3183. doi:10.1128/AEM.71.6.3179-3183.2005

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., Edrington, T.S., 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiol. 8, 125. doi:10.1186/1471-2180-8-125

Durso, L.M., Harhay, G.P., Smith, T.P.L., Bono, J.L., DeSantis, T.Z., Harhay, D.M., Andersen, G.L., Keen, J.E., Laegreid, W.W., Clawson, M.L., 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Appl. Environ. Microbiol. 76, 4858–4862. doi:10.1128/AEM.00207-10

Ebentier, D.L., Hanley, K.T., Cao, Y., Badgley, B.D., Boehm, A.B., Ervin, J.S., Goodwin, K.D., Gourmelon, M., Griffith, J.F., Holden, P.A., Kelty, C.A., Lozach, S., McGee, C., Peed, L.A., Raith, M., Ryu, H., Sadowsky, M.J., Scott, E.A., Domingo, J.S., Schriewer, A., Sinigalliano, C.D., Shanks, O.C., Van De Werfhorst, L.C., Wang, D., Wuertz, S., Jay, J.A., 2013. Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. Water Res. 47, 6839–6848. doi:10.1016/j.watres.2013.01.060

Eckburg, P.B., 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638. doi:10.1126/science.1110591

Fremaux, B., Boa, T., Yost, C.K., 2010. Quantitative Real-time PCR assays for sensitive detection of Canada goose-specific fecal pollution in water sources. Appl. Environ. Microbiol. 76, 4886–4889. doi:10.1128/AEM.00110-10

Green, H.C., Dick, L.K., Gilpin, B., Samadpour, M., Field, K.G., 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. Appl. Environ. Microbiol. 78, 503–510. doi:10.1128/AEM.05734-11

Hattori, M., Taylor, T.D., 2009. The human intestinal microbiome: a new frontier of human biology. DNA Res. 16, 1–12. doi:10.1093/dnares/dsn033

Hagerdorn, C., K. Porter, A.H. Chapman. 2003. Bacterial source tracking to identify sources of fecal pollution in the Potomac and Anacosita rivers and Rock creek, Washington, D. C. Final project report. EPA/600-R-05-064

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., Isaacson, R.E., 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. Vet. Microbiol. 153, 124–133. doi:10.1016/j.vetmic.2011.05.021

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., Sayler, G., 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

Layton, B.A., Cao, Y., Ebentier, D.L., Hanley, K., Ballesté, E., Brandão, J., Byappanahalli, M., Converse, R., Farnleitner, A.H., Gentry-Shields, J., Gidley, M.L., Gourmelon, M., Lee, C.S., Lee, J., Lozach, S., Madi, T., Meijer, W.G., Noble, R., Peed, L., Reischer, G.H., Rodrigues, R., Rose, J.B., Schriewer, A., Sinigalliano, C., Srinivasan, S., Stewart, J., Van De Werfhorst, L.C., Wang, D., Whitman, R., Wuertz, S., Jay, J., Holden, P.A., Boehm, A.B., Shanks, O., Griffith, J.F., 2013. Performance of human fecal anaerobe-associated

PCR-based assays in a multi-laboratory method evaluation study. Water Res. 47, 6897–6908. doi:10.1016/j.watres.2013.05.060

Leecaster, M.K., Weisberg, S.B., 2001. Effect of sampling frequency on shoreline microbiology assessments. Mar. Pollut. Bull. 42, 1150–1154. doi:10.1016/S0025-326X(01)00130-8

Lu, J., Domingo, J.W.S., Hill, S., Edge, T.A., 2009. Microbial diversity and host-specific sequences of Canada goose feces. Appl. Environ. Microbiol. 75, 5919–5926. doi:10.1128/AEM.00462-09

McQuaig, S.M., Scott, T.M., Harwood, V.J., Farrah, S.R., Lukasik, J.O., 2006. Detection of human-derived fecal pollution in environmental waters by use of a PCR-based human polyomavirus assay. Appl. Environ. Microbiol. 72, 7567–7574. doi:10.1128/AEM.01317-06

Odagiri, M., Schriewer, A., Hanley, K., Wuertz, S., Misra, P.R., Panigrahi, P., Jenkins, M.W., 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. Sci. Total Environ. 502, 462–470. doi:10.1016/j.scitotenv.2014.09.040

Parkerson Mill Creek Watershed Management Plan. 2010. http://www.auburn.edu/administration/facilities/organization/university-architect/cpsm/campus-planning/landscape-master-plan/documents/PMCPLAN_LR.pdf

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., Holden, P.A., Ebentier, D.L., Jay, J.A., Wang, D., Boehm, A.B., Aw, T.G., Rose, J.B., Balleste, E., Meijer, W.G., Sivaganesan, M., Shanks, O.C., 2013. Comparison of PCR and quantitative real-time PCR methods for

the characterization of ruminant and cattle fecal pollution sources. Water Res. 47, 6921–6928. doi:10.1016/j.watres.2013.03.061

Ridley, C.M., Jamieson, R.C., Hansen, L.T., Yost, C.K., Bezanson, G.S., 2014. Baseline and storm event monitoring of *Bacteroidales* marker concentrations and enteric pathogen presence in a rural Canadian watershed. Water Res. 60, 278–288. doi:10.1016/j.watres.2014.04.039

Rogers, S.W., Donnelly, M., Peed, L., Kelty, C.A., Mondal, S., Zhong, Z., Shanks, O.C., 2011. Decay of bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-amended soils. Appl. Environ. Microbiol. 77, 4839–4848. doi:10.1128/AEM.02427-10

Roslev, P., Bukh, A.S., 2011. State of the art molecular markers for fecal pollution source tracking in water. Appl. Microbiol. Biotechnol. 89, 1341–1355. doi:10.1007/s00253-010-3080-7

Schill, W.B., Mathes, M.V., 2008. Real-time PCR detection and quantification of nine potential sources of fecal contamination by analysis of mitochondrial cytochrome b targets. Environ. Sci. Technol. 42, 5229–5234. doi:10.1021/es800051z

Scott, T.M., Jenkins, T.M., Lukasik, J., Rose, J.B., 2005. Potential use of a host associated molecular marker in *Enterococcus faecium* as an index of human fecal pollution. Environ. Sci. Technol. 39, 283–287. doi:10.1021/es035267n

Shanks, O.C., Santo Domingo, J.W., Lamendella, R., Kelty, C.A., Graham, J.E., 2006. Competitive metagenomic dna hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72, 4054–4060. doi:10.1128/AEM.00023-06

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., Haugland, R.A., 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

Shanks, O.C., White, K., Kelty, C.A., Hayes, S., Sivaganesan, M., Jenkins, M., Varma, M., Haugland, R.A., 2010. Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 76, 1359–1366. doi:10.1128/AEM.02033-09

Shuval, H., 2003. Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment. J. Water Health 1, 53–64.

Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R., Palmer, C.J., 2000. Sources of *Escherichia coli* in a coastal subtropical environment. Appl. Environ. Microbiol. 66, 230–237. doi:10.1128/AEM.66.1.230-237.2000

Tambalo, D.D., Fremaux, B., Boa, T., Yost, C.K., 2012. Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Res. 46, 2891–2904. doi:10.1016/j.watres.2012.02.048

U. S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA A440/5-84-002. U. S. Environmental Protection Agency, Washington, D. C.

U. S. Environmental Protection Agency. 2005. Microbial source tracking guide document. U.S. Environmental Protection Agency. EPA/600-R-05-064

U. S. Environmental Protection Agency. 2002. Method 1603: *Escherichia coli* (*E. coil*) in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC)

U. S. Environmental Protection Agency. 2012. Recreational water quality criteria. Office of Water 820-F-12-058

Wade, T.J., Sams, E., Brenner, K.P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C.C., Love, D., Li, Q., 2010. Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environ. Health 9, 1–14.

Table 2.1 End-point PCR and real-time PCR primers and amplification conditions used in the study

Assay Name	Animal species	Gene Target	primer sequence (5'-3')	Size (bp)	Annealing T (°C)	Reference
AllBac296F AllBac412R	General	General *Bacteroides*	GAGAGGAAGGTCCCCCAC CGCTACTTGGCTGGTTCAG	106	60	Layton et al. 2006
HF183F HF265R	Human	HF8 cluster, HF74	ATCATGAGTTCACATGTCCG TACCCCGCCTACTATCTAATG	82	60	Bernhard and Field, 2000a
CowM3F CowM3R	Cow	HD superfamily hydrolase	CCTCTAATGGAAAATGGATGGTATCT CCATACTTCGCCTGCTAATACCTT	122	60	Shanks et al. 2008
CowM2F CowM2R	Cow	HDIG domain protein	CGGCCAAATACTCCTGATCGT GCTTGTTGCGTTCCTTGAGATAAT	92	63	
GFD F GFD R	Canada goose	Unclassified *Helicobacter spp.*	TCGGCTGAGCACTCTAGGG GCGTCTCTTTGTACATCCCA	123	57	Green et al. 2012
GFC F GFC R	Canada goose	*Catellicoccus marimammalium*	CCCTTGTCGTTAGTTGCCATCATTC GCCCTCGCGAGTTCGCTGC	162	69	
EF447F EF990R	Deer	*Bacteroides* 16S	AATAACACCATCTACGTGTAGA GCCTGTCCAGTGCAATTTAA	663	62	Dick et al. 2005
DF475F Bac708R	Dog	*Bacteroides* 16S	CGCTTGTATGTACCGGTACG CAATCGGAGTTCTTCGTG	251	62	

Table 2.2 Comparison of sensitivities and specificities of the host-associated genetic markers

Fecal source	No.	CowM3	CowM2	GFD	Elk	Dog
Cattle	11	11/11 (100%)	9/11 (81.8%)	0/11 (0%)	3/11 (27.3%)	0/11 (0%)
Canada Goose	26	0/26 (0%)	9/26 (34.6%)	22/26 (84.6%)	0/26 (0%)	0/26 (0%)
Cat	12	0/12 (0%)	0/12 (0%)	0/12 (0%)	0/12 (0%)	0/12 (0%)
Chicken	13	0/13 (0%)	7/13 (53.8%)	3/11 (27.3%)	0/13 (0%)	0/13 (0%)
Deer	26	2/26 (7.7%)	0/26 (0%)	0/26 (0%)	26/26 (100%)	0/26 (0%)
Dog	21	0/21 (0%)	15/21 (71.4%)	0/21 (0%)	0/21 (0%)	12/22 (54.5%)
Duck	9	0/9 (0%)	6/9 (66.7%)	0/8 (0%)	0/8 (0%)	0/8 (0%)
Goat	3	0/3 (0%)	0/3 (0%)	0/3 (0%)	3/3 (100%)	0/3 (0%)
Horse	12	0/12 (0%)	0/12 (0%)	0/14 (0%)	0/14 (0%)	14/14 (100%)
Sewage sample	10	2/10 (20%)	5/10 (50%)	0/10 (0%)	0/10 (0%)	10/10 (100%)

Table 2.3 Detection of host associated genetic markers in water samples collected from Parkerson Mill Creek

(The unit for real-time PCR: $\log_{10}$ copies/100 ml water)

Date	Site	AllBac		HF183		CowM3		CowM2		GFD	Elk	Dog
		PCR	qPCR	PCR	qPCR	PCR	qPCR	PCR	qPCR	PCR	PCR	PCR
12-Apr	I	+	4.49	-	UN	-	UN	-	3.13	+	-	+
	B	+	5.24	+	>6.00	-	UN	-	3.12	+	-	+
	Q	+	4.70	+	1.97	-	UN	-	2.95	+	-	+
19-Apr	I	+	4.22	-	UN	-	UN	-	2.70	-	-	-
	B	+	4.72	+	DNQ	-	UN	-	3.43	+	-	-
	Q	+	4.48	-	DNQ	-	UN	-	2.92	-	-	-
26-Apr	I	+	4.59	-	UN	-	UN	-	3.26	-	-	-
	B	+	5.78	+	4.39	-	UN	-	2.97	+	-	+
	Q	+	4.81	+	DNQ	-	UN	-	2.81	+	-	-
3-May	I	+	4.64	-	UN	-	UN	-	2.87	-	-	-
	B	+	5.05	+	DNQ	-	UN	-	3.01	+	-	-
	Q	+	4.59	-	UN	-	UN	-	2.77	-	-	-

UN signifies undetected and DNQ signifies detected but not quantifiable

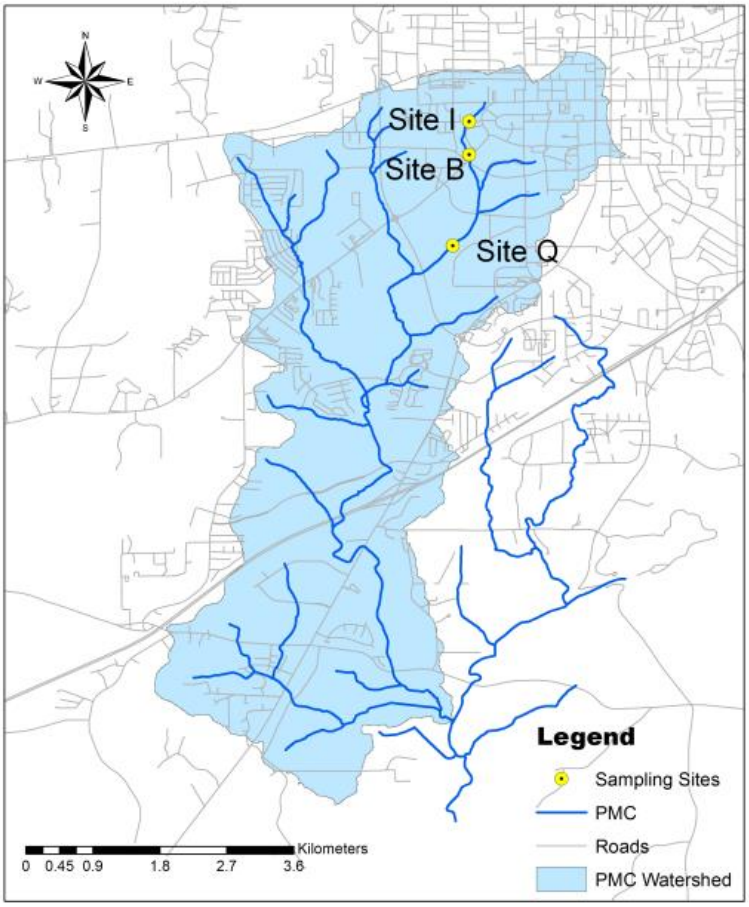


Fig. 2.1 Sampling sites in the Parkerson Mill Creek watershed

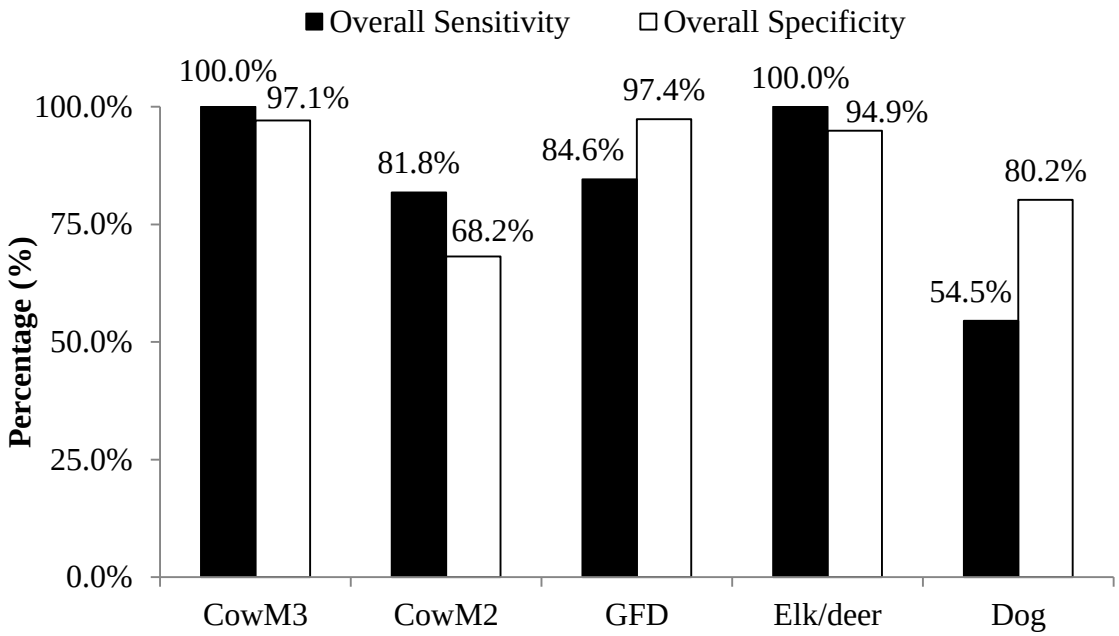


Fig. 2.2 Overall sensitivity and specificity for different markers tested

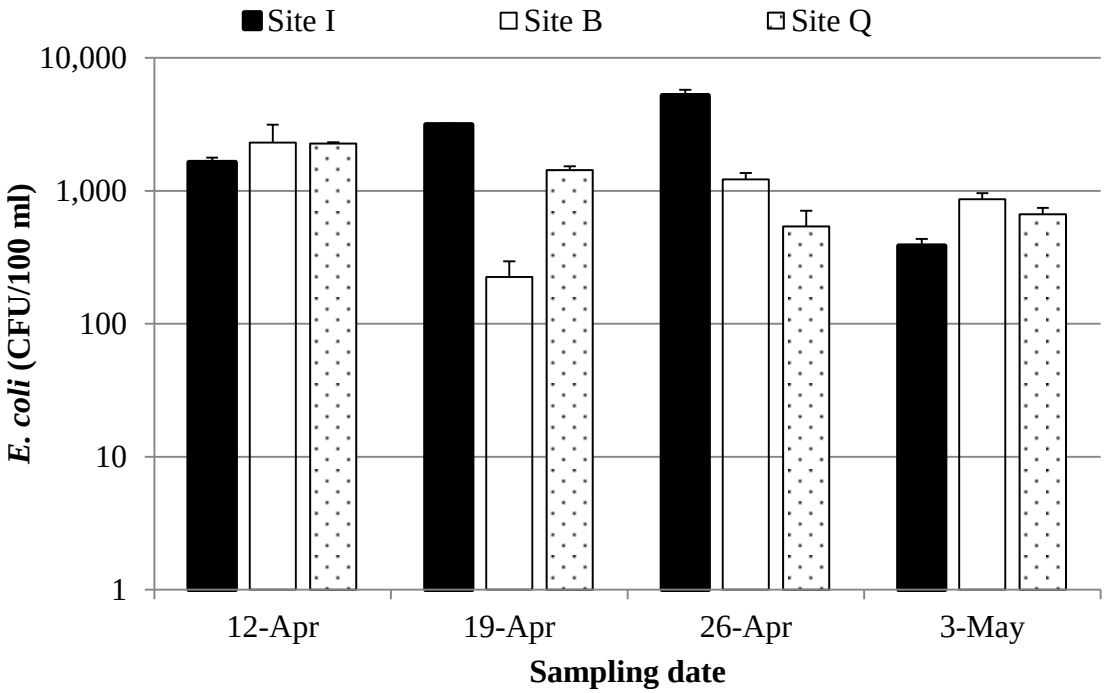


Fig. 2.3 *E. coli* concentrations (CFU/100 ml) at all sampling locations for different dates

Chapter 3

Assessment of the performance of two host-associated *Bacteroidales* PCR assays for tracking cattle fecal contamination in water

3.1 Abstract

End-point and qPCR were used to assess two cattle-associated genetic markers, CowM2 and CowM3, by targeting the *Bacteroidales* genes encoding an HDIG domain protein and a sialic acid-specific 9-O-acetylesterase secretory protein, respectively. A collection of DNA extracts from 143 individual fecal samples representing nine animal species and waste water samples were tested for the CowM2 and CowM3 markers using an end-point PCR assay. The sensitivity and specificity for CowM2 were 81.8% and 68.2%, respectively. CowM2 was found to cross react with Canada Geese (34.6%), duck (66.7%), sewage (50%), canine (71.4%), and chicken (53.8%) samples. In contrast, the CowM3 marker exhibited 100% sensitivity and 97.1% specificity. The CowM2 and CowM3 markers were then applied to environmental water samples (n=12) collected at three different sites along Parkerson Mill Creek in Auburn, AL. CowM2 was quantified in all the water samples, with concentrations ranging between 2.70 and 3.26 $\log_{10}$ copies per 100 ml of water; no amplification was observed for the CowM3 marker. Analytical and process LLOD and LLOQ of the cattle-associated genetic markers in qPCR assays were measured experimentally. These results clearly indicate that for this location, CowM3 exhibits superior performance characteristics compared to CowM2 and highlight the need for robust marker validation before a genetic marker is selected for the assessment of recreational water quality in a specific locality.

3.2 Introduction

Pathogens associated with cattle fecal pollution, such as *E. coli* O157:H7, *Campylobacter spp.*, *Salmonella enterica*, *Cryptosporidium spp.*, and *Giardia lamblia*, can pose a serious threat to human health (Nayak et al., 2015; Webster et al., 2013; Hutchison et al., 2004). Recent studies have suggested that careful consideration is needed for the management of recreational waters that are impacted by cattle sources as well as by human sources (Soller et al., 2010). The United States Department of Agriculture (USDA) has estimated there to be 92 million head of cattle and calves in the United States on Jan. 1st 2016, up 3% from 89.1 million in 2015 (USDA/NASS 2014). Fecal pollution from cattle is usually caused by the inappropriate land application of manure (Pell, 1997), direct contact between livestock and water ways (Burt et al., 2013), and livestock grazing (DeRamus, 2004; Roche et al., 2013). For decades, the microbiological quality of water has been assessed by enumerating fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*) and enterococci in recreational waters (USEPA, 2000a, 2000b). However, FIB are not specific to individual fecal sources and this method requires at least 24 h to obtain results (Boehm et al., 2002; Leecaster and Weisberg, 2001). It is also possible for *E. coli* and enterococci to survive and regrow in the sediment after being released into the environment, so FIB concentrations do not correlate well with levels of bacterial and viral pathogens (Anderson et al., 2005; Desmarais et al., 2002; Solo-Gabriele et al., 2000).

Increasing interest is now being shown in the use of library- and cultivation-independent microbial source tracking (MST) methods based on polymerase chain reaction (PCR) techniques that target host-specific molecular markers (Meays et al., 2004; Scott et al., 2002; Stoeckel and Harwood, 2007; Krentz et al., 2013). Combining these MST methods with end-point PCR and qPCR will provide results that reflect recent fecal pollution and identify the sources, thus making

it possible to develop better strategies to protect humans against the health risks posed by polluted water (Wade et al., 2010; Ryu et al., 2012). Members of the order *Bacteroidales* have been used as a promising fecal indicator to monitor microbial water quality due to its host specificity, broad geographic stability, and abundance (it makes up 20% to 40% of the total fecal bacteria, and about 10% of the fecal mass) in the gastrointestinal tracts of warm-blooded animals (Bernhard and Field, 2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006; Wexler, 2007).

In this study, the cattle-associated *Bacteroidales* genetic markers CowM2 and CowM3 were evaluated to determine their sensitivity, specificity, lower limit of detection (LLOD) and quantification (LLOQ) by using samples collected in Auburn, Alabama (USEPA, 2009). Existing qPCR assays such as the one developed by Shanks et al. (2010) supply information on the concentrations of source-specific fecal pollution, a useful practical application that can be utilized in developing management plans to keep public water supplies safe. However, an effective host-associated genetic marker should ideally be specific and sensitive enough for use in a practical MST application at each watershed of interest, as suggested by Shanks et al. (2010). Several studies have already evaluated the performance characteristics of CowM2 and CowM3 in different areas, with variable results (Boehm et al., 2013; Ebentier et al., 2013; Raith et al., 2013; Tambalo et al., 2012). Hence, local validation is strongly recommended before their application in a particular geographical location since their sensitivity and specificity appear to be subject to geographic constraints (Ebentier et al., 2013; Gawler et al., 2007; Shanks et al., 2008; Tambalo et al., 2012).

As part of the investigation of this issue, a comprehensive multiple-laboratory MST method evaluation study was conducted among twenty-seven labs (Boehm et al., 2013). The

various definitions and interpretations of LLOD and LLOQ applied by the different laboratories affected the method's sensitivity and specificity significantly. A number of studies (Ebentier et al., 2013; Haugland et al., 2005; Kildare et al., 2007; Layton et al., 2013; Raith et al., 2013; Shanks et al., 2008; Tambalo et al., 2012) have proposed methods to measure LLOD and LLOQ using a wide range of different approaches, for example by extrapolating the regression line and designating the lowest point in the calibration curve as LLOQ, or using LOD plus two standard deviations, LOD_{50} (the level at which 50% of the tests are positive), or LOD_{95} (the level at which 95% of the tests are positive) (Stewart et al., 2013). This lack of standardization regarding the measurement of LLOD and LLOQ limits the utility of these MST methods for large scale applications in research laboratories, so the current challenge is to develop a statistically valid and more accurate, practical, and reliable method to determine LLOD and LLOQ in the lab.

Here, we compared the performance characteristics of CowM2 and CowM3 using both end-point and qPCR assays. The markers' applicability in the field was examined by testing environmental water samples. The key objective of this study was to develop an accurate and comparable method for calculating LLOD and LLOQ for each marker in a more practical way. The comparison between different markers not only improves the performance characteristics of this proposed approach, but also will offer a better understanding of the potential threat to local water resources.

3.3 Materials and methods

Sample collection

To determine the host-specificity and sensitivity of cattle-associated genetic markers, a total of 133 individual fecal samples and 10 wastewater samples from wastewater treatment plants were collected in and around three Alabama cities: Auburn, Opelika and Montgomery.

Fecal specimens represented a total of nine animal species likely to affect the local watershed were collected, namely cattle (n=11), Canada geese (n=26), cat (n=12), dog (n=21), horse (n=12), goat (n=3), chicken (n=13), duck (n=9), and deer (n=26). The fecal samples were collected with sterile wooden spatulas and placed in sterile polyethylene tubes. Sewage samples were collected in sterile 1-liter plastic bottles. All samples were immediately stored on ice and transported to the lab on the day of collection. After vigorous shaking, raw sewage samples were transferred to sterile 50 ml centrifuge tubes and centrifuged at 5,000 g at 4°C for 15 minutes to concentrate the solid materials 10 fold. The fecal samples were then stored at -80°C until use. Environmental water samples were collected on four days (April 12th, 19th, 26th, and May 3rd 2013), from three different sites (Site I, Site B, and Site Q) in the Parkerson Mill Creek watershed, in a total of 12 samples. On each occasion the surface water was collected in sterile 1-liter plastic bottles. A portion of each water sample (500 ml) was vacuumed filtered through a 0.45-µm-pore-size, 47-mm-diameter nitrocellulose membrane filter (Thermo Fisher Scientific, Waltham, MA) to collect bacterial cells and these membrane filters were then stored at -20°C prior to DNA extraction.

DNA extraction

All DNA extractions were performed using the PowerSoil^{TM} DNA Isolation kit (Mo Bio Laboratories, Inc., Carlsbad, CA). Following the manufacturer's instructions, 0.25 g of fecal samples or 300 µl of concentrated sewage were used for the DNA extraction. DNA from the water samples were extracted from membrane filters, which were cut into small pieces prior to extraction. The DNA extract yields were quantified with a NanoDrop ND-1000 UV spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

End-point PCR and qPCR assays

Over 430 end-point PCRs were performed on a $TG_{RADIENT}$ thermal cycler (Whatman Biometra®, Germany). Each 25 µl reaction mixture contained 5 µl of 5x colorless GoTaq® Flexi buffer, 1.5 mM of $MgCl_2$ solution, and 0.2 mM of dNTPs, and the forward and reverse primer concentrations for the AllBac and CowM3 markers (Table 1) were 0.5 µM each; for the CowM2 marker, the primer concentration was 0.3 µM, 0.4 mg/ml of bovine serum albumin (BSA), 0.08 unit/µl GoTaq® DNA polymerase, 2.0 µl template DNA, and an appropriate volume of PCR grade water. The thermal cycling parameters for each PCR assay were 94°C for 2 min, followed by 30 cycles of 94°C for 60 s, 60°C (for the different markers' annealing temperatures please see Table 1) for 45 s, and 72°C for 60 s, followed by 72°C for 7 min. The end-point PCR products were examined by 1.5% agarose gel electrophorese and viewed under UV light to verify the absence or presence of the gene target. No template controls (NTC) containing PCR grade water only and positive controls were included in each instrument run for quality control. qPCR assays (AllBac, CowM3, and CowM2) were performed using the StepOne real time PCR instrument (Applied Biosystems, NY); the reaction mixture (15 µl) contained 1x SsoAdvanced™ SYBR® Green Supermix (BIO-RAD, CA), 0.7 µg/µl BSA, 0.2 µM of each primer, and 5 µl of the template DNA. All reactions were performed in duplicate and began with a hold at 95°C for 10 min, followed by 40 cycles of 95°C for 15s, 60°C for 30 s, and 72°C for 30 s. Each set of experiments included a no template control with two replicates and a calibration curve with concentrations spanning the range from 10 to 10^6 gene copies per reaction was constructed in duplicate.

Construction of plasmid DNA standards

Plasmid DNA standards were constructed for the AllBac, CowM3, and CowM2 markers. For the plasmid DNA preparation, PCR using each primer set was performed according to the

conditions described in the previous section. The PCR product was examined by agarose gel electrophoresis and followed by purification using the DNA Clean & Concentrator™ kit (The Epigenetics Company™, CA). The purified PCR products were then ligated into the pCR™ 2.1-TOPO® vector cloning system (Invitrogen by Life Technologies, Grand Island, NY) and each clone was sub-cultured on a Luria Broth (LB) agar plate containing 50 µg/ mL ampicillin, after which the plasmids were extracted from the transformed One Shot® Mach1™-T1^{R} competent *E. coli* strains. The plasmid DNA was purified with DNA Clean & Concentrator™ prior to sequencing. The cloning products were sequenced using M13 primers on an ABI 3100 DNA Genetic Analyzer. The sequencing results were confirmed by searching the NCBI website using the nucleotide Basic Local Alignment Search Tool (BLAST). The gene copy numbers for the plasmid can be calculated using the following equation:

$$Gene\ Copy\ No. = \frac{6.02 \times 10^{23} \times Plasmid\ DNA\ concentration\ (g/\mu l) \times Volume\ used\ in\ qPCR\ (\mu l)}{(MW\ of\ TOPO\ vector + MW\ of\ insert)\ (g/mol)}$$

The molecular weight (MW) of the TOPO-TA plasmid is 2,486,846 g/mol and the MW of the DNA insert can be calculated from the web site (www.basic.northwestern.edu/biotools/oligocalc.html). Calibration standard curves covering the range from 10 to 10^6 gene copies per reaction were prepared using serial dilutions of plasmid DNA extracted from pure cultures of competent *E. coli* strains.

Sensitivity and specificity

Presence and absence data generated from the end-point PCR assays was used to estimate the sensitivity and specificity of the CowM2 and CowM3 assays. Specificity is the ability of a PCR assay to discriminate between the target fecal samples (cattle fecal sample) and other animal sources and is expressed as follows: specificity % = TNC / (TNC + TPI), where TNC

represents the total number of negative samples that test negative correctly, and TPI is the total number of samples that test positive incorrectly. The sensitivity of the genetic marker is defined as the ability of a PCR assay to identify positive samples correctly from target fecal samples, and is expressed as follows: sensitivity % = TPC / (TPC + TNI), where TPC represents the total number of positive samples that test positive correctly, and TNI is the total number of negative samples that test negative incorrectly (USEPA, 2005).

Determination of the lower limit of detection (LLOD) and lower limit of quantification (LLOQ)

Analytical LLOD (A_{LLOD}) refers to the number of gene copies that can reliably be detected in qPCR. Fecal DNA extracted from composite cattle fecal material was used in this study. The composite sample was prepared by combining fecal matter from seven individual animals (5 g each) that was then mixed well in a conical tube. The initial gene copies in the composite fecal DNA sample were quantified using qPCR, then a serial dilution of fecal DNA containing 2, 5, 8, 10, 25, 50, 75, 100, 1000 and 4000 gene copies per reaction (5 µL) were prepared, after which qPCR was performed with four replicates for each concentration. The samples were determined to be positive (detectable) if amplification was observed in at least three of the four replicates. A_{LLOD} was determined for both end-point PCR and qPCR.

Analytical LLOQ (A_{LLOQ}) for both markers was determined by analyzing the composite fecal DNA sample. The same DNA samples that had been used for determining LLOD were also used to determine LLOQ. The standard deviations (SD) of C_T of the replicates for each concentration were used to plot a scatter graph to observe the correlation between gene copies and SD value and a simple linear regression method was used to fit the qPCR testing results. An SD of C_T less than 0.25 was used as the criteria for determining the LLOQ of a qPCR assay

(Applied Biosystems by Life Technologies, 2011). A_{LLOQ} was determined using plasmid DNA in a similar manner.

Process LLOD (P_{LLOD}) refers to the smallest amount of fecal samples that can be subjected to the entire sample preparation process, from dilution in water and filtration through DNA extraction (incorporating the loss of target associated with these manipulations), and still be reliably detected. Fresh composite cattle feces (10 g) was mixed with 1-L distilled water in an autoclaved plastic bottle to obtain fecal slurry at a concentration of 10,000 mg/L. Various volumes of fecal slurry were then added to 200 mL distilled water to obtain a series of concentrations across a range of 0.1 to 100 mg/L. The fecal suspension was filtered through 0.45-µm-pore-size, 47-mm-diameter nitrocellulose membrane and DNA extracted from the filter membrane using PowerSoil DNA Isolation Kit. End-point PCR and qPCR (CowM2 and CowM3) were performed using four replicates. A sample was determined to be positive (detectable) if amplification was observed in at least three of the four replicates.

Process LLOQ (P_{LLOQ}) refers to the smallest amount of fecal matter that can be subjected to the entire sample preparation process, from dilution in water and filtration through DNA extraction (incorporating the loss of target associated with these manipulations), and still can be quantified with a reasonable precision. Amplification must be observed in all four replicate reactions, with an SD of C_T less than 0.25 among all the replicates. Regression analysis was used to observe the relationship between fecal concentration and SD of the C_T value between replicates.

Data analysis

The amplification efficiencies (AE) were calculated based on the following equation: $E= 10^{(-1/slope)} - 1$. The related statistical analyses were performed using SAS® 9.3 software. The

regression analysis used to determine the LLOQ was generated using Microsoft Excel. The results were fitted by a power function ($Y=AX^{\alpha}$), where X represents the number of gene copies, Y is the SD of C_T, and A is a dimensionless parameter. The R^2 value for each regression model was used as an indication of goodness of fit.

3.4 Results

Specificity and sensitivity of the cattle-associated *Bacteroidales* markers

The specificity and sensitivity of the cattle-associated *Bacteroidales* markers were tested on 10 target and 132 non-target fecal DNA samples using end-point PCR (Table 3.2). The CowM3 *Bacteroidales* marker showed 100% sensitivity and 97.1% specificity, hence the CowM3 assay met the 80% benchmark suggested by the USEPA (2005) for both specificity and sensitivity. However, false positive amplification was also observed in four non-target DNA samples (two sewage and two deer samples). The CowM2 marker was present in 9 of the 11 cattle fecal samples, resulting in 81.8% sensitivity. The CowM2 marker cross reacted with 42 non-target fecal DNA samples: 34.6% (n=9) Canada Geese, 66.7% (n=6) duck, 50% (n=5) sewage, 71.4% (n=15) dog, and 53.8% (n=7) chicken, resulting in a 68.2% specificity.

QPCR assays

Two cattle-associated *Bacteroidales* markers were used in the SYBR Green based qPCR assays in this study. Plasmid DNA for CowM2 and CowM3 were prepared to generate a calibration curve for the qPCR. Sequencing results were confirmed on the NCBI website using BLAST. The calibration curve equations and performance characteristics of the qPCR assays are shown in Table 3. The slopes and intercepts for the cattle-associated genetic markers were -3.38 and 33.92 (CowM3) and -3.21 and 39.42 (CowM2), respectively. The R^2 values for all three

calibration curves were ≥ 0.99, and the amplification efficiencies for AllBac, CowM2, and CowM3 were 99.2%, 104.8% and 97.6%, respectively.

A_{LLOD} and A_{LLOQ}

The LLOD was determined when three out of the four replicates at the lowest concentration exhibited amplification. Applying this criterion, the A_{LLOD} vallues for CowM2 and CowM3 were both 5 gene copies per reaction based on the qPCR results (Table 3.4). The relative gene copy numbers for the CowM2 and CowM3 marker in the composite fecal DNA (63.3 ng/μL) were 4.17 and 5.68 $\log_{10}$ gene copies per reaction, respectively. The higher abundance of the CowM3 marker in each DNA sample led to a lower LLOD for CowM3 compared with the CowM2 genetic marker; for CowM3 the LLOD was 0.003 ng composite cattle fecal DNA/ reaction, but 0.108 ng composite cattle fecal DNA/reaction for CowM2. The A_{LLOQ} of the corresponding gene copies for an amplification with an SD of C_T was equal to 0.25 among the four replicates (Fig. 3.1). Here, the SD of C_T between replicates was observed to decrease as the gene copies per reaction increased, gradually reaching a constant level. The relationship between these two parameters was fit well by the power function curve ($Y=AX^{\alpha}$). From the regression results, A_{LLOQ} for CowM2 and CowM3 were 78 and 195 gene copies per reaction, respectively; the R^2 values of the regression lines for CowM2 and CowM3 were 0.9126 and 0.8092, respectively. Although the qPCR results showed CowM3 had a higher A_{LLOQ} compared with CowM2, the actual DNA concentrations for CowM2 and CowM3 were 1.685 and 0.136 ng composite cattle fecal DNA/reaction, respectively. This resulted in an A_{LLOQ} for CowM3 that was 12 times lower than that for CowM2. The plasmid DNA of CowM2 (R^2=0.8549) and CowM3 (R^2=0.8639) was also used to calculate A_{LLOQ} for both genetic markers (Fig. 3.2), with

the A_{LLOQ} being 46 gene copies per reaction for CowM2, and 20 gene copies per reaction for CowM3.

P_{LLOD} and P_{LLOQ}

The process limit of quantification (P_{LLOQ}) is the smallest amount of fecal matter that can be quantified with reasonable precision. Both end-point and qPCR assays were conducted to determine P_{LLOD} and P_{LLOQ}. The minimum amount of fecal matter that could be detected on the membrane filter using end-point PCR assay was 8 mg for CowM2, as 45 µl of DNA was extracted from each filter and a 2 µl DNA template was used in each PCR reaction, so converting all the factors found there to be 356 µg composite feces in each 25 µL reaction. The P_{LLOD} for CowM3, on the other hand, was 0.8 mg fecal matter, which is equivalent to only 35.6 µg fecal matter in each reaction, indicating that for CowM3 the assay was 10 times more sensitive than for CowM2. A process control sample that contained every reagent except for the composite cattle fecal matter and went through all procedures was used for both markers to ensure the sensitivity of the PCR assays.

The results for qPCR show that the P_{LLOD} for CowM2 was 0.4 mg cattle fecal matter on each membrane filter, equivalent to 44.4 µg in each 15 µl reaction; for CowM3 there was 0.02 mg cattle fecal matter, equivalent to 2.22 µg in each 15 µl reaction (Table 3.4). In general, CowM3 had a lower P_{LLOD} than CowM2 in both end-point and qPCR, at 10 and 20 times lower, respectively. Compared with the end-point PCR results, real-time PCR had roughly 8 and 16 times lower detection limits for the CowM2 and CowM3 markers, respectively. The regression analysis method was used to determine P_{LLOQ} for both genetic markers, revealing a P_{LLOQ} for CowM2 of 3.172 mg fecal matter (R^2 =0.946), equivalent to 352.4 µg per reaction (Fig. 3.3); that

for CowM3 was about 10 times lower at 0.308 mg (R^2 =0.9168), equivalent to 34.2 µg per 15 µl reaction.

A large SD means a wider range in the results for the replicates in each treatment and is accompanied by a wide 95% confidence interval. As the range of results exhibited by the replicates overlapped with other treatments, it was not possible to reliably discriminate between those samples with similar gene copies (Figs. 3.5, 3.6, and 3.7). The ranges of the replicates shown in the figures decreased significantly as the DNA concentration increased. When the DNA concentration was above the LLOQ calculated above, the replicates converged and could be confidently discriminated from similar samples. The samples used for calculating P_{LLOQ} (Fig. 3.6) showed larger variations than the DNA samples used for calculating A_{LLOQ} (Figs. 3.5 and 3.7).

Marker performance in environmental water samples

The CowM2 and CowM3 markers were examined for environmental water samples (n=12) collected from three different sites in the Parkerson Mill Creek watershed. Both end-point and qPCR were conducted for all 12 environmental water samples (Table 3.5). The general *Bacteroidales* marker (AllBac) was detected in all the water samples using both PCR assays, with their concentrations ranging between 4.22 and 5.78 $\log_{10}$ copies per 100 ml of water. The coefficient of variation of the replicates for the AllBac marker ranged from 0.4% to 35.8%. Although there was no amplification for the two cattle markers using end-point PCR assays, CowM2 was quantified in all the water samples at concentrations ranging between 2.70 and 3.26 $\log_{10}$ copies per 100 ml of water by using qPCR.

3.5 Discussion

CowM2 and CowM3 are well-developed cattle-associated *Bacteroidales* genetic markers that have been widely used by MST researchers. CowM2 was reported to have a better performance than CowM3 by Raith et al. (2013), but a lower sensitivity for CowM2 (50%) has been reported elsewhere (Odagiri et al., 2015). However, the applicability of those results to other regions may be limited due to factors such as host diet, climate and geographic location. Interestingly, our results indicate a much lower specificity for CowM2 (68.2%) compared with previous studies, which have reported values of over 98% (Ebentier et al., 2013; Raith et al., 2013; Tambalo et al., 2012). There are several possible reasons for this finding. First, geographical differences could affect the host-associated *Bacteroidales* markers significantly due to the different diet and animal digestive tract physiology involved. Research conducted by Dick et al. (2005) showed that *Bacteroidales* spp. specificity reflects animal digestive tract physiology and diet rather than host animal phylogeny, while *Bacteroidales* 16S rRNA gene sequences obtained from the gastrointestinal tract of swine have been shown to be more closely related to *Bacteroidales* 16S rRNA gene sequences obtained from human than bovine samples, even though swine and bovine are in the same order of *Artiodactyla* (Layton et al., 2006). We therefore hypothesized that the false amplifications of the CowM2 marker in Canada Geese, duck, dog, and chicken samples were due to similar food sources for those animals. The evaluation results from the same set of samples may also be different when examined from both presence-absence and quantitative perspectives; where there was a cross reaction with non-target feces in the PCR assays, this was usually at a relatively low concentration level compared with the target feces, so although this would be classified as a false positive in end-point PCR evaluation it may not necessarily do so in a qPCR evaluation (Boehm et al., 2013). Although

qPCR offers several advantages over end-point PCR, end-point PCR remains one of the most widely used PCR assays in field applications, at least partly because it is more cost effective (A to Z PCR; Applied Biosystems by Life Technologies. 2011; Shanks et al., 2008). Third, the decay rate for host-associated markers in the environment may be different due to their size and function (Rogers et al., 2011). CowM2 targets a 437-bp fragment encoding an HDIG domain protein involved in energy metabolism and electron transport, while CowM3 targets a 569-bp fragment encoding a sialic acid-specific 9-O-acetylesterase secretory protein involving cell envelope biosynthesis and the degradation of surface polysaccharides and lipopolysaccharides (Shanks et al., 2006; Shanks et al., 2008), so the decay rates of these two proteins in environmental water after their release from local animal tracts or variations in the abundance of the proteins in fecal samples may explain why this discrepancy exists. Sampling season, sample type, and precipitation may also affect the performance of genetic markers.

CowM3 marker, on the other hand, exhibited overall specificities of 97.1 % and 100% for sensitivity, which is consistent with the findings of previous studies that reported CowM3 as having both a broader target host distribution and greater stability (Raith et al., 2013; Shanks et al., 2006).The relative abundance of the host-associated genetic markers for CowM3 was 32.6 times greater than that of CowM2 marker concentrations in the same DNA samples, comparing favorably with the results previously reported by Shanks et al. (2010). The amount of target gene in each cell may explain the different target copies detected in the same DNA samples by different markers. Ridley et al. (2014) pointed out that the CowM2 marker targets a single copy gene involved in energy metabolism. Our results suggest that there may be two or more CowM3 target genes involved in cell envelope biosynthesis and the degradation of surface polysaccharides and lipopolysaccharides, which means that not only is CowM3 more specific,

but it also has a higher sensitivity and lower detection limit than CowM2 due to this large abundance. Thus, it will be necessary to validate a specific genetic marker in each different geographic location since the performance characteristics may change, affecting the evaluation results.

The test results for the environmental water samples revealed that the AllBac genetic marker that targets the general *Bacteroidales* spp. from a range of mammals (Layton et al., 2006) was detected in all of the water samples by both PCR assays. However the CowM3 signal was detected by neither end-point PCR nor qPCR. The discrepancy in the results of the end-point and qPCR assays for the CowM2 marker is probably due to a cross-reaction between the CowM2 marker and non-target feces since a low target concentration was detected. Even though SYBR Green based qPCR may exhibit lower precision compared with TaqMan based qPCR due to their different operating principles, this disadvantage could be compensated for by analyzing the melting curve for the specific marker amplification so that any false amplification of the primer dimer can be removed; this technique's lower cost also makes it popular with researchers.

Another objective of this study was to develop a comparable method for determining the analytical and process limits of detection and quantification for the practical application of host-associated molecular markers to fecal source identification in the lab. Assessing LLOD and LLOQ for the qPCR based MST method at different levels will help to improve method sensitivity and offer better performance on field applications (Staley et al., 2012). In our study, the results for A_{LLOD}, which refers to the number of gene copies that can reliably be detected in qPCR, were consistent with previously reported values (Lee et al., 2008; Bernhard and Field, 2000b; Lamendella et al., 2007). P_{LLOD}, which refers to the smallest amount of fecal samples that can be subjected to the entire sample preparation process, for both genetic markers indicated a

five order of magnitude higher than A_{LLOD}. Possible explanations for this including: a. P_{LLOD} provides information about the smallest amount of fecal matter rather than the number of DNA molecules; b. *Bacteroidales* only makes up 20% to 40% of the amount of total fecal bacteria, corresponding to about 10% of the fecal mass; c. DNA recovery from filtering and DNA extraction would also increase LLOD due to target loss. Previous research has shown that DNA recovery from the DNA extraction protocol varies from 16.5 % to around 91 %, and the filtration protocol only recovers 79% of the DNA target (Staley et al., 2012; unpublished results from our group). Other factors such as the decay of *Bacteroidales* and any inhibitors present in the fecal samples would also adversely affect P_{LLOD} (Bell et al., 2009). Given that P_{LLOD} provides direct evidence of the quantity of fecal mass in the environmental water, combining A_{LLOD} and P_{LLOD} for a specific genetic marker will not only improve the method's detection limit, but also have important implications for regulators dealing with fecal pollution events in the field.

LLOQ is defined as the level above which quantitative results can be obtained or the lowest reliably detected concentration that meets the criterion for precision (Stewart et al., 2013). Various methods of determining LLOQ have been reported, including signal-to-noise, blank determination, linear regression, and the use of the lowest point of a standard curve as LLOQ (Carden, 1998; Shrivastava and Gupta, 2011; Bustin et al., 2009; Tambalo et al., 2012; Raith et al., 2013). Research has shown that even slight differences in the LLOQ of a specific marker will affect the final report and produce a significant effect on the method's reproducibility (Layton et al., 2013; Ebentier et al., 2013). The USEPA has therefore suggested that each PCR assay should establish a specific LLOQ so that the performance characteristic can be compared across assays (USEPA, 2009). Similarly, other researchers have considered a sample to be quantifiable by applying a C_T with an SD of less than 1 as the criteria when determining the LLOQ of a qPCR

assay (Staley et al., 2012). Here, we chose to use SD=0.25 as the criterion, because to be able to quantify a 2 fold dilution with a 95% confidence, the SD must be less than 0.25 (Applied Biosystems by Life Technologies, 2011) and any samples identified as detected but not quantifiable (DNQ) would create uncertainty in the final results. It has been argued that assigning DNQ as negative rather than positive would increase specificity, but at the cost of reducing the sensitivity (Layton et al., 2013; Schriewer et al., 2013). In the current study, samples that were below A_{LLOQ} contained 0.108 ~ 1.685 ng DNA per reaction for CowM2 and 0.003 ~ 0.136 ng DNA per reaction for CowM3 (Table 3.4). Target copies detected in the DNQ samples were low and a significant difference was observed between detected gene copies and the expected values for each reaction (Fig. 3.4). It was also difficult to determine whether DNQ samples should be assigned as positive or negative without further exploration of the relationship between gene marker concentration and human health risk. However, it should be noted that setting SD =0.25 as the criteria rather than one narrows the gap between LLOD and LLOQ, thus helping to prevent large numbers of samples being classified as DNQ. The data fit in the power function curve very well (the R2 were all above 0.8), providing a mathematical method for calculating the exact quantifiable DNA concentration limit and suggesting that this method provides a more reliable LLOQ with higher precision for qPCR assays.

A box graph based on the SDs of the replicates for each treatment was plotted to further interpret how the method developed in the present study functions (Figs. 3.5, 3.6, and 3.7). Here, a smaller SD indicates that the data points tend to be close to the expected value, while a high value of SD indicates that the data points are spread out over a wider range of values. When there were fewer than 10 target copies in each reaction (Fig. 3.5), the mean value, indicated by the bar inside the box, could not discriminate between them, but a better resolution was observed

when plasmid DNA was used to determine the LLOQ (Fig. 3.7). Compared with plasmid DNA, environmental DNA may contain one or more inhibitory factors that adversely affect the qPCR detection results, for example, humic acids, sodium dodecyl sulfates (SDS), Triton X-100, bile salts (sodium cholate plus deoxycholate), or tannic acids (Kreader, 1996). When feces were used directly in calculating P_{LLOQ} (Fig. 3.6), there is a larger variation compared with the other box plots, probably because the DNA extraction and filtration procedure introduces a larger variation into the final results. However, although the primary advantage of qPCR over end-point PCR is its ability to quantify gene markers with reasonable precision, the relationship between target copies and fecal matter in environmental water remains unclear. Bernhard and Field (2000b) assumed that five 16S rRNA operons are contained in each *Bacteroidales* cell. However simply quantifying the gene copies in each reaction will provide only minimal information for water pollution management. P_{LLOQ}, on the other hand, provides information regarding the smallest amount of fecal matter that can be quantified with reasonable precision. In order to quantify the relative contribution of each fecal source that is present in a body of water, we therefore recommend the use of both A_{LLOQ} and P_{LLOQ} as necessary parameters to assess the utility and validate the use of a specific host-associated marker.

3.6 Conclusions

In general, this study proposes a successful method for creating an accurate LLOQ for a specific host-associated *Bacteroidales* 16S rRNA genetic marker. In combination with environmental water testing results, the validation of two cattle-associated markers against target and non-target fecal samples enables researchers to have more confidence in their measurements. The key findings of this study are as follows:

1. The performance characteristics for two genetic markers were evaluated using end-point PCR. The CowM3 genetic marker showed a higher sensitivity and specificity than the CowM2 marker for this approach, indicating a better field application for CowM3.
2. The limits of detection and the quantification of the two cattle-associated markers using two different methods were also established. The CowM3 marker had generally lower LOD and LOQ values compared with those for the CowM2 marker.
3. The combination of A_{LLOD}, A_{LLOQ} and P_{LLOD}, P_{LLOQ} analysis methods provide more accurate and reliable assessments when using host-associated genetic markers for the prevention of fecal pollution in water.
4. Robust marker validation is vital for each locality when utilizing genetic marker techniques for the assessment of recreational water quality.

3.7 References

Allsop, K., Stickler, D. j., 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. Journal of Applied Bacteriology 58, 95–99. doi:10.1111/j.1365-2672.1985.tb01433.x

Anderson, K.L., Whitlock, J.E., Harwood, V.J., 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Applied and Environmental Microbiology 71, 3041–3048. doi:10.1128/AEM.71.6.3041-3048.2005

Applied Biosystems by Life Technologies. 2011. Real-time PCR: Understanding C_T.

Bell, A., Layton, A.C., McKay, L., Williams, D., Gentry, R., Sayler, G.S., 2009. Factors influencing the persistence of fecal *Bacteroides* in stream water. Journal of Environment Quality 38, 1224. doi:10.2134/jeq2008.0258

Bernhard, A.E., Field, K.G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Applied and Environmental Microbiology 66, 4571–4574. doi:10.1128/AEM.66.10.4571-4574.2000

Bernhard, A.E., Field, K.G., 2000b. Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Applied and Environmental Microbiology 66, 1587–1594. doi:10.1128/AEM.66.4.1587-1594.2000

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E., 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. Environ. Sci. Technol. 36, 3885–3892. doi:10.1021/es020524u

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D., Weisberg, S.B., 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Research 47, 6812–6828. doi:10.1016/j.watres.2012.12.046

Burt, C., Bachoon, D.S., Manoylov, K., Smith, M., 2013. The impact of cattle farming best management practices on surface water nutrient concentrations, faecal bacteria and algal dominance in the Lake Oconee watershed: Cattle farming best management practices. Water and Environment Journal 27, 207–215. doi:10.1111/j.1747-6593.2012.00343.x

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 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

Carden, K.M., 1998. Definition and Procedure for the Determination of the Method Detection Limit. http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=4f1ef92406…e1ed83a0edc5e&rgn=div9&view=text&node=40:23.0.1.1.1.0.1.7.2&idno=40

DeRamus, H.A., 2004. Grazing management of ruminant animals in sustainable agriculture. Outlook on Agriculture 33, 93–100. doi:10.5367/000000004773973073

Desmarais, T.R., Solo-Gabriele, H.M., Palmer, C.J., 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Applied and Environmental Microbiology 68, 1165–1172. doi:10.1128/AEM.68.3.1165-1172.2002

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

genetic markers for fecal source identification. Applied and Environmental Microbiology 71, 3184–3191. doi:10.1128/AEM.71.6.3184-3191.2005

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., Edrington, T.S., 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiology 8, 125. doi:10.1186/1471-2180-8-125

Durso, L.M., Harhay, G.P., Smith, T.P.L., Bono, J.L., DeSantis, T.Z., Harhay, D.M., Andersen, G.L., Keen, J.E., Laegreid, W.W., Clawson, M.L., 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Applied and Environmental Microbiology 76, 4858–4862. doi:10.1128/AEM.00207-10

Ebentier, D.L., Hanley, K.T., Cao, Y., Badgley, B.D., Boehm, A.B., Ervin, J.S., Goodwin, K.D., Gourmelon, M., Griffith, J.F., Holden, P.A., Kelty, C.A., Lozach, S., McGee, C., Peed, L.A., Raith, M., Ryu, H., Sadowsky, M.J., Scott, E.A., Domingo, J.S., Schriewer, A., Sinigalliano, C.D., Shanks, O.C., Van De Werfhorst, L.C., Wang, D., Wuertz, S., Jay, J.A., 2013. Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. Water Research 47, 6839–6848. doi:10.1016/j.watres.2013.01.060

Eckburg, P.B., 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638. doi:10.1126/science.1110591

Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Martin Walters, J., Meijer, W.G., 2007. Validation of host-specific *Bacteriodales* 16S rRNA genes as markers to

determine the origin of faecal pollution in Atlantic rim countries of the European Union. Water Research 41, 3780–3784. doi:10.1016/j.watres.2007.01.028

Hattori, M., Taylor, T.D., 2009. The human intestinal microbiome: a new frontier of human biology. DNA Research 16, 1–12. doi:10.1093/dnares/dsn033

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., Dufour, A.P., 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Research 39, 559–568. doi:10.1016/j.watres.2004.11.011

Hutchison, M. l., Walters, L. d., Avery, S. m., Synge, B. a., Moore, A., 2004. Levels of zoonotic agents in British livestock manures. Letters in Applied Microbiology 39, 207–214. doi:10.1111/j.1472-765X.2004.01564.x

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., Wuertz, S., 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. Water Research, Identifying Sources of Fecal Pollution 41, 3701–3715. doi:10.1016/j.watres.2007.06.037

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., Isaacson, R.E., 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. Veterinary Microbiology 153, 124–133. doi:10.1016/j.vetmic.2011.05.021

Kreader, C.A., 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. Applied and Environmental Microbiology 62, 1102–1106.

Krentz, C.A., Prystajecky, N., Isaac-Renton, J., 2013. Identification of fecal contamination sources in water using host-associated markers. Canadian Journal of Microbiology 59, 210–220. doi:10.1139/cjm-2012-0618

Lamendella, R., Domingo, J.W.S., Oerther, D.B., Vogel, J.R., Stoeckel, D.M., 2007. Assessment of fecal pollution sources in a small northern-plains watershed using PCR and phylogenetic analyses of *Bacteroidetes* 16S rRNA gene: Assessment of fecal pollution sources. FEMS Microbiology Ecology 59, 651–660. doi:10.1111/j.1574-6941.2006.00211.x

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., Sayler, G., 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

Layton, B.A., Cao, Y., Ebentier, D.L., Hanley, K., Ballesté, E., Brandão, J., Byappanahalli, M., Converse, R., Farnleitner, A.H., Gentry-Shields, J., Gidley, M.L., Gourmelon, M., Lee, C.S., Lee, J., Lozach, S., Madi, T., Meijer, W.G., Noble, R., Peed, L., Reischer, G.H., Rodrigues, R., Rose, J.B., Schriewer, A., Sinigalliano, C., Srinivasan, S., Stewart, J., Van De Werfhorst, L.C., Wang, D., Whitman, R., Wuertz, S., Jay, J., Holden, P.A., Boehm, A.B., Shanks, O., Griffith, J.F., 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. Water Research 47, 6897–6908. doi:10.1016/j.watres.2013.05.060

Lee, Y.-J., Molina, M., Santo Domingo, J.W., Willis, J.D., Cyterski, M., Endale, D.M., Shanks, O.C., 2008. Temporal assessment of the impact of exposure to cow feces in two

watersheds by multiple host-specific PCR assays. Applied and Environmental Microbiology 74, 6839–6847. doi:10.1128/AEM.00601-08

Leecaster, M.K., Weisberg, S.B., 2001. Effect of sampling frequency on shoreline microbiology assessments. Mar. Pollut. Bull. 42, 1150–1154. doi:10.1016/S0025-326X(01)00130-8

Meays, C.L., Broersma, K., Nordin, R., Mazumder, A., 2004. Source tracking fecal bacteria in water: a critical review of current methods. Journal of Environmental Management 73, 71–79. doi:10.1016/j.jenvman.2004.06.001

Nayak, B., Weidhaas, J., Harwood, V.J., 2015. LA35 poultry fecal marker persistence is correlated with that of indicators and pathogens in environmental waters. Appl. Environ. Microbiol. 81, 4616–4625. doi:10.1128/AEM.00444-15

Odagiri, M., Schriewer, A., Hanley, K., Wuertz, S., Misra, P.R., Panigrahi, P., Jenkins, M.W., 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. Science of the Total Environment 502, 462–470. doi:10.1016/j.scitotenv.2014.09.040

Pell, A.N., 1997. Manure and microbes: Public and animal health problem? J. Dairy Sci. 80, 2673–2681.

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., Holden, P.A., Ebentier, D.L., Jay, J.A., Wang, D., Boehm, A.B., Aw, T.G., Rose, J.B., Balleste, E., Meijer, W.G., Sivaganesan, M., Shanks, O.C., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Research 47, 6921–6928. doi:10.1016/j.watres.2013.03.061

Ridley, C. M., Jamieson, R. C., Hansen, L. T., Yost, C. K., Bezanson, G. S., 2014. Baseline and storm event monitoring of *Bacteroidales* marker concentrations and enteric pathogen presence in a rural Canadian watershed. Water Research, 60, 278–288. http://doi.org/10.1016/j.watres.2014.04.039

Roche, L.M., Kromschroeder, L., Atwill, E.R., Dahlgren, R.A., Tate, K.W., 2013. Water quality conditions associated with cattle grazing and recreation on national forest lands. PLoS ONE 8, e68127. doi:10.1371/journal.pone.0068127

Rogers, S. W., Donnelly, M., Peed, L., Kelty, C. A., Mondal, S., Zhong, Z., Shanks, O. C., 2011. Decay of bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-amended soils. Applied and Environmental Microbiology, 77(14), 4839–4848. http://doi.org/10.1128/AEM.02427-10

Ryu, H., Griffith, J.F., Khan, I.U.H., Hill, S., Edge, T.A., Toledo-Hernandez, C., Gonzalez-Nieves, J., Santo Domingo, J., 2012. Comparison of gull feces-specific assays targeting the 16S rRNA genes of *catellicoccus marimammalium* and *streptococcus* spp. Applied and Environmental Microbiology 78, 1909–1916. doi:10.1128/AEM.07192-11

Schriewer, A., Goodwin, K.D., Sinigalliano, C.D., Cox, A.M., Wanless, D., Bartkowiak, J., Ebentier, D.L., Hanley, K.T., Ervin, J., Deering, L.A., Shanks, O.C., Peed, L.A., Meijer, W.G., Griffith, J.F., SantoDomingo, J., Jay, J.A., Holden, P.A., Wuertz, S., 2013. Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. Water Research 47, 6909–6920. doi:10.1016/j.watres.2013.03.062

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial source tracking: current methodology and future directions. Applied and Environmental Microbiology 68, 5796–5803. doi:10.1128/AEM.68.12.5796-5803.2002

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., Haugland, R.A., 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

Shanks, O.C., Santo Domingo, J.W., Lamendella, R., Kelty, C.A., Graham, J.E., 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Appl. Environ. Microbiol. 72, 4054–4060. doi:10.1128/AEM.00023-06

Shanks, O.C., White, K., Kelty, C.A., Hayes, S., Sivaganesan, M., Jenkins, M., Varma, M., Haugland, R.A., 2010. Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 76, 1359–1366. doi:10.1128/AEM.02033-09

Shrivastava, A., Gupta, V., 2011. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. Chronicles of Young Scientists 2, 21. doi:10.4103/2229-5186.79345

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J.E., Ashbolt, N.J., 2010. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. 44, 4674–4691. doi:10.1016/j.watres.2010.06.049

Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R., Palmer, C.J., 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

Staley, C., Gordon, K.V., Schoen, M.E., Harwood, V.J., 2012. Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. Applied and Environmental Microbiology 78, 7317–7326. doi:10.1128/AEM.01430-12

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T.-T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K., Weisberg, S.B., 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. Water Research 47, 6829–6838. doi:10.1016/j.watres.2013.04.063

Stoeckel, D.M., Harwood, V.J., 2007. Performance, design, and analysis in microbial source tracking studies. Applied and Environmental Microbiology 73, 2405–2415. doi:10.1128/AEM.02473-06

Tambalo, D.D., Fremaux, B., Boa, T., Yost, C.K., 2012. Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Research 46, 2891–2904. doi:10.1016/j.watres.2012.02.048

USDA/NASS 2014 State Agriculture Overview for California [WWW Document], n.d. URL http://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=CALIFORNIA (accessed 1.27.16).

U. S. Environmental Protection Agency. 2005. Microbial source tracking guide document. U.S. Environmental Protection Agency. EPA/600-R-05-064

U. S. Environmental Protection Agency. 2009. Method validation of U. S. Environmental Protection Agency microbiological methods of analysis.

U. S. Environmental Protection Agency. 2000a. Assigning values to nondetected/ non-quantified pesticide residues in human health food exposure assessments.

U. S. Environmental Protection Agency. 2000b. Improved enumeration methods for the recreational water quality indicators: Enterococci and *Escherichia coli* https://www.michigan.gov/documents/deq/wrd-beaches-rvsdman_449691_7.pdf

Wade, T.J., Sams, E., Brenner, K.P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C.C., Love, D., Li, Q., 2010. Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environmental Health 9, 1–14.

Webster, L.F., Graves, D.A., Eargle, D.A., Chestnut, D.E., Gooch, J.A., Fulton, M.H., 2013. Assessment of animal impacts on bacterial water quality in a South Carolina, USA tidal creek system. Environmental Monitoring and Assessment 185, 7749–7756. doi:10.1007/s10661-013-3132-4

Wexler, H.M., 2007. *Bacteroides*: the good, the bad, and the nitty-gritty. Clinical Microbiology Reviews 20, 593–621. doi:10.1128/CMR.00008-07

Table 3.1 Primers used in qPCR assays to detect host-associated *Bacteroidales*

Assay Name	Target	Sequence (5'-3')	Size (bp)	Annealing temperature (°C)	Reference
AllBac296F	General	GAGAGGAAGGTCCCCCAC	106	60	Layton et al. 2006
AllBac412R	*Bacteroidales*	CGCTACTTGGCTGGTTCAG			
CowM3F	Cow	CCTCTAATGGAAAATGGATGGTATCT	122	60	Shanks et al. 2008
CowM3R	*Bacteroidales*	CCATACTTCGCCTGCTAATACCTT			
CowM2F	Cow	CGGCCAAATACTCCTGATCGT	92	63	
CowM2R	*Bacteroidales*	GCTTGTTGCGTTCCTTGAGATAAT			

Table 3.2 Comparison of sensitivities and specificities of the cattle-associated *Bacteroidales* genetic markers

Fecal source	No.	CowM3	CowM2
Cattle	11	11/11 (100%)	9/11 (81.8%)
Canada Geese	26	0/26 (0%)	9/26 (34.6%)
Cat	12	0/12 (0%)	0/12 (0%)
Chicken	13	0/13 (0%)	7/13 (53.8%)
Deer	26	2/26 (7.7%)	0/26 (0%)
Dog	21	0/21 (0%)	15/21 (71.4%)
Duck	9	0/9 (0%)	6/9 (66.7%)
Horse	12	0/12 (0%)	0/12 (0%)
Goat	3	0/3 (0%)	0/3 (0%)
Sewage sample	10	2/10 (20%)	5/10 (50%)
Overall Sensitivity	11 (target)	11/11(100%)	9/11 (81.8%)
Overall Specificity	132 (Non-target)	4/132 (97.1%)	42/132 (68.2%)

Table 3.3 Performance characteristics of qPCR assays

Assay	**Reagent**	**Calibration Equation**	**Amplification Efficiency (%)**	**R^2**
AllBac	SYBR	Y=34.87-3.34x	99.21	1.00
CowM2	SYBR	Y=39.42-3.21x	104.77	0.99
CowM3	SYBR	Y=33.92-3.38x	97.63	1.00

Table 3.4 Lower limits of detection (LLOD) and quantification (LLOQ) for CowM2 and CowM3 markers at the analytical and process levels for qPCR assays

Genetic markers		Analytical limits		Process limits	
		Marker copies per reaction (Gene copies)	DNA concentration (ng/reaction)	Amount of fecal matter on each filter (mg)	DNA concentration (µg/reaction)
CowM2	LLOD	5	0.108	0.400	44.4
	LLOQ	78	1.685	3.172	352.4
CowM3	LLOD	5	0.003	0.020	2.2
	LLOQ	195	0.136	0.308	34.2

Table 3.5 Field samples tested with three markers using end-point and qPCR

(Unit for qPCR: $\log_{10}$ copies/100 ml water)

Markers		AllBac			CowM3		CowM2	
		end-point PCR	qPCR	CV%	end-point PCR	qPCR	end-point PCR	qPCR
12-Apr	I	+	4.49	3.98	-	ND	-	3.13
	B	+	5.24	35.79	-	ND	-	3.12
	Q	+	4.70	3.49	-	ND	-	2.95
19-Apr	I	+	4.22	11.51	-	ND	-	2.70
	B	+	4.72	0.42	-	ND	-	3.43
	Q	+	4.48	2.52	-	ND	-	2.92
26-Apr	I	+	4.59	11.32	-	ND	-	3.26
	B	+	5.78	1.05	-	ND	-	2.97
	Q	+	4.81	5.49	-	ND	-	2.81
3-May	I	+	4.64	6.75	-	ND	-	2.87
	B	+	5.05	13.82	-	ND	-	3.01
	Q	+	4.59	3.83	-	ND	-	2.77

Note: ND denotes “Undetected”.

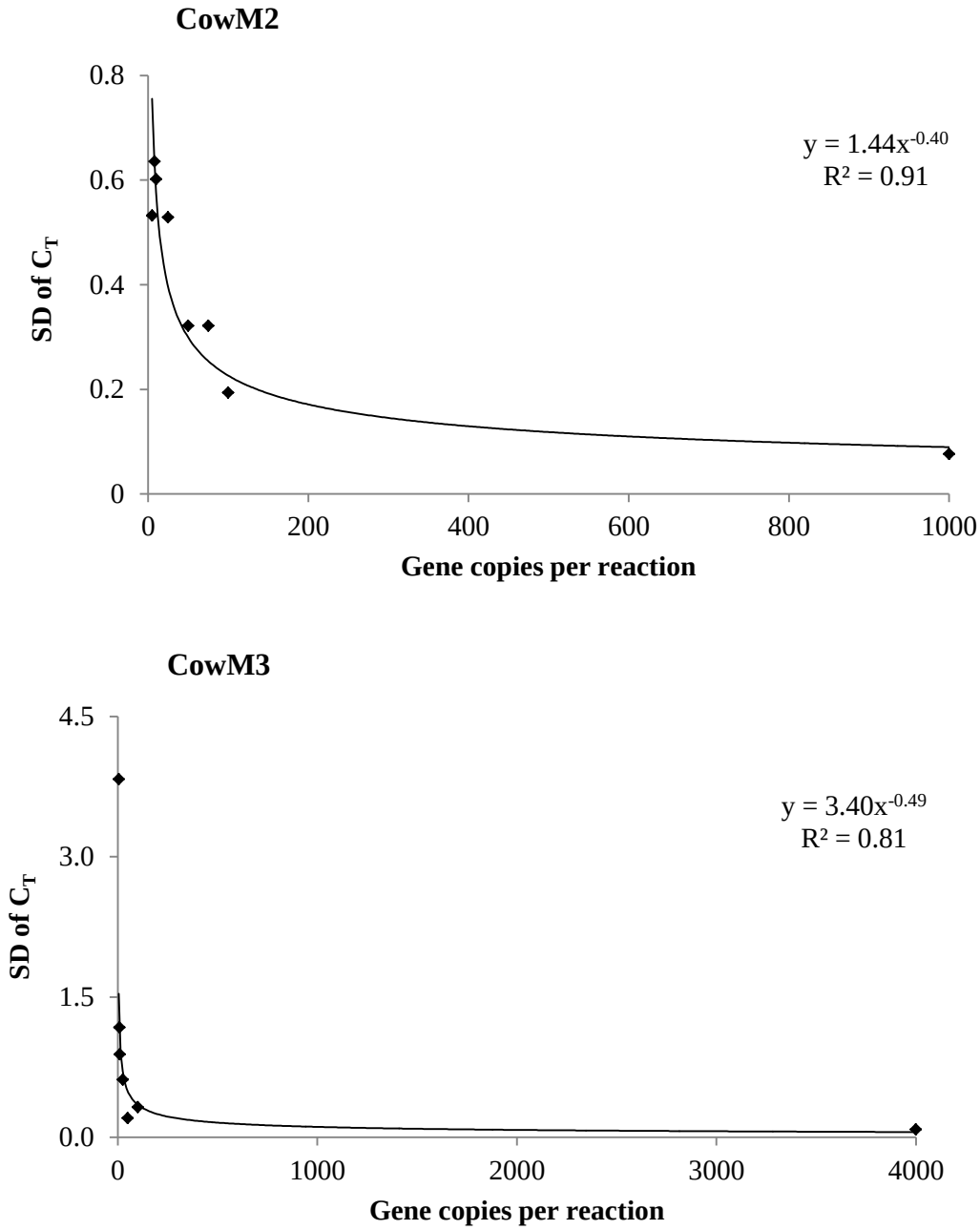


Fig. 3.1 Relationship between standard deviation (SD) of cycle threshold (C_T) among replicates and gene copies of the composite fecal DNA for CowM2 and CowM3 genetic markers.

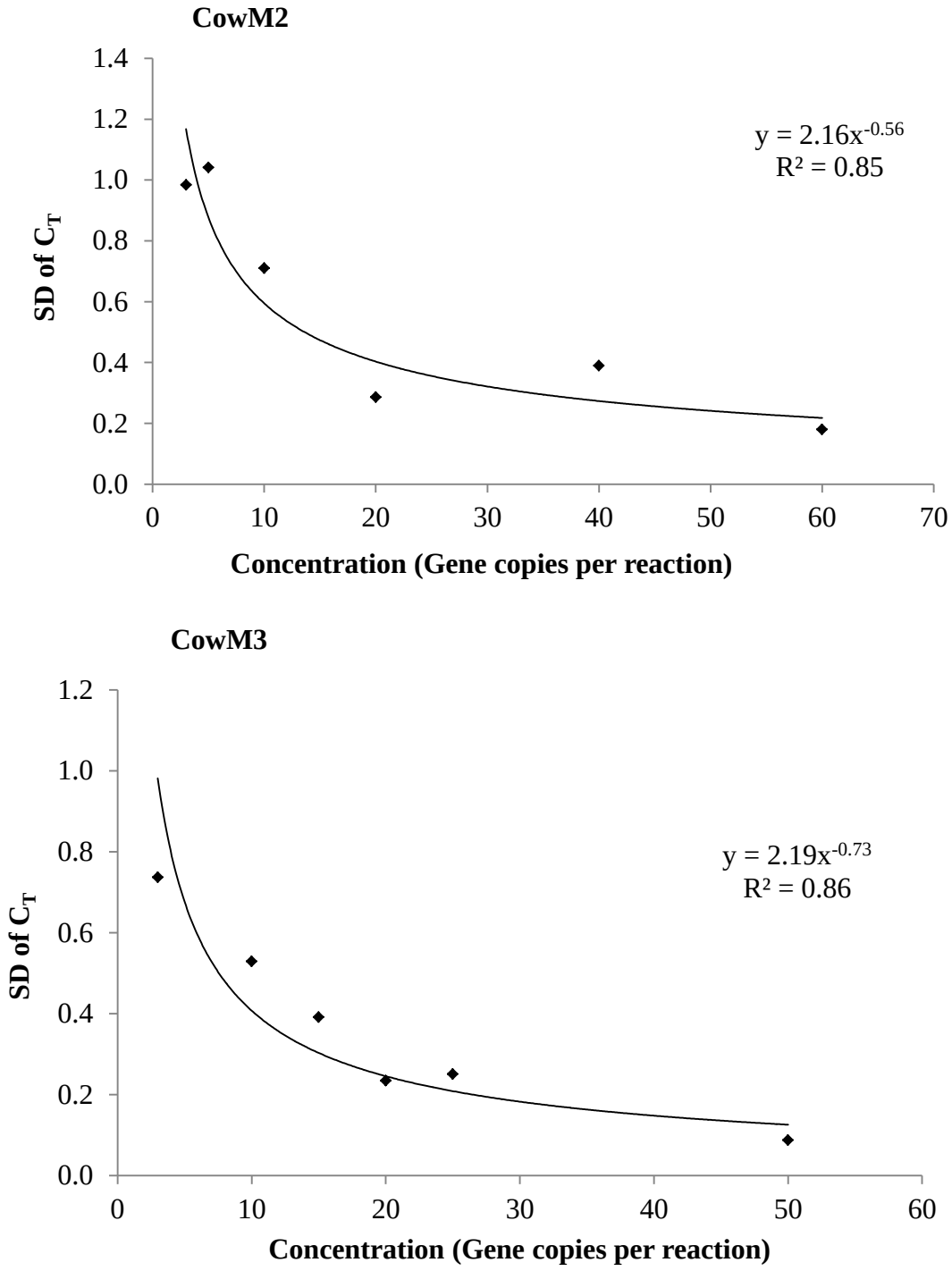


Fig. 3.2 Relationship between standard deviation (SD) of cycle threshold (C_T) among replicates and gene copies concentration of the plasmid DNA for CowM2 and CowM3 genetic markers.

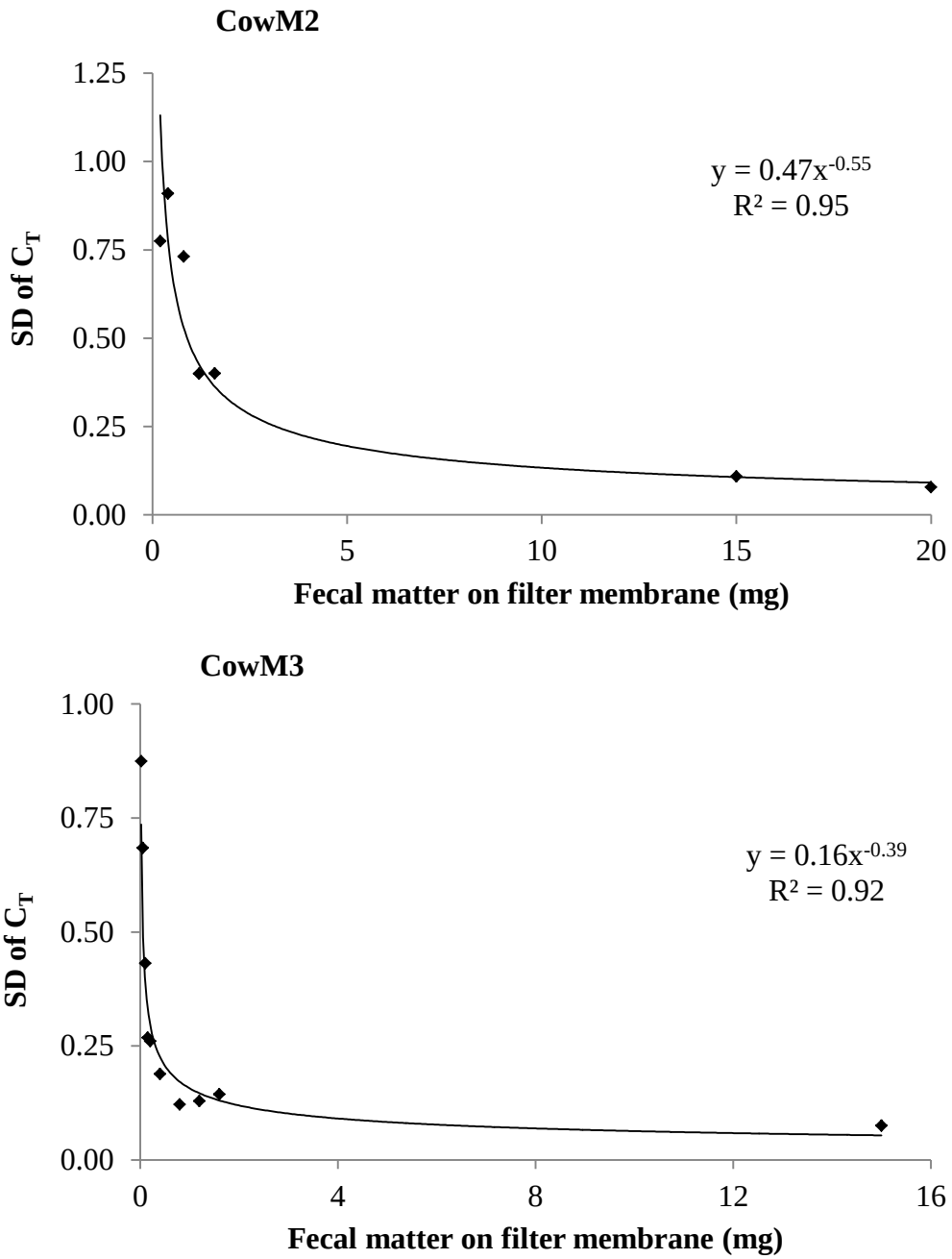


Fig. 3.3 Relationship between standard deviation (SD) of cycle threshold (C_T) among replicates and the amount of fecal matter on each membrane filter for CowM2 and CowM3 genetic markers.

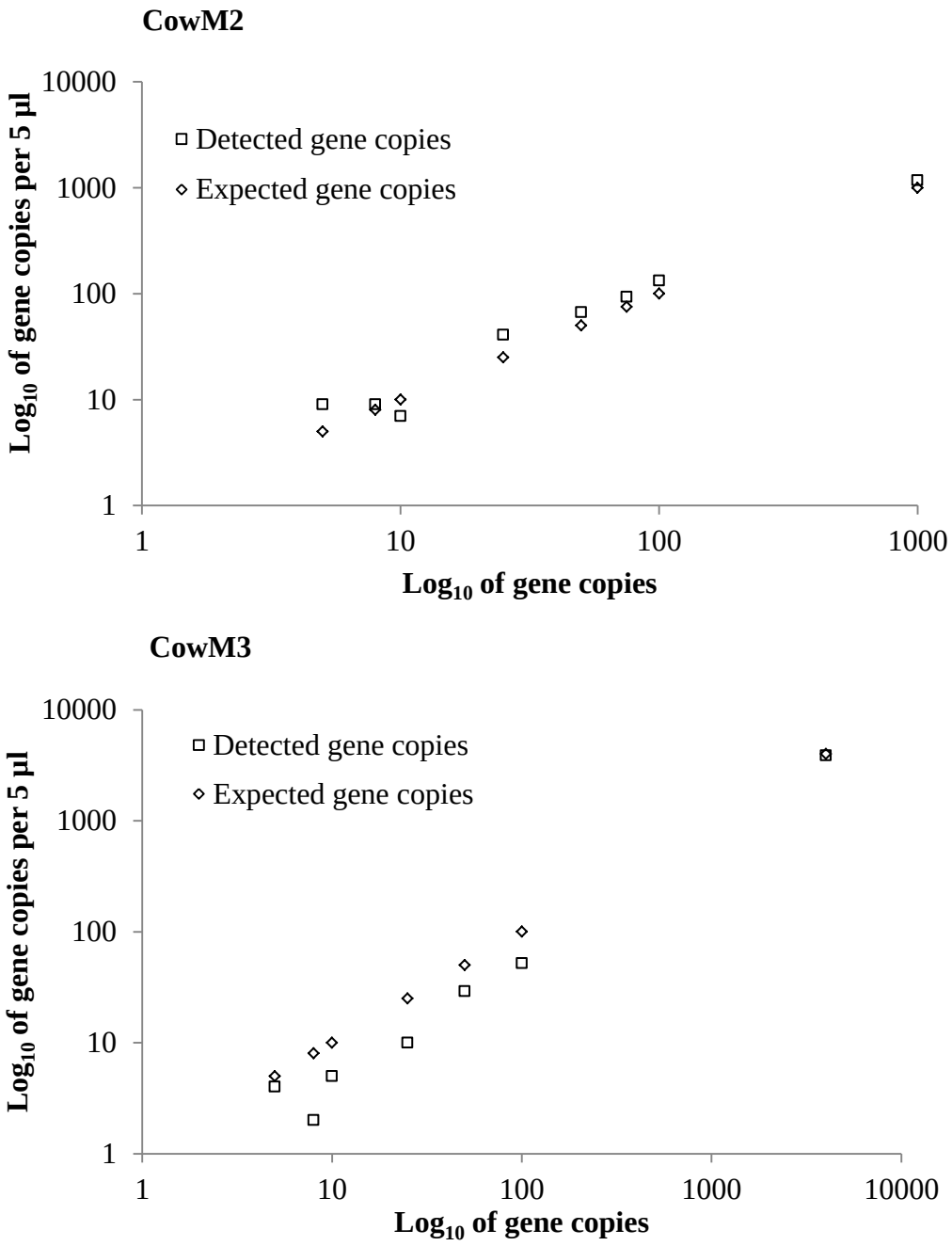


Fig. 3.4 Enumeration of cattle-associated genetic marker in fecal DNA by comparing the expected gene copies contained in each reaction with detected gene copies in each reaction

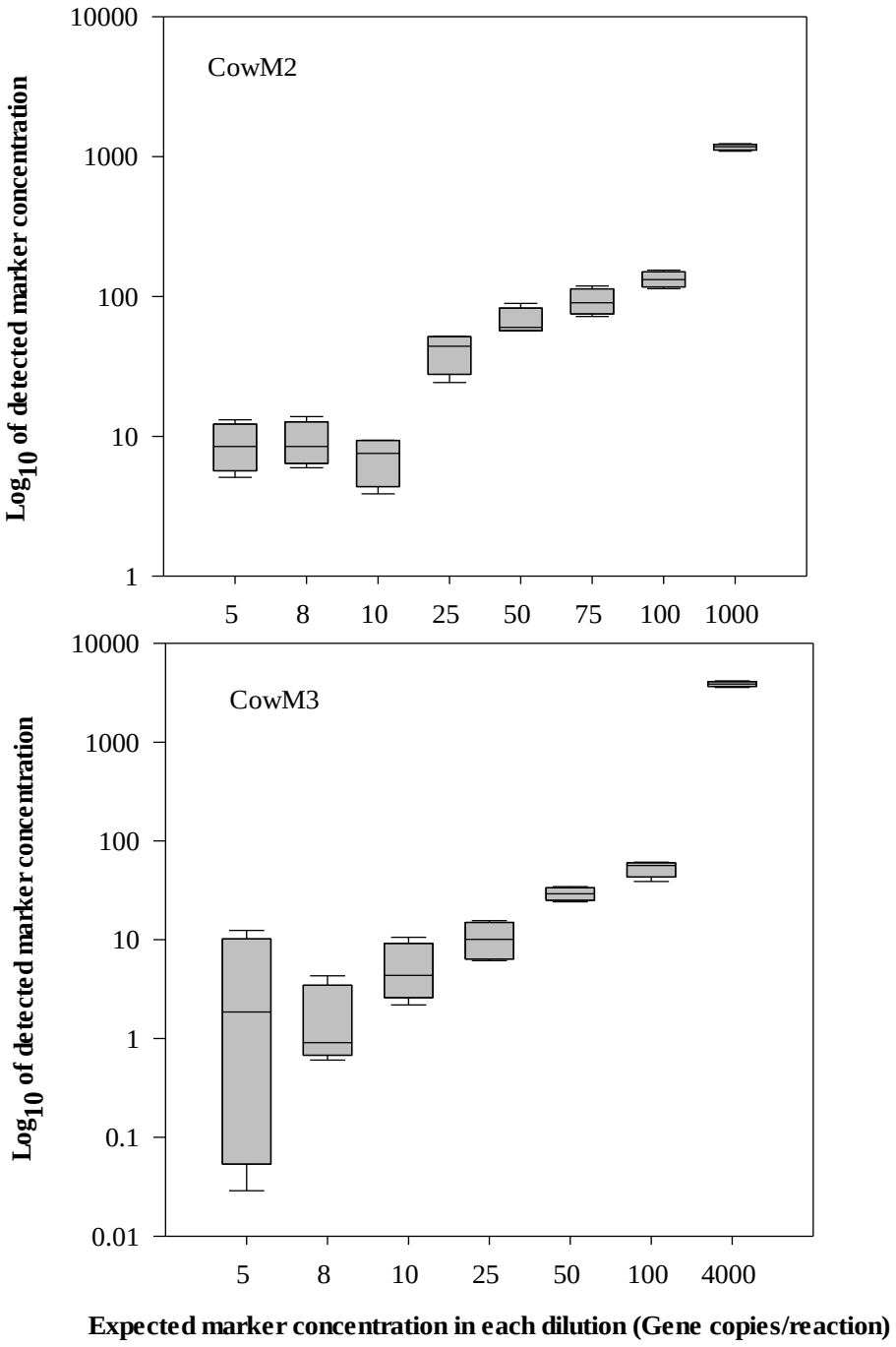


Fig. 3.5 Box plot showing marker concentration detected in each reaction of fecal DNA sample for CowM2 and CowM3 genetic markers. Whisker caps represent the maximum and minimum concentration of the markers. The bar inside box represents the mean concentration and box lengths represent 5^{th}/95^{th} percentiles.

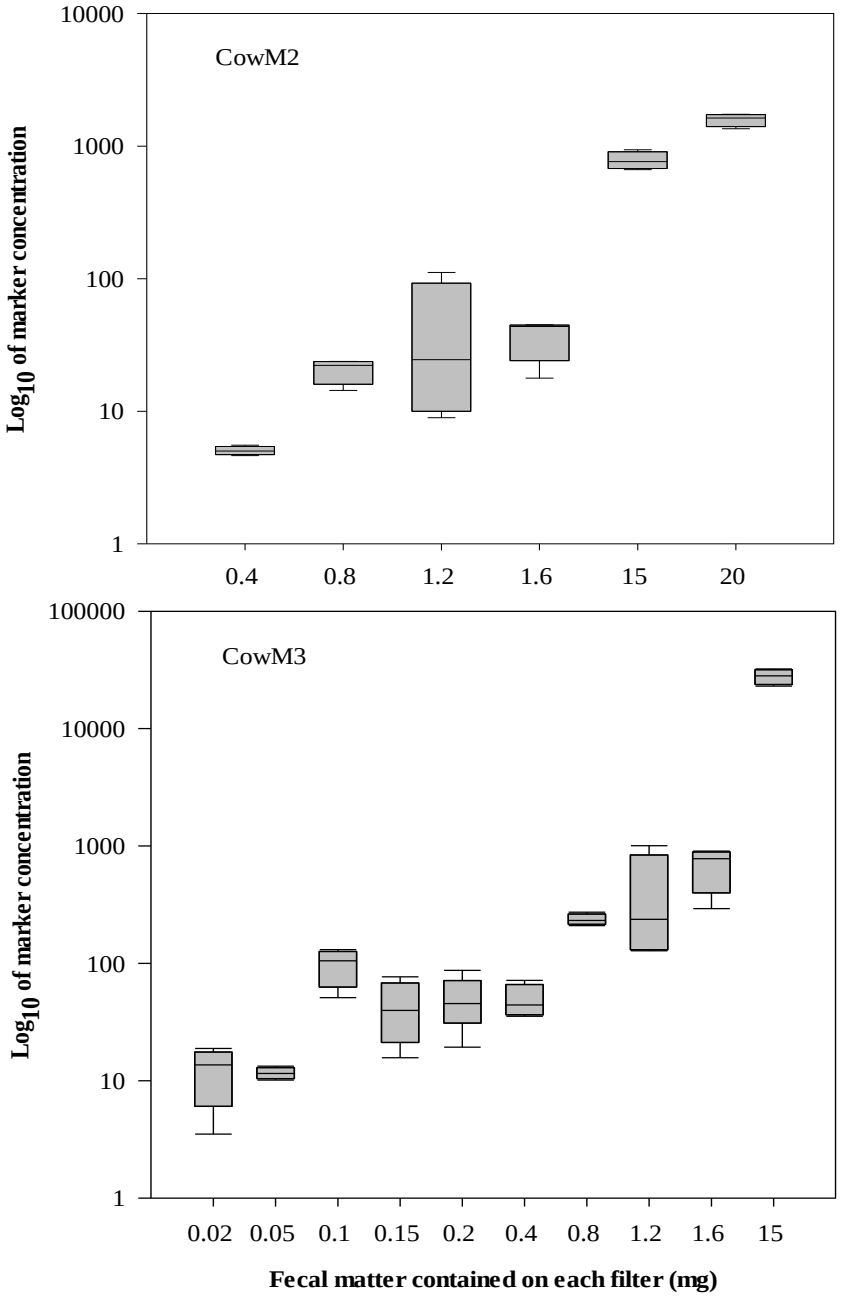


Fig. 3.6 Box plot showing marker concentration detected in serial dilutions of composite fecal sample for CowM2 and CowM3 genetic markers. Whisker caps represent the maximum and minimum concentration of the markers. The bar inside box represents the mean concentration and box lengths represent 5^{th}/95^{th} percentiles.

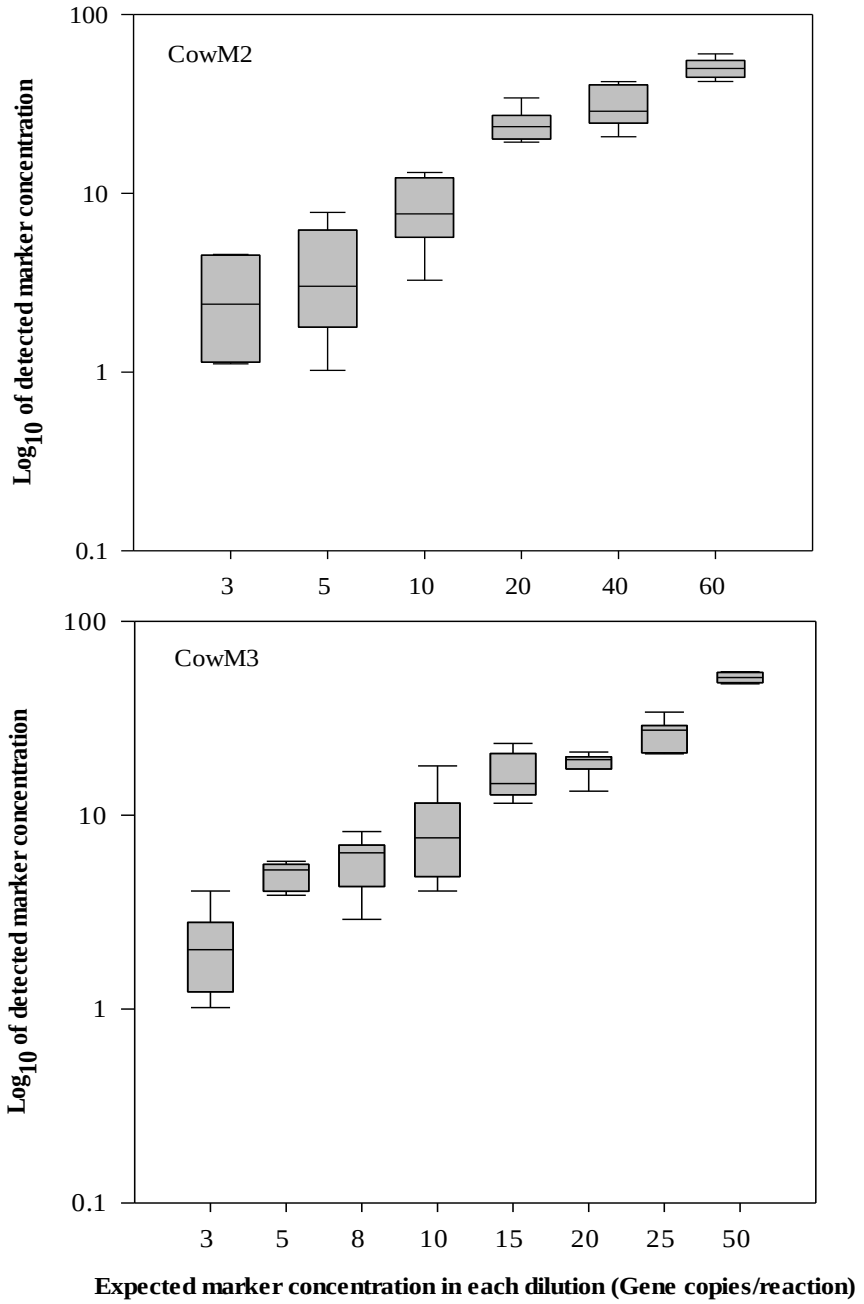


Fig. 3.7 Box plot showing marker concentration detected in each dilution of plasmid DNA sample for CowM2 and CowM3 genetic markers. Whisker caps represent the maximum and minimum concentration of the markers. The bar inside box represents the mean concentration and box lengths represent 5^{th}/95^{th} percentiles.

Chapter 4

Determination of fecal pollution sources at an Alabama beach

4.1 Abstract

Part of the eastern shore of Mobile Bay has been included on Alabama's 303(d) list of impaired waters due to elevated enterococci concentrations. Accurate identification of fecal pollution sources is necessary to develop effective pollution control strategies, so this study sought to determine whether human and/or bovine sources contribute to this fecal pollution. Water samples were collected weekly over a three-month period at nine locations along the shoreline. Enterococci were enumerated on mEI media following membrane filtration and the DNA extracted from the water samples was subjected to quantitative polymerase chain reactions (qPCR) targeting general (AllBac), human- (HF183) and cattle-associated (CowM3) *Bacteroidales*. Enterococci were found in all water samples, at concentrations ranging from 2 to 8000 CFU/100 ml. In particular, high concentrations of enterococci were observed after significant rainfall events. The AllBac marker was detected in all samples, with concentrations ranging from 2.73 to 8.05 $\log_{10}$ gene copies/100ml; there was a positive correlation between the levels of enterococci and AllBac marker recorded. The HF183 marker was detected in 49 out of 101 samples of which only nine had concentrations sufficiently high for quantification, and the CowM3 marker was detected in two, both at levels too low to be quantified. These results suggest that fecal contamination from cattle was minimal during the study period and that due to humans originated from diffuse sources. Further research is needed to determine whether waterfowl, seabirds, and other animal wastes in stormwater runoff are important sources of fecal contamination.

4.2 Introduction

Fecal contamination of coastal water is a serious environmental problem, especially for recreational beaches where human health may be at risk (Nayak et al., 2015; Webster et al., 2013). Recently, a significant portion of the eastern shore of Mobile Bay has been included on Alabama's 303(d) list of impaired waters due to the elevated enterococci concentrations measured in its coastal waters. Officials from the City of Fairhope have sought to determine the cause of this contamination, especially in the area along the beach around Fairhope Pier, where high bacterial counts often force them to close the beach. The possible sources of this pollution include storm water runoff, leaky sewer lines, waterfowl in and around the duck pond on the north beach, cattle farms, wildlife, and agricultural runoff.

For decades, the microbiological quality of water has been assessed by enumerating fecal indicator bacteria (FIB) such as *Escherichia coli* (*E. coli*) and enterococci in recreational waters (USEPA, 2008). However, FIB are not specific to any fecal source and the enumeration method requires at least 24 h to obtain results (Boehm et al., 2002; Leecaster and Weisberg, 2001). It is also possible for *E. coli* and enterococci to survive and regrow in sediment after being released into the environment (Anderson et al., 2005; Desmarais et al., 2002; Solo-Gabriele et al., 2000). The problems persist in spite of the considerable effort the city has made to minimize fecal input into local waters, largely due to the inability to identify non-point fecal pollution sources. In such cases, a rapid and accurate way to determine these sources is necessary if the city is to develop effective pollution control strategies.

Increasing interest is now being directed towards the use of library- and cultivation-independent microbial source tracking (MST) methods based on polymerase chain reaction (PCR) methods that target host-specific molecular markers. Combining MST methods with end-

point PCR and qPCR will provide results that reflect the most recent fecal pollution and identify the sources, thus facilitating the implementation of better strategies to protect humans against the health risks posed by polluted water (Wade et al., 2010). Members of the order *Bacteroidales* have been used as a promising fecal indicator to monitor microbial water quality due to their host specificity, broad geographic stability, and abundance in the gastrointestinal tract of warm-blooded animals (Bernhard and Field, 2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006).

In this study, a fecal pollution source tracking experiment was conducted for a municipal beach in the City of Fairhope in Baldwin County, Alabama. Both fresh and marine water samples were collected from city beaches and nearby tributaries and analyzed to assess the applicability of host-specific *Bacteroidales* genetic markers in identifying sources of fecal pollution and to determine if human and bovine feces were responsible for the poor microbial water quality in the area. The relationship between enterococci and genetic marker concentrations was investigated and a validation of marker sensitivity and specificity for the CowM3 marker carried out to assess its applicability in this particular geographical location, since the sensitivity and specificity are thought to be geographically constrained (Ebentier et al., 2013; Gawler et al., 2007; Shanks et al., 2008; Tambalo et al., 2012).

4.3 Materials and methods

Sample collection

The City of Fairhope, population 15,300, is located on the east shore of Mobile Bay in Baldwin County, Alabama. For this study, environmental water samples were collected weekly for three months in the spring of 2012 from January 17 to April 9 at nine different locations in and around the city's waterfront by city employees and shipped on ice to our laboratory

overnight. The nine sampling sites, shown in Figure 4.1, were: the Orange St. Pier (Site 1), the boat ramp at the city pier (Site 2), the duck pond (Site 3), Big Mouth Gully (Site 4), Fly Creek at the boat pump (Site 5), Fly Creek near the Fly Creek Café (Site 6), Fly Creek at the Scenic 98 road bridge (Site 7), Fly Creek at Woodland (Site 8), and Rock Creek (Site 9). In all, 13 sets of environmental water samples were collected during the sampling period. Water samples were collected from the Rock Creek site (site 9) when there was no water in the Big Mouth Gully (site 4). Among these nine sampling sites, Sites 1, 2, 5 and 6 consist of marine water, while the remaining samples are fresh water. For each, surface water was collected and stored in sterile 1-liter containers. After collection, water samples were kept on ice during transport to the lab, at which point the samples were examined to ensure they were in good condition with no ice inside the bottles. Each water bottle was then shaken vigorously 25 times to distribute the bacteria uniformly and a 500 ml water sample was extracted and filtered through a 0.45 µm-pore-size membrane filter. After filtration, sterile forceps were used to aseptically fold each of the membrane filters, which were placed in separate plastic Whirl-Pak bags (Nasco WHIRL-PAK^{TM}) and stored at -20°C until the DNA extraction.

Enumeration of enterococci

Enterococci concentrations were measured based on standard methods (USEPA Method 1600). To produce 20-60 colonies on the membranes, sample volumes of 3 ml, 10 ml, and 30 ml (1 ml, 3 ml, and 10 ml after significant rain fall events) were used for each site. In order to spread the small volume uniformly, phosphate buffered saline (PBS) was used to dilute the water samples and rinse the sides of the funnels used. Each water sample was filtered through a 0.45 µm-pore-size, 47-mm-diameter nitrocellulose membrane (Thermo Fisher Scientific, Waltham, MA) filter. After filtration sterile forceps were used to aseptically remove the membrane filter

from the filter base, and roll it onto the modified mEI agar. The petri dish was inverted and placed in a Whirl-Pak bag then incubated at 41±0.5°C for 24 ±2 hours. After incubation, colonies on the membrane filters were counted and recorded; growths that were at least 0.5 mm in diameter with a blue halo, regardless of colony color, were deemed to be enterococci.

DNA extraction from water samples

Genomic DNA was isolated from the 0.45 µm-pore-size, 47-mm-diameter nitrocellulose membranes through which the 500 ml water samples had been filtered (described above) using the PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Inc., Carlsbad, CA) according to the following procedure. To maximize the DNA extraction efficiency, the membrane filters were cut into small pieces with sterile scissors and the DNA was quantified with a NanoDrop ND-1000 UV spectrophotometer (Thermo Scientific, Wilmington, USA). The DNA samples were kept in 1.5 ml sterile centrifuge tubes and stored at -20°C for future use.

PCR and real time PCR assays

A total of 101 environmental water samples containing DNA collected from the nine sites were analyzed. The end-point polymerase chain reaction (PCR) assays (AllBac, HF183, and CowM3) were performed on a $TG_{RADIENT}$ thermal cycler (Whatman Biometra®, PA). Each 25 µl reaction mixture contained 5 µl of 5x colorless GoTaq® Flexi buffer, $MgCl_2$ solution (1.5 mM), 0.2mM of dNTPs, 0.5 µM of forward and reverse primers, 0.4 mg/ml of bovine serum albumin (BSA), 0.08 unit/µl GoTaq® DNA polymerase, and 2.0 µl of template DNA, made up to a total volume of 25 µl using PCR grade water. The thermal cycling parameters for each PCR assay were 94°C for 2 min, followed by 30 cycles of 94°C 60 s, 60°C 45 s, and 72°C 60 s, then 72°C for 7 min, after which the PCR was halted and held at 4 °C. The end-point PCR products were examined using 1.5% agarose gel electrophorese and viewed under UV light to verify the

absence or presence of the target gene. No template controls (NTC) containing only PCR grade water and positive controls were included in each instrument run for quality control.

Quantitative qPCR (qPCR) assays for the AllBac and HF183 genetic markers were performed using the Applied Biosystems StepOne real time PCR system (Applied Biosystems, NY). Here, the reaction mixture (15 µl) contained 1x SYBR green PCR master mix (BIO-RAD, CA), 0.7 µg/ µl BSA, 0.2 µM of each primer, and 5 µl of the template DNA. All reactions were performed in duplicate and began with a hold at 95°C for 10 min, followed by 40 cycles of 95°C 15s, 60°C 30 s, and 72°C 30 s. A calibration curve with concentrations spanning the range from 10 to 10^6 gene copies per reaction, with two replicates, was constructed. Duplicate no-template controls (NTC) were included in each run.

Construction of plasmid DNA standards

Plasmid DNA standards were constructed for the general *Bacteroidales* marker (AllBac), cattle-associated *Bacteroidales* marker (CowM3), and human associated *Bacteroidales* marker (HF183). For the plasmid DNA preparation, PCR were performed using each pair of primers according to the conditions described in the next section. The PCR product was examined by agarose gel electrophoresis, followed by purification using the DNA Clean & Concentrator™ kit (The Epigenetics Company™, CA). The purified PCR products were then ligated into the pCR™ 2.1-TOPO® vector cloning system (Invitrogen by Life Technologies, Grand Island, NY). Each clone was sub-cultured on a Luria Broth (LB) plate containing 50 µg/ mL ampicillin and the plasmids extracted from the transformed One Shot® Mach1™-T1^R competent *E. coli* strains. The plasmid DNA was purified with DNA Clean & Concentrator™ prior to sequencing. The cloning products were sequenced using M13 primers on an ABI 3100 DNA Genetic Analyzer (Applied biosystems, NY). The sequencing results were confirmed via the NCBI website using

the nucleotide Basic Local Alignment Search Tool (BLAST). The gene copy numbers for the plasmid were calculated using the equation below. The molecular weight (MW) of the TOPO-TA plasmid is 2,486,846 g/mol and the MW of the insert DNA was calculated using an online tool (www.basic.northwestern.edu/biotools/oligocalc.html). The calibration standard curves covering the range from 10 to 10^6 gene copies per reaction were prepared using serial dilutions of plasmid DNA extracted from a pure culture of competent *E. coli* strains.

$$\text{Gene Copy No.} = \frac{6.02 \times 10^{23} \times \text{Plasmid DNA concentration (g/µl)} \times \text{Volume used in qPCR (µl)}}{(\text{MW of TOPO vector} + \text{MW of insert})\ \text{(g/mol)}}$$

Data analysis

The amplification efficiencies (AE) were calculated based on the following equation:

$$AE = 10^{(-1/slope)} - 1$$

Statistical analyses were performed using SAS® 9.3 software. ArcGIS 10.2 software was used to generate the sampling map and the related spatial analysis of *Bacteroidales* marker concentrations across the nine sites in Fairhope. The R^2 value for each regression model was used to provide an indication of goodness of fit.

4.4 Results

A total of 101 water samples were analyzed over the 13-week sampling period. Enterococci was found in all water samples at levels ranging from 2 to 8000 CFU/100ml (Table 4.1). The highest enterococci concentrations among all the sites were observed for site 4, even though there were only two samples from that site after significant rainfall events. Samples collected from site 3 also had constantly high enterococci concentrations, with a geometric mean of 453 CFU/100ml over the 13-week sampling period (Fig. 4.3). Among the marine water sites

(sites 1, 2, 5, and 6), the geometric means for the enterococci concentration only exceeded the USEPA geometric mean criterion of 104 CFU/100ml at site 3. After significant rainfall events, enterococci counts were generally higher; the samples collected on Jan. 23rd, Feb. 20th, Mar. 12th, and Apr. 9th showed higher concentrations of enterococci compared with the samples collected on sunny days (Fig. 4.3).

SYBR green based qPCR assays were used to analyze AllBac for the general *Bacteroidales* marker, HF183 for the human-associated *Bacteroidales* marker, and CowM3 for the cattle-associated *Bacteroidales* marker in this study. Plasmid DNA for AllBac, HF183, and CowM3 were prepared to generate a calibration curve for the qPCR. Sequencing results were again confirmed by consulting the website NCBI using BLAST. The calibration curve equations and performance characteristics of the qPCR assays are shown in Table 3. The slopes and intercepts for the genetic markers were -3.19 and 34.14 (AllBac), -3.29 and 33.75 (HF183), and -3.38 and 33.92 (CowM3), respectively. The R^2 values for all three calibration curves were $\geq$ 0.99, and the amplification efficiencies for AllBac, HF183, and CowM3 were 99.21%, 97.76% and 97.62%, respectively. Concentrations of the general *Bacteroidales* marker designated as AllBac ranged from 2.73 to 8.05 $\log_{10}$ gene copies/100ml (Table 4.4). The concentrations for the general *Bacteroidales* marker exhibited similar trends to those of the concentrations of the enterococci counts (Fig. 4.4). An extremely high concentration of AllBac marker was observed in the sample collected from site 4 on March 12, at 8.05 $\log_{10}$ gene copies/100 ml. Water samples collected from site 3 had consistently high concentrations of AllBac marker, yielding an average value of 6.25 $\log_{10}$ gene copies/100 ml over the 13-week sampling period. Samples collected at site 1 and 5 had the lowest average concentrations (4.04 $\log_{10}$ gene copies/100 ml) for general *Bacteroidales* markers. In general, the regression analysis of the enterococci CFU and AllBac

marker concentrations showed a significant positive correlation (P-value< 0.0001), with an overall correlation coefficient r of 0.58 (Fig. 4.2).

The HF183 marker was quantifiable in only nine samples (8.9% of the total), although it was detected in 49 of the 101 samples (48.5%). Four of the seven samples collected on January 17 had high human marker concentrations, ranging from 3.02 $\log_{10}$ gene copies/100 ml at site 3 to 4.68 $\log_{10}$ gene copies/100 ml at site 1 (Table 4.5). The CowM3 marker was detected in only two samples (2% of the total), but the concentrations were too low to be quantified accurately. These two samples were collected from site 8 on February 6 and from site 9 on March 26.

The host-specificity and sensitivity of human-associated marker (HF183) have previously been reported by others, who found 50-100% sensitivity and 85-98% specificity, respectively (Ahmed et al., 2009; Carson et al., 2005). The specificity and sensitivity of the CowM3 marker has also been tested in the studies described in the earlier chapters of this dissertation, which examined 10 target and 132 non-target fecal DNA samples using end-point PCR. Here, the CowM3 *Bacteroidales* marker showed 100% sensitivity and 97.1% specificity, thus easily meeting the 80% benchmark recommended by the USEPA (2005) for both specificity and sensitivity. However, false positive amplification was also observed in four non-target DNA samples (two sewage and two deer samples).

In the present study, the water chemical analyses showed that the pH values of the water samples for each site averaged between 6.6 and 7.7, with the marine water samples having higher pH values than the fresh water samples. Not surprisingly, the electrical conductivity (EC) measured for the marine water were much higher than those for the freshwater, varying by an order of magnitude during the 13-week sampling period for sites 1, 2, 3, 5, and 6. The turbidity values averaged by site ranged from 3.6 NTU at site 8 to 44.6 NTU at site 4. The overall

correlation between enterococci concentrations and turbidity was low (r=0.20); however, positive correlations between these two parameters were observed at site 2 (r=0.71) and site 9 (r=0.79).

4.5 Discussion

Enterococci has been identified as a successful fecal indicator bacterial in both marine and freshwater samples as it directly correlates with gastroenteritis illness rates in swimmers exposed to fecal contaminated waters (Haugland et al., 2005). However, this culture based method requires at least 24 hours to obtain results and the potential for enterococci to persist and regenerate in the environment limits its utility (Anderson et al., 2005; Byappanahalli and Fujioka, 2004; Byappanahalli et al., 2006; Lessard and Sieburth, 1983; Sinton et al., 2002, 2007). Water quality in the area changed rapidly, especially after heavy rain events, and storm water runoff could be a significant source of fecal indicator bacteria. The resulting delay in reporting fecal contamination in the water could also pose a significant threat to the health of swimmers. As the data presented in Table 1 shows, the geometric mean of the enterococci CFU concentrations at site 1 was the lowest of all the sites tested, at ~30-fold lower than site 3. Similarly, the average AllBac marker concentration at site 1 was ~150-fold lower than at site 3. In general, the Allbac marker measurements obtained by qPCR were ~1000-fold higher than the enterococci CFU measured by the membrane filter method. Comparing the indicator concentrations from site 1 and site 3 using two different methods, the qPCR assays were more sensitive and were also able to detect DNA from both culturable and non-culturable or dead organisms. Previous studies have obtained similar results, suggesting that DNA molecular from dead cells could persist longer in natural environments (Haugland et al., 2005). The higher AllBac concentrations detected using qPCR not only suggest a lower limit of detection compared with culture-based methods, but also indicates that in addition to humans and cattle, other

potential pollution sources may exist. During the course of the three months study period, the enterococci CFU exhibited large variations and followed similar trends associated with precipitation for all the sampling sites (Fig. 4.3). This is likely because the runoff water originating from the city area after heavy rainfall events introduces waste pollution to nearby water bodies. Previous researchers have demonstrated the possibility that fecal indicator organisms could persist in natural environments by binding to soil particles and may be released by washing (Anderson et al., 2005; Wheeler et al., 2003). Our results for the enterococci concentrations observed at site 4 represent the effect of storm water runoff on fecal indicator concentrations. Only two water samples were collected at this site during the entire experimental period and both were after significant rainfall events. Storm water runoff is receiving more attention from public officials due to the growing awareness of its tendency to pick up and transport a variety of chemicals, pesticides, and human or animal fecal wastes as it washes over land (Liao et al., 2015; Panasiuk et al., 2015; Ridley et al., 2014). Our results support this, as the water quality from all the sampling sites was affected by storm water runoff during rainfall events. However, it is not possible to obtain information on pollution sources from those results, which limits their utility for city managers seeking to improve the water quality along Fairhope municipal beach.

It is highly recommended that specific genetic markers should be validated for use in different geographic locations since their performance characteristics may change markedly, thus affecting the evaluation results. CowM3 is a well-developed cattle-associated *Bacteroidales* genetic markers that has been widely used in many MST research studies. However, the applicability of those results to other regions is potentially limited due to factors such as host diet, climate and geographic location. For example, Dick et al. (2005) showed that *Bacteroidales*

spp. specificity reflects animal digestive tract physiology and diet rather than host animal phylogeny and the *Bacteroidales* 16S rRNA gene sequences obtained from swine have been reported to be more closely related to *Bacteroidales* 16S rRNA gene sequences obtained from humans than to bovines, even though swine and bovine are in the same order of Artiodactyla (Layton et al., 2006). Experiments have been conducted to test the sensitivity and specificity of the CowM3 marker, as described in earlier chapters, with the results showing the CowM3 marker to have an overall specificity of 97.1 % and 100% sensitivity. This result is consistent with previous studies that reported CowM3 to have a broader target host distribution and greater stability than the CowM2 marker (Raith et al., 2013; Shanks et al., 2006).

The regression analysis of the geometric mean of the enterococci CFU densities, which was determined by the membrane filter method (Method 1600), vs. the AllBac marker concentration, determined by the qPCR method, for the samples collected from the sites in and around Fairhope's municipal beach are shown in the scatter plot in Figure 4.2. The strong positive correlation between these two methods indicates that the qPCR method shows great promise as a way to monitor the health risks posed by fecal contamination in both fresh- and marine water environments. Since the two methods used target different indicator organisms, the samples that had low enterococci CFU but high AllBac concentrations may suggest a greater abundance of *Bacteroidales* organisms in environmental water samples. The results shown in Figure 4.2 also suggest that there may be constant loading of a stable fecal pollution from surrounding sources. Thus, a wider effort to identify the major pollution sources contributing to the beach contamination could be useful when developing a better management strategy designed to improve the beach's water quality. Although not present in all the samples, HF183 signals were detected at all the study sites, even though most concentrations were below the

quantification detection limit. The cattle associated marker was only detected in 2% of the samples (once each at site 8 and site 9), but again neither of them were at quantifiable levels. The sampling map shows that site 8 and site 9 are both far from the city center and close to large areas of farmland, thus increasing the likelihood of detecting cattle signals. The human signal detected in samples collected from sites 1, 2, 3, 4, 5 and site 6 were probably because they are all located relatively close to the center of Fairhope, where the larger population and greater frequency of human activities compared with other sampling locations might be the reason why these sites generate more human signals. In general, this result indicates that human sources of fecal pollution might be the major problem near the municipal beach, while fecal contamination from cattle appeared to be minimal during the study period. In order to improve the biological water quality at the Fairhope municipal beach, city managers would be well advised to reduce storm water runoff discharges by applying best management practices.

The limitations of the present study seeking to identify the sources of fecal pollution include the relatively limited samples and markers utilized. Future research is needed to determine whether sediment at the municipal beach is a source of the fecal indicator organisms found in the water, to determine if other pollution sources from seabirds or other animal wastes are contributing to the fecal contamination, and to lengthen the monitoring period so that a better understanding of the water conditions along Fairhope's municipal beach can be obtained.

4.6 Conclusions

1. Very high concentrations of enterococci were found in water samples collected after significant rainfall events. Storm water and surface runoff contributed to the elevated levels of enterococci at the Fairhope municipal beach.

2. Elevated levels of enterococci were found consistently in the water samples collected from site 3, the duck pond. This strongly suggests that the duck pond may contribute to the fecal contamination at the municipal beach.
3. General *Bacteroidales* markers were detected in all the water samples. There was a positive correlation between enterococci and general *Bacteroidales* marker concentrations.
4. Significant amounts of human-specific markers were found on one sampling day at Sites 1, 2, 3, and 5 during the 13-week sampling period, indicating the presence of human fecal sources.
5. Other fecal sources, such as waterfowl, seabirds, and other animal wastes in storm water runoff, and sediments may also contribute to the fecal contamination at the municipal beach.

4.7 References

Ahmed, W., Goonetilleke, A., Powell, D., Chauhan, K., Gardner, T., 2009. Comparison of molecular markers to detect fresh sewage in environmental waters. Water Research 43, 4908–4917. doi:10.1016/j.watres.2009.09.047

Anderson, K.L., Whitlock, J.E., Harwood, V.J., 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Applied and Environmental Microbiology 71, 3041–3048. doi:10.1128/AEM.71.6.3041-3048.2005

Bernhard, A.E., Field, K.G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Applied and Environmental Microbiology 66, 4571–4574. doi:10.1128/AEM.66.10.4571-4574.2000

Bernhard, A.E., Field, K.G., 2000b. Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Applied and Environmental Microbiology 66, 1587–1594. doi:10.1128/AEM.66.4.1587-1594.2000

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., Wellman, D.E., 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. Environ. Sci. Technol. 36, 3885–3892. doi:10.1021/es020524u

Byappanahalli, M., Fujioka, R., 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water Sci. Technol. 50, 27–32.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Ting, W.T.E., Tseng, C.C., Nevers, M.B., 2006. Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. Journal of Water and Health 4, 313–320. doi:10.2166/wh.2006.018

Carson, C.A., Christiansen, J.M., Yampara-Iquise, H., Benson, V.W., Baffaut, C., Davis, J.V., Broz, R.R., Kurtz, W.B., Rogers, W.M., Fales, W.H., 2005. Specificity of a *Bacteroides thetaiotaomicron* marker for human Feces. Applied and Environmental Microbiology 71, 4945–4949. doi:10.1128/AEM.71.8.4945-4949.2005

Desmarais, T.R., Solo-Gabriele, H.M., Palmer, C.J., 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Applied and Environmental Microbiology 68, 1165–1172. doi:10.1128/AEM.68.3.1165-1172.2002

Dick, L.K., Bernhard, A.E., Brodeur, T.J., Santo Domingo, J.W., Simpson, J.M., Walters, S.P., Field, K.G., 2005. Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Applied and Environmental Microbiology 71, 3184–3191. doi:10.1128/AEM.71.6.3184-3191.2005

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., Edrington, T.S., 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiology 8, 125. doi:10.1186/1471-2180-8-125

Durso, L.M., Harhay, G.P., Smith, T.P.L., Bono, J.L., DeSantis, T.Z., Harhay, D.M., Andersen, G.L., Keen, J.E., Laegreid, W.W., Clawson, M.L., 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Applied and Environmental Microbiology 76, 4858–4862. doi:10.1128/AEM.00207-10

Ebentier, D.L., Hanley, K.T., Cao, Y., Badgley, B.D., Boehm, A.B., Ervin, J.S., Goodwin, K.D., Gourmelon, M., Griffith, J.F., Holden, P.A., Kelty, C.A., Lozach, S., McGee, C., Peed, L.A., Raith, M., Ryu, H., Sadowsky, M.J., Scott, E.A., Domingo, J.S., Schriewer, A., Sinigalliano, C.D., Shanks, O.C., Van De Werfhorst, L.C., Wang, D., Wuertz, S., Jay, J.A., 2013. Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. Water Research 47, 6839–6848. doi:10.1016/j.watres.2013.01.060

Eckburg, P.B., 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638. doi:10.1126/science.1110591

Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Martin Walters, J., Meijer, W.G., 2007. Validation of host-specific *Bacteriodales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. Water Research 41, 3780–3784. doi:10.1016/j.watres.2007.01.028

Hattori, M., Taylor, T.D., 2009. The human intestinal microbiome: a new frontier of human biology. DNA Research 16, 1–12. doi:10.1093/dnares/dsn033

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P., Dufour, A.P., 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Research 39, 559–568. doi:10.1016/j.watres.2004.11.011

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., Isaacson, R.E., 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of

commercial pigs. Veterinary Microbiology 153, 124–133. doi:10.1016/j.vetmic.2011.05.021

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

Leecaster, M.K., Weisberg, S.B., 2001. Effect of sampling frequency on shoreline microbiology assessments. Mar. Pollut. Bull. 42, 1150–1154. doi:10.1016/S0025-326X(01)00130-8

Lessard, E.J., Sieburth, J.M., 1983. Survival of natural sewage populations of enteric bacteria in diffusion and batch chambers in the marine environment. Applied and Environmental Microbiology 45, 950–959.

Liao, H., Krometis, L.-A.H., Hession, W.C., Benitez, R., Sawyer, R., Schaberg, E., von Wagoner, E., Badgley, B.D., 2015. Storm loads of culturable and molecular fecal indicators in an inland urban stream. Sci. Total Environ. 530, 347–356. doi:10.1016/j.scitotenv.2015.05.098

Nayak, B., Weidhaas, J., Harwood, V.J., 2015. LA35 poultry fecal marker persistence is correlated with that of indicators and pathogens in environmental waters. Applied and Environmental Microbiology 81, 4616–4625. doi:10.1128/AEM.00444-15

Panasiuk, O., Hedstrom, A., Marsalek, J., Ashley, R.M., Viklander, M., 2015. Contamination of stormwater by wastewater: A review of detection methods. J. Environ. Manage. 152, 241–250. doi:10.1016/j.jenvman.2015.01.050

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J.S., Holden, P.A., Ebentier, D.L., Jay, J.A.,

Wang, D., Boehm, A.B., Aw, T.G., Rose, J.B., Balleste, E., Meijer, W.G., Sivaganesan, M., Shanks, O.C., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Research 47, 6921–6928. doi:10.1016/j.watres.2013.03.061

Ridley, C.M., Jamieson, R.C., Hansen, L.T., Yost, C.K., Bezanson, G.S., 2014. Baseline and storm event monitoring of *Bacteroidales* marker concentrations and enteric pathogen presence in a rural Canadian watershed. Water Res. 60, 278–288. doi:10.1016/j.watres.2014.04.039

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

Shanks, O.C., Santo Domingo, J.W., Lamendella, R., Kelty, C.A., Graham, J.E., 2006. Competitive metagenomic DNA hybridization identifies host-specific microbial genetic markers in cow fecal samples. Applied and Environmental Microbiology 72, 4054–4060. doi:10.1128/AEM.00023-06

Sinton, L.W., Braithwaite, R.R., Hall, C.H., Mackenzie, M.L., 2007. Survival of indicator and pathogenic bacteria in bovine feces on pasture. Applied and Environmental Microbiology 73, 7917–7925. doi:10.1128/AEM.01620-07

Sinton, L.W., Hall, C.H., Lynch, P.A., Davies-Colley, R.J., 2002. Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. Applied and Environmental Microbiology 68, 1122–1131. doi:10.1128/AEM.68.3.1122-1131.2002

Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R., Palmer, C.J., 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

Tambalo, D.D., Fremaux, B., Boa, T., Yost, C.K., 2012. Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. Water Research 46, 2891–2904. doi:10.1016/j.watres.2012.02.048

U. S. Environmental Protection Agency. 2008. National primary drinking water regulations: revisions to the total coliform rule. EPA-HQ-OW-2008-0878

Wade, T.J., Sams, E., Brenner, K.P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C.C., Love, D., Li, Q., 2010. Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environmental Health 9, 1–14.

Webster, L.F., Graves, D.A., Eargle, D.A., Chestnut, D.E., Gooch, J.A., Fulton, M.H., 2013. Assessment of animal impacts on bacterial water quality in a South Carolina, USA tidal creek system. Environmental Monitoring and Assessment 185, 7749–7756. doi:10.1007/s10661-013-3132-4

Wheeler Alm, E., Burke, J., Spain, A., 2003. Fecal indicator bacteria are abundant in wet sand at freshwater beaches. Water Research 37, 3978–3982. doi:10.1016/S0043-1354(03)00301-4

Table 4.1 Summary of the results for enterococci and *Bacteroidales* markers

Site	Enterococci Range CFU/100ml		*Bacteroidales* markers Positive/Total		
	Range	Geometric Mean	AllBac	Human	Cow
Site 1	2-342	15	13/13	4/13	0/13
Site 2	7-600	57	13/13	6/13	0/13
Site 3	140-8000	453	13/13	7/13	0/13
Site 4	1010-8000		2/2	0/2	0/2
Site 5	2-535	27	13/13	7/13	0/13
Site 6	4-1335	47	13/13	8/13	0/13
Site 7	10-250	45	12/12	9/12	0/12
Site 8	2-267	24	13/13	7/13	1/13
Site 9	43-455	103	9/9	2/9	1/9

Table 4.2 PCR primers used in this study

Assay	Target	primer sequence (5'-3')	Size (bp)	Annealing T (°C)	Reference
AllBac296F AllBac412R	General *Bacteroidales*	GAGAGGAAGGTCCCCCAC CGCTACTTGGCTGGTTCAG	106	60	Layton et al. 2006
HF183F HF265R	Human *Bacteroidales*	ATCATGAGTTCACATGTCCG TACCCCGCCTACTATCTAATG	82	60	Bernhard and Field, 2000
CowM3F CowM3R	Cow *Bacteroidales*	CCTCTAATGGAAAATGGATGGTATCT CCATACTTCGCCTGCTAATACCTT	122	60	Shanks et al. 2008

Table 4.3 Performance characteristics of qPCR assays for genetic markers

Assay	Reagent	Cal. Equation	Amplification efficiency (%)	R^2
AllBac	SYBR	y=34.14-3.19X	99.21	1.00
HF183	SYBR	y = 33.75-3.29x	97.76	1.00
CowM3	SYBR	y=33.92-3.38X	97.62	1.00

Table 4.4 Concentrations of general *Bacteroidales* markers at all sampling sites for the three month study period

Site Name	General *Bacteroidales* marker (log_{10} gene copies/100ml)												
	17-Jan	23-Jan	30-Jan	6-Feb	13-Feb	20-Feb	27-Feb	5-Mar	12-Mar	19-Mar	26-Mar	2-Apr	9-Apr
Site 1	4.92	3.49	3.37	4.50	3.41	3.30	3.72	3.19	3.89	3.47	3.38	3.32	2.73
Site 2	4.34	3.61	4.20	4.38	4.38	4.32	4.51	4.09	4.30	3.35	3.88	3.39	3.64
Site 3	5.24	5.86	5.63	5.84	5.27	5.22	5.88	5.32	7.26	5.89	5.85	4.89	5.05
Site 4		5.16							8.05				
Site 5	4.40	4.38	4.00	3.78	3.42	3.43	4.33	4.28	3.64	4.10	3.97	3.32	3.45
Site 6	3.80	4.37	4.31	3.76	3.96	3.91	4.37	4.44	4.22	4.46	4.13	3.65	3.94
Site 7	4.21	3.61	3.79	3.34	4.12	4.10	4.44	4.26	*	4.67	4.37	3.71	3.83
Site 8	3.57	4.33	4.12	3.22	4.19	4.13	4.41	4.43	4.72	4.52	4.72	4.28	4.23
Site 9				3.95	3.80	3.70	4.79	4.53		4.43	4.62	3.67	4.10

A blank cell indicates no water sample was collected.
*The bottle was broken during shipping.

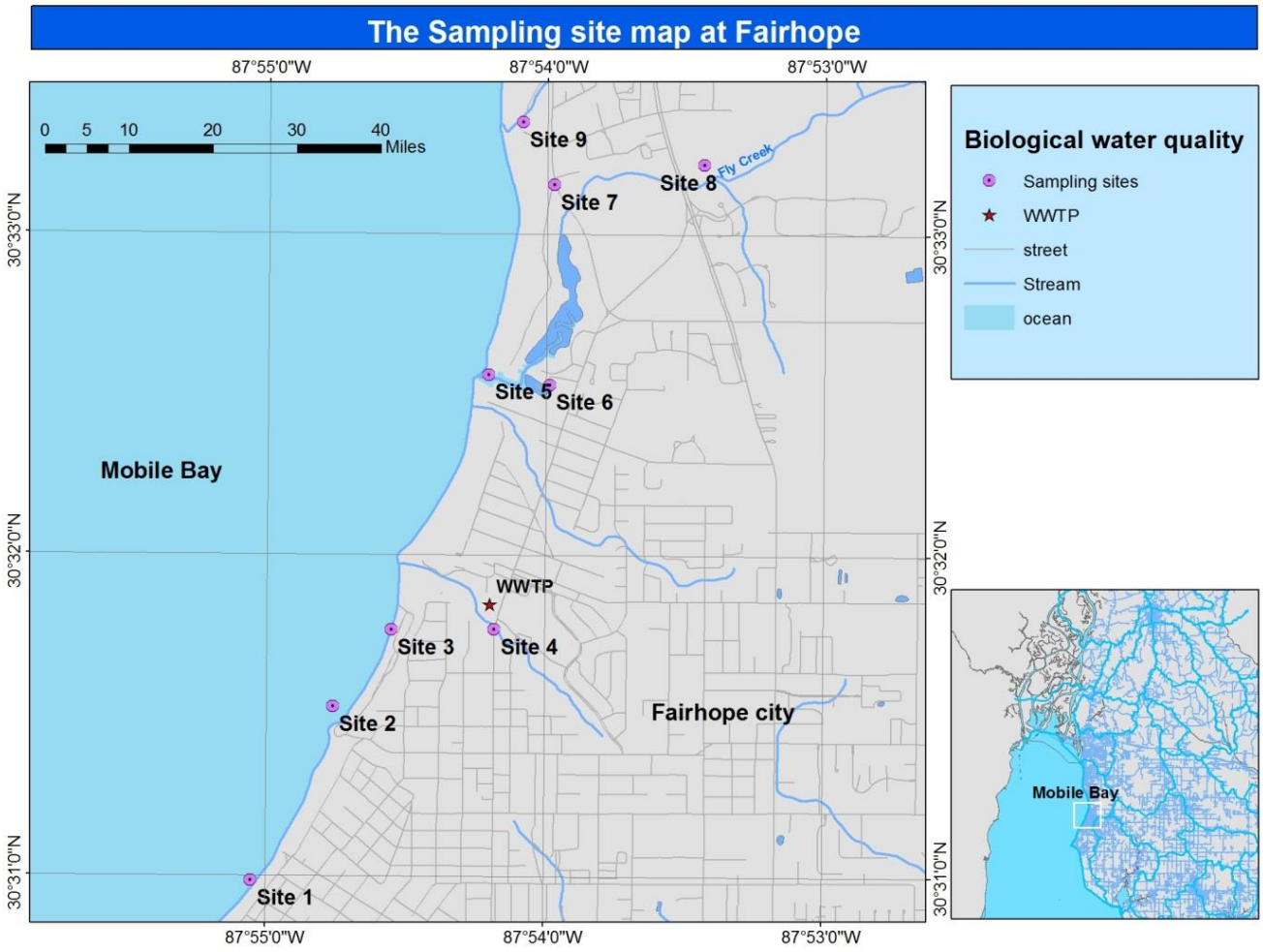
Table 4.5 Concentrations of human-associated *Bacteroidales* markers at all sampling sites for the three month study period

Site	Human-specific Bacteroidales marker (Log10 gene copies/100 ml)												
name	17-Jan	23-Jan	30-Jan	6-Feb	13-Feb	20-Feb	27-Feb	5-Mar	12-Mar	19-Mar	26-Mar	2-Apr	9-Apr
Site 1	4.68	0	+	+	0	0	+	0	0	0	0	0	0
Site 2	3.47	0	2.01	+	2.22	0	+	0	+	0	0	0	0
Site 3	3.02	0	+	+	0	0	+	0	0	0	+	0	+
Site 4		0							0				
Site 5	4.08	0	+	+	0	0	+	1.94	0	1.90	+	0	0
Site 6	+	+	+	+	0	0	+	+	0	+	+	0	0
Site 7	+	+	+	+	0	+	+	+	*	+	0	0	+
Site 8	+	+	+	+	0	0	+	+	0	0	0	0	+
Site 9				0	2.23	0	0	0		+	0	0	0

A blank cell indicates no water sample was collected.

*The bottle was broken during shipping.

+: Detected but too low to be quantified



1

2 Fig. 4.1 Sampling locations in and around Fairhope municipal beach on the eastern shore of

3 Mobile Bay, Alabama.

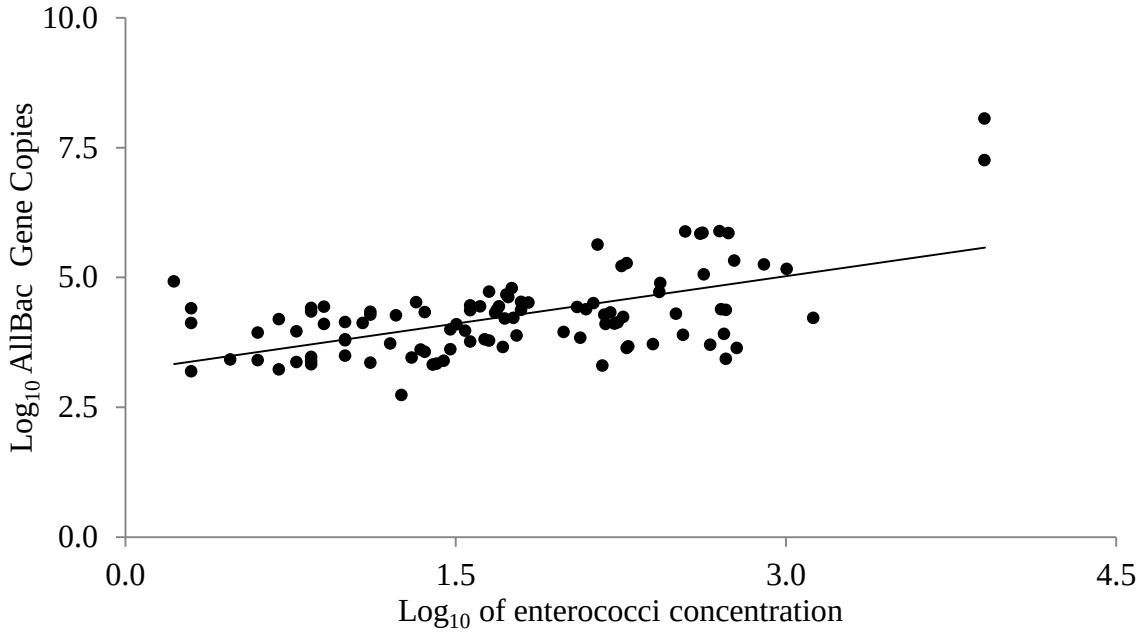


Fig. 4.2 Overall correlation between enterococci and general *Bacteroidales* marker during the sampling period.

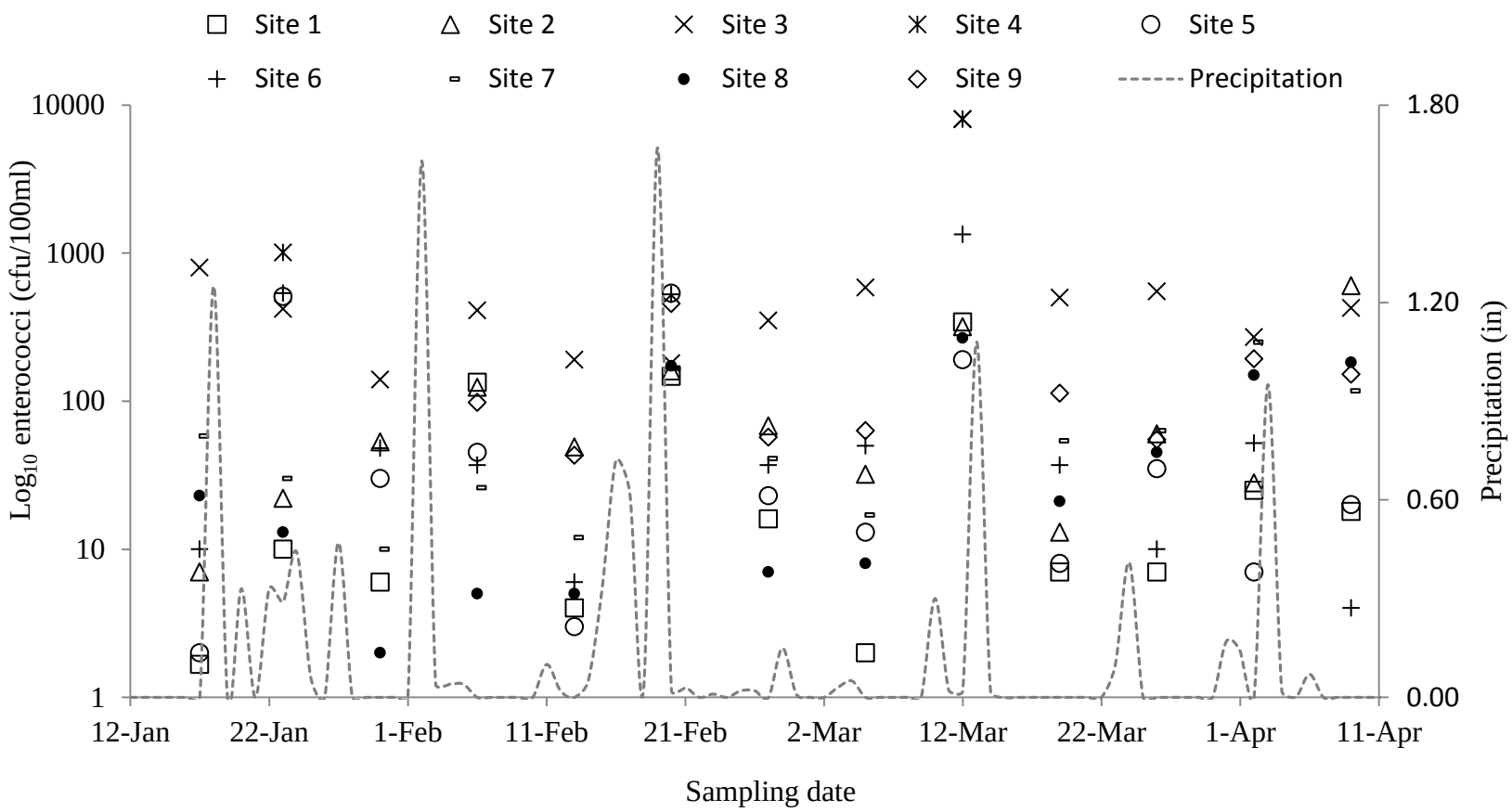


Fig. 4.3 Concentrations of enterococci at the Fairhope sampling sites from January 17 to April 9, 2012. The daily precipitation data were collected by NOAA at Point Clear, AL.

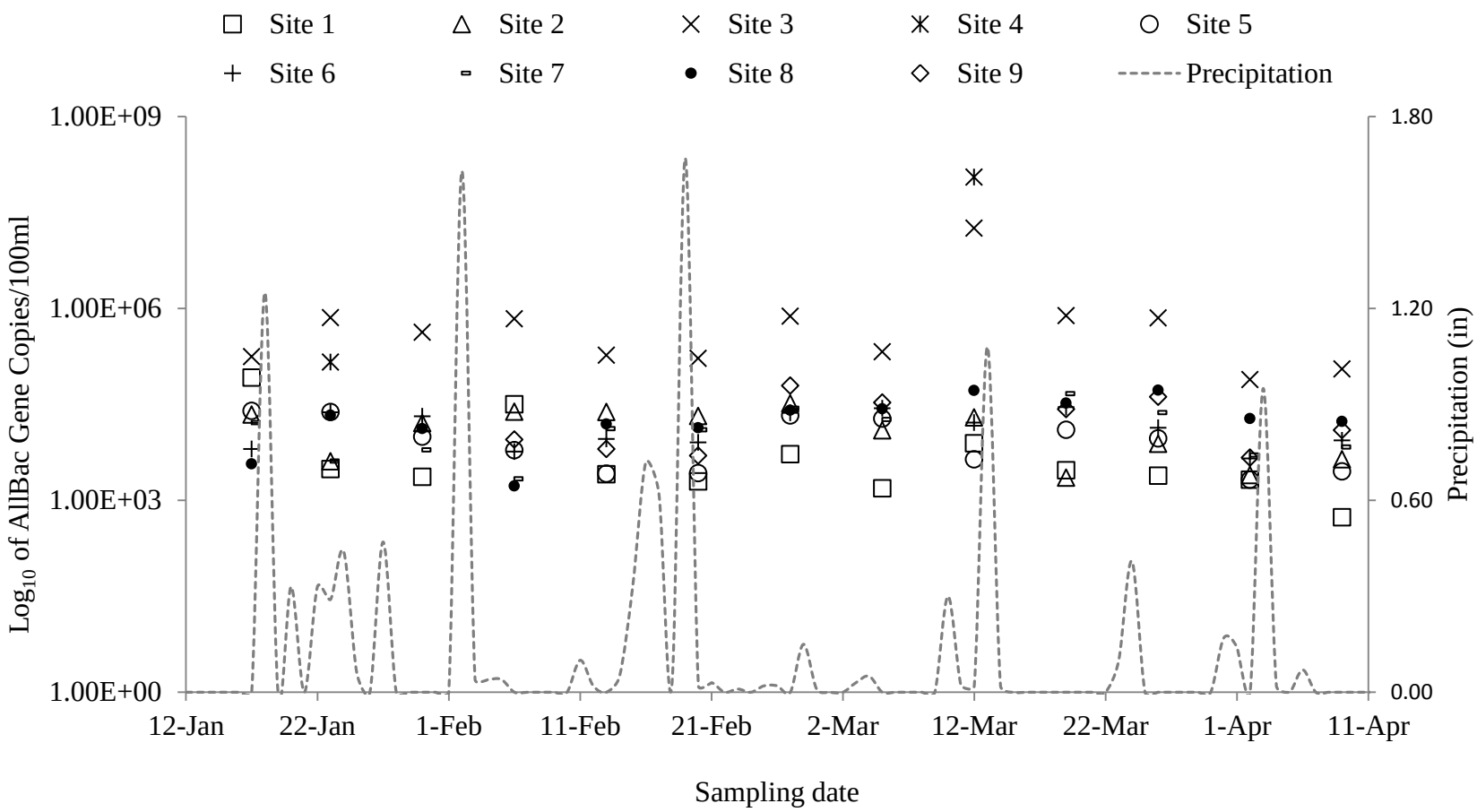


Fig. 4.4 Concentrations of general *Bacteroidales* markers at the Fairhope sampling sites from January 17 to April 9, 2012. The daily precipitation data were collected by NOAA at Point Clear, AL.

Chapter 5

Determination of adsorption and desorption of general *Bacteroidales* genetic marker on freshwater and marine sediments by quantitative PCR

5.1 Abstract

The adsorption of DNA by sediment prolongs the persistence of free DNA in the aquatic environment, often leading to ambiguity in the identification of recent fecal pollution sources when nucleic acid based methods are used. This study quantified the adsorption and desorption of DNA molecules on both freshwater and marine sediments using qPCR. A minimum of 36 hours was needed for the sorption to reach equilibrium. The sorption kinetics of DNA extracted from raw sewage and purified PCR products exhibited different trends. More DNA was adsorbed on both sediments in stream water than in 5 mM NaCl solution and DNA adsorption on both sediments increased in the presence of Ca^{2+} and Mg^{2+}. Clay content in the sediments also affected the DNA adsorption capacity. Adsorption data were fitted with the Langmuir and Freundlich equations. The observed DNA adsorption capacity was higher than the maximal capacity estimated from the Langmuir equation, suggesting the presence of multilayer adsorption. For the various solutions tested, between 5–22% of adsorbed DNA was desorbed. The findings of this study indicate that more DNA molecules were adsorbed on sediment through ligand bonding than electrostatic bonding and that DNA desorption can potentially complicate the results of microbial source tracking.

5.2 Introduction

Extracellular DNA in soil and sediment can persist due to binding on sediments, soil clays, and humic substances, be degraded by nuclease, or be transferred into other bacteria cells by transformation (Levy-Booth et al., 2007; Pietramellara et al., 2009). The transfer of genetic material among bacteria in aquatic and terrestrial ecosystems has long been considered a potentially serious risk to the environment and the persistence of DNA caused by its ability to bind to clay minerals and humic acids has been studied and reviewed extensively (Cai et al., 2006a; Cai et al., 2006b; Crecchio et al., 2005; Yu et al., 2013). Few studies have sought to analyze any differences in adsorption and desorption between environmental DNA and pure linear DNA molecule in aquatic sediment, which is largely composed of sand. Sediments are considered to be a good reservoir for extracellular DNA in a water environment (Pietramellara et al., 2009; Staley et al., 2012) and DNA may be protected from UV light and microbial DNases when it is adsorbed on soil particles (Levy-Booth et al., 2007). Most of the experiments reported have been carried out under controlled or artificial conditions using pure commercial DNA molecules and pure soil clays or humic acids (Crecchio et al., 2005; Pietramellara et al., 2009; Saeki et al., 2011). Real world environmental soil and sediments are likely to contain various components of living organisms other than DNA, such as cell wall fragments, proteins, and other types of DNA molecules. In addition, the detection instrument used in the previous research has generally been a spectrophotometer operating at 260 nm, which could be affected by the humic acids or other UV-absorbing chemicals present during the DNA extraction procedure.

Recently, molecular methods such as polymerase chain reaction (PCR) based microbial source tracking (MST) methods have been introduced to monitor surface water quality to address the problems caused by serious fecal contamination in environmental water (Bernhard and Field,

2000a, 2000b; Dowd et al., 2008; Durso et al., 2010; Eckburg, 2005; Hattori and Taylor, 2009; Kim et al., 2011; Layton et al., 2006). Genetic markers that target members of the order *Bacteroidales* are being increasingly used as alternative fecal indicator bacteria (FIB) due to their host specificity, sensitivity, and high abundance characteristics (Wexler, 2007). Compared with cultivation based methods (USEPA, 2000), the qPCR based MST method not only quantifies but also identifies fecal pollution sources in surface water (Scott et al., 2002). However, PCR is capable of detecting DNA from culturable cells, nonviable intact cells, viable but non-culturable (VBNC) cells, and extracellular free DNA (Rogers et al., 2011), so fecal contamination may be detected long after it first occurs (Josephson et al., 1993; Wolffs et al., 2005). Although the application of propidium monoazide (PMA) prior to qPCR has been shown to be a useful way of discriminating between DNA from live and dead cells, this method is less effective when dealing with untreated environmental samples, especially when particles such as sediment are involved (Kim et al., 2011; Nocker et al., 2007).

Naked DNA decays rapidly in a water environment but the binding of DNA molecules on soils and sediments extends the persistence of DNA in the environment. *Bacteroidales* markers and naked DNA not attached to soil particles can be damaged by sunlight (Green et al., 2011; Ravanat et al., 2001), but under the protection of sediment, molecular markers may persist for weeks or months longer than would be the case for live cells under certain conditions (Dejean et al., 2011; Kindler et al., 2006). Also, the abundance of genetic markers indicating fecal pollution in the water column have been positively correlated with effluent turbidity (Eichmiller et al., 2013) and the re-suspension of sediment may thus contribute significantly to FIB concentrations in shallow water (Staley et al., 2012). Therefore, a better understanding of the adsorption and

desorption of molecular markers in aquatic sediment is vital for the accurate assessment of fecal pollution in the environment.

This study explored the mechanisms involved in DNA adsorption and desorption on aquatic sediment using qPCR systems targeting *Bacteroidales* genetic markers. Unlike earlier research in this area, environmental sediment was used instead of pure clays. DNA from two sources, raw sewage (RS) and purified PCR product (PPP), were used in order to better understand the interactions between pure or environmental DNA and aquatic sediment. The adsorption of DNA from *Bacteroidales* on two sediments with different background solutions and desorption characteristics of DNA from sediment were investigated.

5.3 Materials and Methods

Sediment and water samples

Sediments were collected from two different locations. The fresh water sediment was collected from Parkerson Mill Creek (PMC) in Auburn, Alabama, and the marine sediment was collected from a municipal beach in the city of Fairhope (FH), Alabama. The fresh water sample was collected using sterilized plastic bottles at a knee depth of water and prior to the sediment collection to minimize any disturbance of the water column. The sediments and stream water (SW) sample were kept on ice and immediately transported to the laboratory, where they were stored at 4°C until processing. The sediment samples were air dried and autoclaved (121°C and 15 psi) prior to use. Sediment particle size analysis was conducted using the Bouyoucos hydrometer method (Bouyoucos, 1936). Surface water was filtered through a 0.45-µm-pore-size, 47-mm-diameter nitrocellulose membrane (Thermo Fisher Scientific, Waltham, MA) and then autoclaved at 121ºC for 20 min. The elemental analysis of the surface water was performed using

inductively coupled plasma (ICP) spectroscopy (Spectro Ciros CCD, side on plasma, SPECTRO Analytical Instruments Inc., Kleve, Germany).

DNA preparation

Two types of DNA were prepared: raw sewage DNA (RS) extracted from sewage samples and a 676-bp fragment of *Bacteroidales* 16S rRNA gene amplified from raw sewage DNA (described below). Raw sewage collected from the influent of a wastewater treatment plant in Auburn, AL was used for the DNA extraction. The raw sewage was stored in a 1 L sterilized plastic bottle and immediately transported to the lab on ice. The raw sewage was centrifuged at 5,500 g for 10 min at 4°C in 50 ml sterilized centrifuge tubes to concentrate the sewage slurry ten-fold, from 50 ml to 5 ml. After concentrating, the sewage slurry was mixed thoroughly and kept in sterilized 1.5 ml micro-centrifuge tubes and stored at -20°C prior to DNA extraction. All DNA extractions were performed using the PowerSoil^{TM} DNA Isolation kit (Mo Bio Laboratories, Inc. Carlsbad, CA) following the manufacturer's instructions; concentrated sewage with a volume of 500 µl was used for the DNA extraction.

End-point PCR using Bac32F and Bac708R primers (Bernhard and Field, 2000) targeting general *Bacteroidales* 16S rRNA gene was performed to generate PPP DNA using a $TG_{RADIENT}$ thermal cycler (Whatman Biometra®, Germany). Each 50 µl reaction mixture contained 25 µl of colorless GoTaq® master mix, 0.5 µM of forward and reverse primers, 0.4 mg/ml of bovine serum albumin (BSA) and 1.5 µl of template DNA, and was made up to the final volume of 50 µl by adding PCR grade water. The thermal cycling parameters for each PCR assay were 94°C for 2 min, followed by 30 cycles of 94°C 60 s, 60°C 45 s, and 72°C 60 s, finishing with 72°C for 7 min. The PCR products were examined by 1.5% agarose gel electrophorese under UV light to verify the absence or presence of the target. No template controls (NTC) composed of PCR grade

water and positive controls were included in each instrument run for quality control. The PCR product was collected into a 1.5 ml centrifuge tube and purified using the Wizard® SV gel and PCR clean-up system (Promega corporation, Madison, WI). DNA extracts and purified PCR products were quantified using a NanoDrop ND-1000 UV spectrophotometer. The raw sewage DNA and purified PCR product were stored at -20°C until used in the downstream DNA adsorption experiment.

DNA adsorption

For the adsorption kinetics experiment, two different sediments from PMC and FH were used. The autoclaved sediment (0.25 g) in each tube was mixed with 1 ml of 5 mM NaCl solution and appropriate amounts of raw sewage DNA extracted from waste water samples or purified PCR product (PPP) were added. The final concentrations for the raw sewage DNA experiments were 15 µg/ml for the PMC sediment and 1.0 µg/ml for the FH sediment based on the results of preliminary experiments. For the PPP experiments, a concentration of 15 µg/ml was used for the PMC sediment and 2.5 µg/ml for the FH sediment. For each treatment, two replicates were included and had been mixed on an end-over-end rotator at 40 rpm at 22 ± 1℃ for 48 hours. For each sample collection, 20 µl of supernatant was sampled at 0, 1, 2, 4, 6, 8, 12, 24, 36, and 48 hours after centrifugation at 13,500 g for 1 min. The samples were stored at -20℃ prior to the qPCR analysis (described below).

For the adsorption experiments, 0.25 g of autoclaved sediment was mixed with 1 ml of 5 mM NaCl solution or stream water. In order to study the effect of background solution on DNA adsorption, two additional sterile salt solutions were also used, they were 0.75 mM $CaCl_2$ and 0.75 mM $MgCl_2$. Appropriate volumes of DNA stock solution were added to obtain the desired initial DNA concentrations, ranging from 0.5 to 25 µg mL^{-1}. The mixture was then mixed in an

end-over-end rotator at 40 rpm at 22 ± 1℃ for 48 hours. The supernatant was collected at time 0 and 48 hours after centrifugation at 13,500 g for 1 min. As before, the samples were stored at -20℃ prior to the qPCR analysis.

DNA desorption

Two different DNA desorption experiments were performed to study the wash efficiency of three different wash solutions and the effect of equilibrium concentration on the DNA desorption. For the first objective, appropriate volumes of PPP DNA stock solution were added to 0.25 g PMC or FH sediments to provide initial DNA concentrations ranging from 1 to 15 µg/ml and the samples were equilibrated by mixing on an end-over-end rotator at 40 rpm at 22 ± 1℃ for 48 hours with 1 ml of 5 mM NaCl solution or stream water. As much of the supernatant was collected as possible after centrifugation at 13,500 g for 1 min and the residue was washed with 5 mM NaCl or stream water, 100 mM NaCl, and 100 mM phosphate (Na_3PO_4) at pH 7.0, sequentially. For each wash, 1 ml solution was used and the tube shaken thoroughly by hand for about 1 minute, after which the sample was centrifuged at 13,500 g for 1 minute and the supernatant collected in a sterile 1.5 ml centrifuge tube. The process was repeated for a total of six times for each solution, with the supernatants collected in separate tubes to quantify the DNA desorbed in each desorption cycle. In order to study the effect of equilibrium concentration on the DNA desorption, experiments were performed on both PMC and FH sediments by adding different amounts of PPP DNA, which was added to 1 ml of 5 mM NaCl or SW containing 0.25 g sediments to obtain a range of different equilibrium concentrations (shown in Table 3 as initial DNA concentrations). The samples were mixed on an end-over-end rotator at 40 rpm at 22 ± 1℃ for 48 hours, after which as much as possible of the supernatant was collected after centrifuging the tube at 13,500 g for 1 minute. The residue was washed with 5 mM NaCl and 100 mM

Na_3PO_4 at pH 7.0, sequentially. The same method as that described above was used to wash the residue for a total of six times and the supernatants combined into a single sample, labelled and stored at -20°C until used for the qPCR analysis.

Quantitative PCR assays

QPCR assays (AllBac) were performed using the StepOne real time PCR system (Applied Biosystems, NY). The reaction mixture (15 µl) contained 1x SsoAdvanced™ SYBR® Green Supermix (BIO-RAD, CA), 0.7 µg/ µl BSA, 0.2 µM of each primer, and 5 µl of the template DNA. All reactions were performed in duplicate and began with a hold at 95°C for 10 min, followed by 40 cycles of 95°C 15s and 60°C 30 s, and finishing with 72°C for 30 s. A melting curve step was added in order to check the purity of the PCR product. This step consisted of ramping the temperature from 60 to 95°C in increments of 0.5°C per step. Two NTC and standard curves with six points, where each point consisted of two replicates, were generated for each qPCR assay spanning the range from 10 to 10^6 gene copies per reaction. In order to eliminate the inhibition effect, samples were diluted between 10^2 and 10^4 fold based on the concentrations in the supernatant. In order to determine whether there was any inhibition effect, a blank sample spiked with known amounts of plasmid standard was included in each run.

Data analysis

The reproducibility of the adsorption and desorption experiments was confirmed by two independent experiments, each performed in duplicate. The amplification efficiencies (AE) were calculated based on the following equation: $E= 10^{(-1/slope)} – 1$. Statistical analyses were performed using SAS® 9.3 software (SAS Institute Inc., Cary, NC, USA). DNA adsorption data were fitted using both Langmuir and Freundlich equations. The Langmuir equation is $Y = \frac{X_m KC}{(1+KC)}$, where X_m is the maximum adsorption capacity of the sediment (µg DNA per mass unit), K is the Langmuir

constant, and C is the equilibrium concentration (μg/ml). The Freundlich equation is ($\frac{x}{m} = K_f C^{\frac{1}{n}}$), where K_f is the sorption coefficient, C is the equilibrium concentration (μg/ml), and x/m is the amount adsorbed per unit mass (μg/g). The correlation coefficients (R^2) and P-value for each regression model serve as an indication of goodness of fit.

5.4 Results

Characteristics of sediments and stream water

Two sediments from Parkerson Mill Creek (PMC) and Fairhope (FH) were analyzed according to the Bouyoucos hydrometer method. The PMC sediment was composed of 93.0% sand, 1.2% silt and 5.8% clay, and the FH sediment contained 94.8% sand, 1.2% silt and 4.0% clay. Carbon contents were measured and there were 0.05% and 0.27% of carbon in FH and PMC sediments, respectively. The pH value of the stream water (SW) and 5 mM NaCl were 7.90 and 6.92, respectively. Elemental analysis was conducted on stream water using ICP. The concentrations of Ca^{2+}, Mg^{2+} and Na^{+} in stream water were 27.83 ppm, 13.91 ppm, and 11.86 ppm, respectively. Among all the elements tested, Ca^{2+}, Mg^{2+} and Na^{+} present in significantly higher concentration than the other elements, such as P, K, Cu, S, and Zn.

DNA adsorption kinetics

The DNA adsorption kinetic curves for both the PMC and FH sediments and two different types of DNA molecules (RS and PPP DNA) are shown in Figs. 5.1a and 5.1b. A preliminary experiment found higher sorption capacity for PMC sediment than for FH sediment, therefore different amounts of DNA were added to each of the two sediments in the sorption kinetic experiment. The results showed that 36 to 48 hours were needed for adsorption to reach equilibrium and different adsorption patterns were observed for RS and PPP DNA. The RS DNA was initially quickly adsorbed on both sediments over the period between 2 and 12 hours, after

which the rates of adsorption gradually decreased from 12 hours to 48 hours, when a constant level was achieved. In contrast, the adsorption of PPP DNA on both sediments proceeded relatively slowly to begin with, from 2 to 8 hours, after which the DNA adsorption on both sediments increased from 12 to 36 hours. At 48 hours, about 75% of raw sewage DNA and 60% PPP DNA had been adsorbed on the PMC sediment, while for the FH sediment over 90% of the RS DNA molecules and 60% of the PPP DNA had been adsorbed.

DNA adsorption

Four types of solutions (5 mM NaCl, SW, 0.75 mM $MgCl_2$, and 0.75 mM $CaCl_2$) were used to carry out the adsorption experiments (Figs. 5.3 and 5.4). The adsorption of both RS DNA and PPP DNA on a constant amount of sediment increased with increasing of DNA equilibrium concentration, gradually reaching a constant level. However, for RS DNA adsorbed on FH sediment, the DNA adsorption decreased significantly as the DNA equilibrium concentration increased. DNA adsorption in stream water (SW), $MgCl_2$, and $CaCl_2$ followed similar trends to those observed for the NaCl solutions. These adsorption results fit both equations well, with the R^2 values ranging from 0.65 to 0.96 and from 0.30 to 0.96 for the Langmuir and Freundlich equations, respectively. The P-values showed that most of the fitted regression lines were significant ($P<0.05$), except for one Langmuir and three Freundlich plots (Table 5.1). Both the Langmuir constants (K) and the Freundlich adsorption coefficients (K_f) showed that DNA adsorption in 5 mM NaCl solution had the lowest affinity. The adsorption capacity in the NaCl solution was the lowest for both the PMC and FH sediments (Figs. 5.3 and 5.4), at about one order of magnitude lower than for the other three solutions when FH sediment was used. This result also agrees with the Langmuir parameter X_m, which represents the theoretical maximum amount of DNA that can be adsorbed. The K value in the Langmuir equation showed that the

0.75 mM $CaCl_2$ solution had the greatest affinity for both sediments with both types of DNA molecules. Both sediments showed highest adsorption capacity (X_m) in stream water with PPP DNA, while the adsorption of RS DNA on PMC sediment showed the highest adsorption capacity in 0.75 mM $MgCl_2$ solution. Except for the case of the NaCl solution, the DNA adsorption capacity in the other three solutions was similar.

In general, the fraction of DNA adsorbed in each treatment decreased as the DNA equilibrium concentration increased because the DNA adsorption capacity for a constant amount of sediment was fixed. The PMC sediment consistently exhibited a higher DNA adsorption capacity for both types of DNA in all the solutions compared with the FH sediment, although for both sediments the least amount of DNA was adsorbed when the 5 mM NaCl solution was used compared with the other three solutions. Most of the DNA could be adsorbed (> 80%) on PMC sediments when the DNA concentration was lower than 20 µg/mL and 15 µg/mL for PPP DNA and RS DNA, respectively. A higher percentage of RS DNA was adsorbed than PPP DNA at the same concentration in NaCl and $MgCl_2$ solution when the PMC sediment was used, but the opposite results (RS < PPP DNA) were obtained when FH sediment was used, except for one concentration of $CaCl_2$ solution. Among the four solutions, the largest amount of PPP DNA was adsorbed by both sediments in stream water, but the adsorption of RS DNA on PMC sediment was higher in the $MgCl_2$ and $CaCl_2$ solutions than in stream water. Similar results were observed for RS DNA on FH sediment, where the highest amount of adsorption occurred in the $CaCl_2$ solution.

Desorption experiments

Fig. 5.5 shows the DNA desorption achieved by sequential washing with 5 mM NaCl, stream water, 100 mM NaCl, and 100 mM Na_3PO_4 solutions. A series of six washes were

performed for each solution. In general, the amount of DNA desorbed from the sediment was higher in the first two washes with each solution, after which the percentage of DNA desorbed decreased with each successive wash, gradually reaching a constant level. In order to compare any differences in the effect of washing with 5 mM NaCl and stream water, the adsorption treatment in 5 mM NaCl solution was washed with NaCl solution and the other two treatments in stream water were washed with stream water. The results revealed that more DNA molecules were desorbed by washing the PMC sediment with NaCl solution, but there was no difference for the FH sediment. Four percent of the total adsorbed DNA was desorbed in the first wash in the PMC sediment with 5 mM NaCl solution. The second wash solution was 100 mM NaCl, and the results showed that the amounts of DNA desorbed from the sediments were much smaller for all four of the subsequent treatments. The total amount of DNA desorbed in the six washes ranged from 0.065 to 1.6%. However, for the treatments of both sediments in stream water, the DNA desorption was larger than for either of the two treatments in 5 mM NaCl solution. For the third wash solution we used 100 mM Na_3PO_4; here, the DNA desorption from the four treatments again followed a similar pattern, with the first two washes accounting for the majority of the DNA desorbed by the total washes with Na_3PO_4, which ranged from 4.00 to 16.86%. The total amount of DNA desorbed by six washes with 100 mM Na_3PO_4 ranged from 4.84 to 20.46%.

Table 3 shows the effects of the equilibrium concentration on DNA desorption. These results indicate that more DNA molecules were desorbed by washing with 100 mM Na_3PO_4 than 5 mM NaCl, which is similar to the results obtained for the six individual washes for each solution. The percentage of DNA desorbed ranged from 0.16 to 8.01% and 4.46 to 16.51% for the 5 mM NaCl and 100 mM Na_3PO_4, respectively. The actual DNA desorbed by the 5 mM NaCl and 100 mM Na_3PO_4 ranged from 0.002 to 2.16 µg and 0.09 to 3.62 µg, respectively. As

DNA adsorption rates followed the order of (PMC sediment) > (FH sediment) and (stream water) > (5 mM NaCl solution), the total DNA desorption rates were in the order of (PMC 25 SW) > (PMC 15 SW) > (PMC 15 NaCl) > (PMC 25 NaCl) > (FH 7.5 SW) > (FH 15 SW) > (FH 1 NaCl) > (FH 2.5 NaCl) based on the actual DNA (μg) desorbed, where the numbers represent the initial DNA concentrations used (μg/ml).

5.5 Discussion

There is increasing evidence to suggest that sediments serve as a reservoir for pathogenic microorganisms of fecal origin, such as the fecal indicator bacteria (FIB) *E. coli* and *Enterococcus* (Badgley et al., 2010; Wheeler Alm et al., 2003). Although the adsorption and desorption of these microorganisms on sediments have been studied extensively using culture-based methods (Kim et al., 2011; Shelton et al., 2014; Staley et al., 2012), this provides no information on the pollution sources; the ability of the FIB to survive and regenerate in sediment can also lead to false positive results. qPCR based MST methods have been widely used by researchers due to their specificity and sensitivity, but given that MST markers generally target genetic materials rather than viable organisms, their sorption kinetics may be different. Moreover, extracellular DNA have been shown to be capable of transferring among bacteria cells in the environment by conjugation, transduction, or transformation (Pietramellara et al., 2009; Yin and Stotzky, 1997). In order to gain a better understanding of the fate of extracellular DNA in soil, some research has focused on the adsorption and desorption of DNA molecules on clay minerals and soil particles in recent years, although previous studies have determined the DNA content at A_{260} using spectrophotometric methods (Alvarez et al., 1998; Cai et al., 2006a; Crecchio et al., 2005), which could be affected by the condition or impurities contained in the samples collected such as the effect of any humic acids or other chemicals present on the DNA

extraction procedure. The relatively low sensitivity and selectivity also limit this method's utility. The qPCR used in this study, on the other hand, benefits from a detection limit that can be as low as 1-10 fg DNA (Luna et al., 2012). Since fecal indicator bacteria can cause human infections at very low concentrations (USEPA, 1986), this low detection limit makes the qPCR assay the best choice for assessing biological water quality.

In this study, two different types of DNA molecules were used in the adsorption kinetic experiments to reveal any differences between the behavior of pure (PPP) DNA and environmental (RS) DNA. As expected, differences were observed between the two. The RS DNA results showed that the maximum adsorption occurred at 36 hours on both the types of sediments used, after which no significant increase in adsorption was observed. Previous studies have reported that the maximum DNA adsorption on sand particles was achieved in as little as 1 or 2 hours (Blum et al., 1997; Pietramellara et al., 2009; Pietramellara et al., 2001). This discrepancy is probably due to the relatively high detection limit for spectrophotometric methods; most previous experiments were also done under artificial conditions with pure DNA molecules and pure adsorbing matrices (Cai et al., 2006b; Crecchio and Stotzky, 1998; Pietramellara et al., 2009). In the present study, the adsorption of the PPP DNA was slow or even level from 2 to 8 hours, which is probably why previous researchers assumed that the adsorption process had reached its equilibrium. However, we found that after 12 hours the DNA adsorption sped up and no equilibrium was observed by the 48 hour mark. There may be several reasons for this phenomenon. First, it is possible that in the early stages of the adsorption, for the first few hours both the ends and the middle part of the DNA molecules attach to sediment. During the rotation treatment, however, the middle part of the DNA molecule becomes detached from the sediment particle and only the ends of the DNA molecule are adsorbed due to the richer negative

charges at the ends (Stein et al., 1995), which offer more binding sites and thus allow more DNA molecules to be adsorbed by the sediment. Second, the rotation treatment could split some of the sediment particles into smaller particles and thus increase the surface area, thus offering more DNA binding sites. Third DNA degradation may also explain this phenomenon, as DNA adsorption and degradation can be affected by different types of DNA molecules (Pietramellara et al., 2007) and the presence of other components (Cai et al., 2007a; Cai et al., 2007b). The natural RS DNA has a higher structural complexity and molecular size than PPP DNA and also contains many other kinds of DNA molecules since it was extracted from raw sewage slurry. This likely explains why the DNA adsorption of RS DNA by sediment particles proceeds more strongly and the effect of degradation is weak compared with PPP DNA.

Although both the PMC and FH sediments consisted largely of sand (93.0% for the PMC sediment and 94.8% for the FH sediment), the PMC sediment showed a markedly higher DNA adsorption capacity than the FH sediment. The slightly higher clay content in PMC sediment (5.8%) may account for its higher DNA adsorption capacity. Even though the difference appears small, only 1.8% higher than the 4% in the FH sediment, this corresponds to a 45% increase in the clay content, which has been shown to provide the majority of binding sites where DNA can bind in sediment (Blum et al., 1997; Levy-Booth et al., 2007). In general, the surface area of clay is at least three orders of magnitude greater than that of sand (Slater et al., 2006).

The adsorption capacity followed an order of SW>$CaCl_2$>$MgCl_2$>NaCl for PMC sediment with PPP DNA, SW>$MgCl_2$>$CaCl_2$>NaCl for PMC sediment with RS DNA, SW>$CaCl_2$>$MgCl_2$>NaCl for FH sediment with PPP DNA, and $CaCl_2$>SW>$MgCl_2$>NaCl for FH sediment with RS DNA. The results of the DNA adsorption in different solutions indicate that the divalent cations are much more effective than the monovalent cations in increasing the

adsorption of DNA on sand (Fig. 5.2) (Levy-Booth et al., 2007; Romanowski et al., 1991) because divalent cations such as Ca^{2+} and Mg^{2+} are more effective at forming bridges than monovalent cations like Na^{+}. The adsorption of DNA on sediment proceeds primarily through cation bridges, since both DNA and sediment particles carry negative charges when pH>5 (Yu et al., 2013). Additionally, these results suggest that the adsorbed DNA layer is more compact and rigid with divalent than monovalent cations (Nguyen and Elimelech, 2007). As the Ca^{2+} and Mg^{2+} cation concentrations in SW were about 0.7 and 0.6 mM, respectively, we prepared adsorption solutions that contained similar concentrations (0.75 mM) of Ca^{2+} and Mg^{2+}. Their DNA adsorption results were similar to the adsorption in SW, which suggests that both Ca^{2+} and Mg^{2+} play a very import role in enhancing DNA adsorption. Compared with the DNA adsorption capacity for FH sediment in NaCl solution, the DNA adsorption capacity with the other three solutions went up from 3 to 32-36 μg g^{-1}, over an order of magnitude higher. Interestingly, the DNA adsorption capacity for PMC sediment increased by only 40-110% for the same comparison, that was probably due to the different characteristics of the two sediments.

The DNA adsorption results on both sediments in all solutions were analyzed by both the Langmuir and Freundlich equations (Table 5.1) following the methods recommended by Schulthess and Dey (1996) and Rochette et al. (1996). The theoretical maximal amount of DNA adsorbed (X_m) by the PMC and FH sediments calculated using the Langmuir equation were lower than those observed from the adsorption isotherms (Fig. 5.3). The higher value shown in the isotherms suggests that some multilayer adsorption may also be taking place. Other researchers have also reported evidence indicating that the adsorbed DNA molecules may be more likely to form aggregates or precipitates on the surface of the sand particles (Crecchio and Stotzky, 1998; Saeki et al., 2011; Schulthess and Dey, 1996). The results for the Freundlich

equationwere similar to those obtained for the Langmuir equation (Table 5.1) except for three curves with non-significant P-values above 0.05 (PPP on PMC sediment in $MgCl_2$ solution, PPP on PMC sediment in $CaCl_2$ solution, and RS on FH sediment in NaCl solution). The K_f values for DNA adsorption in 5 mM NaCl solution on both sediments were significantly lower than the other three solutions, demonstrating its lower capacity for DNA adsorption. The data show that although DNA adsorption was affected by the ionic strength, even using a higher concentration of NaCl solution than was required for $MgCl_2$ and $CaCl_2$ did not boost the sorption capacity of the sediments as the divalent cations (Mg^{2+} and Ca^{2+}) are 70-100 times more efficient than monovalent cations when binding the same amount of DNA on sand (Romanowski et al., 1991). The DNA adsorption isotherms show the DNA adsorption capacity of the $CaCl_2$ solution to be higher than that of $MgCl_2$ which is consistent with the results reported by Paget et al. (1992) and Kalra et al. (2003). However, one discrepancy was observed: for RS DNA adsorbing on PMC sediment, the DNA sorption capacity in $MgCl_2$ solution was slightly higher than in $CaCl_2$. Another interesting phenomenon revealed by an examination of the adsorption isotherms was that as the equilibrium concentration increased, in some of the curves DNA adsorption capacity actually decreased. This was probably because as more DNA stock solution was added to the reaction tubes, the ionic strength was reduced due to the change in total volume.

DNA desorption from soil clays and humic acids has been studied by a number of previous researchers but with variable results. Saeki et al. (2011) found that less than 2% of the total DNA adsorbed was desorbed from humic acids by washing with four different solutions, but Crecchio et al. (1997) and Cai et al. (2006b) reported that in total, 20-30% and 25-90% of the adsorbed DNA was desorbed from humic acids and soil clays, respectively. DNA is thought to be adsorbed on humic acids and soil clays predominantly by electrostatic/ionic, ligand, and

hydrophobic bonding (Cai et al., 2006b; Saeki et al., 2011). Given that hydrophobic bonds are only present at pH levels of below 5 (Saeki et al., 2011), in the present study only the DNA adsorbed through electrostatic and ligand bonds was investigated, with the NaCl solution being employed to extract ionically bound DNA molecules and the DNA released by phosphate (Na_3PO_4) regarded as the fraction adsorbed by ligand exchange (Haynes et al., 1994). The results shown in Fig. 5.5 indicate that significantly more DNA molecules were bound to sand particles through ligand bonding than by ionic bonds in each treatment, which is consistent with the results reported by other researchers (Cai et al., 2006b). Interestingly, we observed higher amounts of DNA molecules were desorbed in the second wash from the PMC SW and FH SW treatments than for the same sediments in NaCl (Fig. 5: 100 mM NaCl), though the rates were low (1.3% and 1.6%). This is probably because in the first wash SW could not desorb the ionically bound DNA from the sand particles, so that it was not until the second wash in 100 mM NaCl that the ionically bound DNA could be desorbed at all, assuming that washing by 100 mM NaCl solution did not increase the DNA desorption from the sand particles. Since the rate of desorption observed with 100 mM NaCl solution was low, having a very similar effect to the 5 mM NaCl solution, we discontinued the use of this solution in the desorption experiment to observe the effect of different equilibrium concentrations on DNA desorption. Even though the DNA molecules were desorbed mostly from the first two washes, we did observe DNA targets in the remaining four washes at a low but constant level using qPCR assays. Although this continuing release of DNA molecules from the sand particles into a water environment is at odds with the results reported by others (Crecchio and Stotzky, 1998), this is likely due to the much lower detection limit of qPCR assays. These results do suggest that sediment could be a huge

reservoir for DNA molecules and it should therefore be analyzed before any application of MST in the field.

The results of the desorption experiments conducted using different equilibrium concentrations, shown in Table 5.2 indicate that 78-95% of the adsorbed DNA was not desorbed by sequential washing. This suggests that DNA molecules may be strongly bound to sand particles, rather than the weak bonding proposed by others (Levy-Booth et al., 2007). Fewer DNA molecules were desorbed from the higher equilibrium concentrations for nearly all the treatments, with the sole exception being the PMC sediment in SW. It is reasonable to assume that higher DNA concentrations would lead to more densely packed and stronger bonds between DNA and sand particles, including mechanisms such as hydrophobic interaction, aggregation, or precipitation (Saeki et al., 2011). However, the results of the current research provide no detailed information that either supports or discounts this conjecture and further experiments are needed to explore DNA adsorption and desorption on sand under more complex conditions using qPCR technology.

5.6 Conclusions

DNA adsorption and desorption on both fresh (PMC) and marine (FH) sediments were explored using qPCR based MST methods in the present study. The lower detection limit and higher specificity of qPCR assays provided a better way to investigate the DNA adsorption and desorption mechanisms on sand particles. The study's major findings are as follow:

1. The clay content played a major role in determining the DNA adsorption capacity in sand, even though it made up only a small portion of the composition. In the lab, environmental water had a higher adsorption capacity than 5 mM NaCl solution.

2. At least 36 hours were needed for the sorption experiment on sand to reach equilibrium. RS DNA was adsorbed relatively quickly from two to 12 hours, after which it gradually slowed until finally reaching equilibrium at 36 hours. In contrast, PPP DNA was only slow adsorbed from 2 to 8 hours, but then increased in speed from 12 to 36 hours.
3. DNA adsorption isotherms were fitted using both the Langmuir and Freundlich equations, with similar results being obtained for both. The DNA adsorption capacity shown in the sorption isotherms, which was higher than the theoretical maximum capacity (X_m) predicted by the Langmuir equation, suggested the possible presence of multilayer adsorption.
4. More DNA molecules were bound to sand particles through ligand binding than electrostatic binding, with the overall desorption results showing 5-22% of the adsorbed DNA was desorbed. The continuing release of small amounts of DNA in the later washing steps suggests that this may adversely impact the application of MST in the field.

5.7 References

Alvarez, A.J., Khanna, M., Toranzos, G.A., Stotzky, G., 1998. Amplification of DNA bound on clay minerals. Mol. Ecol. 7, 775–778. doi:10.1046/j.1365-294x.1998.00339.x

Badgley, B.D., Nayak, B.S., Harwood, V.J., 2010. The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. Water Res. 44, 5857–5866. doi:10.1016/j.watres.2010.07.005

Bernhard, A.E., Field, K.G., 2000a. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66, 4571–4574. doi:10.1128/AEM.66.10.4571-4574.2000

Bernhard, A.E., Field, K.G., 2000b. 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. doi:10.1128/AEM.66.4.1587-1594.2000

Blum, S. A. E., Lorenz, M.G., Wackernagel, W., 1997. Mechanism of retarded DNA degradation and prokaryotic origin of DNases in nonsterile soils. Syst. Appl. Microbiol. 20, 513–521.

Bouyoucos, G.J., 1936. Directions for making mechanical analyses of soils by the hydrometer method. Soil Sci. 42, 225–229. doi:10.1097/00010694-193609000-00007

Cai, P., Huang, Q.-Y., Zhang, X.-W., 2006a. Interactions of DNA with clay minerals and soil colloidal particles and protection against degradation by DNase. Environ. Sci. Technol. 40, 2971–2976. doi:10.1021/es0522985

Cai, P., Huang, Q., Zhang, X., Chen, H., 2006b. Adsorption of DNA on clay minerals and various colloidal particles from an Alfisol. Soil Biol. Biochem. 38, 471–476. doi:10.1016/j.soilbio.2005.05.019

Cai, P., Huang, Q., Chen, W., Zhang, D., Wang, K., Jiang, D., Liang, W., 2007a. Soil colloids-bound plasmid DNA: Effect on transformation of *E. coli* and resistance to DNase I degradation. Soil Biol. Biochem. 39, 1007–1013. doi:10.1016/j.soilbio.2006.11.010

Cai, P., Huang, Q., Zhu, J., Jiang, D., Zhou, X., Rong, X., Liang, W., 2007b. Effects of low-molecular-weight organic ligands and phosphate on DNA adsorption by soil colloids and minerals. Colloids Surf. B Biointerfaces 54, 53–59. doi:10.1016/j.colsurfb.2006.07.013

Crecchio, C., Stotzky, G., 1998. Binding of DNA on humic acids: Effect on transformation of Bacillus subtilis and resistance to DNase. Soil Biol. Biochem. 30, 1061–1067. doi:10.1016/S0038-0717(97)00248-4

Crecchio, C., Ruggiero, P., Curci, M., Colombo, C., Palumbo, G., Stotzky, G., 2005. Binding of DNA from on montmorillonite–humic acids–aluminum or iron hydroxypolymers. Soil Science Society of America Journal 69, 834. doi:10.2136/sssaj2004.0166

Crecchio, C., Ruggiero, P., Curci, M., Colombo, C., Palumbo, G., Stotzky, G., 2005. Binding of DNA from on montmorillonite–humic acids–aluminum or iron hydroxypolymers. Soil Sci. Soc. Am. J. 69, 834. doi:10.2136/sssaj2004.0166

Crecchio, C., Stotzky, G., 1998. Binding of DNA on humic acids: Effect on transformation of *Bacillus subtilis* and resistance to DNase. Soil Biol. Biochem. 30, 1061–1067. doi:10.1016/S0038-0717(97)00248-4

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., Miaud, C., 2011. Persistence of environmental DNA in freshwater ecosystems. PLoS ONE 6, e23398. doi:10.1371/journal.pone.0023398

Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeehan, T., Hagevoort, R.G., Edrington, T.S., 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA

bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiol. 8, 125. doi:10.1186/1471-2180-8-125

Durso, L.M., Harhay, G.P., Smith, T.P.L., Bono, J.L., DeSantis, T.Z., Harhay, D.M., Andersen, G.L., Keen, J.E., Laegreid, W.W., Clawson, M.L., 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. Appl. Environ. Microbiol. 76, 4858–4862. doi:10.1128/AEM.00207-10

Eckburg, P.B., 2005. Diversity of the human intestinal microbial flora. Science 308, 1635–1638. doi:10.1126/science.1110591

Eichmiller, J.J., Hicks, R.E., Sadowsky, M.J., 2013. Distribution of genetic markers of fecal pollution on a freshwater sandy shoreline in proximity to wastewater effluent. Environ. Sci. Technol. 130321163549009. doi:10.1021/es305116c

Green, H.C., Shanks, O.C., Sivaganesan, M., Haugland, R.A., Field, K.G., 2011. Differential decay of human faecal *Bacteroides* in marine and freshwater: Differential decay of *Bacteroides*. Environ. Microbiol. 13, 3235–3249. doi:10.1111/j.1462-2920.2011.02549.x

Hattori, M., Taylor, T.D., 2009. The human intestinal microbiome: a new frontier of human biology. DNA Res. 16, 1–12. doi:10.1093/dnares/dsn033

Haynes, C., Sliwinsky, E., Norde, W., 1994. Structural and electrostatic properties of globular-proteins at a polystyrene water interface. J. Colloid Interface Sci. 164, 394–409. doi:10.1006/jcis.1994.1182

Josephson, K.L., Gerba, C.P., Pepper, I.L., 1993. Polymerase chain reaction detection of nonviable bacterial pathogens. Appl. Environ. Microbiol. 59, 3513–3515.

Kalra, S., Pant, C.K., Pathak, H.D., Mehata, M.S., 2003. Studies on the adsorption of peptides of glycine/alanine on montmorillonite clay with or without co-ordinated divalent cations.

Colloid Surf. A-Physicochem. Eng. Asp. 212, 43–50. doi:10.1016/S0927-7757(02)00288-1

Kim, H.B., Borewicz, K., White, B.A., Singer, R.S., Sreevatsan, S., Tu, Z.J., Isaacson, R.E., 2011. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. Vet. Microbiol. 153, 124–133. doi:10.1016/j.vetmic.2011.05.021

Kindler, R., Miltner, A., Richnow, H., Kastner, M., 2006. Fate of gram-negative bacterial biomass in soil—mineralization and contribution to SOM. Soil Biol. Biochem. 38, 2860–2870. doi:10.1016/j.soilbio.2006.04.047

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., Sayler, G., 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

Levy-Booth, D.J., Campbell, R.G., Gulden, R.H., Hart, M.M., Powell, J.R., Klironomos, J.N., Pauls, K.P., Swanton, C.J., Trevors, J.T., Dunfield, K.E., 2007. Cycling of extracellular DNA in the soil environment. Soil Biol. Biochem. 39, 2977–2991. doi:10.1016/j.soilbio.2007.06.020

Luna, G.M., Dell'Anno, A., Pietrangeli, B., Danovaro, R., 2012. A new molecular approach based on qPCR for the quantification of fecal bacteria in contaminated marine sediments. J. Biotechnol. 157, 446–453. doi:10.1016/j.jbiotec.2011.07.033

Nguyen, T.H., Elimelech, M., 2007. Plasmid DNA adsorption on silica: Kinetics and conformational changes in monovalent and divalent salts. Biomacromolecules 8, 24–32. doi:10.1021/bm0603948

Nocker, A., Sossa-Fernandez, P., Burr, M.D., Camper, A.K., 2007. Use of propidium monoazide for live/dead distinction in microbial ecology. Appl. Environ. Microbiol. 73, 5111–5117. doi:10.1128/AEM.02987-06

Paget, E., Monrozier, L., Simonet, P., 1992. Adsorption of DNA on clay-minerals - protection against dnasei and influence on gene-transfer. FEMS Microbiol. Lett. 97, 31–39.

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M.T., Guerri, G., Nannipieri, P., 2009. Extracellular DNA in soil and sediment: fate and ecological relevance. Biol. Fertil. Soils 45, 219–235. doi:10.1007/s00374-008-0345-8

Pietramellara, G., Ascher, J., Ceccherini, M.T., Nannipieri, P., Wenderoth, D., 2007. Adsorption of pure and dirty bacterial DNA on clay minerals and their transformation frequency. Biol. Fertil. Soils 43, 731–739. doi:10.1007/s00374-006-0156-8

Pietramellara, G., Franchi, M., Gallori, E., Nannipieri, P., 2001. Effect of molecular characteristics of DNA on its adsorption and binding on homoionic montmorillonite and kaolinite. Biol. Fertil. Soils 33, 402–409. doi:10.1007/s003740100341

Ravanat, J.-L., Douki, T., Cadet, J., 2001. Direct and indirect effects of UV radiation on DNA and its components. J. Photochem. Photobiol. B 63, 88–102.

Rochette, E.A., Koskinen, W.C., 1996. Supercritical carbon dioxide for determining atrazine sorption by field-moist soils. Soil Sci. Soc. Am. J. 60, 453–460.

Rogers, S.W., Donnelly, M., Peed, L., Kelty, C.A., Mondal, S., Zhong, Z., Shanks, O.C., 2011. Decay of bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-amended soils. Appl. Environ. Microbiol. 77, 4839–4848. doi:10.1128/AEM.02427-10

Romanowski, G., Lorenz, M., Wackernagel, W., 1991. Adsorption of plasmid DNA to mineral surfaces and protection against Dnase-I. Appl. Environ. Microbiol. 57, 1057–1061.

Saeki, K., Ihyo, Y., Sakai, M., Kunito, T., 2011. Strong adsorption of DNA molecules on humic acids. Environ. Chem. Lett. 9, 505–509. doi:10.1007/s10311-011-0310-x

Schulthess, C.P., Dey, D.K., 1996. Estimation of Langmuir constants using linear and nonlinear least squares regression analyses. Soil Sci. Soc. Am. J. 60, 433–442.

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial source tracking: current methodology and future directions. Appl. Environ. Microbiol. 68, 5796–5803. doi:10.1128/AEM.68.12.5796-5803.2002

Shelton, D.R., Pachepsky, Y.A., Kiefer, L.A., Blaustein, R.A., McCarty, G.W., Dao, T., 2014. Response of coliform populations in streambed sediment and water column to changes in nutrient concentrations in water. Water Res. doi:10.1016/j.watres.2014.04.019

Slater, L., Ntarlagiannis, D., Wishart, D., 2006. On the relationship between induced polarization and surface area in metal-sand and clay-sand mixtures. Geophysics 71, A1–A5. doi:10.1190/1.2187707

Staley, C., Reckhow, K.H., Lukasik, J., Harwood, V.J., 2012. Assessment of sources of human pathogens and fecal contamination in a Florida freshwater lake. Water Res. 46, 5799–5812. doi:10.1016/j.watres.2012.08.012

Stein, V., Bond, J., Capp, M., Anderson, C., Record, M., 1995. Importance of coulombic end effects on cation accumulation near oligoelectrolyte B-DNA - a demonstration using Na-23 NMR. Biophys. J. 68, 1063–1072.

U. S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA A440/5-84-002. U. S. Environmental Protection Agency, Washington, D. C.

U. S. Environmental Protection Agency. 2000. Assigning values to nondetected/ non-quantified pesticide residues in human health food exposure assessments.

Wexler, H.M., 2007. *Bacteroides*: the good, the bad, and the nitty-gritty. Clin. Microbiol. Rev. 20, 593–621. doi:10.1128/CMR.00008-07

Wheeler Alm, E., Burke, J., Spain, A., 2003. Fecal indicator bacteria are abundant in wet sand at freshwater beaches. Water Res. 37, 3978–3982. doi:10.1016/S0043-1354(03)00301-4

Wolffs, P., Norling, B., Rådström, P., 2005. Risk assessment of false-positive quantitative real-time PCR results in food, due to detection of DNA originating from dead cells. J. Microbiol. Methods 60, 315–323. doi:10.1016/j.mimet.2004.10.003

Yin, X.M., Stotzky, G., 1997. Gene transfer among bacteria in natural environments, in: Neidleman, S.L., Laskin, A.I. (Eds.), Advances in Applied Microbiology, Vol 45. Elsevier Academic Press Inc, San Diego, pp. 153–212.

Yu, W.H., Li, N., Tong, D.S., Zhou, C.H., Lin, C.X. (Cynthia), Xu, C.Y., 2013. Adsorption of proteins and nucleic acids on clay minerals and their interactions: A review. Appl. Clay Sci. 80-81, 443–452. doi:10.1016/j.clay.2013.06.003

Table 5.1 Langmuir and Freundlich parameters for the adsorption of DNA on aquatic sediments

Sediment	DNA type	Solution	Langmuir				Freundlich			
			X_m (μg DNA g^{-1})	K	R^2	P-value	K_f μg g^{-1} (μg mL^{-1})$^{-1/n}$	1/n	R^2	P-value
PMC sediment	PPP DNA	NaCl	46.01	0.88	0.89	0.0048	22.64	0.25	0.81	0.0152
		SW	98.91	14.67	0.96	0.0036	78.83	0.20	0.87	0.0197
		$MgCl_2$	72.56	22.80	0.96	0.0041	57.50	0.12	0.71	**0.0718**
		$CaCl_2$	76.33	27.98	0.88	0.0194	58.43	0.10	0.61	**0.1207**
	RS DNA	NaCl	49.33	0.21	0.89	0.0046	26.32	0.40	0.96	0.0006
		SW	70.91	8.44	0.90	0.0043	52.15	0.17	0.95	0.0010
		$MgCl_2$	80.65	13.21	0.84	0.0098	60.57	0.23	0.94	0.0012
		$CaCl_2$	71.67	77.38	0.94	0.0013	58.15	0.16	0.93	0.0016
FH sediment	PPP DNA	NaCl	3.09	15.35	0.93	0.0019	2.78	0.21	0.91	0.0034
		SW	36.50	103.50	0.94	0.0012	32.33	0.13	0.89	0.0047
		$MgCl_2$	25.17	42.14	0.78	0.0203	21.05	0.13	0.82	0.0136
		$CaCl_2$	30.23	169.33	0.74	0.0272	26.83	0.10	0.68	0.0422
	RS DNA	NaCl	3.05	14.89	0.65	**0.1002**	2.60	0.07	**0.30**	**0.3369**
		SW	/	/	/	/	/	/	/	/
		$MgCl_2$	/	/	/	/	/	/	/	/
		$CaCl_2$	39.55	613.53	0.94	0.0063	40.69	0.11	0.88	0.0193

PPP DNA refers to the purified PCR product; RS DNA is raw sewage DNA; NaCl, SW, $MgCl_2$ and $CaCl_2$ signify that the adsorption experiment was performed in 5 mM NaCl solution, stream water, 0.75 mM $MgCl_2$ solution and 0.75 mM $CaCl_2$ solution, respectively

Note: The adsorption results in SW and $MgCl_2$ solution for the FH sediment with RS DNA could not be fitted by either the Langmuir or Freundlich equations.

Table 5.2 Percentage of DNA desorbed from PMC and FH sediments by using NaCl and Na_3PO_4 sequentially

Treatment	5 mM NaCl		100 mM Na_3PO_4		Total	
	µg	(%)	µg	(%)	µg	(%)
PMC 15* NaCl**	0.64	4.41	1.25	8.63	1.89	13.04
PMC 25 NaCl	0.51	2.30	0.99	4.46	1.51	6.77
PMC 15 SW	0.94	5.58	2.22	13.22	3.16	18.80
PMC 25 SW	2.16	8.01	3.62	13.42	5.77	21.43
FH 1 NaCl	0.002	0.32	0.11	16.51	0.11	16.83
FH 2.5 NaCl	0.003	0.16	0.09	4.58	0.09	4.74
FH 7.5 SW	0.42	5.71	0.78	10.67	1.20	16.38
FH 15 SW	0.19	1.42	0.66	5.02	0.85	6.44

(PMC 15 NaCl): PMC sediment and 5 mM NaCl solution, initial DNA concentration of 15 µg/ml.

(PMC 25 NaCl): PMC sediment and 5 mM NaCl solution, initial DNA concentration of 25 µg/ml.

(PMC 15 SW): PMC sediment and stream water, initial DNA concentration of 15 µg/ml.

(PMC 25 SW): PMC sediment and stream water, initial DNA concentration of 25 µg/ml.

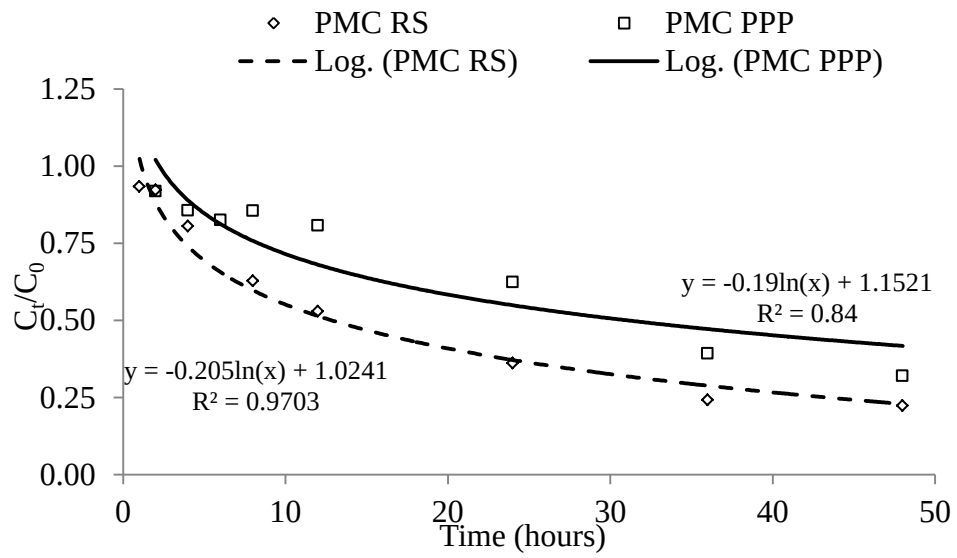
(FH 1 NaCl): FH sediment and 5 mM NaCl solution, initial DNA concentration of 1 µg/ml.

(FH 2.5 NaCl): FH sediment and 5 mM NaCl solution, initial DNA concentration of 2.5 µg/ml.

(FH 7.5 SW): FH sediment and stream water, initial DNA concentration of 7.5 µg/ml.

(FH 15 SW): FH sediment and stream water, initial DNA concentration of 15 µg/ml.

A



B

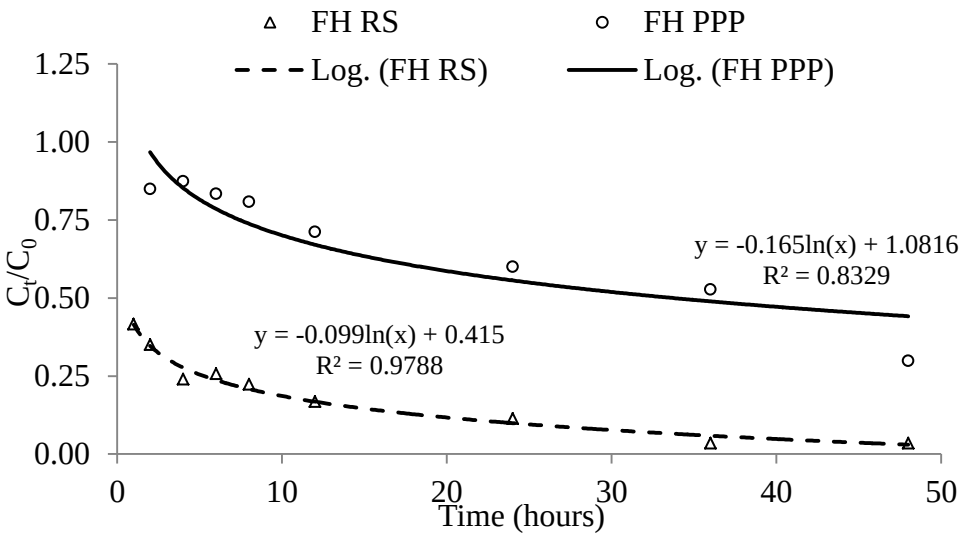


Fig. 5.1 Raw sewage DNA and purified PCR product sorption kinetics on PMC (A) and FH (B) sediments in 5 mM NaCl solution

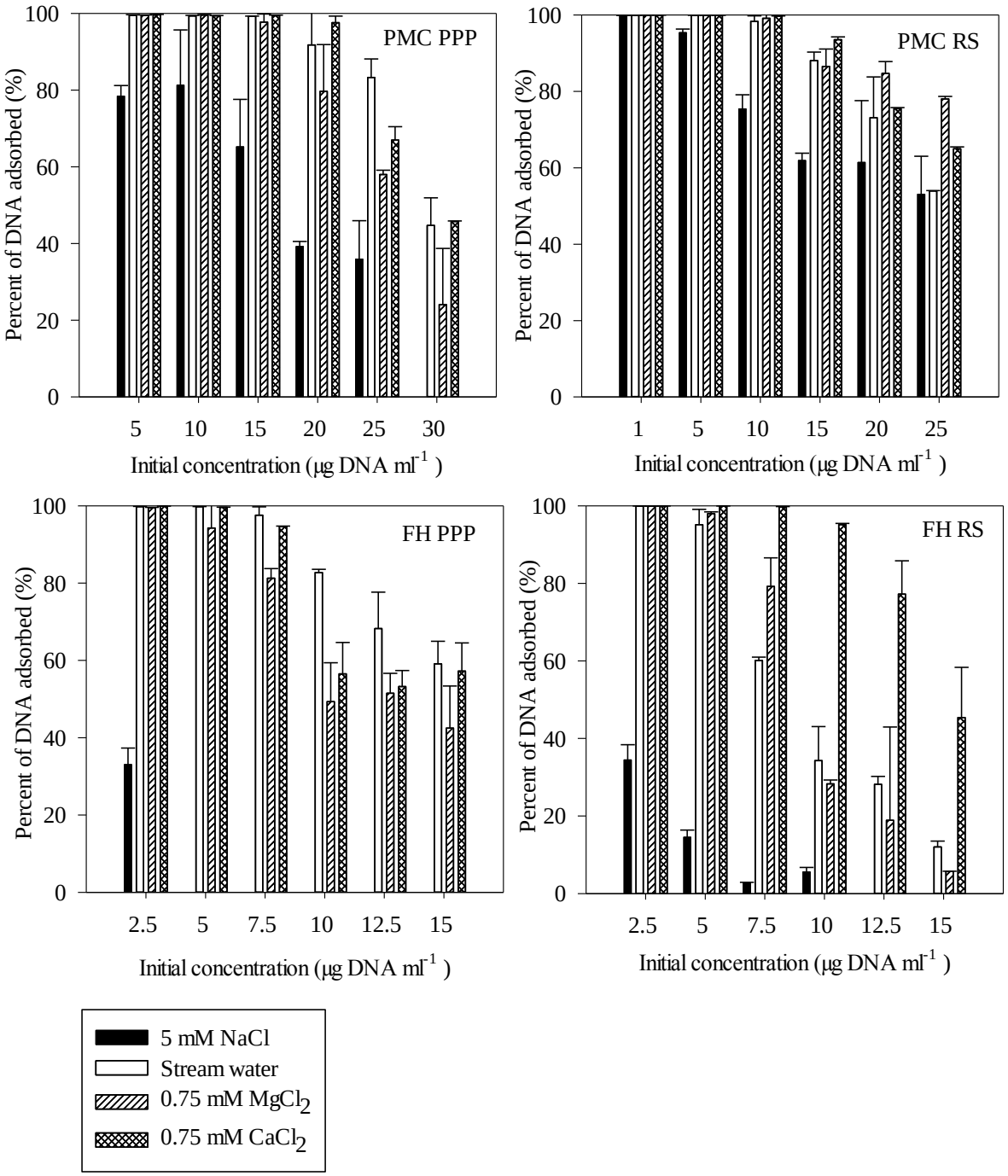


Fig. 5.2 Percent of raw sewage DNA and PPP DNA adsorbed on PMC and FH sediments in NaCl, stream water, $MgCl_2$, and $CaCl_2$ with different DNA concentrations

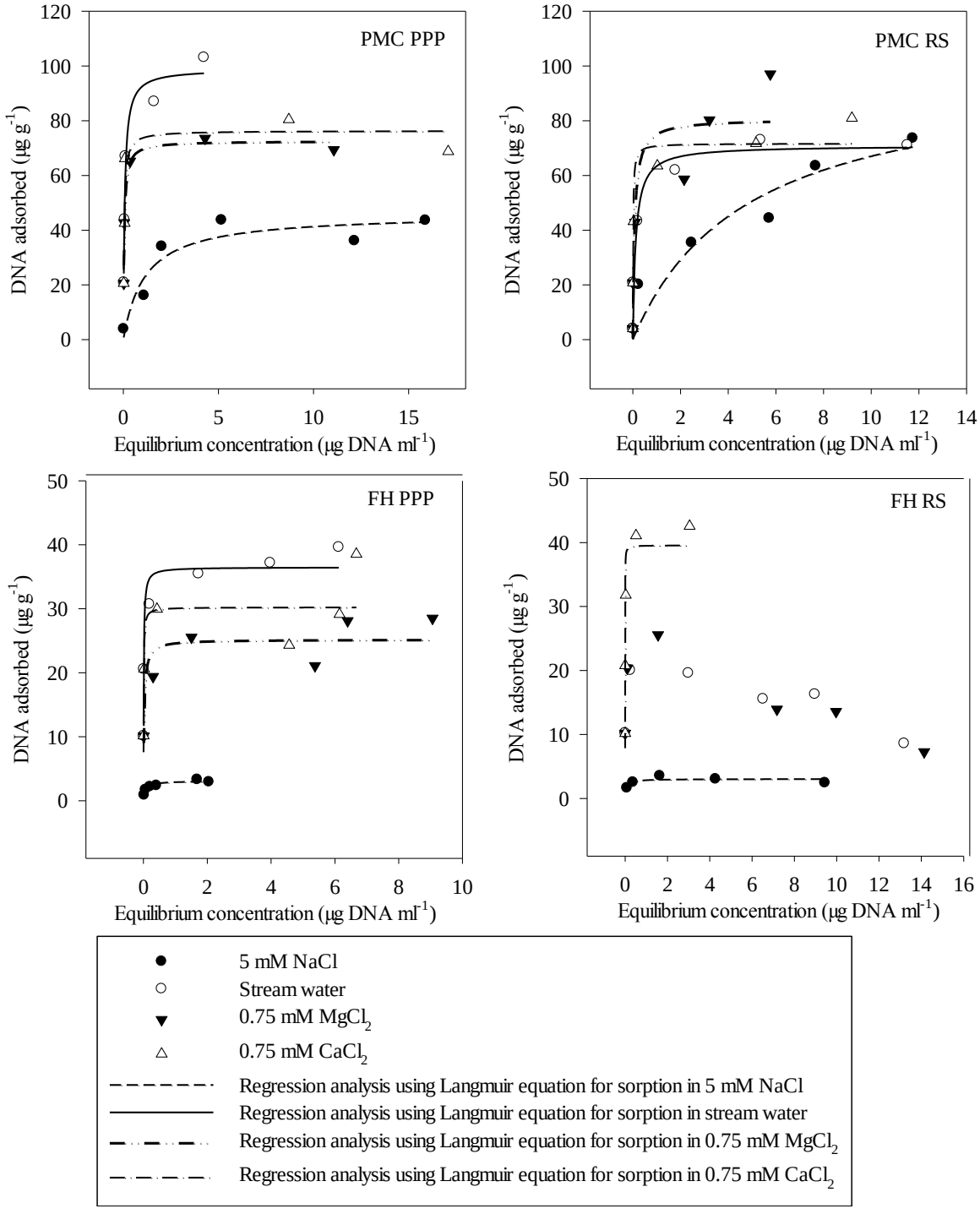


Fig. 5.3 Raw sewage and PPP DNA adsorption isotherms fitted using the Langmuir equation with PMC and FH sediment in different solutions

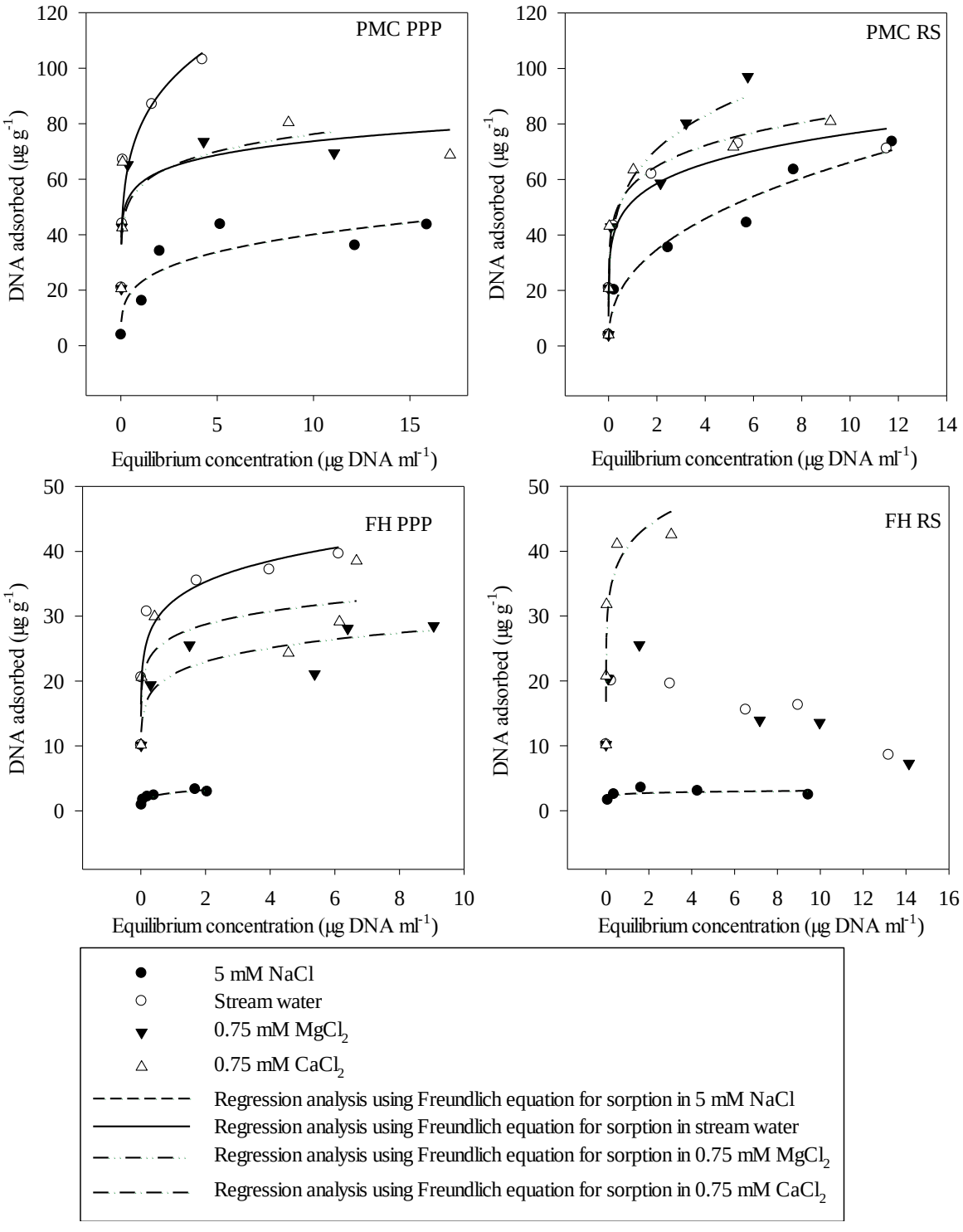


Fig. 5.4 Raw sewage and PPP DNA adsorption isotherms fitted using the Freundlich equation with PMC and FH sediment in different solutions

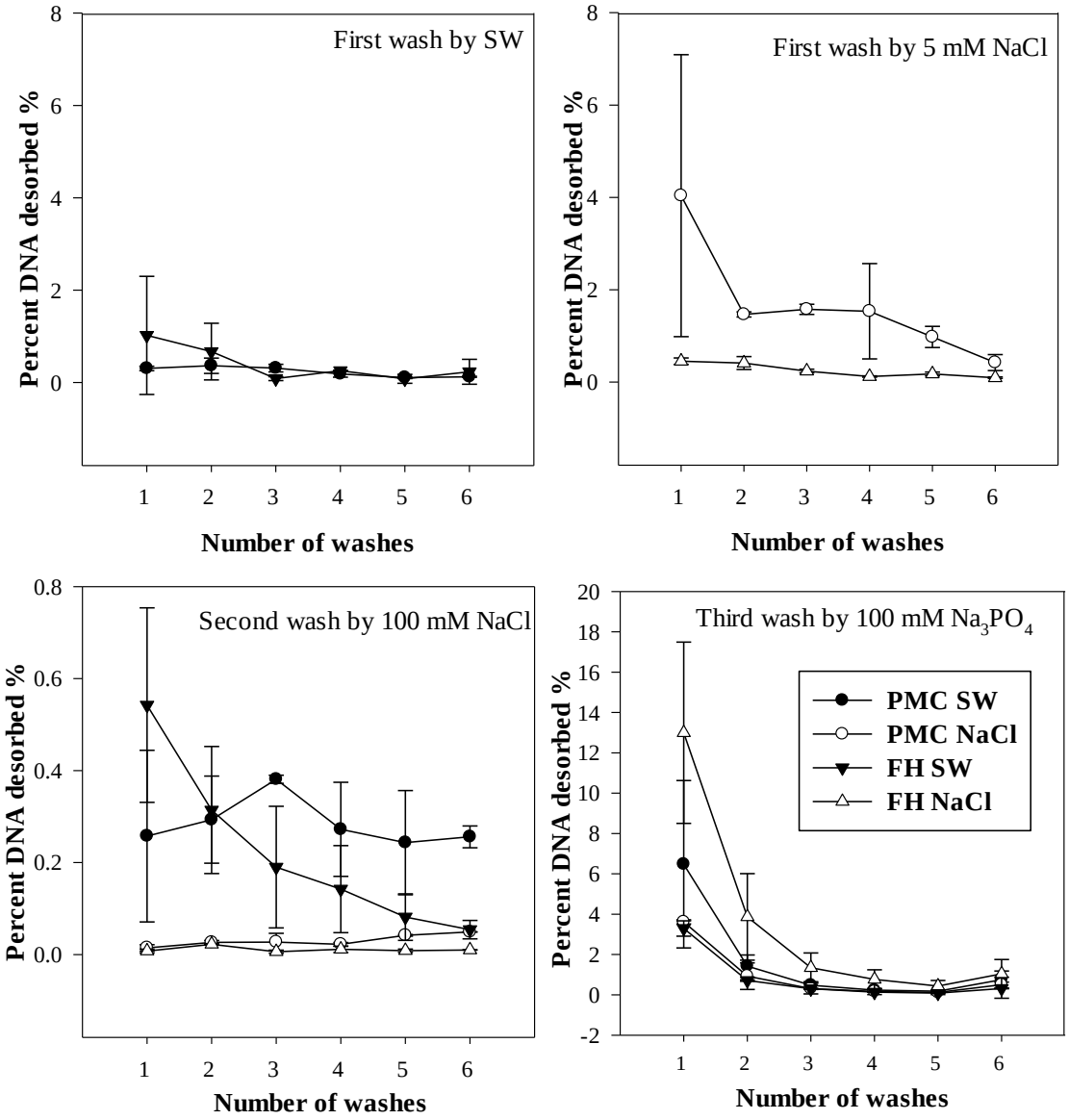


Fig. 5.5 Desorption of DNA from PMC and FH sediments.

(PMC SW): PMC sediment and stream water used for DNA equilibration prior to experiment

(PMC NaCl): PMC sediment and 5 mM NaCl solution used for DNA equilibration prior to experiment

(FH SW): FH sediment and stream water used for DNA equilibration prior to experiment

(FH NaCl): FH sediment and 5 mM NaCl solution used for DNA equilibration prior to experiment

Chapter 6

Summary and future directions

This study was conducted to improve and refine qPCR-based MST methods. Among host-associated genetic markers evaluated, one of the two cattle and goose markers and the elk/deer marker had acceptable specificity and sensitivity and can be used in tracking sources of fecal contamination in Alabama. A more practical and reliable approach was developed to determine the lower limit of detection and quantification for qPCR assays. When qPCR-based methods are used in MST, we recommend that specificity and sensitivity should be determined for the study area, and LLOD and LLOQ at both analysis and process levels be included in the marker performance evaluation.

Adsorption of DNA by sediment increases the persistence of free DNA in the aquatic environment and thus causes ambiguities in the identification of recent fecal pollution sources when qPCR-based methods are used. This study showed that the chemical nature of the surface water and clay contents in sediments influenced DNA adsorption, and that significant amounts of DNA can be desorbed from the sediments. Future research is needed to better understand the influence of sediments on the outcome of MST studies.

Although qPCR holds great promise for MST, current challenges in its application to a variety of environments and situations are numerous. Further work is needed to develop better and more host-specific markers. Multiple host-specific genetic markers for the same target species may be needed in future MST studies. Studies have shown that recent contamination events had a stronger correlation with risks associated with exposure to polluted waters compared with older pollution events. Future efforts should be directed toward the detection of live fecal bacteria in surface waters, which reflect the most recent fecal pollution.