8-2015

Microbial Source Tracking of Nonpoint Source Pollution in New Jersey Rivers

Matthew John Newton
Montclair State University

Abstract

The bodies of water of New Jersey serve many different purposes from recreational to agricultural and drinking to waste water treatment. Due to the vast diversity of usage and the many people that rely on these water bodies for everyday life it is essential that the health and quality of the water bodies is monitored and maintained regularly. Sources of pollution that affect the health of rivers, lakes and streams include biological, microbial, physical and chemical contaminants. All of these pollutions can impact the health of the water body, the organisms living in it and those who come into contact with the contaminated water. This research focuses on the microbial contaminants of the bodies of water using microbial source tracking (MST) techniques to determine the presence or absence of fecal matter contamination from different species-specific sources. Primer development and optimization lead to the utilization of a PCR based-assay and a real-time PCR (qPCR) based-assay and species-specific primers, which were used to determine a relation between the land cover and land use by the sources of contaminants that were found in the water bodies tested in specific areas representing different types of land usages (agricultural, urban and forested). We found that nonpoint source pollution is higher during rain events. The results obtained identified that the agricultural land use is a higher contributor to nonpoint source pollution than urban and forested land uses. Also, we were able to identify nonpoint source pollution from Canadian goose, cow, deer, dog, horse and human throughout the sampling areas tested. Using qPCR based-assay and a copy number equation we were able to quantify the most dominant sources of contaminants in the agricultural area. In this study, horse was found to be the most dominant.

MONTCLAIR STATE UNIVERSITY

Microbial Source Tracking of Nonpoint Source Pollution in New Jersey Rivers

by

Matthew Newton

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

August 2015

College/School: College of Science and Mathematics

Department: Biology

Thesis Committee:

[redacted]

Meiyin Wu, Ph.D.
Thesis Sponsor

[redacted]

Lee Lee, Ph.D.
Committee Member

[redacted]

John Gaynor, Ph.D.
Committee Member

MICROBIAL SOURCE TRACKING OF NONPOINT SOURCE POLLUTION IN NEW JERSEY RIVERS

A THESIS

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

by

MATTHEW JOHN NEWTON

Montclair State University Montclair, NJ

2015

TABLE OF CONTENTS

List of Figures

List of Tables

Introduction

Since the legislation of the Clean Water Act in 1972, much of the water systems in the United States have been monitored for various kinds of contamination. Pollution sources have been identified and there are now several different agencies overseeing those types of contamination. There are three major categories of concern when analyzing water quality and they are chemical, physical and microbiological. The chemical quality of water can be susceptible to changes when there are industrial or agricultural sources found in a watershed. Agricultural sources can contribute animal wastes, commercial fertilizers, pesticides and herbicides, and nutrients from feeds and waste production. Some of the physical issues of water quality concern are suspended solids that cause turbidity and reduce light penetration and thermal pollution. The microbiological issues are the presence of pathogenic microbes in the water supply. The microorganisms, including the indicator bacteria, are associated with fecal material from humans and other animals and if detected, they can signal that there may be enteric pathogens as well (Ashbolt, 2001). Although chemical and organic pollutions are important to monitor and control there has been an increased interest in pathogenic microbe, which is most important for human recreational use, drinking and aquaculture. Deterioration of water quality by pathogenic microbes primarily occurs due to human and animal fecal contamination, as well as surface runoff from agricultural and rural regions (Murugan et al., 2012).

Contamination of the rivers, lakes and streams is very common in the U.S. today even under the implementation of the Clean Water Act. For example, approximately 35%, 45% and 44% of assessed river, lakes and estuaries, respectively, have been

classified as impaired based on the pollutant levels (USEPA, 2000). Across the U.S. the condition of water varies from pristine to highly polluted, the Clean Water Act was passed as a way to try to eliminate waterborne microbial disease but more still needs to be done. New Jersey's water bodies have been assessed by the Environmental Protection Agency as well. In 2010, the EPA found that 17,089.0 miles of rivers and streams are classified as impaired; 45,307.5 acres of lakes, reservoirs and ponds are impaired; 664.9 square miles of bays and estuaries are impaired and 514.6 square miles of ocean and near coastal waters are impaired in New Jersey (USEPA, 2014). Compared to the water statistics of the Nation, which has 559,784 miles of impaired rivers and streams, 12,224,883 acres of impaired lakes, reservoirs and ponds, 26,120 square miles of impaired bays and estuaries and 1,059 square miles of impaired ocean and near coastal waters (USEPA, 2014). The EPA has also listed what they believe to be the causes of impairment for certain water bodies, and in their findings they show that for rivers and streams, of the 17,089 miles of impaired water in NJ that 5,198.3 miles are impaired by fecal coliform (USEPA, 2014). Another proposed finding by the EPA were the probable sources of impairment for rivers and streams of New Jersey, found that 13,690.6 miles have a probable source of contamination by urban runoff, 11,572.9 miles by agriculture, 1,641.8 miles by municipal point source discharges and 695.4 miles by natural sources or wildlife (USEPA, 2014).

The EPA has published a protocol for managing impairments of different water bodies by using total maximum daily loads (TMDL) of pathogens (USEPA, 2001). TMDL is a calculation of the maximum amount of a pollutant that a water body can receive and still meet water quality standards (USEPA, 2014). These TMDLs can also be

utilized to establish contaminant load allocations among point and load allocations among nonpoint pollutant sources (USEPA, 2014). It is very difficult to determine if the causative agent originated from a nonpoint pollution source such as, runoffs from agriculture, forestry, wildlife and urban landscapes. However, as difficult as it may be to pinpoint a specific source, if a potential source could be identified, management and remediation efforts of those water bodies would be more cost effective with resources allocated correctly.

Basic detection methods for pathogenic microbes in water sources includes; the culturing and enumeration of fecal indicator bacteria such as fecal coliforms, *E. coli* and fecal enterococci. However, there has been pressing interest in the field of microbial source tracking, which is a way of molecular fingerprinting not only for detection but also for identification of fecal contamination sources.

What makes it possible for microbial source tracking to be effective? Bacteria are among the most common biological pollutants affecting assessed rivers and streams. The presence of these bacteria may provide evidence on the presence of fecal contamination that can have significant human health impacts. These bacterial indicators can provide insight as to what the water is and is not safe to be used for. For instance, a presence of indicators would show that the waters could be contaminated; therefore any recreational or potable usage of that water should be halted. Several methods have been established for microbial source tracking including a TaqMan assay (Kildare et al., 2007). This assay is based on the amplification of fecal 16S rRNA marker sequences from uncultured cells of *Bacteroidales*. These *Bacteroides* order bacteria are exclusively found in fecal material, animal rumen and other cavities of human and animals (Kildare et al., 2007).

While it is important to determine whether the contamination is from point or nonpoint source pollution, the challenge is in pinpointing a specific source within those various nonpoint pollution sources. Some of the potential nonpoint pollution sources that are important for microbial source tracking are farmlands, pastures for livestock and animal feedlots (Edwards et al., 2000). Other important potential sources include urban runoff, combined sewer overflows, broken sewage lines and faulty septic tanks (Schiff et al., 2001). Another important nonpoint pollutant source is wildlife and domesticated animals. These sources could release substantial contamination into aquatic systems, however, these are among the hardest to identify and remediate (Blankenship, 1996 and Simmons et al., 2000). Researchers are able to identify cluster-specific primer sets that are useful in discriminating from human and other species (Kildare et al., 2007).

There are a few different methods of microbial source tracking, the methods can be molecular or biochemical. The method chosen will determine the target of the study whether it is a certain microbe or a chemical such as caffeine. Of the molecular methods, which will be the focus in this study, there are four different bacteria that are typically targeted in MST studies: *Bifidobacterium, Bacteroides, Enterococcus and Escherichia* (Simpson et al., 2002). Using one of these bacterium and nucleic acid based methods along with polymerase chain reaction (PCR) based-assay it is possible to determine their presence in the environment. In 2005, the EPA released a guide to microbial source tracking (Edge et al., 2005). In this guide the agency builds upon the history of cooperative work among the many different studies performed by academic partners, and other state and federal agencies to assess the best way in which to perform microbial source tracking tests (Edge et al., 2005). The EPA states that there are a multitude of

different tests being used to identify fecal sources such as chemical analyses and the testing of fecal constituents however, these methods are met by some with disapprobation due to issues in reliability and sensitivity (Edge et al., 2005). A number of microbial source tracking methods have been developed to track animal sources within natural waters (Edge et al., 2005). With the use of these methods it is assumed that, given the appropriate method and source identifier the source of pollution can be detected using molecular techniques to amplify a genetic marker via PCR step (Edge et al., 2005).

Among the most crucial aspects of microbial source tracking studies is the specificity of the methods being used. Of the host specific primers used in the PCR assay, many were published in scientific literature and were tested to ensure specificity (Bernhard et al., 2000). However, not all the published primers had reproducible results and therefore, a very important part of this study is to develop and design host specific primers utilizing DNA sequencing techniques in conjunction with bioinformatics (Ebentier et al., 2013). The obtained primers will be used to analyze the environmental samples, to determine the specificity and sensitivity of the primers by using PCR based-assay. These primers will be proven to amplify the target DNA by testing the primers using a positive control, fecal samples taken directly from the desired host animals. The PCR products will then be sequenced, run through a homology test through the National Center for Biotechnology Information nucleotide BLAST software. After a presence has been established for a given host at a certain site the next step is to determine a quantity or the level of contamination by using qPCR methods. Using quantitative real time PCR technologies and the designed specific primers, the amount of DNA can be quantified and

this can give an approximation as to how much contamination from a specific host is present.

As outlined previously, water bodies currently being used for recreational, potable or agricultural uses are continuing to have growing concern of bacterial, specifically pathogenic bacteria, contamination. Determining the specific hosts in certain areas would be a useful tool in the management efforts for those resources. Different land uses and land covers should have different potential pollution sources associated with them. Another purpose of this study along with the presence, absence and quantifying of host specific contamination, is to examine the dominant types of host contamination associated with the type of land use and land cover. This information will be useful in pollution eradication efforts and environmental management practices.

Research Objectives

The objectives of this study include to 1) develop and optimize a species specific primers for PCR based-assay and real-time PCR based-assay 2) use PCR based-assay to identify dominant contaminant sources associated with various types of land use and land cover 3) compare and contrast the results of the PCR assay and the real-time PCR assay and 4) use the real-time PCR assay to determine copy number estimates of contaminants by species.

Methods & Materials

1. Study Sites

The sampling sites were chosen in three different sub-basins, thirteen sites in an agriculture dominant watershed with a county population density of 301.6 people per square mile (Musconetcong watershed), five sites in urban dominant watershed with a county population density of 6,241.4 people per square mile (Third River watershed) and six sites in a forest dominant watershed with a county population density of 284.1 people per square mile (Flat Brook watershed) (US Census Bureau, 2013).

Map of Sampling Sites

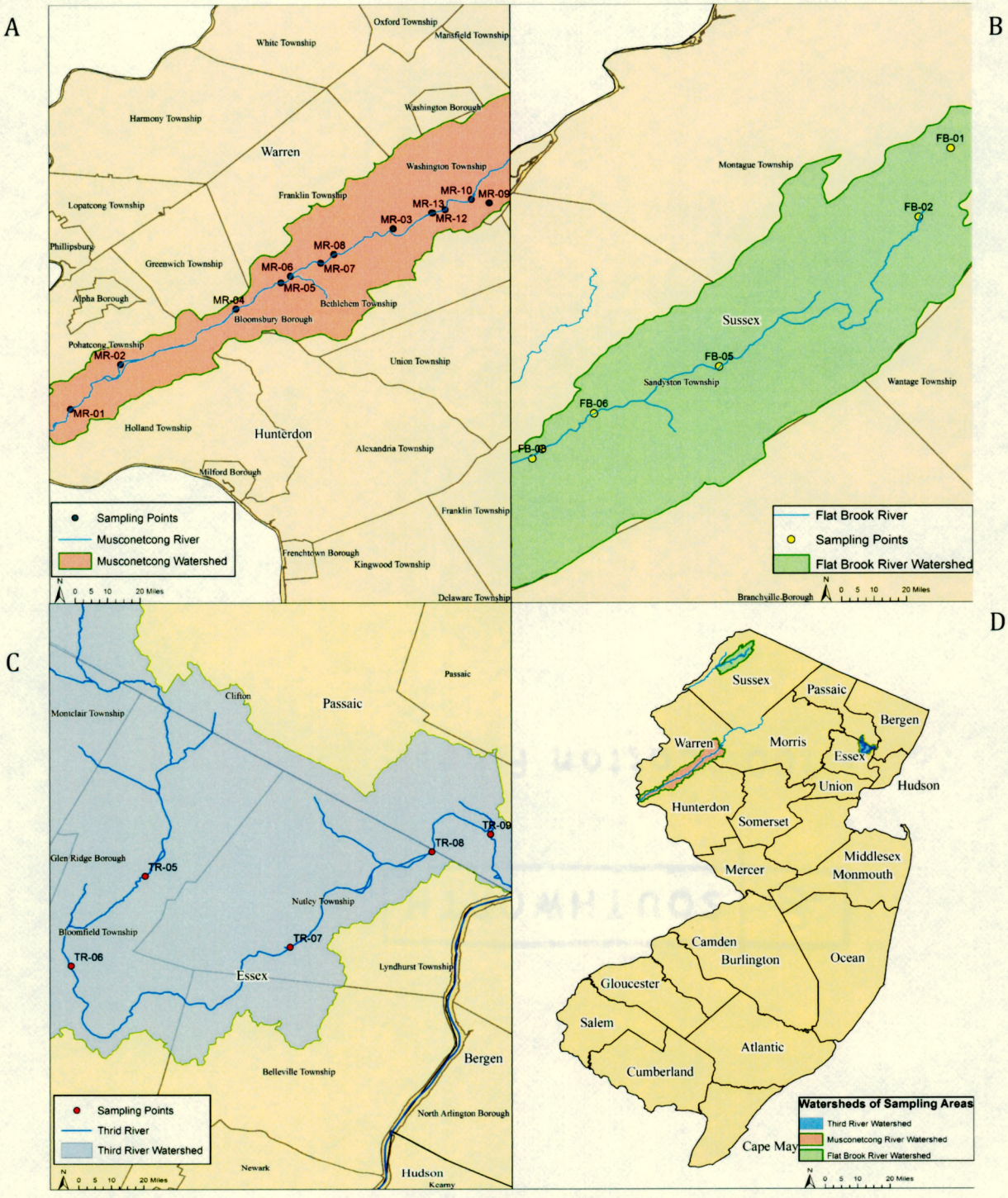


Figure 1. Map of the sampling areas for A) Musconetcong River, B) Flat Brook River and C) The Third River. D) Depicts the rivers and their watersheds and the location of each river in the state.

2. Sample Collection

The collection of water samples was divided into two separate days, a rain event and a dry or non-rain event. Water samples were collected by facing upstream and placing the 500 mL bottle at least six inches, where permissible, below the surface allowing the water to flow into the bottle until full. While collecting the samples, gloves were worn as to not contaminate the water. The samples were collected, stored on ice, transported back to the laboratory and processed within six-hours. The sites were selected at a previous time and the GPS coordinates were recorded to ensure that any additional samples were taken from the same location.

3. Filtration and Culturing

Once the samples were brought back to the lab, 100 mL of the samples were filtered using a vacuum filtration apparatus, including a vacuum filtering flask, Millipore filter holder, clamp, base and stopper. The filters used were Millipore Mixed Cellulose Ester Gridded 0.45 μm HAWG. The filters were then removed using aseptic techniques and placed onto agar plates that support the growth of coliform bacteria. These plates were then incubated for 24 hours at 38°C. After the incubation period, the filters were loaded into tubes containing nutrient broth to be kept as stocks for extraction and kept in refrigeration at 4° C.

4. DNA Extraction

The extraction of fecal DNA followed a Nutrient Broth Exclusion Chelex Extraction Method. Sterile 1.5 mL eppendorf tubes for each sampling site were labeled with the site code. The stock culture tubes were then mix on the vortex at 10,500 rpm for 30 seconds to re-suspend the culture. Next, 100 μL was transferred from each culture

tube into the corresponding eppendorf tube. The eppendorf tubes with the sample were then centrifuged at 14,000 rpm for 10 minutes in 4° C. Once complete, the supernatant was removed and discarded, while keeping the pellet in the tube intact. The pellet was then re-suspended in 100 µL of sterile deionized water and the tube was vortexed for 30 seconds on 14,000 rpm. Each tube then received 100 µL of 5% InstaGene Chelex 100 resin and vortexed again for 30 seconds at 10,500 rpm. A Sherlock cap was then placed on each tube before being placed in a water bath of 100° Celsius for 10 minutes. After the allotted time, the tubes were placed on ice and allowed to cool to room temperature, the next step was to centrifuge at 14,000 rpm for 5 minutes. The supernatant was then transferred into new sterile 1.5 eppendorf tubes with the same label, this supernatant was spun down for 2 minutes in the centrifuge at 14,000 rpm to ensure no Chelex beads or debris were transferred with the supernatant. Once the supernatant was isolated a recording of the DNA concentration and purity at 260/280 were recorded using the NanoDrop machine and the samples were stored in the freezer at -20° C.

5. PCR Assay Methods

The previous method described the extraction of DNA from coliform bacteria from each of the sampling sites. This DNA was then used to perform Presence/Absence tests using a PCR based-assay and gel electrophoresis. The PCR based-assay amplifies the host specific DNA according to the primers used in the reaction. The steps for this method included labeling sterile PCR tubes with the codes for each site and host being tested. Next, the reagents were added as follows: 5.0 µL water, 1.0 µL forward primer, 1.0 µL reverse primer, 12.5 µL GoTaq Colorless master mix and 5.5 µL of DNA (previously extracted). The PCR run parameters varied depending on the optimal

temperatures for each of the primer sets for specific hosts, but the basic foundation of the assay had a denature stage of one cycle of 94° C for four minutes, the annealing stage has 40 cycles of 94° C for one minute, varied temperature for 40 seconds, and 72°C for one minute, the final extension stage was 72°C for five minutes.

6. Positive Control DNA Extraction

The positive controls, used in the Presence/Absence gel electrophoresis assays, were taken directly from a specific host's fecal waste. The fecal material from human, dog, horse, cow, deer, pig and goose were collected from local parks and farms or were volunteered into the study. DNA from these positive control samples were extracted using a QIAmp DNA Stool Mini Kit for DNA purification from stool samples. This kit provides a quick and efficient purification of total DNA from fresh of frozen stool samples, ideal for PCR and other enzymatic reactions. The kit requires a stool sample weighing approximately 220 mg but the sample weight can vary. A larger sample size is recommended to start with if the DNA is not distributed homogeneously throughout the stool or is at low concentration. The first steps of the protocol involve weighing out your stool sample and placing it into a 2 mL centrifuge tube. Next, 1.4 mL Buffer ASL was added to the tube and vortexed immediately until the sample and buffer are thoroughly homogenized. The sample is then incubated for five minutes at 70° C and vortexed again for fifteen seconds followed by centrifuge for one minute at 14,000 rpm to pellet stool particles. 1.2 mL of the supernatant was transferred into a new 2 mL centrifuge tube and one InhibitEX tablet was added to each tube. The tubes were vortexed immediately for one minute or until the tablet was completely dissolved, then the tubes were incubated at room temperature for another minute to allow for inhibitors to absorb to the InhibitEX

matrix. The supernatant and tablet solution was then centrifuged at full speed for three minutes to pellet the inhibitors. The supernatant from this mixture is then transferred into a new centrifuge tube and centrifuged again for another three minutes. 15 µL of Proteinase K into a new 1.5 mL centrifuge tube and the supernatant from the previous step was added to that. Next, 200 µL of Buffer AL was added and vortexed for fifteen seconds. The tubes were then incubated for ten minutes at 70° C. 200 µL of ethanol 96-100% was added to the lysate and was mixed by vortexing. The lysate was transferred into a QIAamp spin column and a 2 mL collection tube and the contents were centrifuged for one minute at 14,000 rpm. The spin column was placed into another collection tube and 500 µL of Buffer AW1 was applied to the filter and centrifuged for another minute. The same spin column received 500 µL of Buffer AW2 and centrifuged for 3 minutes and finally the spin column was transferred into a 1.5 mL centrifuge tube and 200 µL of Buffer AE was pipetted directly onto the QIAamp membrane. The column was incubated for one minute at room temperature and then centrifuged at 14,000 rpm for one minute to elute the DNA.

7. Gel Electrophoresis

A 1X agarose gel was prepared from Tris-Acetate-EDTA, 1X solution and ThermoScientific TopVision agarose with Ethidium Bromide. Ten lane combs were used to make wells for the samples to be loaded into the gel in the Fisher Scientific Electrophoresis Systems Gel Box. There was a positive, negative, HiLo DNA Marker and samples for each site in the gels. The gels at 90 Volts and were then analyzed under UV light and recorded. When a DNA band matched the band of the positive control that site was deemed positive for fecal contamination by that specific host.

8. qPCR Quantitation Methods

The primers used for these tests were the same as those used in the PCR based-assay for most species, goose, human, dog, deer and horse. The real time PCR assay measures the amount of DNA after each cycle via the fluorescent dyes that produce a fluorescent signal in direct proportion to the number of amplicons generated ("Real-time PCR Handbook", 2015).

After all primers were acquired the assay could be performed on an Applied Biosystems Veriti real-time PCR thermocycler. Using the primers, water and SYBR Select Master Mix, a total master mix was made for all individual species to eliminate more possibilities for contamination or error. The final input into each well, excluding the negatives, of the 96 well plates used was 2 uL of DNA whether that be from a standard or from a sample site and 23 uL of the Master Mix solution. The standard curve control was run in a 1:10 dilution beginning at 1.0 nM thru 1.0 x 10^{-4}nM. The negative controls received the 23 uL of Master Mix and 2 uL of water. The experimental steps were the same as the PCR based-assay in respect to the steps and temperatures used with the addition of a melt curve at 1 cycle of 95°C for 15 seconds, 60°C for 1 minute and 95°C for 15 seconds. The method for the this qPCR test was a quantitation method where the resulting threshold cycle (Cq) and melting temperature (Tm) would be used to determine if a sample contained (presence) or lacked (absent) the DNA that was to be amplified by the primers for the specific host.

9. Copy Number Calculation

The qPCR process generated a standard curve for each positive control from the previously amplified positive control from each species in the 1:10 dilution previously

stated (Appendix A). Gene copy numbers were estimated by taking the standard curve equation and extrapolating the quantity of the sample. This calculation is based on the assumption that the average weight of a base pair is 660 Daltons. This means that one mole of a base pair weighs 660 g and that the molecular weight of any double stranded DNA template can be estimated by taking the product of its length and 660. The inverse of the molecular weight is the number of moles of template present in one gram of material. Using Avogadro's number, 6.022×10^{23} molecules/mole, the number of molecules of the template per gram can be calculated: mol/g * molecules/mol = molecules/ g. Finally, the number of molecules or number of copies of template in the sample can be estimated by multiplying by $1*10^{9}$ to convert to ng and then multiplying by the amount of template.

Equation: Number of copies (molecules) = $\frac{ng * 6.0221 \times 10^{23} \text{ molecules/mole}}{(\text{Basepairs} * 660 \text{ g/mole}) * 1 \times 10^{9} \text{ ng/g}}$

10. Statistical Analysis

A statistical analysis, analysis of variance or ANOVA was run on the type of land cover and *E. coli* enumeration using JMP Pro 11 software.

Results and Discussion

Primer Optimization

A key component of MST practices is the specificity and optimal performance of the primers being used. For this project there was many steps taken in order to optimize the performance of the primers for our purpose. To begin all the primers were run at 55°C to first detect any amplification. For those that did not appear after staining the temperature was dropped and rerun until a single band appeared in the gel. The other samples that produced non-specific binding were run in a temperature gradient of one-degree differences until a single band at the location of interest was acquired. Due to the fact that the concentrations of extracted DNA in the environmental samples were high, the volume of sample being used in the gel electrophoresis had to be lowered in order to obtain clean single bands in the gels. Another parameter that was modified throughout the process was the extension time. The extension time was modified in order to target our sample and trying to eliminate amplifying excess DNA in the environment, this was accomplished by lowering the extension time of the PCR process. Finally, to test the specificity of the primers each primer was run against all the positive samples and only amplified the host it was designed for.

Positive Control Homology

Before performing the PCR based-assay for microbial source tracking the positive controls had to be verified by sequencing. Using the host specific primers found in literature and developed in the laboratory, DNA extracted from the feces of the respective species was amplified in a PCR reaction and sequenced. These sequences were then run

through GenBank to search for homologies in other known samples. It is important to remember that the DNA being tested is not chromosomal DNA but the DNA from the bacterial normal flora residing within the species digestive tract, gut and fecal matter.

The homologies found for human samples matched to sequences found in GenBank for *Bacteroides dorei* at 96% homologous to the original sequence. The *Bacteroides dorei* bacterium is a part of the normal flora of the human gut metagenome. In a study by Francesc Peris-Bondia et al., when they tested the microbial makeup of the human gut and tested it through fecal samples they found that *Bacteroidetes* was the second most abundant phyla of bacterium found. Also, they found that when testing fecal samples two families of *Bacteroides* were found in all samples, *Bacteroides dorei* and *Bacteroides uniformis* (Peris-Bondia et al., 2011). For the cow positive control, there were two separate matches from two different isolation sources. The first was a 91% homology to *Bacteroides Prevotella* taken directly from cow feces and the other was a 92% homology to *Bacteroides* bacterium isolated from cow manure. This shows that the cow primers selected for this identification and tracking process are specific to the cow but yet broad enough to be found in multiple cow related sources. The horse positive control sequence showed similar results to that of the cow. Where 97% homology was found between the positive controls sequence and a bacterium isolate from horse feces and 98% to isolate from horse manure. The GenBank results from the dog positive control sequence were quite interesting to see. When first discussing the parameters of this project at the mention of dog samples, the first thought was your typical domesticated dog. The domesticated dog is commonly walked in the parks along the streams and rivers and the owner leaves the feces on the ground where it is washed in the

water body. However, the notion of wild dogs, foxes, coyotes and wolves did not occur to me until after the sequencing results. The positive control matched 99% to sequences from wild dog feces, wolf feces and to domesticated dog-fecal bacterium. This is important for analysis of the land use, land cover data later. The pig positive control matched 97% to bacterium from pig feces and to *Bacteroides* bacterium from pig slurry, which is a mixture of animal waste, in this case from pigs, to be used as fertilizer. The deer positive control is different from all the previous positive controls. The target gene for the deer is the cytochrome b gene, which is mitochondrial DNA. Therefore, the sequence match that was found in GenBank of 100% match to a protein found in the white-tailed deer is in line with what is being searched for in the environmental samples. Finally, the Canadian Goose positive control sequence was run against the whole genome sequence of the cytochrome B gene of the Canadian goose and was found to be a 100% match.

Human Positive Matches

One of the human positive controls was 96% homologous, with an E value of $3e^{-49}$, to a sequence found in *Bacteroides dorei* isolate taken from human gut metagenome, "High abundance of *Bacteroides dorei* in the human gut precedes and predicts type 1 diabetes autoimmunity" by A.G. Davis-Richardson (Davis-Richardson et al., 2014).

The human positive control was also 96% homologous, with an E value of $3e^{-49}$, to a sequence found in *Bacteroides dorei* isolate taken from human gut metagenome in a different study, "The methylome of the gut microbiome: disparate Dam methylation patterns in intestinal *Bacteroides dorei*" by M.T. Leonard et al. (Leonard et al., 2014).

Cow Positive Matches

A cow positive control was 91% homologous with an E value of $3e^{-137}$, to a sequence found in *Bacteroides Prevotella* taken from cow feces, "Quantification of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater" by S. Okabe et al. (Okabe et al., 2007).

The cow positive was also 92% homologous with an E value of $2e^{-155}$, to a sequence found in *Bacteroides* bacterium isolated from bovine manure, "Phylogenetic analysis of *Bacteroidales* 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR" by Sophie Mieszkin et al. (Mieszkin et al., 2010).

Horse Positive Matches

A horse positive control was 97% homologous with an E value of $5e^{-26}$, to a sequence obtained from uncultured bacterium isolated from horse feces, "Changes in fecal bacteria associated with an all-haylage diet in trotting horses" by B. Willing et al. (Willing et al., 2009).

The horse positive control was also 98% homologous with an E value of $2e^{-30}$, to a sequence obtained from uncultured bacterium isolated from equine manure sample, "Assessment of equine fecal contamination: the search for alternative bacterial source-tracking targets" by J.M. Simpson et al. (Simpson et al., 2004).

Dog Positive Matches

The dog positive control was 99% homologous with an E value of $2e^{-40}$, to a sequence obtained from uncultured bacterium isolated from a dhole or wild dog feces, "Phylogenetic analysis of 16S rRNA gene sequences reveals distal gut bacterial diversity in dhole (*Cuon alpinus*)" by H. H. Zhang et al. (Zhang et al.).

The positive was also 99% homologous with an E value of $2e^{-40}$, to a sequence obtained from uncultured bacterium isolated from Canis lupus feces, "Phylogenetic analysis of 16S rRNA gene sequences reveals distal gut bacterial diversity in wild wolves (*Canis lupus*)" by H. Zhang and L. Chen (Zhang et al., 2010)

The positive control was also 99% homologous with an E value of $2e^{-40}$, to a sequence obtained from uncultured *Bacteroides sp.* isolated from canine feces, "*Bacteroidales* diversity in ring-billed gulls (*Laurus delawarensis*) residing at Lake Michigan beaches" by S. N. Jeter et al. (Jeter et al., 2009).

Pig Positive Match

The pig positive control was a 97% homologous with an E value of 0.0, to a sequence obtained from uncultured bacterium isolated from pig feces, "Molecular diversity of *bacteroidales* in fecal and environmental samples and Swine-associated subpopulations" by R. Lamendella et al. (Lamendella et al., 2013).

The pig positive was also 97% homologous with an E value of 0.0, to a sequence obtained from uncultured *Bacteroides* bacterium isolated from pig slurry, "Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific *Bacteroidales* 16S rRNA genetic markers" by Sophie Mieszkin et al. (Mieszkin et al., 2009).

Deer Positive Match

The deer positive control was 100% homologous with an E value of $1e^{-23}$, to a small sequence of cytochrome B protein from *Odocoileus virginianus*, or the white-tailed deer, found throughout much of the Northeastern United States. The sequence was found in, "Mitochondrial and Nuclear Phylogenies of Cervidae (Mammalia, Ruminantia):

Systematics, Morphology, and Biogeography" by Clement Gilbert et al. (Gilbert et al., 2006).

Goose Positive Match

The Canadian goose positive control sequence was tested against the whole genome of the cytochrome b gene found in *Branta canadensis*. The sequence was homologous 100% with an E value of 0.40, to a small piece of the sequence of the whole genome. The cytochrome B gene sequence was found in GenBank from, "Comparision of the mitochondrial genomes of the Canada Goose and the White Fronted Goose" by J.C. Snyder et al. (Snyder et al.).

Method Comparison

The method comparison was conducted between the PCR and gel electrophoresis assay and the qPCR assay for the 26 samples collected from Musconetcong River. Comparing the presence and absence results between the two tests, conclusions can be made about which test is better to use in regards to sensitivity and efficiency. Comparing the results from the PCR based-assay and the real-time PCR based-assay one can see that many of the results correlate well. The species that see 100% matches for all of the samples include, Canadian goose, deer, horse and pig. There are only a few sites for each of the other species, human, dog and cow that do not match completely. Of the total comparison only 12.09% of the tests did not match. Of those that did not match 54.55% were positive in the PCR based-assay and negative for qPCR based-assay and 45.45% were positive in the qPCR based-assay and negative in the PCR based-assay. There are some plausible explanations has to why the two tests showed different results at these sites. For example when looking at the cow samples, all of the non-rain results produced

a negative result, however, for the rain event samples two produced a positive result in the PCR based-assay. One possible cause for this could be nonspecific binding to material in the environment flushed into the system that resulted in a band in the gel at a length close to that of the positive but was not in fact the same. Another explanation for different results from the two tests is the sensitivity that the two tests provide. The PCR based-assay produces results at a much lower sensitivity than those produced by the real-time PCR. The qPCR is able to detect a much lower quantity of DNA that is amplified during the PCR process and therefore, those results where the PCR based-assay produced a negative result but the qPCR result was positive, it can be assumed that the amplified product was not enough to show up in the PCR based-assay but was found in the qPCR based-assay. These results are only proven to be valid when correlated to the Tm profile that matched the positive control in the standard curve.

As for the efficiency of the two tests in regards to time needed the comparisons are as follows. The PCR based-assay entails a DNA extraction and culturing before a test can be run that requires approximately three hours for the extraction and twenty-four hours for the culturing of the bacteria. The PCR in the thermocycler then runs for another three hours followed by a 45-minute gel electrophoresis before the results can be obtained. The real-time PCR requires a DNA extraction that takes approximately three hours followed by the experiment set up which takes about an hour. The next step is the run of the real-time PCR, which again is approximately three hours and the results are obtained directly after the run is completed. Therefore, the time allotted for each of the tests, for the PCR based-assay and gel electrophoresis is thirty hours and 45 minutes and for the real-time PCR based-assay the total time is approximately seven hours. Although

the qPCR based-assay is more expensive the run time of the experiments, sensitivity and results provided by the test, in my opinion, far outweigh the difference in cost to make it the more advantageous test.

Table 1. Non-rain Event Presence and Absence of the Musconetcong River sites in both the PCR based-assay and the real-time PCR based-assay

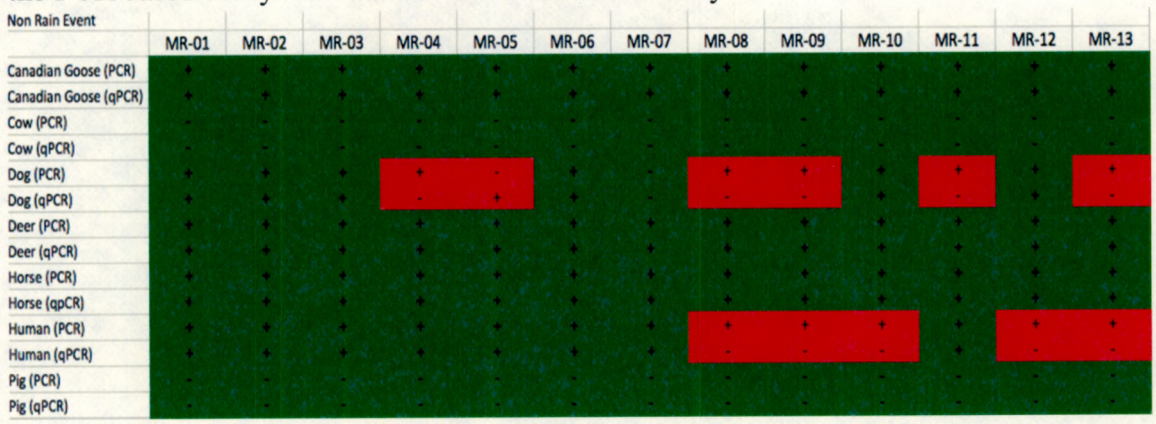
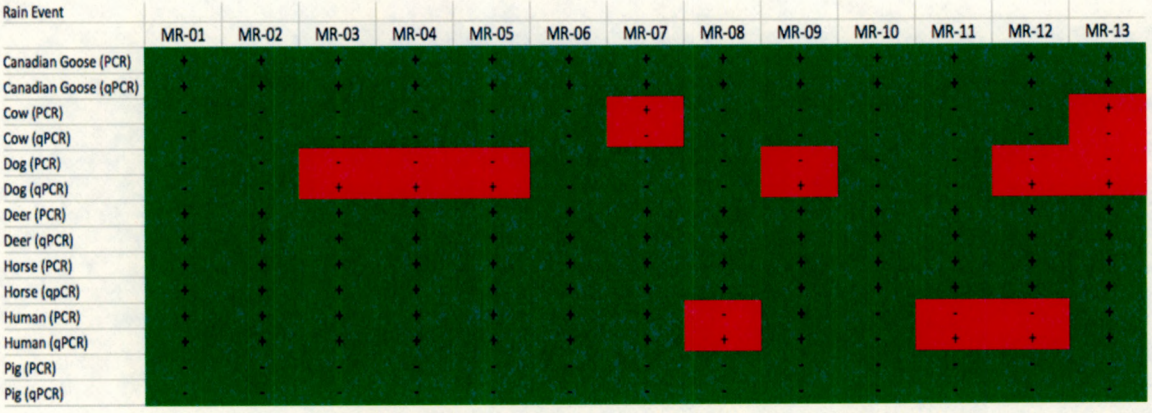


Table 2. Rain Event Presence and Absence of the Musconetcong River sites in both the PCR based-assay and the real-time PCR based-assay



Tables 2 and 3 show the presence (+) and absence (-) of the Musconetcong River sites in both the PCR based-assay and the real-time PCR (qPCR) based-assay. The cells that are green for each species show that both assays produced the same result. Those cells that are red show where the two assays produced opposite results. 87.91% of the

results match between the two assays however we do see some differences (12.09%) in the tests for human and dog in the non-rain event and human, dog and cow in the rain event.

Implications of *E. coli* Enumeration

E. coli acts as an indicator bacterium for fecal contamination in water bodies. Within the family of coliform bacteria the multiple genera of bacteria seen are *Klebsiella*, *Enterobacter* and *Citrobacter* including *E. coli*. However, *E. coli* is the only coliform that is an undoubted inhabitant of the gastrointestinal tract (Dufour, 1977). In an experiment performed by A.P. Dufour (1977), in 28 fecal samples tested the percentage of coliforms that were present in fecal samples, 96.8% were *E. coli*, 1.5% were *Klebsiella* and 1.7% were *Enterobacter* and *Citrobacter* group. These tests are supported by data from Prescott, Winslow and McCrady who found in their study that among the millions of coliform organisms per gram in feces, well over 95% of those are *E.coli* (Prescott et al. 1947). Therefore, having data that quantifies the number of *E. coli* colonies grown from the water samples gives clues as to how much fecal contamination could be expected in each sample at each site.

The *E. coli* colony count for each of the sites at the sampling areas provides insight on the general contamination of the area. *E. coli* is used in biological sampling as an indicator bacterium because it is easy to test for and its presence signals that there is a more than likely chance that other species of pathogenic bacteria are present in the waters as well; *E. coli* is the best indicator to predict health risk for recreational water contact. Therefore, by quantifying the indicator bacteria at each site, one could draw some information on the overall quality of the water before performing any additional tests.

Weather has an impact on bacteria present in the water. Non-rain event samples are collected when there has been no rainfall for 48 hours, rain event samples are collected during rainfall events. During a non-rain event, or dry sampling, the indicator bacteria counts at the sampling sites were relatively low from 63 to 165 *E. coli*/ 100 mL with the MR-08 at 482 *E. coli*/ 100 mL and MR-10 at 1174 *E. coli*/ 100 mL (Figures 2). However, the rain event counts were significantly higher in most of the same sampling sites. Where the lowest count was 83 *E. coli*/ 100 mL and most sites ranged between the 200 to 1010 *E. coli*/100 mL. MR-10 did see a significant decrease in *E. coli* count from 1174 *E. coli*/100 mL in the non-rain to 245 *E. coli*/ 100 mL in the rain event (Figure 2). The same pattern can be seen in the other sites for the Flat Brook river and the Third River where the count of *E. coli*/ 100 mL increases from the non-rain event sampling to the rain event sampling (Figure 3 & 4). This increase between the two types of sample events, non-rain and rain event is usually due to the rainfall washing fecal material, litter and other pollution sources into water bodies as runoff. Another contributor to an increase of bacterial cells is that the rain mixes up the sediment on the bottom of the water bodies releasing any trapped bacterial cells. This type of data can give a general outlook on the quality of water but it does not give any description as to what or who is contributing to the pollution. It also does not signify the how many different contributing sources of pollution there are for any one site. For more in depth information such as that more tests need to be run on the water samples.

The non-rain event data shows that the water samples collected from the agricultural sites contained the most *E. coli* compared to the other sites. The highest colony count from the agricultural sites was 1174 colonies at MR-10 and the lowest count

was 63 colonies at MR-12 (Figure 2). For the forested sites the highest colony count was 10 colonies at FB-06 and the lowest count was two colonies at FB-05 (Figure 3). The *E. coli* colony counts for the urban were the highest 62 colonies at TR-06 and the lowest was 10 at TR-07 (Figure 4). The rain event data for the agricultural area shows an obvious increase in almost all the sites, where the lowest count was 85 colonies at MR-02 and the highest colony count, 1010 colonies, at MR-09. The sites in the forested area all had an increase in colonies with the lowest count being 4 colonies at FB-08 and the highest count was 17 at FB-06. The colony counts for the urban area sites also all increased. The highest count was found at site TR-06, at 201 colonies, and the lowest count was 38 colonies at site TR-07.

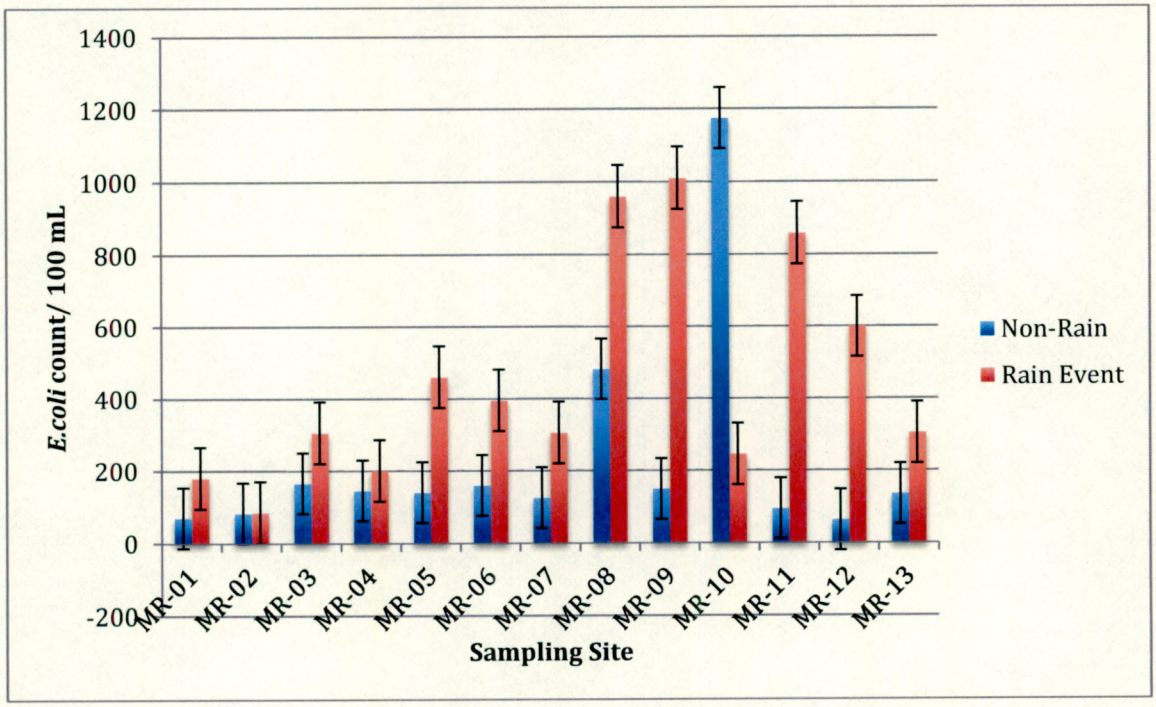


Figure 2. *E. coli* colony count per 100 mL at sampling sites of Musconetcong River during non-rain event and rain event in summer of 2014

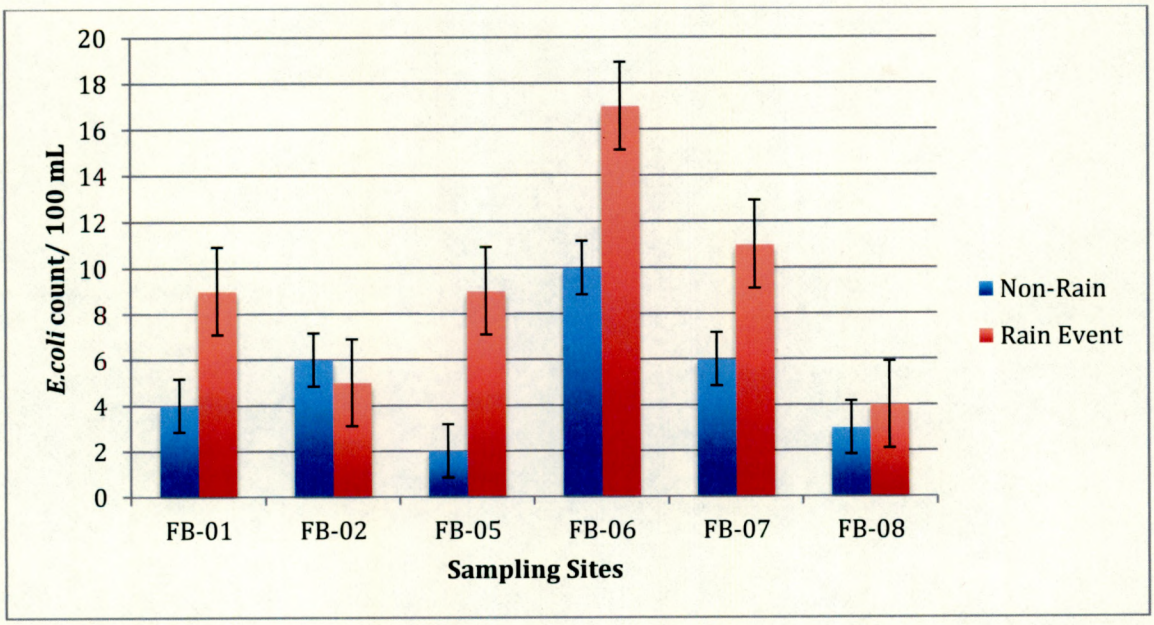


Figure 3. *E. coli* colony count per 100 mL at sampling sites of Flat Brook River during non-rain event and rain event in summer of 2014

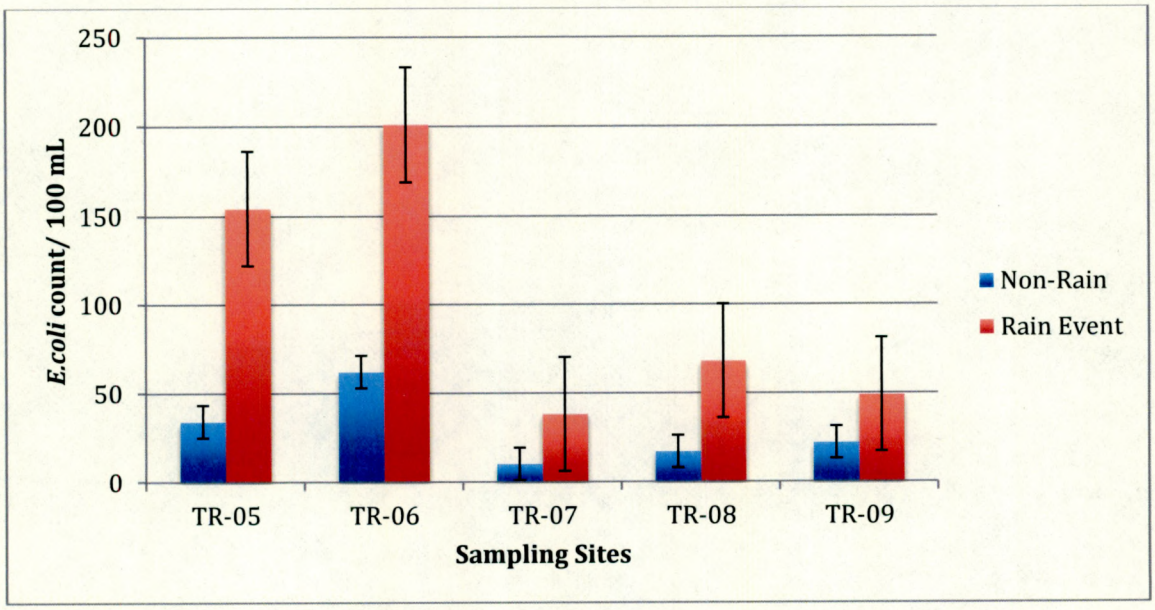


Figure 4. *E. coli* colony count per 100 mL at sampling sites of Third River during non-rain event and rain event in summer of 2014

Land Use - Land Cover

The Musconetcong River area was selected to represent agricultural land usage, the Third River was chosen to represent urban land usage and finally the Flat Brook River was selected to represent a forested/wild land usage (Figure 5, 6 and 7). To determine the statistical significance of the *E. coli* colony counts, an analysis of variance was run comparing the land area cover, agricultural, forested and urban, to the type of sampling event non-rain or rain event. For the non-rain event sampling, the colony counts in the agricultural area was significantly different from the colony counts in the urban and forested/wild land area ($p=0.0346$). For the rain event sampling, the colony counts in the urban area was found to be significantly different from the agricultural and the forested area types ($p=0.0005$ and $p=0.0283$ respectively). The species that were seen in agricultural area in both the non-rain event sampling and rain event were Canadian goose, cow, deer, horse and human (Table 3). Dog contamination was seen in all but two non-rain sampling sites and not present in the rain event samples (Table 3). For the urban area non-rain samples, the species contamination that were found to be present were human, dog, Canadian goose and horse (Table 4). For the urban area rain event samples, the species present were human, Canadian goose, dog and deer (Table 4). Dog was not present in rain event samples from site TR-06 and TR-08 and Canadian goose was not present in TR-07 during the rain event sampling. Finally, the forested area during non-rain sampling had contamination from deer, Canadian goose, horse and human present, however, human was absent in FB-02 and FB-06 (Table 5). For the rain event samples the species seen to be present were deer, Canadian goose, horse and human, human in only FB-01, FB-06 and FB-08 (Table 5).

The agricultural land use and land cover were found to contribute to the contamination seen in those sites. The farmlands that are seen throughout the area have horses on the land, which provides an explanation for why that contamination was found in the water. Canadian geese were found throughout almost all the sites for every area, this was also evident by the vast overpopulation seen throughout New Jersey. According to the New Jersey Department of Environmental Protection (NJDEP) in 2010 the resident population of Canadian geese in New Jersey consisted of 76,190 birds (NJDEP, 2011). Similar trends were found in deer populations in NJ. The deer is an edge species; it thrives in habitats that are broken up into parts much like an agricultural area or urban area such as our study sites. The NJDEP has stated that the deer populations in New Jersey vary geographically but if they were to be evenly distributed throughout the State's approximate 7,417 square miles, the 2010 population would have been an average of 15 deer per square mile (NJDEP, 2011). The dog contamination that was found at some of the agricultural sites might be due to domesticated dog-fecal matter in the area however; the distance covered makes that unlikely. Another explanation could be wild dogs, either coyote or wolves. New Jersey has seen an increase in wild dog populations in Warren, Sussex and Hunterdon counties, all which are surrounding areas of the Musconetcong River, have reported sightings or evidence of coyotes ranging from 31 to 111+ times (NJDEP, 2014). The human contamination is not surprising because the residents of the area have issued concerns over old septic systems that they fear may be leaching fecal matter into their waters and soils.

For the contamination at the urban area sampling sites, sources for Canadian goose and deer contamination were similar to the agricultural area. However, horses are

not found in the sampling areas, so where is that contamination coming from? When the results of horse presence first were observed there was immediate concern over the specificity of the primers used. However, after testing the primers against the other positive controls with no cross contamination false positives, a map of the river was analyzed. The source of horse contamination was found to be nearer the origin of the Third River. In the town of Woodland Park in Passaic County there are two parks, Rifle Camp Park and Garret Mountain Reservation both of which have water bodies, the Great Notch Reservoir and Barbour Pond, which flow into the Third River. Garret Mountain is home to the Garret Mountain Equestrian Center where they board horses for lessons and trail rides throughout 550 acres of the park. This could possibly be the source of horse contamination in the urban river. The dog contamination seen in the urban area is more likely due to domesticated dog feces not being tended to in the parks and streets than from wild dogs. Old sewage lines near the river that were suspected to be cracked and leaching contaminated sewer water into the river might be the source of the human contamination that was observed.

The Flat Brook River sees many of the same contamination results as the other areas, Canadian goose, deer, horse and human. Canadian goose and deer, once again, can be attributed to the wide range and overpopulation of those species. The presence of horse and human contaminations are the two that raise eyebrows at first glance. The horse contamination seems strange because much of the river flows through High Point State Park and Stokes State Forest, which are government managed areas. However, throughout Sussex County, where the river resides there are 19 different equestrian centers. This could have an impact on the river much like was seen with the Third River

and Garret Mountain Equestrian Center. Human contamination in the forested area is intriguing because of land use around the river. Much of the land is untouched forest, however, there are places along the river and its tributaries where trails and campsites are found. Hikers and campgrounds could be a source of possible contamination into the waters. Also, the few villages and towns found in the area are supported by aged septic tank systems that may be leaching contaminants.

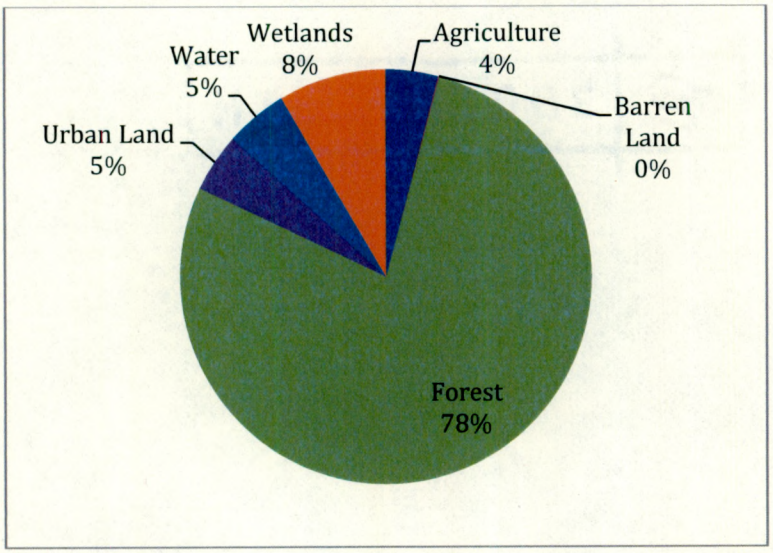


Figure 5. Land cover for the Middle Delaware Sub-basin

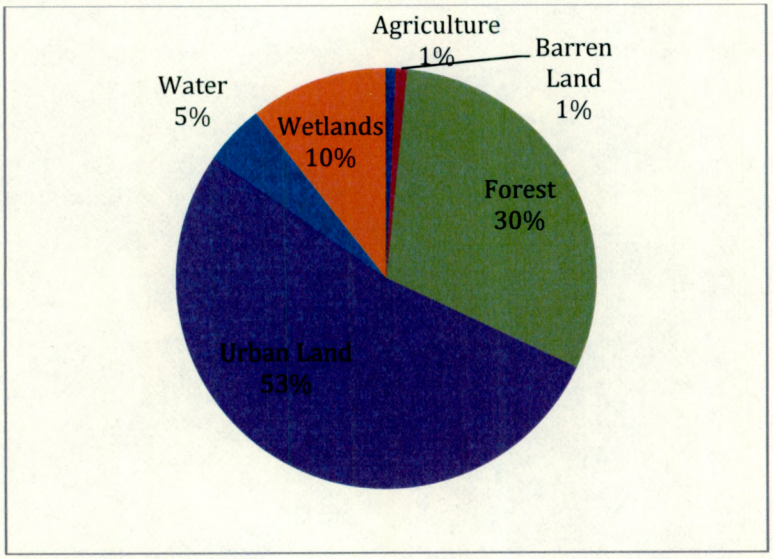


Figure 6. Land cover for the Hackensack-Passaic Sub-basin

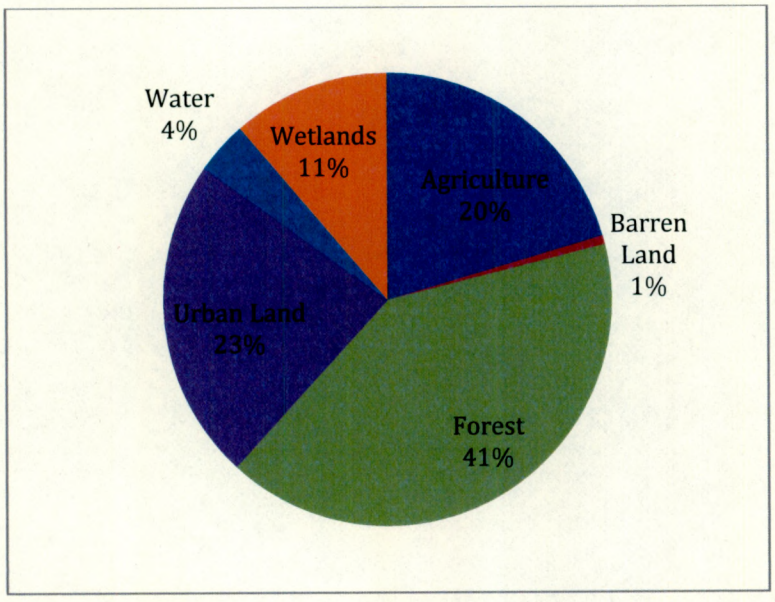


Figure 7. Land cover for the Musconetcong Sub-basin

Table 3. Contamination presence for each sampling site via gel electrophoresis for Musconetcong River

	MR-01		MR-02		MR-03		MR-04		MR-05		MR-06		MR-07		MR-08		MR-09		MR-10		MR-11		MR-12		MR-13		Percent Present
	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	
Canadian Goose	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	100%
Cow	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	7.69%
Deer	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	100%
Dog	+	-	+	-	+	-	+	-	-	-	+	-	-	-	+	-	+	-	+	-	+	-	+	-	+	-	42.31%
Horse	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	100%
Human	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	100%
Pig	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0%

Table 4. Contamination presence for each sampling site via gel electrophoresis for the Third River

	TR-05		TR-06		TR-07		TR-08		TR-09		Percent Present
	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	
Canadian Goose	+	+	+	+	+	-	+	+	+	+	90%
Cow	-	-	-	-	-	-	-	-	-	-	0%
Deer	-	+	-	+	-	+	-	+	-	+	50%
Dog	-	+	+	-	+	+	+	+	+	-	70%
Horse	+	-	+	-	+	-	+	-	+	-	50%
Human	+	+	+	+	+	+	+	+	+	+	100%
Pig	-	-	-	-	-	-	-	-	-	-	0%

Table 5. Contamination presence for each sampling site via gel electrophoresis for Flat Brook River

	FB-01		FB-02		FB-05		FB-06		FB-07		FB-08		Percent Present
	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	Non-Rain	Rain Event	
Canadian Goose	+	+	+	+	+	+	+	+	+	+	+	+	100%
Cow	-	-	-	-	-	-	-	-	-	-	-	-	0%
Deer	+	+	+	+	+	+	+	+	+	+	+	+	100%
Dog	-	-	-	-	-	-	-	-	-	-	-	-	0%
Horse	+	+	+	+	+	+	+	+	+	+	+	+	100%
Human	+	+	-	-	+	-	-	+	+	-	+	+	58.33%
Pig	-	-	-	-	-	-	-	-	-	-	-	-	0%

qPCR Copy Number Estimates of Contamination

Using a 1:10 serial dilution ranging from 1:1 to 1 x 10^{-4} a standard curve was produced for each species using their respective positive control DNA material. From this standard curve a slope equation is generated which is then in turn used to extrapolate quantities for each of the samples determined to be positive for contamination. From the quantity number, using the aforementioned copy number equation, copy number can be calculated for each of samples (Figures 8-20). The results demonstrated that the most dominant contamination source is the horse, reaching a high of $2.12E^{9}$ copy numbers, followed closely by deer and Canadian goose who were the next dominant in contaminant amounts in the samples. There was detection of human and dog contaminants in the range of $3.11E^{1}$ – $1.41E^{3}$ copies in the samples; however, these contaminants were detected at far lower numbers than those seen in the horse, deer and goose. These results suggest that the horse, being loose on pastures or in stables, when their feces is left unattended to and allowed to be introduced into the soil. Also, the area farms commonly apply horse manure onto the fields as fertilizer. Eventually, the bacteria from the manure found its way into the river. Human contaminants were also detected in the water. The older sewage infrastructure and septic systems in this area may be faulty and leaching contaminants into the river.

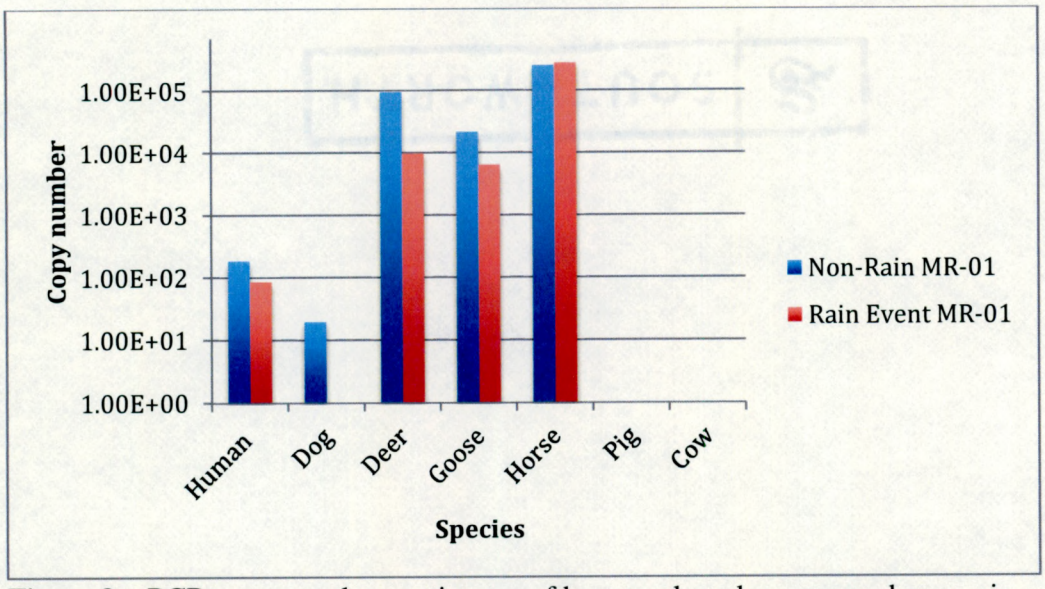


Figure 8. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-01

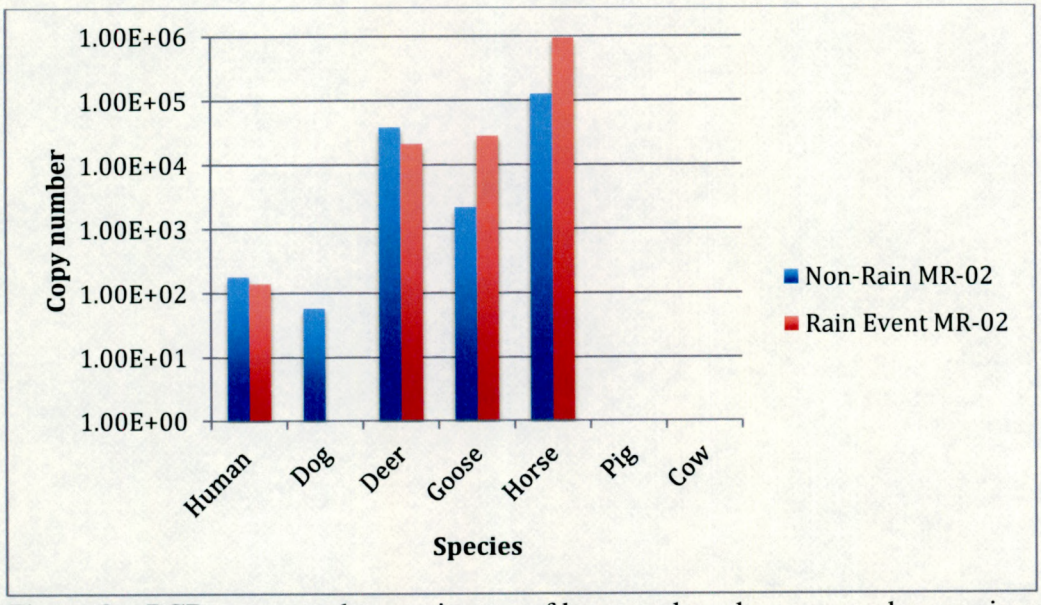


Figure 9. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-02

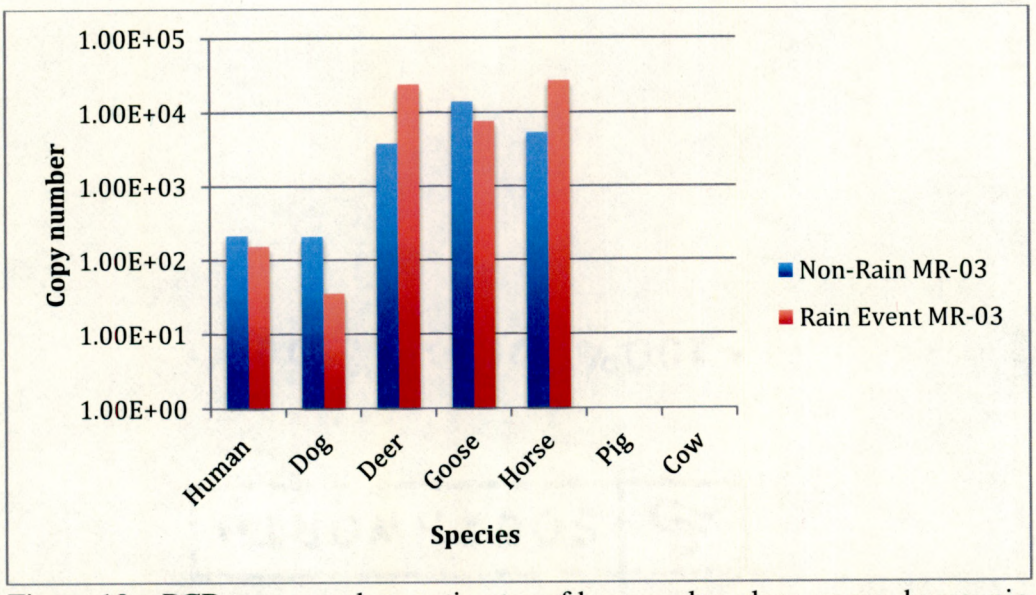


Figure 10. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-03

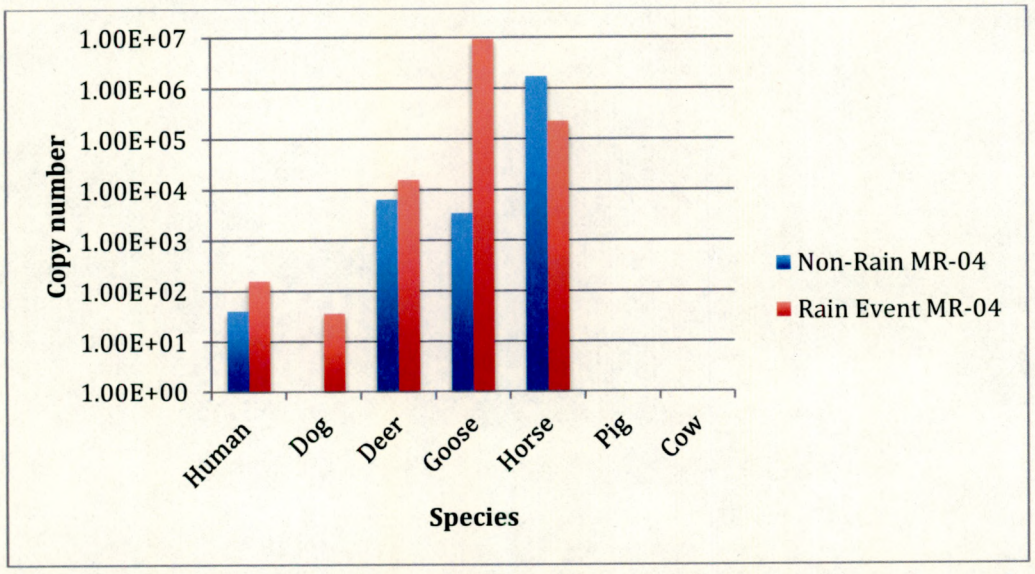


Figure 11. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-04

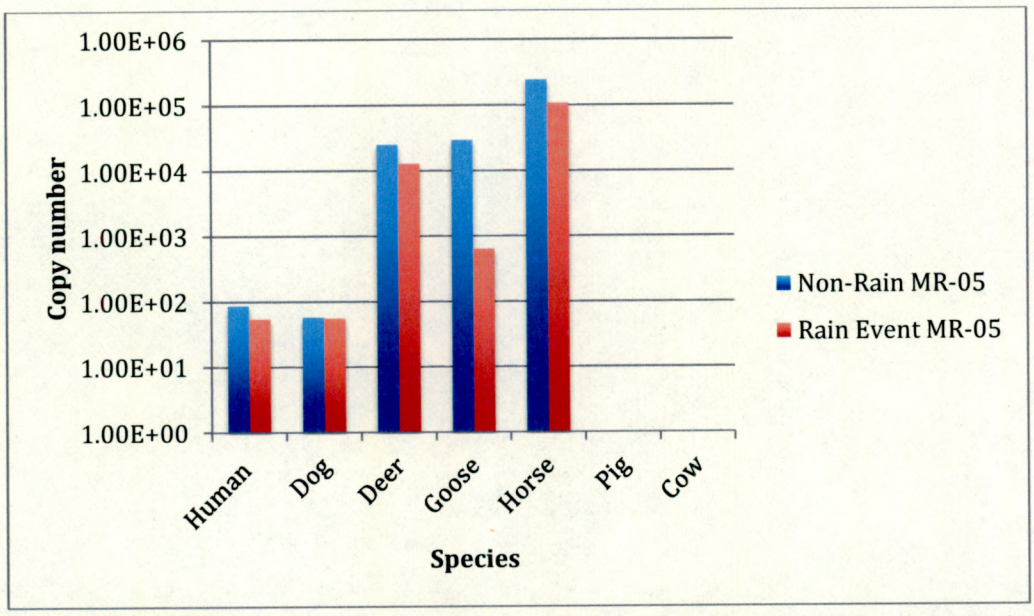


Figure 12. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-05

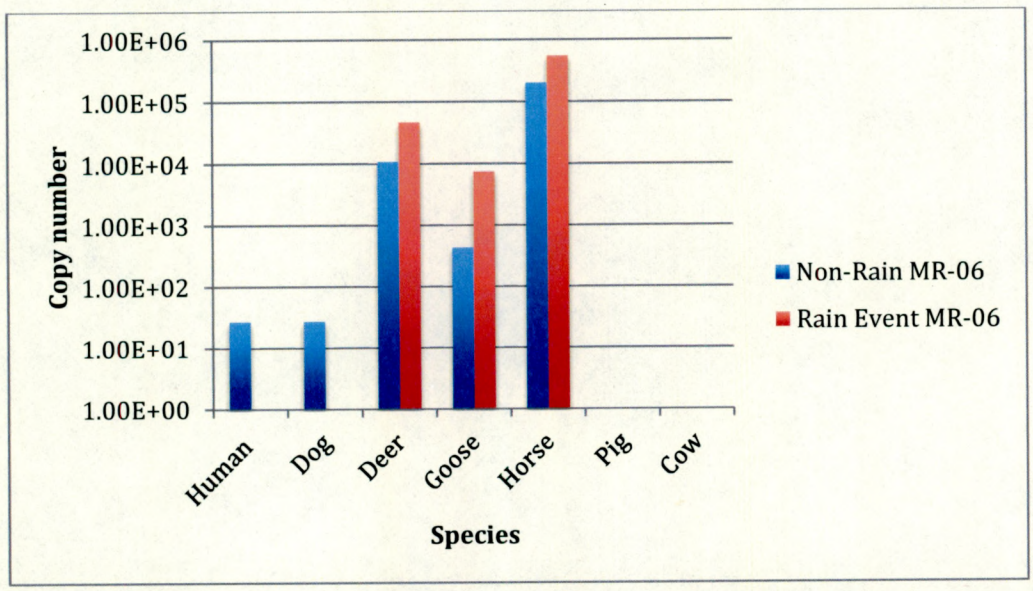


Figure 13. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-06

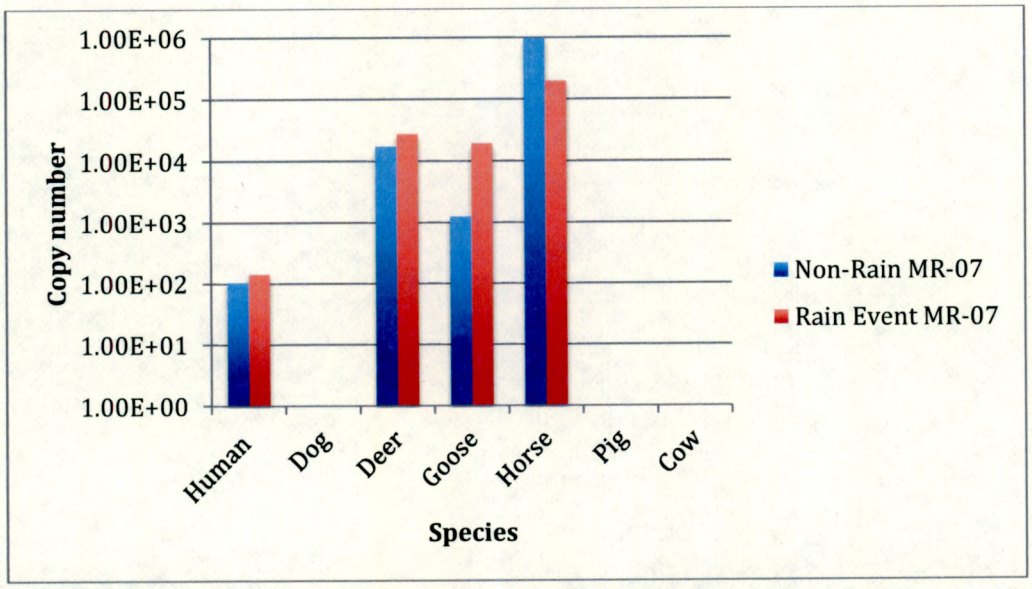


Figure 14. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-07

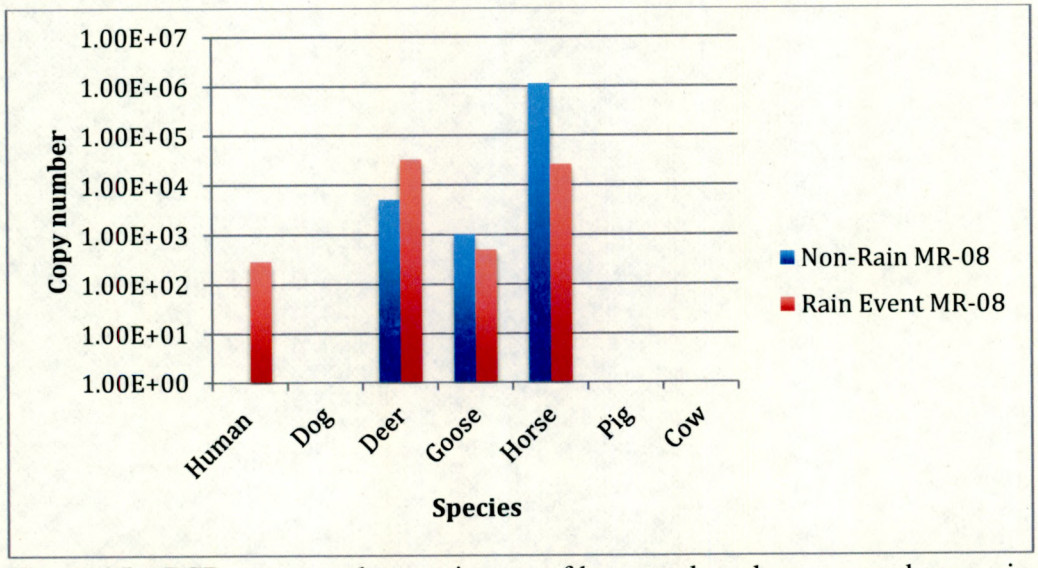


Figure 15. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-08

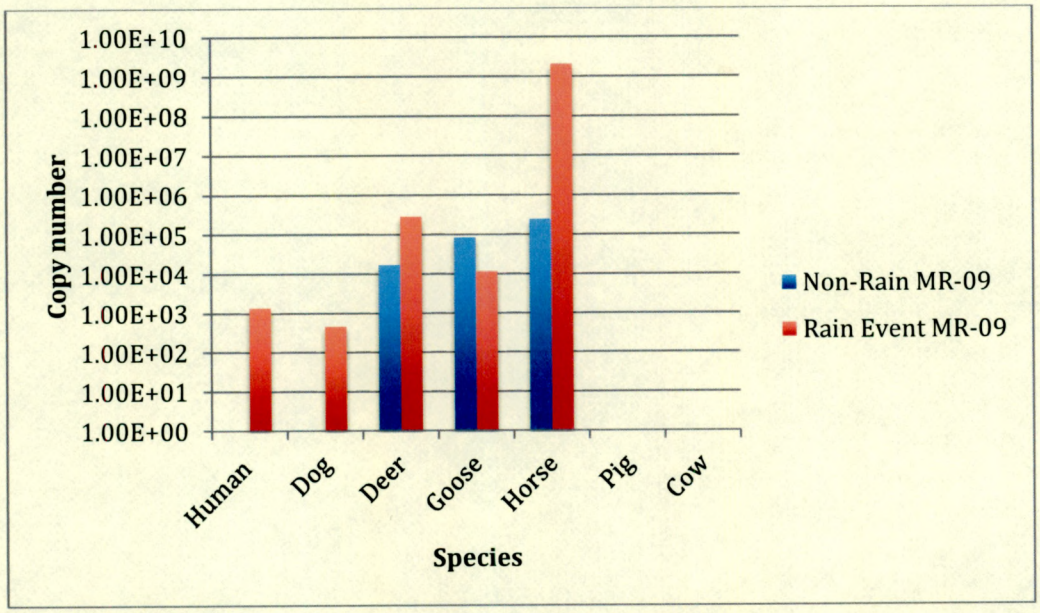


Figure 16. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-09

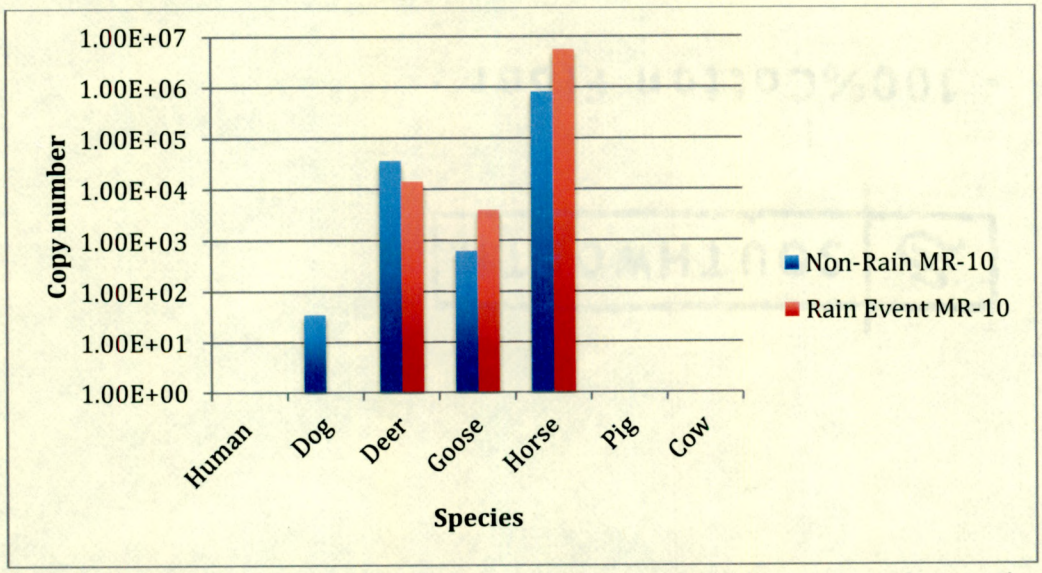


Figure 17. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-10

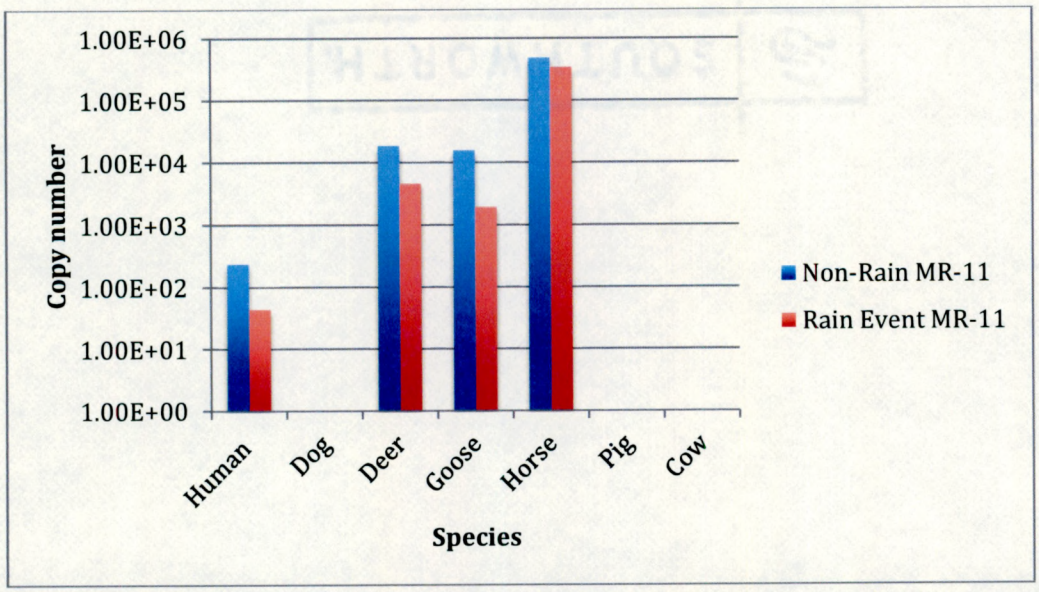


Figure 18. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-11

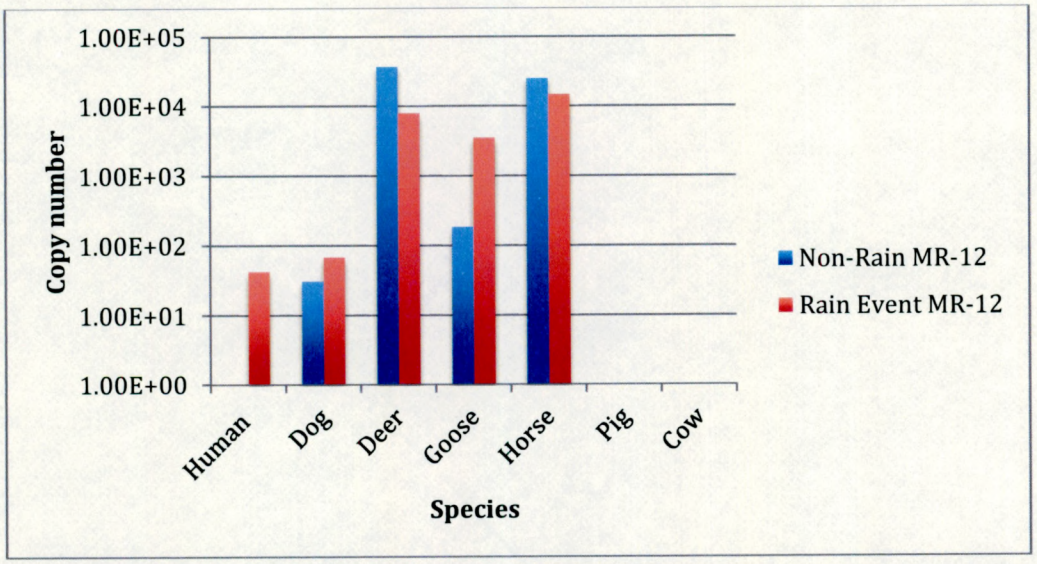


Figure 19. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-12

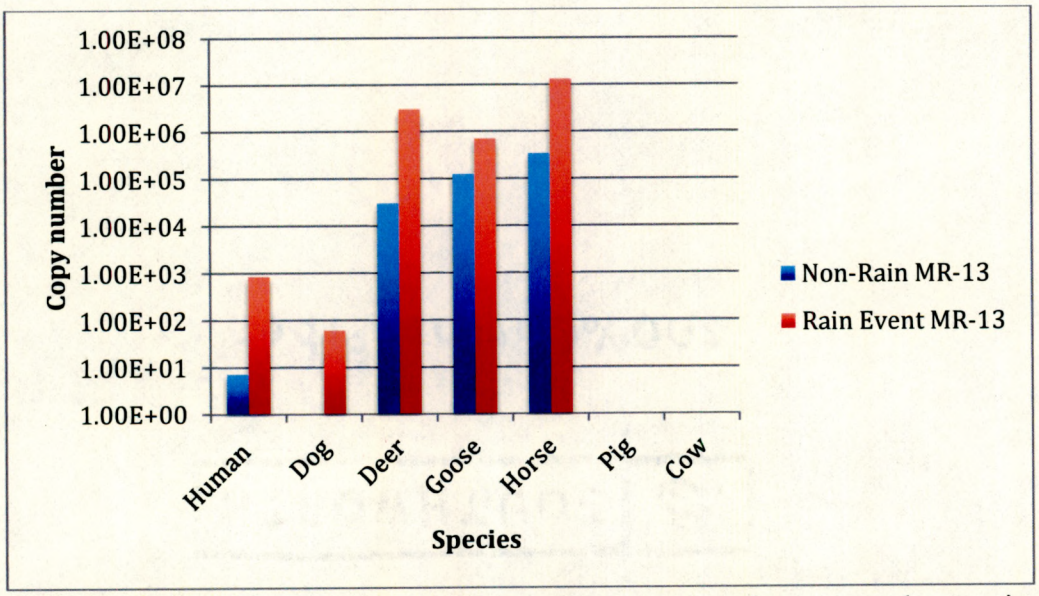


Figure 20. qPCR copy numbers estimates of human, dog, deer, goose, horse, pig and cow at site MR-13

Conclusion

The health of the water systems in New Jersey is a growing concern among many. The use of microbial monitoring techniques can help to determine the quality of the waters. Culturing indicator bacteria on nutrient agar plates to determine *E. coli* presence and therefore the possible presence of other potential pathogenic bacteria is a good way to initially determine the health of a water body. However, once an *E. coli*'s presence is found, microbial source tracking methods can be utilized to determine the ultimate source of the contaminants.

The PCR based-assay was able to provide a microbial map of the different sources of contaminants from the three different study sites characterized by various dominant land use and land cover types. Comparing results from the PCR based-assay and the real-time PCR based-assay it is clear that the qPCR based-assay is a much more sensitive test that will detect lower quantities of contaminants in the waters. Using data gathered from the qPCR tests copy numbers, which represent the approximate number of bacterial cells in 100 mL of water at the time the samples were taken, the most abundant source of contaminant can be identified. After comparing the two tests for sensitivity and efficiency, the results indicated that the qPCR based-assay is the better of the two tests.

References

1) Ashbolt, Nicholas J., Willie O.K. Grabow, and Mario Snozzi. "Indicators of Microbial Water Quality." *Water Quality: Guidelines, Standards and Health*. London, UK: IWA, 2001. Print.

2) "Assessed Waters of United States." *U.S. Environmental Protection Agency*. U.S. Environmental Protection Agency, 12 Dec. 2014. Web. 12 Dec. 2014.

3) Bernhard, A. E., and K. G. Field. 2000. "Identification of Nonpoint Sources of Fecal Pollution in Coastal Waters by Using Host-specific 16S Ribosomal DNA Genetic Markers from Fecal Anaerobes." Appl. Environ. Microbiology 66:1587-1594.

4) Blankenship, Karl. "DNA Library Would Give Investigators Inside Poop on Pollution Sources." *Bay Journal*. Chesapeake Media Service, 1 Sept. 1996. Web. 07 Nov. 2014.

5) Davis-Richardson AG, Ardissone AN, Dias R, Simell V, Leonard MT, Kemppainen KM, Drew JC, Schatz D, Atkinson MA, Kolaczkowski B, Ilonen J, Knip M, Toppari J, Nurminen N, Hyöty H, Veijola R, Simell T, Mykkänen J, Simell O and Triplett EW (2014) *Bacteroides dorei* dominates gut microbiome prior to autoimmunity in Finnish children at high risk for type 1 diabetes. Front. Microbiol. 5:678. doi: 10.3389/fmicb.2014.00678

6) Dufour, A.P. Escherichia coli: The Fecal Coliform, Bacterial Indicators/ Health Hazards Associated With Water, ASTM STP 635, A.W. Hoadley and B.J. Dutka, Eds., American Society for Testing and Materials, 1977, pp. 48-58. Print.

7) Ebentier, Darcy L., Kaitlyn T. Hanley, Yiping Cao, Brian D. Badgley, Alexandria B. Boehm, Jared S. Ervin, Kelly D. Goodwin, Michèle Gourmelon, John F. Griffith, Patricia A. Holden, Catherine A. Kelty, Solen Lozach, Charles Mcgee, Lindsay A. Peed, Meredith Raith, Hodon Ryu, Michael J. Sadowsky, Elizabeth A. Scott, Jorge Santo Domingo, Alexander Schriewer, Christopher D. Sinigalliano, Orin C. Shanks, Laurie C. Van De Werfhorst, Dan Wang, Stefan Wuertz, and Jennifer A. Jay. "Evaluation of the Repeatability and Reproducibility of a Suite of QPCR-based Microbial Source Tracking Methods." *Water Research* 47.18 (2013): 6839-848. Print.

8) Edge, Thomas, John Griffith, Joel Hansel, Valerie Harwood, Michael Jenkins, Alice Layton, Marirosa Molina, Cindy Nakatsu, Robin Oshiro, Michael Sadowsky, Jorge S. Domingo, Orin Shanks, Gerald Stelma, Jill Stewart, Donaldd Stoeckel, Bruce Wiggins, and Jayson Wilbur. "Microbial Source Tracking Guide Document." Ed. Jorge S. Domingo. (June 2005): n. pag. U. S. Environmental Protection Agency Office of Research and Development. Web. 17 May 2015.

9) Edwards, D. R., B. T. Larson, and T. T. Lim. "Runoff Nutrient And Fecal Coliform Content From Cattle Manure Application To Fescue Plots." *Journal of the American*

Water Resources Association 36.4 (2000): 711-21. Print.

10) Gilbert, Clément, Anne Ropiquet, and Alexandre Hassanin. "Mitochondrial and Nuclear Phylogenies of Cervidae (Mammalia, Ruminantia): Systematics, Morphology, and Biogeography." *Molecular Phylogenetics and Evolution* 40.1 (2006): 101-17. Web.

11) Jeter, S. N., C. M. Mcdermott, P. A. Bower, J. L. Kinzelman, M. J. Bootsma, G. W. Goetz, and S. L. Mclellan. "*Bacteroidales* Diversity in Ring-Billed Gulls (Laurus Delawarensis) Residing at Lake Michigan Beaches." *Applied and Environmental Microbiology* 75.6 (2009): 1525-533. Web.

12) Kildare, Beverly J., Christian M. Leutenegger, Belinda S. Mcswain, Dustin G. Bambic, Veronica B. Rajal, and Stefan Wuertz. "16S RRNA-based-assays for Quantitative Detection of Universal, Human-, Cow-, and Dog-specific Fecal *Bacteroidales*: A Bayesian Approach." *Water Research* 41.16 (2007): 3701-715. Print.

13) Lamendella, R., K. C. Li, D. Oerther, and J. W. Santo Domingo. "Molecular Diversity of *Bacteroidales* in Fecal and Environmental Samples and Swine-Associated Subpopulations." *Applied and Environmental Microbiology* 79.3 (2013): 816-24. Web.

14) Leonard, Michael T., Austin G. Davis-Richardson, Alexandria N. Ardissone, Kaisa M. Kemppainen, Jennifer C. Drew, Jorma Ilonen, Mikael Knip, Olli Simell, Jorma Toppari, Riitta Veijola, Heikki Hyã¶Ty, and Eric W. Triplett. "The Methylome of the Gut Microbiome: Disparate Dam Methylation Patterns in Intestinal *Bacteroides Dorei*." Frontiers in Microbiology 5 (2014): n. pag. Web.

15) Mieszkin, S., J.-F. Yala, R. Joubrel, and M. Gourmelon. "Phylogenetic Analysis Of 16S RRNA Gene Sequences from Human and Animal Effluents and Assessment of Ruminant Faecal Pollution by Real-time PCR." Journal of Applied Microbiology 108.3 (2010): 974-84. Web.

16) Mieszkin, S., J.-P. Furet, G. Corthier, and M. Gourmelon. "Estimation of Pig Fecal Contamination in a River Catchment by Real-Time PCR Using Two Pig-Specific *Bacteroidales* 16S RRNA Genetic Markers." *Applied and Environmental Microbiology* 75.10 (2009): 3045-054. Web.

17) Murugan, Kasi, Perumal Prabhakaran, Saleh Al-Sohaibani, and Kuppusamy Sekar. "Identification of Source of Faecal Pollution of Tirumanimuttar River, Tamilnadu, India Using Microbial Source Tracking." *Environmental Monitoring and Assessment* 184.10 (2012): 6001-012. Print.

18) "New Jersey Water Quality Assessment Report." *U.S. Environmental Protection Agency*. U.S. Environmental Protection Agency, 12 Dec. 2014. Web. 12 Dec. 2014.

19) New Jersey Department of Environmental Protection. *Range Expansion of the Eastern Coyote in New Jersey. NJ Upland Wildlife and Furbearer Research Project*. New Jersey

Department of Environmental Protection, n.d. Web. 11 Aug. 2014.

20) Okabe, Satoshi, Noriko Okayama, Olga Savichtcheva, and Tsukasa Ito. "Quantification of Host-specific *Bacteroides–Prevotella* 16S RRNA Genetic Markers for Assessment of Fecal Pollution in Freshwater." Applied Microbiology and Biotechnology 74.4 (2007): 890-901. Web.

21) Peris-Bondia, Francesc, Amparo Latorre, Alejandro Artacho, Andrés Moya, and Giuseppe D'auria. "The Active Human Gut Microbiota Differs from the Total Microbiota." Ed. Markus M. Heimesaat. *PLoS ONE* 6.7 (2011): Web.

22) Prescott, S.c., C.A. Winslow, and M.H. McCrady. "Water Bacteriology." New York: John Wiley & Sons. 1947. P.26.

23) "Real-time PCR Handbook." *Real-time PCR Handbook Single-tube Assays* (2014): n. pag. Web. 2 Apr. 2015.

24) Schiff, Kenneth, and Patrick Kinney. "Tracking Sources of Bacterial Contamination in StorMRter Discharges to Mission Bay, California." *Water Environment Research* (2001): 534-542.

25) Simmons, G. M., D. F. Waye, S. Herbein, S. Myers, and E. Walker. 2000. "Estimating Nonpoint Fecal Coliform Sources in Northern Virginia's Four Mile Run Watershed." p. 248-267. *In* T. Younos and J. Poff (ed.), Abstracts, Virginia Water Research Symposium 2000, VWRRC Special Report SR-19-2000, Blacksburg.

26) Simpson, Joyce M., Jorge W Santo Domingo, and Donald J. Reasoner. "Assessment of Equine Fecal Contamination: The Search for Alternative Bacterial Source-tracking Targets." *FEMS Microbiology Ecology* 47.1 (2004): 65-75. Web.

27) Simpson, Joyce M., Jorge W. Santo Domingo, and Donald J. Reasoner. "Microbial Source Tracking: State of the Science." *Environmental Science & Technology* 36.24 (2002): 5279-288. Print.

28) Snyder, J.C., Snider, A.R., Senecal, A.J., Disantis, E.J., Mackaness, C.A., Vaughn, B.P., Huber, J.P., Sopher, M.R., Disantis, R.S., Tobelmann, P.E., Hecht, N.M., Barry, P.M., Show, M.D., Speering, L.H., Sanchez, J.R., Brenner, F.J. and Ray, D.B. "Comparison of the mitochondrial genomes of the Canada Goose and the White Fronted Goose". Web.

29) US Census Bureau. "2012 Population Density: New Jersey Counties." (n.d.): n. pag. New Jersey Department of Labor and Workforce Development, June 2013. Web. 27 May 2015. <http://lwd.dol.state.nj.us/labor/lpa/content/maps/Popdenmap.pdf>.

30) USEPA. *The Quality of Our Nation's Waters*. Washington, D.C.: USEPA, 2000. Print.

31) USEPA. *Protocol for developing pathogen TMDLs*. U.S. Environmental Protection Agency, Office of Water: Washington, DC, 2001. Print.

32) USEPA. "What Is a TMDL?" *United States Environmental Protection Agency*. N.p., 11 Sept. 2013. Web. 07 Nov. 2014.

33) *Wildlife Populations: Canada Goose*. 1st ed. NJDEP, 2011. Web. 23 Apr. 2015.

34) *Wildlife Populations: White-tailed Deer*. 1st ed. NJDEP, 2011. Web. 23 Apr. 2015.

35) Willing, B., A. Vörös, S. Roos, C. Jones, A. Jansson, and J. E. Lindberg. "Changes in Faecal Bacteria Associated with Concentrate and Forage-only Diets Fed to Horses in Training." *Equine Veterinary Journal* 41.9 (2009): 908-14. Web.

36) Zhang, H. H., L. Chen, and G. S. Liu. "Phylogenetic Analysis of 16S RRNA Gene Sequences Reveals Distal Gut Bacterial Diversity in Dhole (Cuon Alpinus)." (n.d.): Web.

37) Zhang, Honghai, and Lei Chen. "Phylogenetic Analysis of 16S RRNA Gene Sequences Reveals Distal Gut Bacterial Diversity in Wild Wolves (Canis Lupus)." *Molecular Biology Reports* 37.8 (2010): 4013-022. Web.

Appendix A

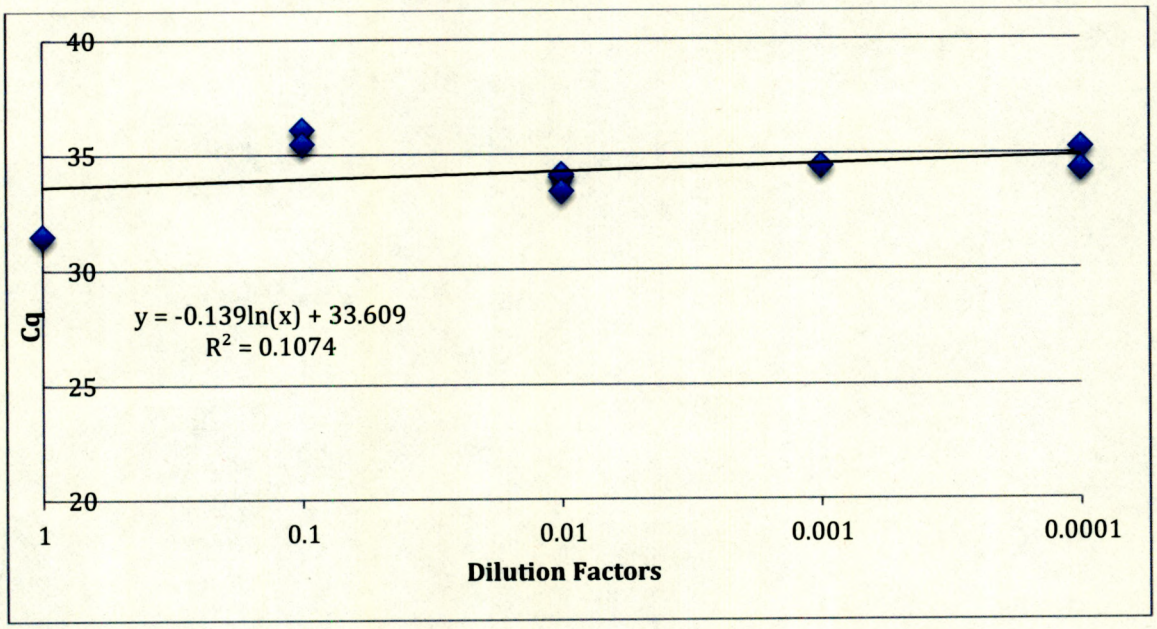


Figure 21. Cow Positive Control Standard Curve

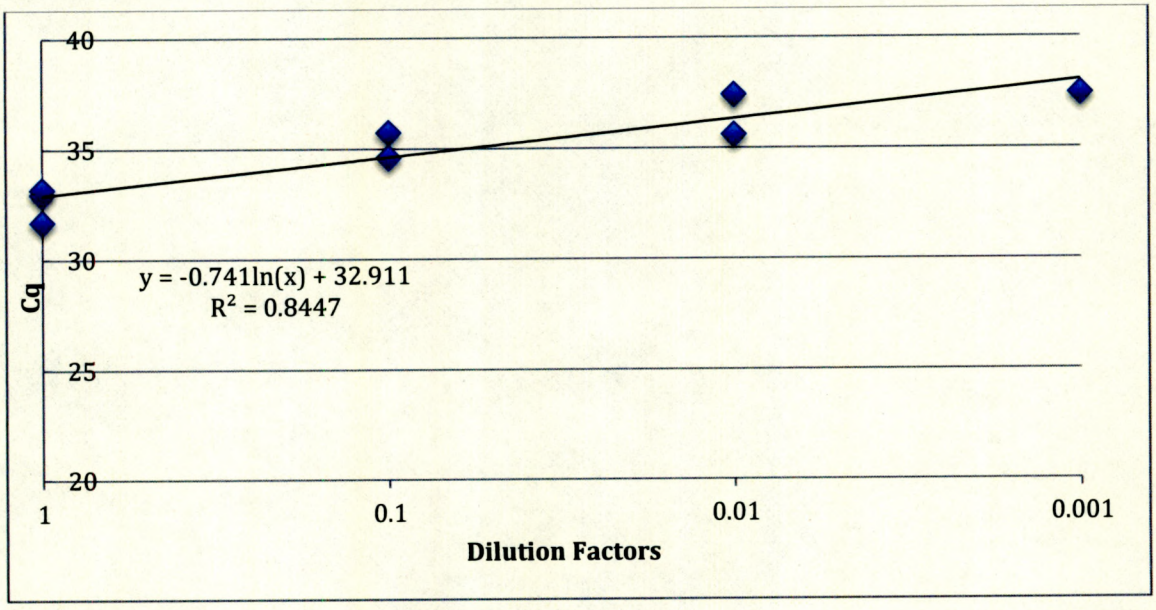


Figure 22. Deer Positive Control Standard Curve

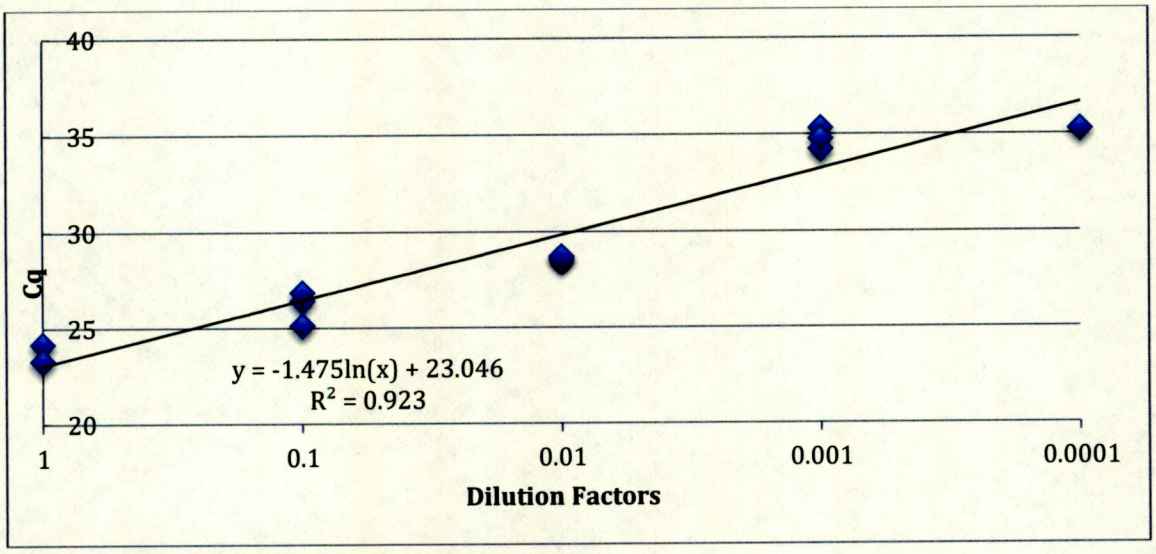


Figure 23. Dog Positive Control Standard Curve

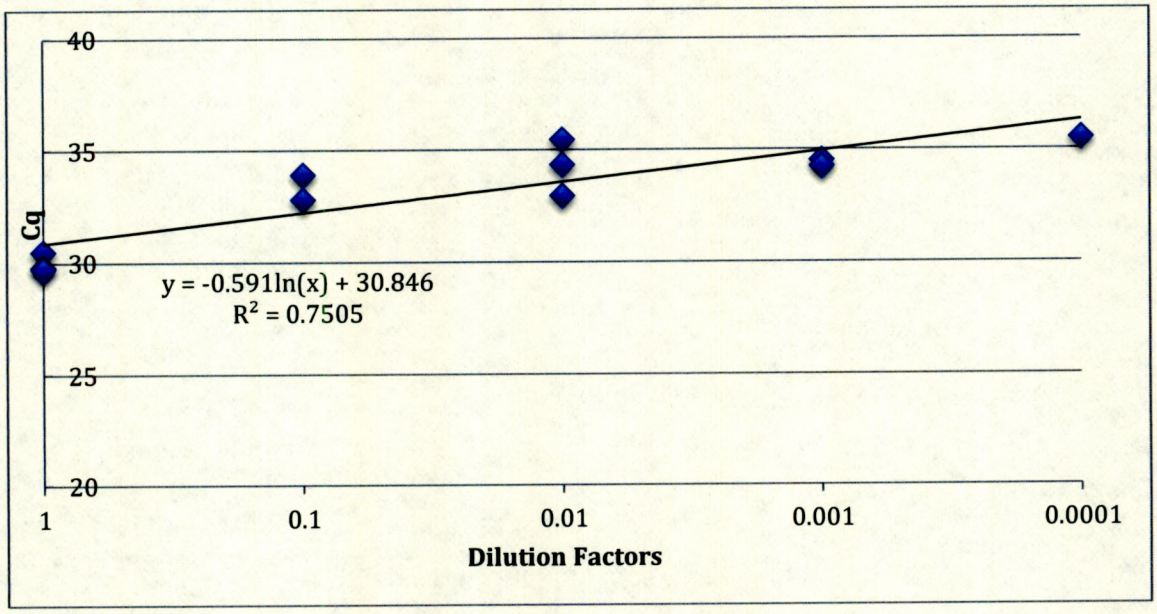


Figure 24. Goose Positive Control Standard Curve

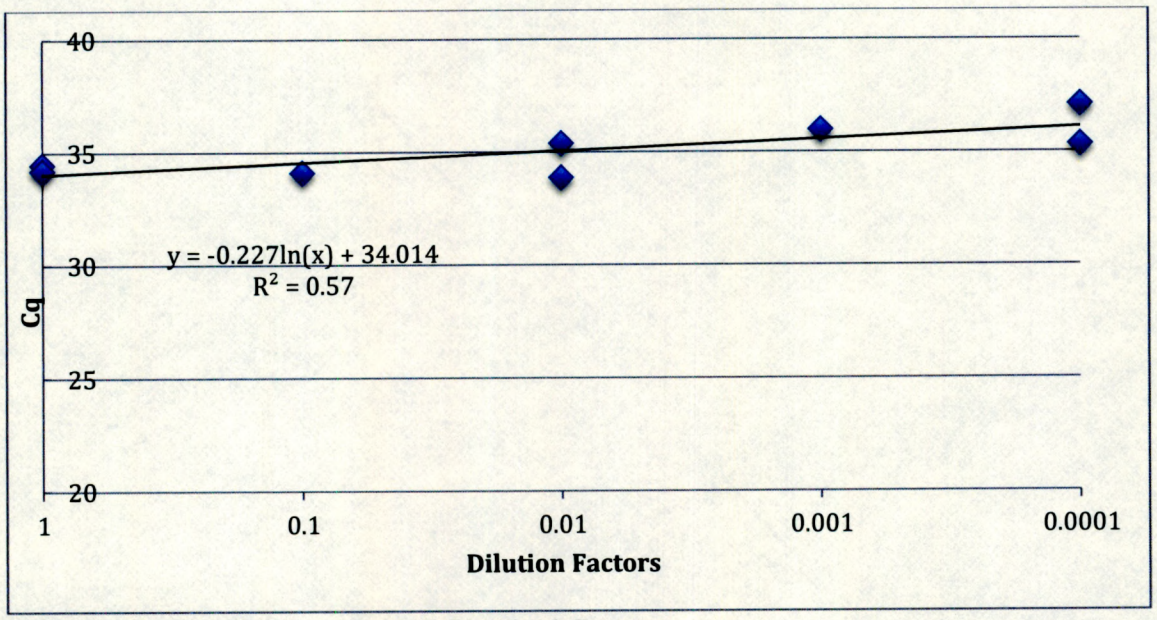


Figure 25. Horse Positive Control Standard Curve

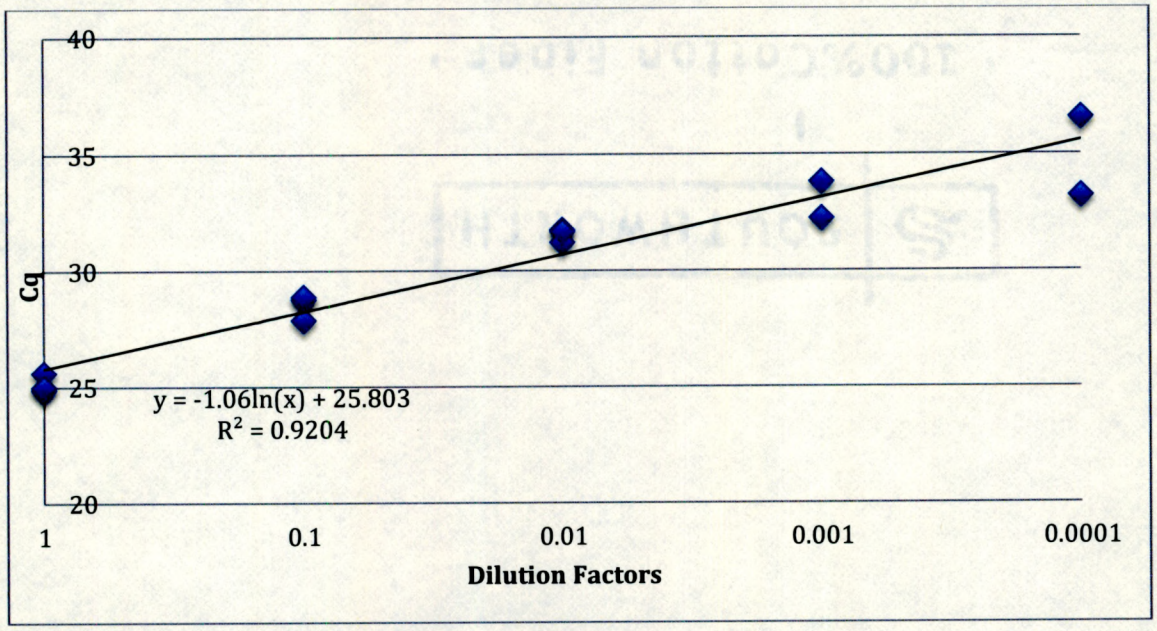


Figure 26. Human Positive Control Standard Curve

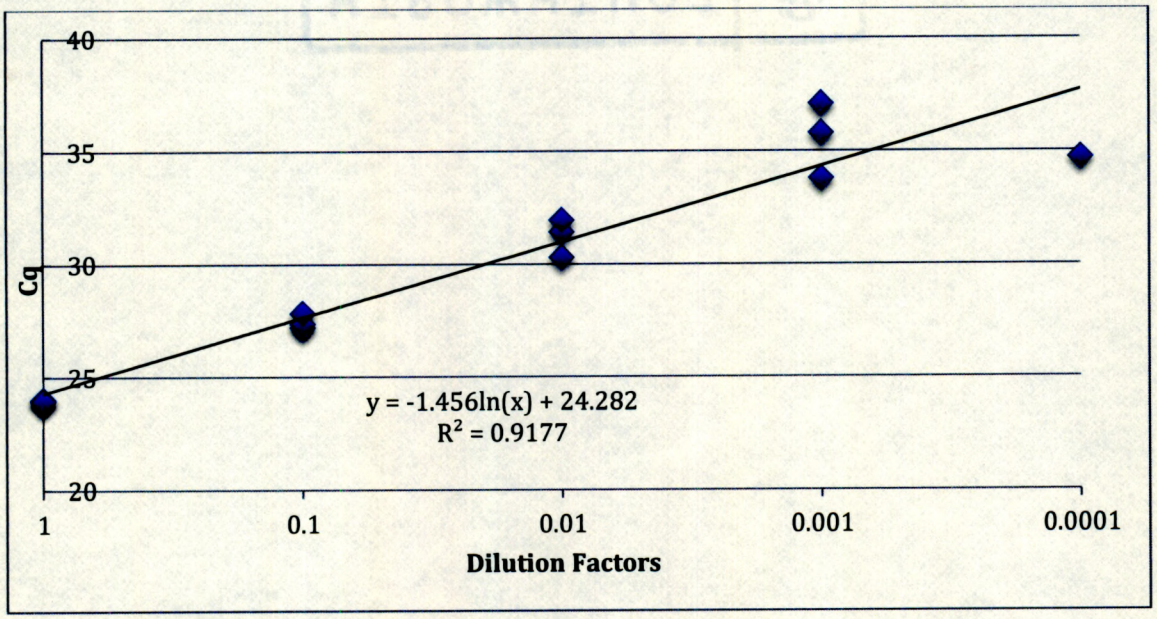


Figure 27. Pig Positive Control Standard Curve

Appendix B

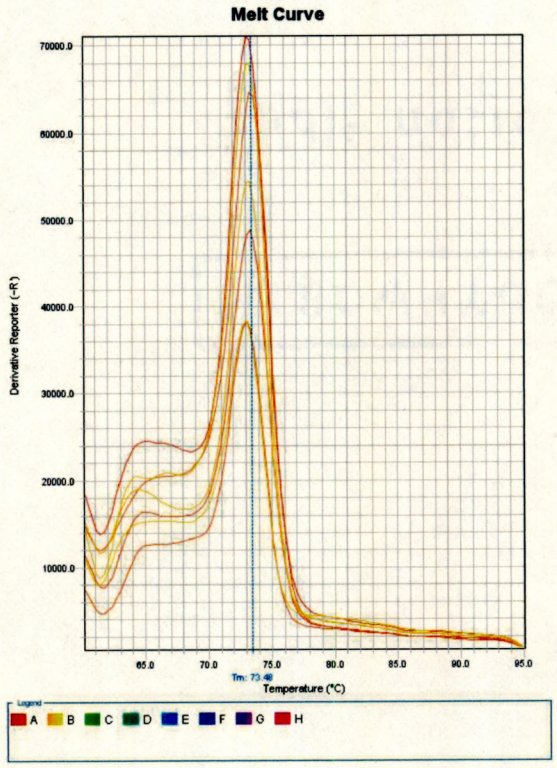


Figure 28. Cow Standard Melt Curve

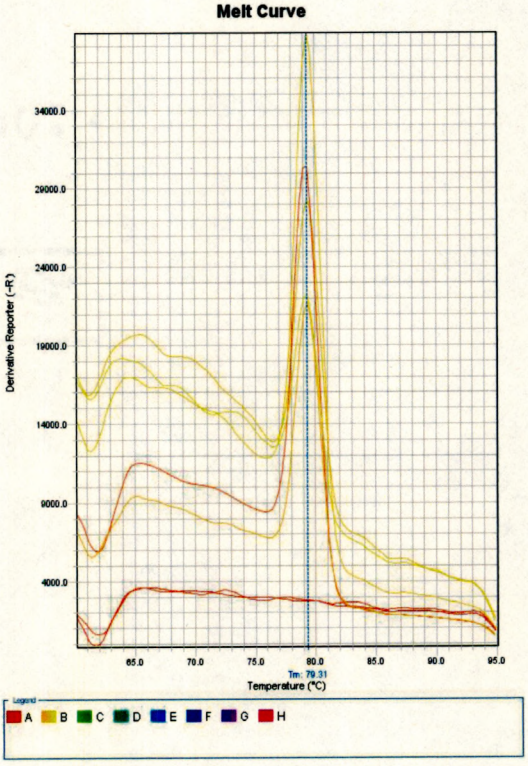


Figure 29. Deer Standard Melt Curve

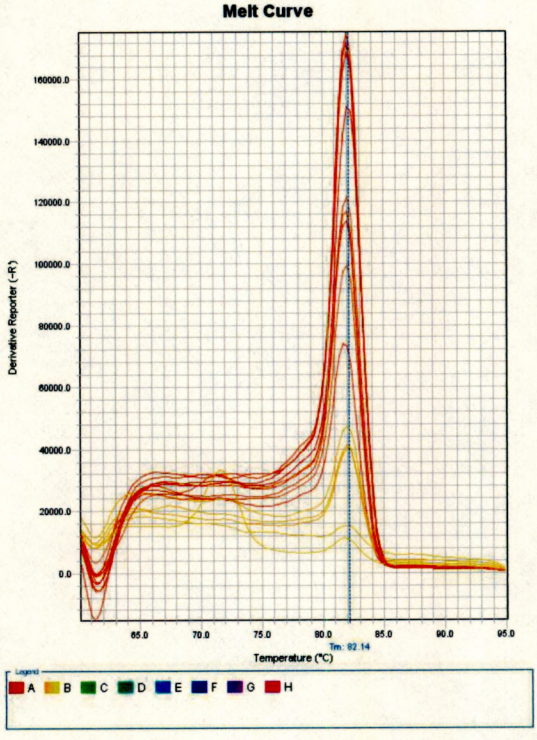


Figure 30. Dog Standard Melt Curve

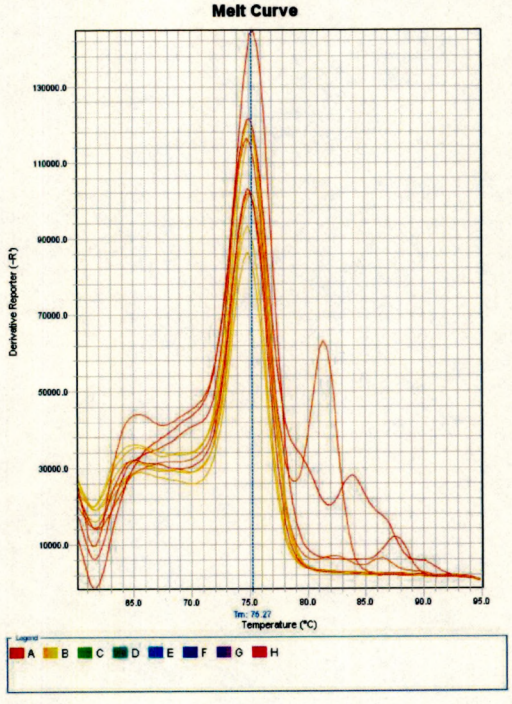


Figure 31. Goose Standard Melt Curve

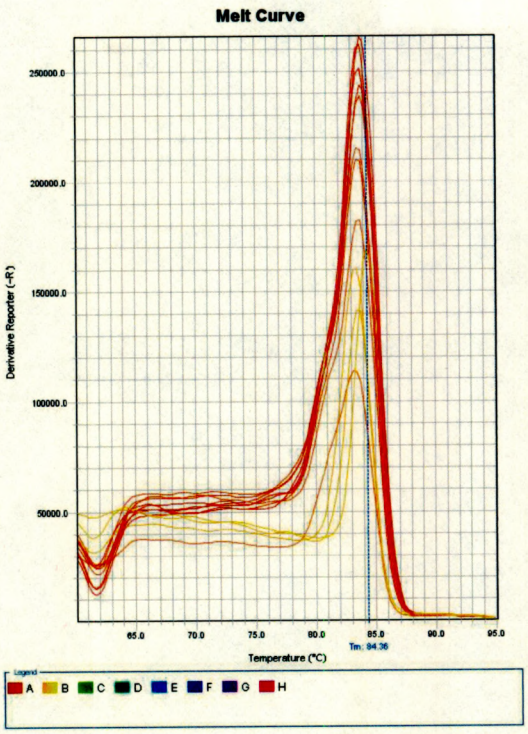


Figure 32. Pig Standard Melt Curve

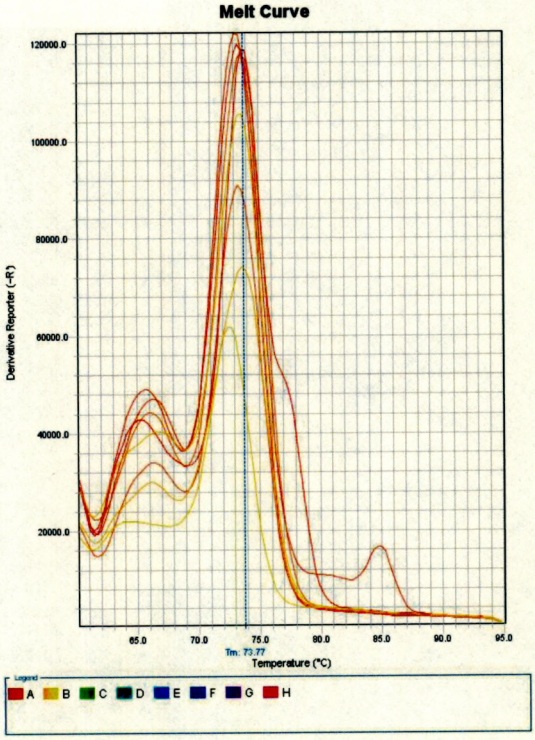


Figure 33. Horse Standard Melt Curve

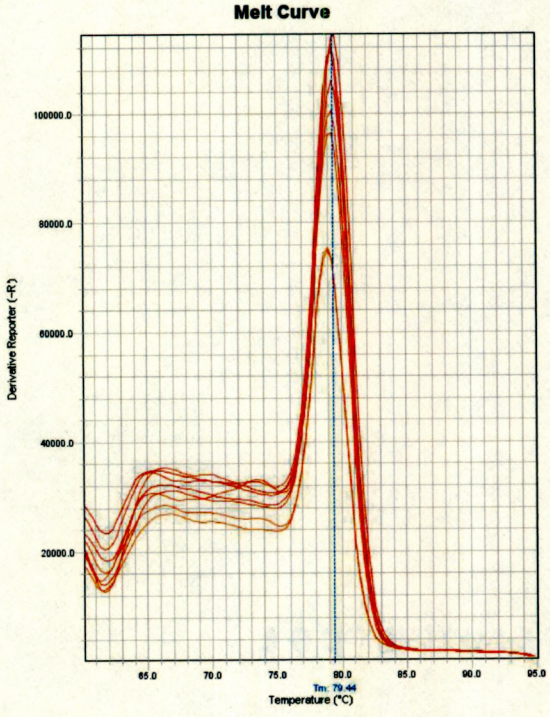


Figure 34. Human Standard Melt Curve