Don’t Get in That Water: Bacteriophages as Indicators of Viruses in Tanyard Creek.

Marissa Thongdy
Abstract

Don’t Get in That Water: Bacteriophages as Indicators of Viruses in Tanyard Creek.

By

Marissa Thongdy

January 2019

INTRODUCTION: Tanyard Creek is one of the urban creeks in metro Atlanta that make up the large urban sub-watersheds sending huge volumes of storm water draining into the Chattahoochee River. The creek is considered impaired, with large visible signs that warn do not play, swim or fish in the creek: As an urban creek and is subject to sewage overflows and runoff contamination. Urban runoff can carry contaminants, such as sewage runoff, animal waste, chemical pollutants, and pesticides to the creek, creating health risks for those who have access to it. To better understand what kinds of contaminants are in the creek, we can look for organisms such as bacteria and viruses. One such virus is bacteriophage MS2. If MS2 is present in this creek, it is an indication that it is possibly human and animal fecal pollution present. Therefore, it is essential to understand the pattern of these indicators as it relates to waterborne illnesses.

AIM: This research will determine the trends of MS2 and E.coli Tanyard Creek if they differ spatially from sampling site to sampling site and temporally. Additional goals include understanding the natural variability and the relationship with rainfall. Also, the relationship between MS2 and E. coli will be examined.

METHODS: Water samples were collected weekly at ten sites downstream from Tanyard Creek CSO located off Collier Road at Ardmore Park which is considered part of the Atlanta Beltline. All samples collected from the creek were then brought to the lab for analysis of microorganisms through membrane filtration and viral assay.

RESULTS: Both MS2 and E. coli are present in Tanyard Creek, at levels higher than U.S. Environmental Protection Agency standards. This data indicates temporal trends; during the August-September, there are higher counts of E.coli. When data were presented on a spatial level, it was discovered that the higher numbers of E. coli were present after the beaver dam that is in the considered to be the middle or the halfway point of Tanyard Creek Park.
DON’T GET IN THAT WATER: BACTERIOPHAGES AS INDICATORS OF VIRUSES IN TANYARD CREEK.

by

Marissa Thongdy

B.S., UNIVERSITY OF LOUISIANA AT MONROE

A Thesis Submitted to the Graduate Faculty of Georgia State University in Partial Fulfillment of the Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA
30303
DON’T GET IN THAT WATER: BACTERIOPHAGES AS INDICATORS OF VIRUSES IN TANYARD CREEK.

by

MARISSA THONGDY

Approved:

Dr. Lisa Casanova
Committee Chair

Dr. Matthew Magee
Committee Member

January 11, 2019
Date
Acknowledgments

I would first like to express my gratitude to my thesis committee chair Dr. Lisa Casanova for her guidance, and constructive criticism throughout the completion of this thesis. I am grateful to my committee member Dr. Mathew Magee for his essential comments and views on issues related to my research.

I would also like to appreciate the faculty and staff of the School of Public Health for their guidance and support throughout the program. Also, I am gratefully indebted to my family and friends for their unwavering support.
In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

Marissa Thongdy
Signature of Author
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................................................ iv  

LIST OF TABLES ........................................................................................................................................ vii  

INTRODUCTION ............................................................................................................................................. 1  
  1.1 Background ........................................................................................................................................ 1  
  1.2 Research Aims and Hypothesis ........................................................................................................ 3  

REVIEW OF THE LITERATURE ..................................................................................................................... 5  
  2.1 Water-borne Pathogens ..................................................................................................................... 5  
  2.2 Detection Methods .......................................................................................................................... 6  
  2.3 Epidemiological Relationships ...................................................................................................... 9  

METHODOLOGY .......................................................................................................................................... 11  
  3.1 Research Areas .................................................................................................................................. 11  
  3.2 Data Analysis ..................................................................................................................................... 14  

RESULTS ...................................................................................................................................................... 15  
  4.1 Results of E.coli .................................................................................................................................. 15  
  4.2 Results of MS2 Bacteriophage ........................................................................................................ 19  

DISCUSSION AND CONCLUSION .................................................................................................................. 24  
  5.1 Discussion of Research Questions ................................................................................................... 24  
  5.2 Study Limitations and Next Steps .................................................................................................. 26  
  5.3 Conclusion ......................................................................................................................................... 28
List of Graphs

Graph 1.1. E. coli Across 10 Sampling Sites by Date
Graph 1.2. E. coli Across 10 Sampling Sites by Date (Log10)
Graph 2.1. E. coli Across All Dates by Sites
Graph 2.2. E. coli Across All Dates by Sites (Log10)
Graph 3.1. MS2 Across 10 Sampling Sites by Date
Graph 3.2. MS2 Across 10 Sampling Sites by Date (Log10)
Graph 4.1. MS2 Across All Dates by Sites
Graph 4.2. MS2 Across All Dates by Sites (Log10)
Graph 5.1. Rainfall Averages
Graph 6.1. Sample Mean of E.coli by Sites
Graph 7.1. Sample Mean of E.coli by Dates
Graph 8.1. All Samples of E.coli Colonies
List of Figures

Figure 1.1. CSO Master Plan
Figure 2.1. Spatial Analysis Output
Figure 3.1. 2012 Recreational Water Quality Criteria (RWQC)
CHAPTER I – Introduction

1.1 Background

Tanyard Creek is one of several urban creeks in Atlanta that eventually drain into the Chattahoochee River. Large urban sub-watersheds send vast volumes of storm water downstream to the river. With large visible signs that warn do not play, swim or fish in creek, Tanyard is an urban creek and is subject to sewage overflows and runoff contamination. Identifying MS2 in this creek can give us an idea of whether there is human and animal fecal pollution present. In an urban creek like Tanyard, urban runoff is a source of pollution. Urban runoff occurs because of rain during wet weather; during dry weather it can be waste that flows from urban landscapes into storm drain systems that lead to the creek. Urban runoff carries contaminants, such as sewage runoff, animal waste, industrial pollutants, and pesticides to the creek creating health risks for those who have access to it.

Sewage runoff was a significant source of pollution in Atlanta creeks. In 1999, the City of Atlanta entered into a consent decree to improve its combined sewer system due to the constant sewer discharges being in violation of the federal Clean Water Act and the Georgia Water Quality Control Act (Hunter and Sukenik, 2007; US Environmental Protection Agency, 1999). An investment was required to make renovations, including separating the combined sewers into distinct sewer and storm water lines and the construction of off-line storage facilities. It is understood that with remediation of storage tunnels there can still have the possibility of generating discharges.

In July 2001, after a 3-year process of study and citizen input the EPA and state of Georgia's Environmental Product Declaration (EPD) approved the City of Atlanta’s plan to reduce water
quality violations from combined sewer overflows (CSOs). The City’s plan involved a combination of tunnels and separation of selected sewer areas. The City submitted a revised proposal to EPA and EPD that would increase the water quality benefits of proposed sections of the program and reduce the lengths of the proposed CSO tunnels. The storage and treatment system involve capturing and storing combined sewer overflows. The overflows are stored in large underground tunnels in bedrock similar to the rock of Stone Mountain. After the rainfall is over, the captured CSO volume is conveyed to a separate treatment system for removal of pollutants and ultra-violet disinfection before discharge to the waters.

In the US, urban water pollution due to effluents emanating from combined sewer facilities is considered a major source of water impairment, and a significant human health concern (Tibbetts, 2005; US Environmental Protection Agency, 2009). Under dry conditions, the mix of precipitation and sewage is channeled to a treatment plant before being discharged into water bodies. During heavy precipitation, storm and waste water exceeding a treatment plant’s processing capacity are discharged into local surface waters, a process known as a combined sewer overflow (CSO) (Tibbetts, 2005). These overflows have several different possible health impacts, including spreading waterborne and vectorborne disease. In Atlanta, GA, Culex quinquefasciatus is the main urban vector of West Nile virus (Vazquez-Prokopec et al., 2010). CSO-affected streams provide optimal habitat for Cx. quinquefasciatus (Calhoun et al., 2007; Chaves et al., 2011, 2009; Nguyen et al., 2012)(Chaves et al., 2011).

The United States Environmental Protection Agency (EPA) mandates the use of fecal indicator bacteria, including E. coli, as a way to detect fecal contamination in water and assess the quality of drinking and recreational waters in the U.S. Fecal indicator bacteria have been used to
determine if there is sewage contamination in water to help protect the public from waterborne pathogens, including bacteria and viruses that spread through human and animal feces. To determine if water is polluted by human or animal feces, we look for the presence of fecal indicators, which are bacteria and viruses that humans and animals carry in their intestines. Bacteriophages, a type of viruses that infect bacteria, are one example of a fecal indicator. Bacteriophages are considered the most abundant form of “life” on earth and can be found in all environments where bacteria grow, including in soil, water, and inside other larger organisms (e.g., humans) harboring host bacteria (e.g., E. coli) (Clokie et al., 2011; Dutilh et al., 2014; Díaz-Muñoz and Koskella, 2014). EPA has conducted many epidemiological studies in both marine and freshwaters to evaluate the relationship between fecal indicators and recreational swimming-associated illnesses in surface waters. The incidence of symptoms associated with gastrointestinal, eye, ear, and respiratory illnesses has been found to be higher in swimmers than in non-swimmers in ambient waters. These studies indicate that bacteriophages in water are related to the risk of illness in swimmers. Fecal indicators like bacteriophages can be used to look for pollution in many different kinds of water, including urban rivers and creeks. While creeks like Tanyard have improved since the consent decree and the upgrades to the city’s sewer systems, there are still sources of pollution. This research will look for patterns of fecal indicators such as bacteriophages in the creek to determine whether fecal pollution might be present in Tanyard creek.

1.2. Research Aims and Hypothesis

This research will use two fecal indicators, one bacteria (E. coli) and one virus (bacteriophage MS2). The overall goals will be to explain the trends of these indicators in Tanyard Creek, including if they differ spatially site by site and temporally over different weeks. We will also
look at the variability of organisms and whether rainfall is a contributing factor. Also, the relationship between MS2 and E. coli will be examined.

**Aim 1**: Compare MS2 bacteriophage levels present in all water sample of Tanyard Creek and how they change spatially (from site to site) and temporally (January 2018 - Present).

**Hypothesis 1**: MS2 levels will indicate that the urban creek is considered to be heavily polluted by urban runoff and sewer outflows.

**Aim 2**: Determine and compare the relationship between MS2 bacteriophage and E. coli levels across all sampling dates by sites.

**Hypothesis 2**: Concurrent discovery of E coli and MS2, compared to only MS2, will be associated noncompliance with United State Environmental Protection Agency.
CHAPTER II - Review of the Literature

2.1. Water-borne Pathogens

The increasing interest in controlling water-borne pathogens in water resources has been evident by a large number of recent publications that indicates the need for studies that synthesize knowledge from multiple fields covering comparative aspects of pathogen contamination, and how to unify them to present and address the problem as a whole. Indicator organisms are commonly used to assess the levels of pathogens in water resources; i.e., water-borne pathogen footprints of water resources. Monitoring the levels of indicator organisms (such as fecal coliforms, E. coli) is a common approach for quantifying the potential pathogen loads in ambient water bodies. For decades, public health officials/scientists have evaluated water quality by enumerating fecal coliforms and E. coli levels in rivers, lakes, estuaries, and coastal waters (Malakoff [2002]; Pandey et al. [2012a]; Pandey et al. [2012b]; Pandey and Soupir [2013]). The EPA defines acceptable recreational limits as those that will result in eight or fewer swimming-related gastrointestinal (GI) illnesses out of every 1,000 swimmers (U.S. EPA [1986]). The current U.S. EPA fresh water quality criteria for E. coli is a geometric mean not exceeding 126 CFU/100 ml, or no samples exceeding a single sample maximum of 235 CFU/100 ml (U.S. EPA [2001]). Criteria were developed based on the U.S. EPA measurements of total and Highly Credible Gastrointestinal Illnesses (HCGI), which correlated with E. coli densities (r = 0.804) in fresh recreational waters (Dufour [1984]). Multiple studies have identified trends between indicator organisms in water and GI illness in humans, including vomiting, diarrhea, and fever (Cabelli [1983]; Wade et al. [2006]).

Coliphages can be inactivated, or made noninfective by various environmental factors, including temperature (Feng et al., 2003), pH (Feng et al., 2003), salinity (Sinton et al., 2002), sunlight
(Sinton et al., 1999), and ultraviolet (UV) light (Sang et al., 2007). Viral inactivation occurs when viral components (nucleic acids, proteins, lipids) are disintegrated. Therefore, characteristics that influence survival include coliphage morphology, including size and surface properties (Jończyk et al., 2011). Coliphages are nonenveloped and are resistant to environmental degradation and chemical inactivation similar to other enteric nonenveloped viruses (Havelaar, 1987; Havelaar et al., 1990; Yahya and Yanko, 1992; Nasser et al., 1993; Gantzer et al., 1998; Sinton et al., 2002; Hot et al., 2003; Ackermann et al., 2004; Bitton, 2005; Lodder and de Roda Husman, 2005; Pillai et al., 2006; Jończyk et al., 2011; Bertrand et al., 2012; Seo et al., 2012; Silverman et al., 2013). Romero et al. (2011) indicated differences in solar inactivation rates between MS2 and rotavirus to their different protein capsid structure and genomes. While there are differences in survival among viruses of different families, there are also differences in survival among viruses within the same family (Sobsey and Meschke, 2003; Nappier et al., 2006). Also, coliphages within the same family and with similar structural similarities do not necessarily share the same survival characteristics (Jończyk et al., 2011). For example, results from laboratory studies showed that different F-specific RNA coliphages differ in their survival in water (Brion et al., 2002; Schaper et al., 2002b; Long and Sobsey, 2004; Nappier et al., 2006).

2.2. Detection Methods

Currently there is a variety of methods available to detect bacteriophages. These include culture-based methods and “rapid” methods (less than or to 24 hours) which can be immunology and molecular-based. Plaque assays are a typical culture-based technique used for identifying infectious virus particles (ISO, 1995, 2000, 2001; Grabow, 2001; U.S. EPA, 2001a, b; Eaton et al., 2005; Rodríguez et al., 2012a). There are three bacteriophage methods published by the
International Organization for Standardization (ISO) for F-specific RNA bacteriophages, somatic coliphages, and bacteriophages infecting Bacteroides fragilis (B. fragilis) (ISO, 1995, 2000, 2001). Rapid methods include immunology based methods (i.e., culture latex agglutination and typing [CLAT]), molecular methods (multiple types of PCR), and Fast Phage (a modified rapid version of EPA Method 1601) (Brussaard, 2004, 2009; Fong and Lipp, 2005; Kirs and Smith, 2007; Love and Sobsey, 2007; Gentilomi et al., 2008; Salter et al., 2010; Rodríguez et al., 2012b). Eaton et al., 2005). The ISO methods have been optimized and tested through interlaboratory comparison (Mooijman et al., 2001, 2002, 2005; Muniesa and Jofre, 2007). The ISO Standard Method 9224A-F provides protocols for detecting or enumerating coliphages (Eaton et al., 2005). Two methods for coliphage monitoring in groundwater were approved by EPA in 2001 (U.S. EPA 2001a, b). These methods include EPA Method 1601 (two-step enrichment process) and EPA Method 1602 (single agar layer [SAL] method). EPA Methods 1601 and 1602 have undergone multi-laboratory validation (U.S. EPA 2003a, b). The results of these inter-laboratory comparisons support the use of these methods in the determination and enumeration of F-specific and somatic coliphages in groundwater (U.S. EPA, 2003a, b). These methods are approved in 40 Code of Federal Regulations Part 136 and can be used for detection of coliphages in wastewater. These culture-based methods have been applied to rivers, drinking water, surface water, storm water, and wastewater (Havelaar, 1987; Davies et al., 2003; Borchardt et al., 2004; Lucena et al., 2004; Sobsey et al., 2004; Ballester et al., 2005; Lodder and de Roda Husman, 2005; Nappier et al., 2006; Stewart-Pullaro et al., 2006; Bonilla et al., 2007; Locas et al., 2007, 2008; Gomila et al., 2008; Love et al., 2010; Francy et al., 2011; Rodríguez et al., 2012a). EPA Method 1601 describes a qualitative two-step enrichment procedure for coliphages and was developed to help determine if groundwater is affected by fecal
contamination (U.S. EPA, 2001a). However, this validated procedure determines the presence or absence of F-specific and somatic coliphages in groundwater, surface water, and other waters (U.S. EPA, 2003a). The Method 1601 protocol directs that a 100 mL groundwater sample be enhanced with a log-phase host bacteria E. coli Famp for F-specific coliphages. After an overnight incubation, samples are put on to a patch of host bacteria specific for each type of coliphage, incubated, and examined for circular lysis zones. If circular lysis are present this indicates coliphages in the sample. For quality control purposes, both a coliphage positive reagent water sample, control, and a negative reagent water sample are analyzed for each type of coliphage from each sample. This method is considered more sensitive than EPA Method 1602, a SAL procedure discussed below (U.S. EPA, 2001a), due to the larger sample volumes used in 1601 (100 mL to 1 L) compared to Method 1602 (100 mL). In total, EPA Method 1601 requires 28 to 40 hours for a final result, depending on incubation times (Salter et al., 2010). The EPA Method 1602 SAL procedure can be used to quantify coliphages in a sample. The Method 1602 protocol directs that a 100 mL water sample may be assayed by adding the log-phase host bacteria E. coli Famp for F-specific coliphage and 100 mL of tryptic soy agar to the sample. The sample is then thoroughly mixed and the total volume is poured into multiple plates. After an incubation of 16 to 24 hours, circular lysis zones (plaques) are counted and summed for all plates from a single sample. The quantity of coliphages in a sample is expressed as PFU per 100 mL. For quality control purposes, both a coliphage-positive reagent water sample and a negative reagent water sample are analyzed for each type of coliphage with each water sample. In total, EPA Method 1602 typically requires an overnight incubation (18 to 24 hours) up to 3 days, but results can be obtained in as few as 8 to 10 hours (Salter et al., 2010). There are also methods for coliphage detection that use membrane filters to concentrate coliphages from a water sample.
Sobsey et al., 1990; Sobsey et al., 2004; Eaton et al., 2005). Coliphages can then be taken off the filter and used in one of the standard assays above, or they can be enumerated directly on the membrane filter (Eaton et al., 2005). For direct filter assays, a single assay dish is utilized for each coliphage-adsorbed filter. One study evaluated the use of a single E. coli host (Escherichia coli host strain CB390) for the detection of both somatic and F-specific coliphages at the same time (Guzmán et al., 2008). This host could be useful for detecting total coliphages. However, more independent and multi-laboratory validation of this method is needed. Rose et al. (2004) used E. coli C-3000 (ATCC #15597), which they report can host both somatic and F-specific coliphages.

2.3 Epidemiological Relationships

Since the 1950s, epidemiological studies have been performed to evaluate relationships between fecal indicators and recreational swimming-associated illnesses in surface waters. The occurrence of symptoms associated with gastrointestinal, eye, ear, and respiratory illnesses has been found to be higher in swimmers than in non-swimmers in ambient waters (Prüss, 1998; Wade et al., 2003; Zmirou et al., 2003). Throughout the years, EPA has conducted a plethora of epidemiological studies in both marine and freshwaters to evaluate the relationship of water quality indicators and human health risks. The results of an epidemiological study conducted by Cabelli et al. (1982) found that densities of enterococci in marine and freshwaters correlated with incidences of swimming-associated gastrointestinal illness, whereas densities of E. coli were correlated with swimming-associated gastrointestinal illness only in freshwaters. EPA’s NEEAR study found that the occurrence of gastrointestinal illness in swimmers was positively associated with exposure to levels of enterococci calculated by EPA’s Enterococcus qPCR Method 1611 in marine and freshwater (Wade et al., 2008, 2010; U.S. EPA, 2012). The odds of gastrointestinal
illness was higher among swimmers compared to non-swimmers on days were coliphages were detected, but the associations did not achieve statistical significance (Wade et al., 2010). In 1982, Cabelli et al. suggested that viruses were a primary cause of gastrointestinal illness, in agreement with quantitative microbial risk assessment (QMRA) modeling that used data from the NEEAR freshwater study (Soller et al., 2015). QMRA modeling demonstrated that the illnesses reported during the NEEAR study were consistent with a virus that had an incubation period similar to NoV (Soller et al., 2015). A consistent association between FIB (E. coli and enterococci) and illness has not been reported at all beaches where epidemiological studies have been conducted (Colford et al., 2007). This may be due partially to the fact that FIB in surface waters can come from sources other than wastewater, such as rainfall, plants, runoff, animals, and human shedding. In some subtropical and temperate climates, bacteria, such as E. coli and enterococci, can multiply in the environment, giving a false impression of an increase in fecal pollution (Solo-Gabriele et al., 2000; Yamahara et al., 2009). Additionally, compared to non-spore-forming FIB, human enteric viruses have been found to be more persistent in water environments and more resistant to physical antagonism, such as heat (55°C) (Lee and Sobsey, 2011). Numerous studies have been conducted to determine whether both somatic and F-specific coliphages are associated with fecal contamination (Chung and Sobsey, 1993; Mocé-Llivina et al., 2005; Love and Sobsey, 2007). Only a few epidemiological studies have evaluated the use of coliphages as an indicator of human fecal contamination in recreational water.
CHAPTER III-Methodology

3.1. Research Areas

The study investigates two different research areas (RA) linked to the research questions:

1. RA.1 Compare MS2 bacteriophage levels present in all water sample of Tanyard Creek and how they change spatially (from site to site) and temporally (January 2018- Present).

2. RA.2 Compare the relationship between MS2 bacteriophage and E. coli levels across all sampling dates by sites.

The primary data collection consisted of weekly water samples from Tanyard Creek CSO located off Collier Road at Ardmore Park which is considered part of the Atlanta Beltline. All samples collected from the creek were then brought to the lab for analysis of microorganisms, 10 different sites are collected for analysis purposes. Parts of the combined sewer system, the same system can be used to treat storm water runoff from the urban core portion of the CSO area, includes the central part of Atlanta. Storm water is a significant source of pollution in the streams and rivers.
For analysis of E. coli, the Membrane Filtration (MF) method is used to estimate bacterial populations in water that is low in turbidity. This method is especially useful for large sample volumes or for many daily tests. Using the membrane filter technique, sample is passed through the membrane using a filter funnel and vacuum system. Any organisms in the sample are concentrated on the surface of the membrane. The membrane, with its trapped bacteria, is then
placed in a petri dish containing the medium RAPID'E.coli 2 chromogenic medium. The agar medium provides direct enumeration, of E. coli in water samples. This test is designed for the simultaneous detection and enumeration of E. coli and total coliforms in water. RAPID'E.coli 2 is based on detection of β-D-glucuronidase (GLUC) and β-D-galactosidase (GAL) activities. Coliforms (GAL+/GLUC-) form green colonies, whereas for E. coli (GAL+/GLUC+) the combined GAL and GLUC (pink) enzyme activities result in violet colonies. After filtration plates are incubated at 44.5°C for 24 hours. Purple colonies (E. coli) are counted and expressed as colony forming units (CFU)/100 mL. When the results are read, and the total number of colonies exceeds 200 per membrane or the colonies are too indistinct for accurate counting, it was reported as “Too numerous to count” (TNTC).

For MS2 bacteriophage analysis Easyphage commercial test kit was used to analyze water samples for coliphage. This also known as a plaque assay that works for bacteriophages that have the ability to form plaques on a host like male specific coliphages (MS2). This requires 100mL sample water, 350 μL bacterial stain, 5 mL of prepared E. Coli Famp host that has been shaking/incubated overnight to select for coliphages. The process starts with adding antibiotic Streptomycin/Ampicillin to Easy Phage plates to keep another organism from the sample. Each Easy Phage kit is for one site; each site is poured up into 10 plates and there are 10 sites. After pouring up the Easy Phage gel/agar and allowing it to sit on bench for an hour, they were incubated at ~37°C for 18-24 hours and then read. While they are incubated viruses in the mixture attach to cells and begin to process the infection. Reading the plates is a direct count method for assaying virus infectivity. The plaques are visible clear spots that are counted to calculate the amount of virus present in each sample.
3.2. Data analysis

Data analysis was conducted using Graph Pad Prism 5. This data set included 10 different sites for 37 weeks. For multiple sites and dates, the program calculated an average of each site and date to be displayed on box and whisker plots. Box and Whisker plots are used to understand the variability, spread, and trends of E.coli and MS2 as well as the central tendency. The primary focus of the data analysis will be sites by dates and dates by site specific to MS2 and E.coli across all sampling dates by site for comparison. The statistical analysis will conducted on a spatial level to examine if there was statistical significance at the p < .05 level when comparing the group of sites to the current EPA water quality standard.
CHAPTER IV – Results

4.1. Results of E.Coli

Graph 1.1. E.coli Across 10 Sampling Sites By Date

Parts of the combined sewer system, the same system can be used to treat storm water runoff from the urban core portion of the CSO area, includes the central part of Atlanta. Storm water is a significant source of pollution in the streams and rivers.

All results are expressed as colony forming units per 100 milliliters. For E. Coli across the 10 sampling sites by date there is a spike in the counts during the warmer months (May-September)
(Graph 1.1.) With the smallest amount of e.coli being 20.00 CFU/100mL and the largest count at 30,300.00 CFU/100mL.

Graph 1.2. E.coli Across 10 Sampling Site by Date (Log10)

**E. coli Across 10 Sampling Sites By Date**

Graph 1.2. displays the log10 transformed data, which is done to normalize the distribution.

There appear to be higher counts during the summer. The ends of each ‘box’ in the box-plot are the upper and lower quartiles (25 percent of the sites are either higher or lower than these values). The top and bottom ‘whiskers’ represent the highest and lowest value. The middle line
of the box represents the median (middle) data point (half the sites are above and half below this value).

Analysis was conducted on Graph Pad Prism as a One sample t and Wilcoxon Test to compare two groups, the spatial data collected and analyzed from Tanyard Creek in comparison to the current Recreational Water Quality Criteria of 200 CFU/100ml. Indicating that all sites of Tanyard Creek are statistically significant when (P<0.05) and can be considered an impaired body of water.

Figure 2.1. Spatial Analysis Output

<table>
<thead>
<tr>
<th>One sample t and Wilcoxon test</th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 3A</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 6A</th>
<th>Site 7</th>
<th>Site 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Theoretical mean</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
<td>200.0</td>
</tr>
<tr>
<td>2 Actual mean</td>
<td>827.6</td>
<td>827.6</td>
<td>2186</td>
<td>1744</td>
<td>2526</td>
<td>1651</td>
<td>1961</td>
<td>3932</td>
<td>2435</td>
<td>1879</td>
</tr>
<tr>
<td>3 Number of values</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>35</td>
<td>34</td>
<td>32</td>
<td>33</td>
<td>21</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 One sample t test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 t, df</td>
<td>t=2.446, df=33</td>
<td>t=2.795, df=33</td>
<td>t=2.625, df=33</td>
<td>t=2.982, df=34</td>
<td>t=3.152, df=33</td>
<td>t=2.829, df=31</td>
<td>t=2.930, df=32</td>
<td>t=2.659, df=20</td>
<td>t=2.069, df=31</td>
<td>t=2.144, df=32</td>
</tr>
<tr>
<td>7 P value (two tailed)</td>
<td>0.0199</td>
<td>0.0068</td>
<td>0.0129</td>
<td>0.0053</td>
<td>0.0034</td>
<td>0.0081</td>
<td>0.0062</td>
<td>0.0175</td>
<td>0.0470</td>
<td>0.0397</td>
</tr>
<tr>
<td>8 P value summary</td>
<td>*</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
<td>&quot;    &quot;</td>
</tr>
<tr>
<td>9 Significant (alpha=0.05)?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>10 How big is the discrepancy?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Discrepancy</td>
<td>236.5</td>
<td>627.6</td>
<td>1986</td>
<td>1544</td>
<td>2326</td>
<td>1451</td>
<td>1761</td>
<td>3732</td>
<td>2235</td>
<td>1679</td>
</tr>
<tr>
<td>12 SD of discrepancy</td>
<td>129</td>
<td>1314</td>
<td>4381</td>
<td>3063</td>
<td>4303</td>
<td>2901</td>
<td>3452</td>
<td>6603</td>
<td>6112</td>
<td>4488</td>
</tr>
<tr>
<td>13 SEM of discrepancy</td>
<td>219.3</td>
<td>225.4</td>
<td>747.9</td>
<td>517.7</td>
<td>738.0</td>
<td>512.9</td>
<td>661.0</td>
<td>1441</td>
<td>1068</td>
<td>783.0</td>
</tr>
<tr>
<td>14 95% confidence interval</td>
<td>90.3 to 982.7</td>
<td>169.2 to 1086</td>
<td>444.3 to 3487</td>
<td>491.9 to 2596</td>
<td>824.8 to 3828</td>
<td>404.6 to 2497</td>
<td>536.9 to 2985</td>
<td>726.6 to 6736</td>
<td>31.38 to 4439</td>
<td>83.70 to 3274</td>
</tr>
<tr>
<td>15 R squared (partial eta squared)</td>
<td>0.1535</td>
<td>0.1903</td>
<td>0.1731</td>
<td>0.2074</td>
<td>0.2214</td>
<td>0.2052</td>
<td>0.2116</td>
<td>0.2512</td>
<td>0.1213</td>
<td>0.1256</td>
</tr>
</tbody>
</table>
Table 2.1. E.coli Across All Dates by Sites

![E. coli Across All Dates By Sites](chart1)

Table 2.2. E.coli Across All Dates by Sites (Log10)

![E. coli Across All Dates By Sites](chart2)
From the output of the graphs from site 5 through site 8 have a more considerable larger variation in e.coli counts compared to the other sites. Site 1 is the first site downstream from the CSO. Site 4 is right before an active beaver dam and site 5 is immediately after. Site 5 onward would include inputs from both the CSO and from the beaver dam. Directly after the beaver dam found between site 4-5 is where the highest number of viruses was detected. The sites after have higher numbers compared to the sites before the beaver dam. Some studies have found associations of E.coli with beaver dams (Fenwick [2006]; Steinmann et al. [2006]); they may influence the chemical and biological properties of the stream water itself (Margolis et al., 2001 & Rosell et al., 2005)

4.2. Results of MS2 Bacteriophage
MS2 and Sites by Dates Characteristics

Results for bacteriophage MS2 are expressed as plaque forming units per 100 milliliters.

Table 3.1. MS2 Across 10 Sampling Site by Dates
Table 3.2. MS2 Across 10 Sampling Site by Dates (Log10)

**MS2 Across 10 Sampling Sites By Dates**

![Boxplot showing MS2 levels across 10 sampling sites by dates, with Log10 PFU/virus values for each date from 08/02/18 to 09/13/18.]

**MS2 and dates by sites characteristics**

Table 4.1. MS2 Across 10 Sampling Dates By Sites

**MS2 Across 10 Sampling Dates By Sites**

![Bar chart showing MS2 levels across 10 sampling dates by sites, with Log10 PFU/virus values for each site.]

Table 4.2. MS2 Across 10 Sampling Dates By Sites (Log10)
Graph 5.1. Rainfall Averages

Retrieved from the National Weather Service to show the “normal” amount of rainfall in the United States compared the averages in Atlanta. From this chart it is noticeable that Atlanta’s
rainfall average is twice the national average during the months of July and August, with the normal inches averaging 4.26 inches compare to Atlanta’s average which is 5.13 inches.

Graph 6.1. Sample Means of E.coli by Sites

![Sample Means of E.coli by Sites](image)

This chart shows the sample means of E.coli by sites on a Log10 scale, site 3A has the largest amount 2.65 log10 cfu/100mL and the smallest is amount is at site6A with 1.69 log10 cfu/100ml.

Graph 7.1. Sample Means of E.coli by Dates

![Sample Means of E. coli By Dates](image)
The overall mean E. coli (or MS2) across all samples by dates is 2.343 Log10 CFU/100ml and the median E. coli of all samples by dates is 2.222 Log10CFU/100ml. With a standard deviation of 0.8674 across all samples.

Graph 8.1. All Samples of E.coli Colonies

The months with the highest and lowest means: April, August, and September all had one site with 0 colonies. The highest mean score was found in July and following closely behind is August. What is also interesting about this data once plotted is that, Site 8 had both the highest and lowest means. The overall mean E. coli (or MS2) across all samples was 2.64 Log10 CFU/100mL. The median E. coli (or MS2) across all samples is 2.54 Log10 CFU/100ml. As for the standard deviation across all samples it was calculated to be 0.75. The max value is 4.481 Log10 CFU/100ml from 7/12/18 at site 7 and min value 0.959 at site 6 on 4/13/18.
CHAPTER V-Discussion

5.1. Discussion of Research Question

The purpose of this thesis was to evaluate the spatial and temporal patterns in Tanyard Creek of E. coli and MS2. E. coli is a common bacterium found in the digestive system of humans and warm-blooded animals, making it a sign of the presence of fecal contamination from people. MS2 Bacteriophages are useful indicators of viruses that come from human feces, because viruses can act differently compared to bacteria when they end up in the environment.

Some possible sources of fecal contamination in a creek like Tanyard include: agricultural runoff, wildlife that uses the water as their natural habitat, and wastewater treatment plants. Heavy precipitation may cause these organisms to be washed into creeks, rivers, streams, lakes, or ground water. If this water is used as a source of drinking water and is not treated, it may result in illness. Diseases obtained from contact with contaminated water can cause gastrointestinal illness, respiratory, and wound infections. The most commonly reported symptoms are stomach cramps, diarrhea, nausea, vomiting, and fever. When E. coli exceeds the permissible level in recreational water, it results in the closing of beaches, lakes, and swimming and fishing areas. There are lower thresholds for levels of bacteria in drinking water from public water systems, that have been set by the Safe Drinking Water Act. The acceptable level of E. coli is determined by risk analysis based on statistics to protect human health. Drinking water should have no E. coli after treatment. E. coli levels at designated swimming beaches should not exceed 88 per 100 milliliter (mL) in any one sample, or exceed a three-sample average over a 60-day period of 47/100mL. Recreational waters that are not designated beaches should not have more than 406 E. coli/100mL in any one sample, or more than 126/100mL in a 60-day, three-sample average. Occasional higher numbers are not unusual, particularly after storms and where urban
or agricultural runoff occurs. These levels are generally not considered unsafe unless investigation indicates the source to be sewage. The 1986 criteria document includes EPA recommendations to use enterococci for marine and fresh recreational waters (a GM of 33 enterococci cfu per 100 mL in fresh water and 35 enterococci cfu per 100 mL in marine water) and E. coli for fresh recreational waters (a GM of 126 E. coli cfu per 100 mL) (U.S. EPA, 1986).

Figure 3.1. 2012 Recreational Water Quality Criteria (RWQC)

<table>
<thead>
<tr>
<th>Criteria Elements</th>
<th>Estimated Illness Rate (NGI): 36 per 1,000 primary contact recreators</th>
<th>Estimated Illness Rate (NGI): 32 per 1,000 primary contact recreators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicator</td>
<td>GM (cfu/100 mL)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>STV (cfu/100 mL)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Enterococci – marine and fresh</td>
<td>35</td>
<td>130</td>
</tr>
<tr>
<td>OR</td>
<td>126</td>
<td>410</td>
</tr>
</tbody>
</table>

Duration and Frequency: The waterbody GM should not be greater than the selected GM magnitude in any 30-day interval. There should not be greater than a ten percent excursion frequency of the selected STV magnitude in the same 30-day interval.

Consumption or contact with water contaminated with feces of warm-blooded animals can cause many illnesses. Gastrointestinal discomfort is probably the most common symptom; however, pathogens that may cause only minor sickness in some people may cause serious conditions in others, especially in the very young, the elderly, or those with a weak immune systems.

The presence of E. Coli and MS2 in aquatic environments may indicate that the water has been contaminated with the fecal material of man or warm blooded animals. This signifies that the source of water may have been contaminated by pathogens or disease producing bacteria or
viruses. The significance of fecal coliform bacteria indicates the presence of sewage contamination of a waterway and the possible presence of other pathogenic organisms. The current EPA recommendations for body-contact recreation is fewer than 200 colonies/100 mL; for fishing and boating, fewer than 1000 colonies/100 mL; and for domestic water supply, for treatment, fewer than 2000 colonies/100 mL. The drinking water standard is less than 1 colony/100ml. What was discovered at Tanyard Creek has exceeded these recommendations for body-contact recreation of 200 colonies/100mL.

**MS2 Variation with E. coli**

In this study MS2 Bacteriophage and E.coli. levels do not appear to be related. Some studies have reported an association between the presence of coliphages and human viruses, while other studies have found no association between their presence in environmental waters (epa.gov). The results are strongly influenced by the environments in which the studies are conducted. For example, an association between indicators and pathogens has more often been reported for brackish and saline water than for freshwater (epa.gov).

**5.2. Study Limitations and Next Steps**

**Challenges to Assay**

Although, the Easyphage Test Kit helped with time constraints providing the pre-made nutrient medium that has a type of agar to form a gel during the incubation process. The method and testing kit, which is a relatively new product for testing water for MS2, had many challenges that can classified as poor-quality assurance such as how they were packaged, delivered, and sealed. The lack of consistency with the kits was also present after the incubation period when plates still do not solidify after 24 hours or the dye did not evenly spread throughout the dish when it was time to be read.
Wider Implications

A waterbody is considered impaired if it does not attain water quality standards. Standards may be violated due to different types of pollutants or an unknown cause of impairment. A waterbody is considered threatened if it currently reaches water quality standards but is predicted to violate standards by the time the next 303(d) list is submitted to EPA (des.hh.gov). The 303(d) list is a comprehensive public accounting of all impaired or threatened waterbodies, not concerned of the cause or source of the impairment or threat. The time should not have to wait until the list is submitted to the EPA, such standards should be kept up to date every day. Priority for regulation indicates how EPA has prioritized a waterbody for regulatory controls under the Clean Water Act. To bring waterbodies into compliance with water quality standards, EPA calculates the maximum amount of a pollutant that a waterbody can receive and still meet water quality standards. This Total Maximum Daily Load (TMDL) calculation provides the basis for permitting decisions under the CWA. A TMDL specifies the reductions needed to meet water quality standards and allocates those reductions among the pollution sources in the watershed. The objective of the TMDL process is to systematically identify impaired or threatened waterbodies and the pollutant(s) causing the impairment and ultimately establish a scientifically-based strategy for correcting the impairment or eliminating the threat and restoring the waterbody. From the results of this data Tanyard Creek would be classified as an impaired water body as it does not meet the recommended standard for body contact recreation.

Intervention and Solution

The best solution would be preventing Tanyard Creek from becoming even more contaminated than it is now. Contacting local agencies responsible for the pollution with revised methods of ensuring that the standards are being met. Finding different alternatives that are cost-effective to
reduce the pollution like local clean-up groups, high school extra-curricular clubs, and on the collegiate level such as capstone, thesis, and senior experience courses. This would allow for a new policy and standard for how recreational water is monitored and controlled. Possibly implementing an active surveillance especially during the warmer months. The idea of active surveillance will help detect the contamination issue before it’s too late since the only monitoring for e. coli being currently implemented is a passive one. A considerable number of studies on pathogen contamination have been conducted on a laboratory-scale. However, there should be more emphasis given to field-scale studies for enhancing the understanding of pathogen interactions in the environment. Integrating knowledge from multiple fields can increase the understanding of pollution levels and help create long-term strategies to improve water quality. This would include a national database easily assessible to those part of the environmental network. As well as, to finding better alternatives other than the monthly park clean-ups for the pollution in the water that contributes to the high counts of E. coli and MS2.

5.3. Conclusion

This data indicates that there is a trend among the water samples on a temporal level. During the warmer months of year August-September there is a higher count of E. coli. When data was presented on a spatial level it was discovered that the higher counts of e. coli were present after the beaver dam that is in the considered to be the half way point of Tanyard Creek Park. Further studies and water samples are needed to confirm these findings. Moving forward, this study can add how rain and low creek levels affect E. coli levels. With limited research on its potential health effects, policymakers, especially at the state level should understand this lack of knowledge is a notable barrier not only to scientific understanding but also to the improvement of public policy and public awareness.
WORKS CITED


