RECOVERY OF DNA AND RNA FROM MICROORGANISMS IN WATER SAMPLES

Catherine M. Dougherty

May 2020

Submitted to the Department of Biological and Allied Health Sciences,
Fairleigh Dickinson University, in partial fulfillment of the
requirements for graduation with University Honors in Biology.

Mentor: Dr. Sara E. Reynolds, Ph.D.

ACKNOWLDGEMENTS

This research was made possible through the cooperative efforts of Fairleigh Dickinson University, the Department of Biological and Allied Health Sciences, and the University Honors Program. I would like to express my sincerest appreciation to my mentor, Dr. Sara Reynolds for her knowledge and guidance over the past three years and to Dr. James Salierno for his feedback and assistance collecting water throughout my experiment. Special thanks to my parents, Jim and Ellen, and my siblings Sarah, Jimmy, and Emma, for their constant support and encouragement.

TABLE OF CONTENTS

ABSTRACT

This study utilized a systematically narrowing method of detection to identify the presence and diversity of three genera of bacteria within the Enterobacteriaceae family: *Escherichia, Shigella,* and *Salmonella*, in a local source of surface water fed by wastewater effluent. To do this, a technique was developed to isolate DNA and RNA from microorganisms in those samples, and use polymerase chain reactions (PCR) to amplify regions of the genetic material using species-specific primers. Preliminary research found that adapting the commercial DNA/RNA AllPrep Kit (Qiagen) to incorporate DNeasy PowerWater bead-beating technology (Qiagen) should provide the greatest amount of quality DNA and RNA from water samples with consideration of both cost and ease of use. Initial trials successfully isolated DNA from *E. coli* spiked water using the PowerWater Kit, as confirmed through gel electrophoresis. Due to the low biomass of environmental samples polymerase chain reactions were investigated as a method of nucleic acid amplification and species identification. Published literature was searched for previously validated species-specific primers targeting organisms in the Enterobacteriaceae family. Selected primer pairs targeted the *lacZ*, *lamB, tuf, eaeA* and the *SLT-I* genes to differentially detect *Escherichia, Shigella,* and *Salmonella* spp. All primers were tested using *E. coli* spiked samples, though only those for the *lacZ* and *lamB* genes successfully produced PCR products identifying the target organism. No primers successfully amplified DNA for environmental samples collected in the Whippany River surrounding the Morristown Wastewater Treatment Plant in Morristown, New Jersey. As such, future studies aim to establish a limit of detection for the PCR to determine the minimum concentration required to visualize a PCR product through gel electrophoresis. Primers will also be optimized and the DNA/RNA All Prep Kit will be tested so the novel protocol can be implemented. Ultimately, this will then be used to monitor fluctuations in freshwater biodiversity over time with respect to environmental and anthropogenic variables.

INTRODUCTION

Aquatic Diversity

Freshwater ecosystems are essential for the growth and fitness of some of the most rich and abundant organisms in the biosphere. These species are organized into three domains: Bacteria, Archaea, and Eukarya, based on differences in the cell's ribosomal RNA nucleotide sequence, membrane lipid structure, and sensitivity to antibiotics (Kaiser, 2018). Eukaryotic organisms contain cells with a nucleus, membrane-bound organelles, and membranes made of unbranched fatty acid chains connected to glycerol by ester linkages (Kaiser, 2018). Eukaryotes are further divided into four kingdoms: Animalia, Plantae, Protista, and Fungi. Multicellular, eukaryotic organisms found in the kingdom Animalia are often most familiar, including species of fish, amphibians, turtles, and aquatic insects. Multicellular, eukaryotic organisms in the kingdom Plantae are equally as well known, including duckweed, elodea, and cattails.

Surveys conservatively estimate that freshwater ecosystems provide suitable habitat for at least 126,000 plant and animal species (Balian et al., 2007). Of the freshwater animal species, the vast majority are insects, followed by vertebrates, crustaceans, arachnids and mollusks, in descending order. The remaining percentage is divided amongst rotifers, annelids, nematodes, flatworms, and other minor groups. In terms of freshwater macrophytes, or aquatic plants, there are an estimated 2,614 species distributed over 412 genera, which equates to approximately 1% of the total number of vascular plants known to date, a significant percentage given that macrophytes are typically found on land (Balian et al., 2007). These plants are vital to the success of freshwater ecosystems, as they provide a habitat and food source for other aquatic

consumers. Together with freshwater animals, these macroorganisms also provide a variety of services for humans, including flood protection, food, water filtration, and carbon sequestration (Collen et al., 2013).

This assessment of freshwater diversity focuses solely on animals and vascular plants in the eukaryotic domain. Other eukaryotic organisms, including protists and fungi; prokaryotic organisms, including archaea and bacteria; and viruses, which are neither eukaryotic nor prokaryotic, also play an essential role in freshwater ecology (Balian et al., 2007). Such individuals are classified as microorganisms, a broad label which encompasses species that require a host to survive and reproduce, as well as those that are self-sufficient. As implied by their name, they range in size from less than 100 nanometers to one millimeter in length (Batt, 2016). Unfortunately, due to their small size, microorganisms are difficult to monitor and thus are drastically under-studied in environmental surveys. As the biogeochemical importance of microorganisms in an ecosystem becomes increasingly apparent, it is imperative that assessments begin to focus on the diversity and fitness of these species in order to create a more complete image of freshwater biodiversity over time, especially with respect to environmental and anthropogenic variables (Balian et al., 2007; Newton et al., 2011).

Research to date has touched on the species richness of the remaining two eukaryotic kingdoms, Fungi and Protista. Fungi, such as mushrooms, molds, and yeast, are either multicellular or unicellular and have cell walls made of chitin. They obtain nutrients through absorption as decomposers, or though symbiotic relationships with other plant species (Kaiser, 2018). Protists are simple, unicellular eukaryotes typically classified by their mode of movement such as flagellates, ciliates, amoeboids, and

sporozoans. Common examples include algae, euglenoids, and protozoans (Kaiser, 2018). Preliminary surveys on the diversity in these kingdoms estimate that there are over 3,000 species of aquatic fungi, of which approximately 2,000 are likely restricted to freshwater (Balian et al., 2007; Shearer et al., 2007). Similarly, it is estimated that just under 2,400 species of freshwater protozoans exist and that sixteen phyla of protists contain free-living freshwater protozoan species (Balian et al., 2007; Finlay & Esteban,1998). These organisms are essential in bodies of freshwater as they graze on microbes commonly found in aquatic habitats.

The two domains found outside the eukaryotic kingdom, Archaea and Bacteria, are much more difficult to investigate due to the need to collect samples and utilize molecular biology tools to examine them (Newton et al., 2011). These domains include prokaryotic microorganisms formed from cells that lack a true nucleus and membrane bound-organelles. Unlike bacteria and eukarya, archaea are composed of branched hydrocarbon chains attached to glycerol by ether linkages and contain cells walls that lack peptidoglycan (Kaiser, 2018). Due to their composition, many archaea are considered extremophiles, meaning they can withstand life in extreme salinity (halophiles), temperature (thermophiles), and pH (acidophiles). Recent research has established that archaea are widespread, found in almost every habitat on the planet. It has also shown that these species have a profound impact on biogeochemical processes including methanogenesis, the formation of methane (Thauer et al., 2008); methane oxidation, the production of energy and assimilation of carbon from methane (Valentine, 2002); sulphate reduction, anaerobic respiration which uses sulfate as the terminal electron acceptor (Stams & Plugge, 2009); and ammonia oxidation, the initial and rate-

determining step of nitrification during which ammonia (NH_3) is converted into nitrate (NO_2^-) (Lehtovirta-Morley, 2018). Specifically in freshwater lake environments, archaea function in methane release and nitrogen transformation, which contributes approximately 6-16% of total natural methane emission on a global scale (Ma et al., 2016; Spring et al., 2000).

Lastly, the final prokaryotic domain, Bacteria, includes the first and most abundant life forms on Earth, and can be found in nearly every environment on the planet. Most bacteria are either harmless or beneficial as they facilitate nutrient cycling (Nixon, 1981), aid in digestion (Kim & Gilliland, 1983; Long et al., 2017), and even produce antibiotic metabolites (Abdulkadir & Waliyu, 2012; Mannanov & Sattarova, 2001; Raaijmakers & Mazzola, 2012). About 5% of bacteria are pathogenic, meaning they cause serious illness, disease, or death. The main bacterial diseases transmitted through drinking water include cholera, gastroenteritis caused by vibrios, typhoid fever and other serious salmonellosis, bacillary dysentery or shigellosis, and acute diarrheas and gastroenteritis (Cabral, 2010). Like archaea, bacteria play a crucial role in stimulating and replenishing nutrients in freshwater food webs. As primary decomposers and mineralizers, bacteria convert organic compounds into their inorganic components, thus cycling biologically active elements throughout the aquatic ecosystem (Newton et al., 2011). They also contribute to biomass production, creating a nutrient sink which can be utilized by other organisms for energy (Newton et al., 2011; Sanderman & Amundson, 2003), and to trophic coupling, by which microbial trophic interactions are linked to those of aquatic macroorganisms (Moore et al, 2018; Newton et al., 2011). Because of

their impact on biological relationships and carbon fluctuations, bacteria can be utilized to monitor elemental fluctuations, water quality, and climate change over time.

All bacteria are anatomically similar in that their cell walls contain peptidoglycan, a substance made of complex polysaccharide chains interlinked with short peptides that coat the outside of the plasma membrane. Bacteria can be broadly divided into two groups, gram-positive and gram-negative, based on differences in this cell wall structure. Gram-positive bacterial cell walls contain a thick layer of peptidoglycan, as shown in Figure I B, thus they retain crystal violet dye and turn a blue/purple color when subjected to a Gram stain (Sizar & Unakal, 2019). Gram-negative bacterial cell walls have a much thinner layer of peptidoglycan and are surrounded by an outer membrane, as shown in Figure I A. As a result, these types of bacteria do not hold the crystal violet dye in a Gram stain and appear clear under a microscope (Sizar & Unakal, 2019).

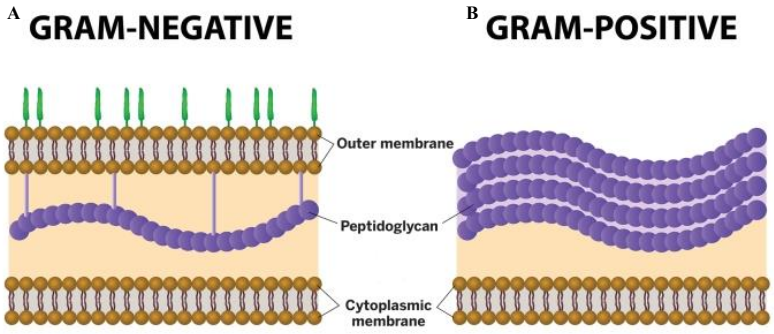

Figure I: Diagram depicting the difference in cell wall structure of gram-negative and gram-positive bacteria. (A) Gram-negative bacteria contain a thin layer of peptidoglycan surrounded by an outer membrane coated with liposaccharides (green). (B) Gram-positive bacteria have a thick layer of peptidoglycan and lack an outer membrane. (Image obtained from Berg, 2015).

Gram-negative bacteria can be further classified as coliform or non-coliform. Coliform bacteria are members of the family Enterobacteriaceae and are considered facultatively anaerobic, non-spore-forming rods, that contain the lacZ gene for β-galactosidase, the enzyme used to ferment lactose with the production of acid and gas (Gerba, 2015; Octavia & Lan, 2014). Non-coliform bacteria lack this gene and thus cannot cleave lactose into glucose and galactose (Gerba, 2015). Total coliform bacteria comprise a wide variety of generally harmless bacteria found in soil, water, vegetation, and animal or human waste (Figure II). Common species include *Escherichia, Citrobacter, Enterobacter,* and *Klebsiella* (Gerba, 2015; U.S. Environmental Protection Agency, 2012). While the presence of total coliform bacteria cannot confirm fecal contamination, it does indicate contamination from an outside source, and thus is utilized as an indicator of water quality. Fecal coliforms are a subdivision of total coliforms commonly found in the intestines and in animal and human waste (Figure II). Presence of fecal coliforms in drinking water suggests potential fecal contamination which poses a significant health threat to consumers (U.S. Environmental Protection Agency, 2012).

Escherichia coli presents a species of fecal coliform, commonly found in the gastrointestinal tract of mammals, exists in a mutualism with its host (Figure II). For instance, in humans *E. coli* aid in metabolizing food, provide essential vitamins such as vitamin K, and help protect against chronic diseases of the gut. In exchange, the human intestines provide nutrients to sustain the bacteria (Haque & Haque, 2017; Kaper et al., 2004). Despite this positive relationship, some *E. coli* clones obtained virulence factors which led to the development of new pathotypes of bacteria which can cause disease. These pathotypes are described by Kaper et al. (2004) and include enteropathic *E. coli*

(EPEC), enterohaemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC). The most common of these pathotypes is enterohemorrhagic *E. coli* serotype O157:H7, identified by shared O (lipopolysaccharide) and H (flagellar) antigens (Figure II) (Kaper et al., 2004). O157:H7 produces shiga-like cytotoxin (Imtiaz et al., 2013; Mead & Griffin, 1998) which results in serious illnesses including diarrhea; hemorrhagic colitis, or abdominal cramps followed by bloody diarrhea (Cohen & Giannella, 1992); hemolytic uremic syndrome, a condition which destroys platelets, drives anemia, and can cause kidney damage (Canpolat, 2015); and, in the most severe cases, death (Mead & Griffin, 1998).

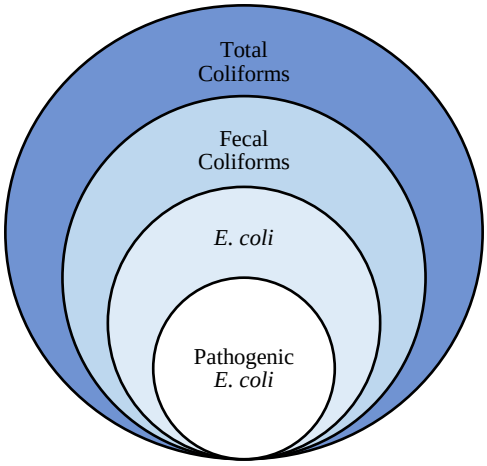

Figure II: Nested diagram illustrating the relationship between types of coliform bacteria indicators. Total coliforms are generally harmless, live in a wide variety of environments, and are used as general indicators of water quality. Fecal coliforms are intestinal bacteria often found in animal and human waste, thus their presence in water typically indicates fecal contamination. *E. coli*, a subset of fecal coliforms, are typically mutualistic bacteria in the human body but can also include certain disease-causing pathotypes.

Two other genera within the Enterobacteriaceae family commonly found in aquatic environments are *Shigella* and *Salmonella*. *Shigella* are gram-negative, facultatively anaerobic, non-spore-forming rods, similar to total coliforms (Hale & Keusch, 1996). The genus is broken into four species: *S. dysenteriae* (Group A) *S. flexneri* (Group B), *S. boydii* (Group C), and *S. sonnei* (Group D) (Hale & Keusch, 1996). Research has shown that the nucleotide sequences of *Shigella* and *E. coli* are an estimated 80-90% similar, thus they are often treated as a single genetic species and cannot be distinguished based on DNA sequences alone (Brenner et al., 1972; Devanga Ragupathi et al., 2018; Maheux et al., 2009, 2011). Despite the many similarities, *E. coli* and *Shigella* remain in two separate genera based on differences in biochemical and pathogenicity tests (Maheux et al., 2011). For instance, *Shigella* are nonmotile due to a deletion in the operon that codes for flagella. Additionally, *Shigella* cannot metabolize lactose because most species do not contain the *lacZ* gene for β-galactosidase. Though *S. sonnei* has the gene, it is unable to ferment due to a failure in functionality of the permease enzyme (Devanga Ragupathi et al., 2018). Shigella can be transmitted to water through fecal contamination and, if consumed, the pathogen can cause shigellosis or bacillary dysentery, an intestinal infection which leads to severe diarrhea with blood or mucus in the stool (Devanga Ragupathi et al., 2018; Hale & Keusch, 1996).

Like *Shigella*, *Salmonella* are gram-negative, facultatively anaerobic, non-spore-forming rods that present a major cause of foodborne illness (Kasturi & Drgon, 2017). The genus is divided into two species *S. enterica* and *S. bongori*. The former is further classified into six subspecies: *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtenae* (IV), and *indica* (VI). *Bongori* was formally known as subspecies V

before it was reclassified as an independent species within *Salmonella* (Brenner et al., 2000). These subspecies can then be broken into either serogroups, based on solely O (liposaccharide) antigens, or serovars (like serotypes in *E. coli*), based on O and H (flagellar) antigens (Brenner et al., 2000; Kaper et al., 2004). The consumption of raw or undercooked meat or eggs, as well as drinking water contaminated with *Salmonella* can lead to a variety of serious illnesses including gastroenteritis, colloquially known as the stomach flu; enteric fevers, such as Typhoid fever; bacteremia, the presence of bacteria in the blood stream; septicemia, also known as sepsis, or bacterial blood poisoning; and focal infections, bacterial infections localized to a particular organ such as the gastrointestinal tract, liver, bone, or meninges (Bush & Perez, 2020; Giannella, 1996).

In conclusion, surveys and research have established the vast diversity of macroorganisms found in freshwater ecosystems. Microorganisms, though less studied, have also been found to play substantial roles in natural water sources through nutrient cycling, trophic coupling, and stimulating aquatic food webs (Newton et al., 2011). This microbial activity emits and absorbs greenhouse gases and fluctuates with environmental changes, thus making microorganisms vital to the study of climate change (Dutta & Dutta, 2016). Furthermore, microbes have also been used to maintain an accurate measurement of water quality in the ecosystem. More specifically, Enterobacteriaceae presents a family of bacteria commonly used as indicators in environmental microbiology. As described previously, total coliforms such as *Escherichia* are used to indicate contamination from an outside source, while specific species and pathotypes can suggest fecal contamination or particular waterborne pathogens (U.S. Environmental Protection Agency, 2012). Similarly, non-coliforms such as *Shigella* and *Salmonella* can

be transmitted through the water and cause serious diseases. These attributes allow these genera of microorganisms to be ideal measures of the efficacy of wastewater treatment. As such, this research aims to monitor the presence of Enterobacteriaceae in local sources of surface water fed by wastewater effluent, and to map out fluctuations in diversity due to climate change and anthropogenic effects (Figure III).

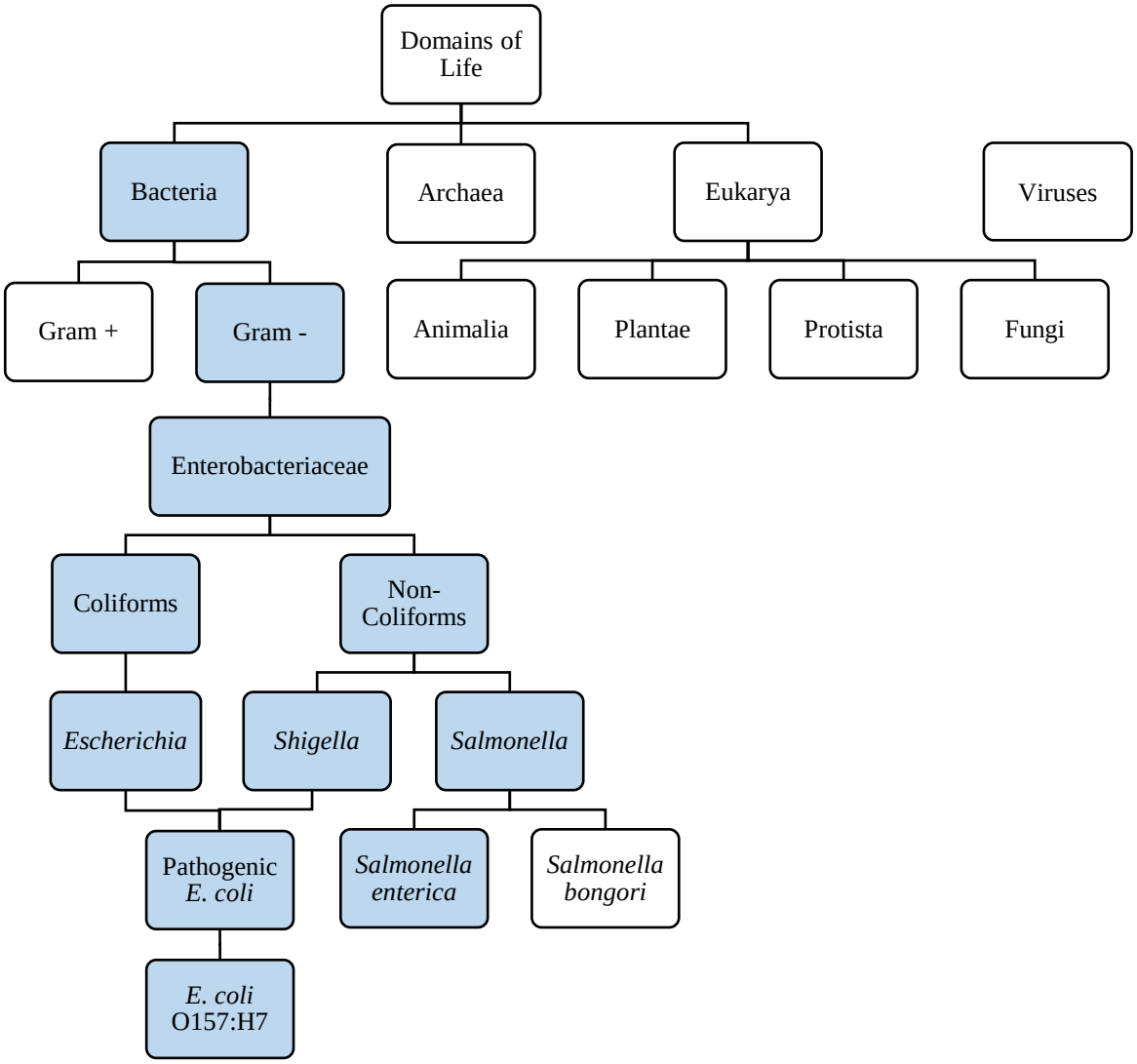

Figure III: Reverse pyramid showing the systematically narrowing method of detection utilized to identify the diversity of organisms found in water samples. This study focuses on testing for the presence of three genera of bacteria within the Enterobacteriaceae family: *Escherichia, Shigella,* and *Salmonella*, in a local source of surface water fed by wastewater effluent, as highlighted in blue in the flowchart.

LITERAURE REVIEW

Meta-Analysis of Extraction Methods

Research on the diversity of waterborne organisms in natural sources has revealed a myriad of microorganisms, however work remains to be done to better understand how this diversity varies over time due to environmental and anthropogenic impacts. Because these microorganisms are so small, it can prove especially difficult to survey and monitor for signs of contamination. Meanwhile, the need to develop a cost-effective, replicable methodology becomes increasingly dire. As of 2019 it was reported that approximately 2.1 billion people, or 29% of the world, do not have access to safe drinking water, an issue which is responsible for an estimated 1.2 million deaths each year (Ritchie & Roser, 2019). This is particularly serious for low-income countries where 6% of deaths are the result of water contamination yet they cannot afford to monitor water quality over extended periods of time, even with assistance from nonprofits (Ritchie & Roser, 2019).

To overcome the size challenge microorganisms pose, scientists utilize extraction techniques which isolate nucleic acids from samples and use this genetic material to characterize species of microorganisms and to hypothesize community relationships with macroorganisms. However, traditional methods for DNA or RNA extraction are time-consuming, require large quantities of reagents, and are prone to contamination (Yang et al., 2010), while many modern methods, though effective, are expensive and complex to perform (Tan & Yiap, 2009). Given this information, the first objective of this research was to perform a meta-analysis of common extraction techniques and evaluate each with respect to sample type, cost, ease of use, and nucleic acid extraction ability. The ultimate goal was to identify a method that allowed for the simultaneous extraction of DNA and

RNA to identify a wide variety of microorganisms, particularly Enterobacteriaceae in a single trial. This would minimize time, resources, and the likelihood of cross-contamination.

Methods of extraction are typically divided into four sections including pretreatment of the sample to increase quality and yield, lysis of cells to make nucleic acids accessible, isolation of nucleic acids, and purification of nucleic acids. The primary concern with water samples is that tend to be low biomass which can decrease yield. To improve this samples are filtered prior to analysis and pretreatment is avoided to minimize premature cell lysis (Giovannoni et al., 1990; Purdy, 2005). The lysis step of extraction is vital for maximum recovery and varies based on the extraction procedure. Chemical lysis is one technique by which a detergent is added to destruct the cell membrane followed by phenol as a lytic agent, and chaotrophic salts to disrupt proteins (Purdy, 2005). Common chemical lysis procedures include organic solvent-phenol-chloroform (Guillén-Navarro et al., 2015; Tan & Yiap, 2009), alkaline extraction using sodium dodecyl sulfate (SDS), and extraction with cetyltrimethylammonium bromide (CTAB), a non-ionic detergent (Tan & Yiap, 2009), amongst others. Enzymatic lysis presents a second technique by which specific enzymes are utilized to deconstruct the cell walls of certain types of microorganisms. For example, lysozyme is often used to break apart gram-negative bacteria (Guillén-Navarro et al., 2015; Purdy, 2005).

In addition to adding powerful chemicals or enzymes to samples, other means to lyse cells include mechanical lysis such as bead-mill homogenization, when cells are physically broken apart using glass beads that split cells open when shaken at high speeds (Moré et al., 1994; Purdy, 2005; Yu & Mohn, 1999), as well as freeze/thaw cycles,

freeze/boil cycles, microwave heating, and mortar mill grinding (Moré et al., 1994; Purdy, 2005). Improved technology has also fueled the popularity of automated systems, particularly in purifying DNA in the final step of extraction. Magnetic bead-based nucleic acid purification, for example, removes particles with a magnetic charge from the sample by applying a magnetic field (Tan & Yiap, 2009; Triant & Whiteheard, 2009; Yang et al., 2010). Column purification including Elutip-d, Elutip-r, microcon ultracentrifugation, and Qiagen spin columns have also grown popular (Purdy, 2005).

Through detailed analyses of each of these nucleic extraction protocols, it has been concluded that the most widely accepted extraction methods employ multiple lysis and purification techniques. Unfortunately, these procedures continue to present other challenges. Chemical and enzymatic lysis steps are prone to incomplete phase separation, cross-contamination, or carry-over into downstream applications such as PCR, thus decreasing purity and yield (Tan & Yiap, 2009; Yang et al., 2010). Manual column purification utilizes reagents and elutions of high volumes and can lead to filter clogging (Yang et al., 2010). The solid-phase nucleic acid extraction using bead-beating technology eliminates many of the concerns involved with contamination or excessive reagents, but the magnet itself is very expensive, approximately $800 (Qiagen), which fails to meet the cost-effective aspect of this research. Given that the most effective protocols incorporate many aspects of nucleic acid lysis and purification, the most straightforward, user-friendly approach may be to adopt a commercial extraction kit which includes all of the reagents for a single fee.

Work by Ficetola et al. (2008) utilized a commercial kit to confirm that DNA fragments are preserved in aquatic environments and can be used to reliably represent a

particular species' presence. In his experiment a bullfrog (*Rana catesbeiana*) acted as the species of interest and was evaluated in a controlled laboratory and in natural field conditions. DNA was extracted using the QIAamp Tissue Extraction Kit (Qiagen) and amplified using PCR with species-specific primers. The results showed that DNA was successfully amplified from water samples in an aquarium and in natural ponds where the bullfrog was present, while it was not amplified in water samples in an aquarium and in natural ponds where the bullfrog was not present (Ficetola et al., 2008). This reveals that environmental DNA can be utilized to confirm the presence of a macroorganism in wetlands and can differentiate between the absence and presence at different densities. This study demonstrates the role of environmental DNA extraction in the detection and quantification of harmful, invasive, or threatened species. It also highlights a series of factors that impact the amount of DNA extracted from water samples, including volumes of water and size and density of organisms present or absent (Ficetola et al., 2008).

It can be incredibly difficult to obtain adequate quality DNA in sufficient quantities from environmental samples, as explained above. Most existing extraction methods are susceptible to contaminants, and the physical or chemical changes in the ecosystem where the sample was collected can influence the results obtained. Guillén-Navarro et al. (2015) analyzed the effectiveness of three extraction techniques: phenol-chloroform, enzymatic lysis, and a lysozyme method, and compared the yields to results obtained from commercial DNA extraction kits: QIAamp DNA Stool Mini Kit and ZR Soil Microbe DNA Miniprep. Samples were collected from decaying coffee pulp and mangrove sediments, representing unstable environments, as well as compost and soil. All extraction methods produced DNA, the quantity of which was analyzed using gel

electrophoresis to estimate the amount recovered (Guillén-Navarro et al., 2015). The quality of DNA was analyzed through PCR with species-specific primers. Ultimately, the extraction method that resulted in the most DNA and allowed for PCR amplification of markers for all microbial groups was a lysis procedure combining physical, chemical, and enzymatic steps, thus demonstrating that extraction methods are optimized by combining different purification procedures (Guillén-Navarro et al., 2015). Though the extraction kits did not have the highest quantity yield, they were still found to successfully isolate DNA from unstable samples, thus showing potential for use extracting DNA from microorganisms in freshwater samples.

The prior two studies demonstrated the potential that commercial extraction kits present as a cost-effective, straightforward, and well-researched technique to isolate DNA from tissue, sediment, and soil samples, but they have also been optimized to simultaneously extract both DNA and RNA from water samples as well. Scientific methods that allow for the extraction of high-quality DNA and RNA from the same biological sample prove vital for molecular and microbiological studies that identify and monitor microbial diversity. These methods are especially important for low biomass samples, such as water, that usually show low nucleic acid yield. Unfortunately, common techniques for simultaneous extraction of DNA and RNA, including commercial kits, tend to yield either high quality DNA or RNA, but not both. As such, Triant & Whitehead (2009) developed an optimal technique for simultaneous extraction of high-quality DNA and RNA from fish embryos. To do this, total genomic DNA or RNA was isolated from embryos first as a baseline.

Following these initial tests, three possible simultaneous DNA and RNA extraction methods were compared. First, DNA extraction after RNA extraction was tested using a ChargeSwitch gDNA Tissue Kit from Invitrogen, substituting phenol phase and interphase for lysate. Second, DNA and RNA extraction from split portions of homogenate was tested by dividing homogenate immediately after homogenization and allocating a portion of it for DNA extraction with a Qiagen DNeasy Kit and the other for RNA extraction. Lastly, DNA and RNA extractions from commercial kits were tested using three other methods: the Qiagen AllPrep DNA/RNA Micro Kit; the magnetic bead technology of the MagMAX-96 total RNA isolation kit, dividing samples during final washing steps and applying DNase or RNase; and the bead-based sample disruption of the MagMAX Total Nucleic Acid Isolation Kit, dividing samples during the final washing steps and applying DNase or RNase (Triant & Whitehead, 2009).

Results of this experiment showed that all methods had successful isolation of RNA, with DNA also isolated and PCR amplified from all methods except ChargeSwitch. The Qiagen AllPrep DNA/RNA Micro Kit consistently produced high-quality RNA and DNA products easily visualized through gel electrophoresis, though the concentration of both nucleic acids was low. Examining the overall extraction and isolation of both DNA and RNA, it was found that the second option, dividing chaos buffer homogenate prior to nucleic acid purification, then separately extracting DNA and RNA from each portion, most consistently obtained high-quality DNA and RNA at high concentrations (Triant & Whitehead, 2009). Despite the use of fish embryos as samples, this study most closely relates to the objective of this experiment. It examines a comparison of different methods for isolating DNA and RNA, and shares the same end goal: development of an optimal

technique for simultaneous extraction of high-quality DNA and RNA. Additionally, it tested commercial kits including Qiagen AllPrep DNA/RNA Micro Kit and visualized DNA on agarose gels. Lastly, the results suggest that the protocol could be applied to small samples that typically do not yield high concentrations of nucleic acids, such as freshwater samples collected for this research (Triant & Whitehead, 2009).

A second study by Yang et al. (2010) compared six automated nucleic acid extraction systems and one manual kit for their ability to recover nucleic acids from human nasal wash specimens spiked with five respiratory pathogens representing Gram-positive bacteria, Gram-negative bacteria, DNA viruses, segmented RNA viruses, and non-segmented RNA viruses. Commercial kits were compared with respect to nucleic acid recovery, degree of inhibition, carry-over contamination, reproducibility, and linearity of protocols, as well as cost per extraction and the maximum number extractions per run (Yang et al., 2010). The results of this research showed that all methods exhibited similar performance in DNA extraction from adenovirus and *L. pneumophila*. The MagNA Pure Compact extracted more DNA from *S. pyogenes* than other methods but it was less efficient in RNA purification from RSV and influenza A virus. Additionally, the KingFisher mL and EasyMag recovered RNA from both viruses more efficiently than the EZ1, the MDX, and the AllPrep Kit. From this it was concluded that the structure and composition of bacterial cell walls or viral capsules can influence lysis efficiency. No single protocol was found to be superior for all of the agents tested; protocol performance was determined to be pathogen specific (Yang et al., 2010).

Comparison of Commercial Kits

Ultimately the DNA and RNA extraction techniques investigated revealed that there may be no single kit or extraction method that meets all of the needs of this research objective. It also provided potential evidence to support the development of a novel protocol that meets all the requirements of this research. The studies concluded that the most popular and straightforward techniques to extract nucleic acids utilize commercial extraction kits. Based on this, the eight most common commercial kits for extraction from environmental samples were investigated to determine which would yield the best results with consideration of cost, ease, and ability to extract both DNA and RNA from water samples. All kits examined are shown in Table I.

The DNeasy and RNeasy PowerSoil Kits and the QIAamp DNA Kit were eliminated as options as they are designed to analyze soil or tissue samples. Other kits in the running, such as the DNeasy and RNeasy PowerWater Kits specialized in water samples, but only extracted one nucleic acid. Because the objective of this research was to isolate both DNA and RNA, more kits were investigated to better meet these requirements. The MagAttract PowerWater DNA/RNA Kit and AllPrep DNA/RNA Kit allow for the simultaneous extraction of both DNA and RNA from samples, however, the former requires a PowerMag magnet, a powerful separator designed to work with automated liquid handling systems. This piece of equipment, combined with the physical kit, proved very costly, thus it was not selected as the ideal option.

Commercial Kit	Sample Type			Nucleic Acid		Avg Cost (per prep) (50-100 preps/kit)	Suitable for Downstream Applications (i.e. PCR)?	Additional Considerations
	Soil	Cells/ Tissue	Water	DNA	RNA			
DNeasy PowerSoil Kit	X		X	X		$5	Yes	Easy to follow protocol
RNeasy PowerSoil Kit	X				X	$6	Yes	High RNA sensitivity
DNeasy PowerWater Kit			X	X		$9	Yes	Optimized to increase yields from low biomass samples
RNeasy PowerWater Kit			X		X	$10	Yes	High RNA sensitivity
MagAttract DNA Kit		X		X		$4	Yes	Requires ClearMag technology (Magnet ~$780)
QIAamp DNA Kit		X		X		$3	Yes	Manual processing (not bead-beating)
MagAttract PowerWater DNA/RNA Kit			X	X	X	$6	Yes	Simultaneous extraction; Requires ClearMag technology (Magnet ~$780)
AllPrep DNA/RNA Kit		X		X	X	$11	Yes	Simultaneous extraction

Table I: Commercial DNA and/or RNA isolation and extraction kits were analyzed and compared with regard to sample type, cost, ease of use, and nucleic acid extraction ability. The above pool of kits depicts the most commonly used methodologies that fulfill at least one aspect of this experiment's objective: to develop a logical and effective protocol for the simultaneous extraction of DNA and RNA from water samples in a cost-effective manner. Average cost was calculated by dividing the total kit cost, if purchased online from Qiagen, by the number of preps each kit can prepare.

Ultimately, it was confirmed that no single kit allows from the simultaneously extraction of both DNA and RNA from water samples in a cost-effective manner, as stated by Yang et al. (2010). Based on these findings a novel approach was proposed that combines the protocols of the AllPrep DNA/RNA Kit (Qiagen) and the DNeasy PowerWater Kit (Qiagen). The AllPrep DNA/RNA Kit allows for the simultaneous extraction and purification of DNA and RNA isolated from tissue samples (Figure IV A). Through this procedure cell or tissue samples are first lysed by a denaturing buffer which deactivates DNases and RNases. The lysate then passes through an AllPrep spin column which allows for the selection and binding of genomic DNA from the sample. This can then be washed and eluted to isolate pure DNA. RNA is purified from the AllPrep column flow-through using a RNeasy Mini spin column (Qiagen, 2005).

Following a similar protocol, the DNeasy PowerWater Kit also isolates genomic DNA but from water samples (Figure IV B). The sample is first filtered onto a filter membrane. This is then added to a specialized PowerWater bead-beating tube where the filter and sample are lysed through vortexing in a lysis buffer designed to enhance microorganism isolation. Total genomic DNA can be gathered in an MB spin column where it is washed and eluted to complete the pure DNA extraction process (Qiagen, 2017). Because the AllPrep DNA/RNA Kit allows for simultaneous extraction of DNA and RNA, and the PowerWater Kit is optimized for isolation from water samples, one possible methodology that will be considered for this experiment is adapting the AllPrep Kit to incorporate the PowerWater bead-beating tube which could be used to break apart the filter and sample prior to extraction following the AllPrep protocol.

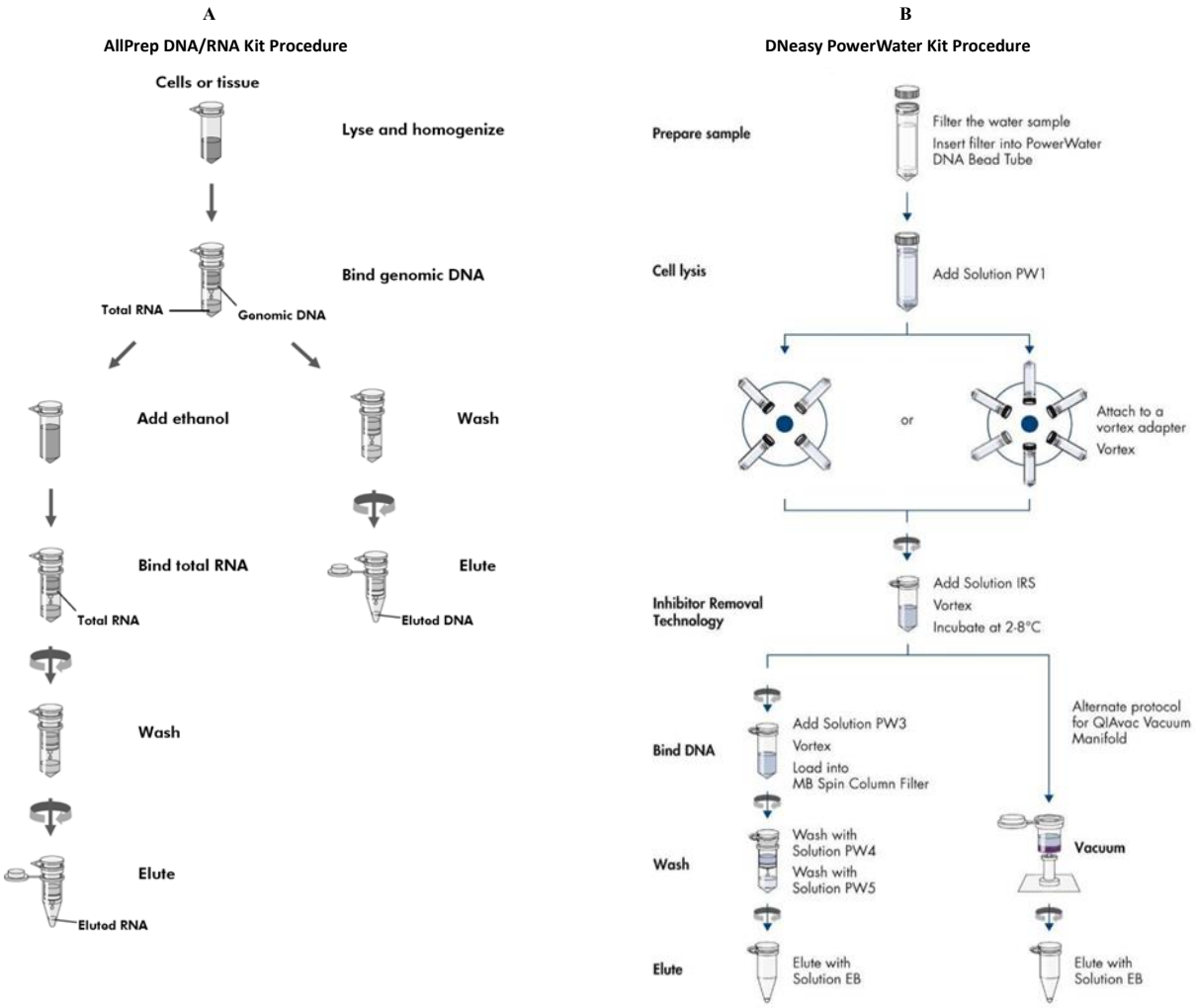

Figure IV: Schematic of extraction procedures using the (A) AllPrep DNA/RNA Kit and the (B) DNeasy PowerWater Kit, both by Qiagen.

Polymerase Chain Reaction

Following nucleic acid extraction from environmental samples, particularly water, the isolated DNA or RNA needs to be amplified to make larger samples to use in downstream applications. Polymerase chain reactions (PCR) presents a technique for DNA amplification in which DNA polymerase synthesizes a new strand of DNA complementary to the template strand (PCR, 2017). This technique poses many advantages including quick speed, low-cost, and high sensitivity and specificity (Higgins et al., 2001). However, it is important to consider that PCR detects both viable and dead cells and requires a clean starting sample (Nurliyana et al., 2018; Purdy, 2005). Additionally, DNA polymerase requires a 3'-OH group to begin replication, thus primers or short, single-stranded nucleic acids complementary to the target sequence are required in the reaction mixture to initiate synthesis of the new strand (PCR, 2017). Because primers are target sequence-specific, they allow researchers to identify a certain region of the DNA strand they wish to amplify, resulting in over 1 billion copies, or amplicons, of that particular region, which can then be used to detect organisms of interest and perform further genome analyses (PCR, 2017).

DNA replication utilizes a thermocycler, an instrument that increases and decreases the temperature for specific increments of time to allow proper denaturation and annealing of templates and primers. The PCR reaction mixture contains the following components: the DNA sample to be replicated; a forward and reverse primer to initiate duplication; four deoxynucleotide triphosphates (dATP, dTTP, dCTP, and dGTP) used as DNA building blocks; *Taq* polymerase, the enzyme required to synthesize new DNA strands; and a buffer used to create the ideal chemical environment for the reaction

(Figure V A). The PCR process is broken into three steps that repeat in multiple cycles (Figure V B). During the denaturation step, the isolated DNA sample containing the target sequence is heated to break the hydrogen bonds between complementary DNA strands (PCR, 2017). Once separated, the temperature is decreased to a primer-specific annealing temperature to allow primers bind to target DNA. This signals *Taq* polymerase to begin synthesizing new dNTPs at the end of the primers. In the final extension step the temperature is again decreased to allow *Taq* polymerase to elongate the new strand of complementary DNA in the 5' to 3' directions, following the template strand (PCR, 2017). This three-part process constitutes a single cycle of a polymerase chain reaction; after 25-35 cycles, billions of amplicons are produced and can be utilized in downstream applications (PCR, 2017).

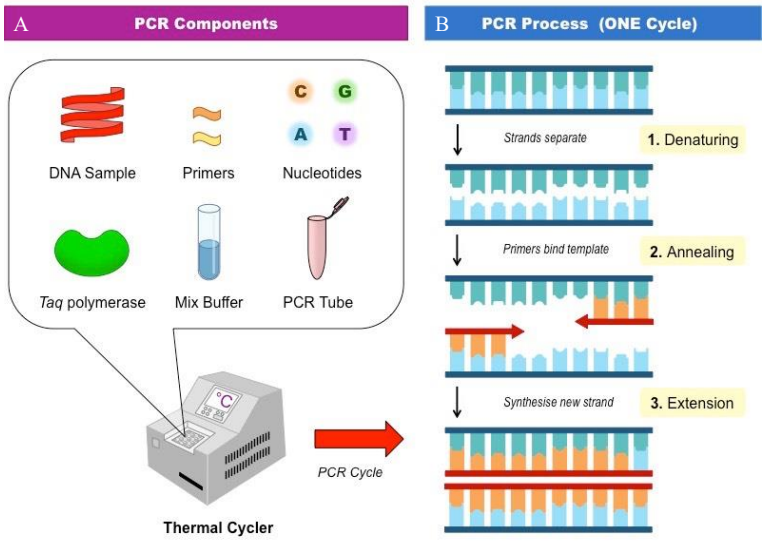

Figure V: Schematic of (A) PCR components and (B) procedure using a thermal cycler and Platinum *Taq* DNA polymerase (ThermoFisher Scientific). (Image obtained from Banerjee, 2020).

Identification of Primers

Given the assumed low biomass of environmental water samples, DNA and RNA yield from extraction is expected to be low. As such polymerase chain reactions are required to amplify the nucleic acids prior to analysis. As established previously, PCR utilizes species-specific primer sequences complementary to a certain template to initiate synthesis of a new strand (PCR, 2017). Having already established three genera of bacteria within the Enterobacteriaceae family: *Escherichia, Shigella, and Salmonella*, as the species of interest for this research (Figure III), it was then necessary to identify possible primer sequences that would target those species. As a BSL1 teaching facility, our laboratory lacks access to pathogenic species for PCR validation, so published literature was investigated in search of previously validated species-specific primers.

Coliform bacteria present members of the Enterobacteriaceae family commonly used as an indicator of water quality, as presence of coliforms indicates potential contamination from an outside source (U.S. Environmental Protection Agency, 2012). As previously mentioned, coliform bacteria are able to ferment lactose because they contain the *lacZ* gene for β-galactosidase, while non-coliform bacteria lack this gene and thus cannot break down lactose (Gerba, 2015; Octavia & Lan, 2014). As such, common primer sequences used as a broad detector of coliform bacteria target the activity of β-galactosidase on the *lacZ* gene. Using a PCR primer annealing temperature of 50°C, Bej et al. (1990) confirmed that amplification of a segment of the coding region on *lacZ* of *Escherichia coli* can be used to successfully detect *E. coli* and *Shigella* spp., but not *Salmonella* spp. and other non-coliform bacteria. Though *Shigella* are not lactose-fermenting, their genome is 80-90% similar to that of *E. coli*, thus they are treated as a

single genetic species and cannot be differentiated (Brenner et al., 1972). As such, primers targeting the *lacZ* gene have been identified as strong markers to monitor total coliform bacteria in freshwater samples (Figure VI). In the same study Bej et al. (1990) also suggested a primer that targets a region of the *lamB* gene in *E. coli*. This gene codes for the protein maltoporin which is involved in the transport of maltose and maltodextrins and acts as a receptor for several bacteriophages including lambda (Gene Blast, 2020). Using an annealing temperature of 50°C, Bej et al. (1990) successfully detected *E. coli, Shigella,* and *Salmonella* spp. using a primer amplifying a region of the *lamB* gene. This suggests that primers targeting the *lamB* gene can be utilized to selectively detect the three genera of Enterobacteriaceae of interest in this research (Figure VI).

Given that *Escherichia coli* and *Shigella* are not different enough to reliably detect independently, the next primer needed to selectively differentiate between this pair of genera and *Salmonella.* Maheux et al. (2009) analyzed nine potential primer sets designed to detect *E. coli* and *Shigella*. These primers targeted either the *uidA* gene which encodes β-glucuronidase, an enzyme that catalyzes reactions and is involved in carbohydrate and protein binding, or the *tuf* gene which codes for the elongation factor Tu and promotes the GTP-dependent binding of tRNA to the ribosome during protein synthesis (Gene Blast, 2020). When tested on a series of *E. coli* and *Shigella* strains, it was found that, of five primer sequences targeting the *uidA* gene, only three amplified DNA from all *E. coli* strains and none of those three also amplified DNA from all *Shigella* strains. Only the primer sequence targeting the *tuf* gene successfully identified all *E. coli* and *Shigella* strains tested (Maheux et al., 2009). From this it was concluded

that primer sequences targeting the *tuf* gene could be used in this research to selectively amplify DNA from the Enterobacteriaceae genera *Escherichia* and *Shigella* (Figure VI).

While Maheux et al. (2009) focused on identifying *E. coli* and *Shigella*, Kasturi and Drgon (2017) investigated primers to identify *Salmonella*. The primer they proposed amplified a segment of the *invA* gene in *Salmonella*, a gene which codes for the invasion protein invA which invades cells of the intestinal epithelium (Gene Blast, 2020). The exclusivity of the primer was tested on 12 control *Salmonella* organisms and then extended to 328 *Salmonella* isolates from food and environmental samples. Findings were confirmed using 22 non-*Salmonella* target organism including species of *E. coli* and *Shigella* for comparison (Kasturi & Drgon, 2017). Results of the experiment found that the *Salmonella*-specific primer set successfully identified all 329 *Salmonella* isolates that consisted of 126 serovars belonging to all subspecies of *S. enterica*. None of the non-*Salmonella* target organisms were identified (Kasturi & Drgon, 2017). Based on these results, primer sequences targeting the *invA* gene can be utilized to selectively amplify DNA from the Enterobacteriaceae genus *Salmonella* (Figure VI).

The presence of wastewater effluent in surface water suggests human pathogenic bacteria may be present in samples collected during this research. Barak et al. (2004) compared a series of primers with the goal of identifying one that gave specific and sensitive detection of pathogenic *E. coli*. One of the main gene targets for the primer sets tested was the *eae* gene in *E. coli*. This gene encodes a protein for gamma-intimin which functions in cell adhesion (Gene Blast, 2020). All primers were tested on numerous samples of *E. coli* O157:H7, other pathogenic *E. coli*, and one non-pathogenic serotype as a control. Results found that the primer set named conceaeA, targeting the *eaeA* gene,

successfully detected all O157 strains, as well as strains from the serotypes O111:H8 and O55:H7 (Barak et al., 2004). When compared to other primer sets tested, conceaeA exhibited lower detection limits of 10^0 CFU and 10^1 CFU for *E. coli* O157:H7 and O55:H7, respectively (Barak et al., 2004). For this reason, the primer set conceaeA, targeting the *eaeA* gene proves most sensitive for detection of pathogenic *E. coli* from water samples (Figure VI).

Lastly, research has shown that the *E. coli* serotype O157:H7 can cause serious illnesses and even death, as discussed previously. Immediate indication of its presence in contaminated water could minimize or prevent sickness and bodily damage, thus highlighting the importance of finding a primer sequence to specifically detect this serotype. O157:H7 produces shiga-like cytotoxin which causes illness (Mead & Griffin, 1998). With this knowledge, Imitaz et al. (2013) tested a primer set that targeted the *E. coli* virulence gene *SLT-I*, along with a second set targeting the *E. coli* gene *O157* using spiked water samples. Results shows that the second primer set failed to differentiate the O157:H7 serotype from other O157 isolates, while the *SLT-I* primer set was more successful. These findings suggest that primers targeting the *SLT-I* gene can be used to amplify DNA of the pathogenic *E. coli* serotype O157:H7 (Figure VI).

Ultimately, this literature investigation identified six genes of interest which can be used to test for the presence of the Enterobacteriaceae genera *Escherichia, Shigella,* and *Salmonella*, in a local source of surface water fed by wastewater effluent. The systematically narrowing approach of detection can help identify and monitor the diversity of organisms found in water samples. Figure VI shows a summary of the six genes and their respective target species.

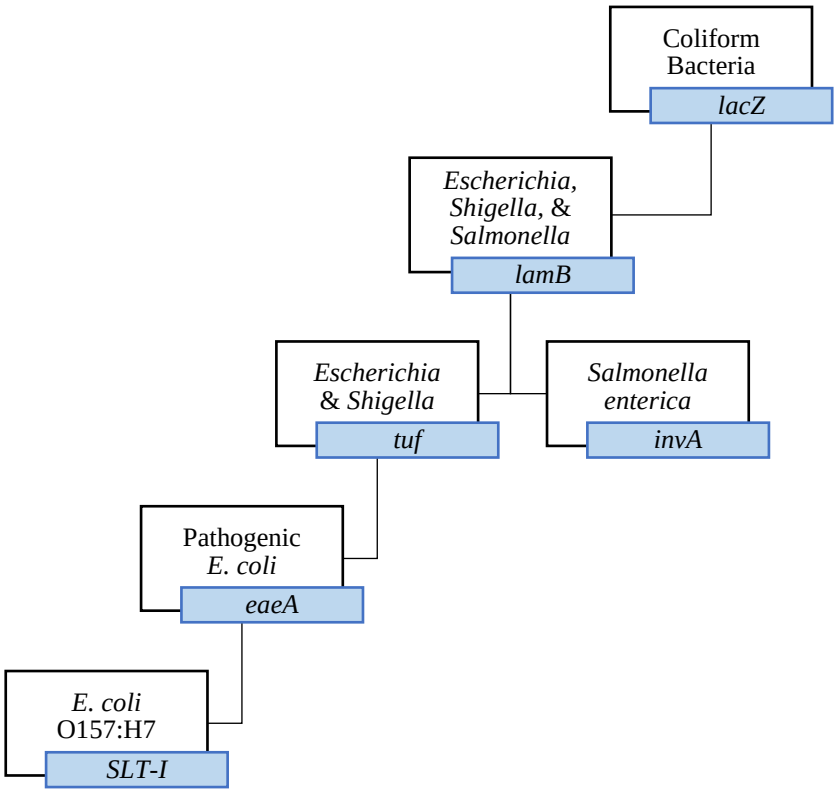

Figure VI: Flow chart illustrating results of a literature investigation of previously validated species-specific primers targeting organisms in the Enterobacteriaceae family. Genes of interest are shown in blue, with their respective target species shown in white.

MATERIALS AND METHODS

Water Samples

Experiments were performed using both water spiked with *E. coli* in a blind study and natural water samples collected from the Whippany River. Control samples were prepared by spiking deionized water with *Escherichia coli* HB101 K12 at a concentration high enough to ensure a positive result. Environmental water samples were collected in mid-September 2018 and early November 2019 during rainfall. Effluent samples were collected directly from the surface water of the source, an output pipe running from the Morristown Wastewater Treatment Plant in Morristown, New Jersey and emptying into the river. Upstream samples were collected 50 feet upstream of the source, directly underneath a road overpass. Downstream samples were collected 50 feet downstream of the source in the middle of the river. All surface water was gathered in sterile 500 ml Pyrex media bottles and stored at 4°C overnight until nucleic acid extraction.

Recovery of DNA

Control and environmental DNA extractions were performed using the DNeasy PowerWater Kit (Qiagen) following the manufacturer's recommendations using 0.45 µm filters, as outlined in Figure IV B. Following the initial filtration step filter membranes from all environmental samples were lifted from opposite edges using sterile forceps and placed face down on appropriately labelled TSA or HEK plates to visualize growth. Membranes were lightly patted to ensure transfer of microorganisms; plates were incubated at 30°C or 37°C, respectively for 24 hours then stored at 4°C. At the conclusion of the protocol, extracted DNA from all samples were frozen at -20°C in 20 µl aliquots.

PCR Amplification

PCR amplification was performed using a DNA thermal cycler and Invitogen Platinum *Taq* DNA polymerase (ThermoFisher Scientific) following the manufacturer's suggested protocol for 25 µl reaction mixtures. Briefly, the PCR master mix was prepared using 2.5 µl 10x PCR reaction buffer, 0.75 µl 50 mM $MgCl_2$, 0.5 µl 10 mM dNTP mix, 0.1 µl platinum *Taq* polymerase, 1 µl isolated DNA, 0.5 µl of both forward and reverse primers, and RNase-free water to fill a total 25 µl mixture. Primer sequences and their respective gene and species targets are shown in Table II. In order to identify multiple species in a single run of the thermal cycler, annealing temperatures were identified between those reported in literature for the primers of interest in each trial. All PCR mixtures were subject to an initial denaturation at 94°C for 3 minutes. This was followed by 25 PCR cycles of the following conditions: denaturation at 94°C for 30 seconds; primer annealing at 50°C (*lacZ* and *lamB* gene primer pairs) and 56°C (*tuf, eaeA,* and *invA* gene primer pairs) for 30 seconds; and DNA extension at 72°C for 2 minutes. PCR products were stored at -20°C prior to analysis with gel electrophoresis.

Visualization of Products

PCR-amplified products were visualized using gel electrophoresis. 0.9% agarose gel was prepared by mixing 0.90 g of agarose with 100 ml of 1X TAE buffer and stained with 10 µl SYBR DNA gel safe stain. PCR reaction mixtures were thawed and 4.5 µl of 6x loading dye were added to each. 12 µl of each mixture were loaded into the wells and the gel was run in 1X TAE buffer at 200 V and 110 mA for approximately 45 minutes. Final gels were visualized under UV transilluminator and photographed.

Gene Target	Primer Name	Sequence (5` to 3`)	Target Species	Annealing Temp (°C)	T_m (°C)	Extension Time	Product Length (bp)	Reference
lacZ	ZL-1675	ATGAAAGCTGGCTACAGGAAGGCC	*Coliforms*	50.0	64.1	1-2 min	264	Bej et al., 1990
	ZR-2025	GGTTTATGCAGCAACGAGACGTCA			59.6			
lamB	BL-4910	CTGATCGAATGGCTGCCAGGCTCC	*E. coli, Shigella,* & *Salmonella*	50.0	64.0	1-2 min	309	Bej et al., 1990
	BR-5219	CAACCAGACGATAGTTATCACGCA			56.8			
tuf	TEcol553	TGGGAAGCGAAAATCCTG	*E. coli* & *Shigella*	58.0	52.6	30 sec	258	Maheux et al., 2009
	TEcol754	CAGTACAGGTAGACTTCTG			48.9			
invA	InvA F	AGCGTACTGGAAAGGGAAAG	*Salmonella*	60.0	54.7	Not in literature	115	Kasturi, et al., 2017
	InvA R	ATACCGCCAATAAAGTTCACAAAG			53.9			
eaeA	ConceaeA F	GACCCGGCACAAGCATAAGC	Pathogenic *E. coli*	65.0	59.2	45 sec	384	Barak et al., 2005
	ConceaeA R	CCACCTGCAGCAACAAGAGG			58.9			
SLT-I	SLT-I F	TGTAACTGGAAAGGTGGAGTATAC	*E. coli* O157:H7	55.5	53.9	30 sec	210	Imitaz et al., 2013
	SLT-I R	GCTATTCTGAGTCAACGAAAAATAAC			52.7			

Table II: Comparison of species-specific primers, gene and species targets, and sequence details gathered from literature.

RESULTS AND DISCUSSION

Preliminary research focused on studying and comparing methodologies for nucleic-acid extraction and purification. Because the most popular and efficient techniques utilize commercial extraction kits, the most common were investigated to determine which would yield the best results with consideration of cost, ease, and ability to extract both DNA and RNA from water samples over an extended period of time. Unfortunately, from the final list depicted in Table I, it was concluded that no single kit allows for the simultaneous extraction of DNA and RNA from water samples in a cost-effective manner. The two most viable options included the AllPrep DNA/RNA Kit (Qiagen) and the DNeasy PowerWater Kit (Qiagen). The AllPrep Kit allows for the simultaneous extraction of both DNA and RNA from samples which proves cost-effective and increases the ease of the experiment (Table I). However, it is designed to isolate material from cell and tissue samples, not water.

Alternatively, the PowerWater Kit is optimized to increase yields from low biomass samples and presents an easy-to-follow protocol which allows for the isolation of large quantities of high-quality DNA from samples, though it does not allow for the simultaneous extraction of both DNA and RNA (Table I). As such, it was determined that the best long-term option would be a novel extraction technique which combines the protocols of the AllPrep DNA/RNA and DNeasy PowerWater Kits to meet all of the extraction criteria for this experiment. With consideration of the protocols for each kit (Figure II), it was planned to adapt the AllPrep Kit to incorporate the PowerWater bead-beating tube to break apart the filter and sample prior to extraction using the AllPrep protocol, thus effectively extracting both DNA and RNA from water samples.

Spiked Samples

Initial tests focused on perfecting a single nucleic acid extraction kit prior to experimenting with a novel technique. As such, the first experiment tested the efficacy of the DNeasy PowerWater Kit using deionized water spiked with *Escherichia coli* HB101 K12 in a blind test. One sample was spiked with *E. coli* while the other was left unaltered. The extraction was performed following the manufacturer's instructions and the results were visualized using gel electrophoresis (Figure VII). Only lanes filled with DNA isolated from Sample A showed fragments, which indicates that Sample A was

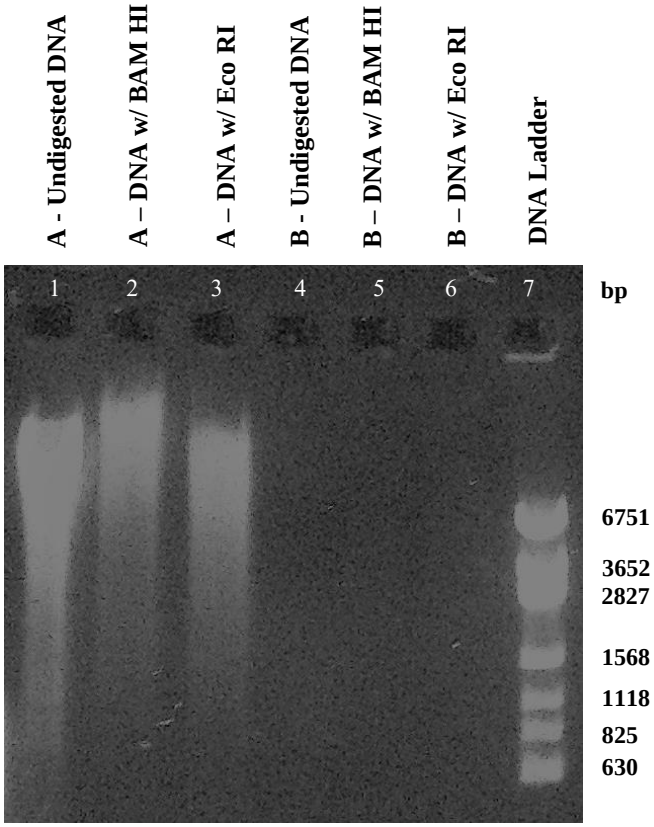

Figure VII: Agarose gel electrophoretic analysis of DNA isolated from water samples spiked with HB101 *E. coli* in a blind test. DNA was extracted from two samples, A (Lanes 1-3) and B (Lanes 4-6), following the DNeasy PowerWater Kit protocol. Restriction enzyme digests were performed to help visualize DNA. Lane 7 shows the DNA ladder.

spiked with such a high concentration of *E. coli* HB01 K-12 that the gel was unable to separate out the fragments, resulting in single a blurry band for both digested and undigested DNA (Lanes 1-3). Lanes with material from Sample B showed no DNA fragments, confirming that Sample B was not spiked with *E. coli*. These results confirm that the DNeasy PowerWater Kit effectively extracts DNA from high concentrations and support the decision to incorporate this protocol into the novel extraction technique.

While the spiked sample contains a very high concentration of bacteria, environmental samples are likely much more dilute, and thus require amplification through polymerase chain reactions prior to downstream applications. The literature investigation identified six well-researched genes of interest which can be used to test for the presence of three genera within the Enterobacteriaceae family: *Escherichia, Shigella,* and *Salmonella* in water samples (Figure VI). The selected primer pairs were: ZL-1675 and ZR-2025 (Bej et al., 1990), used to target the *lacZ* gene and detect total coliform bacteria; BL-4910 and BR-5219 (Bej et al., 1990), used to target the *lamB* gene and detect *Salmonella, Escherichia,* and *Shigella*; TEcol553 and TEcol754 (Maheux et al., 2009), used to target the *tuf* gene and detect *Escherichia* and *Shigella*; InvA F and InvA R (Kasturi, et al., 2017), used to target the *invA* gene and detect *Salmonella enterica*; ConceaeA F and ConceaeA R (Barak et al., 2005), used to target the *eaeA* gene and detect pathogenic *E. coli*; and SLT-I F and SLT-I FR (Imitaz et al., 2013), used to target the *SLT-I* gene and detect the pathogenic serogroup of *E. coli* O157:H7 (Table II).

To test the effectiveness of these primer sets a PCR was performed on the DNA isolated from Sample A in the blind *E. coli* study, and products were visualized through gel electrophoresis (Figure VIII). Sample B was included as a control. As anticipated, no

PCR products were observed from material extracted from Sample B (Lanes 2-12). Two fragments less than 500 bp were noted from Sample A amplified using the ZL-1675 and ZR-2025, and BL-4910 and BR-5219 primer pairs (Lanes 2-3), consistent with the expected 264 and 309 bp lengths (Table II). The PCR products in Lane 2 demonstrate that the *lacZ* gene was successfully targeted and confirms that the water was spiked with a coliform bacterium. Likewise, the amplified DNA in Lane 3 demonstrates that the *lamB* gene was also successfully targeted and narrows the options of potential species to a

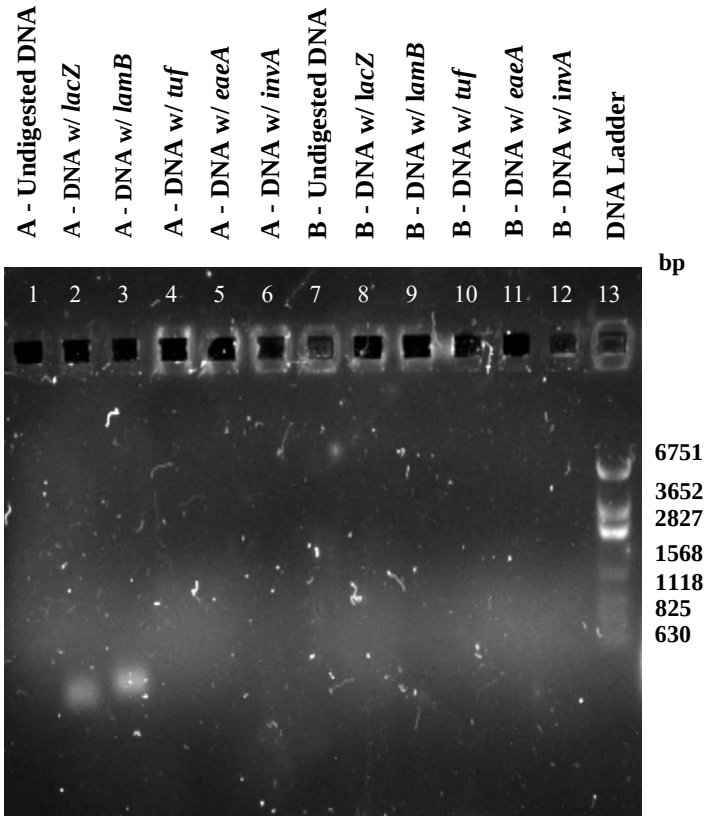

Figure VIII: Agarose gel electrophoretic analysis of PCR-amplified DNA products from bacterial DNA of HB101 K12 *E. coli* spiked samples using primers targeting *lacZ, lamB, tuf, eaeA,* and *invA* genes and undigested DNA for comparison. Trials in Lanes 1-6 were completed using DNA from Sample A. Lanes 7-12 were completed using Sample B for comparison. Lane 13 shows the DNA ladder.

member of the genera *Salmonella, Escherichia,* or *Shigella*. This agrees with the known target species of *Escherichia coli* HB101 K12. With this in mind, the *tuf* gene should have also been targeted by the TEcol553 and TEcol754 primers, thus showing positive detection of *Escherichia* and *Shigella*. Contrary to this expectation, no bands appeared in Lane 4, showing that no DNA was amplified using primer pair targeting the *tuf* gene. Additionally, no DNA was amplified using primer pairs targeting the *eaeA* or *invA* genes in Sample A (Lanes 5-6); this was expected given the ability of these genes identify *Salmonella enterica* and pathogenic *E. coli*, neither of which describe the known target species. It is important to consider that, though these genes are not found in *E. coli* HB101 K12, it can only be assumed based on literature that this is the reason bands were not observed; in order to confirm that this is the true cause the primers would need to be validated through other spiked samples with known concentrations of the appropriate bacteria to guarantee a positive test. At this time such tests cannot be performed as this is a BSL1 teaching facility which lacks access to pathogenic species for PCR validation.

In order to identify multiple species in a single run of the thermal cycler, annealing temperatures were identified between those reported in literature for the primers of interest in each trial. The annealing temperature tested for TEcol553 and TEcol754 was 56°C, though the reported ideal annealing temperature for this primer pair was 58°C (Maheux et al., 2009). This presents one factor that can influence the specificity and sensitivity of DNA amplification. If the annealing temperature is too high, primers fail to bind to the template, and if it is too low, the primers may bind non-specifically to the template (Bio-Rad Laboratories, 2020; Sipos et al., 2007). The annealing temperature should be relatively close to the melting temperature (T_m) of the

primers. The T_m for TEcol553 and TEcol754 is approximately 52.6°C and 48.9°C, respectively (Table II). Based on this information, the PCRs of DNA isolated from water spiked with *Escherichia coli* HB101 K12 using *tuf* gene primers were repeated using temperatures surrounding the ideal temperature and the melting temperatures of the primers. Results were visualized through gel electrophoresis (Figure IX). No PCR

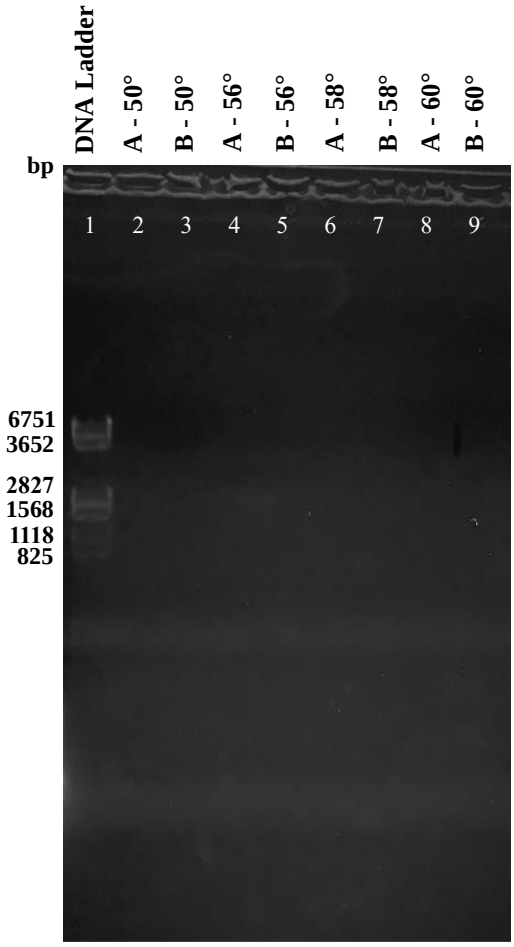

Figure IX: Agarose gel electrophoretic analysis of PCR-amplified DNA products from bacterial DNA of HB101 K-12 *E. coli* spiked samples using primers targeting the *tuf* gene. Trials tested annealing temperatures surrounding the reported ideal value (58°C) and the T_m for forward and reverse primers (52.6°C and 48.9°C, respectively). Trials in even lanes were completed using DNA from Sample A. Trials in odd lanes were completed using Sample B for comparison. Lane 1 shows the DNA ladder.

products were found for samples run with an annealing temperature of 50°C (Lanes 2-3), 58°C (Lanes 6-7), 60°C (Lanes 8-9) nor for a repeat trial using 56°C (Lanes 4-5). These results indicate that the annealing temperature likely was not the factor impeding detection of *Escherichia* and *Shigella*. Other variables that impact DNA amplification through polymerase chain reaction include primer concentration, and Mg^{2+} concentration, and the type of DNA polymerase used (Imtiaz, 2013). Future studies will examine these factors to determine if any impeded the ability of TEcol553 and TEcol754 primers to target the *tuf* gene and correctly identify *Escherichia* and *Shigella* as the target species.

Environmental Studies

Having confirmed the efficacy of the DNeasy PowerWater Kit, as well as the accuracy of the primer sets, the method could then be used to extract and analyze nucleic acids from environmental samples. Freshwater samples were gathered in mid-September 2018 and early November 2019 from three sites in the Whippany River. Effluent samples were collected from an output pipe running from the Morristown Wastewater Treatment Plant into the river, while the remaining samples were collected 50 feet upstream and downstream from the source. Nucleic acids were extracted from all samples using the DNeasy PowerWater Kit. Following filtration, filter membranes were inverted on TSA plates and bacterial growth was observed. After incubation at 30°C for 24 hours, it can be seen that both upstream and source samples collected in September 2018 showed bacterial colonies that were too numerous to count (Figure X). The same process was repeated for samples collected in December of 2019 (Figure XI). Again, colonies from upstream, source, and downstream samples are too numerous to count. Population size and diversity do not appear to differ between the three locations from both collection

dates, thus justifying the use of polymerase chain reactions to gain insight on variations between microbial populations.

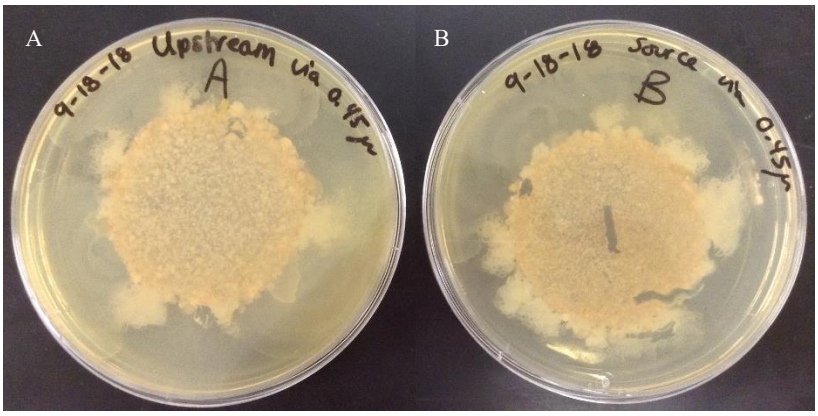

Figure X: Microbial growth 24 hours after filter membranes from environmental samples were inverted on TSA plates. Samples were collected Sept. 2018 from the (A) upstream site and (B) effluent source.

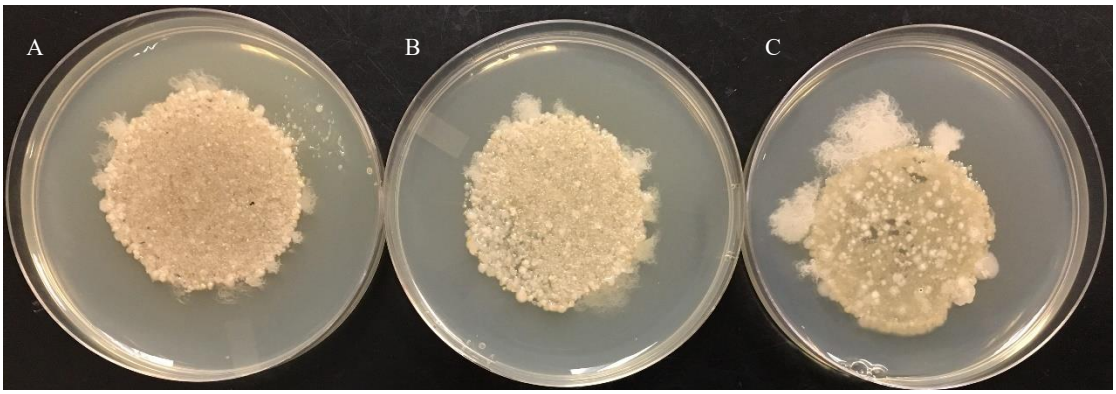

Figure XI: Microbial growth 24 hours after filter membranes from environmental samples were inverted on TSA plates. Samples were collected Nov. 2019 from the (A) upstream site, (B) effluent source, and (C) downstream site.

In addition to inversion on TSA plates, the downstream sample collected in December 2019 was also tested on Hektoen enteric agar (HEK), a selective culture medium used to detect the presence of *Salmonella* and *Shigella* (Figure XII). HEK contains bile salts which inhibit the growth of Gram-positive microorganisms. Because of their outer membrane, Gram-negative bacteria are resistant to the effects of the salts and thus will grow on the plate (King & Metzger, 1968). Agar is prepared using a lactose substrate along with bromothymol blue and acid fuchsin as indicators of fermentation; acid biproducts react with the indicator and produce orange colonies (King & Metzger, 1968). The medium also contains ferric ammonium citrate and sodium thiosulfate which react in the presence of hydrogen sulfide produced by non-lactose fermenting bacteria and create a visible black dot in the center of the colony (Dickinson, 2013). Given this information, the downstream plate showed a potential *Salmonella enterica* colony, as indicated by the arrow in Figure XII A. However, after propagation on a clean HEK plate (Figure XII B), the colony was confirmed not to be *Salmonella* as the agar turned orange, indicating a Gram-negative lactose-fermenting bacterial species.

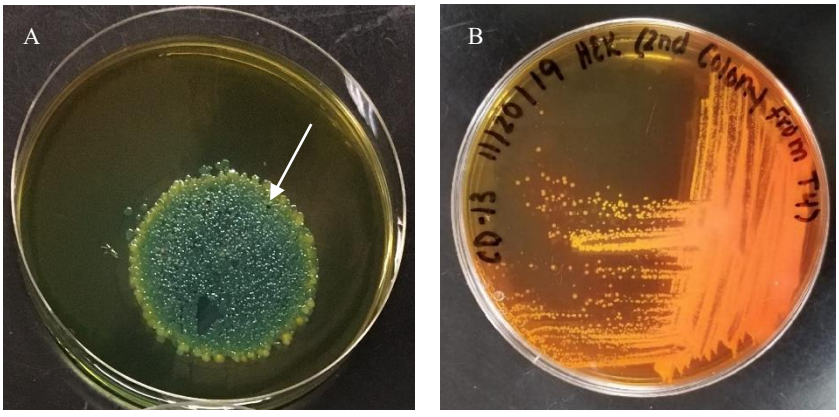

Figure XII: (A) Microbial growth 24 hours after the downstream filter membrane from the Nov. 2019 environmental sample was inverted on a HEK plate**.** Arrow indicates suspected *Salmonella enterica* colony. (B) Propagation of suspected *S. enterica*.

Having confirmed the presence of microorganisms on the petri dishes, polymerase chain reactions were run to amplify the DNA of the microorganisms using the previously tested primers, and results were visualized using gel electrophoresis (Figure XIII). The *Escherichia coli* HB101 K12 spiked water (Sample A) from the earlier trails was used as a positive control for all primers (Lanes 2, 7, 12, 17, and 22). Upstream and effluent samples from September 2018 (Lanes 3-4, 8-9, 13-14, 18-19, and 23-24) and November 2019 (in Lanes 5-6, 10-11, 15-16, 20-21, and 26-27) were also tested. Two fragments less than 500 bp were noted using the primer pair ZL-1675 and ZR-2025 (Lane 2, circled) and BL-4910 and BR-5219 (Lane 7, circled). Both observed fragments were consistent with the expected 264 and 309 bp lengths (Table II). These amplicons were expected as these lanes contained DNA amplified from the *E. coli* spiked positive control, confirming the ability of the primer pairs to target the *lacZ* and *lamB* genes and identify the spiked target species as a coliform, and *E. coli, Shigella,* or *Salmonella*, respectively.

Again, the *tuf* primers failed to amplify any DNA from the positive control, contrary to what was expected, thus further optimization of the TEcol553 and TEcol754 primer set is necessary. Lastly, no environmental samples collected from upstream, downstream, or directly from the effluent source showed PCR products using any of the primers (Lanes 3-6, 8-11, 13-16, 18-21, 23-26). Based on the colonial growth observed on the agars (Figures X-XII), this is likely due to the low biomass of water samples, rather than a lack of microbes in the samples or the sensitivity of the primers. Moving forward a limit of detection for the PCR should be investigated to determine the minimum concentration required to visualize a PCR product through gel electrophoresis.

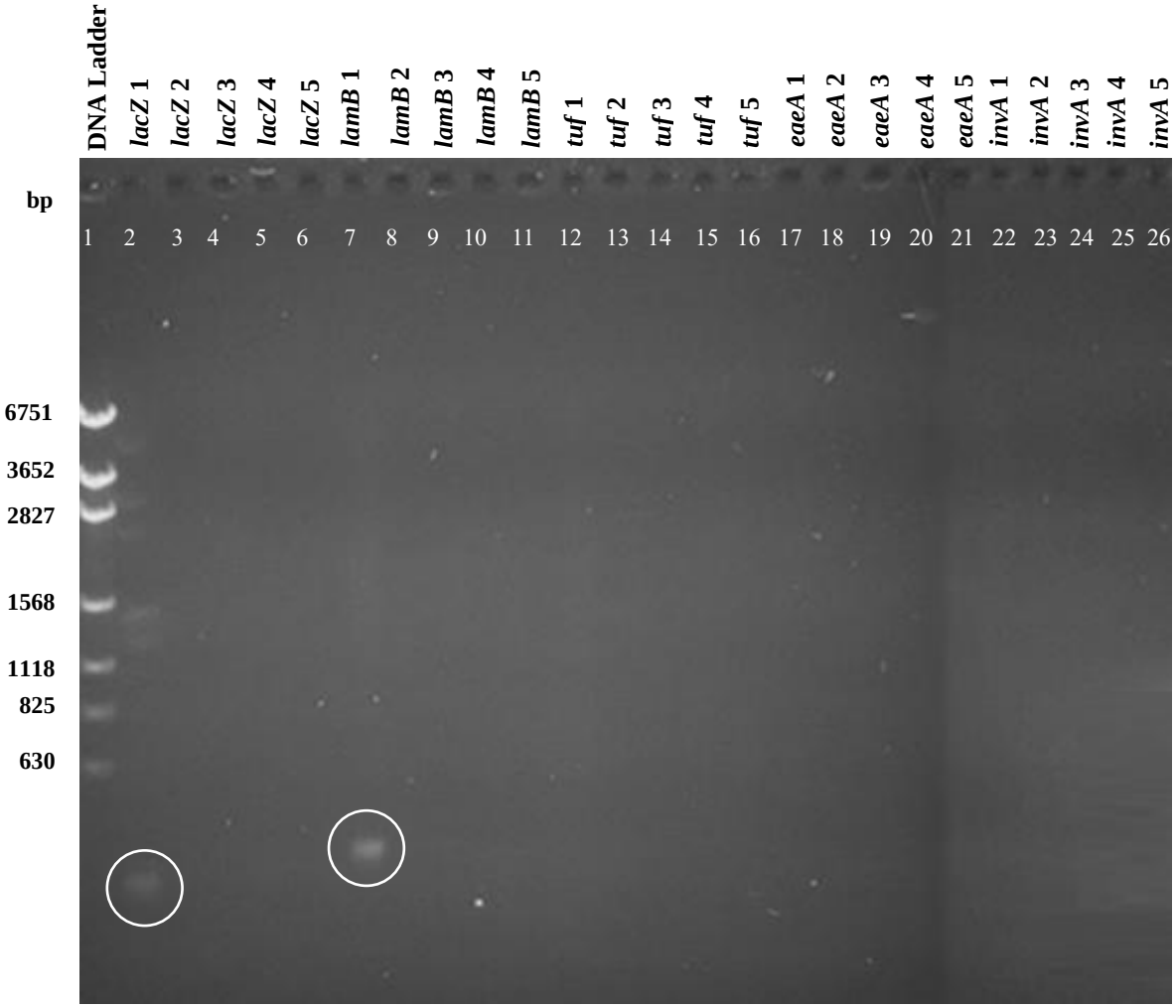

Figure XIII: Agarose gel electrophoretic analysis of PCR-amplified DNA products from bacterial DNA of control and environmental samples using primers targeting *lacZ, lamB, tuf, eaeA,* and *InvA*. The numbers listed after the gene target define the origin of the sample. Trials in Lanes 2, 7, 12, 17, and 22 (labeled 1) were completed using confirmed HB101 *E. coli* from Sample A of the spiked trial, included as a positive control. Trials in Lanes 3, 8, 13, 18, and 23 (labeled 2) were completed using the environmental upstream sample collected Sept. 2018. Trials in Lanes 4, 9, 14, 19, and 24 (labeled 3) were completed using the environmental effluent sample collected Sept. 2018. Trials in Lanes 5, 10, 15, 20, and 25 (labeled 4) were completed using the environmental upstream sample collected Nov. 2019. Trials in Lanes 6, 11, 16, 21, and 26 (labeled 5) were completed using the environmental effluent sample collected Nov. 2019. Lane 1 shows the DNA ladder. Circles indicate PCR products.

FUTURE STUDIES

Ultimately this research confirmed the ease and efficacy of the DNeasy PowerWater Kit (Qiagen) using spiked and environmental samples. Because of the low biomass of water samples, it was concluded that polymerase chain reactions need to be performed to amplify nucleic acids and determine differences in bacterial diversity. Six primer sets were identified to target genes found in Enterobacteriaceae genera *Escherichia, Shigella,* and *Salmonella*. The ZL-1675 and ZR-2025 primer pair was confirmed to target the *lacZ* gene and detect total coliform bacteria, and the BL-4910 and BR-5219 primer pair was confirmed target the *lamB* gene and detect *Salmonella, Escherichia,* and *Shigella* using water spiked with *Escherichia coli* HB101 K12. Contrary to what was expected, the primer pair TEcol553 and TEcol754, used to target the *tuf* gene and detect *Escherichia* and *Shigella*, failed to show a PCR amplification in the initial spiked test. Further trials using different annealing temperatures confirmed that this was not the cause. As expressed earlier, other factors including primer concentration, Mg^{2+} concentration, number of PCR cycles, and the type of polymerase used may impact the specificity and sensitivity of the PCR (Imtiaz et al., 2013). For this reason, future studies should analyze these variables to optimize the *tuf* gene primers.

Additionally, other primer pairs may be investigated detect *Escherichia* and *Shigella*. Bej et al. (1991) found that bacteria associated with fecal contamination in water can be detected through PCR amplification using primers targeting the *uidA* and *uidR* genes which encode β-glucuronidase, an enzyme which catalyzes reactions and is involved in carbohydrate and protein binding (Gene Blast, 2020). Similarly, Riyaz-Ul-Hassan et al. (2009) also found that the *uidA* gene can be targeted for detection for

Escherichia and *Shigella* in milk samples which are compositionally similar to water samples. Though Maheux et al. (2009) found primers targeting the *uid* gene to be less selective than those targeting *tuf*, other studies provide enough support to justify further investigation of both primers on spiked and environmental samples.

Results of this research also highlighted the difficulty of identifying microbial species from low biomass samples. Moving forward tests should be performed to estimate the concentration of the environmental samples and to establish a limit of detection for the PCR. One potential technique includes multiple tube fermentation (MTF) to detect the presence of coliform bacteria in a sample and estimate their numbers. Through the presumptive test of this method serial dilutions of sampled water are placed in lauryl-sulfate-tryptose-lactose broth. The presence of coliforms (lactose-fermenters) can be noted by a color change of the broth from red to yellow as well as through gas production seen in a Durham tube (Gerba, 2015; Nurliyana et al., 2018). Concentrations of coliforms can then be identified by counting the number of positive tubes at each dilution and comparting the values to an MPN index which calculates the most probable number (MPN) of coliform bacteria per 100 ml of solution. Following the presumptive test, the confirming test of MTF can then verify the presence of coliforms by inoculating bacteria from positive tubes onto selective bacterial agars and observing growth, similar to the HEK plates observed in earlier trials of this experiment (Gerba, 2015). Another method used to estimate the total number of culturable bacteria in water samples is through plate count enumeration in which samples of microbes are diluted, placed on an agar media, and incubated overnight. Colonies that form can then be counted and the total population size can be calculated using the known dilution factor (Nurliyana et al., 2018).

After establishing the limit of detection of the PCR, other studies can begin to test the AllPrep DNA/RNA Kit (Qiagen) on cell or tissue samples to become familiar with the protocol. Once optimized, the PowerWater bead-beating technology can be incorporated into the AllPrep Kit to test the efficacy of the novel DNA/RNA extraction method using water samples. Each of these steps will progress this research towards its objective to monitor the presence of Enterobacteriaceae in local sources of surface water fed by wastewater effluent, and to map out fluctuations in diversity due to climate change and anthropogenic effects. Ultimately, partnering projects will examine other types of waterborne species, such as Gram-positive bacteria, archaea, and viruses, thus forming connections between the vast diversity of aquatic microorganisms. In conclusion, not only will this research help identify potentially harmful microbes found in the water and indicate potential pathogenic outbreaks before they cause mass illness, it also has the ability to increase environmental awareness and can empower society to take necessary actions to improve the quality of water worldwide.

REFERENCES

Abdulkadir, M., & Waliyu, S. 2012. Screening and Isolation of the Soil Bacteria for Ability to Produce Antibiotics. *European Journal of Applied Sciences 4*(5): 211-215.

Balian, E. V., Segers, H., Lévèque, C., & Martens, K. 2007. The Freshwater Animal Diversity Assessment: An Overview of the Results. *Hydrobiologia, 595*(1), 627–637.

Banerjee. S. 2020. Polymerase Chain Reaction (PCR). Retrieved from https://thebiologynotes.com /polymerase-chain-reaction-pcr/

Barak, J. D., Sananikone, K., & Delwiche, M. J. 2005. Comparison of Primers for the Detection of Pathogenic *Escherichia coli* using Real-Time PCR. *Letters in Applied Microbiology, 41*(2), 112-118.

Batt, C. A. 2016. Microorganisms. In: *Reference Module in Food Science.* Retrieved from https://www.sciencedirect.com/science/article/pii/B9780081005965034363

Bej, A. K., DiCesare, J. L., Haff, L., & Atlas, R. M. 1991. Detection of *Escherichia coli* and *Shigella* spp. in Water by Using the Polymerase Chain Reaction and Gene Probes for *uid*. *Applied and Environmental Microbiology, 57*(4), 1013-1017.

Bej, A. K., Steffan, R. J., DiCesare, J. L., Haff, L., & Atlas, R. M. 1990. Detection of Coliform Bacteria in Water by Polymerase Chain Reaction and Gene Probes. *Applied and Environmental Microbiology, 56*(2), 307-314.

Berg, E. G. 2015. A New Spin on the Old Gram Stain. *C&EN.* Retrieved from https://cen.acs.org/articles/93/web/2015/04/New-Spin-Old-Gram-Stain.html

Bio-Rad Laboratories. 2020. PCR Troubleshooting. Retrieved from https://www.bio-

rad.com/en-us/applications-technologies/pcr-troubleshooting?ID=LUSO3HC4S.

Brenner, D. J., Fanning, G. R., Steigerwalt, A. G., Orskov I, & Orskov F. 1972. Polynucleotide Sequence Relatedness among Three Groups of Pathogenic *Escherichia coli* Strains. *Infection and Immunity*. *6*(3), 308-315.

Brenner, F. W., Villar, R. G., Angulo, F. J., Tauxe, R., & Swaminathan, B. 2000. Salmonella Nomenclature. *Journal of Clinical Microbiology, 38*(7), 2465-2467.

Bush, L. M. & Perez, M. T. 2020. Nontyphoidal Salmonella Infections. *Merck Manual: Professional Version.* Retrieved from https://www.merckmanuals.com /professional/infectious-diseases/gram-negative-bacilli/nontyphoidal-salmonella-infections?query=salmonella.

Cabral, J. P. 2010. Water Microbiology. Bacterial Pathogens and Water. *International Journal of Environmental Research and Public Health, 7*(10), 3657-3703.

Canpolat N. (2015). Hemolytic Uremic Syndrome. *Turk pediatri arsivi*, *50*(2), 73–82.

Cohen M. B., & Giannella, R. A. 1992. Hemorrhagic Colitis associated with *Escherichia coli* O157:H7. *Advances in Internal Medicine*. *37*, 173-195.

Collen, B., Whitton, F., Dyer, E. E., Baillie, J. E. M., Cumberlidge, N., Darwall, W. R. T., Pollock, C., Richman, N. I., Soulsby, A., -M., & Böhm, M. 2013. Global Patterns of Freshwater Species Diversity, Threat and Endemism. *Global Ecology and Biogeography, 23*(1), 40–51.

Devanga Ragupathi, N. K., Muthuirulandi Sethuvel, D, P, Inbanathan, F.Y., Veeraraghavan, B. 2017. Accurate Differentiation of *Escherichia coli* and *Shigella* Serogroups: Challenges and Strategies. *New Microbes New Infections*. *21*, 58-62.

Dickinson, B. 2013. Instructions for Use: Ready-to-Use Plated Media. BD Hektoen

Enteric Agar (HE Agar). Retrieved from https://www.bd.com /resource.aspx?IDX=8970.

Dutta, H., & Dutta, A. 2016. The Microbial Aspect of Climate Change. *Energy, Ecology and Environment, 1*(4), 209-232.

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. 2008. Species Detection using Environmental DNA from Water Samples. *Biology Letters, 4*(4), 423-425.

Finlay B.J. & Esteban G.F. 1998. Freshwater Protozoa: Biodiversity and Ecological Function. *Biodiversity and Conservation* 7, 1163–1186.

Gene Blast [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2020. Retrieved from https://blast.ncbi.nlm.nih.gov/Blast.cgi

Gerba, C. P. 2015. Indicator Microorganisms. In: *Environmental Microbiology*, 3rd ed. London: Academic Press. pp. 485-499.

Giannella, R.A. 1996. Salmonella. In: Baron, S., editor. *Medical Microbiology*, 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston. Chapter 21.

Giovannoni, S. J., DeLong, E. F., Schmidt, T. M., & Pace, N. R. 1990. Tangential Flow Filtration and Preliminary Phylogenetic Analysis of Marine Picoplankton. *Applied and Environmental Microbiology*, *56*(8), 2572–2575.

Guillén-Navarro, K., Herrera-López, D., López-Chávez, M. Y., Cancino-Gómez, M., & Reyes-Reyes, A. L. 2015. Assessment of Methods to Recover DNA from Bacteria, Fungi and Archaea in Complex Environmental Samples. *Folia Microbiologica, 60*(6), 551-558.

Hale, T. L. & Keusch, G. T. 1996. Shigella. In: Baron, S., editor. *Medical Microbiology*,

4th ed. Galveston (TX): University of Texas Medical Branch at Galveston. Chapter 22.

Haque, S. Z., & Haque, M. 2017. The Ecological Community of Commensal, Symbiotic, and Pathogenic Gastrointestinal Microorganisms - An Appraisal. *Clinical and Experimental Gastroenterology*, *10*, 91–103.

Higgins, J. A., Jenkins, M. C., Shelton, D. R., Fayer, R., & Karns, J. S. 2001. Rapid Extraction of DNA From *Escherichia coli* and *Cryptosporidium parvum* for Use in PCR. *Applied and Environmental Microbiology, 67*(11), 5321-5324.

Imtiaz, J., Hashmi, I., Saeed, A., Qazi, I. A., & Arshad, M. 2013. Development of PCR Protocol for Detection of *Escherichia coli* in Drinking Water. *Water Resources Management VII, 171*, 225-234.

Kaiser, G. E. 2018. Classification: The Three Domain System. Retrieved from http://faculty.ccbcmd.edu/courses/bio141/lecguide/unit1/3domain/3domain.html

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

Kasturi, K. N., & Drgon, T. 2017. Real-Time PCR Method for Detection of *Salmonella* spp. in Environmental Samples. *Applied and Environmental Microbiology, 83*(14).

Kim, H. S., & Gilliland, S. E. (1983). *Lactobacillus acidophilus* as a Dietary Adjunct for Milk to Aid Lactose Digestion in Humans. *Journal of Dairy Science, 66*(5), 959–966.

King, S. & Metzger, W. I. 1968. A New Plating Medium for the Isolation of Enteric Pathogens. *Applied Microbiology, 16*(4), 577-578.

Lehtovirta-Morley, L. E. (2018). Ammonia Oxidation: Ecology, Physiology, Biochemistry and Why They Must All Come Together. *FEMS Microbiology Letters, 365*(9).

Long, S. L., Gahan, C. G. M., and Joyce, S. A. 2017. Interactions Between Gut Bacteria and Bile in Health and Disease. *Molecular Aspects of Medicine 56*, 54–65.

Ma, Y., Liu, F., Kong, Z., Yin, J., Kou, W., Wu, L., & Ge, G. 2016. The Distribution Pattern of Sediment Archaea Community of the Poyang Lake, the Largest Freshwater Lake in China. *Archaea (Vancouver, B.C.)*, *2016*, 9278929.

Maheux, A. F., Bissonnette, L., Boissinot, M., Bernier, J.-L. T., Huppé, V., Picard, F. J., Bérubé, È., & Bergeron, M. G. 2011. Rapid Concentration and Molecular Enrichment Approach for Sensitive Detection of Escherichia coli and Shigella Species in Potable Water Samples. *Applied and Environmental Microbiology, 77*(17), 6199-6207.

Maheux, A. F., Picard, F. J., Boissinot, M., Bissonnette, L., Paradis, S., & Bergeron, M. G. 2009. Analytical Comparison of Nine PCR Primer Sets Designed to Detect the Presence of *Escherichia coli/Shigella* in Water Samples. *Water Research, 43*(12), 3019-3028.

Mannanov, R. N., & Sattarova, R. K. 2001. *Chemistry of Natural Compounds, 37*(2), 117–123.

Mead, P. S. & Griffin, P. M. 1998. *Escherichia coli* O157:H7. *Lancet*. *352*(9135), 1207-1212.

Moore, M. V., De Stasio, B. T., Huizenga, K. N., & Silow, E. A. 2018. Trophic Coupling

of the Microbial and the Classical Food Web in Lake Baikal, Siberia. *Freshwater Biology.*

Moré, M. I., Herrick, J. B., Silva, M. C., Ghiorse, W. C., & Madsen, E. L. 1994. Quantitative Cell Lysis of Indigenous Microorganisms and Rapid Extraction of Microbial DNA from Sediment. *Applied and Environmental Microbiology*, *60*(5), 1572-1580.

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. 2011. A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiology and Molecular Biology Reviews, 75*(1), 14-49.

Nixon S.W. 1981. Remineralization and Nutrient Cycling in Coastal Marine Ecosystems. In: Neilson B.J., Cronin L.E., editors. *Estuaries and Nutrients. Contemporary Issues in Science and Society.* Totowa (NJ): Humana Press. pp. 111-138.

Nurliyana, M. R., Sahdan, M. Z., Wibowo, K. M., Muslihati, A., Saim, H., Ahmad, S. A., Sari, Y., & Mansor, Z. 2018. The Detection Method of Escherichia coli in Water Resources: A Review. *Journal of Physics: Conference Series, 995,* 012065.

Octavia S., & Lan R. 2014. The Family *Enterobacteriaceae*. In: Rosenberg E., DeLong E.F., Lory S., Stackebrandt E., Thompson F., editors. *The Prokaryotes*. Berlin, Heidelberg: Springer. pp. 225-286.

Polymerase Chain Reaction. [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2017. Retrieved from https://www.ncbi.nlm.nih.gov/probe/docs/techpcr/.

Purdy, K. J. 2005. Nucleic Acid Recovery from Complex Environmental Samples. *Methods in Enzymology*, *397*, 271-292.

Qiagen. 2005. AllPrep DNA/RNA Mini Handbook. Retrieved from https://www.qiagen.com/us/resources/resourcedetail?id=bbd50261-3b80-4657-ad58-6a5a97b88821&lang=en

Qiagen. 2017. DNeasy PowerWater Kit Handbook. Retrieved from https://www.qiagen.com/us/resources/resourcedetail?id=75765ef9-2a6f-4f5d-a36b-dbd9beb43079&lang=en

Raaijmakers, J. M., & Mazzola, M. 2012. Diversity and Natural Functions of Antibiotics Produced by Beneficial and Plant Pathogenic Bacteria. *Annual Review of Phytopathology, 50*(1), 403–424.

Ritchie, H. & Roser, M. 2019. Clean Water. *Our World in Data*. Retrieved from https://ourworldindata.org/water-access

Riyaz-Ul-Hassan, S., Syed, S., Johri, S., Verma, V., & Qazi, G. N. (2009). Application of a Multiplex PCR Assay for the Detection of *Shigella, Escherichia coli* and Shiga Toxin-Producing *Esch. coli* in Milk. *Journal of Dairy Research, 76*(2), 188.

Sanderman, J., & Amundson, R. 2003. Biogeochemistry of Decomposition and Detrital Processing. *Treatise on Geochemistry,* 249–316.

Shearer, C. A., Descals, E., Kohlmeyer, B., Kohlmeyer, J., Marvanová, L., Padgett, D., Porter, D., Huzefa, R., Schmit, D., Thorton, H., & Voglymayr, H. 2007. Fungal Biodiversity in Aquatic Habitats. *Biodiversity and Conservation, 16*(1), 49–67.

Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialigeti, K., & Nikolausz, M. 2007. Effect of Primer Mismatch, Annealing Temperature and PCR Cycle Number on 16S rRNA Gene-Targetting Bacterial Community Analysis. *FEMS Microbiology Ecology, 60*(2), 341-350.

Sizar, O. & Unakal, C. G. Gram Positive Bacteria. 2019. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK470553/

Spring, S., Schulze, R., Overmann, J., & Schleifer, K. 2000. Identification and Characterization of Ecologically Significant Prokaryotes in the Sediment of Freshwater Lakes: Molecular and Cultivation Studies. *FEMS Microbiology Reviews*, *24*(5), 573–590.

Stams, A. J. M., & Plugge, C. M. (2009). Electron Transfer in Syntrophic Communities of Anaerobic Bacteria and Archaea. *Nature Reviews Microbiology, 7*(8), 568–577.

Tan, S. C., & Yiap, B. C. 2009. DNA, RNA, and Protein Extraction: The Past and The Present. *Journal of Biomedicine and Biotechnology, 2009*, 1–10.

Thauer, R. K., & Shima, S. (2008). Methane as Fuel for Anaerobic Microorganisms. *Annals of the New York Academy of Sciences, 1125*(1), 158–170.

Triant, D. A., & Whitehead, A. 2008. Simultaneous Extraction of High-Quality RNA and DNA from Small Tissue Samples. *Journal of Heredity, 100*(2), 246-250.

U.S. Environmental Protection Agency. 2012. Water: Monitoring & Assessment. Retrieved from https://archive.epa.gov/water/archive/web/html/vms511.html

Valentine, D. L. 2002. Biogeochemistry and Microbial Ecology of Methane Oxidation in Anoxic Environments: A Review *Antonie van Leeuwenhoek, 81*(1), 271–282.

Yang, G., Erdman, D. E., Kodani, M., Kools, J., Bowen, M. D., & Fields, B. S. 2010. Comparison of Commercial Systems for Extraction of Nucleic Acids from DNA/RNA Respiratory Pathogens. *Journal of Virological Methods, 171*(1), 195-199.

Yu, Z., & Mohn, W. W. 1999. Killing Two Birds with One Stone: Simultaneous Extraction of DNA and RNA from Activated Sludge Biomass. *Canadian Journal of Microbiology, 45*(3), 269-272.