SYSTEMATIC ANALYSIS OF ANTIBIOTIC RESISTANCE GENES (ARGs) IN WATER AND ENVIRONMENTAL SYSTEM

by

Sol Park

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Civil and Environmental Engineering

Charlotte

2021

Approved by:

_____
Dr. Mariya Munir

_____
Dr. Cynthia Gibas

_____
Dr. Way Sung

_____
Dr. Olya Keen

_____
Dr. Juan Vivero-Escoto

ABSTRACT

SOL PARK. Systematic Analysis of Antibiotic Resistance Genes (ARGs) in Water and Environmental System
(Under the direction of DR. MARIYA MUNIR)

The global research community is aware of the prevalence and ubiquity of antibiotic resistant genes (ARGs) which is increasing over time due to human activities (Allen et al. 2010; Baquero, Martínez, and Cantón 2008; Bergeron et al. 2016; Di Cesare et al. 2015; Huang et al. 2019; Knapp et al. 2010; Pruden et al. 2006; Rizzo et al. 2013; Tacconelli et al. 2018). The main issue of emerging antibiotic resistance (AR) is causing casual infections untreatable and can cause high socioeconomic costs as well as health care burden (CDC 2019; L. G. Li et al. 2020). According to the World Health Organization (WHO), we lack global surveillance, and this poses threat to not only human health but also has risks of system overloading, failure, or indirect hazards(Sanitation Safety Planning, Greywater and Excreta. 2016; WHO 2014). Studies have shown that ARGs can be present even in a pristine environment like the sea bed of the deep ocean (B. Chen et al. 2013). Others have suggested most of the antibiotic resistance and metal resistances are coming from environmental stress in the urban area (Medeiros et al. 2016).

My dissertation focuses on targeting ARGs abundance, distribution, and their fate in the water environment impacted by human activities. It consists of four sections: 1) Introduction on antibiotic resistance (AR) and methods of quantitative ARG analysis 2) Method comparison and optimization of qPCR and ddPCR methods for ARGs detection

3) ARGs and pathogen quantification in the water system under the influence of anthropogenic activities,

4) ARG tracking and metagenomic study of microbiome under the effects of Flue Gas Desulfurization (FGD) bioreactors

The first chapter focuses on the background of Antibiotic Resistance and describes the method development and optimization of ARG quantification using quantitative polymerase chain reaction (qPCR). DNA is extracted from an environmental sample that potentially has targeted genes and was screened for the presence of ARGs by PCR and gel electrophoresis. Once the target is identified, target genes were cloned into competent cells. The cells were cultured on antimicrobial plates, and once survived passing the blue/white screening, its grown culture DNA can serve as a standard for qPCR analysis. Standards were developed through this method were further used in the studies as described in subsequent chapters.

The second chapter focused on the following criteria for comparing both technologies, 1) efficiency, 2) range of detection and limitations under different disciplines and gene targets, (3) optimization, and (4) status on antibiotic resistance genes (ARGs) analysis in the literature. Then droplet digital polymerase chain reaction (ddPCR) is developed and optimized to detect ARGs in environmental samples and compare quantification and detection performances along with qPCR. The performance of ddPCR to qPCR was compared on ARGs (tetA, ereA) standard and samples, and the correlation between lab results and trend of ARGs reported in the literature was discussed.

The third chapter describes our study that shows how the antibiotic resistance genes abundance varies in three different water systems across the U.S. with various

types of water bodies such as lakes, rivers, underground recharge, and wetlands, particularly with discharge from the wastewater treatment plants (WWTP). The objective of the study is to 1) find effects of WWTP effluents on ARGs abundance in the water system, 2) investigate how ARGs concentration changes throughout different types of water bodies, and 3) find connections of other anthropogenic sources of ARGs inflow into the water system. The WWTP effluents were the major contributor of increased ARG concentration for most of the ARG analyzed except for qnrB. We found that specific types of water systems have a different distribution of ARGs: underground recharge contaminated with qnrA compared to others, the lake was detected to have higher sul3 and qnrB genes compared to others. This study has concluded and is at the writing phase for a journal paper submission

The fourth chapter describes the study on the microbiome of the Flue Gas Desulfurization (FGD) bioreactors, for their genetic functions on contaminant uptake from the FGD wastewater into biomass and the expression of ARGs across all sites. FGD is a process to remove toxic volatile substances from the coal combustion process of power plants, and this process wastewater as a byproduct. The wash water of the scrubbing tower is sourced from a nearby water source such as a river or lake, and this makes the presence of ARGs to be expected, because of the known ubiquity of ARGs even in pristine water. The wastewater created from the process also contains high concentrations of heavy metals including Se, Cd, Cr, Hg, As, Pb, Ni, Cu, etc. along with sulfates, chlorides, nitrates (Ma et al. 2019; Martini and Vadbunker 2016; W. Zhang et al. 2019). In the treatment of the contaminated FGD wastewater, two-stage bioreactors which can biologically harness sulfates and other heavy metals were applied with sludge

recycling with little or no waste. The microbiome that is present in the bioreactors showed various functional pathways to harness contaminants, as well as antibiotic resistance of different mechanisms. The co-occurrence of ARGs and metal resistance genes (MRGs) has been previously reported (Xiaomin Wang et al. 2021), and the result from this chapter helps understanding and containment of ARGs under stressed environments.

DEDICATION

This dissertation is dedicated to:

1) my parents who firmly believed and supported me and my decision to start my journey of being a graduate student, and
2) to my husband, Juyeong Choi, who has been sharing love and encouragement to follow my dream along this long-long way.

ACKNOWLEDGEMENTS

I would like to thank my Ph.D. advisor, Dr. Mariya Munir for her help and support throughout my study. It has been an extraordinary, exciting, and heartwarming learning experience thanks to her kindness, wisdom, and excitement shared throughout this journey. I thank Dr. Cynthia Gibas for her support and guidance in the microbiome study of bioreactors. Her insights helped me to run pipelines and further bioinformatic analysis. I thank Dr. Way Sung for his help and share of insights in comparison study on ddPCR with qPCR. He has guided me towards critical and statistically sound ddPCR optimization and approaches through multiple experiments. I thank Dr. Olya Keen for her guidance throughout the study of ARGs detection in the U.S. water system. Her help improved the data representation and analysis. I thank Dr. Juan Vivero-Escoto for his support in serving on my committee. I thank Dr. Mei Sun for her help and support throughout the NCDOT project of testing stormwater runoff of different roadside matrices even though the materials are not included in my dissertation. I thank all engineers and site workers of Duke Energy and various Water works sector who made the sampling possible for my dissertation. I would like to thank other members of my research group: Adeola Sorinolu, Mohammed Khalid, Ariful Islam Juel, and Nita Khanal. I also thank my kind, loving, and bright colleagues & friends including Anita Rana, Lauren Roppolo Brazell, Kevin Lambirth, Meghana Patel, Savannah & David Tilley, Cristine Quach, Xiaoyu Bai, and Tarini Shukla for their collaboration, encouragement, and help I needed both in and out of the lab/office environments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abt An Acronym for Antibiotic

AR An acronym for Antibiotic Resistance

AMR An acronym for Antibiotic Microbial Resistance

ARB An acronym for Antibiotic Resistant Bacteria

ARG An acronym for Antibiotic Resistance Gene

AWEF An acronym for Advanced Water Treatment Plant Effluent

dPCR An acronym for Digital PCR

ddPCR An acronym for Digital Droplet Polymerase Chain Reaction

DWIn An acronym for Drinking Water Treatment Plant Influent

DWTP An acronym for Drinking Water Treatment Plant

ED An acronym for Electrodialysis

FGD An acronym for Flue Gas Desulfurization

HUMAnN An acronym for HMP Unified Metabolic Analysis Network

MDS An acronym for Multi-Dimensional Scaling

MetaPhlAn An acronym for Metagenomic Phylogenetic Analysis

PCA / PCoA An acronym for Principal Coordinate Analysis

PCR An acronym for Polymerase Chain Reaction

PEAR An acronym for Paired-end Read Merger

PUR An acronym for Post Underground Recharge

qPCR An acronym for Quantitative Polymerase Chain Reaction

SE An acronym for Sampling Event

ShortBRED An acronym for Short, Better Representative Extract Dataset

TCA An acronym for Tricarboxylic Acid

UNCC An acronym for University of North Carolina at Charlotte

WEF An acronym for Wastewater Treatment Plant Effluent

WHO An acronym for World Health Organization

WWTP An acronym for Wastewater Treatment Plant

ZLD An acronym for Zero-liquid Discharge

ZVI An acronym for Zero-valent Iron

CHAPTER 1: Method Development and Optimization for Quantitative Analysis of Antibiotic Resistant Genes Using qPCR

1.1 Introduction

Studies are showing that antibiotic resistant bacteria (ARB) and genes (ARG) pose a threat to human health by inhibition of clinical treatment with antibiotics(Allen et al. 2010; Asfaw, Genetu, and Shenkute 2020; CDC 2019, 2020). According to USEPA Science inventory, recent findings showed ARG concentration is more related to watershed urbanization and agriculture than water chemistry metrics (Hill, R., S. Keely, N. Brinkman, E. Wheaton, S. Leibowitz, M. Jahne, R. Martin 2018). Specifically, their analysis showed sul 1, a type of sulfonamide resistant gene has a positive correlation between urbanization and agriculture. This relationship was higher than the association with turbidity, nutrient concentrations, or suspended solids (Hill, R., S. Keely, N. Brinkman, E. Wheaton, S. Leibowitz, M. Jahne, R. Martin 2018). The research focused on tetracycline and ampicillin-resistant bacteria and genes near the Hudson River estuary. They reported the highest ARG concentration at the nearshore with the positive relation of their fecal origins (Young, Juhl, and O'Mullan 2013). In 2010, the tetracycline-specific ARG abundance was 15 times more than the amount detected in the 1970s (Knapp et al. 2010). This represents that the amount of antibiotics is gradually increasing and leading to an increase in ARGs in the environment. According to the CDC report, the estimated cost of infections caused by antibiotic resistance (AR) reaches $20 to $50 billion or more per year in direct health care, causing negative externalities of up to $35 billion in lost productivity and 8 million additional days of patient's hospitalization (Executive Office of the President President's Council of Advisors on Science and Technology 2014). It is also reported that such AR is costly and hard to treat because ARBs with the resistance

most likely have gained resistance against all of the antibiotics available(CDC 2019). Mitigation efforts on spreading of ARB and ARGs are ongoing, and the method to quantify the ARGs became essential for precise detection in the environment as well as in clinical studies. Addressing the growing concerns of AR requires attention in surveillance and response to acquire better understanding on the fate of AR.

Many studies and global communities noted that the prevalence and ubiquity of antibiotic-resistant genes (ARGs) is increasing over the decades due to human activities (Allen et al. 2010; Baquero, Martínez, and Cantón 2008; Bergeron et al. 2016; Di Cesare et al. 2015; Huang et al. 2019; Knapp et al. 2010; Pruden et al. 2006; Rizzo et al. 2013; Tacconelli et al. 2018). The sources of ARGs are categorized mainly coming from medical, industrial/ pharmaceutical, agricultural, and domestic areas (L. G. Li et al. 2020). Some anthropogenic activities of producing and overdosing antibiotics that affect our water environment are represented in **Figure 1.1.** below (L. G. Li et al. 2020)**.** Antibiotics are widely used in our daily lives. From agriculture to industrial level manufacturing, we often use them to prevent disease spreading among livestock, promote growth in farming, and consume them for individual purposes to treat illness or combined with over-the-counter medicine and hygiene products at home. The wastewater and soil surrounding human, agricultural land, and industrial manufacturing are contaminated with antibiotics and develop antibiotic resistance (AR) progressively. There is a need for better quantification and tracking of ARGs due to their health effects and socioeconomic impacts (Scott et al. 2018).

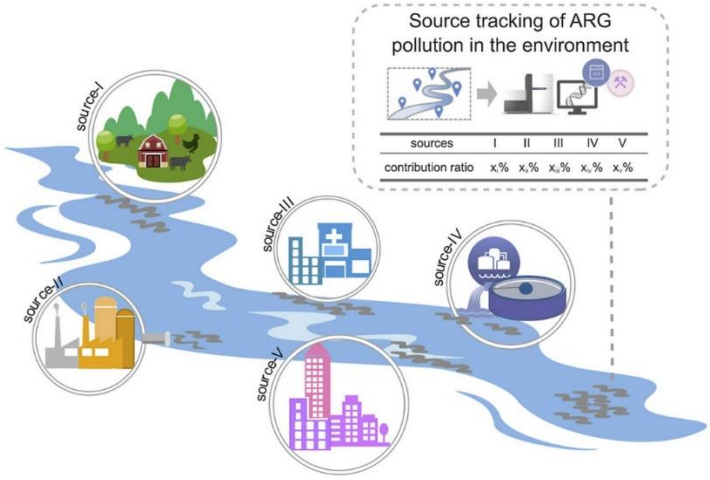


Figure 1.1. **Source Tracking of ARGs in the Environment** (L. G. Li et al. 2020)

According to the World Health Organization (WHO), we are lacking in global surveillance concerning AR which poses threat to not only human health but also has risks of system overloading, failure, or indirect hazards (Sanitation Safety Planning, Greywater and Excreta. 2016; WHO 2014). It urges public awareness, national action plans, optimizing the use of antimicrobials, innovation on research and development, and access, surveillance, and global governance with sustainable development goals (SDGs)(WHO 2019). For example, researchers found that there are anthropogenic antibiotic usage influences the accumulation of antibiotic-resistant bacteria (ARB) and ARGs in sediments of the estuary(Bhattacharyya et al. 2019; Guo et al. 2018), marine and river (Fiorentino et al. 2018; Huang et al. 2019; Pruden, Arabi, and Storteboom 2012b; Reddy and Dubey 2019) environments, as well as throughout the water reuse cycles (Christou et al. 2017). The spread of ARG was mostly predominant by horizontal gene transfer (HGT) with the use of mobile genetic elements (Van Hoek et al. 2011). Other research suggested the use of clinical class 1 integron-integrase gene (intl1) as an

indicator gene for anthropogenic pollution not only for antibiotics but also other contaminants such as pathogenic bacteria and heavy metals because of intl1's genetic relatedness and ability to be transported during the HGT process (Gillings et al. 2014).

By far, conventional methods used for quantitative measurement of ARG in the environment are the following: 1) PCR, 2) real-time PCR (rtPCR) or quantitative PCR (qPCR) for genetic target materials, 3) ddPCR, 4) flow cytometry methods and 5) cell counting for in vitro cultures of bacteria samples (R. Yang et al. 2014). So far, most of the ARG analysis is done with qPCR, and there are more and more applications suggested with increasing attention. The qPCR has been widely applied in the analysis of ARGs in the water system including urban stormwater, municipal and hospital wastewater, and seawater (Garner et al. 2017; J. Li et al. 2015; Preston et al. 2011; Zhu et al. 2018). The objective of this chapter is to develop and optimize the assay using quantitative polymerase chain reaction (qPCR) that can be used to identify ARGs in different environmental samples. Approaches for this chapter is to first develop standards for the analysis of environmental samples using robust reporting metrics that can be further used for universal comparison.

1.2 Materials and Methods

1.2.1. Sample Collection

Environmental samples were collected from Mallard Creek Wastewater Treatment Plant, Charlotte (NC) from activated sludge and transported on an ice cooler to the Environmental Microbiology Research Lab at the University of North Carolina at Charlotte (UNCC). The samples were pelleted with centrifugation to remove water content and processed for long-term storage at - 80°C freezer for further analysis. The sample processing and storage occurred immediately as it arrives.

1.2.2. Sample DNA Extraction

Sample DNA was extracted manually using QIAamp DNA Mini kit (QIAGEN, Cat # 51304) following the manufacture's protocol, in elution volume of 50 µL. The extracted DNA was stored in a -20 °C fridge until future analysis.

The QIAamp DNA Mini kit follows the procedure described as the following in the DNA extraction process., The 25mg of pallet or measured volume of sample or biomass added to 1.5mL centrifuge tube was located in a sample tray. The maximum volume of PBS solution f of 80 µL was used for resuspension of the pallet. The centrifuge tubes were vortexed to homogenize the sample. Into each tube, 100 µL of Buffer ATL is added followed by 20 µL of proteinase K, then vortexed to mix. The tubes were set in a floating rack and incubated at 56 °C for a minimum of 30 minutes. For optimum results, the samples can be incubated for 1~3 hours, shaken occasionally during the incubation period. After incubation, samples are removed, the incubator was set to 70 °C with a timer of 10 minutes for a future step. Centrifuge tubes with samples were

centrifuged briefly to remove any droplets on the upper side of the tube. Then 200 µL of 96% ~ 100% ethanol was added and mixed by pulse-vortex for 15 seconds, and sample tubes were centrifuged briefly. The QIAmp mini-column assembled with an elution tube was prepared and a sample mixture was applied into the column. Each cap of the elution tube is capped with the column inside, then centrifuged at 6000g(8000rpm) for 1 minute. The column was transferred into a clean collection tube, and the filtrate is discarded. 500 µL of buffer AW2 is added to the column then centrifuged at full speed (20,000g/14,000 rpm) for 3 minutes. The filtrate is discarded, a column is put to a new collection tube and centrifuged for 1 minute at full speed. The mini-spin column is put in a clean 1.5mL centrifuge tube, and then 200 µL of Buffer AE or nuclease-free water is added. The tube with the column is incubated at room temperature for 1~5 minutes (longer duration for higher DNA yield) and centrifuged at 6000g (8000rpm) for 1 minute.

1.2.3. Target genes and Primer selection

Based on literature review, target genes and corresponding primers sequences were selected for this study for ARG analysis, shown in **Table 1.1.** Yellow highlighted genes indicate ARGs that are not selected for standard development despite having a successful result in cloning. The genes are stored in cloned culture and may be applied for analysis if needed. Blue highlighted genes did not show a successful cloning result or was not able to extract positive standard from the environmental samples available.

Table 1.1 List of Target genes and primer sets

Target gene	Primer	Primer sequence (5′—3′)	Product size (bp)	Annealing Temp.(°C)	Reference
16S rRNA	1369F	CGGTGAATACGTTCYCGG	123	56	[47]

	1492R	GGWTACCTTGTTACGACTT			
tetA	tetA-F	GCTACATCCTGCTTGCCTTC	210	55	[34]
	tetA-R	CATAGATCGCCGTGAAGAGG			
tetC	tetC-F	CTTGAGAGCCTTCAACCCAG	418	55	[35]
	tetC-R	ATGGTCGTCATCTACCTGCC			
tetW	tetW-F	GAGAGCCTGCTATATGCCAGC	168	60	[40]
	tetW-R	GGGCGTATCCACAATGTTAAC			
tetX	tetX-F	CAATAATTGGTGGTGGACCC	468	58	[39]
	tetX-R	TTCTTACCTTGGACATCCCG			
Sul1	Sul1-F	CGCACCGGAAACATCGCTGCAC	163	55.9	[41]
	Sul1-R	TGAAGTTCCGCCGCAAGGCTCG			
Sul2	Sul2-F	TCCGGTGGAGGCCGGTATATGG	191	60.8	[41]
	Sul2-R	CGGGAATGCCATCTGCCTTGAG			
Sul3	Sul3-F	TCCGTTCAGCGAATTGGTGCAG	128	60	[41]
	Sul3-R	TTCGTTCACGCCTTACACCAGC			
qnrA	qnrA-F	TCAGCAAGAGGATTTCTCA	516	50	[42]
	qnrA-R	GGCAGCACTATGACTCCCA			
qnrB	qnrB-F	TCGGCTGTCAGTTCTATGATCG	469	54	[42]
	qnrB-R	TCCATGAGCAACGATGCCT			
ereA	ereA-F	AACACCCTGAACCCAAGGGACG	420	52	[43]
	ereA-R	CTTCACATCCGGATTCGCTCG			
mphA	mphA-F	AACTGTACGCACTTGC	837	52	[45]
	mphA-R	GGTACTCTTCGTTACC			
blaCTX-M	blaCTX-M-F	ATGTGCAGYACCAGTAARGT	593	50	[46]
	blaCTX-M-R	TGGGTRAARTARGTSACCAGA			

1.2.4. Cloning

1.2.4.A. Preparation of LB Agar Plates

For incubation stage of cloned cells, agar plates with Lysogeny broth (LB) were prepared ahead of time. In a 1L beaker, 20g of premised LB-Agar powder were dissolved to the total volume of 1L with autoclaved distilled water. The LB-Agar solution was transferred in an autoclavable bottle, capped loosely, and autoclaved at 121 °C for 15 min by selecting cycle: LIQ15. Autoclaved bottle was cooled in ambient, sterile environment to reach 50-55 °C. In a biosafety hood with laminar air flow and disinfected surfaces, 50 mg/mL 250 µl of freshly made kanamycin solution was added into the LB-Agar solution and stirred gently. The final concentration of kanamycin reached 50 µg/mL. The kanamycin can be stored in the -20 °C freezer for up to 1 month.

The LB-Agar-kanamycin mixture was poured into sterile petri dishes, 20mL in volume per each plate. Lid was used to cover the surface to prevent impurities entering the plate, then cooled in stack for 1 hour so that the solidified surfaces are even and free of impurities. Plates were transferred to an incubator at 37 °C and placed upside down to remove any dews formed. Then the plates were sealed in wrappers to prevent from light reactions and stored at room temperature. Plates with any growth were autoclaved and discarded accordingly.

1.2.4. A*. Blue-White Screening of Bacterial Colonies Preparation Method

This step can be added to ensure the addition of target gene to E. coli plasmid by showing blue and white colors during the colony identification. For specific target genes that were

hard to be acquired were followed by this step for extra verification purposes. Materials needed for this procedure are X-Gal (GoldBio, Cat. #X4281C), dimethyl sulfoxide, DMSO, distilled water, isopropyl β-D-1-thiogalactopyranoside, IPTG (GoldBio, Cat. #I2481C), screening antibiotic of choice, agar media, and plates.

First, X-Gal solution of 40mg/mL was prepared in DMSO, by weighing 400mg of X-Gal into a 15mL polypropylene centrifuge tube. Under the laminar flow fume hood, 10mL of DMSO was added. The centrifuge tube was capped and vortexed until X-Gal powder is dissolved, and stored at -20 °C. Second, 100mM of IPTG solution was prepared using distilled water. 0.238g of IPTG was weighed and added to 10mL of sterile distilled water, and gently inverted to mix until solids are completely dissolved. Using a syringe and prewetted 0.22 µm syringe filter, the IPTG solution was filtered through for sterilization. Each 1mL aliquot was stored in sterile 1.5mL centrifuge tubes, capped, and stored at -20 °C for up to a year. On the surface of LB-Agar plates, 60 µL of X-Gal (40mg/mL) and 40 µL of 100mM IPTG solution was added and spread over the entire surface using a spreader. The surface coated with X-Gal and IPTG were dried under the biosafety hood for 30 minutes before use.

When making LB-Agar mixture instead of plating on the surface, the blue-white colony reagents was added into the agar mixture before plating and solidification of agar. For X-Gal, 20µL of 40mg/mL X-Gal reagent was added per 2mL of media. Since each media's volume was 20 mL, 100 µL of X-Gal solution was added to each plate. The 10 µL of 100mM IPTG solution was added per 1mL of media to reach a final concentration of 1mM. For a 20mL agar plate, 200 µL of IPTG solution was added.

1.2.4.B. PCR product Preparation & Gel Electrophoresis Assay

With the extracted DNA from the samples, PCR reaction mixes were prepared in triplicates with the target set of primers. PCR cycle was set with a specific annealing temperature of each target gene. Once sample allocation of plate setup is done using the software connected to PCR thermal cycler, the PCR was started.

While waiting for the PCR cycles to finish, gel electrophoresis was prepared for verification of the size of the target gene. To make 2% Agarose gel, 1.6 g of agarose powder was put into 80mL of Tris/Borate/EDTA (TBE) or Tris/Acetate/EDTA (TAE) buffer in a glass beaker. A coloring reagent of either ethidium bromide (EtBr) or Gel Red was added to the agarose mixture with the volume of 5µL and swirled carefully for mixing. The mixture in the beaker was heated in a microwave for 30 seconds, followed by 10 seconds interval until the mixture looks homogeneous and free from any grainy materials within the transparent liquid. The mixture was cooled on a benchtop to reach approximately 60 °C, then poured into assembled agarose gel mold that was assembled and leveled with combs in. The poured agarose gel was solidified for 50 minutes at room temperature, then combs were removed to prepare inserted in a gel electrophoresis bath. After the PCR reaction of cloned plasmids, one part of the amplified DNA is mixed with five-part of 5X loading buffer, then each DNA-dye mix was loaded to the wells while pipette tips are submerged under the surface of the buffer solution. Depending on the size of the target DNA, the DNA size ladder (molecular rulers, BIO-RAD) ranging expected base pairs was loaded on the first well or in appropriate locations. The horizontal gel electrophoresis system (BIO-RAD) was set up at 80V for 50 minutes while the moving speed of DNA is monitored. The PCR products that are identified with the size of the

target gene can be used for success in cloning and can be used for the following steps. All other failed PCR products are discarded properly, and excess successful PCR reactions were saved in the -20 °C freezer.

1.2.4.C. Ligation & Cloning

In preparation for the expedited procedure of cloning to ensure the viability of competent cells and increase the chances of cloning, it is important to set up an incubator and a water bath at the desired temperature. The non-shaking incubator was set to 37 °C and LB plates were pre-heated for 30 min. The shaking incubator was set to 37 °C, and the water bath was set to 42 °C. Surfaces of biosafety hood were disinfected ahead of time with 70% ethanol and DNase Away reagent because cloning and ligation procedure proceeded inside the hood. One-Shot TOP10 chemically competent cells (Invitrogen, Cat. # C404003) as *E. coli* were removed from the -80 °C freezer and immediately placed on ice. Due to the high sensitivity of competent cells, the tube bottoms/base were made sure not to be touched by hand and always kept cold.

The following procedure used reagents Invitrogen TOPO TA Cloning Kit (Thermo Fisher Scientific Inc., Cat # K4500-02). The ligation mixture was prepared in a PCR tube with 4 µL PCR product, 1 µL salt solution, 1 µL pCR 2.1-TOPO vector, and sterile water to reach a total volume of 6 µL. Each ligation mix was gently vortexed to mix and incubated at room temperature for 5-30 minutes. After incubation, 2 µL of ligation product was transferred to thawed competent cells on ice and gently vortexed without any rough mixing or inversion. The remaining ligation product was saved and stored at 4°C until the cloning reaction is verified. The ligation product can be used for troubleshooting if the cloning procedure fails.

Competent cells with ligation mix were incubated on ice for 5-30 minutes while super optimal broth with catabolite repression (SOC) medium is thawed. Then the competent cell mixture was heat-shocked at 42°C for exactly 30 seconds in a hot water bath without shaking. This is the most crucial step which requires accuracy of temperature and time. After incubation, the tubes were immediately transferred back to the ice bed. In each tube, 250 µL of room-temperature SOC medium was gently pipetted to the side of the tube. In a shaking incubator set at 200rpm and 37°C, each tube was capped tightly and incubated horizontally for 1 hour.

The next step was followed by spreading cloned cell mixture onto agarose-LB plates to find competency of bacterial survival and its function against antibiotic. Ensuring the spreading of colony in plate incubation, general recommendation for number of plate and volume to plate is plating four different volume amounts of cloned cell mixture onto agarose-LB plates, by putting each of the 20, 50, 80, and 100 µL of cell mixture. The plating was done under biosafety hood and covered with plate tops and the rims were sealed with parafilm. The packaged plates were put upside down as the surface of gel is facing the bottom and incubated at 37 °C for 16 – 18 hours. After incubation, plates were taken out and observed for colonies formed on the surface of each plates. Desired colony for target gene collection should be clearly visible and large enough to be harvested using a loop. The colonies that are too small or not showing are not touched and properly discarded after disinfection.

1.2.5. Colony PCR

1.2.5.A. Colony Extraction and Inoculation

The colony PCR method allows conservation of colony with continuous growing so that the essential bacteria with target gene are preserved and standard can be obtained many times as needed. Per each colony that needs to be picked, a set of PCR tube and 15mL sterile centrifuge tube was prepared and labeled properly. In each 15mL centrifuge tube, 10mL of LB broth and 10 µL of 50mg/mL Kanamycin solution were added to reach a final concentration of 50 µg/mL of Kanamycin. For plates that were treated with blue-white screening reagents, a colony that does not have the dark blue tint but a completely white color was chosen for incubation. Using the tip of the pipette, a colony was picked. Then the bacterial colony is transferred to an empty PCR tube by gently rubbing the tip in a circular motion. Then the used pipette tip was ejected into the prepared LB-Kanamycin solution and capped. The 15mL centrifuge tubes with the colony and broth mix were incubated at 37°C for 18~24 hours until the PCR results are identified with which colony has desired positive clones.

Taq 2X master mix (New England BioLabs INC., Cat. # M0270L) was used to prepare colony PCR reaction mix. Each PCR reaction was prepared with 12.5 µL of Taq2X master mix, 11.5 µL of nuclease-free water, and 0.5 µL of forward and reverse primers in a PCR tube with a colony residue. Thermal cycling steps of reaction mix cell ligation and colony PCR was set to 95°C for 10 minutes for cell lysis, then followed by 30 cycles of 95°C for 15 sec, 45-68°C for 15-60 sec, 68°C for 1 minute/kb, and final extension of 68°C for 5 minutes. Then the reaction mixture was cooled to 4 °C until the PCR product is collected.

While waiting for PCR reaction towards completion Agarose gel plate was prepared as described in the section for gel electrophoresis of PCR products. PCR

products were loaded and identified for the target gene at the designed base pair lengths. PCR products that do not have a band of the appropriate size were discarded. Clones without inserts showed bands around 150-200 bp.

1.2.5.B. PCR Product Cleanup

Removal of Taq polymerase, primers, and extra cellular components other than DNA is essential to obtain quality DNA standards. For the cleanup procedure, ExoSAP-IT (GE Healthcare Bio-Sciences Ltd.) was used for the purification process.

Using a pipette, 5 µL of PCR product and 2uL of ExoSAP-IT reagent were put in a new PCR tube. The tube was vortexed gently to mix the contents, then put in the thermal cycle of (i) 37 °C for 15 minutes, (ii) 80 °C for 15 minutes, and (iii) 4 °C for 3 minutes. When the cycling is done, the DNA sample is ready for other applications such as sequencing and qPCR. The DNA completed with the steps was labeled and stored in a capped container at -20 °C until future use.

1.2.6. Plasmid DNA Extraction

1.2.6.A. Plasmid DNA Purification

For plasmid DNA extraction, QIAprep Spin Miniprep Kit High-Yield (Cat. #. 27104, QIAGEN) was used following the manufacturer's protocol. LyseBlue reagent was added to Buffer P1 at a ratio of 1 to 1000. From the 15mL centrifuge tubes with an inoculated bacterial colony that has been incubated overnight or 12-16 hours, 5mL of culture was palleted by centrifugation at >8000 rpm (6800 g) for 3 min at room temperature. The

supernatant was discarded, and the pellet was resuspended in 250 µL of Buffer P1. When completely resuspended, the cells were transferred into a 2mL microcentrifuge tube, and 250 µL of Buffer P2 was added and inverted to mix for 10-12 times to show blue color and incubated for lysis for 5 minutes. Buffer N3 of 350 µL was added and immediately mixed until the solution turns colorless. The mixture was centrifuge for 10 minutes, and the supernatant was loaded to the QIAprep spin column. The column was centrifuged for 60 seconds and flowthrough was discarded. To wash the column from impurities, 500 µL of Buffer PB was added and centrifuged for 30-60 seconds.750 µL of Buffer PE was added to the column and centrifuged for 30-60 seconds to remove flowthrough. After discarding any remaining liquids, the spin column was centrifuged for 1min to remove residual wash buffer. In a clean 1.5 mL microcentrifuge tube, a spin column was placed, and 60 µL of Buffer EB was added to the center of the column filter for DNA elution. The column was incubated at room temperature for 1 min, and centrifuged for 1 min. The plasmid DNA collected at the bottom of the collection tube was stored at -20 °C for further analysis.

1.2.6.B. Plasmid DNA Concentration and Calculations

Extracted plasmid DNA concentration was measured using Nanodrop One microvolume UV-Vis spectrophotometer(Thermo Fisher Scientific, Agawam, MA) with the Buffer EB used as a blank. Each value was recorded. With the known length of the cloned vector and length of the ARGs inserted, copy numbers of plasmid from extracted DNA are calculated. The size of the pCR2.1 -TOPO vector is 393l nucleotides, and the base pair size of target ARGs was used to calculate the final number of the copy number. An

equation in calculating the number of copies is shown below, where N is the length of plasmid fragment in base pairs, and X is the concentration of plasmid measured in nanograms (ng) (Barczak et al. 2015) :

$$Number\ of\ Copies\ (molecules) = \frac{X\ (ng)\times 6.0221\times 10^{23}\left(\frac{molecules}{mole}\right)}{N\times 660\left(\frac{g}{mol}\right)\times 1\times 10^{9} ng/g} \qquad \text{Eq. (1)}$$

1.2.7. Standard Preparation using qPCR

Using the positive stock DNA, eight sequences of 1:10 serial dilution were done for each stock plasmid DNA. After the transfer of 1 part volume, each well was mixed thoroughly to avoid large dilution errors.

Then amplification reaction mix was prepared following the list of primers shown in **Table 1.1** for identifying each target ARG. The SsoAdvanced Universal SYBR Green Supermix (Bio-Rad Laboratories) was used for the entire analysis of ARG quantification. Each sample was prepared in triplicates with 8-point serial dilutions of standard and negative control for a statistically reliable qPCR run.

Plate design for qPCR analysis of samples was done using Bio-Rad CFX Manager software (Bio-Rad Laboratories) on CFX Touch Real-Time PCR Detection System (Bio-Rad Laboratories). The User-defined run type was selected to set the thermal cycling steps. The annealing temperature is set differently depending on the target genes as shown in **Table 1.1**. Sample volume per well was set at 20 µL for all reactions, having 10 µL of SsoAdvanced Universal SYBR Green Supermix (Bio-Rad Laboratories), 0.4 µL of each forward and reverse primers of target gene designed, 2 µL of DNA template, and

7.2 µL of nuclease-free water. The general thermal cycling setting is shown in **Table 1.2**. Once a protocol is set, plate assignment of wells was arranged in the order of standards on a 1st column, then triplicates of samples and blanks on the rest of the columns.

Standards prepared in serial dilution were verified with qPCR. The standard measurement was verified for linearity and robustness by monitoring amplification peaks and R^2 values for each set of dilutions. Dilution series that meets the R^2 value over 0.85 was used for the qPCR analysis of environmental samples. The standards that did not meet linearity standards and caused errors were prepared again starting from the cloning procedure.

Table 1.2. Thermal cycling protocol for target qPCR analysis

Step #	Temperature (°C)	Duration
1	95	3:00
2	95	0:15
3	Anneal temperature for target gene	0:30
4	Go to step #2 repeat 39 more times	
5	Temperature gradient to 95 °C with 0.5 °C increments	0:05
6	4	3:00 or ∞

1.3. Results & Discussion

Standard curves for each targeted ARG were developed based on the copy numbers calculated from cloning. The concentration of the stock standard (X), the size of target ARG sequence (size), the length of the cloned plasmids in base pairs (N), and the final copies/µL concentration of ARGs are shown in **Table 1.3**. Target genes that were cloned and screened for positive signals were used, but not all the positively signaled target genes were successful in standard developments.

Table 1.3. Concentration of Stock Plasmid DNA for target genes

Target	X(ng/uL)	Size	N (bp)	# of copies/uL
16S rRNA	4.09	143	1522	2.45E+09
tetA	2.17	210	4521	4.38E+08
tetW	17	168	4479	3.46E+09
Sul1	1.4	163	4474	2.86E+08
Sul2	2.3	722	5033	4.17E+08
Sul3	4	128	4439	8.22E+08
qnrA	0.4	516	4827	7.56E+07
qnrB	30.3	469	4780	5.78E+09
ereA	3.9	420	4731	7.52E+08
bla ctx-M	33.327	593	4904	6.20E+09

The standard curves created are shown in following **Figures 1.2-1.6**. Each dilution series of ARGs were analyzed with 7-point standard and blank with nuclease free water.

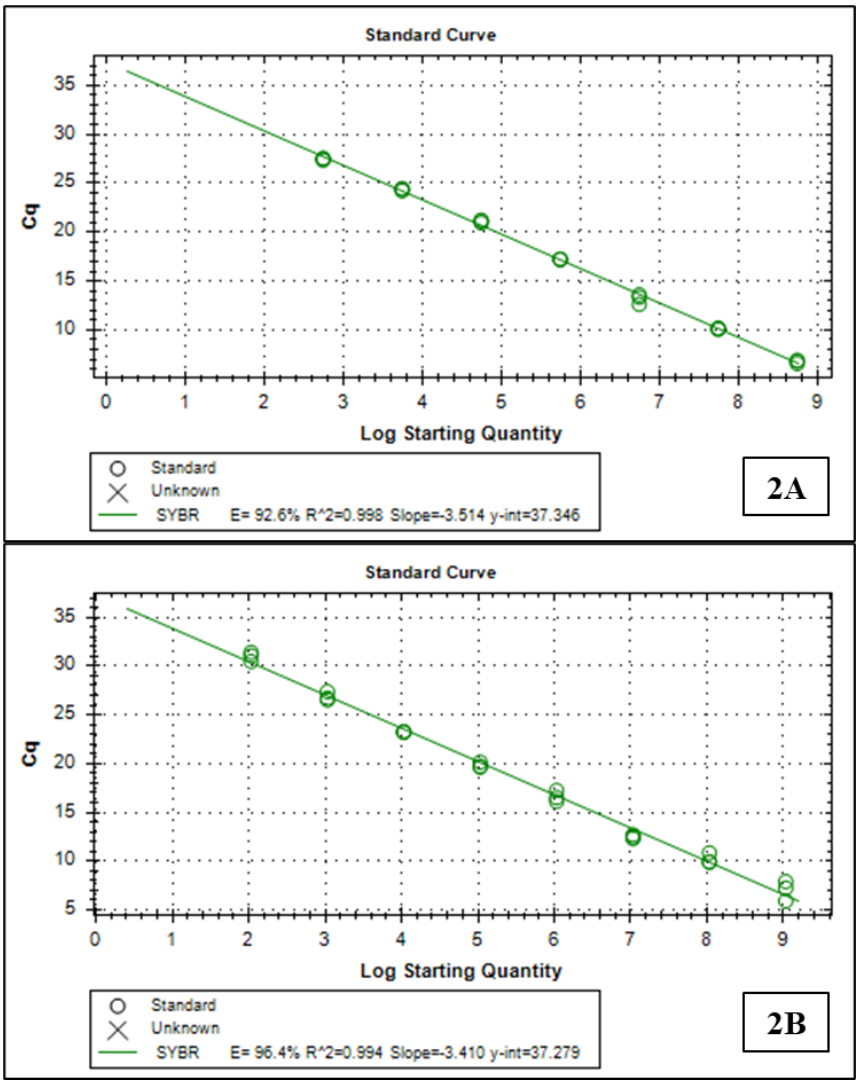


Figure 1.2. Standard for tetA (2A) and tetW (2B)

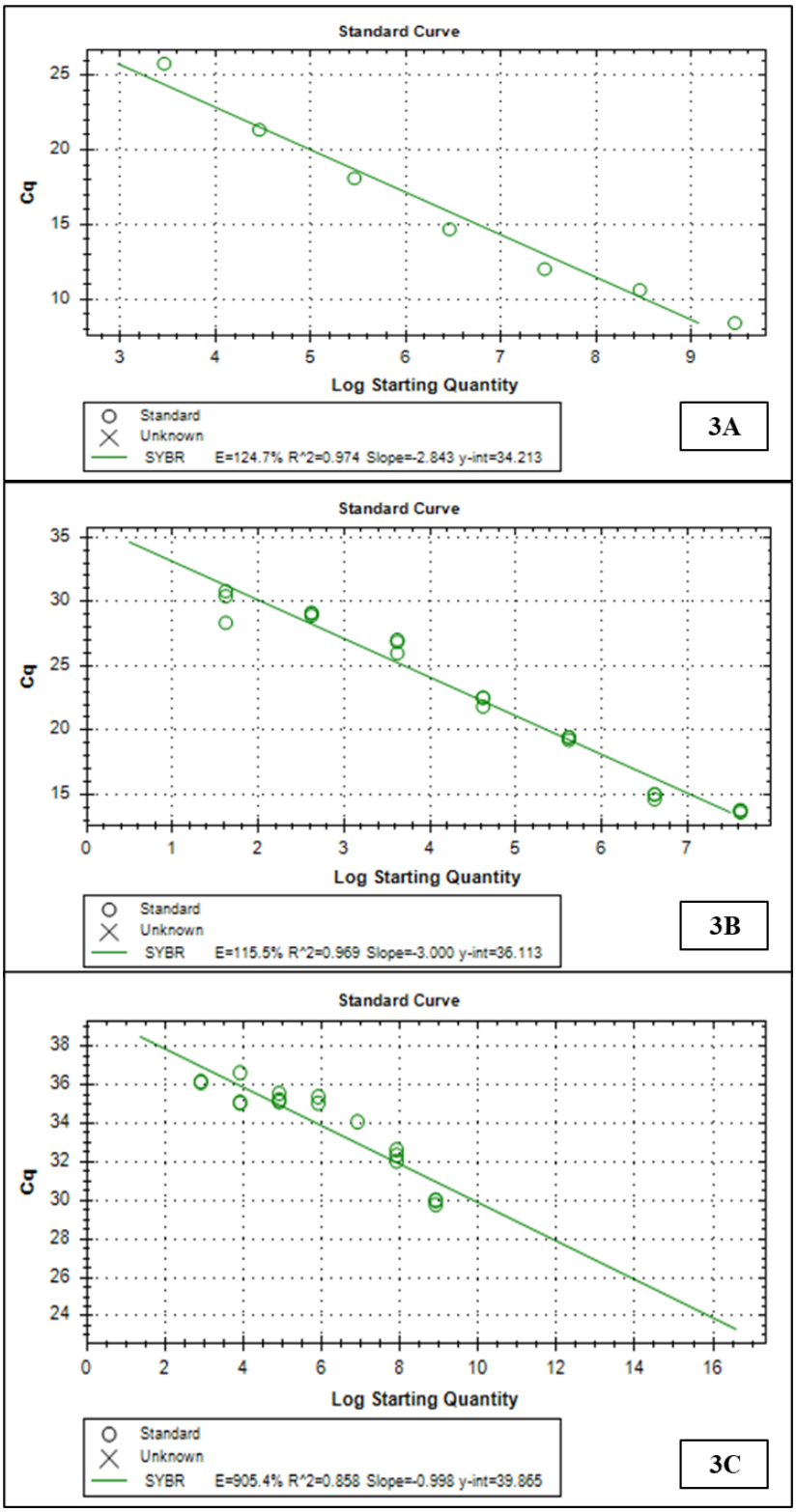


Figure 1.3. Sul 1 (3A), Sul 2 (3B), and Sul 3 (3C) Standard

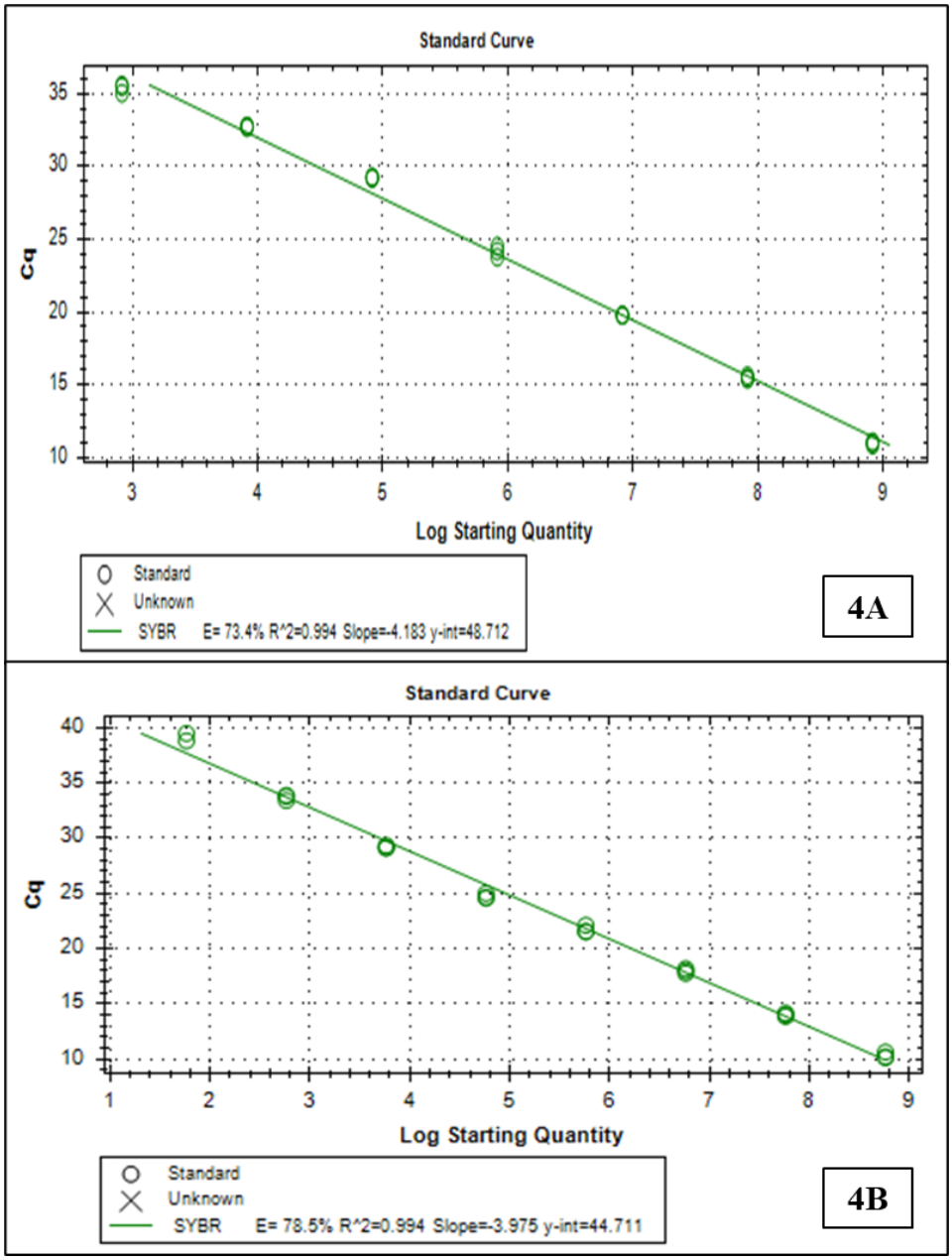


Figure 1.4. qnrA (4A) and qnrB (4B) Standard

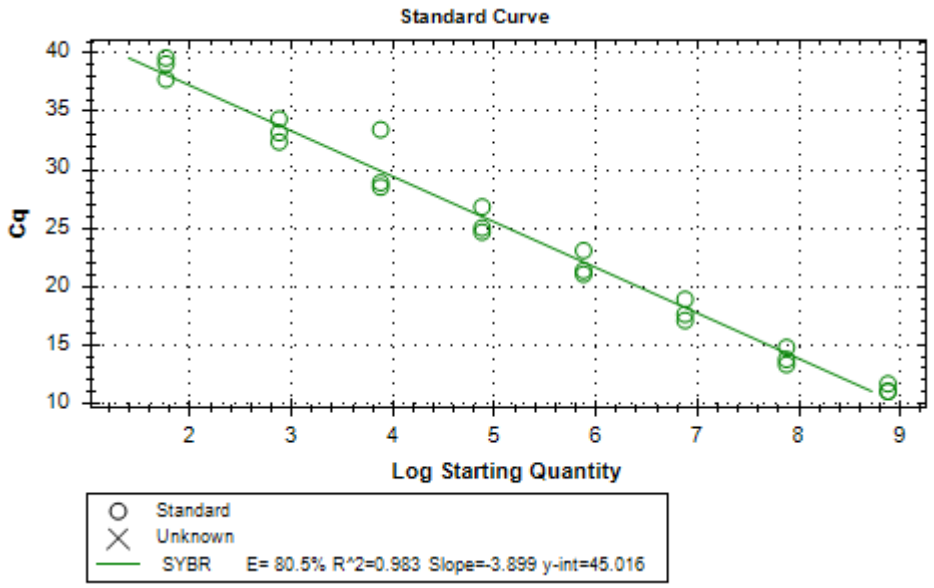


Figure 1.5. ereA Standard

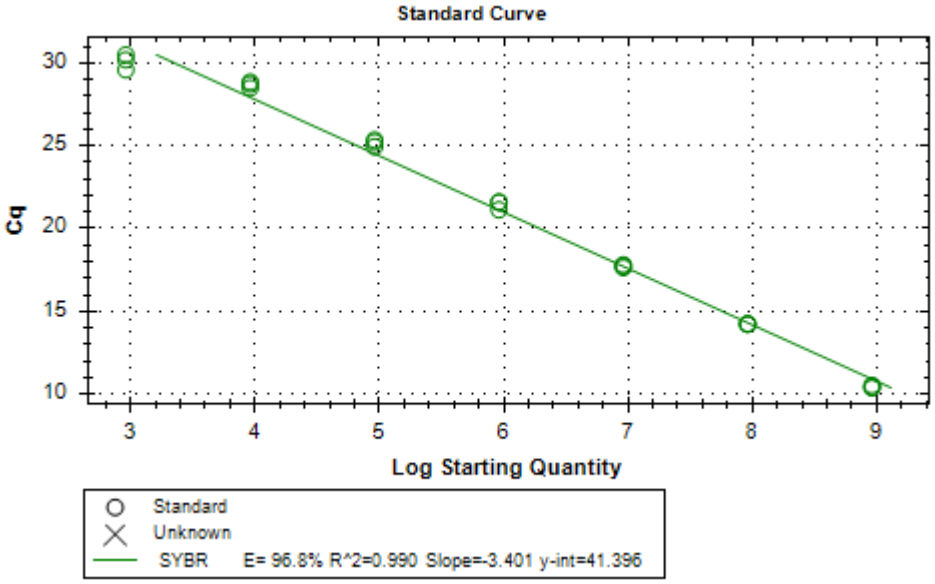


Figure 1.6. 16S rRNA Standard

Most of the standard curves have high linearity to copies of genes in each dilution with R^2 values above 96%. Only in the case of sul 3, the R^2 value was as lower as 85% which is assumed to be caused by target gene length and primer design. During

amplification detection, qPCR may not be suitable for all of the genes because qPCR was losing linearity at the lower and higher ends of the detection. Regarding the dependency of targets, it can be verified by a comparison of other techniques with the same settings.

Some ARGs that showed positive cloning results but were not used for standard development are shown in following **Figures 1.7 – 1.9**. **Figure 1.7** shows some positivity and amplification of mphA gene, but the bands were not shown in **Figure 1.8** through gel electrophoresis. This may be due to having too low genetic amounts to run a gel and need more amplification. In the case of cloning attempts with tetC (**Fig. 1.9**), it was not successful to harness the target gene from the environmental sample used for cloning. **Figure 1.9** shows identified bands of tetA, tetW, and tetX. **Figure 1.10** shows colony PCR product blaCTX_M from different colonies grown from incubation, which showed some positive bands of 6 out of 10 selected.

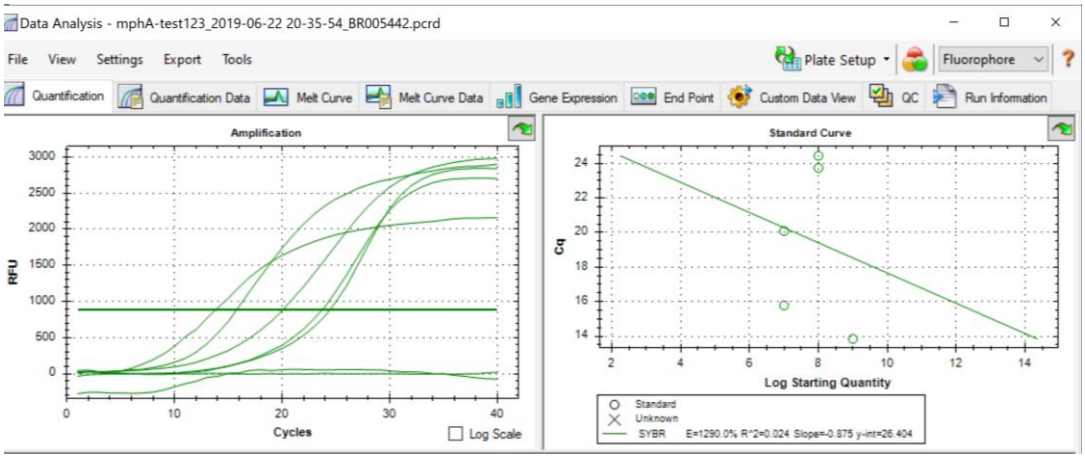


Figure 1.7. mphA colony PCR

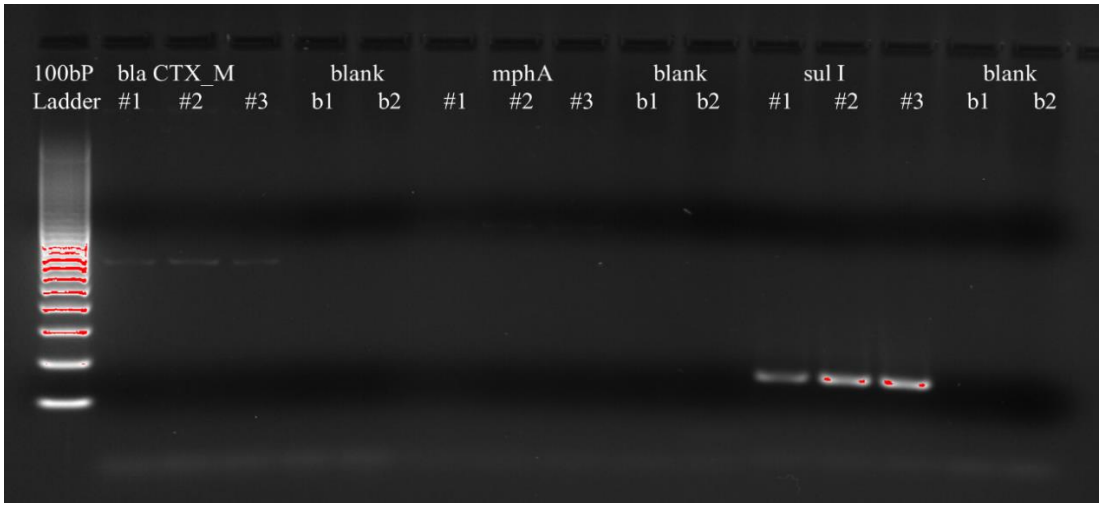


Figure 1.8. Colony PCR of bla CTX_M, mphA, and sul 1

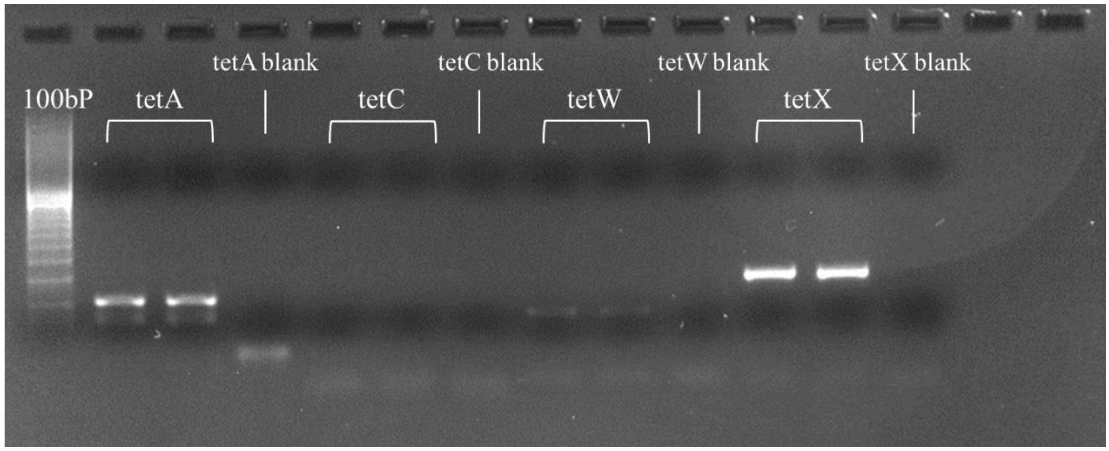


Figure 1.9. PCR products of tetA, tetW, and tetX

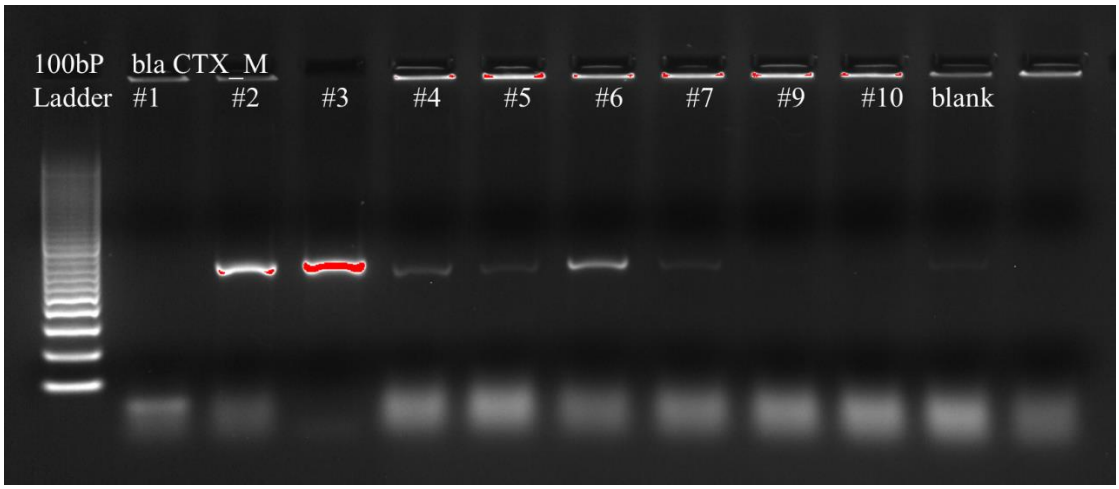


Figure 1.10. Bla CTX-M gel verification of colony PCR at 593 bp

1.4. Conclusion

Through i) primer selection, ii) target DNA selection in environmental sampling, iii) cloning, iv) colony PCR and gel verification, v) serial dilution and standards preparation, the qPCR methods were ready to be used for subsequent studies as described in the following chapters. Out of 13 target genes, 10 were prepared and ready for use in the following projects. Two of the genes, tetX, and mphA were identified and successful in cloning, but not used for the analysis of samples for further projects. The tetC was not able to be identified either through cloning or standard DNA extraction methods.

Based on the MIQE guidelines, all standards were prepared to address each point of quality control recommendation (S. A. Bustin et al. 2009). The MIQE guideline suggests parameters that are considered for targeted qPCR analysis, such as i) analytical sensitivity, ii) analytical specificity, iii) accuracy, iv) repeatability, and v) reproducibility. Most of the samples showed high repeatability and reproducibility throughout the standard development and sample runs, as well as similar results in triplicates. The linear dynamic range seemed robust with high efficiency and high R^2 values most of them being above 96%. Limit of detection (LOD) does not seem to be an issue in bringing bias to analysis because the range of standards was aligned with the standard curve and expected copy number calculations from cloning.

CHAPTER 2: Comparison of Droplet Digital PCR (ddPCR) and Quantitative PCR (qPCR) methods for ARGs detection

2.1. Introduction

The droplet digital polymerase chain reaction (ddPCR) system is considered to be the next-generation technology of the original polymerase chain reaction (PCR). The ddPCR system consists of stepwise compartments of droplet generator, PCR, and droplet reader. The difference between conventional PCR and ddPCR is that the droplet generator captures and insulates a single genetic molecule such as RNA or DNA in oil droplets, giving the ability for researchers to read amplification signals from each droplet. Per one reaction well, the droplet generator can make up to 20,000 droplets making the reading more precise without the need of having a positive control for analysis.

Many studies and global communities noted that the prevalence and ubiquity of antibiotic-resistant genes (ARGs) are increasing over time due to human activities of producing and overdosing antibiotics (Allen et al. 2010; Baquero, Martínez, and Cantón 2008; Bergeron et al. 2016; Di Cesare et al. 2015; Huang et al. 2019; Knapp et al. 2010; Pruden et al. 2006; Rizzo et al. 2013; Tacconelli et al. 2018) Antibiotics are widely used in our daily lives. From agriculture to industrial level manufacturing, we often use them to prevent disease spreading among livestock, promote growth in farming, and consume them for individual purposes to treat illness or combined with over-the-counter medicine and hygiene products at home. The wastewater and soil surrounding human, agricultural land, and industrial manufacturing are contaminated with antibiotics, and developing antibiotic resistance (AR) gradually, cycling and accumulating in the environment and in

organisms (**Fig. 2.1**). According to the World Health Organization (WHO), we are lacking on global surveillance of AR which poses a threat not only to human health but also can lead to overloading and systemic failure of environmental systems (Sanitation Safety Planning, Greywater and Excreta. 2016; WHO 2014). The WHO urges public awareness, national action plans, optimizing the use of antimicrobials, innovation on research and development, and access, surveillance, and global governance with sustainable development goals (SDGs) (WHO 2019). The WHO's guidance is consistent with recent studies showing anthropogenic antibiotic usage influences the accumulation of antibiotic-resistant bacteria (ARB) and ARGs in sediments of estuary (Bhattacharyya et al. 2019; Guo et al. 2018), marine and river (Fiorentino et al. 2018; Huang et al. 2019; Pruden, Arabi, and Storteboom 2012b; Reddy and Dubey 2019) environments, as well as throughout the water reuse cycles (Christou et al. 2017).

Figure 2.1. **Sources of anthropogenic influence of AR in the water system** (created using Biorender).

The ARGs have been hypothesized to spread through horizontal gene transfer (HGT) with the use of mobile genetic elements (MGEs) (Van Hoek et al. 2011), and are often linked to other gene elements associated with heavy metals such as integron-integrase gene (intl1) (Gillings et al. 2014). The MGEs are often tested with ARGs using high throughput qPCR or ddPCR analysis and 16S rRNA sequencing for testing the environmental effects and the fate of ARGs (Gao et al. 2018; Shen et al. 2019). In order to uncover more and more characteristics of AR and its relation to human activity, it is important that we can identify and quantify ARB and ARG in different environments and at different genetic targets.

2.2 Evolution/Expansion of Nucleic Acid Detection Methods for Molecular Targets

2.2.1. The Initial State of Polymerase Chain Reaction (PCR)

The ability to identify, detect, and quantify nucleic acids has significantly improved since the mid-1980s. The groundbreaking first-generation polymerase chain reaction (PCR) device used thermal cycling to amplify specific regions of DNA in the presence of Taq polymerase (de Carvalho and da Fonseca 2017), deoxynucleotide triphosphates (dNTPs), and primers of the desired target. The thermal cycling process denatures DNA, anneals the primers and the desired target, and amplifies DNA using polymerase and dNTPs over multiple cycles (Garibyan and Avashia 2013). This rapid amplification process used in PCR reduced the time and cost of nucleic acid detection in medicine and research, and it boosted most major fields of science. However, PCR has been reported with its downside that it is hard to separate incomplete and damaged genes amplified in the detection of signals (Clewley 1989). In other cases, incomplete DNA template fragments (for example if DNA is extremely degraded) may not be amplified by PCR if the primer site is missing. Other examples of limitation in PCR include presence of primer-dimers (S. Bustin and Huggett 2017) and different levels of inhibition observed depending on target primer binding structures (Chandler, Wagnon, and Bolton 1998). A study in the late 1980s by Clewley summarized the approaches to minimize contamination and erroneous issues of PCR by suggesting practical measures such as using well-characterized controls, improving lab techniques and practices, optimizing primer selection, and testing of plasmids (Clewley 1989). Throughout the years, researchers also evaluated PCR for forensic applications (Romsos and Vallone 2019),

variant identification (Perry et al. 1989), and primer and temperature optimization (Cao et al. 2020; Vishnuraj et al. 2021), and PCR has been continually optimized for various applications (Butler, Ruitberg, and Vallone 2001; Ekman 1999; Wong et al. 2014). What researchers expected as potential applications of the PCR technique have become more general nowadays. The following paragraphs describe new generations of PCR techniques and their uniqueness.

2.2.2. The Second Generation - qPCR

The emergence of real-time PCR allowed for the fluorescent detection of nucleotides during amplification. The real-time amplification fluorescence detection enabled the relative quantification of target genes with the use of known standard genes. As fluorescence signal is detected throughout the multiple cycles of gene amplification, when the intensity is over the threshold, the Ct value is recorded (Ponchel et al. 2003). Based on Ct values of samples, relative genetic quantities can be known using the standard. By the meaning of the possibility of gene quantification, real-time PCR is also known as the quantitative polymerase chain reaction, qPCR.

Currently, qPCR is the most widely used technique for the identification and quantification of genetic matters. This technique gave rise to broader use in scientific analysis in food, agricultural, medical, and environmental fields. Furthermore, multiplex real-time PCR enables the tracking of different genotypes in the target of interest made “categorical analysis” easier. For example, Alia et al. (2020) studied four major foodborne pathogen genotype isolates from meat processing plants (Alía et al. 2020).

With various applications for qPCR, direct comparison of qPCR performance remains difficult due to variations in the initial template amount, partitioning of sample, and sample preparation methods. Still, the downfall of the qPCR method is that it is susceptible to inhibitors from environmental contaminants that can arise from the extraction of nucleic acids and reaction mix compositions (Huggett et al. 2008). Such inhibitors can lead to inaccurate or overestimated results depending on the type of inhibitor, the number of amplicons, or biological sample type (Huggett et al. 2008; Karlen et al. 2007; Suslov 2005). In addition, all PCR application techniques are subject to poor primer/probe conditions requiring multiple rounds of probe and primer optimization. In addition, all PCR application techniques are subject to poor primer/probe conditions requiring multiple rounds of probe and primer optimization.

2.2.3. The Next Generation - ddPCR

Droplet digital PCR (ddPCR) is a third-generation PCR technique that sequesters probes and nucleic acids within droplets in an oil emulsion (Nyaruaba et al. 2019). Using ddPCR can generate up to 20,000 droplets in a single-tube reaction, with a singular PCR reaction occurring in each droplet. The number of droplets with (and without) amplification can be counted allowing for absolute quantification at a molecular level. ddPCR detects fluorescence at the endpoint of amplification, identifying the number of positive droplets and determining total concentration by Poisson distribution. Examining the number of negative and positive droplets together helps obtain the absolute copies per reaction (Biorad 2018).

While the concept of reading the fluorescent signals from amplified genes is the

same compared to qPCR, ddPCR is better at preserving the initial reaction mix conditions and clearly representing amplification progress, minimizing the occurrence of primer-dimers, and false amplification reading due to the processing of the reaction mix for analysis (**Figure 2.2**). ddPCR widely applies to many fields due to its high specificity and lower limit of detection across many applications. It has been used for better detection of copy number variants (CNV) in samples compared to conventional PCR (Q. Liu et al. 2020; Tone and Torunn 2016), gene expression of RNAs from viral targets and pathogens (Pinheiro-de-Oliveira et al. 2018a; Taylor et al. 2015), and food research (Xiaofu Wang et al. 2019; ZHONG et al. 2018). Some clinical study examples are on cancer and tuberculosis research (Nyaruaba et al. 2019), and other examples include product quality determination in the food industry such as olive oil (Scollo et al. 2016), fish (Lin et al. 2020), and meat (Naaum et al. 2018) products. Other studies compared the use of qPCR and ddPCR methods for detecting different targets such as meat products or Cryptosporidium, which reported lower limit of quantification (LoQ) for determining meat product identification and better relative standard deviation (RSD%) to detect Cryptosporidium for ddPCR than qPCR (Floren et al. 2015; R. Yang et al. 2014). Moreover, ddPCR has been used to detect and quantify amounts of H7N9 influenza virus cDNA and was able to do so at a resolution of 3 to 5 copies/uL (Ahrberg, Lee, and Chung 2019; Heyries et al. 2011). On the other hand, qPCR will detect fluorescent signals at the exponential phase of amplification (Quan, Sauzade, and Brouzes 2018), which does not easily allow for absolute quantification and such a low limit of detection (LoD). For this reason, it is the most important benefit of ddPCR that it is free from factors such as

requiring a standard for calibration or PCR efficiency, and the result is only dependent on the molecular interactions within the droplets (Huggett, Cowen, and Foy 2015).

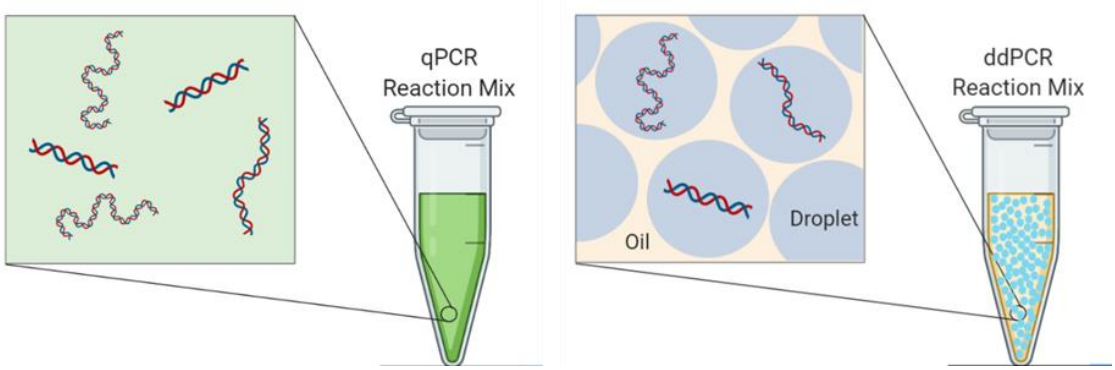


Figure 2.2. Reaction mix comparison of qPCR and ddPCR (image created using Biorender)

Though the results of analysis for both qPCR and ddPCR are represented as copies per microliter (copies/volume), the output of ddPCR is different from the result produced from qPCR. An example is shown in Figure 2. The qPCR produces a real-time amplification curve in a 1D plot for each polymerase amplification reaction (**Figure 2.3A**, adapted from another study) showing relative fluorescence, which also can be represented as the Ct value, of different samples. On the other hand, ddPCR measures fluorescence signals at post-amplification from each droplet, which captures individual target genes and can be presented in 1D or 2D plots. This shows groups of similar amplification, which makes it easier to detect negative and positive populations of droplets. **Figure 2.3B** shows a 1D plot of ddPCR with the automated threshold line shown in pink color by the ddPCR Quantasoft software (adapted from Ibekwe M.A. 2020 (Ibekwe et al. 2020)). The clear numeration of negative and positive, as mentioned earlier, is a key factor for absolute quantification by Poisson distribution.

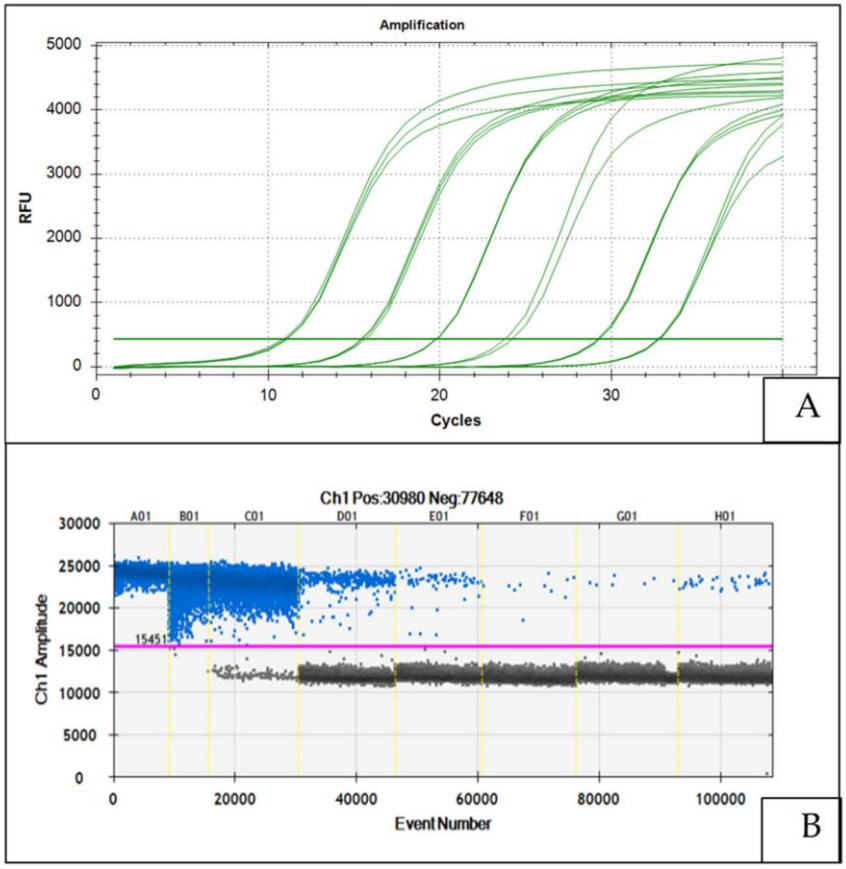


Figure 2.3. Example of qPCR(A) and ddPCR (B) output showing differences in amplification detection analysis.

2.2.3.A. The ddPCR specificity

Furthermore, for rare targets, a single ddPCR sample can be subdivided across a 96-well plate, allowing for the detection of a single target among 2 million (Hindson et al. 2011). One application of ddPCR is single-cell ddPCR (sc-ddPCR), where it is possible to detect a single nucleotide polymorphism (SNP) in an enclosed single cell (Maeda et al. 2020). The sc-ddPCR saves time and costs of the assay for detection of SNPs in mitochondrial DNA because researchers can skip cloning procedures, which takes 2–3 days to complete. The procedure of sc-ddPCR is similar to performing conventional PCR

reactions, which takes a few hours. The high-throughput ddPCR has been used in parallel comparisons with PCR for testing efficiencies of systematic evolution of ligands by exponential enrichment (SELEX) (Takahashi et al. 2016). In this study, PCR showed a bias for C/T-rich nucleotide regions when forming initial RNA libraries. More detailed comparison studies are needed because both PCR and ddPCR showed amplification bias in forming initial RNA libraries. In addition, researchers should pay close attention to the details for such verification because the sequence and structure of a target gene can influence PCR amplification efficiency.

2.2.3.B. Multiplex ddPCR

Multiplexing (targeting multiple loci within a single reaction) is another consideration for accurate and multi-target detection. Multiplex qPCR was verified to provide 100% efficiency and an R2 value of 0.99 for parasite detection, such as Giardia and Cryptosporidium (Marangi et al. 2015), and between 92.5% to 105.8% with an R2 value of >0.98 in detecting Listeria monocytogenes serotype in ready-to-eat meat products (Alía et al. 2020). Multiplex qPCR has been mainly used to identify different species within the sample, with the benefit of enabling detection as low as one organism/40 mL of soil in the case of sand flies (Giantsis and Chaskopoulou 2019). On the other hand, the use of multiplex ddPCR benefits assays of atypical transcripts such as detecting the presence of the fusion transcript BCR-ABL1, a marker for chronic myeloid leukemia (Petiti et al. 2020).

The utilization of ddPCR in multiplexing is important because such mutation and variance are hard to detect using multiplex qPCR, which is capable of targeting up to four

different targets within a single reaction (Taylor et al. 2015). Another study tracked epigenetic subtype markers using methylation-specific multiplex ddPCR (Malic et al. 2019). Application of ddPCR to multiplexing has enabled detecting methylation site variations of the target gene. The results described above suggest a greater potential of multiplex ddPCR over qPCR for application use.

This chapter explores and compares the two PCR application methods, qPCR and ddPCR, to investigate the findings in the suitability of methods to target ARGs, including specificity, benefits, and downfalls of each method from the literature. The purpose of this study is to discover suitable conditions to best utilize both qPCR and ddPCR, in an effort to reduce erroneous research results and to suggest optimized procedures for different sample types and ARGs. Our primary literature search criteria involve the keywords such as ddPCR, qPCR, resistance/antimicrobial, comparison, etc., but careful review criteria used in a comparison of qPCR and ddPCR techniques are as follows: (1) applications, (2) limit of detection and sensitivity, (3) reproducibility, (4) measurement variance, (5) cost, (6) biases, and (7) ARG detection.

2.3. Comparison of qPCR and ddPCR Methods and Their Applications

With qPCR as a conventional method and ddPCR as the emerging technique, findings in comparison of both methods in different types of studies are collected in **Table 2.1**. Most studies were done on human diseases such as HIV and viral infections as well as areas related to bacteria, plants, and food. More of the studies reported that the LoD and sensitivity are better performed in ddPCR compared to qPCR, and accuracy was up to 20 times higher at a lower limit of detection. qPCR was performed at wider ranges of gene concentrations, especially at higher concentration ranges. Some ddPCR detections were more precise and better at LoD but could not cover the broad LoQ.

There are some controversies in the use of ddPCR for genetic analysis and quantification since it is shown that ddPCR sensitivity can vary depending on the sequence of a target material, and the quantification range does not surpass the performance level of qPCR. In HIV-1 and human DNA detection studies, both ddPCR and qPCR performed similarly for the lower limit of detection on target genes 50% of the time (Henrich et al. 2012). A study done with norovirus reported that there is no advantage of choosing either qPCR or ddPCR methods if research focuses on a lower detection limit (Persson et al. 2018). Some research showed that ddPCR may have better LoD but has a limitation in detecting a high number of target genes, and both qPCR and ddPCR showed a similar coefficient of determination of standards in the gyrB gene (Porcellato, Narvhus, and Skeie 2016). Other research with bacterial genetic markers stated that while ddPCR shows better reproducibility for marker detection in fecal composites, qPCR shows a higher sensitivity for markers with environmental and composite samples with less than 10% sensitivity difference (Nshimyimana et al. 2019).

Nevertheless, both methods rely on target amplification, so certain genetic targets may work better in one or the other technology, and marker optimization is critical to both technologies.

Under the presence of inhibitors, the performance of reverse transcriptase ddPCR in the detection of pepper mild mottle virus (PMMoV) was better than that of qPCR under complex matrices of seeds, plants, soil, and wastewater, as reported by Racki N. et al. (Rački, Dreo, et al. 2014). Another study also supported this finding, that ddPCR is less sensitive to inhibitors compared to qPCR in viral detection (Rački, Morisset, et al. 2014). In addition, it was suggested that the multiplexing capability of ddPCR technology is helpful for library preparation and next-generation sequencing applications to accommodate a large number of samples while showing robustness on inhibitors (Gutiérrez-Aguirre et al. 2015).

Research cases mentioned above imply qPCR and ddPCR performances change depending on the type, structure, and initial concentration of samples. One study presented a challenge in ddPCR analysis; it is hard to find validation of the trueness of the sample reading since it is difficult to find representative samples with reference values (Deprez et al. 2016). Depending on the ddPCR setup, the reporting value may change drastically, and it is hard to compare between studies. Therefore, careful comparison studies focusing on the performances of the two qPCR and ddPCR methods need to be explored in varying areas of study.

Table 2.1. Findings in ddPCR and qPCR analysis in LoD, reproducibility, and range of detection.

Author (Year)	***Gene***	**Type**	***LOD***		***LOQ & Range***		***Reproducibility***	
			qPCR	**ddPCR**	**qPCR**	**ddPCR**	**qPCR**	**ddPCR**
Laura Cavé et al. (2016) [74] [#]	*sul1, qnrB*	ARG		+, 10-fold	+	-		

Cesare, Andrea Di et al. (2018) [75] #	*sul2, Intl1*	ARG	-	+				
Ginn O. et al. (2021) [76]	*tetA, qnrB, blaTEM, intl1*	ARG	N/A	+		*		
Kimbell L. et al. (2021) [77] #	*blaTEM, blaSHV, sul1, czcD, copA, intl1*	ARG	-	+	*	*		
Sun Y. e al. (2021) [78]	*quinolones, tetracyclines, sulfonamides, macrolides*	ARG	N/A	*				
Srisutham S. et al. (2021) [79]	*pfmdr1, pfplasmepsin2, pfgch1*	ARG	N/A		*	*		
Xu J. et al. (2021) [80]	*mcr-1, blaCTX-M-14, bla CTX-M-55*	ARG	-	+				
Yang et al. (2014) [53]	*Cryptosporidium Oocysts, 18S rRNA*	Parasite	N/A	N/A			+	+
Weerakoon K.G. et al (2016) [81] #	*S. japoricum, SjR2 and nad1*	Parasite	-	0.05fg, +		-		
	Overall	**ARG & Parasite**	-	+	+	-	*	*
Henrich T.J. et al. (2012) [66] #	*HIV-1, human CCR5 DNA*	Human Diseases	+	+, HIV-1				
Heredia N.J. et al. (2013) [82]	*HER2 (= erbB2), CEP17*	Human Diseases	N/A	+				+
Strain M.C. et al. (2013) [83] #	*HIV, episomal 2-LTR*	Human Diseases		+	+	-		+
Bharuthram A. (2014) [84] #	*CCL4L, CCL4L1 and CCL4L2 encodes HIV-1*	Human Diseases			-	+	-	+
Jones M. et al. (2014) [85]	*HIV-1 from 8E5/LAV cells*	Human Diseases	-	+	+	-, lower target		+
Coudray-Meunier et al. (2015) [86] #	*Hepatitis A, Norovirus*	Human Diseases	*	*	-	+	-	+
Taylor S.C. et al. (2015) [46] #	*H275-WT and H275Y-MUT of H1N1*	Human Diseases	+, mutant	+			-	+
Yan Y. et al (2016) [87]	*H7N9*	Human Diseases	N/A			-		
Yang Q (2017) [88] #	*PRRSV*	Human Diseases	-	+	+, false positive	-		
Link-Lenczowska D. et al (2018) [89]#	*JAK2 mutation on V617F*	Human Diseases	0.12%	0.01%,+	*	*	+	
Persson S. et al. (2018) [67] #	*norovirus GI (GI.4) and GII (GII.4)*	Human Diseases	*	*				+
Pinheiro T.F. et al (2018) [45]	*foot-and-mouth disease virus RNA*	Human Diseases	-	+				
Baume M. et al. (2019) [90] #	*Legionella DNA reference material*	Human Disease		+	*	*		
Zhang Y. et al. (2019) [91] #	*PCV3*	Human Diseases	-	+			-	+
Dong L. et al. (2020) [92]	*Tumor DNA reference material, BRAF V600E*	Human Diseases	N/A	0.02%		0.10%		+
Lin Q. et al (2020) [50] #	*ISKNV*	Human Diseases	-	+	-	+, low		
Petiti J. et al (2020) [64]	*BCR-ABL1 disease marker leukemia*	Human Diseases	N/A	0.001%				
Thwin KKM et al. (2020) [93] #	*NB-mRNAs (CRMP1, DBH, DDC, GAP43, ISL1, PHOX2B, and TH mRNA)*	Human Diseases	-	+	-	+		
	Overall	**Human Disease**	-	+	*	*	-	+
Milbury C.A. et al. (2014) [94]	*EGFR T790M, L858R*	Mutation		+				
Zhao Y. (2019) [95]	*MTRNR1-WT*	Mutation	N/A	+		+	-	+

Liu Q. et al (2020) [44] #	*CNVs causing somatic mosaicism*	Mutation	+	-	*	*	N/A	N/A
	Overall	**Mutation**	*	*		+	-	+
Burns et al. (2010) [96] #	*ERM-AD413 carries Mon810*	Plant, Food	-	+	-, lower range			
Coudray-Meunier et al. (2015) [86] #	*Hepatitis A, Norovirus*	Plant, Food	-	+			+	-, bias
Porcellato D. et al. (2016) [68] #	*gyrB of B. cereus group*	Plant, Food	-	+	+	-	*	*
Scollo F. et al. (2016) [49] #	*11C Chloroplast locus*	Plant, Food	-	+				
Wang X. et al. (2019) [47] #	*transgenic rice line TT51-1*	Plant, Food	-	+				
Demeke et al. (2020) [97] #	*Canola and soybean*	Plant, Food	*	*	*	*		+
	Overall	**Plant & Food**	-	+	*	*	+	*
Pinheiro L.B. et al. (2012) [98]	*Lambda DNA*	Bacteria, Phage		+				+
Xi Z. (2018) [48] #	*16S rRNA of Las*	Bacteria, Phage	-	+	-	+		
Sivagnesan et al. (2018) [99]	*Std1_XhoI insert with M13 E coli plasmid DNA*	Bacteria, Phage	-	+			*	*
Furuta-Hanawa B. et al. (2019) [100] #	*rAAV2RSM, rAAV8RSM*	Bacteria, Phage	*	*	-	+	-	+
Nshimyimana J.P. et al (2019) [69] #	*Bacteroidales, BacHum and B. theta*	Bacteria, Phage	+, Environmental	+, sensitivity				+, fecal
Raurich et al. (2019) [101]	*Bifidobacterium animalis (BAN)*	Bacteria, Phage	+	*	+	-		
Ahn Y. et al.(2020) [102] #	*Burkholderia epacian*	Bacteria, Phage	-	+	-, recovery	+, recovery	-	+
Ibekwe M.A. et al (2020) [58] #	*Shiga toxin-producing E. coli O157:H7*	Bacteria, Phage		+	*	*	-	+
Voegel T.M. et al. (2020) [103] #	*amoA, nirS, nirK, nosZI, nosZII*	Bacteria, Phage	-	+	+	-		
	Overall	**Bacteria & Phage**	-	+	*	*	-	+

[1] (+) represents better performance compared to the other method from the paper. (-) represents worse performance compared to the other method from the paper. The * represents similar efficiency and performance regarding criteria. The # represents literature that made direct comparison analyses of qPCR and ddPCR. Detailed comparison results are summarized and provided in Supplementary Information (SI) Excel spreadsheet. (Tables S1–S5 include Supplementary Tables S1–S5 for each type of target gene, and Table S6 includes a heatmap of all literature compared in this table.).

2.4. Current Findings on ddPCR Analysis on Genetic Targets Compared with Other Methods

2.4.1. Sensitivity

To our knowledge, general reporting on sensitivity in studies using qPCR and ddPCR methods is that the lower limit of detection of ddPCR is 10 times more effective compared to qPCR (Biorad 2018; Cavé et al. 2016; Lin et al. 2020; Raurich et al. 2019; Taylor et al. 2015; Taylor, Laperriere, and Germain 2017; Xiaofu Wang et al. 2019; Y. Zhang et al. 2019). The target genes of such studies vary from bacterial 16S RNA gene and ARGs to HIV and viral genes. Due to the diversity of target genes, the sensitivity performance resulted in a broad spectrum in quantification and detection limits. Therefore, it is asserted that optimization of the target gene will enhance detection at a lower level of genetic materials, such as viral genes that cause disease (Henrich et al. 2012). This means the optimization process of primer/probes and concentration of reagents are required for increased sensitivity, especially for ddPCR. Depending on the target of interest for study, one must optimize testing conditions to fairly compare and assess quantification capabilities of qPCR and ddPCR.

Particularly, higher accessibility and amplifiability govern ddPCR results. The use of restriction enzymes can increase accessibility for PCR amplification and detection, while fragmentation of DNA can decrease positive readings in ddPCR, although such fragmentation can reduce rain effects, where the droplet readings are located in between positives and negatives. The smaller the fragments, the greater the reduction in positive readings (Kline, Romsos, and Duewer 2016). This implies that assay characteristics,

target gene size, and GC contents are important determinants for reproducibility in quantitative analyses.

During their case study using ddPCR and comparing its performance to qPCR, Guitierrez et al. mentioned that sensitivity can vary depending on the assay characteristics, presence of inhibitors, SNPs in the probe annealing region, etc. (Gutiérrez-Aguirre et al. 2015). It is plausible that these factors will affect the testing lower limit range of genes, as increased sensitivity needs improvement in marginal errors. Additionally, Quan et al. (2018) mentioned that sensitivity or lower LoD is governed by partitioned volume, the standard deviation of the volume, and partitions of the reaction mix (Quan, Sauzade, and Brouzes 2018). This is because a single nucleic acid in a partition will be detected in the lowest amount, which can be quantified as the lower limit. For this reason, the study mentioned that ddPCR can reach higher sensitivity if the reaction volume can be adjusted (Quan, Sauzade, and Brouzes 2018) .

It was also stated that ddPCR detection may be troublesome when genes have high GC contents (Q. Liu et al. 2020). This is due to the nature of ddPCR analysis methods of encapsulating genetic materials within oil droplets; genes with higher GC content tend to have a hydrophobic site, staying intact within hydrophobic oil droplets can be unstable because the oil and higher GC region repels each other and carries the risk of incomplete encapsulation throughout amplification cycles. One study on the detection of Leishmania infection using ddPCR found that ddPCR may not be suitable for the assay, as the target sequence has more than 50% higher GC content (Ramírez et al. 2019). A review study on the diagnosis of leishmaniasis reported that qPCR reached a lower detection limit of 10 pg, equivalent to ~120 parasites, in insect vectors, which

verifies the method is precise in quantification and identification of species (Galluzzi et al. 2018). Noticing the differences in methods, curated planning for more effective evaluation and verification is recommended in studies handled with ddPCR and/or qPCR analysis. Each method has its differences in strength and weakness, and proper application of these methods will bring significant developments in the sensitivity of targeting genetic matter and lessen the noise of erroneous results. Importantly, it brings to our attention that assay-specific performances can differ depending on reaction volume, reaction mixes, and target nucleotide content.

2.4.2. Dynamic Range of Detection and Measurement Variance

In most cases, qPCR has a broader dynamic range of quantification, especially higher than 4 or 5 log units of concentration, compared to ddPCR (Strain et al. 2013; Q. Yang et al. 2017). The degree of dynamic range seems to change depending on the types of the gene and analysis settings (Pinheiro et al. 2012; Porcellato, Narvhus, and Skeie 2016; Q. Yang et al. 2017). This may be coming from the number of partitions that ddPCR can make, which is up to 20,000 droplets in the case of ddPCR. When the DNA concentration increases more than the number of partitions it can be in, there is no possible number of negative droplets to conduct the Poisson distribution analysis. Then, the analysis may be no longer valid. Theoretically, based on 95% confidence intervals, the partitioning error significantly increases when the average number of targets per partition, λ, is outside of the range of 0.001 to 65.38, where λ equals the sample concentration (C) multiplied by partition volume (n) (Basu 2017). To prevent this overpopulation of positive over negative droplets, researchers need to dilute the

concentrated samples before processing by measuring the DNA/RNA contents to make sure the experimental plan is suitable for ddPCR application.

While the concentration of the target gene is important for absolute quantification based on Poisson distribution calculations, % coefficient of variance (%CV), a calculated parameter of standard deviation/average*100, may vary depending on the dilution factor. From ddPCR and qPCR assessments, the %CV of the same dilution series ranges from 0 to 8.26, and 1.45 to 12.37 for qPCR and ddPCR, respectively (Taylor, Laperriere, and Germain 2017). Such variances may be coming from inaccurate dilution of residual protein and chemical contaminants (Taylor, Laperriere, and Germain 2017). Digital PCR (dPCR) is regarded as a technique that will have less variance compared to qPCR if it is free of any upstream errors derived from sampling and extraction processes (Huggett, Cowen, and Foy 2015). With the development of technology, having a higher number of partitions of ddPCR will minimize partitioning errors, and much improvement in the precision of detection can be expected.

Multiple studies proved that qPCR can be applied to various types of genes with a broader range of quantification. The qPCR performance is quite competitive compared to ddPCR. For example, the range of quantification was reported to be similar or better for qPCR depending on samples, where Cave et al. (2016) reported a broader quantification range compared to ddPCR (Cavé et al. 2016). Zhang Y. et al. (2019) reported that ddPCR and qPCR readings were highly correlated, by 95% (Y. Zhang et al. 2019).

In the case of the lower limit of detection, ddPCR seems to perform either similar or better with detection. Both ddPCR and qPCR showed 95% similarity in finding the LoD on a norovirus gene (Persson et al. 2018) and showed similar detection capabilities

on genetically modified canola and soybean genes (Demeke, Beecher, and Eng 2020). Depending on the gene types, ddPCR can detect down to 1 log unit better compared to qPCR (Cavé et al. 2016; Deprez et al. 2016; Nshimyimana et al. 2019; ZHONG et al. 2018). To our best knowledge, studies with qPCR showed LoD and LoQ as low as 2 log units and 3 log units, respectively (Raurich et al. 2019; ZHONG et al. 2018). For overall ranges of concentration where both methods can be used, ddPCR shows a positive correlation to qPCR results at an R-value around 0.85 (J. Liu et al. 2019) and shows high linearity (Y. Zhang et al. 2019).

2.4.3. Reproducibility

Reproducibility of an analysis method is important, especially when assaying samples that have low concentrations or are highly unstable, such as RNA. It is particularly important to reduce human error that could result from pipetting, cross-contamination, RNA degradation, or other factors. Establishing an internationally recognized reference system for achieving consistent results between laboratories is critical to ensure that analyses are comparable and informative (Kline, Romsos, and Duewer 2016). To develop good practice, consistent DNA or RNA extraction methods are critical to obtaining trustworthy and reproducible data (Scollo et al. 2016). In addition, the importance of quality and purity control of the target genetic material is stressed enough for testing reproducibility (Demeke, Beecher, and Eng 2020).

The ddPCR platform provides automated robotic droplet generation and sample mix preparation that can filter out human biases. In a study using QX200 automated ddPCR for developing miRNA markers, authors were able to determine the miRNA copy number repeatedly ($p < 0.05$) (Vishnuraj et al. 2021). If there are issues observed with the

expected reproducible rate, one may check if the target genetic concentration is at a proper range, as described in the previous section of measurement variance, because too high or low partitioning may increase errors in measurements. Additionally, proper optimization steps for each method are essential because an approach to detect the same target sample showed varying results without optimization (Sivaganesan et al. 2018).

Overall, both ddPCR and qPCR measurements are highly correlated in many different studies, consistent with high experimental reproducibility and repeatability for both (Arvia et al. 2017). In one study, ddPCR measurements showed better reproducibility for quantification in fecal composites, while qPCR showed a higher reproducibility for environmental and composite samples (Nshimyimana et al. 2019). In another study, ddPCR achieved a higher reproducibility for specific species such as P. falciparum, while it showed similar sensitivity to qPCR in *P. vivax* (Koepfli et al. 2016). Although ddPCR appears to be slightly more reproducible (Taylor, Laperriere, and Germain 2017), and the automated platform can reduce human error, the improvement in reproducibility does not appear to be significant.

2.4.4. Cost

The cost-effectiveness of ddPCR is one of its main concerns (Nshimyimana et al. 2019; Nyaruaba et al. 2019). It is reported that ddPCR costs two times as much as qPCR (R. Yang et al. 2014), and its lesser availability has prevented the distribution of ddPCR technology in developing countries since its introduction in 2011 (Nyaruaba et al. 2019). In addition, despite Sanger sequencing and ddPCR having their own benefits and irreplaceability

in microbiological analyses, when compared, ddPCR is capable of detecting low-frequency mutations better than Sanger sequencing, while the cost to detect mutations was higher in ddPCR (Zhao 2019). For microbiological applications that require direct knowledge of the DNA sequence of the microbe that extends beyond the locus, assaying unknown mutants, or developing primers and probes for detection, Sanger or high-throughput sequencing technologies cannot be replaced. However, if the mutant is known, ddPCR or qPCR can be a cost-effective method for microbiological identification.

In the case of multiplex ddPCR, the cost differential becomes less pronounced. Multiplex ddPCR allows for multiple assays within a single assay (at least four) and does not need standards, while qPCR will require multiple reactions with standards (Link-Lenczowska et al. 2018). The ability to target multiple samples concurrently can level out the higher cost of the ddPCR reaction. High-throughput ddPCR with automated analysis settings also reduces the processing, labor time, and cost for analysis [60]. Development of single-use, low-cost injection molds for ddPCR have begun to be put into use in field applications (Malic et al. 2019). Over time, there is a possibility that cost differences between qPCR and ddPCR methods may be decreased with improvements in technology in reagents and consumables production, increased availability of primers and probes, multiplexing, and price balancing by competition within the markets.

2.4.5. Risk of Bias in qPCR and ddPCR

Although the high correlation and linearity between qPCR and ddPCR methods are true (Petiti et al. 2020; Y. Zhang et al. 2019), there is detection bias in the qPCR method, which stems from its analysis method. The qPCR technique requires standards

developed before sample analysis, which makes the assay very dependent on the reading of the standards. One study on bacterial plasmids in feces reported that quantification errors in the concentration of original standards may lead to reading biases in qPCR (Sivaganesan et al. 2018). In addition, qPCR tends to be more affected by contaminants such as SDS and heparin compared to ddPCR (Pinheiro-de-Oliveira et al. 2018b). It is plausible because ddPCR separates nucleic acids into single droplets, which prevents the reaction from being interrupted by contaminants that may be captured in a different droplet or in the oil emulsion. The absolute quantification of genes makes ddPCR more independent of systematic errors of standard curves, unlike qPCR (Nshimyimana et al. 2019). Therefore, quantification methods using qPCR will need well-developed standard and analysis methods put in place for justifiable analysis and comparison between samples. When the standard and/or sample purity is compromised, test results can be biased even when other types of optimization factors such as primer design, temperatures, and reaction volume are considered.

The ddPCR is not immune to the risks of bias. As discussed in the sensitivity section above, one should consider finding the right types of sample genes with GC contents less than 50%, understanding the physico-chemical nature of ddPCR reactions, and adapting the known information of reaction specificity of primers to genes, temperature cycles, and the ratio of positive/negative counts for fine outcomes. If possible, testing with known positive standards with ddPCR may be necessary to develop a robust analysis on specific target genes. While knowing different types of genes may have different ranges of concentration to suit ddPCR analysis, researchers can pinpoint

experimental conditions and record optimized settings for the prosperity of future research.

2.4.6. Applicability

While ddPCR is used in more and more gene expression analysis, still, qPCR has been the main method used in the field. Transitioning to different technology may require standardization and verification efforts, and this would probably be the cause for qPCR technology being mainstream in analysis, despite ddPCR offering higher sensitivity and better view of gene amplification environments. Application of ddPCR in gene expression analysis was explored in cases of clinical study and detection of targets of low abundance (Heredia et al. 2013; Taylor, Laperriere, and Germain 2017). It would be best used in situations where there are no well-developed positive standards available, or the target gene is present in low concentrations, or when detection attempts using qPCR are not successful. Since ddPCR does not require standards for each run, it is suitable for analysis of many samples of low abundance with fewer preparation steps, which facilitate efficiency and productivity (Giantsis and Chaskopoulou 2019).

2.5. Current Status in ARG Detection Methods with ddPCR

To our best knowledge, as reviewed in this study, there are only a few studies that have analyzed ARG using ddPCR and qPCR methods in comparison. More comparisons of detection methods on ARGs are needed for precise quantification, better understanding of ARG, and optimized conditions for target genes. The conventional methods used for quantitative measurement of ARG in the environment are the following: (1) PCR, (2) real-time PCR (rtPCR) or quantitative PCR (qPCR) for genetic target materials, (3) ddPCR, (4) flow cytometry methods, and (5) cell counting for in vitro cultures of bacteria samples (R. Yang et al. 2014). Out of these methods, ddPCR is the only method that does not need positive controls for quantification, and it is free of bias originating from the quality of the positive control.

By far, most ARG analyses are done with qPCR, and are coupled with sequencing methods such as Illumina, to comprehensively investigate the relationship of the microbiome of targeted environments to the spread of ARG. Some research has been done on the impact of nanoparticles on ARGs and surrounding microbial communities in the estuary (Y. R. Chen et al. 2019), analysis of metagenomic correlation of AR in anaerobic digestors (Fujimoto, Carey, and McNamara 2018), and ARGs in bioaerosols of municipal sewage (Gaviria-Figueroa et al. 2019). One study applied ddPCR and 16S rRNA sequencing to track ARGs, MGEs, and bacterial compositions in the air of composting plants. The reported detection of ARGs ranged from 1 to 7 log units (Gao et al. 2018); within this range, a lower LoD is hard to find if qPCR is used.

Many researchers have been searching for a better method for precise detection and efficiency in such analyses, but each detection method has different limitations.

There is a need for more studies to be done on absolute quantification of ARG using droplet digital PCR (ddPCR) and comparison of techniques on differences in specificity for optimization, the suitability of measurement, and the possibility of usage in broader fields. If optimized appropriately, ddPCR has the potential to be applied in many different fields with greater efficiency and sensitivity.

2.6 Materials and Methods

2.6.1 Sample Collection and Preparation

Water and wastewater samples were collected in different stages of Mallard Creek Wastewater Treatment Plant (MCWWTP) and other environmental locations such as a creek, lake, river, and groundwater. The detailed location of sample collection within the MCWWTP is shown in **Figure 2.4**, and the list of locations that each sample collected is described in detail in **Table 2.2**. Samples were collected and promptly transported to the lab in a cooler with ice packs. Upon arrival to UNCC, samples were stored at 4 ℃ processed within 24 hrs for sample preservation and DNA extraction. The sample DNA was extracted manually using QIAamp DNA Mini kit (QIAGEN, Cat # 51304) following the manufacture's protocol, in elution volume of 80 µL. The extracted DNA concentration was measured using Nanodrop One microvolume UV-Vis spectrophotometer (Thermo Fisher Scientific, Agawam, MA) and Qubit 2.0 fluorometer (Thermo Fisher Scientific, Agawam, MA) to was stored in a -20 ℃ fridge until future analysis.

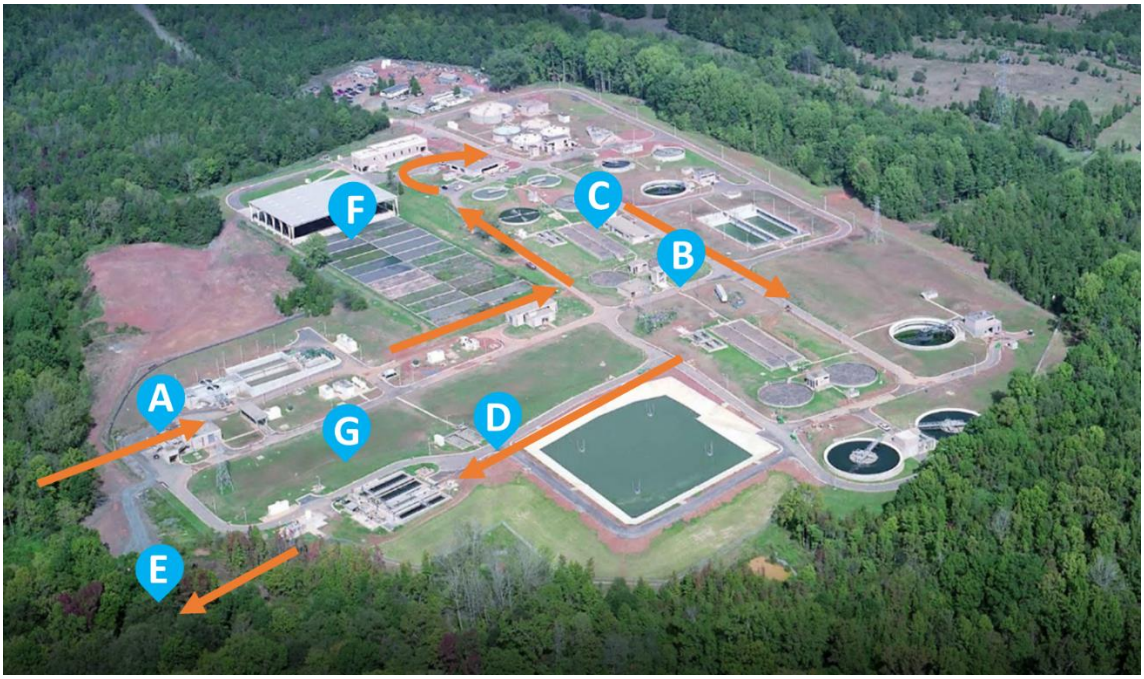


Figure 2.4. Sampling locations in Mallard Creek Wastewater Treatment Plant

Table 2.2. Sample location details from MCWWTP & other water source

Sample #	Sampling Location
A	Influent
B	Raw Sludge (after 1' Clarifier)
C	Aeration Basin Effluent
D	2' clarifier->filter
E	Effluent
F	Biosolids
G	Site Blank
H	Creek
I	River1
J	River2
K	Groundwater
L	Lake

2.6.2. Optimization of ddPCR assay

Primers sequences of target genes (forward and reverse) are shown previously in **Table 1.1**. The QX 200 Droplet Reader (Bio-Rad Laboratories, Inc.), Automated Droplet Generator (Bio-Rad Laboratories, Inc.), C1000 Touch thermal cycler (Bio-Rad Laboratories, Inc.) were used in this study. Depending on the number of samples including duplicates and blanks to run, total reaction volume was calculated. Reaction mix using Evagreen Supermix (Bio-Rad Laboratories, Inc., #1864034) was prepared in an autoclaved 2mL centrifugation tube with an appropriate volume of reagents needed as shown in **Table 2.3** as an example of calculation (9 reactions).

Table 2.3. **ddPCR Reaction Mix**

Number of RXNS	**Components**	Volume per 1RXN (uL)	Total volume of #RXN (uL)
9	**Evagreen Mix (2X)**	11.25	101.25
Total	**Forward Primer**	0.3375	3.0375
	Reverse Primer	0.3375	3.0375
	DI Water	9.225	83.025
	DNA template	1.125	10.125
	Total	22.275	200.475

On a 96-well plate, sample DNA was inserted carefully using only one side of the well. Then reaction mix was distributed evenly to each well preventing the formation of bubbles on an ice bed. The plate was sealed with an aluminum cover the centrifuged for 10 seconds for homogenous mixing and removal of air bubbles. Then the plate was incubated over an ice bed for 3 minutes. Next, an automated Droplet Generator (Bio-Rad Laboratories, Inc.) was set up for cartridges, pipette tips, and power. A clean, empty 96-well plate was inserted on an ice block and put in place for receiving droplet mixture after

droplet generation. The sample plate was inserted and operated following the manufacturer's manual. Once the generation is done, the 96-well plate with droplets of samples was removed and sealed with an aluminum foil cover. The mixture was placed into a PCR thermal cycling machine set up at the appropriate cycle and annealing temperature for amplification. After the PCR run is complete, the rack was removed and placed into the QX 200 Droplet Reader (Bio-Rad Laboratories, Inc.).

The sample plate was set up and saved using Quantasoft software (Bio-Rad Laboratories, Inc.) and sample reading was initiated. The analyzed sample file was recorded with an appropriate name. Target DNA, thermal cycling, gradient amplification of PCR, various ranges of primer, and DNA concentrations were explored to find optimum detection and droplet separation.

2.6.3. Comparison analysis of qPCR and ddPCR

Verification of target ARGs was done by application of qPCR analysis to ddPCR. Selected gene targets for ARGs as the analysis used in this study are shown in **Table 2.4**. The same set of primers, target gene standards, and samples was used for comparison analysis. In the case of ddPCR, the concentration of environmental samples was diluted to meet the detection range. qPCR amplification was done with original concentration for all environmental samples.

Table 2.4. **Target ARGs for qPCR and ddPCR methods comparison**

Target Gene	***Primer Set***	***Primer Sequence***	***Product Size (bp)***	***Annealing Temp. ('C)***
TetA	tetA-F	GCTACATCCTGCTTGCCTTC	210	*55*

	tetA-R	CATAGATCGCCGTGAAGAGG		
ereA	ereA-F	AACACCCTGAACCCAAGGGACG	420	52
	ereA-R	CTTCACATCCGGATTCGCTCG		

To improve detection of target genes, the following options were tested to find optimum conditions for ddPCR amplification: (i) temperature-gradient amplification with primer sets designed, (ii) pairing of different primer sets, (iii) use of high-fidelity DNA restriction enzyme, EcoRI. Especially, the use of restriction enzymes may help excise unnecessary genetic coils within the standard plasmid. However, the effect of the use of restriction enzyme digestion was not clear and deemed ineffective.

2.6.5. Test of Environmental samples

Extracted DNA from samples was measured using Nanodrop One microvolume UV-Vis spectrophotometer (Thermo Fisher Scientific, Agawam, MA) and Qubit 2.0 fluorometer (Thermo Fisher Scientific, Agawam, MA) to detect to the lowest possible measurement. Then the sample DNA was diluted to meet the optimum level for ddPCR analysis. For qPCR, original extracted DNA concentration was used for qPCR analysis within detection ranges set up for standard target gene analysis as described in Chapter 1.

2.7. Results

2.7.1. Optimization of target gene measurement

Testing of EcoRI restriction enzyme was designed to cut off extra gene that might have attached to tetA out of tetA -carrying plasmid DNA but did not clearly showed improvements in detection of ARGs using ddPCR. Rather, the chopped fragments were too short that all droplets were clustered together which made it harder to distinguish negatives and positives. In case of the target DNA of long lengths, using of restriction enzymes will improve the separation of droplets, but having shorter length of target DNA (210 bp for tetA) did not improved the separation at higher concentration as shown in Figure 2.X.

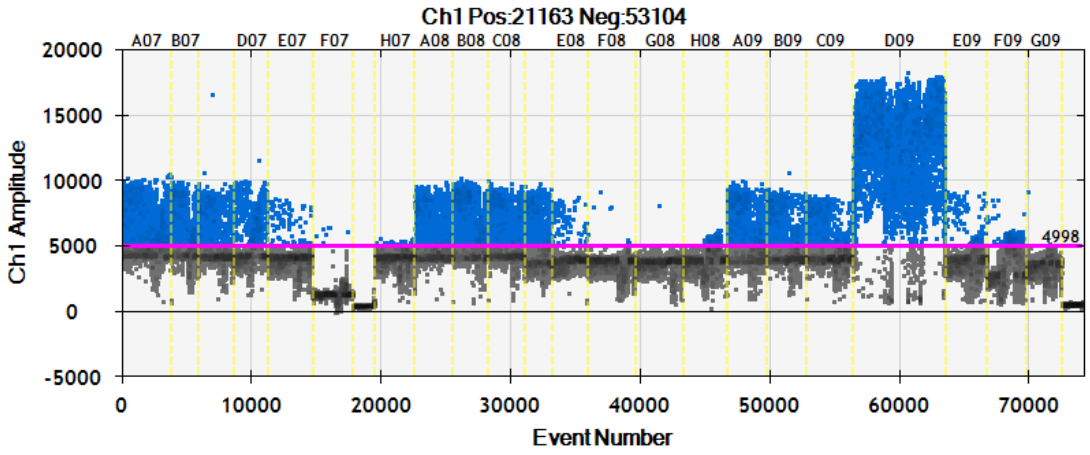


Figure 2.5. EcoRI digested tetA gene using designed primer sets

In order to optimize primer concentration and primer sets, four different combination of primer sets were tested using two different types of forward and reverse primers. The result is shown in **Figure 2.6.** Based on the figure, set of FR and F1R1 seems to have most separation of positive and negative droplets in reading. Blanks for each test seems be trustworthy with very little or no positive droplets. For further analysis, F1R1 pair was

selected based on its separation. Also, the better separation was achieved by further diluting the sample DNA, by lowering the populated number of droplets in a reaction. This helped improving separation issues that was observed previously, and also showed close link to finding measurable range of concentrations for target genes, because ddPCR tend to have sensitive and narrow range of concentration window for trustworthy detection.

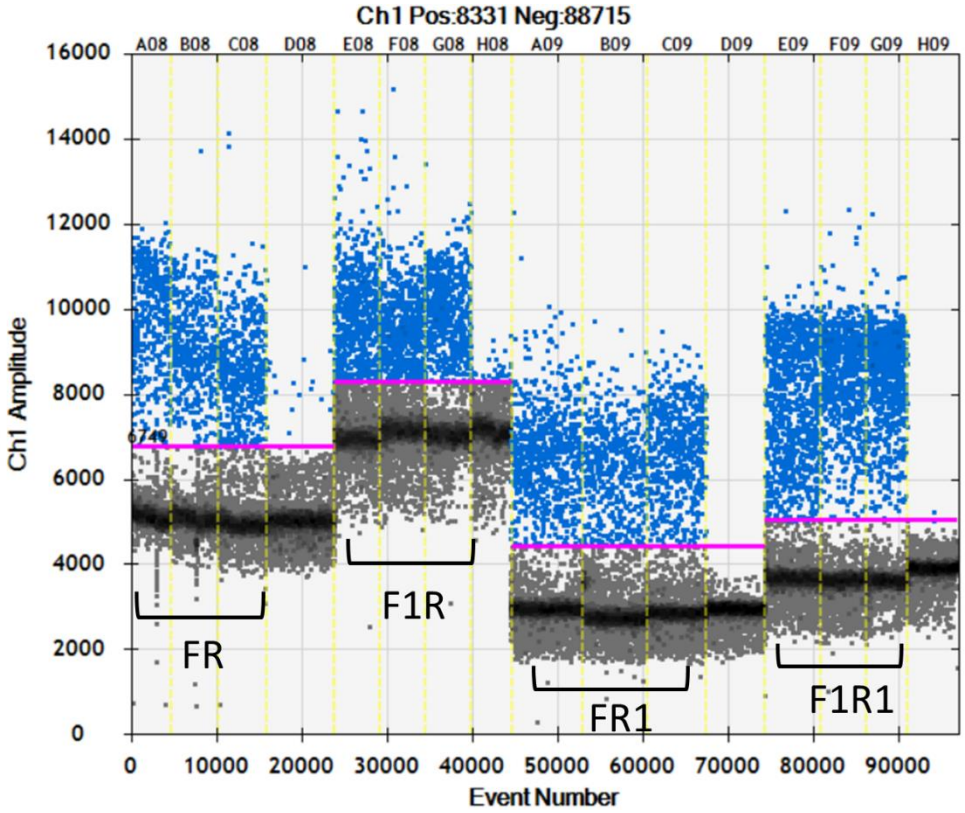


Figure 2.6. Primer combination testing for tetA

2.7.2. Limit of Detection and Limit of Quantification Analysis for ddPCR

Trustworthy detection of target ARGs requires investigation of the sensitivity of target genes to the reaction mix, volume, and the specific methods used between ddPCR and qPCR. Knowing the range of detection is important for ddPCR before the measurement

of samples. This is because the concentration range for detection varies depending on the nature of the reaction mix and target gene composition. The serial dilution prepared with standards of tetA and ereA was measured using a ddPCR setting with an optimized primer set to find the LoD, LoQ, and linearity in ARGs measurement within the measurable range.

To determine the limit of detection (LoD) for qPCR, extensive dilutions up to 16 were tested using qPCR. The stretch towards lower concentration of standard was showing amplification bias and curving of quantification cycle (Cq) values which represents a limitation of qPCR analysis on target genes of lower starting quantity. At the limit of amplification detection, the trend line of standard samples was detected in a trend that has been distorted from the original straight line (**Figure 2.7**).

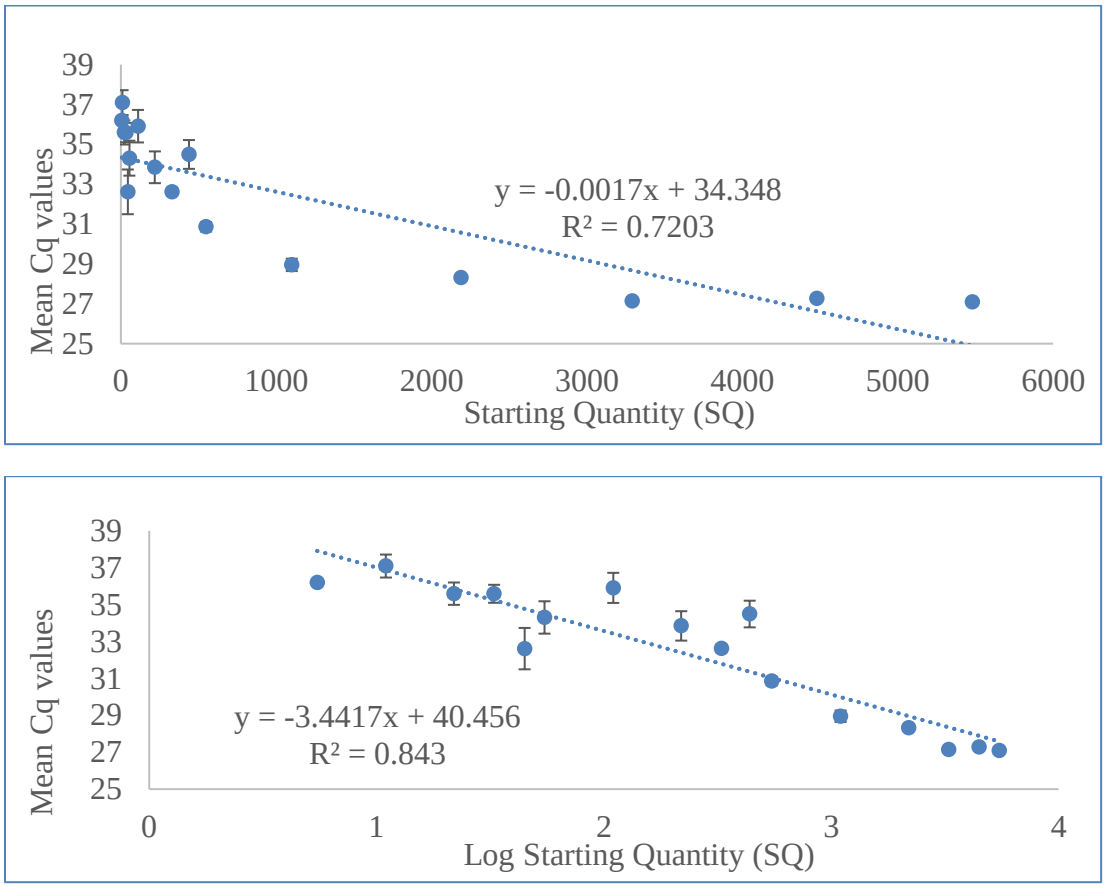
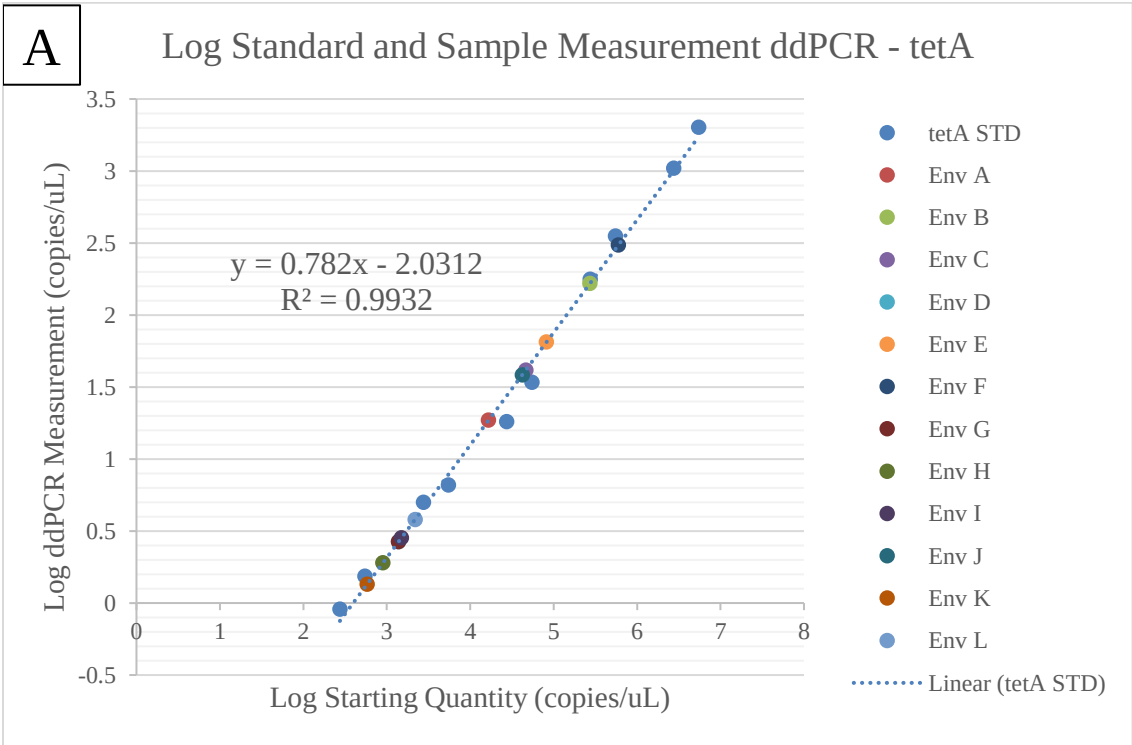


Figure 2.7. Extrapolation of standards dilution measurement for tetA. The top (A) shows the measurement represented in original concentration, and the bottom (B) shows concentration at log level.

2.7.3. Comparison of qPCR and ddPCR technologies on target ARGs

The comparison of ddPCR and qPCR for tetA, showed linearity in the trendline of both qPCR and ddPCR, with a similar R^2 value of 0.9932 and 0.9776, respectively. The result of qPCR and ddPCR are shown in **Figure 2.8**.



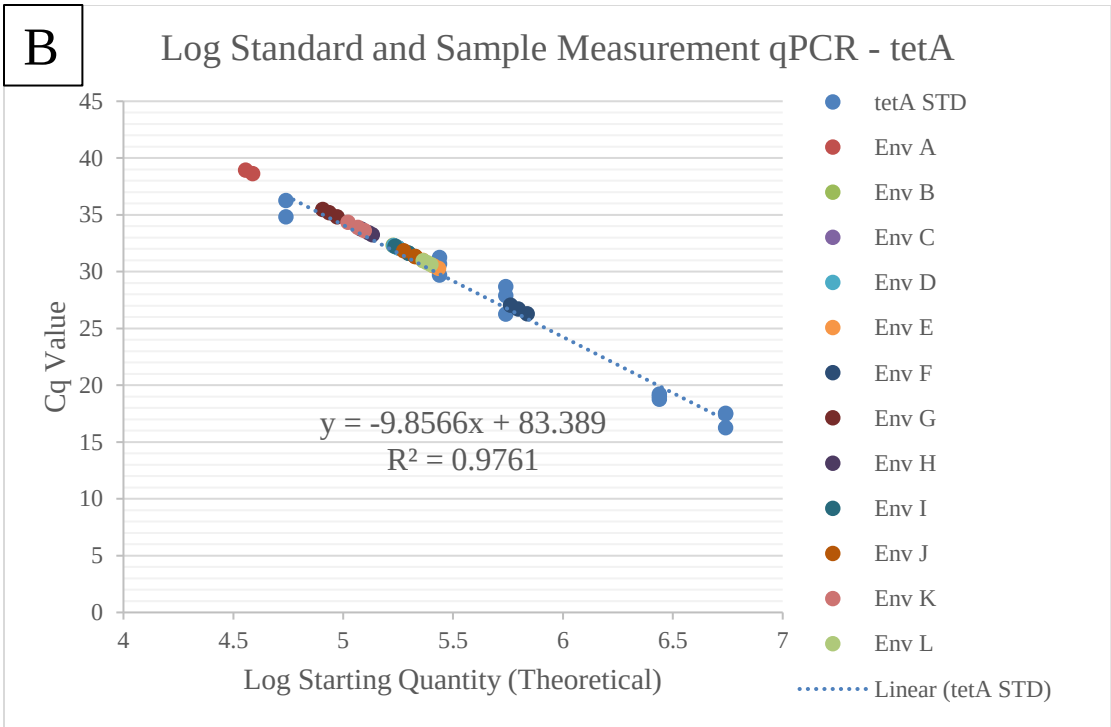


Figure 2.8. Comparison of qPCR and ddPCR on tetA gene. (A) shows the result of ddPCR, and (B) shows the result of qPCR

While ddPCR was able to detect tetA down to 1 log unit, qPCR had a limit of detection at log 3 unit, where the standard curve trendline was curved at that point. The compatible range of detection for tetA for qPCR and ddPCR seems to be ~4 to 6.7 log units, and 2.5 to 7.5 log units for each method, respectively. This showed that ddPCR is better at detecting lower limit of tetA, while qPCR can cover roughly 0.5 log units higher concentration range. In terms of the range of detection, while ddPCR lost its quantification capacity at lower and higher concentrations outside the ranges detected, qPCR was able to detect higher and lower concentrations outside the confirmed range. This means the stretching of the standard line may be possible for qPCR, but ddPCR

requires careful dilution or concentration steps would be required if a sample of interest has a higher or lower range of target gene concentration.

In the case of ereA, the gene was found to be detectable at log 1 to 8 units using qPCR. However, some environmental samples measured were at the lower limit, which the current measurable standard was not able to cover all of the sample concentration range at the lower range as shown in **Figure 2.9**.

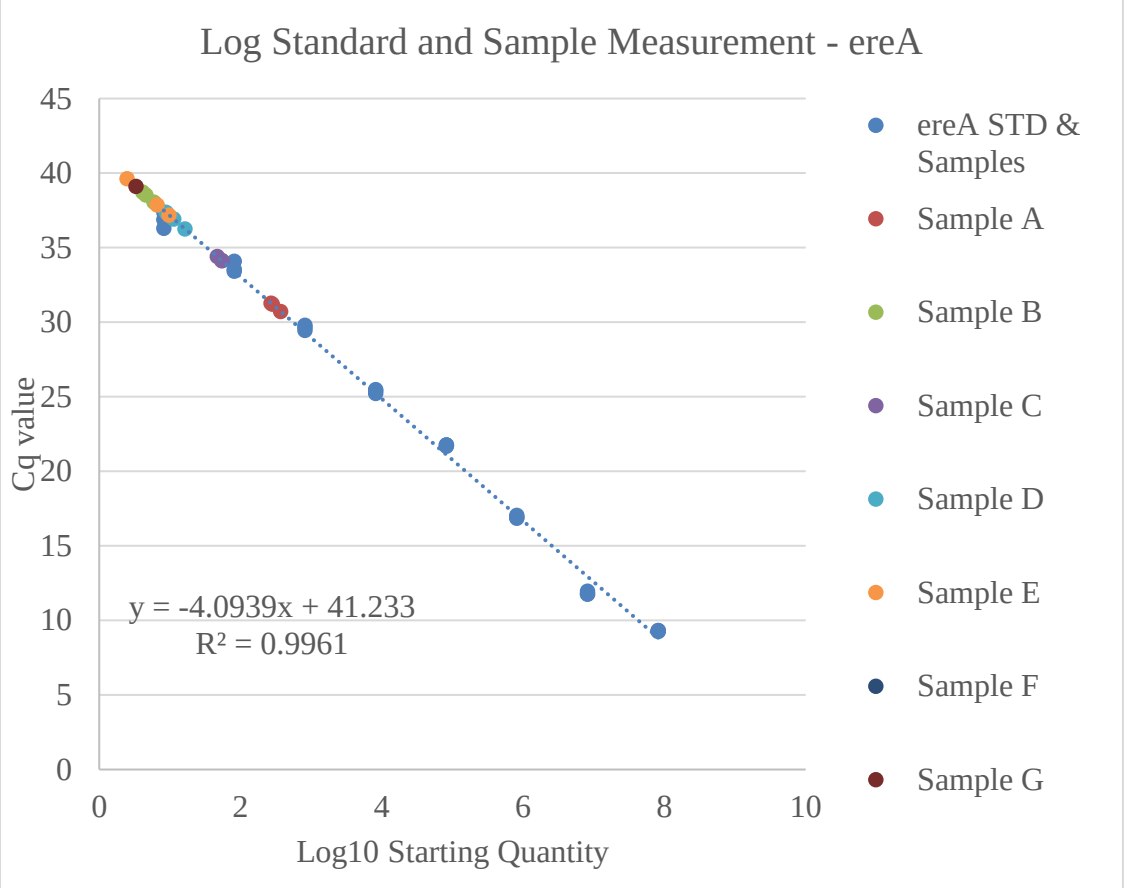


Figure 2.9. Standard and environmental sample measurement on ereA gene

2.8. Discussion

While having multiple issues in operating a droplet generator, some of the earlier data before having errors showed that tetA and ereA genes were possible to be identified in ddPCR with comparative quality to qPCR. In most cases, the level of copy number detection was more sensitive for ddPCR than qPCR while qPCR has a capability to quantify genes at higher concentration range. At lowest dilutions of standard, qPCR showed amplification-detection bias where the regression curve was skewed to detect a higher Cq value than a theoretical trend. This may be due to the limit of detection of fluorescence of the instrument because the simplification signal is too weak where there is not much of the target gene available. Comparative study of ereA is still ongoing.

2.8.1. ddPCR as the Future

ddPCR is reported to show a lower detection limit with 1 log unit better in LoD than qPCR (Scollo et al. 2016; ZHONG et al. 2018); a greater number of samples tested positive (Y. Zhang et al. 2019), and it has better sensitivity compared to qPCR when it is optimized (Deprez et al. 2016; Raurich et al. 2019; Taylor et al. 2015; Taylor, Laperriere, and Germain 2017; Y. Zhang et al. 2019; ZHONG et al. 2018). For this reason, ddPCR has the potential to be applied to a broader and more generalized field, which is supported by a review paper (Quan, Sauzade, and Brouzes 2018). Therefore, depending on its use, ddPCR can be applied to many studies improving the sensitivity of the measurement, determining fine variances between sample sequences, or analyzing samples with low initial DNA concentration.

The unknown outcomes and variances of ddPCR efficiency depending on sample types, DNA/RNA concentrations, gene structures, and assay methods can be defined as the “case specificity” of ddPCR. While having the merit of broad applicability for PCR, ddPCR needs more fine-tuning than qPCR to find the true performances matched to each target analysis. Collective knowledge from more research is needed to discover the advantages and unknown characteristics of ddPCR technology.

ddPCR is a promising method with better performance in detecting transgenic components due to its stability, accuracy, and resistance to PCR inhibitors, and there is no need for reference materials (Xiaofu Wang et al. 2019). If studied in careful experimental design and comparison analyses, there will be numerous potentials to be used in the future. The current global situation of the pandemic on viral pathogens and the increasing amount of ARGs promote the need for a faster, more efficient, and more precise analysis method for genetic materials. The occurrence of pathogens in the environment needs to be tested and studied in a timely manner with continuous mutation/evolution and adaptation of microbes to the ever-changing environment. Careful implementation of new technologies such as ddPCR may bring benefits with its properties of absolute quantification and sensitivity.

2.9. Conclusions and Perspectives

The consensus observed from the literature and this study's results is that qPCR has a strength of having broader detection range of genetic materials, especially at higher range of concentration. The literature review highlighted the benefit of having lower costs, and better specificity to some target genes over ddPCR. The ddPCR method, on the other hand, showed an enhanced lower limit of detection for both target ARG in this study and many different sample types in the literature. It is considered as a powerful tool for other biological assays such as mutation tracking. While both methods require optimization steps, ddPCR requires more optimization to develop a robust analysis reviewing the steps that were described earlier.

Transferring the exact qPCR primers to ddPCR may be possible without losing amplification efficiency if the general requirements of primer design criteria are met, such as length of the target gene, GC contents of binding sites, and avoidance of secondary structures. Additional considerations recommended for ddPCR reaction by the industry include avoiding repeats of G/C longer than three bases, addition of GC repeats at the 3' end of primers, and design of 50–60% of GC content in the target region (Biorad 2018). Experimental conditions such as concentrations of a target gene, primers, and probes (if necessary) may need to be changed depending on the optimization process of detection. If the result of ddPCR is not as expected, the reading may show the potential presence of impurities in the reaction mix, or that the detection setting is not optimized.

The detection and quantification of ARGs require more studies on different types of ARGs to find the suitability of each method on specific types and to determine the optimal settings for a reaction (e.g., DNA concentration, GC contents, reaction volumes,

partitioned number of reactions, and suitability of primers/probes to the analysis). Although there are only a few studies that have employed ddPCR to detect ARGs, ddPCR shows similar or better LoD and sensitivity predictions when compared to qPCR (Table 1). ddPCR excelled in detection for human diseases and viral genes and given time, we expect ddPCR could be a more effective solution for ARG tracing and quantification in the environment.

Recent studies have shown increasing accumulation of ARGs in municipal wastewater treatment plant effluent (Mao et al. 2015)(3–5 log unit copies/mL) as well as in the air (Gao et al. 2018); thus, better surveillance and treatment methods are needed. To track ARGs' spread, peak, and attenuation, the technical power to detect lower gene concentrations is important. Clinical and environmental samples are expected to be found at lower concentrations when compared with cultured samples. Therefore, the ddPCR gene quantification method can be more suitable in the following situations: (1) when more precise detection is required, such as genetic mutation detection; (2) for new types of strains or genes that do not have reliable positive controls; and (3) limited amounts of genetic material can be extracted from collected sample.

However, researchers who choose to use the ddPCR technique must not ignore that ddPCR can have varying results, which can be influenced by the experimental setup and process, such as annealing temperature and threshold (Gerdes et al. 2016). Therefore, the optimization step is necessary, and optimization settings should be adjusted to a level that is most suitable for the characteristic of the target gene. Without this optimization, ddPCR will not outperform qPCR.

Despite its merits, the use of ddPCR still holds some disadvantages, such as higher operation cost, reagent costs, and availability, when compared to conventional methods (Q. Liu et al. 2020; Zhao 2019). Despite these limitations, the high accuracy and resolution of ddPCR have led it to be widely used in food sciences (Cao et al. 2020; Floren et al. 2015; Gerdes et al. 2016; Naaum et al. 2018; Noh et al. 2019; Persson et al. 2018; Pinheiro-de-Oliveira et al. 2018b; Porcellato, Narvhus, and Skeie 2016; Scollo et al. 2016; Xiaofu Wang et al. 2019; Y. Zhang et al. 2019; ZHONG et al. 2018)and disease and evolution studies (Baume et al. 2019; Deprez et al. 2016; Dong et al. 2020; Gao et al. 2018; Henrich et al. 2012; Hulme 2017; Jones et al. 2014; Koch, Jeschke, and Becks 2016; Kyae et al. 2019; Liao et al. 2019; Milbury et al. 2014; Nshimyimana et al. 2019; Pinheiro-de-Oliveira et al. 2018b; Ram et al. 2019; Raurich et al. 2019; Sivaganesan et al. 2018; Talarico et al. 2018; Überbacher et al. 2019; R. Yang et al. 2014). These studies are continually developing the technology and creating an extensive collection and recording of the optimization process for different genes. In the future, with more and more usage, one can organize a database of ddPCR genes and their optimum conditions depending on primer sets, temperature, and probes. This information will help future researchers to reach the best output in a shorter timeframe and at a lower cost.

The global community has experienced a pandemic where a novel pathogen has threatened global health and the economy. Detecting, tracking, and forecasting these microbes and their genes in nature is very hard due to the different characteristics of species. The development of biotechnologies that enhance genetic detection methods with more precision and sensitivity is required. Studies show that with proper optimization steps and verification with positive and negative controls, ddPCR analysis

will allow researchers to capture biological information that could have been missed in conventional methods. There is a strong case for more ddPCR application studies to be done with many other gene targets because these data will contribute to better usage of ddPCR technology and boost future work in this field.

CHAPTER 3: Tracking the Source and Anthropogenic Effects of Antibiotic Resistance Genes (ARGs) in the U.S. Water System

3.1 Introduction

In this chapter, the overall ARG distribution, and quantification throughout three different water systems across the U.S. The definition used as "water system" in this chapter refers to a set of different water body types such as a river, lake, wetland, underground aquifer, including effluents of wastewater treatment plants. Environmental buffers such as a lake, river, wetland, and groundwater aquifer are generally perceived as an attenuator of anthropogenic contaminants. However, the examination to seek their specific roles and effect posed on newly emerging contaminants such as over-the-counter pharmaceuticals and antibiotic-resistant genes (ARGs). Amongst all, growing antimicrobial resistance (AMR) decreases antimicrobial specificity and activity of frequently used drugs, causing the rise of multidrug-resistance (MR) and their genes which threatens public health. Historically, anthropogenic effects increased ARGs, AMR, and MR in the water environments such as rivers, lakes, and ocean (B. Chen et al. 2013; Posada-Perlaza et al. 2019; Pruden et al. 2006; Pruden, Arabi, and Storteboom 2012a).

3.1.1. Origin and Transmission of Antibiotic Resistance in the Environment

Antibiotics are identified as one of the contaminants of emerging concerns (CECs) according to EPA (U.S. EPA 1985). Antibiotics are frequently used to support the treatment of disease, enhance crop yields in agriculture, and prevent diseases in livestock farming. Since the initial application of antibiotics, the amount we need to overcome the

current emergence of new antibacterial species has been increasing over time. This can be due to natural selection where the fittest survives amongst bacteria, which cause human to apply more amounts of antibiotics to meet the expected effects on antibiotics.

With the increasing use of antibiotic use due to human activities, there were increasing concentration of antibiotics and corresponding antibiotic-resistant bacteria (ARB) and antibiotic-resistant genes (ARGs) detected in the U.S. waterways (Gottlieb and Nimmo 2011; Pruden et al. 2006; Pruden, Arabi, and Storteboom 2012a). The use of antibiotics in livestock production such as beef, chicken, and pork intensified the susceptibility of single and multiple drug resistance (Pruden et al. 2013). Inadequate disposal and overuse of over-the-counter antibiotic medicines have contaminated the streams and underground water to accumulate its concentration over time. The natural decay process of pharmaceuticals seems slower than the rate of input and production within our natural system. Remaining amount of antibiotics creates environment for natural bacterial strains to develop resistance over time. The overview of the source and transport of antibiotic resistance is demonstrated well in **Figure 3.1.**

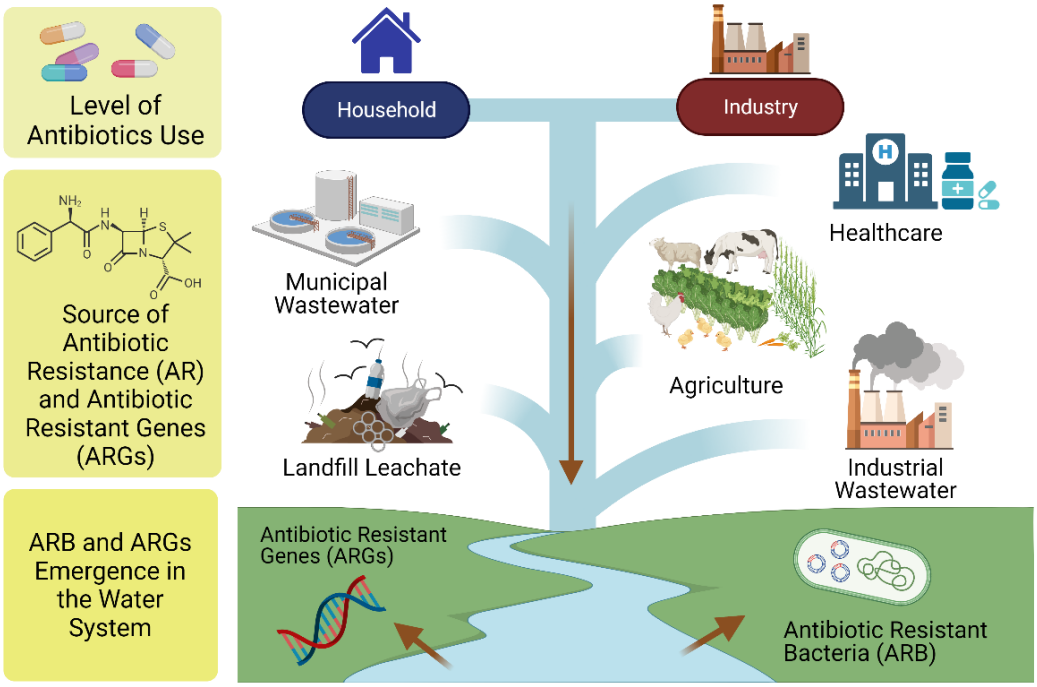


Figure 3.1. An overview of the source, transport, and emergence of antibiotic resistance in the water system.

In order to adapt to this environment, bacteria create defense mechanisms: Inactivation of the foreign chemical, developing pumping out mechanism, modification of the chemical by changing its bond with enzymes, or create impermeability with cell wall so that the antibiotic does not enter inside the bacterial cell (Thermo-Fischer 2016). Then the bacteria with antibiotic resistance will share its evolved genetic element either with horizontal gene transfer or vertical gene transfer (Grabow, Prozesky, and Smith 1974). Horizontal gene transfer can be done in three ways, transformation, transduction, and recombination (Griffiths et al. 2000; Lederberg and Tatum 1946; Parkinson 2016). These paths make the genetic element possible to spread out more quickly. Szczepanowski et al. showed horizontal gene transfer capability of bacteria by recombination capability of bacterial plasmid DNA which carries two different ARG

with implication that mobile genetic element may present in unknown regions of the plasmid (Szczepanowski et al. 2009). As many of the bacterial genetic elements are not fully understood and discovered, bacterial HGT and development of AMR is endless with the current scope of our knowledge.

Growing AR is identified as a growing human health security threat (Peak et al. 2007) which requires global surveillance and attention to decrease (WHO 2014). Also, multidrug resistance to serious diseases such as tuberculosis and malaria became harder to control if AR intervenes in disease progression (WHO 2014). In the near future, some antibiotics may no longer be effective enough to help humans to fight back from disease and bacterial infections due to AR development. Finding the origin of AR with human fecal contamination and differentiating it with livestock contamination, one study tracked two different bacteriophages that are specific to the human gut and animal feces, respectively (Karkman, Pärnänen, and Larsson 2019). Although not being efficient in tracking due to low detection of animal fecal origin, the study showed a positive correlation between anthropogenic origin with the ARG abundance while the selection hotspots seem to be in the sediments contaminated with industrial-level antibiotics pollution (Karkman, Pärnänen, and Larsson 2019). Many other researchers have been working on finding an adequate proxy for human fecal origin to illuminate the path of ARG amount in the environment (Gillings et al. 2014; Gravel, Messier, and Roy 1998).

This chapter investigates the overall ARGs distribution and quantification throughout the selected three different water systems across the U.S. The definition used as “water system” in this study refers to a set of different water body types such as a river, lake, wetland, underground aquifer, including effluents of wastewater treatment

plants. Different types of water bodies such as a lake, river, wetland, and groundwater aquifer are generally perceived as an attenuator of anthropogenic contaminants. However, the examination to seek specific roles of water bodies and the effect of anthropogenic wastewater posed on the expression of ARGs is limited. Amongst all, growing AR decreases antimicrobial specificity and activity of frequently used drugs, causing the rise of multidrug-resistance (MR) and their genes which threatens public health. Historically, anthropogenic effects increased ARGs, AR, and MR in the water environments such as rivers, lakes, and ocean (B. Chen et al. 2013; Posada-Perlaza et al. 2019; Pruden et al. 2006; Pruden, Arabi, and Storteboom 2012a).

3.1.2. ARGs Surveillance and Mitigation in the Water Environment

Studies showed that ARG concentration can increase after storm events with stormwater runoff, while fecal indicators such as E. coli and enterococci were not a significant indicator of ARG of stormwater origin(Garner et al. 2017). Peak et al. showed lagoons with high-exposure tetracycline showed ARG 3 log units greater compared to those that are not (Huang et al. 2019). The overuse of antibiotics was considered to be a source of emerging AMR in the environment (Gottlieb and Nimmo 2011). To mitigate the overuse and oversupply inducers, there was economic modeling study finding interdependency of factors to reinforce sustainable antibiotics use (Lhermie et al. 2019). Such modeling analysis utilized Drivers-Pressure-State-Impacts-Responses (DPSIR) framework, while responses can act a role as a policy indicator that helps prevention, mitigation, restoration, and adaptation of AMR related issues.

According to an FDA report, antibiotics used in 2017 for food-producing animals were over 10000 tons including over 5000 tons of medically important antibiotics (USFDA 2018). In 2016, congress approved the allocation of financial resources for antibiotic resistance solutions initiatives to detect, respond to, and contain emerging resistance (Center for Disease Control 2018). In order to improve antibiotic use, first tracking of the current state of antibiotics in the environment is needed before taking action. With tracking of ARG from pristine water source to a water body with the gradual influence of human activities, sul1 showed a strong correlation with human activities with ARG amount in both bed and suspended sediments, while tetW broadly spread out in the pristine region and more in animal feeding lagoons (Pruden, Arabi, and Storteboom 2012b).

The specific objectives of this study are to (i) quantify ARG concentrations and investigate how ARG quantities vary under different types of water environments, (ii) determine how each water body in the water system have been impacted by the effluents of wastewater treatment plants (WEFs), and (iii) investigate what other anthropogenic and geographical factors influencing the emergence and spread of ARGs on each site. This study is unique because water systems in three different locations are investigated for ARG concentrations by comparing the trend in different water bodies while anthropogenic and geographical factors are considered for ARG expression analysis. Target genes under different types of drug classes were analyzed using qPCR analysis in three different U.S. water systems. Detection of ARGs in various water bodies was compared for water connectivity and anthropogenic influence in the surrounding region. Also, this study may suggest an efficient water connection and treatment system on

ARGs, and the learnings can benefit future AMR surveillance, containment, and control processes.

3.2. Trend of ARGs in Different Water Bodies of the Environment

3.2.1. River

ARG types expressed and attenuated differently depending on environmental factors of being in different water or soil system such as rivers, lakes, underground aquifer, and drinking water or wastewater treatment plants. It was found that the river was serving as a reservoir of ARGs such as beta-lactam resistance genes and mobile genetic elements (MGEs) such as intl1 genes, which was identical to *E. coli* isolates from swine, poultry, and humans (Kim, Kang, and Lee 2008). Another study found that tetW was broadly distributed throughout the river region regardless of pristine and other regions while sul1 was highly correlated to landscape features (Pruden, Arabi, and Storteboom 2012b).

3.2.2. Lake

In the freshwater lake, treated wastewater effluent was identified as a point source, showing up to 200 fold concentration difference in ARGs such as sul1, sul2, tetB, tetM, tetW, and qnrA compared to other regions of the lake that is further away from the point of wastewater effluent discharge (Czekalski, Gascón Díez, and Bürgmann 2014). Another study showed that tetA and sul1 genes are positively correlated with dissolved oxygen while negatively correlated with chlorophyll, while the lake is serving as a reservoir of AR (Di Cesare et al. 2015).

3.2.3. Wetland

Multidrug resistance was found most abundant in the natural wetland sediment, and the progression of ARGs is slow in the wetland that did not have any recent anthropogenic influence (Tang et al. 2021). In the case of a constructed wetland, ARGs detected at the starting point were higher when compared to the ARG quantities detected at maturation,

although relative abundance was increased at effluents (Abou-Kandil et al. 2021). Also, ARG removal was more effective during summer than winter (Abou-Kandil et al. 2021), which conveys microbial activities related to ARGs is better performed during warmer weather.

3.2.4. Underground aquifer

A meta-analysis review study pointed out that the concentration detected in groundwater may not be a human health risks, but consumption of the water may develop AR in humans, and some species of ARGs present in groundwater is putting moderate to low risks to aquatic species (Zainab et al. 2020). The spread of sulfa antibiotics was studied in soil and groundwater, and the ability was higher in the order of sulfamethoxazole, sulfamethyldiazine, sulfamethazine, and sulfathiazole(J. Wang et al. 2016). The study implies that specific sulfa antibiotics are able to travel faster through the water underground and can cause potential risk in development in ARGs. Another study also found that the contamination of tetracyclines and fluoroquinolones penetrated deeper in the alluvial sediment underground, up to 20m (Tong et al. 2017).

3.2.5. WWTPs

In case of the wastewater treatment plants (WWTPs) and their effluents, while the treatment process seems effective, the effluent can also serve as a point source of ARGs when it meets natural water body. Out of 140 clinically relevant ARGs, 123 of them were detected in effluents which included multidrug efflux and small multidrug resistance gene (Szczepanowski et al. 2009).

The **Table 3.1** shows the list of literature reported on antibiotics and ARGs detection in various types of water bodies.

Table 3.1. Summary of literature investigating on emergence of antibiotics and ARGs in different water body types

Category	ARGs or Antibiotics Detected	Method Used	Key Findings	Author
River	N/A	16S rRNA gene sequencing	This study suggest that homogenization of bacterial species happens in streams, and WWTP may contribute to loss of diversity and spread of antibiotic resistance	Clinton S. et al., 2020
River	nine Abt (sulfapyridine (SPD), sulfamethazine (SMZ), sulfadimethoxine (SDM), chlortetracycline (CTC), oxytetracycline (OTC), tetracycline (TC), ciprofloxacin (CIP), norfloxacin (NOR), and roxithromycin (ROX)) and corresponding ARGs (tetA, tetB, tetC, tetM, qnrA, qnrS, gyrA, sul1, and 16S rRNA)	HPLC-MS/MS, qPCR	Relative abundance of qnrA and tetC higher than other ARGs. Heavy metals such as Ni, Cr, and As positively correlated with some ARGs, promoting expression of ARGs	Huang Yu-Hong et al., 2019
River, surface water	ten Abt (sulfamethoxazole, trimethoprim, ciprofloxacin, cephalexin, levofloxacin, amoxicillin, clindamycin, doxycycline, ertapenem, and azithromycin) and ARGs	UHPLC-MS, Shotgun Metagenomic Sequencing (illumina HiSeq 2500)	Abt concentration detected to be highest at final clarifier effluent (FCE). Hospital and residential wastewater showed higher concentration of antibiotics as well as total average ARGs compared to upstream of the water body, showing contribution in increase antibiotics and ARGs in the WWTP influent.	Lambirth K. et al., 2018
River	sul1, tetW, 16S rRNA	qPCR	sul1 and tetW showed different trend throughout the water system where sul1 is carried with intl1, which is increasing with landscape use towards downstream while tetW is more related to animal feeding location but showing no correlation within the watershed.	Pruden Amy et al., 2012
River	B-lactamase resistance (CTX-M-14, TEM-52, and CMY-1), intl1 genes: aadA1, dfr12-orfF-aadA2, and dfr17-aadA5	PCR, Sequencing with Taq DyeDeoxyTerminal cycle-sequencing kit, and pulsed-field gel electrophoresis (PFGE) of *E. coli* isolates	This study suggests that river can act as a reservoir for ARGs and ARBs	Kim J. et al., 2008
Lake	tetA, tetB, tetM, bla TEM, bla CTX-M, bla SHV, strA, strB, qnrA, qnrS, sul2, and sul3	qPCR (rtPCR)	tetA and sul1 gene positively correlated with DO and negatively correlated with chlorophyll. Lake is serving as a reservoir of AR	Cesare A.D. et al., 2015
Lake	sul1, sul2, tetB, tetM, tetW, qnrA	qPCR (rtPCR)	compared to central lake, location near sewage discharge were up to 200fold above levels in ARGs detection	Czekalski Nadine et al. 2014
wetland	16S rRNA, lacZ, intl1, sul1, tetM, ermB (macrolide resistance gene) vanA	qPCR on target ARGs, Metagenome sequencing with Illumina HiSeq 2500	Most ARG belong to bacterial effluex transporter superfamilies. Removal pattern of ARGs were not related to seasonal effect. Heterotrophic bacteria, fecal coliforms, e. coli, 16S rRNA and lacZ gene increased with operation time	Abou-Kandil A., 2021
wetland	N/A	metagenomic analysis	multidrug was the most abundant type, largely by soil samples. Slow process of succession of ARGs in the mudflat possibly due to alkaline pH environment and no anthropogenic activities	Tang Xingyao et al., 2021
Underground	N/A	16S rRNA Sequencing, Illumina Miseq	Groundwater affecting coastal influence with its connectivity showed microbial community resembling ocean microbiome. Actinobacteria found in abundance in	Moore A. et al., 2020

			freshwater but not in hotel discharge. Bacteroidetes present in all inland and coastal sites	
Underground	Review		sulfonamide and tetracycline ARGs are most common in groundwater. The ARG concentration varies depending on country and region.	Zainab S. M. et al., 2020
Underground	Abt (sulfadiazine (SDZ), sulfathiazole (STZ), sulfamerazine (SMR), sulfamethazine (SMZ), sulfamethoxazole (SMX), norfloxacin (NOR), ciprofloxacin (CIP), enrofloxacin (ENR), lomefloxacin (LOM), oxytetracycline (OTC), chlorotetracycline (CTC), doxycycline (DC), clarithromycin (CTM), azithromycin (AZM), roxithromycin (RTM), and erythro- mycin (ETM))	Microwave accelerated reaction system, UHPLC-HRMS	Deeper migration and adsorption tetracyclines and fluoroquinolones into sediment layers, while most of antibiotics has contamination in sediments under 8m depth	Tong L. et al., 2017
Underground	sulfa Abt: sulfatiazole (ST), sulfamethyldiazine (SM), sulfamethazine (SM2), sulfamethoxazole (SMX)	HPLC, UV-vis spectrophotome try	The migration order highest to lowest: SMX>SM>SM2>ST. Fe(3+), Mn(2+), iron oxide, manganese oxide and humic acid aid adsorption of sulfa antibiotics	Wang J. et al., 2016
Underground	Review	Managed Aquifer Recharge (MAR)	Managed Aquifer Recharge (MAR) systems showed efficient removal of antibiotics as the distance and residence time increases.	Maeng S. et al., 2011
WWTP	N/A	Illumina MiSeq 16S rRNA sequencing	ARG concentration was highest at effluent, downstream, then upstream. MGEs positively correlated with ARG density and diversity. Sul1 may be indicator of resistome shifts downstream of water.	Raza Shahbaz et al., 2021
WWTP	192 resistant genes	PCR, sequencing using Millipore	Detected aminoglycoside, β-lactam, chloramphenicol, fluoroquinolone, macrolide, rifampicin, tetracycline, trimethoprim, and sulfonamide resistance gene & multidrug efflux, and small multidrug resistance gene. Demonstrated genetic exchange between clinical and WWTP bacteria	Szczepano wski Rafael et al., 2009
WWTP, Urban stream	Antibiotic resistant bacteria	Illumina HiSeq 16S rRNA sequence analysis	bacterial compositions change throughout the treatment process and ARB counts reduced in downstream and effluent of the wastewater treatment plant	Sorgen Alicia, A. Fodor, C. Gibas., 2021
WWTP, river	13 ARGs and 7 gene markers for facultative pathogenic bacteria (FPB)	qPCR	blaTEM, ermB, tetM, and sul1 genes were detected in all samples. Meeting the WWTP effluent, bla CTXmM, bla oxa-48, bla KPC-3, mcr-1, mecA, and CMY-2 detected for strong or last-resort antibiotics. Downstream was highly correlated with ammoniacal nitrogen and WWTP effluent while upstream was correlated with turbidity, SS, seasonal factors- UVA radiation, macrophytes.	Reichert Gabriela et al., 2021
WWTP-R	intl1, bla_TEM, ermF, mecA, tetA, and 16S rRNA	qPCR on ARGs and 16S rRNA	WWTP significantly reduced ARGs and intl1 in wastewater. The WWTP was still discharging 3.3 log unit of total ARGs after treatment	Thakali Ocean et al., 2020
WWTP-sludge	285 ARG types, 10 MGEs, 16SrRNA gene.	Ct value of 31 used as cutoff	ARGs were at lower abundance in summer and winter compared to other seasons ($p<0.05$, Kruskal-Wallis test)	Xu Sai et al., 2020

*Acronyms: Abt, antibiotics.

3.3. Materials and Methods

3.3.1. Sampling Locations

Three different sites across the U.S. were chosen for analysis. Each site has a unique combination of various types of water bodies such as rivers, lakes, wetlands, underground aquifers, and merging points of drinking/wastewater treatment plants (DWTP and WWTP). Site A contains water flow which starts from a river with an inflow of effluents from the wastewater treatment plant, then to constructed wetland, and the downstream of the river is feeding into an advanced purification system which is run by the county along with the influent of effluents from another wastewater treatment plant. Then treated water is fed into the underground aquifer, then the water body is connected to monitoring well. Site B has a similar layout of a river that has three different wastewater plants' effluents feeding upstream, then a part of the river is fed into an underground aquifer, then to an advanced purification system. Site C large lake flows down into a smaller lake with an influent of wastewater plant effluent, then from the smaller lake to a river. The second lake water is fed into a drinking water treatment plant. The location of the three water systems were marked in **Figure 3.2**.

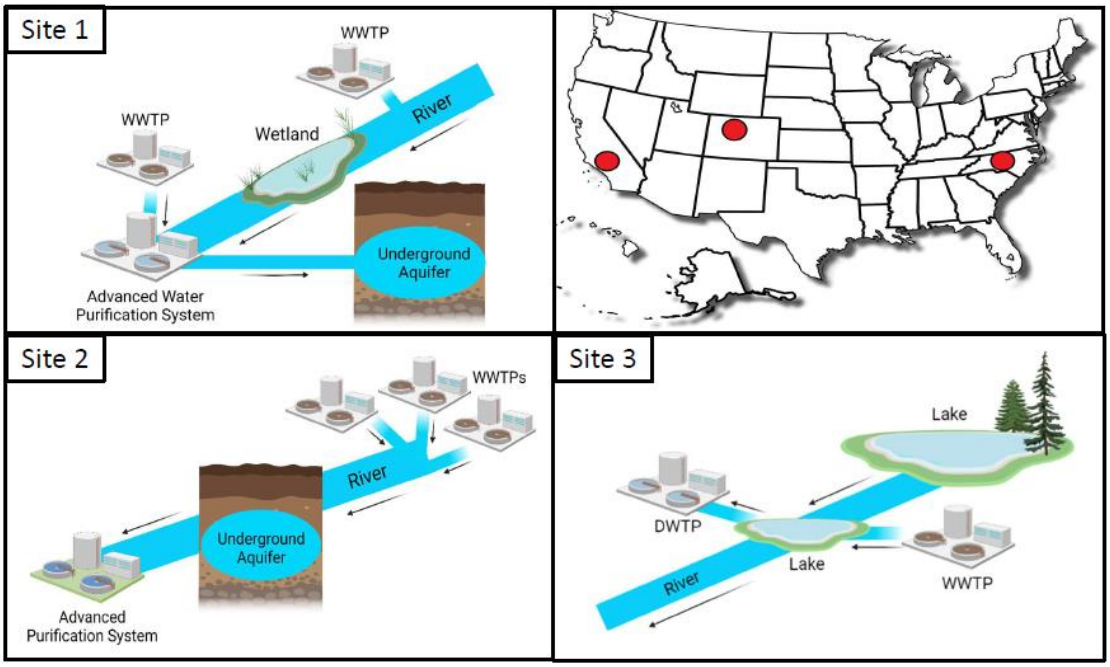


Figure 3.2. **Sampling locations of three water systems across the U.S.** The top right shows the location of the water system in the U.S. map, and starting from top to bottom, each box shows detailed connectivity of water bodies and water treatment systems laid out.

The acronyms of each sample are shown in **Table 3.2.** Samples are called based on an acronym for better representation and understanding later in this chapter.

Table 3.2. Acronyms for samples at three sites

Site #	Location	Acronym
1	River before wetland	**River1**
1	constructed wetland	**Wetland1**
1	WWTP effluent	**WEF1**
1	advanced water treatment effluent	**AWEF**
1	post underground recharge	**PUR1**
2	WWTP effluents	**WEFs**
2	river receiving the WWTP effluents	**River2**

2	post underground recharge	**PUR2**
3	WWTP effluents	**WEF3**
3	Lake	**Lake**
3	DWTP influent	**DWIn3**

3.3.2. Sample Collection and Concentration

Samples were collected on-site and shipped to the University of North Carolina at Charlotte (UNCC). Due to the geographical distance of sites to the research facility, the residence time varied depending on the best possible available transportation methods. Due to the complexity of sites and water sampling locations, date and time samples were collected were recorded, and also wet and dry dates were noted for better comprehension of the result. Borosilicate amber bottles with PTFE lined caps were used to store water samples, and aliquots for microbiological test volume were stored in autoclaved Nalgene bottles. Samples were shipped with ice packs through overnight shipping to ensure low temperature to prevent drastic chemical changes in the natural characteristics of the samples. In the case of the samples which can be accessed through driving, some lake samples at site C were collected by a graduate student and transported the same to be on ice packs as the rest of the samples. Samples were immediately processed upon delivery, by concentrating a large volume of water using filtration to concentrate bacterial contents to result in high DNA yields. If not, the original water samples were stored in a -80 °C freezer to preserve nucleic acids and bacterial composition.

3.3.3. DNA Extraction

Depending on the condition of the experiment, sample DNA was obtained through either automatic DNA extraction or through a manual extraction method. For manual extraction, sample DNA was extracted using QIAGEN Mini DNA Extraction Kit, and the automatic method used QIAamp DNA Mini Kit with the QIAcube instrument (QIAGEN).

3.3.4. Target Gene Selection and Standard Preparation

Specific types of ARGs across five different types of antibiotics were selected based on their occurrence in the environment through extensive literature review. Standards were developed following the steps described in Chapter 1. Targeted genes were cloned into competent cells and cultured to extract synthesized plasmids which have target ARG sequences to serve as a positive standard. The types of ARGs are selected as shown in **Table 3.3**.

Table 3.3. List of target ARGs used in this study

Target Genes	Antibiotics
tetA, tetW	Tetracycline
qnr A, qnr B	Quinolone
bla CTX-M	Beta-lactamase
sul 1, sul 2, and sul 3	Sulfonamide
ereA	Erythromycin / macrolide

Additionally, the 16SrRNA gene was selected to quantify the total amount of bacterial community for the normalization of ARGs data. Also, two types of water pathogen, *Cryptosporidium parvum* and *Giardia Lamblia*, positive standards of both species were donated from Charlotte Water and the DNA of each target was extracted to serve as a positive standard. DNA extraction process was followed the same as sample DNA

extraction. All other types of ARGs went through cloning and verification procedure to develop positive standards as described previously.

3.3.5. ARG Quantification Using qPCR

For amplification reaction mix, the list of the following primers shown in **Table 3.4** are used for identifying each ARGs. The SsoAdvanced Universal SYBR Green Supermix (BIO_RAD) was used for entire analysis of ARG quantification. Each sample was prepared in triplicates with 7-point serial dilutions of standard and negative control for statistically reliable qPCR run. The reaction mix recipe is presented in **Table 3.5**.

Table 3.4. List of target ARGs and primers

Target gene	Primer	Primer sequence (5′—3′)	Product size (bp)	Annealing Temp.(°C)	Reference
16S rRNA	1369F	CGGTGAATACGTTCYCGG	123	56	[47]
	1492R	GGWTACCTTGTTACGACTT			
tetA	tetA-F	GCTACATCCTGCTTGCCTTC	210	55	[34]
	tetA-R	CATAGATCGCCGTGAAGAGG			
tetW	tetW-F	GAGAGCCTGCTATATGCCAGC	168	60	[40]
	tetW-R	GGGCGTATCCACAATGTTAAC			
Sul1	Sul1-F	CGCACCGGAAACATCGCTGCAC	163	55.9	[41]
	Sul1-R	TGAAGTTCCGCCGCAAGGCTCG			
Sul2	Sul2-F	TCCGGTGGAGGCCGGTATATGG	191	60.8	[41]
	Sul2-R	CGGGAATGCCATCTGCCTTGAG			
Sul3	Sul3-F	TCCGTTCAGCGAATTGGTGCAG	128	60	[41]
	Sul3-R	TTCGTTCACGCCTTACACCAGC			
qnrA	qnrA-F	TCAGCAAGAGGATTTCTCA	516	50	[42]
	qnrA-R	GGCAGCACTATGACTCCCA			
qnrB	qnrB-F	TCGGCTGTCAGTTCTATGATCG	469	54	[42]
	qnrB-R	TCCATGAGCAACGATGCCT			
ereA	ereA-F	AACACCCTGAACCCAAGGGACG	420	52	[43]
	ereA-R	CTTCACATCCGGATTCGCTCG			

blaCTX-M	blaCTX-M-F	ATGTGCAGYACCAGTAARGT	593	50	[46]
	blaCTX-M-R	TGGGTRAARTARGTSACCAGA			
Giardia lamblia	β-Giardin P241 -F	CATCCGCGAGGAGGTCAA	74	58	[48]
	β-Giardin P241 -R	GCAGCCATGGTGTCGATCT			
Cryptosporidium parvum	COWP P702 -F	CAAATTGATACCGTTTGTCCTTCTG	150	66	[48]
	COWP P702 -R	GGCATGTCGATTCTAATTCAGCT			

Table 3.5. Reaction mix preparation recipe for qPCR

Components	Volume (uL)
SYBR Green	**10**
Forward Primer	**0.4**
Reverse Primer	**0.4**
DNA template	**2**
DI Water	**7.2**
Total	**20**

3.3.5.A. Plate and Software Setting

Plate design for qPCR analysis of samples was done using Bio-Rad CFX Manager software (BIO-RAD). The User-defined run type was selected to set the thermal cycling steps. The annealing temperature is set differently depending on the target genes and primer set as shown in **Table 3.4**. Sample volume per well was set at 20 µL for all reactions. The general thermal cycling setting is the same as described in Chapter 1. Once the protocol is set, plate assignment of wells was arranged in the order of standards on a 1st column, then triplicates of samples and blanks on the rest of the columns.

3.3.5.B. Reaction Mix Preparation

The qPCR reaction mix was prepared with 10 µL of SYBR Green, 0.4 µL of forward and reverse primer, 2 µL of DNA template/sample, and nuclease-free water to meet the total volume of 20 µL. The reagents ratio of reaction mix followed manufacturer's manual of SsoAdvanced Universal SYBR Green Supermix (BIO_RAD).

3.3.5.C. qPCR Operation, Quantification, and Verification

CFX96 Touch Real-Time PCR Detection System with C1000 Touch (BIO-RAD) and Bio-Rad CFX Manager 3.1. were used for the qPCR thermal cycling and analysis. C1000 Touch was turned on, and thermal cycling protocol for qPCR analysis was set up to be 95 °C for 3 minutes, 39 cycles of 95 °C for 15 seconds and 30 seconds at annealing temperature, then temperature gradient to 95 °C with 0.5 °C increments for 5 seconds, and 4 °C for 3 minutes.

Preparation of reaction mix was operated in the biosafety hood with disinfection of surfaces before and after the preparation procedure. The prepared reaction mix is then loaded in 96-well plates and sealed with transparent film before putting into the thermal cycler. Using Bio-Rad CFX Manager 3.1, target genes were quantified by detecting the fluorescence signal during amplification. Based on a standard curve of each plate run, Cq values of each sample were converted to copy numbers of the target gene.

In the case of the analysis which provided no standard amplification, a new standard was developed through cloning procedure and target gene quantification of such failed samples were redone until they pass the statistically sound number of samples in one run.

(i.e. at least five-point of trustworthy standards with standard curve R^2 value over 0.98, and at least two of three triplicates showing a similar range of quantification). All qPCR result values were converted back to report the ARGs and pathogens concentration in the unit of copies/100mL.

3.4. Results & Discussion

3.4.1. Antibiotic Resistant Genes (ARGs) in the Water System

3.4.1.A. Tetracycline Resistant Genes: tetA and tetW

Two tetracycline-resistant genes, tetA and tetW were quantified in each sample. The results are shown in **Figures 3.4** and **3.5**. The tetA was detected the highest in WWTP effluent of site 1, and tetW was detected the highest in WWTP effluents of site 2. Based on variances between the site location, the concentration and trend of ARG concentration of tetA are dependent on the site location rather than the types of water body.

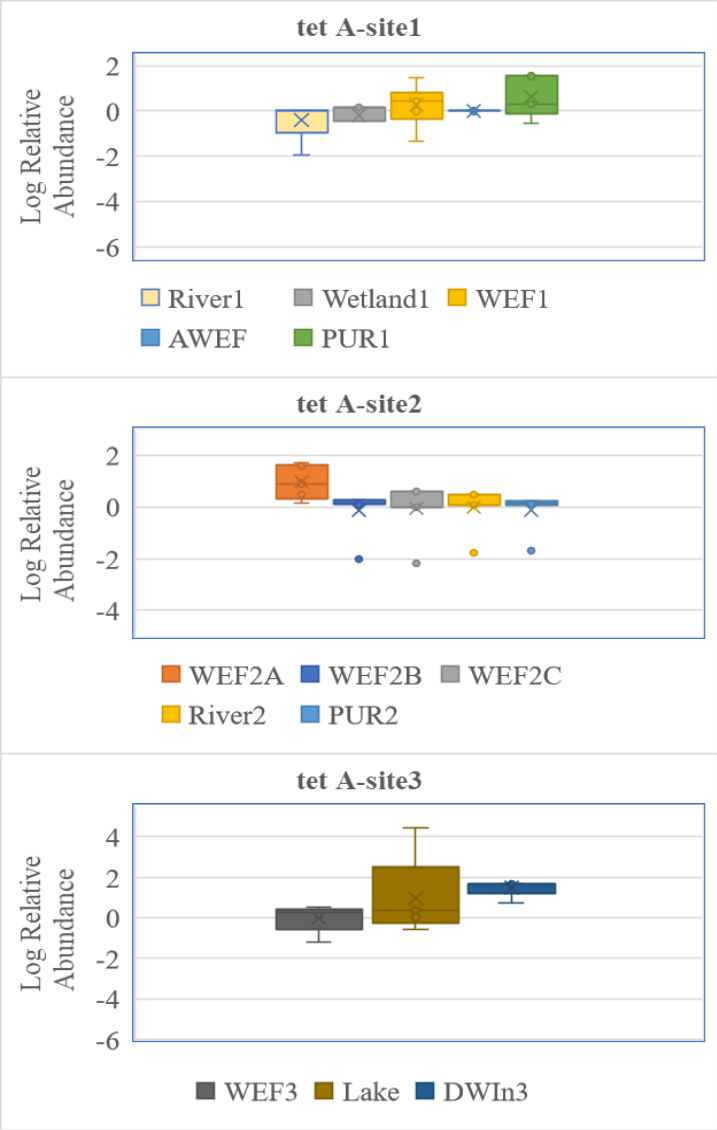


Figure 3.3. Quantification of tetA in three different U.S. Water Systems. From the top, site 1, 2, and 3 is shown.

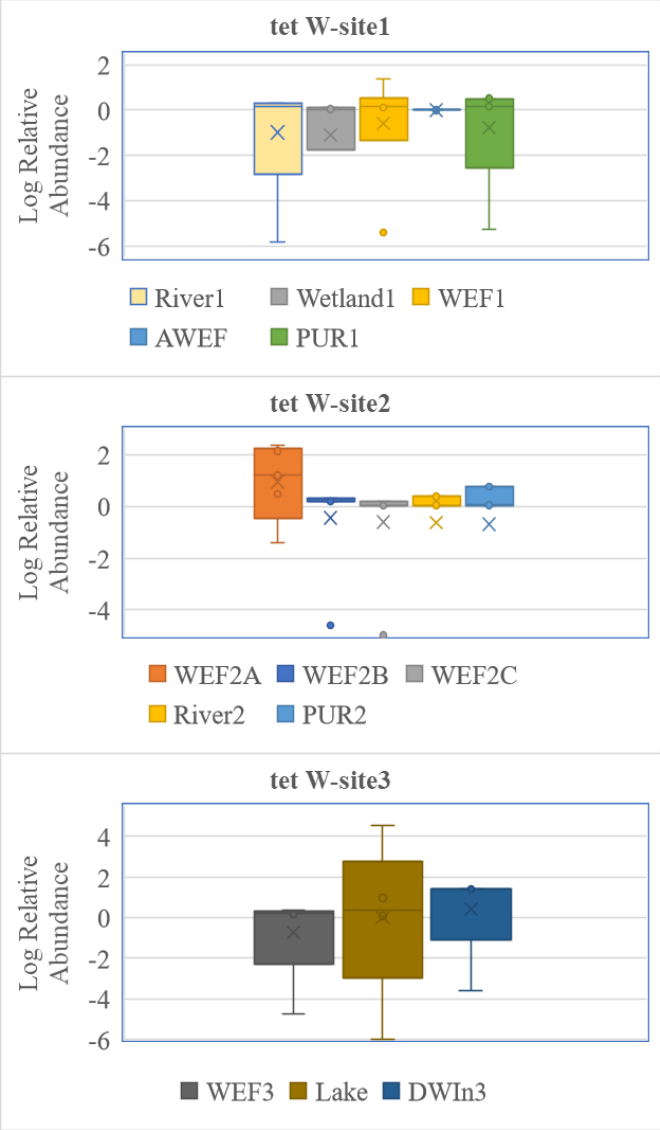


Figure 3.4. Quantification of tetW in three different U.S. Water Systems.

3.4.1.B. Quinolone Resistant Genes: qnrA and qnrB

Quinolone resistant genes were detected highest in post underground recharge of site 2 for qnrA, and mountain lake of site 3 for qnrB. It is found that the qnrA contamination is getting higher as the water goes downstream of site 1 until the advanced water treatment system, then decreases significantly after the treatment removal so that it provides a lower concentration of qnrA into the groundwater. In site 2, there are around 4 log units

of qnrA gene present in the wastewater effluents which gest diluted in the river. However, the concentration of groundwater downstream of the river has the highest concentration, which is assumed to be contaminated over long periods with Quinolone-related antibiotics. Also, site 3 showed an increasing concentration of qnr A towards downstream, which means the environmental buffer is not working. Rather, we can assume that the rate of qnrA accumulation in the water system is higher than the rate of qnrA is degraded by the water system. The result is shown in **Figure 3.5** and **Figure 3.6** for qnrA and qnrB, respectively.

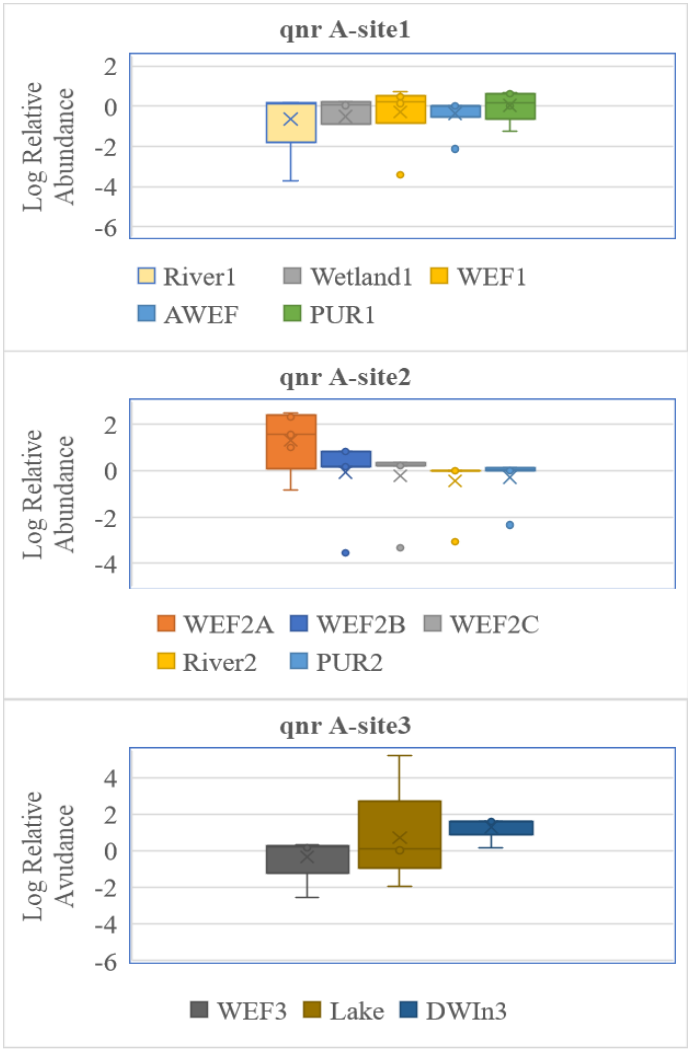


Figure 3.5. Detection of quinolone resistant genes (qnrA) in three U.S. Water Systems.

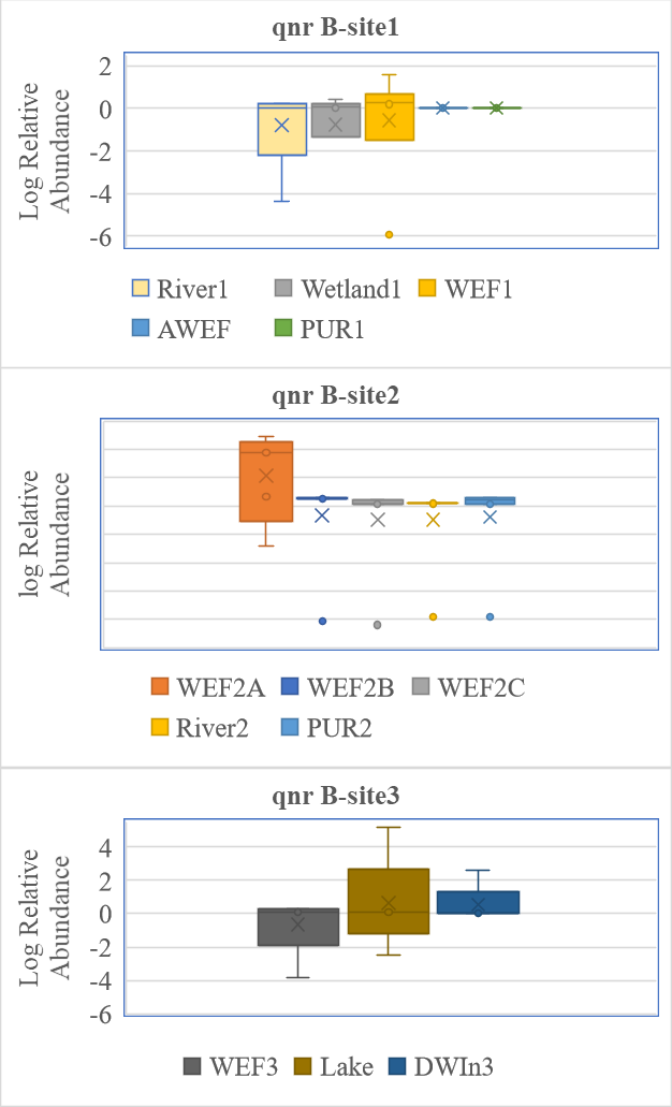


Figure 3.6. Detection of quinolone resistant genes (qnrB) in three U.S. Water Systems.

3.4.1.C. Beta-lactamase Resistant Gene: bla CTX-M

For all three sites, WWTP effluents were the biggest source of blaCTX-M gene, and following water bodies lowered their concentration using an advanced water treatment

system, river, or lake (**Figure 3.7**). Having a wetland also showed some attenuation effect at site 1, having a slightly lower concentration compared to the upstream of a river.

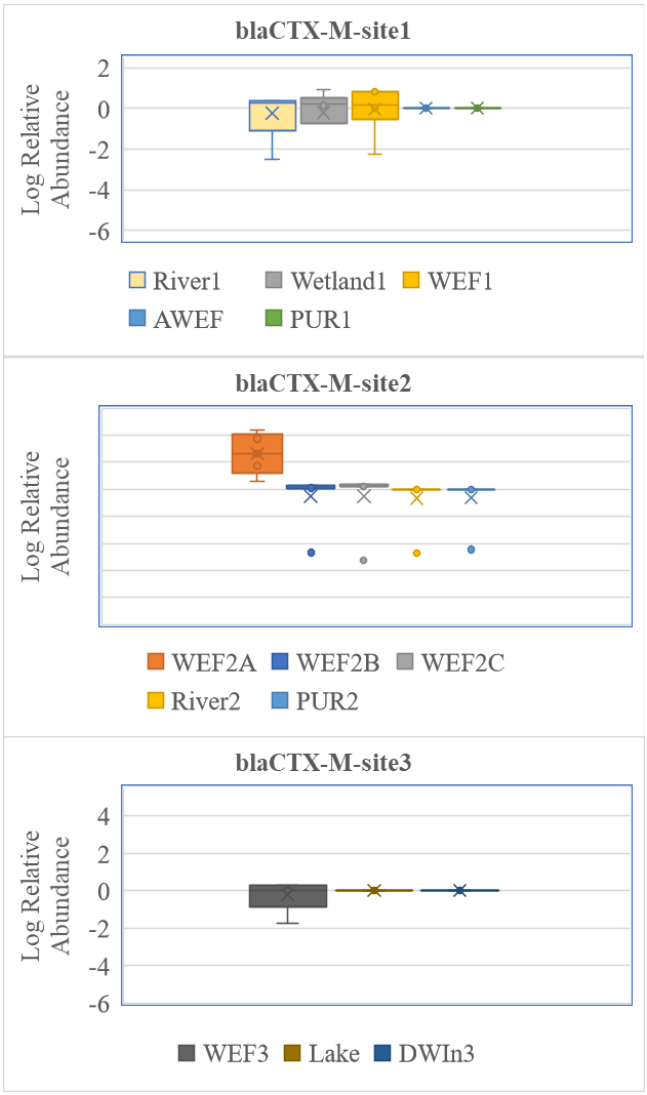


Figure 3.7. Beta-lactam resistance gene (blaCTX_M) gene quantification in three U.S. Water Systems.

3.4.1.D Sulfonamide resistance genes: sul1, sul2, and sul3

Three types of sulfonamide resistant genes were analyzed, and the results are shown in **Figures 3.8, 3.9, and 3.10** for sul1, sul2, and sul3, respectively. While WWTP effluents

in all sites seem to have high ranges of sulfonamide resistant genes, rivers within the water system showed some amount of high concentration, especially in site 2. Especially in site 3, the sul 3 concentration was the highest in the lake.

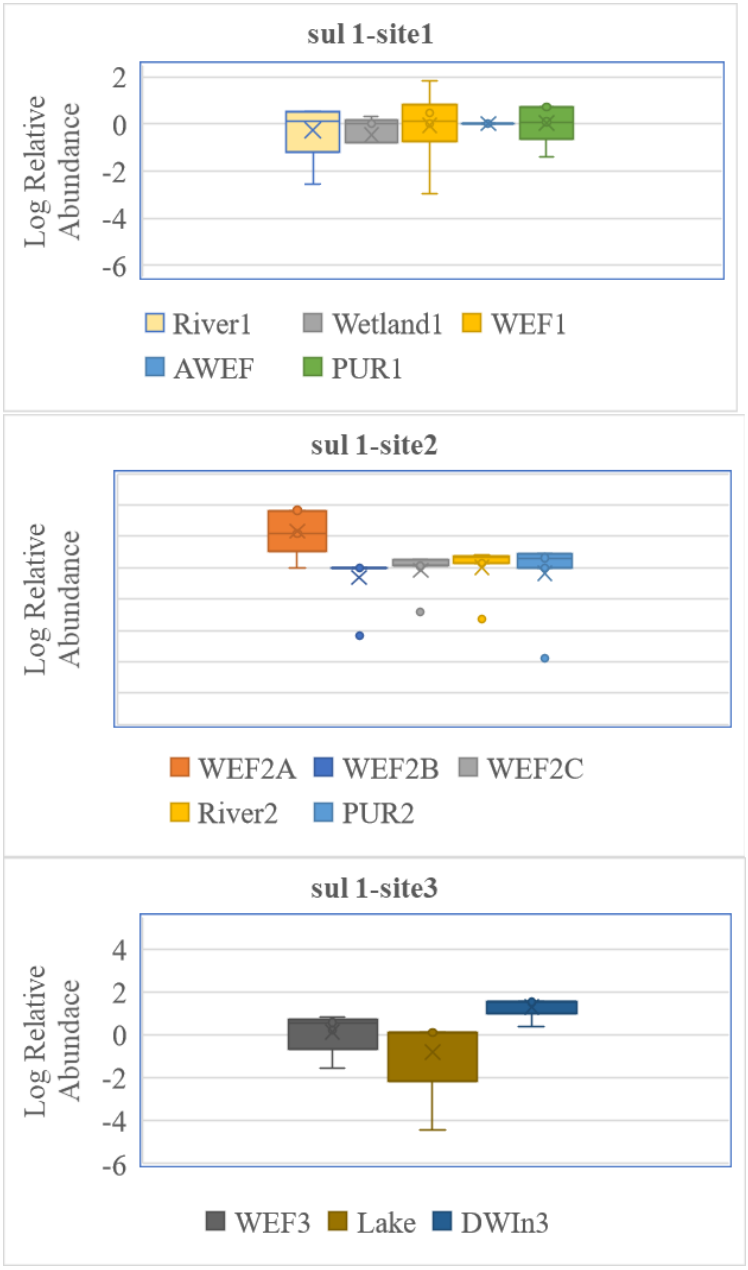


Figure 3.8. Quantification of sul 1 in three different U.S. Water Systems.

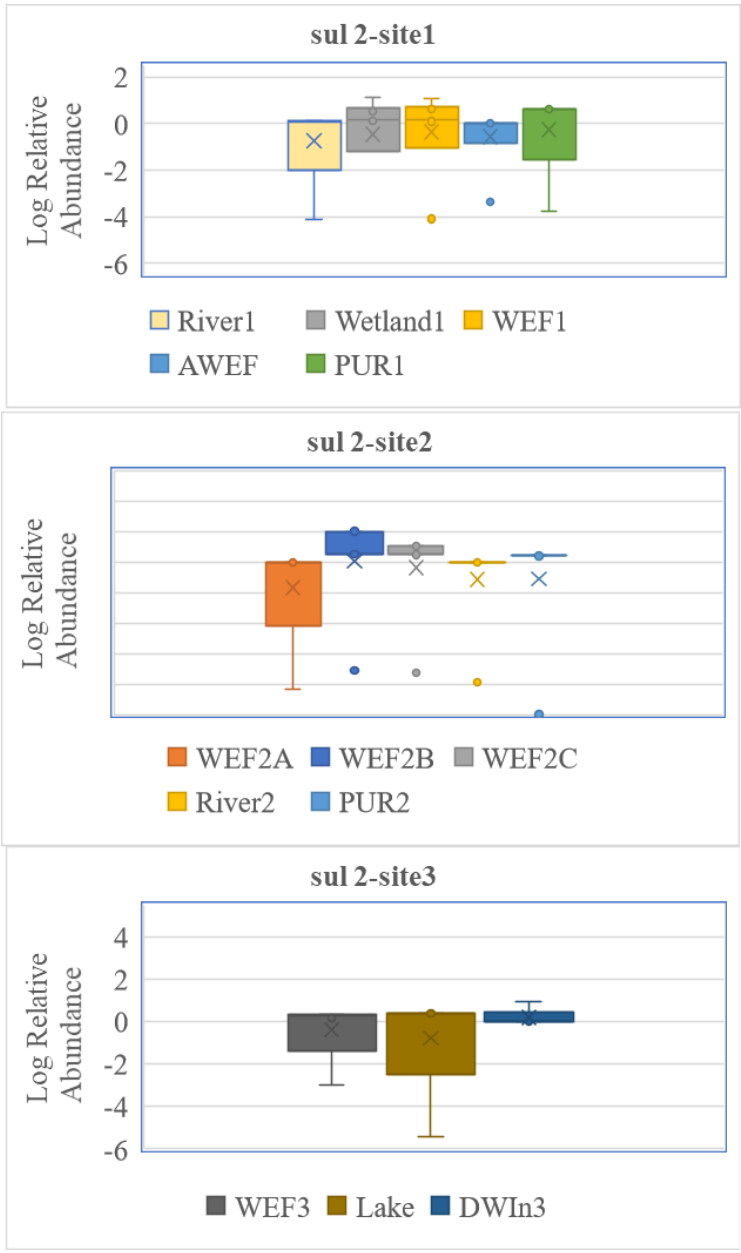


Figure 3.9. Quantification of sul 2 in three different U.S. Water Systems.

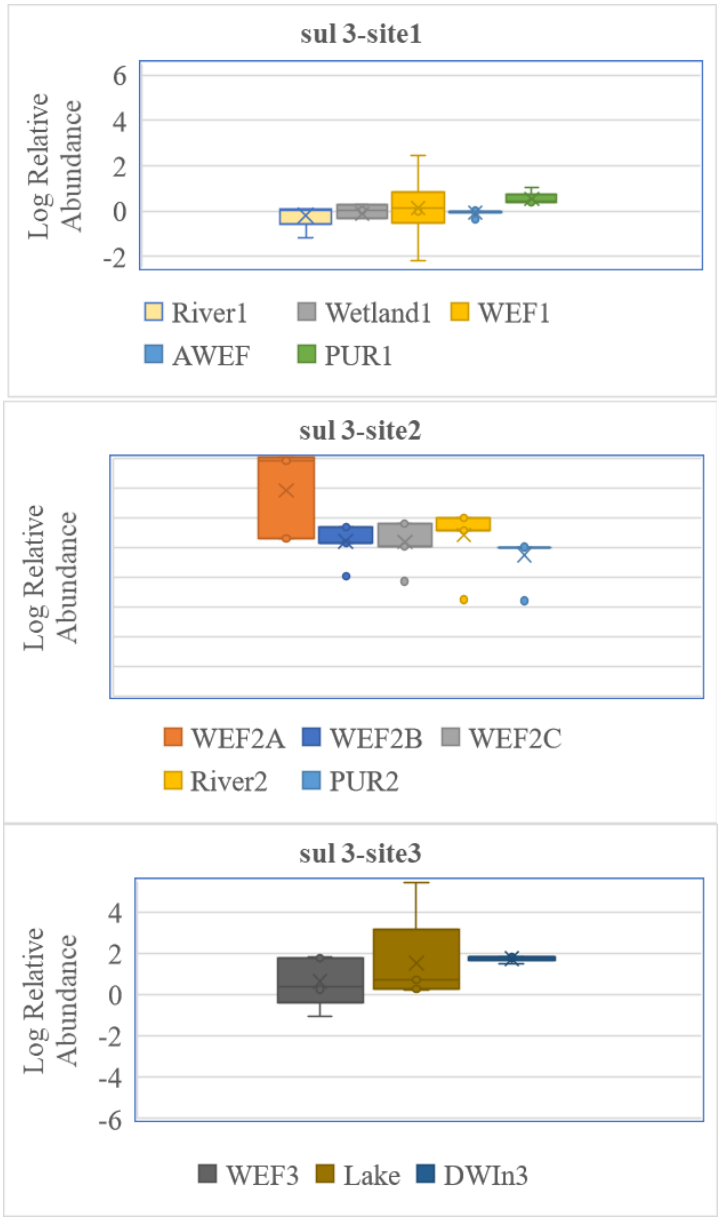


Figure 3.10. Quantification of sul 3 in three different U.S. Water Systems.

The highest contributor of sul 1 gene in the water system seems to be WWTP effluents on all three sites. In site 2, the river accepting one of the WWTP effluents showed a similar concentration of sul 1 gene which may cause concern in the potential spread of the gene because the river was not efficient in providing attenuation effect. On site 1, the WWTP effluent was found to be having the highest concentration of sul 2 gene compared to other water bodies. The sul 2 also detected in relatively higher concentration

in all three WWTP effluents, and the river at downstream of site B showed a higher concentration.

In the case of sul3, the wetland of site1 contained a higher concentration compared to WWTP effluent downstream, while site 2 WWTP effluent was higher than a downstream river. Interestingly, the mountain lake of site 3 showed a relatively higher concentration of sul 3 than any other water buffers, thus accumulation and growth of sul3 gene can be assumed at the specific lake with potential contamination of sulfonamide. In all of the water systems, having an advanced water treatment system helped reduce the concentration of sulfonamide resistant genes below log 1 unit, and introduction to an underground aquifer also reduced the concentration relatively lower compared to the previous water body content.

3.4.1.E. Erythromycin/macrolide resistance genes: ereA

Erythromycin resistant gene was detected the highest in WWTP effluent of site 2. We can assume that the high concentration of ereA is due to the anthropogenic effects at site 2, then the river and underground recharge provides attenuation effect of the contaminants as the water flows downstream. Likewise, in sites 1 and 3, WWTP effluents were detected to be the highest on ereA. It shows that to eliminate erythromycin-resistant genes, humans need to decrease the amount of antibiotics used to remove contaminants and the risks of having the ARG and pathogen emergence. **Figure 3.11** shows the ereA quantification among samples.

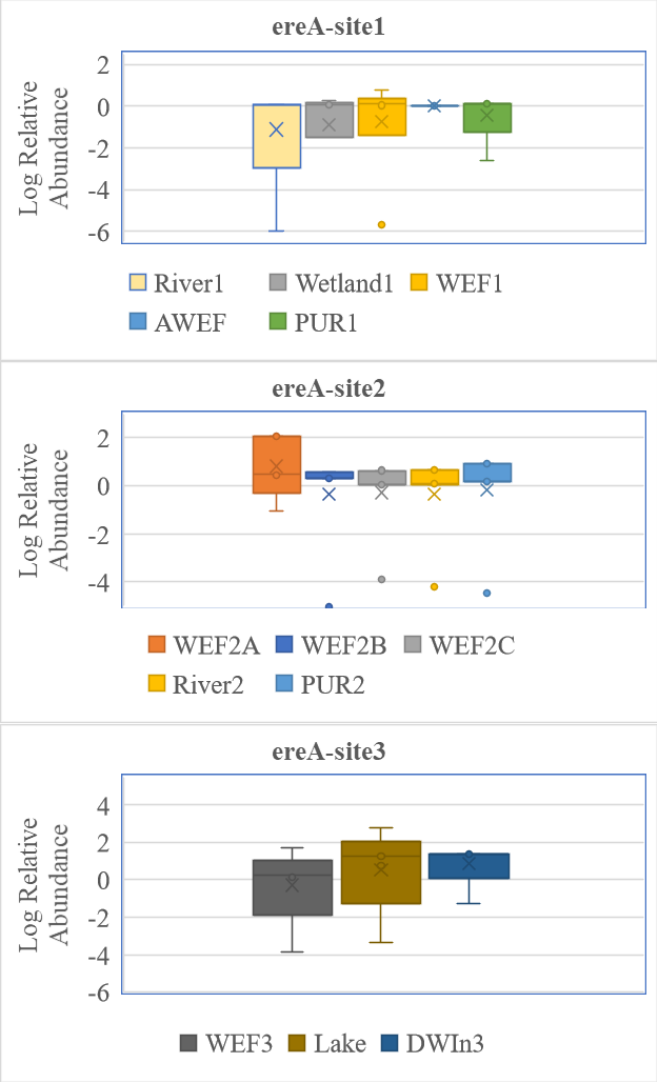


Figure 3.11. ereA gene quantification in three U.S. Water Systems

The ereA was detected the highest from wastewater treatment plan effluents of site 2, which represents anthropogenic activity may be the biggest contributor to ereA genes. It is also found that ereA amount is larger in WWTP effluents than other environmental buffers in site 1 as well.

3.4.2. General Bacterial Quantification using universal bacterial gene, 16S rRNA

The universal bacterial gene, 16S rRNA was quantified from all sampling points of three sites. The quantification result is shown in **Figure 3.12**. The total 16S rRNA gene was detected to be increasing in site A as the water flows from the river to the wetland, and then from getting the WWTP effluent. After advanced water treatment, the water was significantly reduced its bacterial gene to a lower level, as well as the underground recharge seems to be freer from the bacterial quantities.

In the case of site 2, WWTP effluent showed a higher range of bacterial concentration compared to other downstream water bodies of the river and post underground recharge. This showed the attenuation effect of water bodies of the river and underground recharge dilute the bacterial concentration. In the case of site 3, a low amount of WWTP effluent leads to a higher concentration of mountain lake, then a lower concentration of bacterial gene before going into the DWTP through the river. This can potentially imply that the mountain lake has a large hydraulic retention time which may cause a buildup of bacterial genes in the lake. After the mountain lake, the bacterial gene concentration decreased to a level similar or lower compared to the WWTP effluent upstream of the site 3 water system.

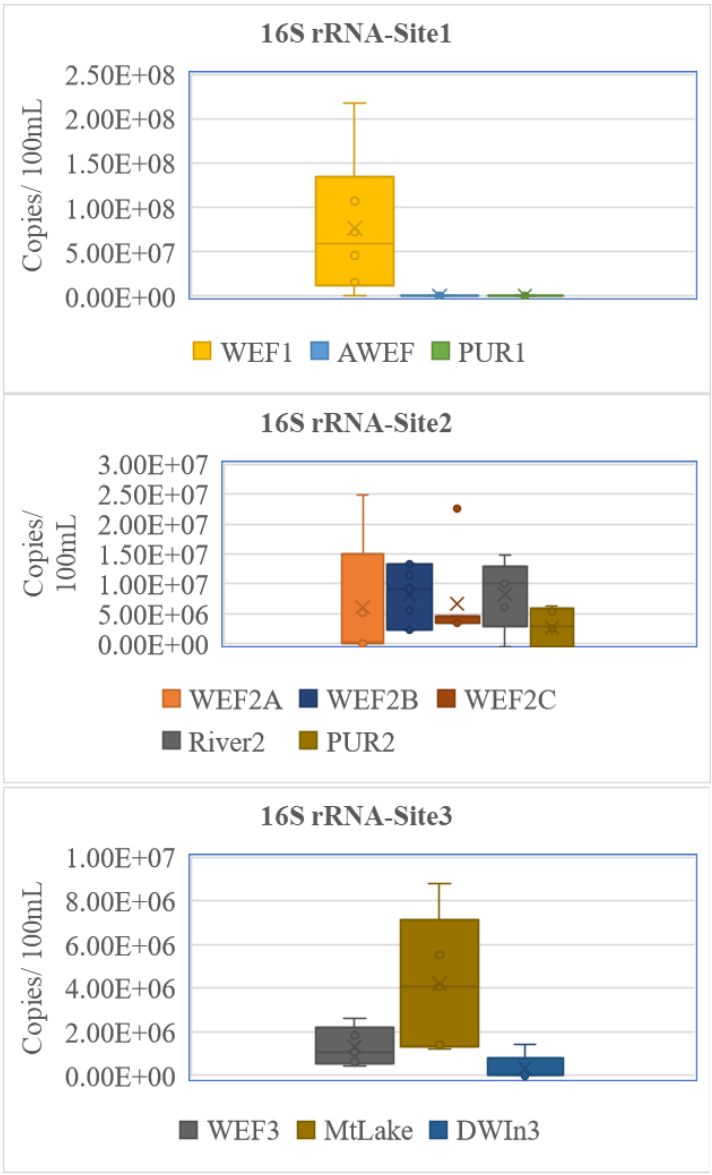


Figure 3.12. The concentration of 16S rRNA gene in three U.S. water systems.

3.4.3. Pathogen Detection in the Water System

In addition to the developed standards of Chapter 1, a result of positive control of parasites prepared with standard culture is shown below. In verification of waterborne pathogens Cryptosporidium and Giardia, Cryptosporidium was not identified from the positive stock samples DNA extraction. **Figure 3.13** shows no positive bands for

Cryptosporidium, while Giardia was identified in some of the samples at 74 bp. Therefore, only Giardia was utilized as a positive control and standard in qPCR analysis and the standard curve developed is shown in **Figure 3.14**.

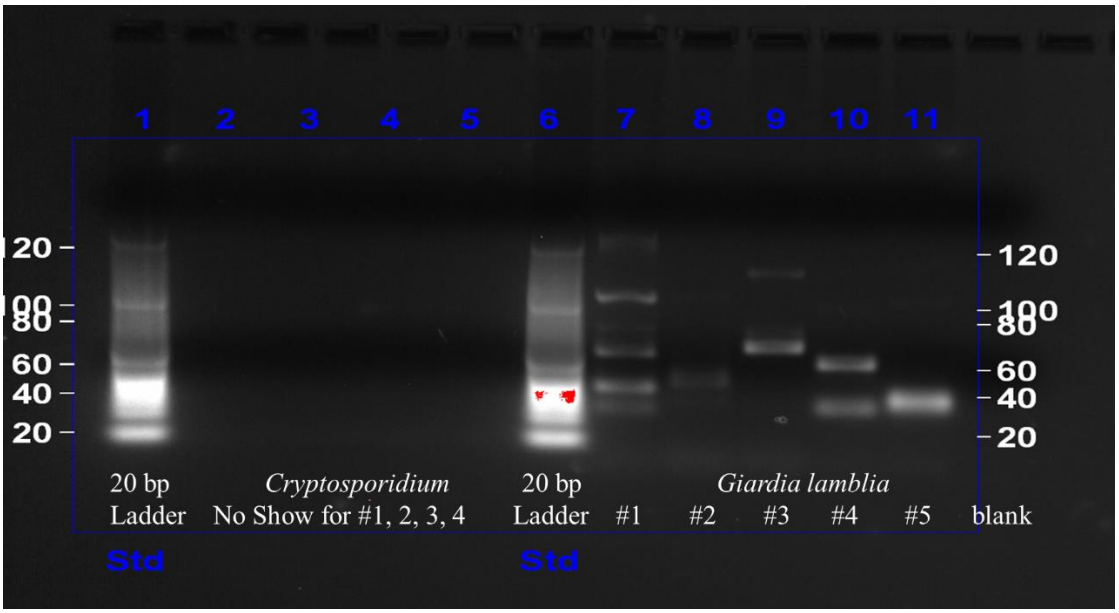


Figure 3.13. Crypto-Giardia PCR product

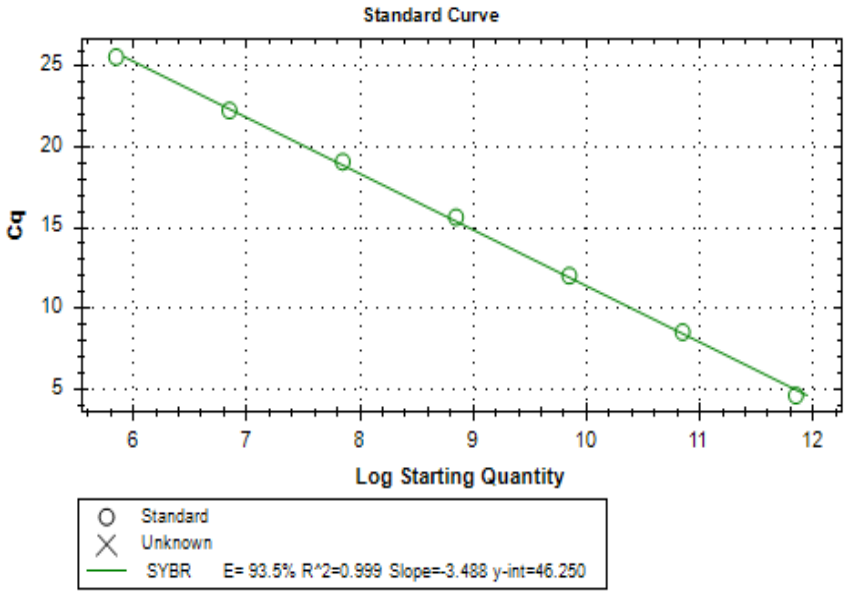


Figure 3.14. Standard Curve of Giardia for qPCR

Based on the copy numbers from the *Giardia* standard readings, the relative abundance of Giardia is represented in **Figure 3.15**. The Giardia is detected to be the most in the wetland of site 1, then in the river of site 2. Because giardia can be present in a dormant state as a cyst or active state, the wetland seems to provide a suitable environment for the increase of Giardia species in the water environment.

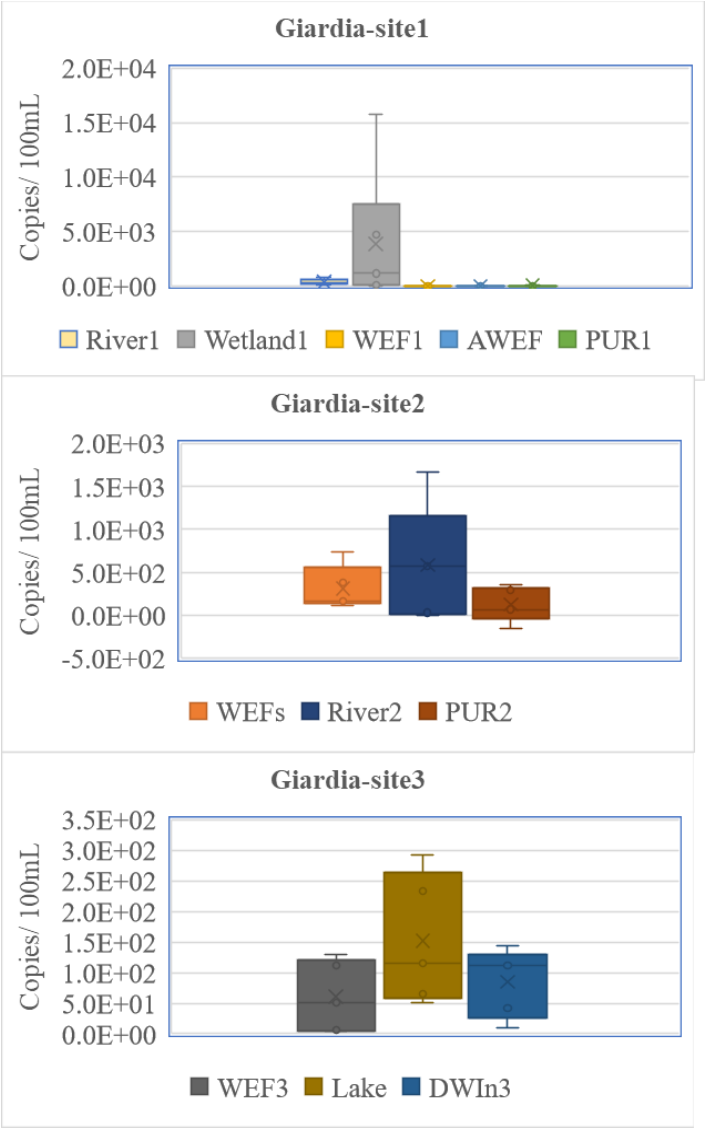


Figure 3.15. Giardia gene quantification in three U.S. Water Systems

Based on the result, the wetland seems to provide a more habitable environment for Giardia in site 1, and the introduction of Giardia from wastewater treatment plant

effluents also contributed to some amount of Giardia in the river which is downstream in site 2.

Multidimensional scaling (MDS) analysis was done to find the level of similarity in ARGs detection by water body types, where seven different types were categorized: AWEF, Wetland, DWI, Lake, PUR, River, and WEF. Site locations from 1-3 were added after the water body types to identify the difference in sites. As shown in **Figure 3.16**, all ARG data at each water body were collected for median value, then transposed to Cartesian space by Euclidean distance. The ARGs trend in the river is similar to each other, while wastewater effluents showed varying levels of divergence across both the MDS1 and MDS2 axis. Two distinctive loci where water body types are close to each other, one is AWEF1 and PUR1, and the other is WEF2 and Wetland 1. The closeness between each locus may suggest that the ARGs pattern and abundance were similar to each other compared to other water body sites.

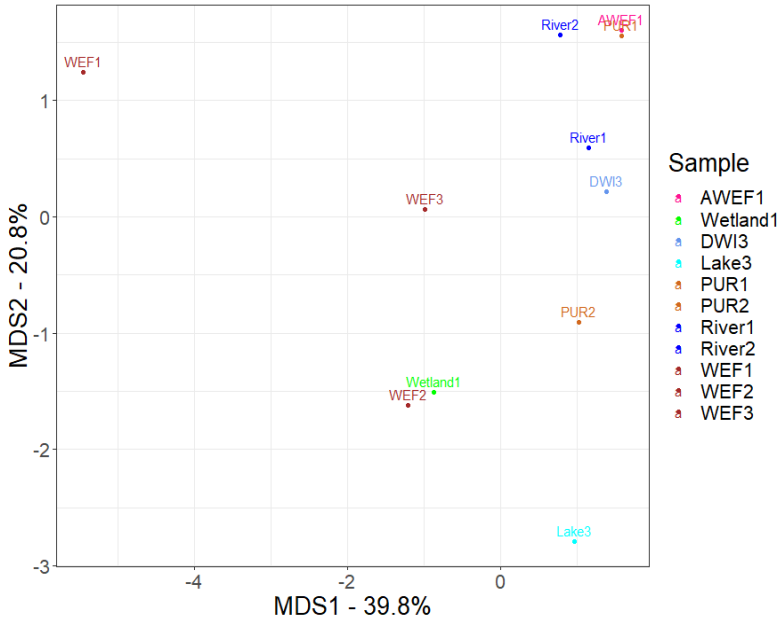


Figure 3.16. Multidimensional scaling analysis of site locations on ARG distribution

3.5. Conclusion

Except for the case of sul 2 and Giardia, river showed significant attenuation effect of ARG concentrations up to 3 log units and seems to be the most effective in decreasing ARGs content with aeration and continuous flow. Advanced wastewater treatment succeeded in removing most of the ARGs in the water, decreasing up to 6 log units to reach log 0 units or below regardless of the ARGs. Therefore, despite the higher cost of treatment, advanced water treatment might be worthwhile for the preservation of water for reuse and recycling. The WWTP effluents, on the other hand, appeared to be contributing to an increase of ARG concentrations compared to connecting water bodies. The increased ARG concentration of WEFs proved that anthropogenic activities introduce increasing ARGs in the water system over time. The analysis result showed the need to treat ARGs at the WWTP site more carefully because higher ARGs discharge may contaminate other clean water sources and cause damages to the water system and the ecosystem including human health.

Referring to the literature review done in **Table 3.1**, this chapter showed a coherent result of tetracycline (tetA, tetW) and fluoroquinolone (qnrA, qnrB) resistance genes having relatively higher concentrations compared to other water bodies at underground aquifer. It is assumed that tet- and qnr- type ARGs can be transported easily and percolate well through the media because they were found to be the most abundant than other ARG types and found deeper in the sediment (Tong et al. 2017; Zainab et al. 2020). Additionally, erythromycin (ereA) and sulfonamide resistance (sul1, sul2) genes were also relatively higher in the underground aquifer compared to the previous water body,

river, at site 2. This may be due to the accumulation of ARGs as the water is kept fed into the groundwater to be treated and reused by humans.

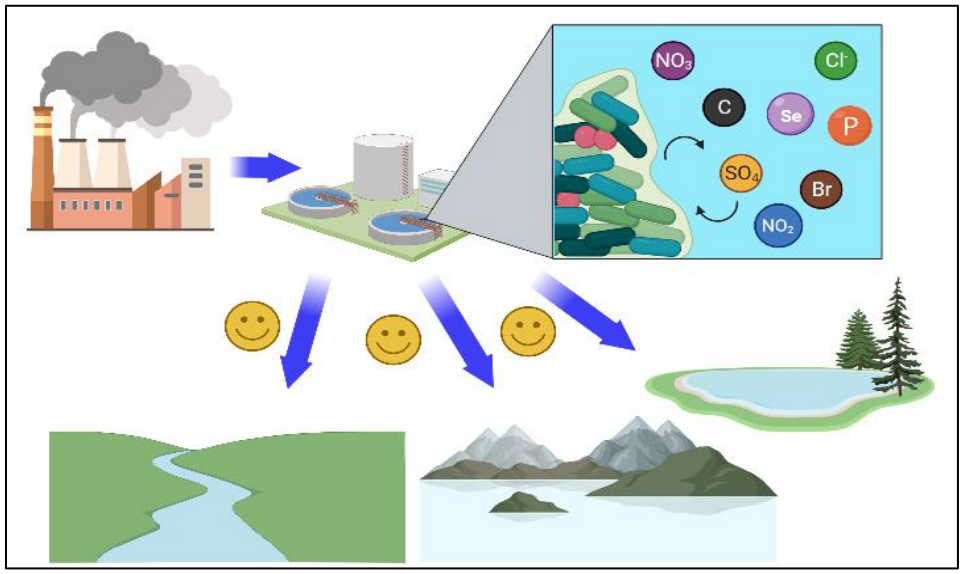
When comparing two locations of PUR of site 1 and site 2 on how precious water bodies (AWEF1 and River2) are related to the PURs, they are very different in the location in **Figure 3.16** because AWEF1 and PUR1 are overlapped while River2 and PUR2 are far away from each other along the MDS2 axis. This may be due to the method in water treatment and how recycled water is fed into the underground aquifer, where site 1 used advanced treatment method to clean water to feed the underground aquifer solely, while site 2 used riverbank filtration method restock the water which may still have some ARGs and antibiotics mixed from WEF upstream remains.

Overall total bacterial concentration with 16S rRNA followed the general trend of ARGs in their relative quantities except for the rivers. This may imply that other types of bacteria are abundant without antibiotic resistance properties in the river, as the water is continuously mixing and flowing receiving different types of water bodies flowing in and out. Also, the river may act as a reservoir of ARGs because it receives various types of water to be merged, mixed, and transported downstream. For this reason, it is important to monitor and regulate discharges into the river. Some future steps for environmental engineering may be setting guidelines on antibiotics and ARGs concentrations to keep contaminants from entering the water system and having the appropriate treatment before release to the environment. Preventing the spread ahead of time before the ARGs enter our water system may be greatly helpful for protecting human health and delaying the emergence of multidrug-resistant bacteria and diseases caused by them.

CHAPTER 4: Microbial Consortia and ARGs Investigation of Flue Gas Desulfurization Bioreactors Through Metagenomic Analysis

4.1. Background & Literature Review

Flue Gas Desulfurization (FGD) is the process of treating sulfur oxide generated in the combustion process of fossil fuels into the aqueous phase, where it can then be treated with alkaline chemicals to increase its low pH range. The overall schematic of FGD wastewater treatment is shown in **Figure 4.1**.



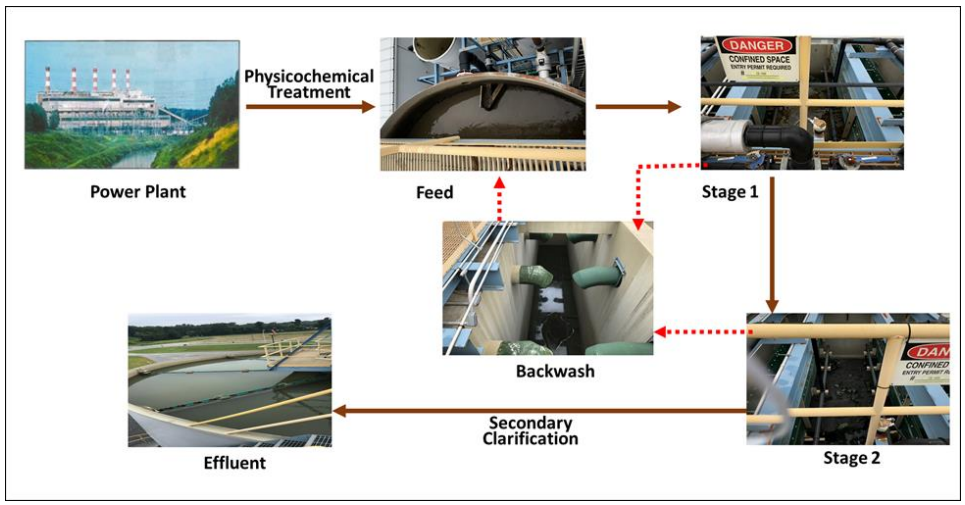


Figure 4.1. **FGD wastewater treatment and importance to the environment (top), and overall schematic of FGD wastewater treatment process (bottom)**

Despite the decreased use of fossil fuels, it is reported that 79% of the domestic energy production in the U.S. comes from fossil fuel consumption (Mcfarland, 2019)The FGD wastewater created during the power generation process can threaten clean environment maintenance and stable ecosystems if not treated properly before its release. Contaminants known to be present in FGD wastewater are sulfate, chloride, nitrate, bromide, and heavy metals like Se, Cd, Cr, Hg, As, Pb, Ni, Cu, etc. (Zhang et al., 2020). Without the treatment of the high sulfate and heavy metal contents, releasing FGD wastewater may lead to water, groundwater, and soil contamination which can threaten living organisms in water as well as human health. This is because the heavy metals cannot be biodegraded easily and cause accumulation in plants, soil, and aquatic lives. Also, high sulfur contents in water create a bad odor and induce corrosion of water pipelines. As sulfate is reduced, sulfide with organic carbon can be toxic to plants and crops which can cause environmental and economic issues.

FGD wastewater can be applied to microbial reactors and other treatment methods, including the use of zero-valent iron (ZVI) (Zhang et al., 2019), use of zero-liquid discharge (ZLD) with cyclone atomizer to eliminate the production of wastewater (Ma et al., 2019), and in-situ removal of liquid-phase selenium into solid gypsum, which prevents the formation of selenate with a removal efficiency of 96.5% (Zou et al., 2020). Out of all other treatment methods (i.e., filtration, chemical precipitation, electrodialysis (ED), or ion exchange), the use of microbial bioreactors may be the most sustainable.

Previous research has analyzed the composition of microbial communities in FGD wastewater in similarly operated tandem bioreactors, placing specific focus on the 16S rRNA eubacterial composition (Lin et al., 2015). This study showed an increased elimination capacity of sulfur vs. inlet load of the bioreactor. However, there are some needs expressed in this area to focus on microorganisms related to efficient removal and capability to handle varying substrates. With high sulfur, chloride, and some heavy metal content, our site of study concerns Se removal and bioaccumulation capacity of two-step bioreactors from three sites.

In this study, we focused on Se and sulfur removal to eliminate complexation effects of other ions and activities of heavy metals. Also, we focused on additional functional pathways that are involved in Se and sulfur removal as needed.

The finding of selenium by Berzelius in the bottom sludge of a sulfuric acid manufacturing system was not a coincidence, as it was later discovered there are similarities between selenium and sulfur utilization (Hatfield D., Berry M., 2011). Selenium-containing protein is considered essential in certain bacterial and mammalian enzyme systems (Stadtman, 1974). The metabolism and biosynthetic pathways of sulfur and selenium have yet to be discovered in detail, but this study will contribute to the identification of species and pathways connected with the microbial activities of these metals. Some bacteria were found to use arsenic and selenium as terminal electron acceptors under anaerobic respiration (Stolz and Oremland, 1999). Converting toxic oxyanions like arsenate and selenite into other forms is beneficial for both bacteria and the environment.

Some researches analyzed the composition of microbial communities in FDG wastewater bioreactors, operated in similar settings of two bioreactors, and analyzed the 16S rRNA eubacterial composition (Lin J., 2015). There are some needs expressed in this area to focus on microorganisms related to efficiency removal and capable of handling more various substrates (Li, 2015). Other contaminants such as ARG may be also detected since there are studies that reported ARG presence even in pristine environments. (Chen Baowei, 2013) Therefore, tracking of ARGs and studying how the microbiome plays a role in the presence of ARG types and quantities. If the water is irrigated from the nearby lakes of rivers, this could be an important point to check since virulence of AR and metal resistance (MR)) and stress response-related genes were detected more in urban areas (Medeirosa, 2016), and often the river and lakes are in the vicinity of human interactions.

To our knowledge, there only a few studies focused on the presence of ARGs and their fate in FGD bioreactors. And industrial wastewater has been on a blind sight for ARG detection. Most metagenomic studies mainly focused on the 16S rRNA gene, looking for bacterial functions on how to modify and transfer contaminants of concern (Ruan et al. 2015). For this reason, the fate of ARGs under the stressed industrial environment of FGD bioreactors and how the ARGs concentration is different between site locations may be helpful to understand ARG activities and prevention of spreading. There is preliminary data on characteristics of FGD wastewater and preserved samples collected from three different coal power plants in NC.

In this study, a various spectrum of the microbiome was found working towards assimilation and utilization of compounds in FGD wastewater, such as sulfates, nitrates, selenium, and other heavy metals. The goal of this project is to further investigate the mutual effects of the microbiome and the presence of ARGs in the FGD bioreactors. The objectives of this study are (i) determination of microbial consortium that contribute to sulfate and heavy metals removal, (ii) investigation of the seasonal effects posed on microbial variations of FGD bioreactors, and (iii) Validate potential sources of ARGs and contribution of the microbiome for varying ARGs quantity at different sites.

4.2. Materials and Methods

The overall schematic diagram of metagenomic analysis is shown in **Figure 4.2**. The main part of the procedure consists of sample collection and processing, DNA extraction, shotgun metagenomic sequencing, and data analysis.

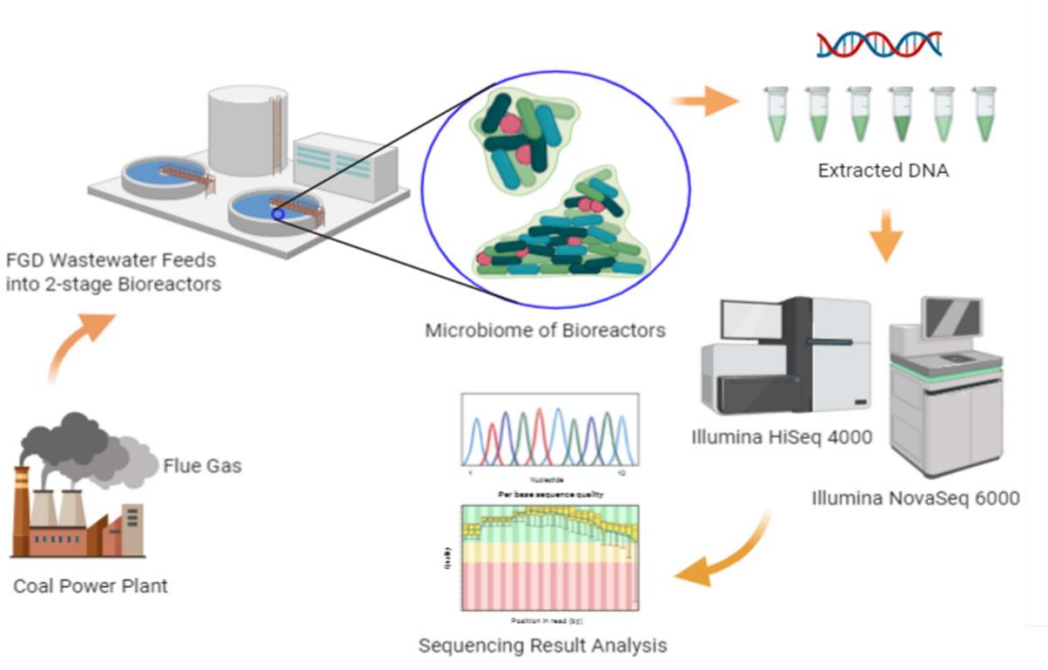


Figure 4.2. Metagenomic analysis procedure of FGD bioreactor

4.2.1. Sample Collection

Samples were collected at three different power generation plants in NC with two bioreactors each per plant during four different time points over two years: July 2018, January 2019, June 2019, and November 2019. Sampling Events 1 and 2 contained samples of all sites, but site C was missing in Sampling Events 3 and 4 due to the retirement of the plant, having no reactor beds available for sample collection. Sample

photos taken upon receipt are shown in Appendix C. **Table 4.1** shows sampling details with site-specific and dates.

Table 4.1. Sampling event details and collection dates

		Sampling event 1 (SE1)	Sampling event 2 (SE2)	Sampling event 3 (SE3)	Sampling event 4 (SE4)
		Jul-18	Jan-19	Jun-19	Oct-19
Site A	A S1	x	x	x	x
	A S2	x	x	x	x
Site B	B S1	x	x	x	x
	B S2	x	x	x	x
Site C	C S1	x	x		
	C S2	x	x		

The samples were delivered to UNC Charlotte for analysis immediately after the sampling events and processed for further analysis within 24 hours.

4.2.2. DNA Extraction

The samples were concentrated by centrifugation of 15mL in volume for sample processing, and 50mL for sample preservation. To avoid experimental/processing errors, solid and liquid parts of samples were collected in equivalent proportions for each sample. DNA was extracted from concentrated pallets of samples using the PowerSoil DNA extraction kit (MO BIO Laboratories, Inc., Carlsbad, CA).

4.2.3. Shotgun Metagenomic Sequencing and Library Preparation

Extracted sample DNA was sealed in a sequencing plate, protected with an aluminum covering, and sent to the Duke University Sequencing Center via overnight shipping, securely kept on ice. Samples from SE1 and SE2 were sequenced by using the Illumina Hiseq 4000 instrument. The fragmented metagenomic DNA was processed as part of the library preparation for HiSeq Whole Genome sequencing. Samples from SE3 and SE4 were sequenced using the Illumina Novaseq 6000 instrument, as the sequencing facility went through an update and retired the Hiseq 4000 system. Novaseq provides efficient and simple data processing while providing more output reads per flow cell, reaching up to 10,000 reads. This makes it suitable for a wide range of applications. Regardless, all sample reads were processed through the same workflow of library preparation, sequencing, data analysis, and interpretation.

4.2.4. DNA Sequence Trimming and Quality Control

The sample reads were quality controlled with the same threshold and pipeline, keeping data processing and handling conditions consistent. Raw sequence data were prefiltered by the sequencing center using the Illumina Hi-Seq HCS pipeline. Sequence reads were removed below an accuracy rate of 99.9%. Additional prefiltering was performed locally using Trimmomatic and PEAR to remove remaining traces of Illumina adapters and low-quality sequences by the UNCC Bioinformatics Services Division at the North Carolina Research Campus.

4.2.5. Metagenomic Classification Analysis using MetaPhlAn2

Metagenomic sequences were analyzed using the MetaPhlAn2 software package. MetaPhlAn2 is used as a method to perform taxonomic classification of shotgun metagenomic sequence reads using a custom database of clade-specific marker genes. Taxonomic abundances of bacterial genera of each sample were compared to each other to find similarities and differences. A bi-directional heatmap was created using RStudio of all samples to visualize the relative abundances and shifts in targeted features. The relatedness of bacterial genera in between the samples was mapped which brings an overview of microbially and genetically related sites and samples. Heatmaps were created for the warmer and colder season, respectively, to capture the differences between the climate effect on the activities of bioreactors.

4.2.6. Pathway Assessment using BioCyc Database Collection

The use of the BioCyc Database helped fine lining the pathways of interest for target contaminant metabolisms and pathways of denitrification, desulfurization, and demineralization cycles. The web-database enabled visualization of metabolic pathways, enzymes, and substrates as well as intermediate products giving an overview of microbial activities. The researched information from the BioCyc database was actively used in functional analysis of the next step using HUMAnN2.

4.2.7. Functional Classification Analysis using HUMAnN2

The collection of whole-genome shotgun metagenomic sequences enables a range of analysis beyond just taxonomic classification. The HUMAnN2 can find the presence/absence and abundance of microbial metabolic pathways in the data. Pathways

can be selected based on keywords and functions of interest, and the sequence data can be stratified to show which organisms contribute to metabolic functions of interest. Each set of graphs was created on the target pathways of interest in two ways- in the order of the pathway abundances, and the order of sampling locations.

4.2.8. Identification and Quantification of Resistant Elements using ShortBRED

The ShortBRED pipeline was used to search for protein families of interest, for instance, a protein that conveys antibiotic resistance, which its custom reference databases exist. ARGs were target groups of this analysis, which the data was obtained through a two-step process: 1) ShortBRED Identify, 2) ShortBRED quantify. The first method allows creating new markers from a set of proteins of interest and reference proteins, or if there is already a previously prepared database. For the analysis of ARGs, the CARD database was used to match AR protein-encoding genes. Then the second step, ShortBRED-quantify was run to match with AR encoding protein genetic markers with the whole genome sequenced sample. The result of the counted marker will be shown in the normalized result, which is expressed in units of Reads Per Kilobase of reference sequence per Million sample reads (RPKM). And the final result was shown as the median of marker normalized count as family level value.

4.2.9. Statistical Methods to find bacterial diversion and significance

Statistical analysis methods of various representations such as Principal Coordinate Analysis (PCoA), stacked bar, and percent relative abundance were used to better represent the data. Using PCoA plots, microbial community compositions and

abundances between samples were studied depending on the site and season they were collected.

4.3. Results

4.3.1. Extracted DNA Quality and Concentration

DNA extracted from samples was initially quantified using Nanodrop One microvolume UV-Vis spectrophotometer (Thermo Fisher Scientific, Agawam, MA) for rough quantity and purity of isolated DNA. The DNA measurement result is shown in **Table 4.2**. Site C showed a comparably low concentration of DNA compared to site A and site B from the same amount of 15mL sample in volume. This may be due to the current operation state of each station, as site C stopped its operation in 2013 and recently retired. The microbial community seems to be more abundant in site A and site B sites, as it is still in an active stage. Although, we can see that the microbial properties from site C are still intact and left with small amounts as some of the DNA remains in low concentration.

Table 4.2. Extracted DNA concentration and sample specification of three sites.

		Site A		Site B		Site C	
		Stage 1	**Stage 2**	**Stage 1**	**Stage 2**	**Stage 1**	**Stage 2**
SE1	**pH**						
	DNA Conc. (ng/uL)	60	40	35	36	42	54
	A260/A280	1.8	1.84	1.82	1.8	1.81	1.9
	A260/A230						
SE2	**pH**	7	7	7	7	7	7
	DNA Conc. (ng/uL)	346.4	69.6	123	98.9	16.9	24.7
	A260/A280	1.88	1.92	1.88	2.02	1.85	1.89
	A260/A230	0.29	0.05	0.1	0.09	0.02	0.02
SE3	**pH**	7.5	7.0	7.5	7.0		
	DNA Conc. (ng/uL)	153.75	54.05	73.25	57.6		
	A260/A280	1.86	1.76	1.78	1.8		

	A260/A230	0.14	0.15	0.17	0.09		
SE4	**pH**	8.0	7.0	7.0	7.0		
	DNA Conc. (ng/uL)	113.8	28.9	121.4	66.2		
	A260/A280	1.85	1.92	1.82	1.87		
	A260/A230	0.14	0.04	0.13	0.07		

4.3.2. Metagenomic Consortium and abundances in FGD Bioreactors

The heatmaps of different sites showed unique microbiomes depending on their location. While several groups of microbes were identified in most of the samples, there were slight shifts amongst stages of bioreactor and site locations observed. For example, continuous horizontal bands of *Thauera, Klebsiella, Alicycliphilus, Pseudomonas, Leptonema, and Yualikevirus* on the heatmap of all samples (**Figure 4.3**) demonstrates some trend in the shared microbiome of FGD bioreactors despite the differences in location. Among samples, those from site C were rather unique when compared with other samples. Since the plant in site C is being retired, it is plausible that the microbiome of site C tends to gravitate towards more of the fermentation-favorable, anaerobic environment compared to other sites.

The classified bacterial consortium represented activities of bioreactors, having sulfate-reducing bacteria, nitrate reducers, and fermenters as majority populations. A mixture of bacterial genera utilizing chlorides, degrading aromatic compounds, and some viral genera were found which aid sulfate and nitrate reduction and natural remediation process. The normalized abundances of the microbiome in each bioreactor sample are shown in **Figure 4.3**, while samples that have phylogenetically similar compositions are located closer to each other. The bacterial consortia of sites 1 and 2 were closer to each

other compared to site C. *Thauera* and *Klebsiella* showed the most relative abundance throughout the samples. Most of the species under the genus Thauera are known as the nitrate-reducers, which was identified as the key microbiome of aromatic compound biodegradation[34]. This study result reinforces the findings that Thauera acts as a key role in the digestion process of nitrate and aromatic compounds under anaerobic conditions. *Klebsiella* is known as facultative anaerobic, fermentative bacteria which seems to promote the biodegradation process.

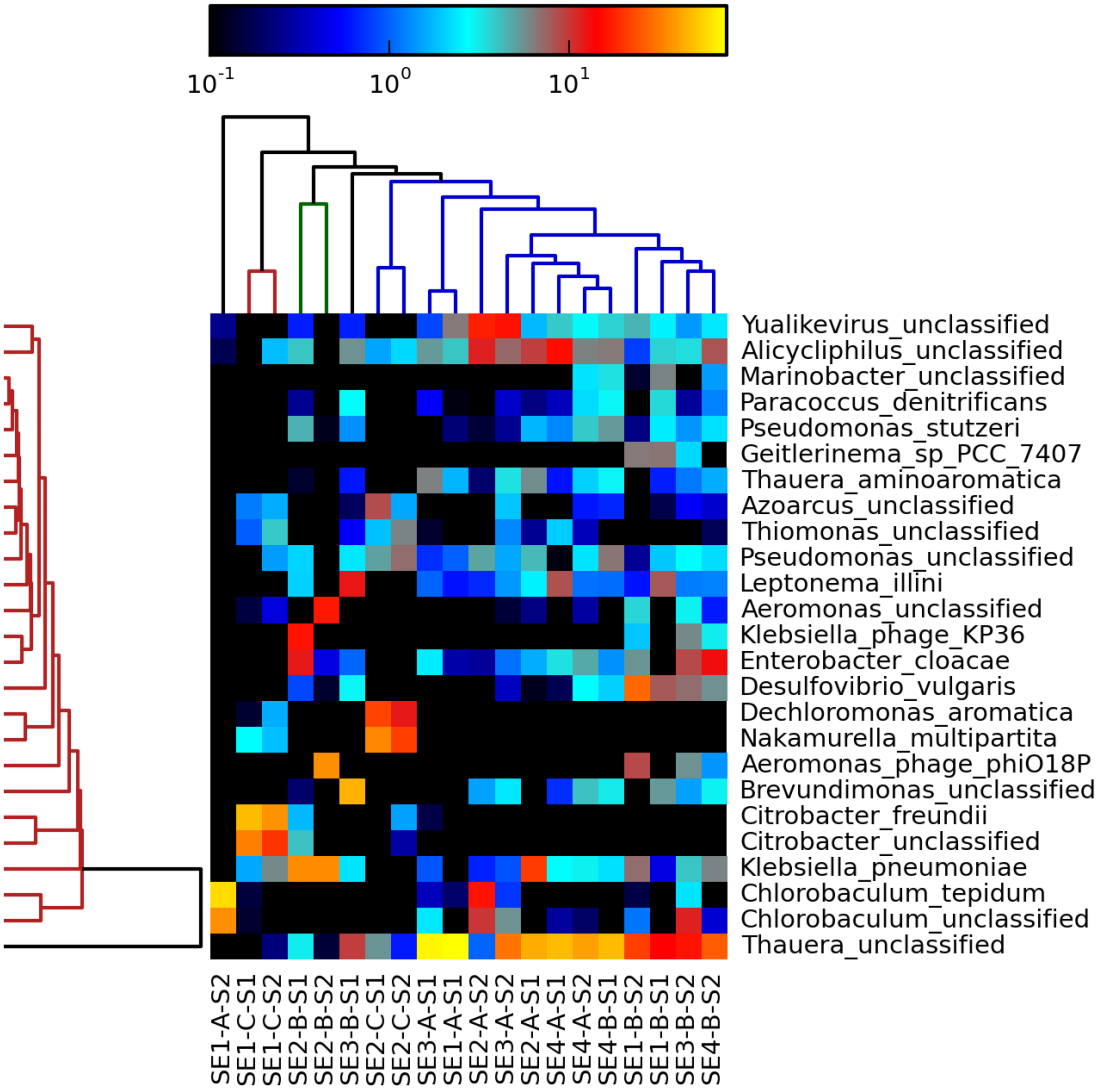


Figure 4.3. Heatmap of all sampling events (SE1 ~ SE4) on the genus level. SE3 and SE4 only have samples from two sites A and B: There are general trends in all sites that *Thauera (unclassified), Klebsiella,* and *Alicycliphilus* are present for most of samples. There are similarities between site A and site B bacterial distribution where higher *Yualikevirus, Enterobacter, Leptomena,* and *Thauera_arminoaromatica* are more abundant than site C. Site C samples showed unique patterns of high measurement of *Dechloromonas, Nakanurella* for SE2, and *Citrobacter* genera for SE1.

4.3.3. Confirmed Seasonal Effects of Microbiome in Bioreactors

The samples collected from warmer months (SE1 and SE3) and colder months (SE2 and SE4) are categorized separately and each figure is presented in **Figure 4.4** and **Figure 4.5**, respectively. During the warmer month, Site C detected a higher abundance of *Citrobacter, Raoultella, Klebsiella*, and some groups of *Azoarcus, Nakamurella*, and *Thiomonas*. Other samples from sites A and B showed a similar distribution of genera over most of the samples, while site B has more species variety.

On the other hand, under the colder months, Site C samples showed a large abundance of *Dechloromonas* and *Nakamurella* along with *Alicycliphilus, Janthinobacterium, Thiomonas, Acidovorax,* and *Azoarcus* genera. In sites A and B samples, *Yualikevirus, Acycliphilus, Klebsiella*, and *Thauera* were found to be abundant, showing very similar patterns among the two sites.

In the warmer months, relatively more abundance and variety of species are present in site B samples compared to site A samples, while in colder months we can see a decrease in bacterial community variety in both sites A and B. In the case of site C, we could see shifts in bacterial community distribution between the warmer and colder months: under warmer weather, site C samples are high in *Citrobacter, Raoultella, Klebsiella,* and some groups of *Azoarcus, Nakamurella,* and *Thiomonas*. Under cold weather, *Dechloromonas*

and *Nakamurella* are in high abundance, and groups of *Alicycliphilus, Janthinobacterium, Thiomonas, Acidovorax,* and *Azoarcus* are present in relatively higher abundance than other genera.

The comparison result showed that extreme weather conditions such as too hot in the summer months or too low temperature in the colder months hinder the bacterial diversity and prosperity, as samples collected in June (SE3) are more diverse compared to the samples collected in late July (SE1), and samples from late October (SE4) are showing a broader spectrum of bacterial distribution than samples from early January (SE2). This implies bioreactor operation conditions are more sustainable in the spring or fall seasons, since summer and winter may cause drastic changes in temperature, upsetting the equilibrium of the ecosystem within the bioreactor, and risking the survival of certain bacterial species benefiting the FGD wastewater remediation process. For example, temperature change during the fermentation process may restrict microbes from performing enzymatic activities in their metabolism. This change can cause a shift in equilibrium in the bacterial consortium that causes microbes to adjust to the new environment. For certain bacteria, it can be time to survive and prosper, while others will be diminishing. This shifting transformation becomes noticeable in the case of *Thauera*, as its relative abundance is larger in the warmer months and descends in the colder months. On the other hand, *Klebsiella* was at reduced relative abundances in the warmer months and then facilitated to grow more in the colder months.

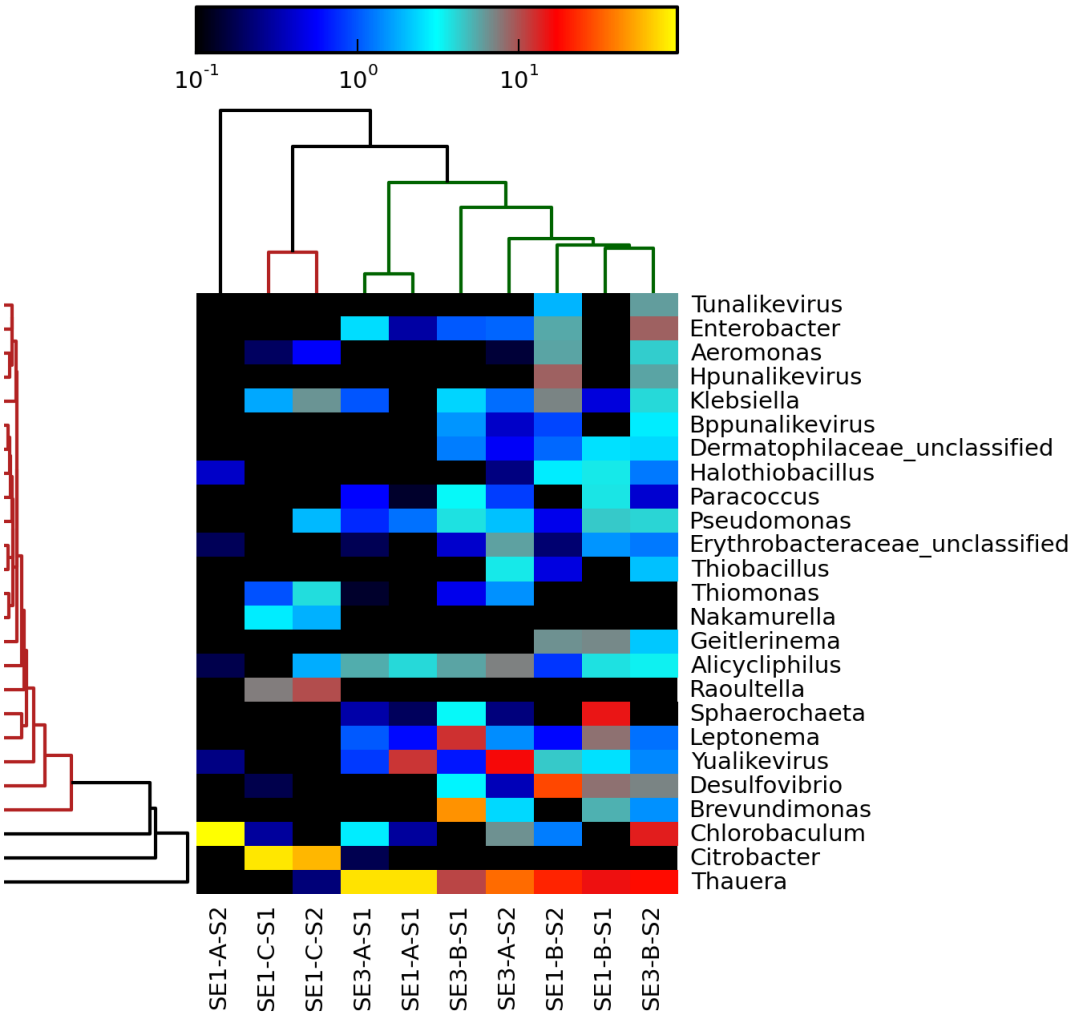


Figure 4.4. Heatmap of samples from warmer months (SE1 and SE3): Site C samples are more abundant in *Citrobacter*, *Raoultella*, *Nakamurella*, and *Thiomonas* compared to the other sites. Other samples from site A and B show a similar distribution of microbial consortia across most of the samples, with *Thauera*, *Chlotobaculum*, *Yualikevirus*, *Leptonema*, and *Enterobacter*.

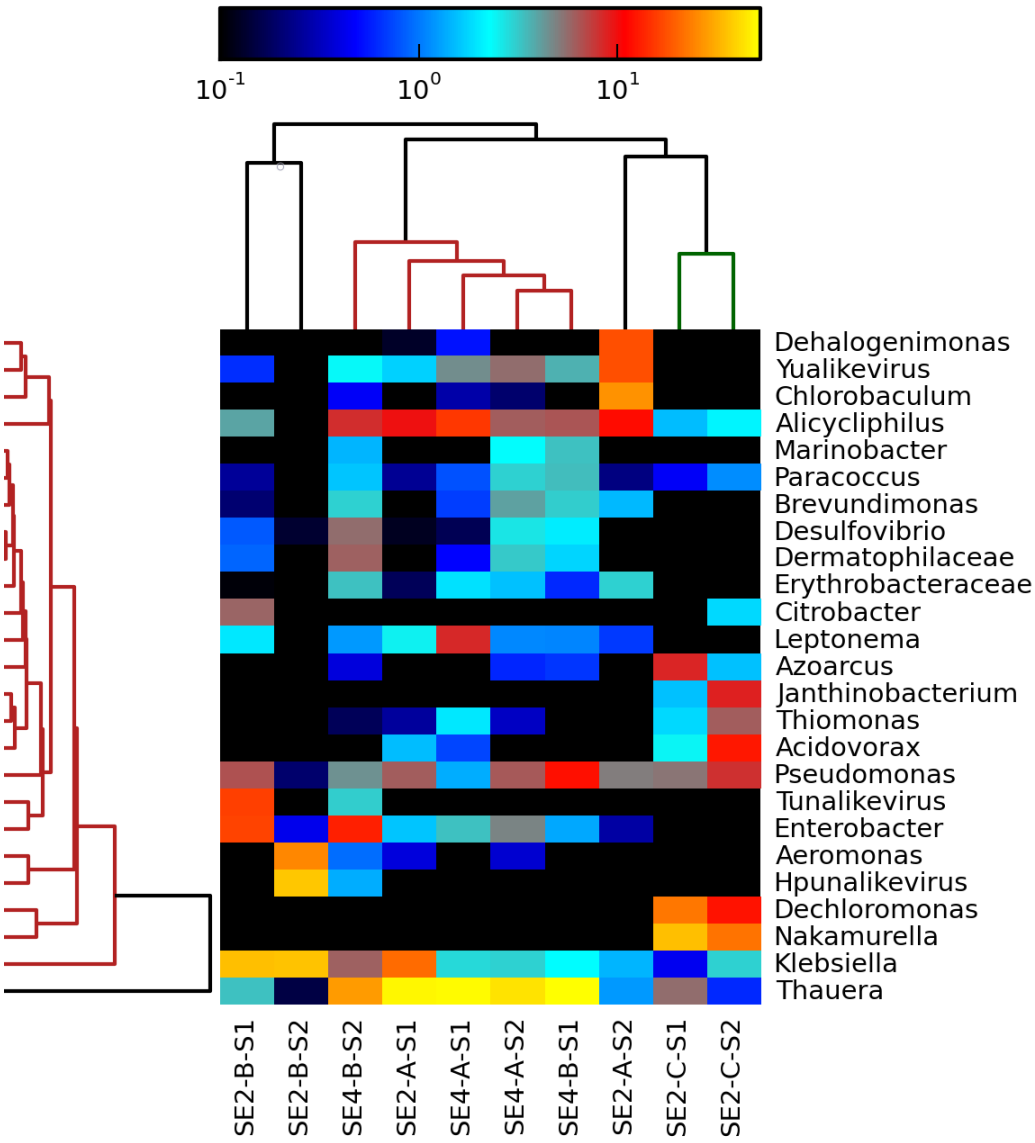


Figure 4.5. Heatmap of samples from colder months (SE2 and SE4): In RB samples, Dechloromonas and Nakamurella are in higher abundance from site C, along with Janthinobacterium, Thiomonas, Acidovorax, and Azoarcus. For site A and B, the showed very similar patterns across the two bioreactor stages, where Alicycliphilus, Thauera, and Pseudomonas are present in relatively higher abundance than other genera.

4.3.4. Pathway Analysis

4.3.4.A Pathways for Sulfur Cycle

With the high concentration of sulfur contents in FGD wastewater, abundant sulfur metabolism pathways were found. Several sulfur-metabolism for cellular energy creation, sulfate reduction, and sulfate assimilation were considered to identify species involved in sulfur cycling in bioreactors. Sulfate reduction I (assimilatory, SO4ASSIM-PWY) and superpathway of sulfate assimilation and cysteine biosynthesis (SULFATE-CYS-PWY) are shown in **Figure 4.6**.

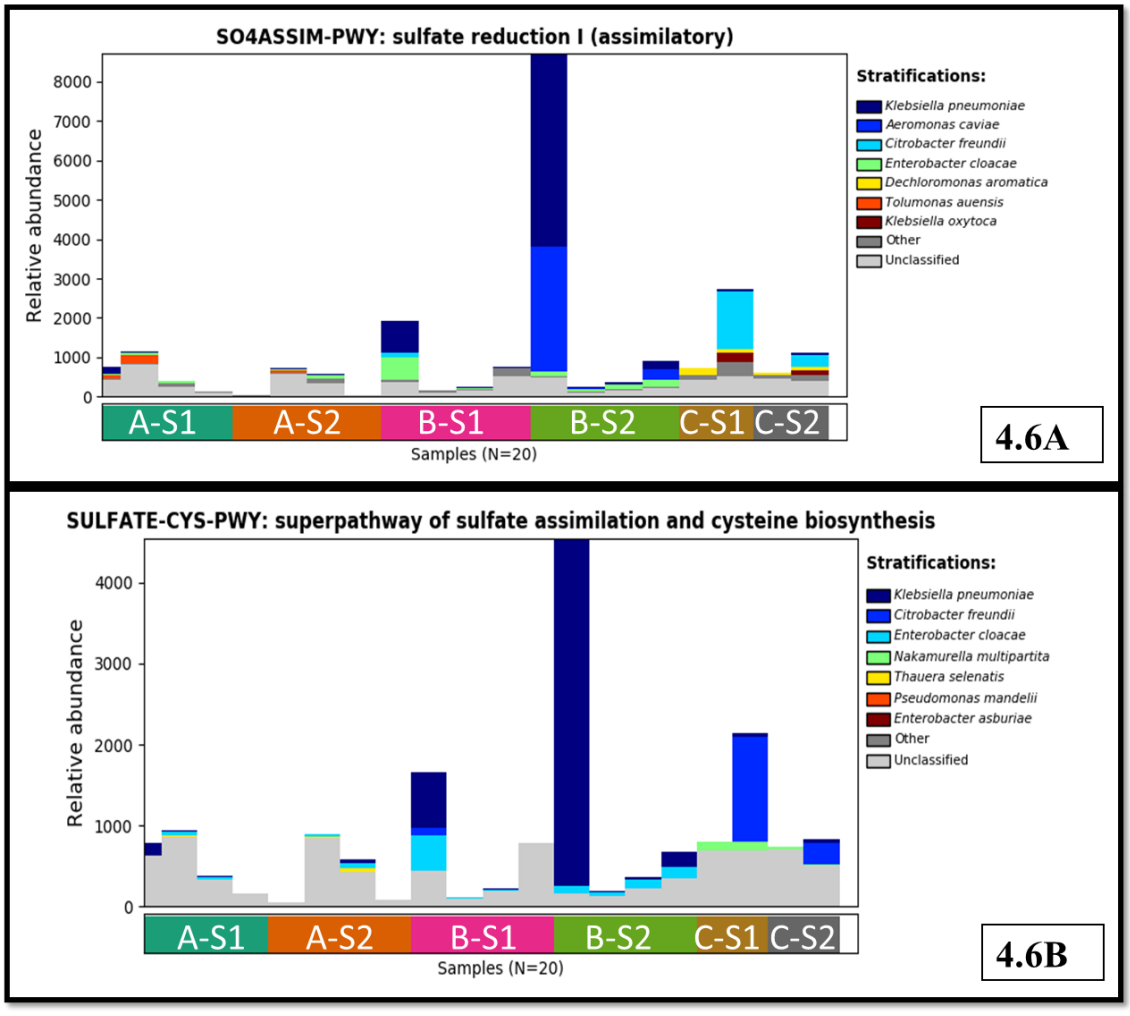


Figure 4.6. Sulfate reduction I (assimilatory, SO4ASSIM-PWY) pathway and sulfate assimilation and biosynthesis superpathway: The top plot (4.6A) represents

abundance of assimilatory sulfate reduction reactions detected in bioreactor samples. The bottom plot (4.6B) shows superpathway which takes sulfate assimilation and biosynthesis of cysteine, which can have sulfur as part of the compound. Both plots represent sulfate utilization by microbes in the bioreactor and showing similar general trends.

Sulfate reduction and sulfate assimilation process were most highly detected in site B stages 1 & 2, and site C stages 1 sample. These samples were dominated by *Klebsiella* and *Aeromonas* taking charge in sulfate reduction, and *Klebsiella* and *Citrobacter* in the assimilation process. As shown in **Figures 4.6**, the sulfate assimilation is not limited to one type of species. Rather, the types of bacteria involved in the process vary depending on the sample sites, as well as between the stages of reactors. Considering the genera presented in **Figure 4.3**, some of the *Paracoccus* species are known for desulfonation and they were detected in most of the samples except for SE1 of both site C and site A samples, and stage 2 of site B samples from SE1 and SE2. Thiobacillus, which was present in most of site A and B samples, is known for the reductive role for sulfur and selenium[35] and *Thiobacillus ferrooxidans* was found to oxidize elemental sulfur and inorganic sulfur to sulfate[7]. *Desulfovibrio* was also found in most of site A and B samples while none in site C, is identified as sulfate reducer which is aerotolerant. It is reported that *Desulfovibrio* is overlooked and underreported due to its nature of overgrowing in mixed cultures and difficulty in identification[36]. The presence of various sulfur processors and their yet-to-be uncovered nature may be due to the complexity of sulfur pathways and their intermediates of the sulfur cycle in FGD wastewater.

4.3.4.B. Pathways for Nitrite and Nitrate

Based on the background study on bioreactors, there are low levels of nitrite and nitrate present in their microbial environments. Multi-step nitrate reduction pathways were found with some major nitrate reducers involved, such as *Thauera, Pseudomonas*, and *Enterobacter* species. However, it is important to note that there is a significantly large proportion of identified pathway were from unclassified bacterial genera, which means there are still largely unknown species that actively participates in the denitrification process. The pathways found through HUManN2 functional classification are as follows: nitrate reduction I (denitrification, DENITRIFICATION-PWY), nitrate reduction V (assimilatory, PWY-5675), nitrate reduction VI (assimilatory, PWY490-3), nitrate reduction VII (denitrification, PWY-6748).

For nitrate reduction VI (assimilatory, PWY490-3), some low ranges of abundances were found with relative abundances around 16~43, however, none of the samples were matched with the known species or genus level. In the case of nitrate reduction VII (denitrification, PWY-6748), the pathway was only found in site B stage 2 reactor at SE1. The specificity of this nitrate reduction VII reaction is the use of a menaquinol-dependent enzyme, while nitrate reduction I reaction uses a cytochrome c-dependent enzyme. Therefore, **Figure 4.7** shows two of the most relevant and abundant pathways, nitrate reduction I and V.

Nitrate reduction is most actively done by *Thauera aminoaromatica* and is more abundant in site A compared to sites B and C. Site C seems to follow a similar trend as site A in denitrification but proceeded with unclassified species, and site B samples showed some activity of *Pseudomonas*, which varies from having *Thauera* in site A and

none in site C samples. One research identified *Thauera* as a core microbiome that showed a positive correlation in denitrification under the anaerobic degradation process[34]. In the case of assimilatory nitrate reduction (**Fig. 4.7B**), most of the bacteria doing such work remains unknown, while *Enterobacter* was identified to be a fraction of classified pathways in site B stages 1 sample. Being facultative anaerobes, Enterobacter can aid in the nitrate reduction process while its other role was identified as selenium-oxyanion reducers[37]. Having the majority of pathways unclassified, an assimilatory nitrate reduction pathway may be conducted with unidentified species, and more studies on microbiome may need to be done to understand the complete relationship. Some traits observed from bioreactor bed from site 1 stage 1 samples was that there were more soft flocs of microbes compared to stage 1 of site 1, site 2 and 3. The stage 1 of site 1 bioreactor sample had unique stinky smell which may indicate more nitrogen and surfur processing bacteria. The pathway analysis result verified that the odor is coming from the nitrous bacteria than sulfur, because site 2 stage 2 showed the highest sulfate reduction pathway compared to site 1 stage 1 sample.

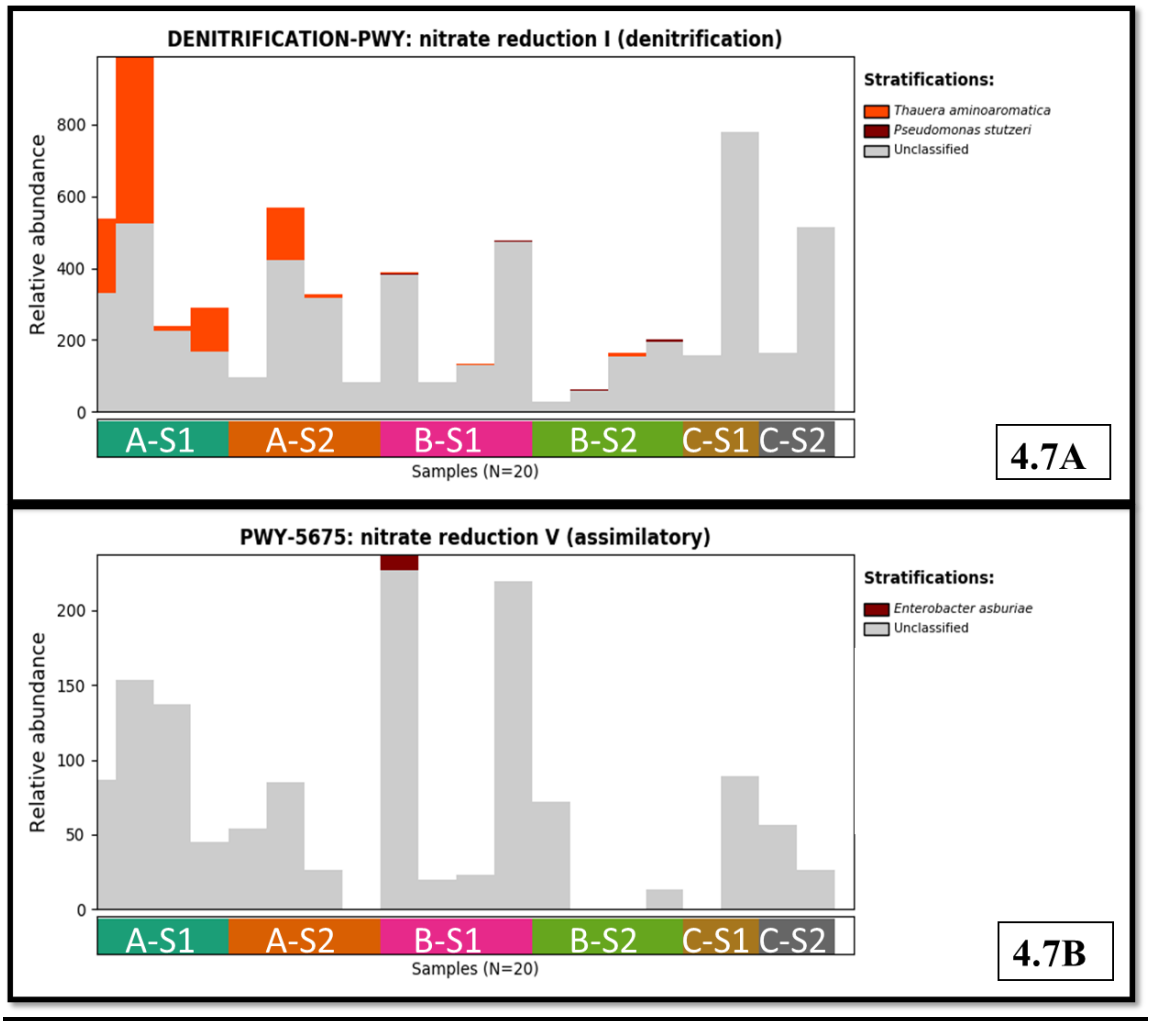


Figure 4.7. Relative abundance on different nitrate reduction pathways: The top figure (4.7A) represents abundances detected in denitrification, and the bottom figure (4.7B) shows assimilatory nitrate reduction pathways. Denitrification activity detected to be the highest in site A stage 1, while the other sites such as site C stage 1, site A stage 2 also showed higher relative abundances. Except for site B, site A and C showed higher denitrification activity in SE2 when samples were collected in colder months. Assimilatory nitrate reduction was the highest in site B stage 1 , followed by site A stage 1.

4.3.4.C. Pathways for Selenium

The most detected activities of selenium metabolism were seleno-amino acid biosynthesis (PWY-6936), two different aerobic and anaerobic heme biosynthesis I and II (HEME-BIOSYNTHESIS-II, HEMESYN2-PWY). **Figure 4.8, 4.9A and 4.9B** were created for each pathway, respectively. Each figure depicts the unique characteristics of each site and the differences that location environment forms the microbial consortium differently in the selenium cycle. More similarity of bacterial genera composition is observed in sites A and B compared to site C, which also represents the status of the FGD plants, demonstrating that sites A and B are active condition while site C is not.

One of the reasons why such a high abundance of selenium biosynthesis can be from the fact that selenate reductase (SerABC) acts with three heterologous subunits, αβγ, and molybdenum, iron, acid-labile sulfide, and heme b as cofactors (Van Hullebusch 2017). Seeing a comparable amount of various selenium compounds such as Se(IV), Se(VI), SeCN, and SeMe in the bioreactor effluents indicates that such active selenium transport and metabolism. Seleno-amino acid metabolism was detected most abundant from the site which has a large volume of production (site B) and then at a site that is assumed to be the most anaerobic environment (site C). It was clear the strains which are the majority from the large volume of the site were different from the dormant site, site C, which are *Klebsiella* and *Citrobacter*, respectively. For site A, there were some samples detected with more than 60% of unclassified genera of the microbiome.

In case of aerobic and anaerobic heme biosynthesis shown in **Figure 4.9A** and **4.9B**, site A and B showed similar trends in consortia in relative abundance and species variations.

Site C was distinctively different where *Citrobacter* were detected under aerobic pathway while *Nakamurella* was detected under anaerobic pathway.

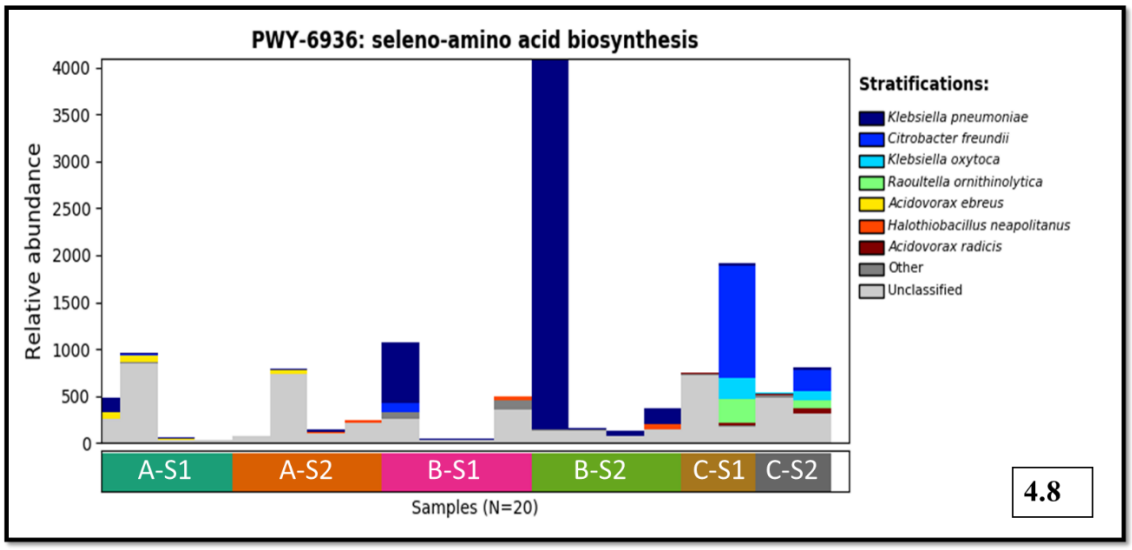


Figure 4.8. Seleno-amino acid biosynthesis (PWY-6936) pathway analysis: The bars represents samples grouped by site reactors and relative abundance. Species identified in samples are stacked in the order of greater abundance to unclassified genera. By the representation of colors, we can see that each site (A, B, and C) has a distinctive combination of microbes acting on seleno-amino acid biosynthesis.

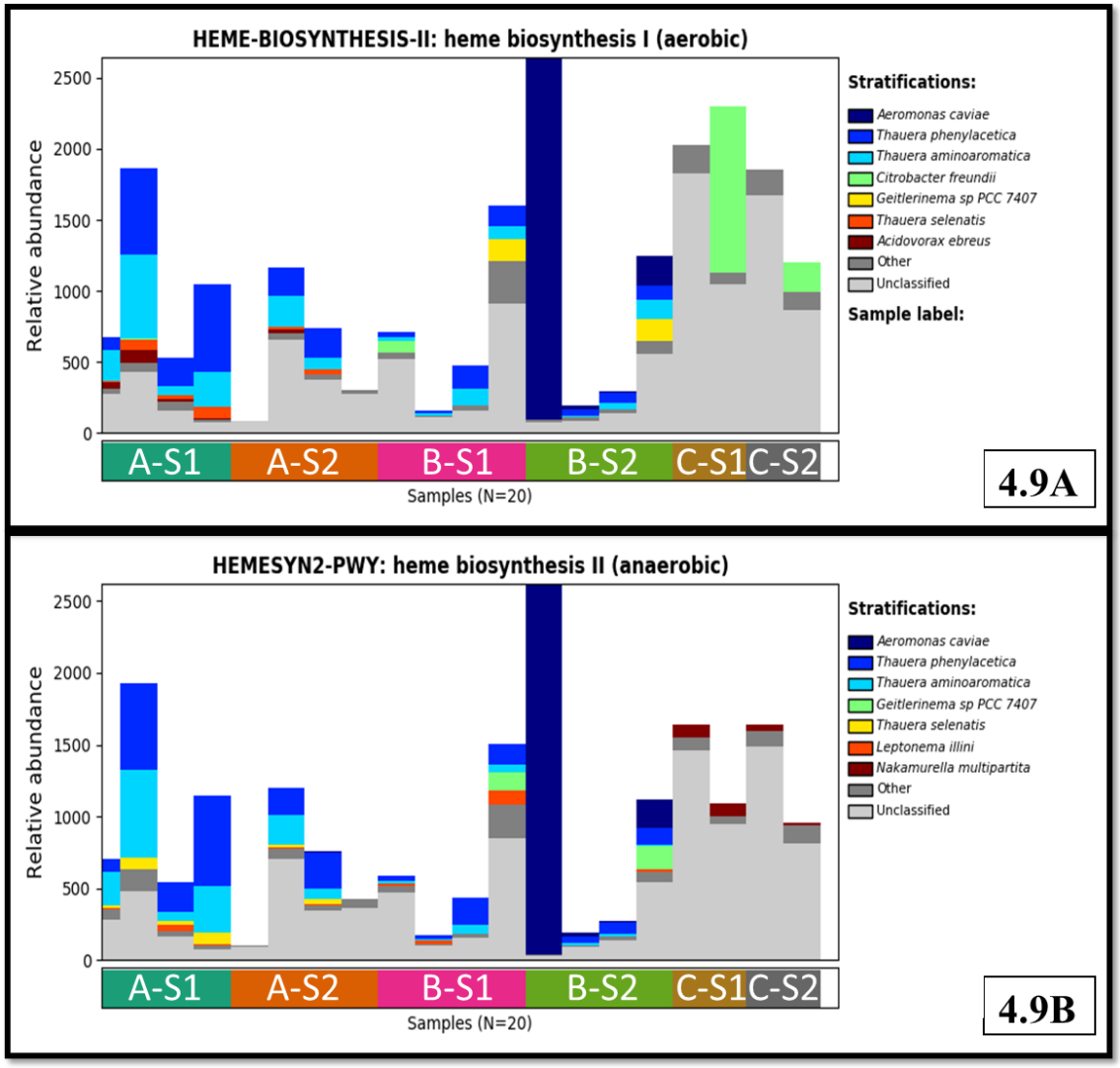


Figure 4.9. Heme biosynthesis I (HEME-BIOSYNTHESIS-II, aerobic) and heme biosynthesis II (HEMESYN2-PWY, anaerobic) pathway analysis: For heme biosynthesis I, *Aeromonas* and *Thauera* are identified to be leading species in sites 1 and 2. In the case of site C, *Citrobacter* was detected in higher abundance which is different from sites A and B. Similar trend persists in Heme biosynthesis II, where *Aeromonas* and *Thauera* are dominant in their relative abundance while in this case *Nakamurella* was seen in site C samples.

4.3.4.D. Pathways for Energy Generation

The TCA cycle, also known as the Krebs cycle, is essential for all organisms to create energy, facilitates the utilization of multiple metabolites, and produces biomolecules that are essential to biosynthetic pathways. For this reason, TCA cycle abundance can be referred to as the amount of biodegradation capacity and viability of a reactor. TCA cycle I (prokaryotic, TCA) and the superpathway of glyoxylate bypass and TCA(TCA-GLYOX-BYPASS) are shown in **Figure 10A** and **10B**, respectively to show microbial energy generation and reducing power formation. The notable difference between TCA cycle I and the superpathway of glyoxylate bypass is that the generation of glyoxylate and succinate by the presence of isocitrate lyase enzyme. The isocitrate lyase enzyme allows bypassing the rest of the TCA cycle and offers a shortcut to faster completion. This enables more accessible energy generation while saving the loss of carbon in the form of CO2. Seeing similar general trends in pathway abundances detected between TCA cycle I and TCA-GLYOX-BYPASS verifies that the bioreactors have less organic carbon throughout the samples and tend to be at similar energy utilization levels. The main genera that were popular in site A and site B are Enterobacter and Citrobacter, while site C TCA cycle is composed of mainly Citrobacter and Nakanurella. This is a clear shift in microbiome which validates that the core acting microbe under inactive plant condition is Citrobacter, and Enterobacter becomes dominant when the plant is under active conditions. Site C showed advances in both TCA activities, which reveals more reduction and energy generation are occurring.

In energy generation and reduction, potentials were higher in all stage 1 sample on site C, A, and B based on TCA cycle relative abundance, which is ranked at the top three, respectively. Activities of *Citrobacter, Enterobacter,* and *Thauera* were detected across these samples. Since stage 1 reactors receive fresh FGD wastewater influent than stage 2 reactors, it is plausible that the microbiome in stage 1 has the potentials to create higher energy, uptake metallic compounds, and provide more reductive potentials at the cellular level. On the other hand, energy-saving fermentation was detected higher in most of the stage 2 samples compared to stage 1 of all sites except for site C. This is understandable since recycled feed of bioreactors (stages) 1 and 2 may already share similar properties in composition and the remainder of compounds needed for energy conservation may be more abundant in reactor 1 than 2.

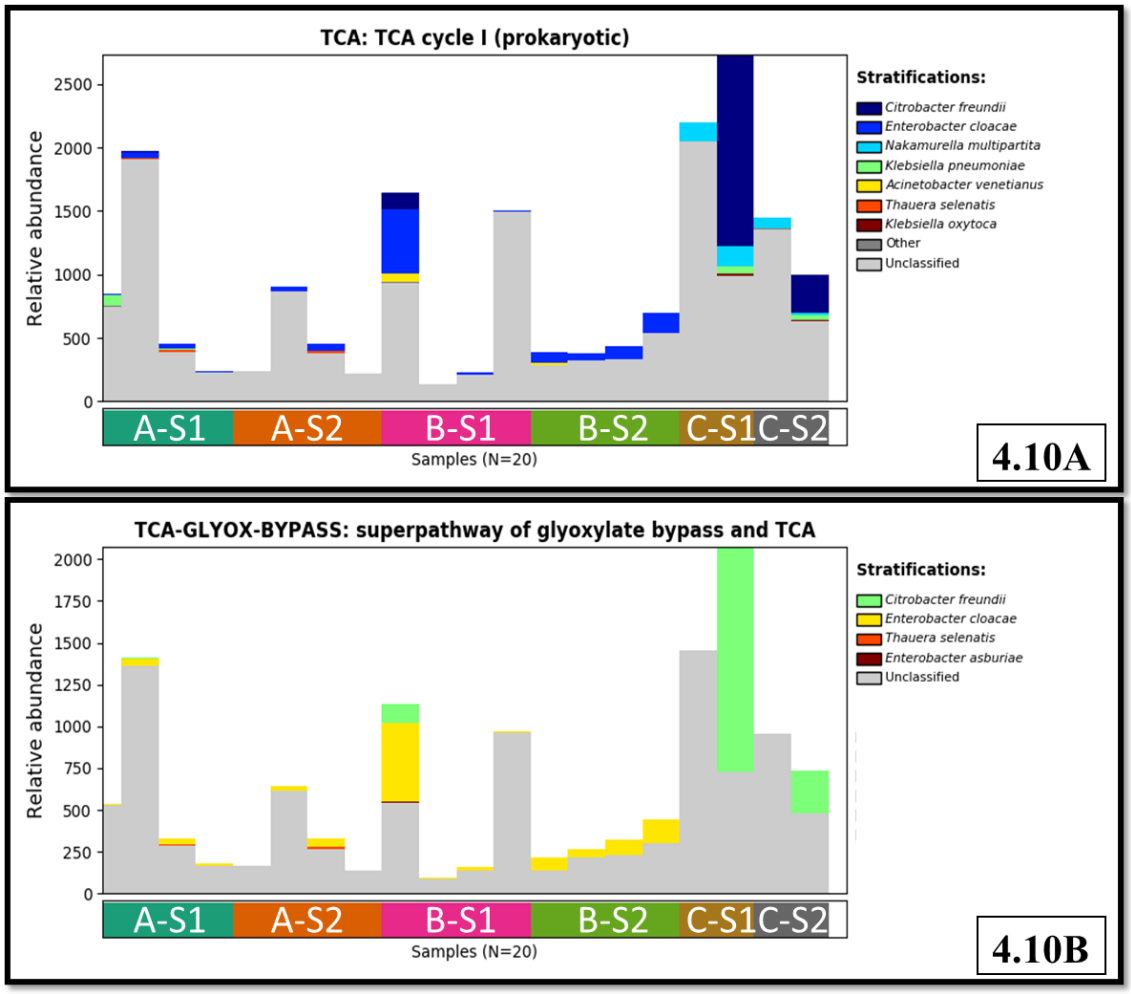


Figure 4.10. **TCA cycle I (prokaryotic, TCA)[10A] and Superpathway of glyoxylate bypass and TCA(TCA-GLYOX-BYPASS)[10B]:** The TCA cycle generates energy in the form of ATP for the cell and generates reducing agents such as NADH. This activity was detected the highest in site C stage 1(C-S1), followed by site A stage 1 (A-S1) and site B stage 1 (B-S1). For glyoxylate bypass pathway, site C stage 1 showed the highest relative abundance with *Citrobacter* detected.

4.3.4.E. Pathways for Fermentation

Among the top 27 identified genera of microbial consortium in bioreactors, the majority of the genera that were identified in most of the samples are facultative anaerobes: *Alicycliphilus, Thauera, Aeromonas, Enterobacter*, and *Raoultella* while *Aeromonas* and *Enterobacter* are identified as fermenters. There are also some genera that are anaerobic such as *Dechloromonas* and *Desulfovibrio*, and bacterial genus identified as anoxic as *Chlorovavulum*. As the microbial consortium is formed favorable to fermenters, denitrification, desulfurization, and demineralization processes are expected to thrive in the FGD bioreactor environment.

The pyruvate fermentation to isobutanol (PWY-7111) shown in **Figure 4.11** was examined for engineering potentials as well as conditions of bioreactors. This pathway occurs when L-valine is unavailable for the TCA cycle (especially in yeast), it uses the Ehrlich pathway to form isobutanol. The capability of bioreactors for such reaction has the potentials to be used as low-cost, low-energy reaction beds for microbial fuel cells because providing L-valine can be expensive. The most abundant isobutanol production was detected in Stage 2 of site A, 2, then both stage 1 and Stage 2 of site C. This implies that Stage 2 bioreactors are more under anaerobic conditions that promote fermentation led by *Chlorobaculum, Klebsiella, Aeromonas, Citrobacter*, and some trace amounts of other bacterial compositions.

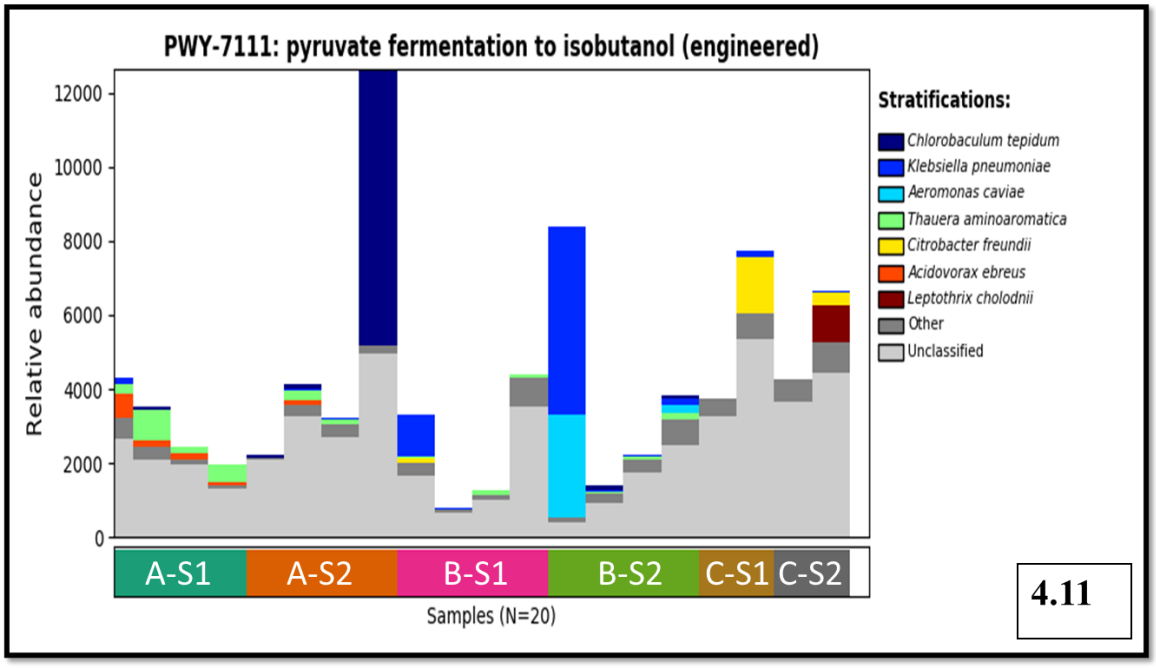


Figure 4.11. Pyruvate fermentation to isobutanol (PWY-7111): The most abundant pyruvate fermentation activity was detected in site A stage 2 (A-S2) followed by site B stage 2 (B-S2). It is estimated that the bioreactor culture can be utilized as a potential source of energy generation.

4.3.4.F. Pathways of sulfhydrylation and sulfide oxidation

Additionally, other types of sulfur utilization at the bacterial levels were detected: superpathways of L-methionine biosynthesis by sulfhydrylation (PWY-5345) and sulfide oxidation by phototrophic sulfur bacteria (PWY-6676) shown in **Figure 12A** and **12B**, respectively. The bacterial species were not identified in any of those pathways, but they are assumed to be other branches of bacterial species that are not classified to current findings. L-methionine is one of the kinds of amino acids that contains sulfur within its chemical structure in the sulfhydrylation pathway. This is important for sustaining cellular life for humans as well as many bacterial species and synthesized using a TCA

cycle intermediate, oxaloacetate. Some microorganisms are assumed to be performing direct sulfhydrylation are *Saccharomyces cerevisiae*, the spirochete *Leptospira meyeri*, *Bacillus subtilis*, *[Brevibacterium] flavum*, *Corynebacterium glutamicum*, *Pseudomonas aeruginosa*, and *Pseudomonas putida* (MetaCyc Pathway: superpathway of L-methionine biosynthesis (by sulfhydrylation) n.d.). It is also found that *E. coli* K-12 strains use transsulfuration pathway (Battley 1988).

In the case of the sulfide oxidation pathway by chemotrophic bacteria, it is known to use inorganic reduced sulfur compounds. Some species that can go through this pathway are *Allochromatium vinosum*, *Chlorobaculum tepidum*, *and Chlorobaculum tepidum TLS*. Such phototrophic bacteria performing sulfide oxidation is highest in the second reactor of site A compared to other sites. Other bioreactors showed a rather alleviated level of sulfide oxidation which implies site A stage 2 had relatively higher sulfide contents and had more oxygen content or had more nitrate or nitrite within the bioreactor to act as electron acceptor (Fuseler et al. 1996).

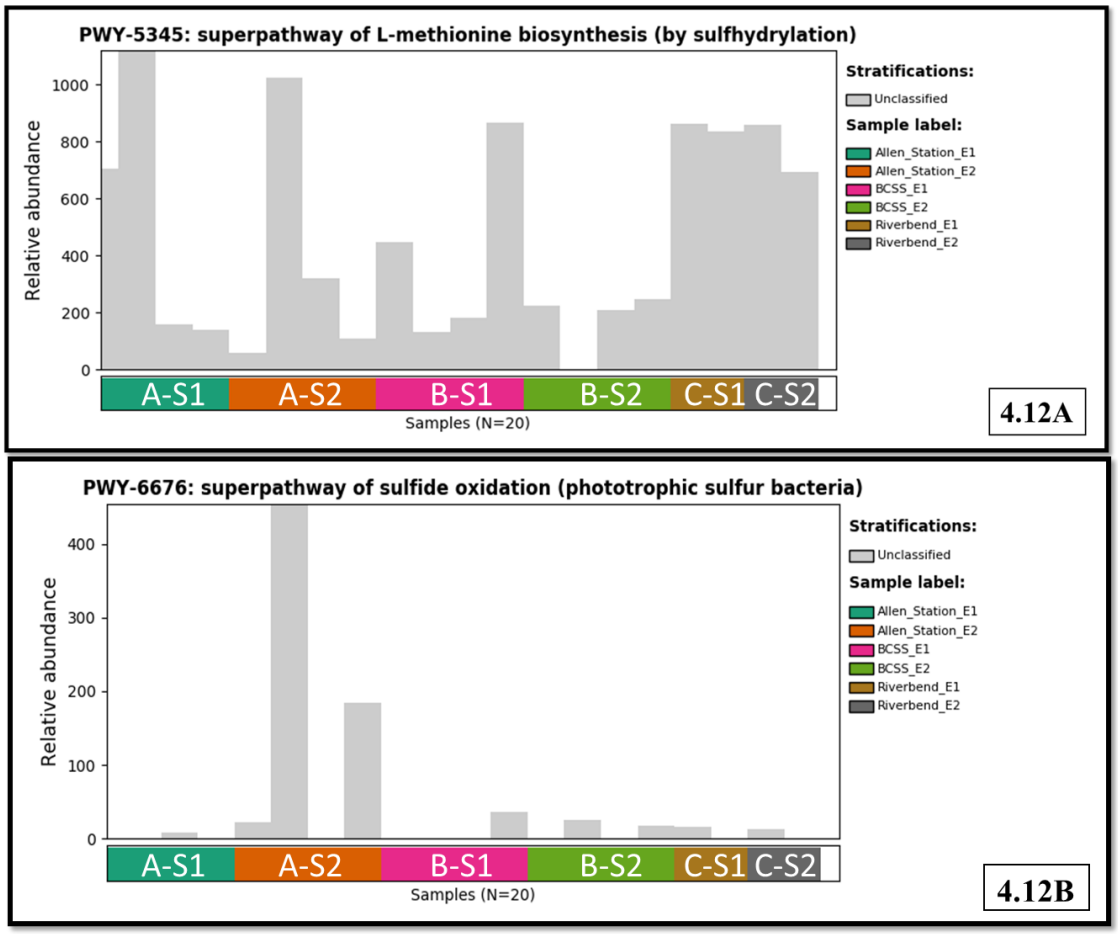


Figure 4.12. Superpathway of L-methionine biosynthesis (by sulfhydrylation, PWY-5345)[12A], and Superpathway of sulfide oxidation (phototrophic sulfur bacteria, PWY-6676) [12B]: Despite the good amount of abundance detected, the pathway does not show any identified species. This may be because L-methionine is essential to every life of microorganisms and ubiquitous. The highest abundance was detected in site A stage 1 (A-S1), and the most consistent, relatively higher level of detection was observed in both reactors from site C. The highest abundance of sulfide oxidation was observed in site A stage 2 (A-S2). Other reactors show the minimum or lower level of such a pathway.

4.3.5. Bacterial Information: Genus-Level, Microbial Description of the identified bacterial community

Bacterial genus from sampling events is presented in alphabetical order. Underlined bacterial genus represents higher abundance from SE3 and SE4.

Acinetobacter: The genus Acinetobacter has 38 identified species. They have strictly aerobic, nonfermentative, gram-negative characteristics.

Aeromonas: *Aeromonas hydrophila* is a gram-negative bacteria. These organisms have an oxidase mechanism and are facultatively anaerobic. One out of six *Aeromonas* species was found to be pathogenic to humans and fish. *A. hydrophila* is considered as heterotrophic organism and opportunistic pathogens. Having the ability to live in aerobic and anaerobic conditions, they can also ferment glucose. It has fimbriae (pili) structure which helps the bacteria attach to host organisms and invade cells. They can reproduce asexually while invading the host. *A. hydrophila* can be found in aquatic environments as well as in foods that are present worldwide. It can cause infections that are fatal in humans, targeting the either intestinal or non-intestinal area. Some examples of the disease that *Aeromonas* species can cause are septicemia, meningitis, pneumonia, and gastroenteritis. The three major wound infections which are caused by *Aeromonas hydrophila* are cellulitis, myonecrosis, and ecthyma gangrenosum. Among these, cellulitis is caused as a result of injury or sepsis, which is most common.

Alicycliphilus: Under the phylum Proteobacteria, *Alicycliphilus* contains a single species of *A. denitrificans* which can do denitrification process.

Azoarcus: *Azoarcus evansii* is a species of proteobacteria. Having a rod-shaped cell body with a polar flagella tail, these bacteria are related to plants. Being aerobic, aromatic

bacterium, they were used to learn about the enzymes utilized in the metabolism process. It has two independent catabolic pathways function to convert benzoate to benzoyl-CoA, which involves 150 Proteins in the metabolism. The suitable growth condition is at 37ºC and pH of 7.8.

Brevundimonas: Under the genus of proteobacteria, *Brevundimonas* are gram-negative, non-fermenting, aerobic. They are ubiquitous in the environment but hard to be isolated in clinical conditions. They are one of the bacterial species which showed a high survival rate under Martian conditions (EOL website, https://eol.org/pages/97684/articles).

Chlorobaculum: *Chlorobium tepidum* is a thermophile, anaerobic phototrophic bacterium. Under the Chlorobiaceae family, it is considered a model bacterium. They can live in anoxic and sulfide-rich waters, mud, sediments, microbial mats. *C. tepidum* and other green sulfur bacteria are anaerobic phototrophs, and scientists considered them to be one of the first photosynthetic organisms that arose on Earth when there was low oxygen in the atmosphere in the archaic periods. The optimal growth is observed at 47–48°C, and they are directly involved in the global sulfur and nitrogen cycle.

Citrobacter: It is ubiquitous in soil, water, wastewater, etc., including the human intestine. They are rarely pathogenic, some cases detected from urinary tract infection, infant meningitis, and sepsis.

Dechloromonas: Under the genus of proteobacteria, Dechloromonas has four species: (1) dissimilatory perchlorate reducing bacterium (DPRB), *D. agitate*; (2) *D. aromatica*, which does anaerobic benzene reductions, perchlorate reduction, and oxidation of chlorobenzoate, toluene, and xylene; (3) *D. denitrificans*, a motile bacterium produces

N2O; (4) *D. hortensis*, a motile, facultatively anaerobic, (per)chlorate reducing bacteria.

Dermatophilaceae_unclassified: The family *Dermatophilaceae* is gram-positive under *Actinomycetales*, found in human skin and fish guts (EOL website, https://eol.org/pages/7904/articles).

Desulfovibrio: They have gram-negative, sulfate-reducing characteristics. They are mostly found in water environments with high organic content, water-logged solids, or fractured rock aquifers. They can form a major community in extreme oligotrophic habitats. Despite being sulfate-reducers, they are not obligate anaerobes. They are aerotolerant. Desulfovibrio is known for its ability in the bioremediation of toxic radionuclides through reductive bioaccumulation (EOL website, https://eol.org/pages/83219/articles).

Erythrobacteraceae_unclassified: *Erythrobacteraceae* is a type of alpha-proteobacteria, having gram-negative, rod-shaped, non-spore-forming characteristics. Also, it is chemo-organotrophic, aerobic, and contains chlorophyll A and carotenoids (Lee et al., 2005).

Thiobacillus: *Thiobacillus* is under a phylum of proteobacteria, class of betaproteobacteria, under the order of Nitrosomonadales and family of Thiobacillacaeae. Species under *Thiobacillus* are *thioparus, denitrificans*, and *thiophilus*, which are all autotrophs that involve the Calvin cycle. *Thiobacillus ferrooxidans* are known to produce Se1 protein. This protein can reduce selenium ions into insoluble selenium, which has selenium binding functions (Floren et al. 2015). *Thiobacillus thiooxidans* and *T.ferrooxidans* oxidize elemental sulfur and inorganic sulfur into sulfate, and their optimum pH ranges are different (Singh 2010; Tacconelli et al. 2018). Based on the

conditions of the 2-stage bioreactor, we can assume that *thioparus* and *denitrificans* activity will be most dominant among other *Thiobacillus* species under ~pH7. Both species are capable of nitrate reduction to nitrite, and only *denitrificans* are capable of nitrate reduction to nitrogen gas.

Thiomonas: It is a gram-negative, non-spore-forming bacteria under the family of Comamonadaceae.

Tunalikevirus: This bacteria can adhere to the host cell's cytoplasm and inject DNA through long pilus, which causes viral replication of their genetic material.

Enterobacter: *Enterobacter* is a nosocomial, opportunistic pathogen that can cause infections, or some of its types are useful to humans. It is motile, rod-shaped, some of which are encapsulated, and has peritrichous flagella. As facultative anaerobes, some *Enterobacter* bacteria can go through fermentation using both glucose and lactose as a carbon source. Gas is produced from the metabolic processes, but not hydrogen sulfide. *Enterobacter* can be found on human skin and plants as well as in soil, water, sewage, intestinal tracts of humans and animals, and some dairy products (Health Canada). When infected with them, possible symptoms include bacteremia, lower respiratory tract infections, skin infections, soft tissue infections, urinary tract infections, UTI, endocarditis, intra-abdominal infections, septic arthritis, osteomyelitis, and ophthalmic infections.

Yualikevirus: It is under the family Siphoviridae and genus viruses. Other bacteria serve as a natural host, and Yualikevirus is transmitted through passive diffusion. Three species were found in this genus, including species *Pseudomonas phage yua*. The virus attaches to the host cell using its terminal fibers and introduces viral DNA into a periplasmic

space of the host cell. The DNA genome is circularized or integrated into the host's chromosome before transcription and translation.

All of the information above in this subsection was retrieved from the Encyclopedia of Life (EOL) bacterial website.[https://eol.org/]

4.3.6. Statistical Analysis of bacterial consortia and ARGs in the FGD Bioreactors

The heatmap of abundance across samples were shown in **Figure 4.13**. Since there are too many of the ARGs detected from CARD databased, the data was sorted to represent the top 20 most abundant ARGs in **Figure 4.14**. One of the most abundant genes was AAC(6')-Ib7 Salmonella enterica, which transfers acetyl group to aminoglycoside antibiotics with Acetyl Co-A, which makes the antibiotics inactive when faced with bacteria such as salmonella. This was the most abundant in sites A and B rather than site C. The second gene ranked is C-reactive protein (CRP) of Escherichia coli K-12 strain. This protein encoded from this gene is made in the liver, in response to acute inflammation of bacterial species such as pneumococcus. This gene was found mostly in site C and second stage bioreactors of site B. The third one in the rank is LCR-1 Pseudomonas aeruginosa, which is a beta-lactamase resistance gene that hydrolyzes the beta-lactam ring and turns it into substituted beta-amino acid This gene was most abundance in stage 1 bioreactors of site A, which showed clear difference compared to other locations. The fourth one, vanXYL is a vanXY variant of the vanL gene cluster, which is resistant to glycopeptide antibiotics which inhibits the formation of the bacterial cell wall. This gene was mostly found in stage 2 bioreactor of site A, which showed that

LCR-1 and vanXYL are predominant in site A, which is considered to be active, and specific site receiving regional river water, while all other location takes the water sourced from regional lakes. The fifth most abundant gene was APH(3')-Ib Pseudomonas aeruginosa, which confers antibiotic resistance of *P. aeruginosa* with an effect of aminoglycoside phosphotransferase. This gene was found to be present in nearly all *P. aeruginosa* bacteria isolated from bloodstream samples of academic medical institutions in a study (Murphy-belcaster et al. 2021).

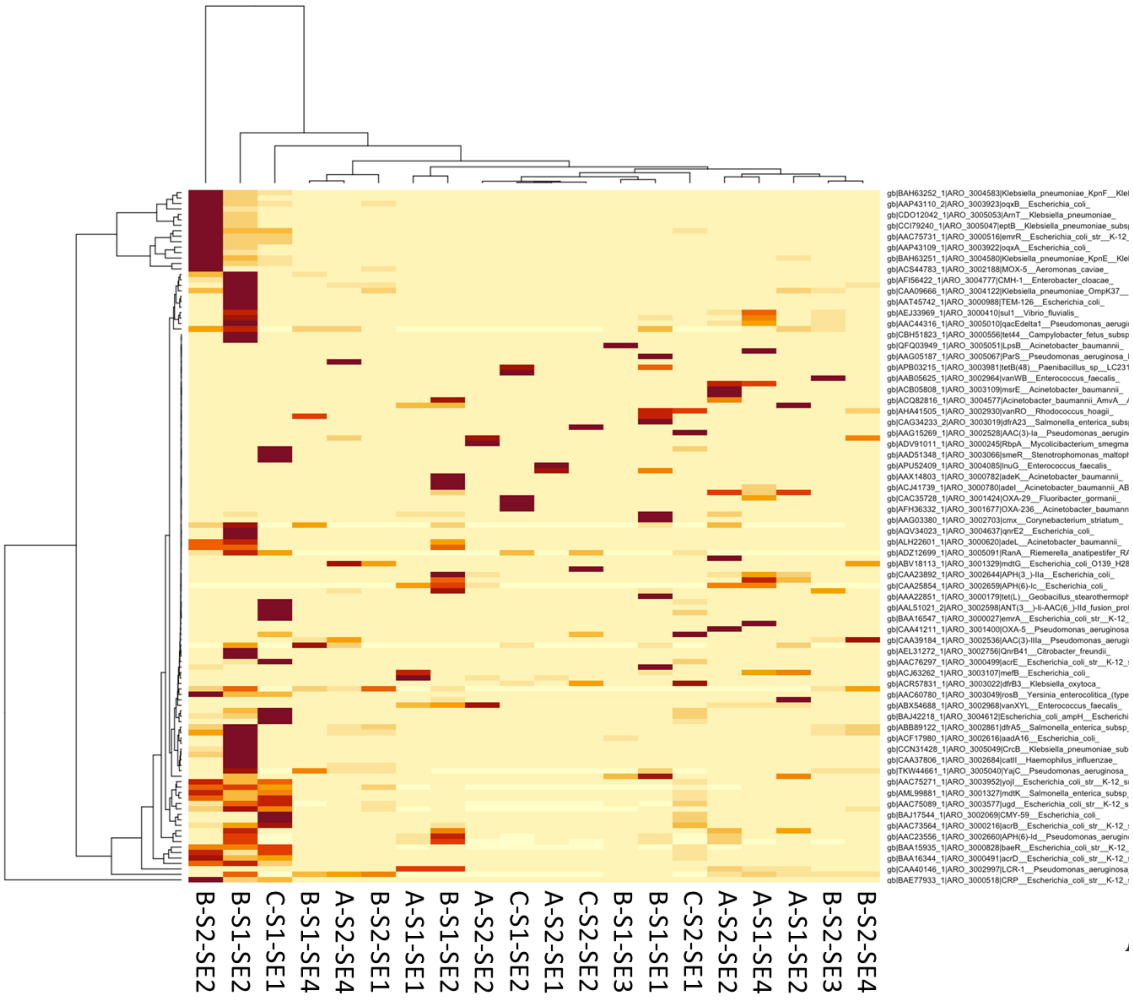


Figure 4.13. Heatmap of ARGs abundance of all genes detected

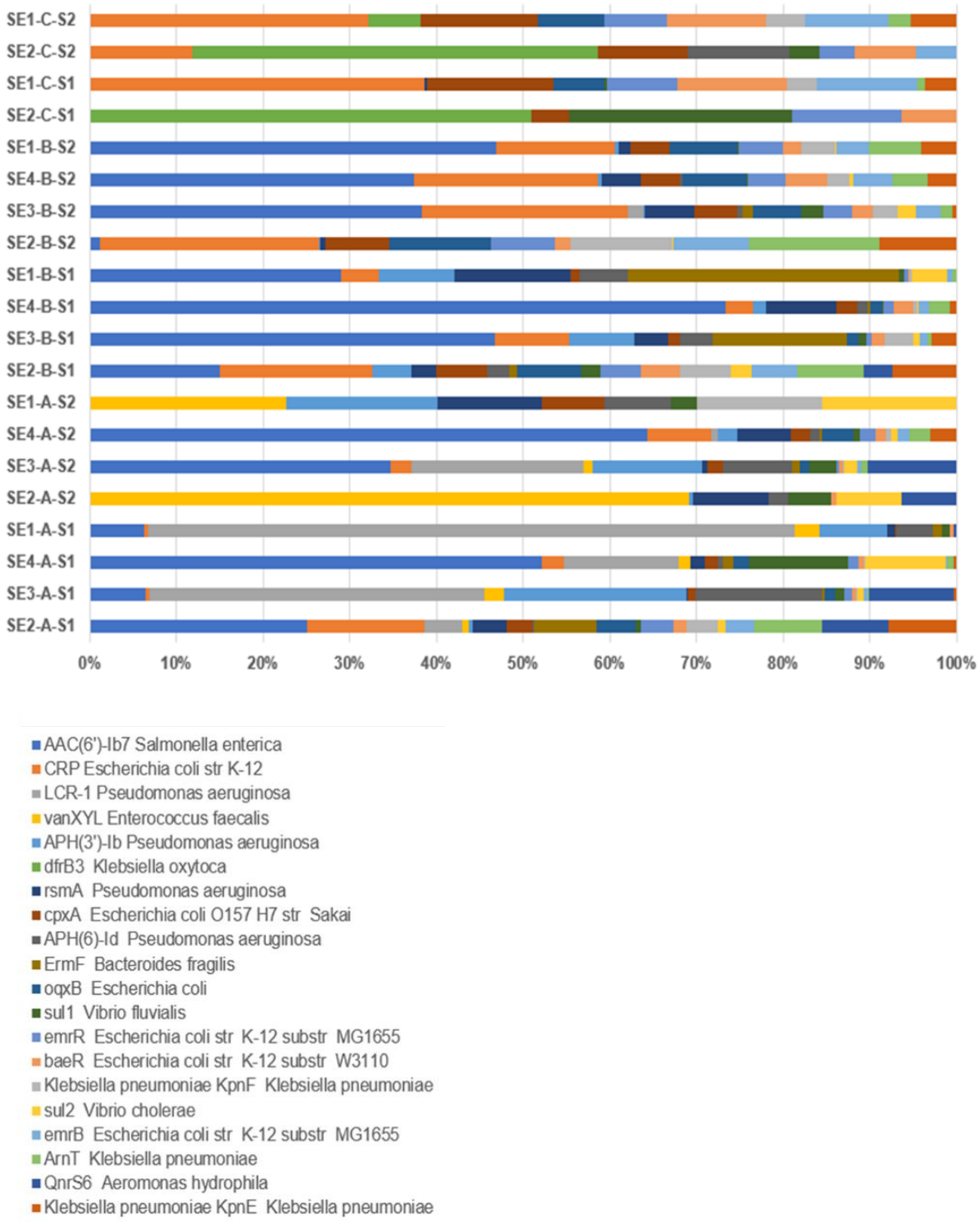


Figure 4.14. Stacked relative abundance bar chart of top 20 ARGs detected across samples

Principal coordinate analysis on bacterial and antibiotic resistance gene abundances across samples was shown in **Figure 4.15** and **Figure 4.16**, respectively. Site A and B showed a more similar pattern of distribution across the plotting plane, where bacterial abundances were spread out across the y-axis, PC2, and ARG abundances were greatly distributed across the x-axis, PC1. There was some seasonal variation observed, but the pattern was less distinctive compared to site locations. This represents that site was a major contributor to the microbiome and antibiotic resistance distribution rather than seasonal effect.

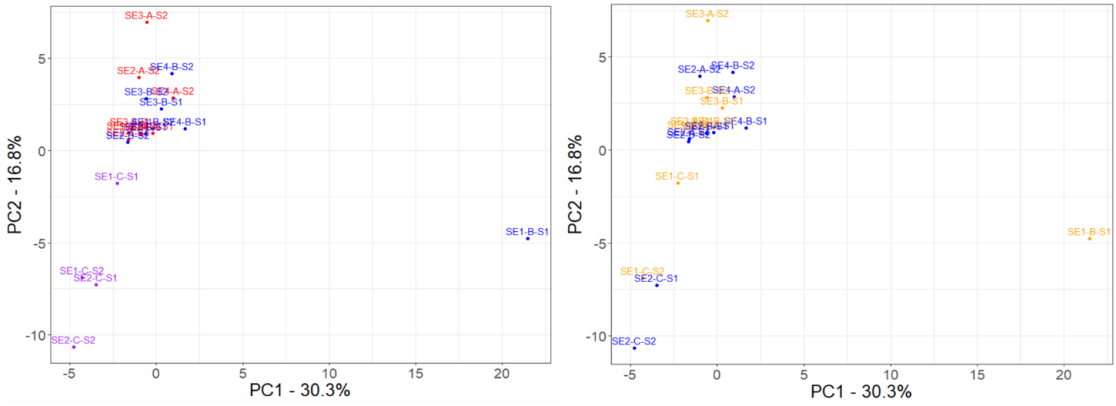


Figure 4.15. Principal coordinate analysis of bacterial distribution across samples by sites (left) and season (right). Site A, B, and C are colored in red, blue, and purple, respectively (left). Samples collected in the summer and winter months are colored in orange and blue, respectively (right).

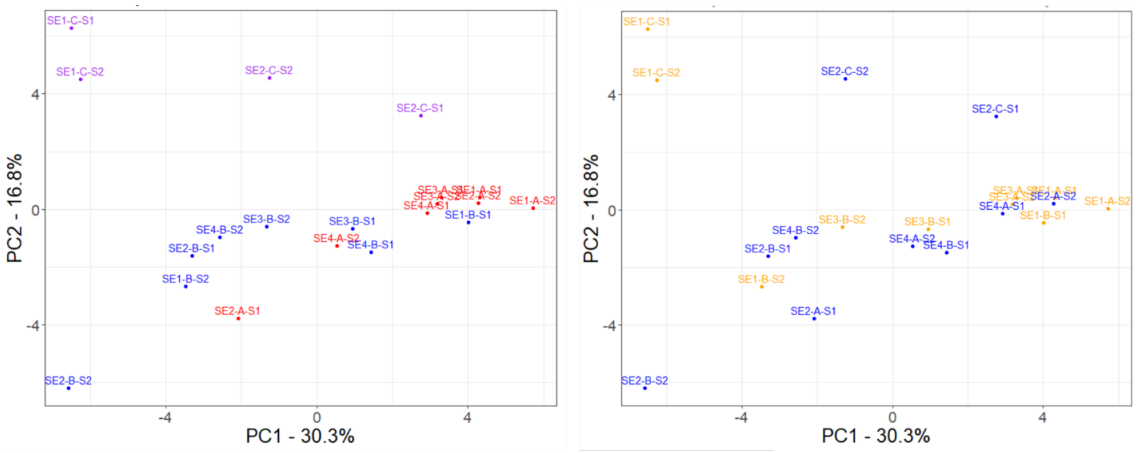


Figure 4.16. PCoA plot of ARGs distribution across samples by sites (left) and season (right). Site A, B, and C are colored in red, blue, and purple, respectively (left). Samples collected in the summer and winter months are colored in orange and blue, respectively (right).

The top 20 most detected ARGs were shown in percent relative abundance in **Figure 4.17** in the visualization of ARGs composition divergence and development by sample. The change in ARG compositions from stage 1 to stage 2 bioreactors at each site showed that expression in bacterial resistome towards antibiotics evolve and adapt to conditions of the bioreactor stage, where the largest abundance of ARG mechanisms can change to different target antibiotics. For example, from site A stage 1 bioreactor, LCR-1, an ARG which confers beta-lactam resistance showed the largest percentage of its amount decreased its presence in stage 2 bioreactor, where vanXYL, a multi-drug, glycopeptide resistance showed one of the largest abundances among ARGs. Likewise, the transition from stage 1 to stage 2 bioreactors in site B showed that stage 1 showed higher abundance in ErmF gene, resistant to macrolides, streptogramin, and lincosamide,

decreasing its abundance in stage 2. The Site B stage 2 bioreactor showed an increased amount of oqxB and emrR genes, which of both confers resistance to fluoroquinolones.

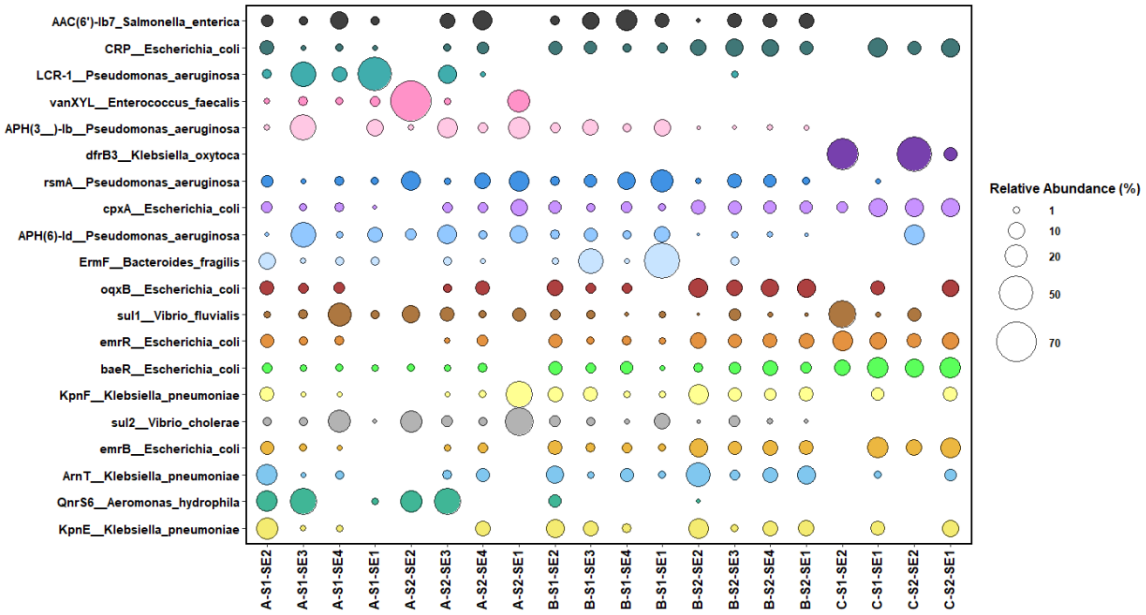


Figure 4.17. The ARG divergence and development in percent relative abundances across samples

Antibiotic resistance resistome from the ShortBRED analysis was then categorized by the mechanism of AR, which is shown in **Figure 4.18**. The 50 most abundant ARGs were analyzed for the presence of ARG and resistance mechanisms. The AR mechanism was categorized into four: *inactivation, modification, pumping out, and reduced permeability*. *Inactivation* includes the transfer of chemical functional groups to an antibiotic such as acetyl or phosphate group to the antibiotic to modify the structure and effect of the antibiotic. *Modification* is change of receptors or receiver of antibiotic target to be altered or replaced so that the antibiotic no longer can bind to its original targeted site. *Pumping out* mechanism creates efflux pump of antibiotics, removing the antibiotic from the

bacterial cell. The last mechanism of AR is creating *Reduced permeability*, which is enhancing cell walls to create barriers toward antibiotics from entering the cytoplasm of bacteria. The mechanism that detected the largest number of ARGs is pumping out that creates efflux of antibiotics. And site B showed the most ARGs variation and amount of resistome compared to other sites and stages.

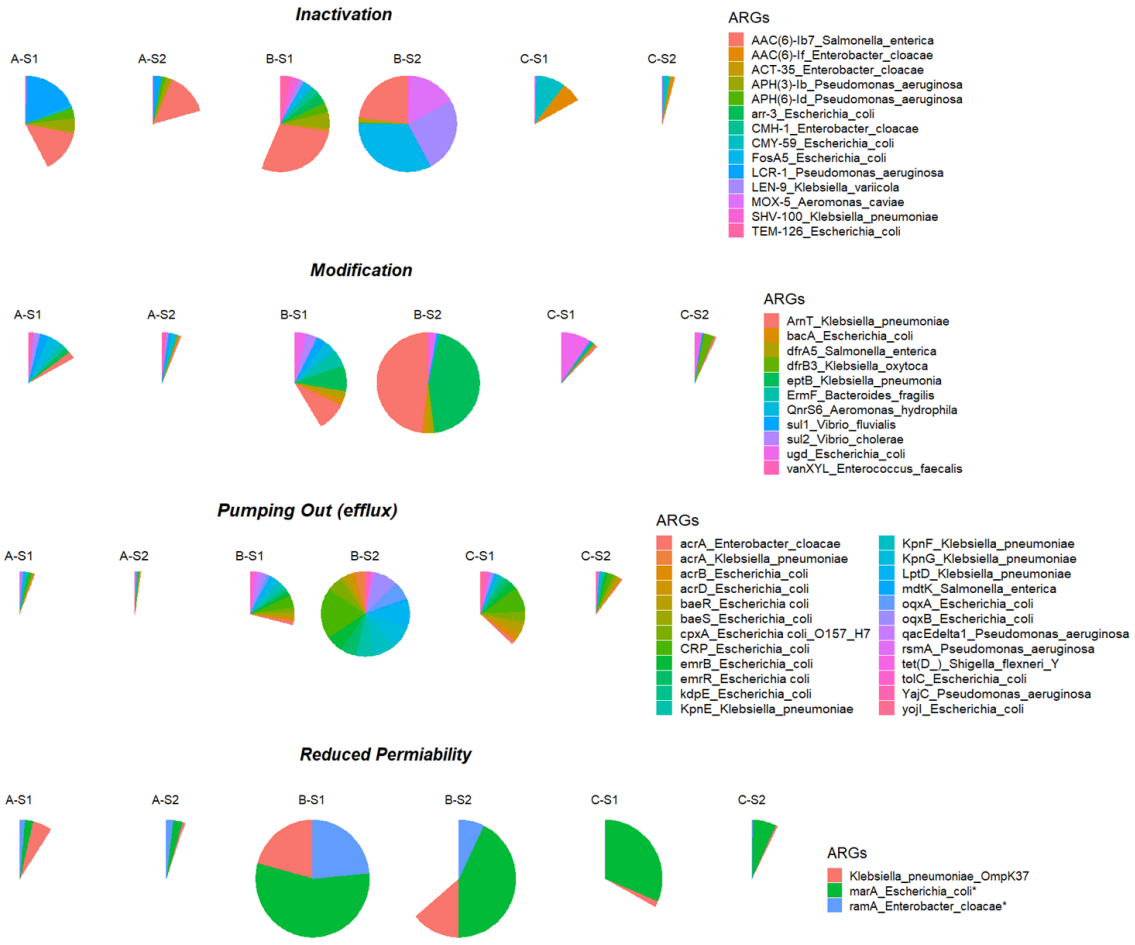


**Figure 4.18. The AR characteristics and Mechanisms of FGD bioreactors at different stages. ***For marA and ramA genes under Reduced Permeability also holds Pumping out (efflux) mechanisms to be expressed in the bacteria.

4.4. **Discussion**

4.4.1. Bacterial consortia variances on geographical, stages, and climate factors

The metagenomic consortia analysis showed unique microbiomes depending on sites and sampling periods in **Figure 4.3**. While having some variances, there were few patterns of the abundant bacterium within the microbiome consistent across all sites. For example, continuous horizontal bands of *Thauera, Klebsiella, Alicycliphilus, Pseudomonas, Leptonema, and Yualikevirus* demonstrate some trend in the shared microbiome of FGD bioreactors despite the difference in location, stages, and climate of sampling. Among samples, Site C showed unique patterns of having compared to other samples. Because the plant site is being retired, it is plausible that the bioreactor microbiome tends to gravitate towards more of the fermentation-favorable, anaerobic environment compared to other sites of site A and site B.

As described in **Figures 4.4 and 4.5**, warmer and colder months showed variation in bacterial species and distribution. Seasonal factors such as temperature, influent of environmental water, or weather conditions seem to be affecting the microbiome of the reactors. While showing more diversity of bacterial genera in warmer months, site C samples presented a great shift in its microbiome. This may be affected by the fermentation process making it hard for certain bacteria to survive and prosper in changing temperature conditions, while site A and site B contain more anaerobic, facultative bacterial species such as *Thauera*.

4.4.2. Energy generation, Heavy mental bioremediation potentials through bacterial pathways

Selenium biosynthesis was confirmed in most of the samples, especially in site B samples utilizing *Klebsiella pneumoniae* and *Citrobacter freundii,* and in site C samples utilizing *Citrobacter freundii* and *Klebsiella oxytoca*. In the case of Site A samples, it seems to be utilizing *Klebsiella pneumoniae* and *Acidovorax ebreus.* Heme biosynthesis was confirmed in both aerobic and anaerobic circumstances to be used as a cofactor for selenium uptake, showing *Citrobacter* and *Thauera* genera as the most abundant workers for aerobic conditions, while two different *Thauera* species were most active under anaerobic conditions.

In the case of the sulfur cycle, sulfate reduction, and sulfate assimilation process, they were most highly detected in site B stage 1 & 2, site C stage 1 samples. These samples were dominated by *Klebsiella* and *Aeromonas* taking charge in sulfate reduction, and *Klebsiella* and *Citrobacter* in the assimilation process. As shown in **Figures 4.6**, the sulfate assimilation is not limited to one type of species. Rather, the types of bacteria involved in the process vary depending on the sample sites, as well as between the stages of reactors.

Nitrate reduction is most actively done by *Thauera aminoaromatica* and is more abundant in Site A sites. Site C seems to follow a similar trend as site A in denitrification, and site B samples showed some activity of *Pseudomonas*, which varies from having *Thauera* in site A and site C samples. In the case of assimilatory nitrate reduction, most of the bacteria doing such work remains unknown, while *Enterobacter* was identified to be a fraction of the site B stage 1 sample.

In energy generation and reduction potentials, the result from stage 1 in site C, site A, and site B was ranked at the top. Activities of *Citrobacter, Enterobacter,* and *Thauera* were

detected across these samples. Since stage 1 reactors receive fresh FGD wastewater influent, it makes sense that the microbiome in stage 1 can create higher energy, potential metal uptake, and reductive potentials at cellular levels. On the other hand, in the case of energy-saving fermentation, most of the stage 2 samples were detected higher for such pathways compared to Stage 1 of all sites except for site C. This is understandable since site C is close to, or already at retirement stage in the sampling timepoints.

While having unclassified and unidentified significant portions of sequenced genes, there are some limitations in finding the exact target metabolism that we want, such as other heavy metal uptake or their end cycle. Certainly, chemical, hydrological, and physical factors of the bioreactors need to be considered for a complete study on FGD bioreactors. That said, this study provided a clear overview of bacterial distribution and abundance in sites of interest, as well as some pathways identified for their interaction with wastewater materials.

4.4.3. Antibiotic Resistance Conferring protein sequence and mechanism analysis

In the previous section 4.3.6, antibiotic resistome distribution changed depending on the site location and weather. Depending on the stage of bioreactors that are fed by the same FGD wastewater in series, the microbial distribution showed some shifts in specific ARG types showing that the ARGs abundance and variance can be expressed differently at each stage of the bioreactor. It showed the potential of environmental stress shaping the microbiome and resistome. Also, the mechanism of ARGs was categorized by four, where most variances and abundance were observed in site B. The **Table 4.3**. shows the detailed mechanism of AR and roles of the top 20 abundant AR conferring genes. One of

the most abundant ARG detected are CRP, AAC(6')-Ib7, ArnT, LptD, and oqxB genes which is three out of five confers multidrug resistance. From the analysis shown in earlier sections, antibiotic resistance along with multidrug resistance are present in the industrial water system, and study of environmental factors that shapes ARGs pattern and resistome evolution is needed.

Table 4.3. Mechanism of AR in top 20 abundant ARGs

Mechanism	Family	Antibiotics	Description
Pumping out (Efflux)	CRP Escherichia coli str K-12	macrolide, penam, fluoroquinolone	global regulator that represses MdtEF multidrug efflux pump expression
Inactivation	AAC(6')-Ib7 Salmonella enterica	Aminoglycoside	Transfers an acetyl group from acetyl-CoA to the antibiotic, modify antibiotic
Modification (Abt target alteration)	ArnT Klebsiella pneumoniae	peptide	polymyxin-class Atb resis./ which is used to treat multidrug resistance
Pumping out (efflux)	LptD Klebsiella pneumoniae	carbapenem, peptide, aminocoumarin, rifamycin	LPS transport in ABC transporter efflux system
Pumping out (efflux)	oqxB Escherichia coli	tetracycline, glycylcycline, nitrofuran, diaminopyrimidine, fluoroquinolone	Efflux pump membrane transporter
Modification (Abt target alteration)	eptB Klebsiella pneumonia	peptide	Kdo(2)-lipid A phosphoethanolamine 7"-transferase
Pumping out (efflux)	KpnF Klebsiella pneumoniae	peptide, tetracycline, cephalosporin, rifamycin, aminoglycoside, macrolide	antibiotic efflux
Inactivation	FosA5 Escherichia coli	fosfomycin	Fosfomycin resistance protein
Pumping out (efflux)	KpnG Klebsiella pneumoniae	peptide, carbapenem, fluoroquinolone, aminoglycoside, penem, penam, cephalosporin, macrolide	antibiotic effluex pump. KpnG consists of ~390 residues and resembles EmrA of E. coli. Disruption of the pump components KpnG-KpnH signficantly decrease resistance to azithromycin, ceftazidime, ciprofloxacin, ertapenem, erythromycin, gentamicin, imipenem, ticarcillin, norfloxacin, polymyxin-B, piperacillin, spectinomycin, tobramycin, and streptomycin
Pumping out (efflux)	cpxA Escherichia coli	aminocoumarin, aminoglycoside	membrane-localized sensor kinase that is activated by envelope stress. It starts a

	O157 H7 str Sakai		kinase cascade that activates CpxR, which promotes efflux complex expression
Pumping out (efflux)	emrB Escherichia coli str K-12 substr MG1655	fluoroquinolone	translocase in the emrB -TolC efflux protein in E. coli. It recognizes substrates including carbonyl cyanide m-chlorophenylhydrazone (CCCP), nalidixic acid, and thioloactomycin.
Pumping out (efflux)	KpnE Klebsiella pneumoniae	peptide, tetracycline, cephalosporin, aminoglcoside, macrolide, rifamycin	KpnE subunit of KpnEF resembles EbrAB from E. coli. Mutation in KpnEF resulted in increased susceptibility to cefepime, ceftriaxon, colistin, erythromycin, rifampin, tetracycline, and streptomycin as well as enhanced sensitivity toward sodium dodecyl sulfate, deoxycholate, dyes, benzalkonium chloride, chlorhexidine, and triclosan
Pumping out (efflux)	oqxA Escherichia coli	tetracycline, nitrofuran, glycylcyline, diaminopyrimidine, fluoroquinolone	RND efflux pump conferring AR to fluoroquinolone
Pumping out (efflux)	emrR Escherichia coli str K-12 substr MG1655	fluoroquinolone	negative regulator for the EmrAB-TolC multidrug efflux pump in E. coli. Mutations lead to EmrAB-TolC overexpression.
Inactivation	LEN-9 Klebsiella variicola	penem (beta-lactam), penam (penicillin)	beta-lactamase
Inactivation	LCR-1 Pseudomonas aeruginosa	Beta-lactam	hydrolyze the active ring of beta-lactam antibiotics
Pumping out (efflux)	acrD Escherichia coli	aminoglycoside	efflux pump which expression induced by indole
Pumping out (efflux)	baeR Escherichia coli str K-12 substr W3110	aminocoumarin, aminoglycoside	response regulator that promotes the expression of MdtABC and AcrD efflux complexes.
Pumping out (efflux)	acrA Klebsiella pneumoniae	tetracycline, phenicol, dlycylcycline, cephalosporin, rifamycin, penam, triclosan, fluoroquinolone	subunit of the AcrAB multidrug efflux system that in K. pneumoniae, which is encoded by the acrRAB operon.
Reduced permeability, pumping out (efflux)	marA Escherichia coli	tetracycline, penam, phenicol, rifamycin, penem, cephamycin, monobactam, cephalosporin, glycylcycline, carbapenem, triclosan, fluoroquinolone	reduced permeability to antibiotic, antibiotic efflux

Inactivation	MOX-5 Aeromonas caviae	cephalosporin, cephamycin, penam	beta-lactamase
Inactivation	APH(3')-Ib Pseudomonas aeruginosa	Aminoglycoside	chromosomal aminoglycoside phosphotransferase gene
Pumping out (efflux)	tolC Escherichia coli	peptide, tetracycline, macrolide, cephalosporin, phenicol, cephamycin, aminoglycoside, penam, carbapenem, fluoroquinolone, rifamycin, glycycline, penem, aminocoumarin, triclosan	multidrug efflux pump protein subunit
Pumping out (efflux)	rsmA Pseudomonas aeruginosa	Phenicol, Diaminopyrimidine, and fluoroquinolone	regulates virulence of Pseudomonas aeruginosa. However, its negative effect on MexEF-OprN overexpression has been noted to confer resistance to various antibiotics
Modification (Abt target alteration)	ugd Escherichia coli	peptide	transfer of 4-amino-4-deoxy-L-arabinose(Ara4N) to lipid A, which converts cationic antibmicrobial peptides and antibiotics
Pumping out (efflux)	acrA Enterobacter cloacae	tetracycline, phenicol, glycycline, rifamycin, penam, cephalosporin, fluoroquinolone, triclosan	risistance-nodulation-cell division (RND) antibiotic efflux pump
Modification (Abt target protection)	QnrS6 Aeromonas hydrophila	fluoroquinolone	plasmid-mediated quinolone resistance protein
Inactivation	APH(6)-Id Pseudomonas aeruginosa	Aminoglycoside	an aminoglycoside phosphotransferase encoded by plasmids, integrative conjugative elements and chromosomal genomic islands in K.
Pumping out (efflux)	yojI Escherichia coli	peptide	resistance to pepide antibiotic, pumping out microcin molecules
Pumping out (efflux)	baeS Escherichia coli	aminocoumarin, aminoglycoside	sensor kinase in the BaeSR regulatory system
Modification (Abt target alteration)	ErmF Bacteroides fragilis	streptogramin, macrolide, lincosamide	confers the Macrolide-lincosamide-streptogramin B (MLSb) phenotype
Pumping out (efflux)	acrB Escherichia coli	tetracycline, rifamycin, glycylcycline, phenocol, cephalosporin, penam, triclosan, fluoroquinolone	multidrug efflux pump- responsible for substrate recognition and energy transduction

Reduced permeability	Klebsiella pneumoniae OmpK37	cephamycin, penem, penam, cephalosporin, carbapenem, monobactam	outermembrane porin protein of Klebsiella pneumoniae
Pumping out (efflux)	qacEdelta1 Pseudomonas aeruginosa	disinfecting agents and intercalating dyes, acridine dye	confer resistance to antiseptics
Inactivation	CMY-59 Escherichia coli	cephamycin	beta-lactamase
Modification (Abt target replacement)	sul1 Vibrio fluvialis	sulfonamide	dihydropteroate synthase of gram-negative bacteria. Linked to class 1 integrons
Modification (Abt target replacement)	sul2 Vibrio cholerae	sulfonamide	sulfonamide resistant dihydropteroate synthase of Gram-negative bacteria, usually found on small plasmids.
Pumping out (efflux)	mdtK Salmonella enterica	fluoroquinolone	multidrug and toxic compound extrusions (MATE) transporter
Reduced permeability, pumping out (efflux)	ramA Enterobacter cloacae	tetracycline, rifamycin, cephamycin, glycycline, cephalosporin, penam, monobactam, penem, carbapenem, phenicol, triclosan, fluoroquinolone	positive regulator of crAB-TolC and leads to high level mutidrug resistance
Pumping out (efflux)	tet(D) Shigella flexneri Y	tetracyline	tetracycline efflux pump exclusive to Gram-negative bacteria
Inactivation	AAC(6_)-If Enterobacter cloacae	aminoglycoside	plasmid-encoded acetyltransferase
Inactivation	arr-3 Escherichia coli	rifamycin	plasmid-encoded ribosyltransferase
Modification (Abt target replacement)	dfrA5 Salmonella enterica	diaminopyrimidine	integron-encoded dihydrofolate reductase
Pumping out (efflux)	kdpE Escherichia coli	aminoglycoside	transcriptional activator, regulates range of virulenxe loci through direct promotor binding
Inactivation	TEM-126 Escherichia coli	penam, cephalosporin, monobactam, penem	extended-spectrum beta lactamase
Modification (Abt target alteration)	bacA Escherichia coli	peptide	recycles undecaprenyl pyrophosphate during cell wall biosynthesis
Inactivation	SHV-100 Klebsiella pneumoniae	cephalosporin, carbapenem, penam	beta-lactamase
Inactivation	CMH-1 Enterobacter cloacae	cephalosporin	class C beta-lactamase

Pumping out (efflux)	YajC Pseudomonas aeruginosa	tetracycline, penam, phenicol, rifamycin, cephalosporin, glycylcycline, fluoroquinolone, triclosan	interacts with AcrAB-TolC efflux pump shows increased fitness in the presence of antibiotics
Inactivation	ACT-35 Enterobacter cloacae	penam, cephamycin, cephalosporin, carbapenem	beta-lactamase

Acronyms: Abt -Antibiotics; AR – Antibiotic Resistance.

4.4.4. Relationship of microbial consortia and antibiotic resistance development

When directly comparing the most abundant bacteria and ARGs detected, there were five distinctive connections on *Klebsiella, Pseudomonas, Enterobacter, Aeromonas*, and *Escherichia* families where most of the top 50 ARGs were connected to species as shown in **Table 4.4**. Additionally, among the top 50 ARGs, there were some genes such as sul1, sul2 from *Vibrio fluvialis* or *chloerae*, and ErmF from *Bacteroides fragilis* were not matched exactly to the genera found from MetaphlAn analysis, but at the higher level of the phylogenic tree, connections of Gammaproteobacteria and Bacteroides were found for sul1 and sul2, and ErmF, respectively. Considering the bacteria's ability to transfer resistance genes within and across their kinds, some of these ARGs could be carried by some other species as tetracycline resistance was found in bacteria such as *E. coli* and *Enterococcus* species (Calva, Sifuentes-Osornio, and Cerón 1996; Domínguez et al. 2002; Francois, Charles, and Courvalin 1997), diaminopyrimidine resistance genes found in *E. coli* species (Lee et al. 2001).

Based on the HUMAnN2 analysis, there were trace amounts of polymyxin resistance pathways found in some samples. This may be due to the introduction of regional environmental water sources used to wash water as FGD bioreactor systems or can be

coming from some community-level bacterial response to the natural synthesis of polymyxins during the bacterial fermentation process (Satlin and Jenkins, 2017). There is potential to study whether FGD bioreactor can remediate such influent of newly emerging contaminants when compared to the effluent, after the microbial treatment activities. To our knowledge, genetic databases available currently are heavily on human-related pathways, and there are myriad genetic segments that are not considered functional from sequencing data. If the current data are revisited after bacterial genetic pathways are more identified and understood more thoroughly past the primary and secondary level of the genome, it might offer new insights on the study where clear pathways of microbial conversion of contaminant processing are found, uncovering some assumptions made at the current stage of technology and development.

Table 4.4. Comparison of bacteria and ARG resistome-baring bacterial families in FGD bioreactors

Bacterial genus-MetaphlAn	**Bacteria Match (Family) - ShortBRED**	**ARGs**
Klebsiella	Klebsiella pneumoniae	ArnT, LptD, etB, KpnF, KpnG, KpnE, LEN-9, acrA, OmpK37, SHV-100
Pseudomonas	Pseudomonas Aeruginosa	LCR-1, APH(3')-Ib, rsmA, APH(6')-Id, qacEdelta1, YajC
Enterobacter	Enterobacter cloacae	ramA, AAC(6')-If, CMH-1, ACT-35
Aeromonas	Aeromonas caviae / hydrophila	MOX-5, QnrS6

Escherichia	Escherichia coli	CRP, oqxA, oqxB, FosA5, cpxA, emrB, emrR, acrD, baeR, marA, tolC, ugd, yojI, baeS, acrB, CMY-59, arr-3, kdpE, TEM-126, bacA,

4.5. Conclusion

While having significant portions of sequenced genes unclassified and unidentified, there are some limitations to finding the exact target metabolism such as other heavy metal uptake and bioaccumulation due to limited findings in the literature. More studies need to be done on various bacterial species with the role of utilizing heavy metals and other contaminants. Certainly, chemical, hydrological, and physical factors of the bioreactors need to be considered for a more comprehensive study on FGD bioreactors. Despite the unknown factors that may reside, this study provided a clear overview of bacterial distribution, classification, and pathway analysis on sulfate, nitrite/nitrate, selenium metabolism, energy generation, and fermentation. Metagenomic analysis of classified consortia of genetic information helped to understand the relationship of seasonal changes and microbiome composition in bioreactors as well as site-specific details. Based on the HUMAnN2 analysis, there were trace amounts of polymyxin resistance pathways found in some samples. This may be due to the introduction of regional environmental water sources used to wash water as FGD bioreactor systems or can be coming from some community-level bacterial response to the natural synthesis of polymyxins during the bacterial fermentation process[41]. There is a potential to study whether the FGD bioreactor can remediate such influence of newly emerging contaminants when compared to the effluent, after the microbial treatment activities. Monitoring the microbiome of two-step bioreactors at three different sites provided multiple symbiotic possibilities of bacterial genera to process contaminants from FGD wastewater. Bacterial distribution showed distinctive patterns depending on site and season when samples were collected, and specific bacterial genus identified from the

bacterial identification was matched to the possible bacterial family which carries ARGs. Such bacterial family were *Klebsiella, Pseudomonas, Enterobacter, Aeromonas, Escherichia, Proteobacteria* and *Bacteroides*.

Furthermore, the long-term effects of the dormant status of the bioreactors at site C were observed was both stages are not receiving influent FGD wastewater feeds anymore. This comparison study of site C to sites A and B can provide essential information to develop a long-term remediation project to further degrade contaminants of interest with the metal-laden sample. The shift in the microbiome demonstrates the ability of bacteria to adapt to their changing environments such as seasonal temperature to fitness and survival. In future analysis, sampling from different depths/regions of bioreactors and metagenomic analysis of both environmental feed water such as lake or river along with sample analysis will be helpful to find loci-driven microbiome distribution along the water cycle at influent, treatment, and discharge stages. This may help discover new knowledge of selenium reducers and other bioprocesses related to heavy metal uptake, and the fate and expression of antibiotic resistome. Having the current knowledge found in the chapter, having further detailed studies will strengthen the understanding of microbial interactions, transfer, and evolution of their traits.

SUMMARY & FINAL REMARKS

In testing the hypothesis laid out, this dissertation study is expected to benefit future studies and fill in the knowledge gap in the following ways:

1) Establishment of ARGs detection method using qPCR and ddPCR and optimization for specific ARGs types for future comparative studies

2) Systematic overview and understanding in ARGs of the water system finding distinctive ARGs expression across various water body types, and how ARGs expression occurs under the anthropogenic influence

3) Addition of knowledge on the microbiome of FGD bioreactors and their functions of contaminant remediation, and ARGs quantities their roles within the bioreactor under receiving of regional water as its source.

The trend of ARGs quantification using ddPCR and qPCR in the literature was observed, and the tested ARGs were showing a similar trend in the range of detection, and limit of quantification between both methods. The ddPCR was more sensitive to measure lower concentration ranges, while qPCR can cover a broader range of quantification at higher concentrations. Careful optimization steps are essential for building robust analysis for ddPCR, and the result obtained from ddPCR can share a more detailed view of reaction conditions, and amplification outcomes.

Studying the ARGs in the U.S. water system showed tetracycline-resistant genes (tetA and tetW) and sul1 seems to be present in the environment and seems to accumulate throughout the water bodies downstream of the water cycle, and showed some increased concentration in underground recharge, and lakes. Effluents from the WWTPs acted as a

source of increased ARGs at the discharge point, and ARGs concentration at the connecting rivers and lake varied depending on the site location. Merged with river showed decreased qnrA concentration of ARG at site 2, while the concentration increased when meeting lake at site 3. Flow measurement data at each sampling location might have helped track the ARGs in water environments because many factors contribute to the expression of ARGs other than the own concentrations. The blaCTX-M was diluted at downstream flow through the water system for all sites. Also, the advanced water treatment facility removed most of the ARGs in the water, which showed close to log 0 units of all ARGs at treatment effluent.

Microbial analysis of FGD bioreactor study showed that VanXYL gene is specific to site A, which indicates the contamination is higher in the river water than lakes. Across sites A and B, the highest ARGs detected was AAC(6')-Ib7, in which aminoglycoside antibiotics contamination can be expected in the regional river and lakes used as the water source. The second-most abundant gene, CRP, is more found in sites B and C, which may have been introduced by the lake water. The majority of ARGs detected in site C was CRP, which may indicate the dormant stage of the bioreactors cause transfer and harboring of these ARGs in the system, along with site-specific ARG, dfrB3, which is resistant to diaminopyridine antibiotics. While there is more to discover on the interactions of the microbiome and FGD wastewater components within the treatment system, the status of bioreactors and feeding water sources shape the microbial consortia in the bioreactors. One can be careful in the design and operation of the FGD wastewater treatment to build a microbial community that is desired and favorable for HM removals, and/or ARGs. More future studies about the detailed analysis of inflow, within the

bioreactor, and outflow microbiome may help better understand the formation and expression of ARGs.

The findings of this study will contribute to improving surveillance and developing mitigation plans for AR and ARGs in the water system globally. Guidelines for having sensitive, optimized ARG detection has been set through the study. By having a comprehensive analysis done on ARGs in the environment especially in the water system, we have more control over ARGs tracking, treatment, and containment protecting the health of humans and the ecosystem. The relationship between ARG and microbial variation found in this study can be utilized to predict the emergence of ARG types in future research. Also, engineering environments for cultivating specific microbial consortia to thrive may help the control of ARG types to minimize pathogens for human health. The optimized method development of ARGs detection using ddPCR can be utilized as a standard method for many future studies to detect ARGs with higher sensitivity and accuracy, particularly in environments where a low concentration of occurrence is predicted.

LIST OF PUBLICATIONS

1. **Park, S.,** Rana, A., Sung, W., and Munir, M. (2021). "Competitiveness of qPCR and ddPCR Technologies, with a Particular Focus on Detection of ARGs". *Applied Microbiology* 1, no. 3: 426-444. https://doi.org/10.3390/applmicrobiol1030028
2. **Park, S.,** Rana, A., Sung, W. and Munir, M. "Comparing qPCR and ddPCR Methods for ARGs analysis from the Environmental Water Sample." [In preparation]
3. **Park, S.,** Roppolo-Brazell, L., Lambirth, K., Handerson, D., Gibas, C., and Munir, M. "Metagenomic Analysis of Flue Gas Desulfurization Bioreactors of Coal-fired Power Stations". [In preparation]
4. **Park, S.,** Brown, X., Keen, O., and Munir, M. "Systematic Analysis of Antibiotic Resistant Genes (ARGs) in U.S. Water System". [In preparation]

REFERENCES

Abou-Kandil, Ammar et al. 2021. "Fate and Removal of Bacteria and Antibiotic Resistance Genes in Horizontal Subsurface Constructed Wetlands: Effect of Mixed Vegetation and Substrate Type." *Science of the Total Environment* 759: 144193. https://doi.org/10.1016/j.scitotenv.2020.144193.

Ahrberg, Christian D., Jong Min Lee, and Bong Geun Chung. 2019. "Microwell Array-Based Digital PCR for Influenza Virus Detection." *BioChip Journal* 13(3): 269–76. http://link.springer.com/10.1007/s13206-019-3302-8.

Alía, Alberto et al. 2020. "Development of a Multiplex Real-Time PCR to Differentiate the Four Major Listeria Monocytogenes Serotypes in Isolates from Meat Processing Plants." *Food Microbiology* 87(November 2019): 103367. https://doi.org/10.1016/j.fm.2019.103367.

Allen, Heather K. et al. 2010. "Call of the Wild: Antibiotic Resistance Genes in Natural Environments." *Nature Reviews Microbiology* 8(4): 251–59.

Arvia, Rosaria et al. 2017. "Droplet Digital PCR (DdPCR) vs Quantitative Real-Time PCR (QPCR) Approach for Detection and Quantification of Merkel Cell Polyomavirus (MCPyV) DNA in Formalin Fixed Paraffin Embedded (FFPE) Cutaneous Biopsies." *Journal of Virological Methods* 246(April): 15–20. http://dx.doi.org/10.1016/j.jviromet.2017.04.003.

Asfaw, Tsegahun, Deribew Genetu, and Demissew Shenkute. 2020. "High Burden of Antibiotic-Resistant Bacteria from Wastewater in Ethiopia: A Systematic Review." *Risk Management and Healthcare Policy* 13: 3003–11.

Baquero, Fernando, José-Luis Martínez, and Rafael Cantón. 2008. "Antibiotics and Antibiotic Resistance in Water Environments." *Current Opinion in Biotechnology* 19(3): 260–65. https://linkinghub.elsevier.com/retrieve/pii/S0958166908000591.

Barczak, Wojciech, Wiktoria Suchorska, Błażej Rubiś, and Katarzyna Kulcenty. 2015. "Universal Real-Time PCR-Based Assay for Lentiviral Titration." *Molecular Biotechnology* 57(2): 195–200.

Basu, Amar S. 2017. "Digital Assays Part I: Partitioning Statistics and Digital PCR." *SLAS Technology* 22(4): 369–86. https://doi.org/10.1177/2472630317705680.

Battley, Edwin H. 1988. "Escherichia Coli and Salmonella Typhimurium. Cellular and Molecular Biology, Volume 1; Volume 2 . Frederick C. Neidhardt , John L. Ingraham , Boris Magasanik , K. Brooks Low , Moselio Schaechter , H. Edwin Umbarger." *The Quarterly Review of Biology* 63(4): 463–64. https://www.journals.uchicago.edu/doi/10.1086/416059.

Baume, Maud et al. 2019. "Quantification of Legionella DNA Certified Reference Material by Digital Droplet PCR." *Journal of Microbiological Methods* 157(December 2018): 50–53. https://doi.org/10.1016/j.mimet.2018.12.019.

Bergeron, Scott et al. 2016. "Presence of Antibiotic Resistance Genes in Different Salinity Gradients of Freshwater to Saltwater Marshes in Southeast Louisiana, USA." *International Biodeterioration and Biodegradation* 113: 80–87. http://dx.doi.org/10.1016/j.ibiod.2016.02.008.

Bhattacharyya, Anish, Anwesha Haldar, Maitree Bhattacharyya, and Abhrajyoti Ghosh. 2019. "Anthropogenic Influence Shapes the Distribution of Antibiotic Resistant Bacteria (ARB) in the Sediment of Sundarban Estuary in India." *Science of the Total*

Environment 647: 1626–39. https://doi.org/10.1016/j.scitotenv.2018.08.038.
Biorad. 2018. "Droplet Digital ™ PCR Droplet Digital ™ PCR Applications Guide." *Biorad*: 145. http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf.
Bustin, Stephen A. et al. 2009. "The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments." *Clinical Chemistry* 55(4): 611–22. https://academic.oup.com/clinchem/article/55/4/611/5631762.
Bustin, Stephen, and Jim Huggett. 2017. "QPCR Primer Design Revisited." *Biomolecular Detection and Quantification* 14(November): 19–28. https://doi.org/10.1016/j.bdq.2017.11.001.
Butler, J. M., Christian M. Ruitberg, and Peter M. Vallone. 2001. "Capillary Electrophoresis as a Tool for Optimization of Multiplex PCR Reactions." *Fresenius' Journal of Analytical Chemistry* 369(3–4): 200–205. http://link.springer.com/10.1007/s002160000641.
Calva, Juan J., José Sifuentes-Osornio, and Cecilia Cerón. 1996. "Antimicrobial Resistance in Fecal Flora: Longitudinal Community-Based Surveillance of Children from Urban Mexico." *Antimicrobial Agents and Chemotherapy* 40(7): 1699–1702.
Cao, Weiwei et al. 2020. "Species Identification and Quantification of Silver Pomfret Using the Droplet Digital PCR Assay." *Food Chemistry* 302(July 2018): 125331. https://doi.org/10.1016/j.foodchem.2019.125331.
de Carvalho, Carla C.C.R., and Manuela M.R. da Fonseca. 2017. "Biotransformations ☆." In *Reference Module in Life Sciences*, Elsevier, 574–85. https://www.sciencedirect.com/science/article/pii/B978012809633809083X (April 5, 2020).
Cavé, Laura et al. 2016. "Efficiency and Sensitivity of the Digital Droplet PCR for the Quantification of Antibiotic Resistance Genes in Soils and Organic Residues." *Applied Microbiology and Biotechnology* 100(24): 10597–608. http://dx.doi.org/10.1007/s00253-016-7950-5.
CDC. 2019. Centers for Disease Control and Prevention *Antibiotic Resistance Threats in the United States*. https://www.cdc.gov/drugresistance/biggest_threats.html.
———. 2020. "From the Federal Task Force on Combating Antibiotic-Resistant Bacteria 2 National Action Plan for Combating Antibiotic-Resistant Bacteria." (October): 2020–25.
Center for Disease Control. 2018. "Antibiotic Use in the United States, 2018: Progress and Opportunities." *US Department of Health and Human Services*: 1–37.
Di Cesare, Andrea et al. 2015. "Constitutive Presence of Antibiotic Resistance Genes within the Bacterial Community of a Large Subalpine Lake." *Molecular Ecology* 24(15): 3888–3900.
Chandler, Darrell P., Christina A. Wagnon, and Harvey Bolton. 1998. "Reverse Transcriptase (RT) Inhibition of PCR at Low Concentrations of Template and Its Implications for Quantitative RT-PCR." *Applied and Environmental Microbiology* 64(2): 669–77.
Chen, Baowei et al. 2013. "Metagenomic Profiles of Antibiotic Resistance Genes (ARGs) between Human Impacted Estuary and Deep Ocean Sediments." *Environmental Science and Technology* 47(22): 12753–60.
Chen, Yu Ru et al. 2019. "Impact of ZnO Nanoparticles on the Antibiotic Resistance

Genes (ARGs) in Estuarine Water: ARG Variations and Their Association with the Microbial Community." *Environmental Science: Nano* 6(8): 2405–19.

Christou, Anastasis et al. 2017. "The Potential Implications of Reclaimed Wastewater Reuse for Irrigation on the Agricultural Environment: The Knowns and Unknowns of the Fate of Antibiotics and Antibiotic Resistant Bacteria and Resistance Genes – A Review." *Water Research* 123: 448–67.

Clewley, J.P. 1989. "The Polymerae Chain Reaction, a Review of the Practical Limitations for Human Immunodeficiency Virus Diagnosis." *Journal of Virological Methods* (25): 179–88. https://www.sciencedirect.com/science/article/abs/pii/0166093489900311.

Czekalski, Nadine, Elena Gascón Díez, and Helmut Bürgmann. 2014. "Wastewater as a Point Source of Antibiotic-Resistance Genes in the Sediment of a Freshwater Lake." *ISME Journal* 8(7): 1381–90.

Demeke, Tigst, Brian Beecher, and Monika Eng. 2020. "Assessment of Genetically Engineered Events in Heat-Treated and Non-Treated Samples Using Droplet Digital PCR and Real-Time Quantitative PCR." *Food Control* 115(December 2019): 107291. https://doi.org/10.1016/j.foodcont.2020.107291.

Deprez, Liesbet et al. 2016. "Validation of a Digital PCR Method for Quantification of DNA Copy Number Concentrations by Using a Certified Reference Material." *Biomolecular Detection and Quantification* 9: 29–39. http://dx.doi.org/10.1016/j.bdq.2016.08.002.

Domínguez, Elena et al. 2002. "Mechanisms of Antibiotic Resistance in Escherichia Coli Isolates Obtained from Healthy Children in Spain." *Microbial Drug Resistance* 8(4): 321–27.

Dong, Lianhua et al. 2020. "Interlaboratory Assessment of Droplet Digital PCR for Quantification of BRAF V600E Mutation Using a Novel DNA Reference Material." *Talanta* 207(August 2019): 120293. https://doi.org/10.1016/j.talanta.2019.120293.

Ekman, Stefan. 1999. "Pcr Optimization and Troubleshooting, with Special Reference to the Amplification of Ribosomal DNA in Lichenized Fungi." *The Lichenologist* 31(5): 517–31. https://www.cambridge.org/core/product/identifier/S0024282999000675/type/journal_article.

Executive Office of the President President's Council of Advisors on Science and Technology. 2014. *REPORT TO THE PRESIDENT ON COMBATING ANTIBIOTIC RESISTANCE*. https://www.cdc.gov/drugresistance/pdf/report-to-the-president-on-combating-antibiotic-resistance.pdf.

Fiorentino, Antonino et al. 2018. "Simulating the Fate of Indigenous Antibiotic Resistant Bacteria in a Mild Slope Wastewater Polluted Stream." *Journal of Environmental Sciences (China)* 69: 95–104. https://doi.org/10.1016/j.jes.2017.04.018.

Floren, C. et al. 2015. "Species Identification and Quantification in Meat and Meat Products Using Droplet Digital PCR (DdPCR)." *Food Chemistry* 173: 1054–58. http://dx.doi.org/10.1016/j.foodchem.2014.10.138.

Francois, B., M. Charles, and P. Courvalin. 1997. "Conjugative Transfer of Tet(S) between Strains of Enterococcus Faecalis Is Associated with the Exchange of Large Fragments of Chromosomal DIMA." *Microbiology* 143(7): 2145–54. https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-

143-7-2145.
Fujimoto, Masanori, Daniel E. Carey, and Patrick J. McNamara. 2018. “Metagenomics Reveal Triclosan-Induced Changes in the Antibiotic Resistome of Anaerobic Digesters.” *Environmental Pollution* 241: 1182–90. https://doi.org/10.1016/j.envpol.2018.06.048.
Fuseler, Knut, Daniel Krekeler, Ulrike Sydow, and Heribert Cypionka. 1996. “A Common Pathway of Sulfide Oxidation by Sulfate-Reducing Bacteria.” *FEMS Microbiology Letters* 144(2–3): 129–34.
Galluzzi, Luca et al. 2018. “Real-Time PCR Applications for Diagnosis of Leishmaniasis.” *Parasites & Vectors* 11(1): 273. https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-018-2859-8.
Gao, Min, Tianlei Qiu, Yanmei Sun, and Xuming Wang. 2018. “The Abundance and Diversity of Antibiotic Resistance Genes in the Atmospheric Environment of Composting Plants.” *Environment International* 116(9): 229–38. https://doi.org/10.1016/j.envint.2018.04.028.
Garibyan, Lilit, and Nidhi Avashia. 2013. “Polymerase Chain Reaction.” *Journal of Investigative Dermatology* 133(3): 1–4. http://dx.doi.org/10.1038/jid.2013.1.
Garner, Emily et al. 2017. “Stormwater Loadings of Antibiotic Resistance Genes in an Urban Stream.” *Water Research* 123: 144–52. http://dx.doi.org/10.1016/j.watres.2017.06.046.
Gaviria-Figueroa, Andrés et al. 2019. “Emission and Dispersal of Antibiotic Resistance Genes through Bioaerosols Generated during the Treatment of Municipal Sewage.” *Science of the Total Environment* 686: 402–12.
Gerdes, Lars, Azuka Iwobi, Ulrich Busch, and Sven Pecoraro. 2016. “Optimization of Digital Droplet Polymerase Chain Reaction for Quantification of Genetically Modified Organisms.” *Biomolecular Detection and Quantification* 7: 9–20. http://dx.doi.org/10.1016/j.bdq.2015.12.003.
Giantsis, I. A., and A. Chaskopoulou. 2019. “Broadening the Tools for Studying Sand Fly Breeding Habitats: A Novel Molecular Approach for the Detection of Phlebotomine Larval DNA in Soil Substrates.” *Acta Tropica* 190(November 2018): 123–28.
Gillings, Michael R et al. 2014. “Using the Class 1 Integron-Integrase Gene as a Proxy for Anthropogenic Pollution.” *The ISME Journal* 9(6): 1269–79. http://dx.doi.org/10.1038/ismej.2014.226.
Gottlieb, Thomas, and Graeme R. Nimmo. 2011. “Antibiotic Resistance Is an Emerging Threat to Public Health: An Urgent Call to Action at the Antimicrobial Resistance Summit 2011.” *Medical Journal of Australia* 194(6): 281–83. https://onlinelibrary.wiley.com/doi/abs/10.5694/j.1326-5377.2011.tb02973.x.
Grabow, W. O.K., O. W. Prozesky, and L. S. Smith. 1974. “Drug Resistant Coliforms Call for Review of Water Quality Standards.” *Water Research* 8(1): 1–9.
Gravel, Annie, Nancy Messier, and Paul H Roy. 1998. “Point Mutations in the Integron Integrase IntI1 That Affect Recombination and / or Substrate Recognition.” 180(20): 5437–42.
Griffiths, Anthony JF et al. 2000. *An Introduction to Genetic Analysis, 7th Edition*. New York: W. H. Freeman.
Guo, Xing pan et al. 2018. “Biofilms as a Sink for Antibiotic Resistance Genes (ARGs) in the Yangtze Estuary.” *Water Research* 129: 277–86.

https://doi.org/10.1016/j.watres.2017.11.029.
Gutiérrez-Aguirre, Ion, Nejc Rački, Tanja Dreo, and Maja Ravnikar. 2015. "Droplet Digital PCR for Absolute Quantification of Pathogens." In *Methods Mol Biol.*, Methods in Molecular Biology, ed. Christophe Lacomme. New York, NY: Springer New York, 331–47. http://link.springer.com/10.1007/978-1-4939-2620-6.
Henrich, Timothy J. et al. 2012. "Low-Level Detection and Quantitation of Cellular HIV-1 DNA and 2-LTR Circles Using Droplet Digital PCR." *Journal of Virological Methods* 186(1–2): 68–72.
Heredia, Nicholas J. et al. 2013. "Droplet Digital™ PCR Quantitation of HER2 Expression in FFPE Breast Cancer Samples." *Methods* 59(1): S20–23. https://linkinghub.elsevier.com/retrieve/pii/S1046202312002526.
Heyries, Kevin A. et al. 2011. "Megapixel Digital PCR." *Nature Methods* 8(8): 649–51.
Hill, R., S. Keely, N. Brinkman, E. Wheaton, S. Leibowitz, M. Jahne, R. Martin, AND J. Garland. 2018. "The Prevalence of Antibiotic Resistance Genes in US Waterways and Their Relationship to Water Quality and Land Use Indicators." In Detroit, MI: Annual Meeting of the Society for Freshwater Science.
Hindson, Benjamin J. et al. 2011. "High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number." *Analytical Chemistry* 83(22): 8604–10.
Van Hoek, Angela H.A.M. et al. 2011. "Acquired Antibiotic Resistance Genes: An Overview." *Frontiers in Microbiology* 2(SEP): 1–27.
Huang, Yu-Hong et al. 2019. "Occurrence and Distribution of Antibiotics and Antibiotic Resistant Genes in Water and Sediments of Urban Rivers with Black-Odor Water in Guangzhou, South China." *Science of The Total Environment* 670: 170–80. https://doi.org/10.1016/j.scitotenv.2019.03.168.
Huggett, Jim F. et al. 2008. "Differential Susceptibility of PCR Reactions to Inhibitors: An Important and Unrecognised Phenomenon." *BMC Research Notes* 1(1): 70. http://bmcresnotes.biomedcentral.com/articles/10.1186/1756-0500-1-70.
Huggett, Jim F., Simon Cowen, and Carole A. Foy. 2015. "Considerations for Digital PCR as an Accurate Molecular Diagnostic Tool." *Clinical Chemistry* 61(1): 79–88.
Van Hullebusch, Eric D. 2017. Bioremediation of Selenium Contaminated Wastewater *Bioremediation of Selenium Contaminated Wastewater*. ed. Eric D van Hullebusch. Cham: Springer International Publishing. http://link.springer.com/10.1007/978-3-319-57831-6.
Hulme, John. 2017. "Recent Advances in the Detection of Methicillin Resistant Staphylococcus Aureus (MRSA)." *Biochip Journal* 11(2): 89–100.
Ibekwe, A. Mark et al. 2020. "Comparative Use of Quantitative Pcr (Qpcr), Droplet Digital Pcr (Ddpcr), and Recombinase Polymerase Amplification (Rpa) in the Detection of Shiga Toxin-Producing e. Coli (Stec) in Environmental Samples." *Water (Switzerland)* 12(12): 1–15.
Jones, Mathew et al. 2014. "Low Copy Target Detection by Droplet Digital PCR through Application of a Novel Open Access Bioinformatic Pipeline, 'Definetherain.'" *Journal of Virological Methods* 202: 46–53.
Karkman, Antti, Katariina Pärnänen, and D G Joakim Larsson. 2019. "Fecal Pollution Can Explain Antibiotic Resistance Gene Abundances in Anthropogenically Impacted Environments." *Nature Communications* 10(1): 80.

http://dx.doi.org/10.1038/s41467-018-07992-3.
Karlen, Yann et al. 2007. “Statistical Significance of Quantitative PCR.” *BMC Bioinformatics* 8: 1–17.
Kim, Jungmin, Hee Young Kang, and Yeonhee Lee. 2008. “The Identification of CTX-M-14, TEM-52, and CMY-1 Enzymes in Escherichia Coli Isolated from the Han River in Korea.” *The Journal of Microbiology* 46(5): 478–81. http://link.springer.com/10.1007/s12275-008-0150-y.
Kline, Margaret C., Erica L. Romsos, and David L. Duewer. 2016. “Evaluating Digital PCR for the Quantification of Human Genomic DNA: Accessible Amplifiable Targets.” *Analytical Chemistry* 88(4): 2132–39. http://link.springer.com/10.1007/s00216-020-02733-2.
Knapp, C W, J Dolfing, P A Ehlert, and D W Graham. 2010. “Evidence of Increasing Antibiotic Resistance Gene Abundance in Archived Soils since 1940.” *Environ Sci Technol* 44(2): 580–87. http://dx.doi.org/10.1021/es901221x.
Koch, Hanna, Alina Jeschke, and Lutz Becks. 2016. “Use of DdPCR in Experimental Evolution Studies.” *Methods in Ecology and Evolution* 7(3): 340–51.
Koepfli, Cristian et al. 2016. “Sensitive and Accurate Quantification of Human Malaria Parasites Using Droplet Digital PCR (DdPCR).” *Scientific Reports* 6(1): 39183. http://www.nature.com/articles/srep39183.
Kyae, Khin et al. 2019. “Level of Seven Neuroblastoma-Associated MRNAs Detected by Droplet Digital PCR Is Associated with Tumor Relapse/Regrowth of High-Risk Neuroblastoma Patients.” *The Journal of Molecular Diagnostics*. https://doi.org/10.1016/j.jmoldx.2019.10.012.
Lederberg, Joshua, and E. L. Tatum. 1946. “Gene Recombination in Escherichia Coli.” *Nature* 158(4016): 558–558. http://www.nature.com/articles/158558a0.
Lee, Je Chul et al. 2001. “The Prevalence of Trimethoprim-Resistance-Conferring Dihydrofolate Reductase Genes in Urinary Isolates of Escherichia Coli in Korea.” *Journal of Antimicrobial Chemotherapy* 47(5): 599–604.
Lhermie, Guillaume et al. 2019. “Tradeoffs between Resistance to Antimicrobials in Public Health and Their Use in Agriculture: Moving towards Sustainability Assessment.” *Ecological Economics* 166(July): 106427. https://doi.org/10.1016/j.ecolecon.2019.106427.
Li, Jianan et al. 2015. “Antibiotic-Resistant Genes and Antibiotic-Resistant Bacteria in the Effluent of Urban Residential Areas, Hospitals, and a Municipal Wastewater Treatment Plant System.” *Environmental Science and Pollution Research* 22(6): 4587–96.
Li, Li Guan, Qi Huang, Xiaole Yin, and Tong Zhang. 2020. “Source Tracking of Antibiotic Resistance Genes in the Environment — Challenges, Progress, and Prospects.” *Water Research* 185: 116127. https://doi.org/10.1016/j.watres.2020.116127.
Liao, Yingyin et al. 2019. “Diagnostic Test Accuracy of Droplet Digital PCR for the Detection of EGFR Mutation (T790M) in Plasma: Systematic Review and Meta-Analysis.” *Clinica Chimica Acta* (September): 0–1. https://doi.org/10.1016/j.cca.2019.11.023.
Lin, Qiang et al. 2020. “Development and Application of a Sensitive Droplet Digital PCR (DdPCR) for the Detection of Infectious Spleen and Kidney Necrosis Virus.”

Aquaculture 529(July): 735697. https://doi.org/10.1016/j.aquaculture.2020.735697.
Link-Lenczowska, Dorota et al. 2018. "A Comparison of QPCR and DdPCR Used for Quantification of the JAK2 V617F Allele Burden in Ph Negative MPNs." *Annals of Hematology* 97(12): 2299–2308. http://link.springer.com/10.1007/s00277-018-3451-1.
Liu, Jinxin et al. 2019. "Dairy Farm Soil Presents Distinct Microbiota and Varied Prevalence of Antibiotic Resistance across Housing Areas." *Environmental Pollution* 254: 113058. https://doi.org/10.1016/j.envpol.2019.113058.
Liu, Qian et al. 2020. "Parental Somatic Mosaicism for CNV Deletions – A Need for More Sensitive and Precise Detection Methods in Clinical Diagnostics Settings." *Genomics* 112(5): 2937–41. https://doi.org/10.1016/j.ygeno.2020.05.003.
Ma, Shuangchen et al. 2019. "Advanced Treatment Technology for Fgd Wastewater in Coal-Fired Power Plantcurrent Situation and Future Prospects." *Desalination and Water Treatment* 167: 122–32.
Maeda, Ryotaro et al. 2020. "High Throughput Single Cell Analysis of Mitochondrial Heteroplasmy in Mitochondrial Diseases." *Scientific Reports* 10(1): 1–10. https://doi.org/10.1038/s41598-020-67686-z.
Malic, Lidija et al. 2019. "Epigenetic Subtyping of White Blood Cells Using a Thermoplastic Elastomer-Based Microfluidic Emulsification Device for Multiplexed, Methylation-Specific Digital Droplet PCR." *Analyst* 144(22): 6541–53.
Mao, Daqing et al. 2015. "Prevalence and Proliferation of Antibiotic Resistance Genes in Two Municipal Wastewater Treatment Plants." *Water Research* 85: 458–66. http://dx.doi.org/10.1016/j.watres.2015.09.010.
Marangi, Marianna et al. 2015. "Multiplex PCR for the Detection and Quantification of Zoonotic Taxa of Giardia, Cryptosporidium and Toxoplasma in Wastewater and Mussels." *Molecular and Cellular Probes* 29(2): 122–25. http://dx.doi.org/10.1016/j.mcp.2015.01.001.
Martini, Diane R., and Sidney K. Vadbunker. 2016. "Biological Treatment for FGD Wastewater: A Few Options." *Power Engineering*: Issue12 Vol.120. https://www.power-eng.com/2016/12/22/biological-treatment-for-fgd-wastewater-a-few-options/#gref.
Mcfarland, Allen. 2019. "Fossil Fuels Continue to Account for the Largest Share of U.S. Energy." *USEIA*. https://www.eia.gov/todayinenergy/detail.php?id=41353 (June 4, 2020).
Medeiros, Julliane Dutra et al. 2016. "Comparative Metagenome of a Stream Impacted by the Urbanization Phenomenon." *Brazilian Journal of Microbiology* 47(4): 835–45. https://linkinghub.elsevier.com/retrieve/pii/S1517838216305391.
"MetaCyc Pathway: Superpathway of L-Methionine Biosynthesis (by Sulfhydrylation)." *BioCyc database*. https://biocyc.org/META/NEW-IMAGE?object=PWY-5345&&redirect=T.
Milbury, Coren A. et al. 2014. "Determining Lower Limits of Detection of Digital PCR Assays for Cancer-Related Gene Mutations." *Biomolecular Detection and Quantification* 1(1): 8–22.
Murphy-belcaster, Megan, Katherine R Murphy, Egon A Ozer, and Alan R Hauser. 2021. "Infections at a United States Academic Hospital."
Naaum, Amanda M. et al. 2018. "Complementary Molecular Methods Detect Undeclared

Species in Sausage Products at Retail Markets in Canada." *Food Control* 84: 339–44. https://doi.org/10.1016/j.foodcont.2017.07.040.

Noh, Eun Soo et al. 2019. "Quantitative Analysis of Alaska Pollock in Seafood Products by Droplet Digital PCR." *Food Chemistry* 275(August 2017): 638–43. https://doi.org/10.1016/j.foodchem.2018.09.093.

Nshimyimana, Jean Pierre, Mercedes C. Cruz, Stefan Wuertz, and Janelle R. Thompson. 2019. "Variably Improved Microbial Source Tracking with Digital Droplet PCR." *Water Research* 159: 192–202. https://doi.org/10.1016/j.watres.2019.04.056.

Nyaruaba, Raphael, Caroline Mwaliko, Kelvin Kimutai Kering, and Hongping Wei. 2019. "Droplet Digital PCR Applications in the Tuberculosis World." *Tuberculosis* 117(April): 85–92. https://doi.org/10.1016/j.tube.2019.07.001.

Parkinson, John S. 2016. "Classic Spotlight : Discovery of Bacterial Transduction." *Journal of Bacteriology* 198(21): 2899–2900.

Peak, Nicholas et al. 2007. "Abundance of Six Tetracycline Resistance Genes in Wastewater Lagoons at Cattle Feedlots with Different Antibiotic Use Strategies." 9: 143–51.

Perry, David J. et al. 1989. "Antithrombin Cambridge, 384 Ala to Pro: A New Variant Identified Using the Polymerase Chain Reaction." *FEBS Letters* 254(1–2): 174–76. http://doi.wiley.com/10.1016/0014-5793%2889%2981033-6.

Persson, Sofia et al. 2018. "Comparison between RT Droplet Digital PCR and RT Real-Time PCR for Quantification of Noroviruses in Oysters." *International Journal of Food Microbiology* 284(July): 73–83.

Petiti, Jessica et al. 2020. "Novel Multiplex Droplet Digital PCR Assays to Monitor Minimal Residual Disease in Chronic Myeloid Leukemia Patients Showing Atypical BCR-ABL1 Transcripts." *Journal of Clinical Medicine* 9(5): 1457.

Pinheiro-de-Oliveira, Tatiana F. et al. 2018a. "Development of a Droplet Digital RT-PCR for the Quantification of Foot-and-Mouth Virus RNA." *Journal of Virological Methods* 259(June): 129–34. https://doi.org/10.1016/j.jviromet.2018.06.015.

———. 2018b. "Development of a Droplet Digital RT-PCR for the Quantification of Foot-and-Mouth Virus RNA." *Journal of Virological Methods* 259(August 2017): 129–34. https://doi.org/10.1016/j.jviromet.2018.06.015.

Pinheiro, Leonardo B. et al. 2012. "Evaluation of a Droplet Digital Polymerase Chain Reaction Format for DNA Copy Number Quantification." *Analytical Chemistry* 84(2): 1003–11. https://pubs.acs.org/doi/10.1021/ac202578x.

Ponchel, Frederique et al. 2003. "Real-Time PCR Based on SYBR-Green I Fluorescence: An Alternative to the TaqMan Assay for a Relative Quantification of Gene Rearrangements, Gene Amplifications and Micro Gene Deletions." *BMC Biotechnology* 3: 1–13.

Porcellato, Davide, Judith Narvhus, and Siv Borghild Skeie. 2016. "Detection and Quantification of Bacillus Cereus Group in Milk by Droplet Digital PCR." *Journal of Microbiological Methods* 127: 1–6. http://dx.doi.org/10.1016/j.mimet.2016.05.012.

Posada-Perlaza, Carlos Eduardo et al. 2019. "Bogotá River Anthropogenic Contamination Alters Microbial Communities and Promotes Spread of Antibiotic Resistance Genes." *Scientific Reports* 9(1): 11764. http://www.nature.com/articles/s41598-019-48200-6.

Preston, Christina M. et al. 2011. “Underwater Application of Quantitative PCR on an Ccean Mooring.” *PLoS ONE* 6(8).

Pruden, Amy et al. 2013. “Management Options for Reducing the Release of Antibiotics and Antibiotic Resistance Genes to the Environment.” 878(8): 878–86.

Pruden, Amy, Mazdak Arabi, and Heather N. Storteboom. 2012a. “Correlation Between Upstream Human Activities and Riverine Antibiotic Resistance Genes.” *Environmental Science & Technology* 46(21): 11541–49. https://pubs.acs.org/doi/10.1021/es302657r.

Pruden, Amy, Mazdak Arabi, and Heather N Storteboom. 2012b. “Correlation Between Upstream Human Activities and Riverine Antibiotic Resistance Genes.” *Environmental Science & Technology* 46(21): 11541–49. https://pubs.acs.org/doi/10.1021/es302657r.

Pruden, Amy, Ruoting Pei, Heather Storteboom, and Kenneth H. Carlson. 2006. “Antibiotic Resistance Genes as Emerging Contaminants: Studies in Northern Colorado †.” *Environmental Science & Technology* 40(23): 7445–50. https://pubs.acs.org/doi/10.1021/es060413l.

Quan, Phenix Lan, Martin Sauzade, and Eric Brouzes. 2018. “DPCR: A Technology Review.” *Sensors (Switzerland)* 18(4).

Rački, Nejc, Tanja Dreo, et al. 2014. “Reverse Transcriptase Droplet Digital PCR Shows High Resilience to PCR Inhibitors from Plant, Soil and Water Samples.” *Plant Methods* 10(1): 1–10.

Rački, Nejc, Dany Morisset, Ion Gutierrez-Aguirre, and Maja Ravnikar. 2014. “One-Step RT-Droplet Digital PCR: A Breakthrough in the Quantification of Waterborne RNA Viruses.” *Analytical and Bioanalytical Chemistry* 406(3): 661–67.

Ram, Ravishankar M. et al. 2019. “Polymorphisms in the Host CYP2C19 Gene and Antibiotic-Resistance Attributes of Helicobacter Pylori Isolates Influence the Outcome of Triple Therapy.” *Journal of Antimicrobial Chemotherapy* 74(1): 11–16.

Ramírez, Juan David et al. 2019. “Development of a Digital Droplet Polymerase Chain Reaction (DdPCR) Assay to Detect Leishmania DNA in Samples from Cutaneous Leishmaniasis Patients.” *International Journal of Infectious Diseases* 79: 1–3. https://doi.org/10.1016/j.ijid.2018.10.029.

Raurich, Sergi et al. 2019. “Optimisation of a Droplet Digital PCR for Strain Specific Quantification of a Probiotic Bifidobacterium Animalis Strain in Poultry Feed.” *Journal of Microbiological Methods* 163(March).

Reddy, Bhaskar, and Suresh Kumar Dubey. 2019. “River Ganges Water as Reservoir of Microbes with Antibiotic and Metal Ion Resistance Genes: High Throughput Metagenomic Approach.” *Environmental Pollution* 246: 443–51. https://doi.org/10.1016/j.envpol.2018.12.022.

Rizzo, L. et al. 2013. “Urban Wastewater Treatment Plants as Hotspots for Antibiotic Resistant Bacteria and Genes Spread into the Environment: A Review.” *Science of The Total Environment* 447: 345–60. http://dx.doi.org/10.1016/j.scitotenv.2013.01.032.

Romsos, Erica L., and Peter M. Vallone. 2019. “Estimation of Extraction Efficiency by Droplet Digital PCR.” *Forensic Science International: Genetics Supplement Series* 7(1): 515–17.

Ruan, Ting et al. 2015. “Methodology for Studying Biotransformation of Polyfluoroalkyl

Precursors in the Environment." *TrAC - Trends in Analytical Chemistry* 67: 167–78. http://dx.doi.org/10.1016/j.trac.2014.11.017.

Sanitation Safety Planning, Greywater and Excreta. 2016. World Health Organization Geneva, Switzerland. http://apps.who.int/iris/bitstream/handle/10665/171753/9789241549240_eng.pdf?sequence=1.

Scollo, Francesco et al. 2016. "Absolute Quantification of Olive Oil DNA by Droplet Digital-PCR (DdPCR): Comparison of Isolation and Amplification Methodologies." *Food Chemistry* 213: 388–94. http://dx.doi.org/10.1016/j.foodchem.2016.06.086.

Scott, Anna Mae et al. 2018. "Is Antimicrobial Administration to Food Animals a Direct Threat to Human Health? A Rapid Systematic Review." *International Journal of Antimicrobial Agents* 52(3): 316–23. https://doi.org/10.1016/j.ijantimicag.2018.04.005.

Shen, Yike et al. 2019. "Pharmaceutical Exposure Changed Antibiotic Resistance Genes and Bacterial Communities in Soil-Surface- and Overhead-Irrigated Greenhouse Lettuce." *Environment International* 131(June): 105031. https://doi.org/10.1016/j.envint.2019.105031.

Singh, Prakash K. 2010. *Recent Trends in Microbial Biotechnology: (Chapter) Desulfurization and Demineralization of Coal with Bacteria: An Ecofriendly Concept for Clean Coal Energy*. ed. Vijai Kumar Bajpai, Vivek; Sharma, Pallavi, Gupta.

Sivaganesan, Mano, Manju Varma, Shawn Siefring, and Richard Haugland. 2018. "Quantification of Plasmid DNA Standards for U.S. EPA Fecal Indicator Bacteria QPCR Methods by Droplet Digital PCR Analysis." *Journal of Microbiological Methods* 152(July): 135–42. https://doi.org/10.1016/j.mimet.2018.07.005.

Strain, Matthew C. et al. 2013. "Highly Precise Measurement of HIV DNA by Droplet Digital PCR" ed. Yuntao Wu. *PLoS ONE* 8(4): e55943. https://dx.plos.org/10.1371/journal.pone.0055943.

Suslov, Oleg. 2005. "PCR Inhibition by Reverse Transcriptase Leads to an Overestimation of Amplification Efficiency." *Nucleic Acids Research* 33(20): e181–e181. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gni176.

Szczepanowski, Rafael et al. 2009. "Detection of 140 Clinically Relevant Antibiotic-Resistance Genes in the Plasmid Metagenome of Wastewater Treatment Plant Bacteria Showing Reduced Susceptibility to Selected Antibiotics." *Microbiology* 155(7): 2306–19.

Tacconelli, Evelina et al. 2018. "Discovery, Research, and Development of New Antibiotics: The WHO Priority List of Antibiotic-Resistant Bacteria and Tuberculosis." *The Lancet Infectious Diseases* 18(3): 318–27. https://linkinghub.elsevier.com/retrieve/pii/S1473309917307533.

Takahashi, Mayumi et al. 2016. "High Throughput Sequencing Analysis of RNA Libraries Reveals the Influences of Initial Library and PCR Methods on SELEX Efficiency." *Scientific Reports* 6(September): 1–14. http://dx.doi.org/10.1038/srep33697.

Talarico, Sarah et al. 2018. "High Prevalence of Helicobacter Pylori Clarithromycin Resistance Mutations among Seattle Patients Measured by Droplet Digital PCR." *Helicobacter* 23(2): 2–9.

Tang, Xingyao et al. 2021. “The Changes in Antibiotic Resistance Genes during 86 Years of the Soil Ripening Process without Anthropogenic Activities.” *Chemosphere* 266: 128985. https://doi.org/10.1016/j.chemosphere.2020.128985.

Taylor, Sean C., Julie Carbonneau, Dawne N. Shelton, and Guy Boivin. 2015. “Optimization of Droplet Digital PCR from RNA and DNA Extracts with Direct Comparison to RT-QPCR: Clinical Implications for Quantification of Oseltamivir-Resistant Subpopulations.” *Journal of Virological Methods* 224: 58–66. http://dx.doi.org/10.1016/j.jviromet.2015.08.014.

Taylor, Sean C., Genevieve Laperriere, and Hugo Germain. 2017. “Droplet Digital PCR versus QPCR for Gene Expression Analysis with Low Abundant Targets: From Variable Nonsense to Publication Quality Data.” *Scientific Reports* 7(1): 1–8. http://dx.doi.org/10.1038/s41598-017-02217-x.

Thermo-Fischer. 2016. “Mechanism of Antibiotic Resistance.” https://www.thermofisher.com/blog/behindthebench/wp-content/uploads/sites/9/2016/04/Mechanism-of-antibiotic-resistance-1.jpg.

Tone, M, and K Torunn. 2016. “Presence and Levels of Antibiotic Resistance Genes in Saliva from Dental Students in Tromsø.” (May).

Tong, Lei et al. 2017. “Distribution of Antibiotics in Alluvial Sediment near Animal Breeding Areas at the Jianghan Plain, Central China.” *Chemosphere* 186: 100–107. http://dx.doi.org/10.1016/j.chemosphere.2017.07.141.

U.S. EPA. 1985. “Contaminants of Emerging Concern Including Pharmaceuticals and Personal Care Products.” https://www.epa.gov/wqc/contaminants-emerging-concern-including-pharmaceuticals-and-personal-care-products.

Überbacher, Christa et al. 2019. “Application of CRISPR/Cas9 Editing and Digital Droplet PCR in Human IPSCs to Generate Novel Knock-in Reporter Lines to Visualize Dopaminergic Neurons.” *Stem Cell Research* 41(November): 101656. https://doi.org/10.1016/j.scr.2019.101656.

USFDA. 2018. *2017 Summary Report On Antimicrobials Sold or Distributed for Use in Food-Producing Animals*.

Vishnuraj, M.R. et al. 2021. “Detection of Giblets in Chicken Meat Products Using MicroRNA Markers and Droplet Digital PCR Assay.” *LWT* 140(September 2020): 110798. https://doi.org/10.1016/j.lwt.2020.110798.

Wang, Jiali et al. 2016. “Research on the Adsorption and Migration of Sulfa Antibiotics in Underground Environment.” *Environmental Earth Sciences* 75(18): 1–9.

Wang, Xiaofu et al. 2019. “Detection of Transgenic Rice Line TT51-1 in Processed Foods Using Conventional PCR, Real-Time PCR, and Droplet Digital PCR.” *Food Control* 98(September 2018): 380–88. https://doi.org/10.1016/j.foodcont.2018.11.032.

Wang, Xiaomin et al. 2021. “Heavy Metal Could Drive Co-Selection of Antibiotic Resistance in Terrestrial Subsurface Soils.” *Journal of Hazardous Materials* 411(December 2020): 124848. https://doi.org/10.1016/j.jhazmat.2020.124848.

WHO. 2014. Bulletin of the World Health Organization *Antimicrobial Resistance: Global Report on Surveillance*.

———. 2019. UN Ad hoc Interagency Coordinating Group on Antimicrobial Resistance *No Time to Wait: Securing the Future from Drug-Resistant Infections*. https://www.who.int/antimicrobial-resistance/interagency-coordination-

group/IACG_final_report_EN.pdf?ua=1.
Wong, Wing Hing et al. 2014. "'Direct PCR' Optimization Yields a Rapid, Cost-Effective, Nondestructive and Efficient Method for Obtaining DNA Barcodes without DNA Extraction." *Molecular Ecology Resources* 14(6): 1271–80. http://doi.wiley.com/10.1111/1755-0998.12275.
Yang, Qi et al. 2017. "The Development of a Sensitive Droplet Digital PCR for Quantitative Detection of Porcine Reproductive and Respiratory Syndrome Virus." *International Journal of Biological Macromolecules* 104: 1223–28. http://dx.doi.org/10.1016/j.ijbiomac.2017.06.115.
Yang, Rongchang, Andrea Paparini, Paul Monis, and Una Ryan. 2014. "Comparison of Next-Generation Droplet Digital PCR (DdPCR) with Quantitative PCR (QPCR) for Enumeration of Cryptosporidium Oocysts in Faecal Samples." *International Journal for Parasitology* 44(14): 1105–13. http://dx.doi.org/10.1016/j.ijpara.2014.08.004.
Young, Suzanne, Andrew Juhl, and Gregory D. O'Mullan. 2013. "Antibiotic-Resistant Bacteria in the Hudson River Estuary Linked to Wet Weather Sewage Contamination." *Journal of Water and Health* 11(2): 297–310.
Zainab, Syeda Maria, Muhammad Junaid, Nan Xu, and Riffat Naseem Malik. 2020. "Antibiotics and Antibiotic Resistant Genes (ARGs) in Groundwater: A Global Review on Dissemination, Sources, Interactions, Environmental and Human Health Risks." *Water Research* 187: 116455. https://doi.org/10.1016/j.watres.2020.116455.
Zhang, Wenlong et al. 2019. "Removal of Heavy Metals by Aged Zero-Valent Iron from Flue-Gas-Desulfurization Brine under High Salt and Temperature Conditions." *Journal of Hazardous Materials* 373(March): 572–79. https://doi.org/10.1016/j.jhazmat.2019.03.117.
Zhang, Yongning et al. 2019. "Development of a Droplet Digital PCR Assay for Sensitive Detection of Porcine Circovirus 3." *Molecular and Cellular Probes* 43(November 2018): 50–57. https://doi.org/10.1016/j.mcp.2018.11.005.
Zhao, Yanchun. 2019. "A Comparative Study of DdPCR and Sanger Sequencing for Quantitative Detection of Low-Frequency Mutation Rate." *IOP Conference Series: Earth and Environmental Science* 332(3): 0–10.
ZHONG, Xi et al. 2018. "Development of a Sensitive and Reliable Droplet Digital PCR Assay for the Detection of 'Candidatus Liberibacter Asiaticus.'" *Journal of Integrative Agriculture* 17(2): 483–87. http://dx.doi.org/10.1016/S2095-3119(17)61815-X.
Zhu, Yijing et al. 2018. "Robust Performance of a Membrane Bioreactor for Removing Antibiotic Resistance Genes Exposed to Antibiotics: Role of Membrane Foulants." *Water Research* 130: 139–50. https://doi.org/10.1016/j.watres.2017.11.067.