

Shotgun metagenomic analysis of antimicrobial resistance in wastewater

Ayella Zorka Maile-Moskowitz

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Civil Engineering

Peter J. Vikesland, Co-Chair
Amy Pruden, Co-Chair
Brian D. Badgley
Helmut Bürgmann

February 3rd, 2023
Blacksburg, VA

Keywords: wastewater, antibiotic resistance, SARS-CoV-2, next-generation sequencing,
wastewater based surveillance

Copyright 2023

Shotgun metagenomic analysis of antimicrobial resistance in wastewater

Ayella Zorka Maile-Moskowitz

ABSTRACT

Antimicrobial resistance (AMR) threatens our modern standard of living with the potential return to a pre-antibiotic condition where deadly infections are no longer treatable. Wastewater treatment plants (WWTPs) are vital components in water sanitation infrastructure and are now also being recognized as valuable monitoring points for antibiotics, antibiotic resistant bacteria (ARB), and antibiotic resistance genes (ARGs) disposed of or excreted by human populations. Hospital waste water is of special interest as a potential focused monitoring point and in general research is needed to establish the benefits of both on-site and community-scale wastewater treatment as important barriers to the disseminators of ARGs into the environment. The research aims described herein examine these components of wastewater treatment and how they relate to AMR indicators identified through metagenomic sequencing. Through monitoring of local WWTPs, it was found that AMR indicators shifted over time and in relation to human behavior that changed due to the COVID-19 pandemic. Hospital wastewater did not measurably impact the microbiome during simulated activated sludge wastewater treatment according to broad-scale metagenomic ARG profiling; however, some clinically-relevant ARGs escaped treatment. Lastly, a study of a transect of WWTPs indicated impacts on the abundance of certain ARGs in downstream riverine receiving environments. Nonetheless, there appeared to be a number of other factors at play, and upstream and downstream resistomes tended to remain similar, calling for further research to delineate impacts of various wastewaters and treatments on ARGs in affected aquatic environments.

Shotgun metagenomic analysis of antimicrobial resistance in wastewater

Ayella Zorka Maile-Moskowitz

GENERAL AUDIENCE ABSTRACT

Antimicrobial resistance (AMR) occurs when bacteria, viruses, and fungi are able to survive in the presence of antibiotics because they carry antibiotic resistance genes (ARGs) encoded in their DNA. AMR is a major public health concern as it makes it so that antibiotics are no longer effective against potentially deadly infections. Wastewater treatment plants (WWTPs) are being discovered as a hub of opportunity for monitoring potential AMR problems in a community. WWTPs receive sewage from homes and various industries. This sewage contains rich information for researchers to examine in terms of which antibiotics, bacteria, and ARGs are circulating in the community. This makes it possible to find out which antibiotics are being consumed in the community and which ARGs might be prevalent. The purpose of this research was to better understand both how WWTPs can be used as monitoring points for AMR and how they can be improved to help reduce ARGs emitted to rivers and streams where treated water is discharged. It was found that the types of ARGs prevalent in wastewater changed over time, especially during the COVID-19 pandemic as people worked from home and changed habits regarding doctors' visits, which impacted antibiotic use. Hospital sewage was studied as a useful indicator of pathogens and ARGs that are challenging a community and also the antibiotics being used. This research explored what happened to ARGs during the treatment of domestic (i.e., from people's homes) wastewater along with hospital wastewater and found that hospital wastewater introduced some ARGs that are typically found in clinical settings, but did not negatively impact the overall wastewater treatment process. Finally, the impact that WWTPs have on rivers to which treated water is discharged was explored. The results indicated that certain ARGs were elevated downstream of the WWTPs. However, when examining all ARGs together, no major shifts due to the treated wastewater were apparent.

DEDICATION

I would like to dedicate this dissertation to my family, especially those whom I was able to tell about starting graduate school, but who are no longer here to tell that I've finished. To my Grandma Mary Ann, who held our family together and always encouraged me to focus on my education. To my Grandpa Jerome, Tante Maria, Dodo, Kate, and Uncle George. And in everything that I do, for you Dad.

ACKNOWLEDGEMENTS

Thank you to my advisors Dr. Peter Vikesland and Dr. Amy Pruden for challenging me and believing that I am capable and deserving to be in this space. Thank you for listening to me when I needed help and providing the assistance and encouragement to succeed. Thank you to Dr. Helmut Bürgmann for the opportunity to work in a lab with the most beautiful view and, thank you to both Dr. Bürgmann and Dr. Brian Badgley for your support and feedback over the years.

Thank you to the current and former Pruden and Vikesland group members for teaching me so much over the years, for helping process samples during a pandemic, and for always being ready to answer my questions. Thank you for eating my baked goods, exploring restaurants with me, and for being my workout buddies and teammates. I cherish your friendships and all of the memories we created together.

To my original UMD poop lab crew, thank you for inspiring me to pursue a career in science/engineering and for your continued help as I followed in your footsteps. I'm grateful for all of your advice, the good food that we have shared, and most importantly all the times you've made me laugh.

Finally, I would like to thank my family. Thank you especially to my Mom, Ansel, and Aimee for being my home away from home, for understanding my crazy schedule and corresponding exhaustion, and for your curiosity; please never stop asking me poop/wastewater/fart related questions. And an extra special thank you for sending me snacks, they helped get these words onto this page. Guckamol Oma, i hab's gschafft!

ATTRIBUTIONS

This research was a collaborative effort across disciplines; the contributions the coauthor made to this dissertation are described below:

Dr. Peter Vikesland (Civil and Environmental Engineering, Virginia Tech). PhD co-advisor, committee co-chair and project co-PI. Co-author of chapters 2, 3, and 4. Contributed guidance on experimental design, data interpretation, and assistance in manuscript preparation and review.

Dr. Amy Pruden (Civil and Environmental Engineering, Virginia Tech). PhD co-advisor, committee co-chair and project co-PI. Co-author of chapters 2, 3, and 4. Contributed guidance on experimental design, data interpretation, and assistance in manuscript preparation and review.

Chapter 2 coauthors contributed to the bioinformatic processing of samples: **Monjura Afrin Rumi** (Genetics, Bioinformatics, and Computational Biology, Virginia Tech) annotated samples against the Comprehensive Antibiotic Resistance Database, **Nazifa Moumi** (Genetics, Bioinformatics, and Computational Biology, Virginia Tech) annotated samples against the Genome Taxonomy Database, and **Dr. Liqing Zhang** (Genetics, Bioinformatics, and Computational Biology, Virginia Tech) was the project co-PI. **Dr. Carla Finkielstein** (Department of Biological Sciences) and **Alessandro Ceci** (Molecular Diagnostics Laboratory, Fralin Biomedical Research Institute, Virginia Tech) extracted RNA and **Dr. Raul Gonzalez** (Hampton Roads Sanitation District) conducted ddPCR analyses. Samples previously collected by **Haniyyah Majeed** were also included in the analyses.

Chapter 3 was a project involving a number of co-authors. I designed and primarily maintained the sequencing batch reactors, however, **Connor Brown** (Genetics, Bioinformatics, and Computational Biology, Virginia Tech) and I processed samples together. Connor performed the bioinformatic data processing and provided insights into the data analysis. **Dr. Latania Logan** (Rush University Medical Center, Chicago, Illinois, USA) provided hospital wastewater. **Dr. Kang Xia** (Department of Crop and Soil Environmental Sciences) analyzed pharmaceutical and antibiotic concentrations.

Chapter 4 was a large scale collaborative project that utilized samples collected by a number of individuals associated with Virginia Tech (Civil and Environmental Engineering) at the time of sampling: **Dr. Maria Virginia Prieto Riquelme**, **Dr. Emily Garner**, **Dr. Jake Metch**, **Dr. Marjorie Willner** and **Matthew Blair**. Co-PIs included: **D.G. Joakim Larsson** (Department of Infectious Diseases, University of Gothenburg), **Dr. Helmut Bürgmann** (Eawag: Swiss Federal Institute of Aquatic Science and Technology), **Dr. Tong Zhang** (Department of Civil Engineering, The University of Hong Kong), **Dr. Indumathi M. Nambi** (Department of Civil Engineering, Indian Institute of Technology), and **Dr. Diana Aga** (Department of Chemistry, University at Buffalo) who also analyzed antibiotic concentrations.

TABLE OF CONTENTS

ABSTRACT.....	ii
GENERAL AUDIENCE ABSTRACT.....	iii
DEDICATION.....	v
ACKNOWLEDGEMENTS.....	v
ATTRIBUTIONS.....	vi
TABLE OF CONTENTS.....	vii
LIST OF FIGURES.....	x
LIST OF TABLES.....	xi
CHAPTER 1: INTRODUCTION.....	1
1.1 Overview and Research Motivation.....	1
1.1.1 Antibiotic Resistance.....	1
1.1.2 COVID-19 Pandemic and Wastewater-Based Surveillance.....	2
1.1.3 AMR and Biological Treatment.....	3
1.1.4 WWTP and the Environmental Resistome.....	4
1.1.5 Methods for examining AMR in the environment.....	4
1.2 Research Objectives.....	5
1.3 Annotated Dissertation Outline.....	5
1.4 References.....	6
CHAPTER 2: SURVEILLANCE OF ANTIMICROBIAL RESISTANCE IN TWO RURAL U.S. WWTPS DURING THE COVID-19 PANDEMIC.....	10
2.1 Abstract.....	10
2.2 Introduction.....	10
2.3 Materials and Methods.....	13
2.3.1 Site Description and Sampling.....	13
2.3.2 SARS-CoV-2.....	13
2.3.3 Sequencing and Bioinformatics.....	14
2.3.4 PPCPs and QACs.....	15
2.3.5 Statistics.....	15
2.4 Results.....	16
2.4.1 WWTP Physiochemical Parameters.....	16
2.4.2 SARS-CoV-2.....	16
2.4.3 QAC Resistance Genes.....	19
2.4.4 Resistome Comparison.....	23
2.4.5 Intrinsic and Extrinsic Factors.....	26
2.5 Discussion.....	27
2.5.1 Pandemic/University Impact.....	27
2.5.2 QAC Resistance Trends.....	30
2.5.3 Wastewater Based Surveillance.....	30
2.5.4 Next Steps.....	31
2.6 Conclusions.....	31
2.7 Acknowledgements.....	32
2.8 References.....	32
APPENDIX A.....	39
A.1 Supporting Tables.....	39

A.2 Supplementary Figures	42
CHAPTER 3: IMPACT OF HOSPITAL WASTEWATER ON SIMULATED ACTIVATED SLUDGE TREATMENT AND ANTIBIOTIC RESISTOME	48
3.1 Abstract	48
3.2 Water Impact.....	48
3.3 Introduction.....	48
3.4 Materials and Methods.....	50
3.4.1 SBR Setup and Operation	50
3.4.2 Sample Collection and DNA Extraction.....	51
3.4.3 Shotgun Metagenomic Sequencing Analysis.....	52
3.4.4 Statistical Analyses	52
3.5 Results.....	53
3.5.1 Feed.....	53
3.5.2 SBR Performance.....	55
3.5.3 Comparison of Mixed Liquors Across the SBR Conditions.....	55
3.5.4 Taxonomic Groups Containing Pathogens	56
3.5.5 Antibiotic Resistance Genes	58
3.5.6 Clinically-Relevant ARGs and Associated Taxa.....	61
3.6 Discussion.....	63
3.6.1 Differences in HWW and SBR Feed	63
3.6.2 Activated Sludge is a Distinct Ecosystem that was Only Subtly Affected by the HWW	63
3.6.3 Impact of Hospital Wastewater.....	64
3.6.4 Limitations	64
3.7 Conclusions.....	65
3.8 Acknowledgements.....	65
3.9 References.....	66
APPENDIX B	71
B.1 Supplementary Tables	71
B.2 Supplementary Figures.....	80
CHAPTER 4: INTERNATIONAL COMPARISON OF ANTIBIOTIC RESISTANCE GENES FROM WASTEWATER TREATMENT PLANT FINAL EFFLUENTS AND THEIR RECEIVING ENVIRONMENT	87
4.1 Abstract	87
4.2 Introduction.....	87
4.3 Materials and Methods.....	88
4.3.1 Site Description.....	88
4.3.2 Sample Collection and Processing.....	89
4.3.3 Molecular Analysis and High Throughput Sequencing.....	89
4.3.4 Statistical Analyses	89
4.4 Results.....	90
4.4.1 qPCR and Metagenomics.....	90
4.4.2 Resistome	90
4.4.3 Taxonomy	97
4.4.4 Antibiotics.....	100
4.5 Discussion	102

4.6	Conclusions.....	104
4.7	Acknowledgements.....	105
4.8	References.....	105
	APPENDIX C	110
C.1	Supplementary Figures.....	110
C.2	Supporting Tables	113
	CHAPTER 5: CONCLUSIONS	116
	APPENDIX D.....	118
D.1	Sub-sewershed SARS-CoV-2 Wastewater Surveillance and COVID-19 Epidemiology Using Building-specific Occupancy and Case Data.....	118

LIST OF FIGURES

Figure 2.1	16
Figure 2.2	18
Figure 2.3	19
Figure 2.4	20
Figure 2.5	22
Figure 2.6	24
Figure 2.7	25
Figure 2.8	28
Figure A.1	42
Figure A.2	43
Figure A.3	44
Figure A.4	45
Figure A.5	46
Figure A.6	46
Figure A.7	47
Figure A.8	47
Figure 3.1	51
Figure 3.2	54
Figure 3.3	57
Figure 3.4	60
Figure 3.6	62
Figure B.1	80
Figure B.2	80
Figure B.3	82
Figure B.4	82
Figure B.5	83
Figure B.6	84
Figure B.7	84
Figure B.8	85
Figure B.9	86
Figure 4.1	91
Figure 4.2	92
Figure 4.3	93
Figure 4.4	94
Figure 4.5	96
Figure 4.6	98
Figure 4.7	99
Figure 4.8	101
Figure 4.9	102
Figure C.1	110
Figure C.2	111
Figure C.3	112
Figure C.4	112

LIST OF TABLES

Table 1.1	3
Table 2.1	18
Table 2.2	21
Table 2.3	27
Table A.1.....	39
Table A.2.....	39
Table A.3.....	40
Table A.4.....	41
Table A.5.....	41
Table 3.1	58
Table B.1.....	71
Table B.2.....	72
Table B.3.....	73
Table B.4.....	73
Table B.5.....	75
Table B.6.....	76
Table B.7.....	77
Table B.8.....	78
Table B.9.....	79
Table B.10.....	79
Table 4.1	95
Table C.1.....	113
Table C.2.....	114
Table C.3.....	114

CHAPTER 1: INTRODUCTION

1.1 Overview and Research Motivation

Wastewater treatment plants (WWTPs) are important engineered systems within the urban water cycle, acting as barriers protecting human health and the environment by preventing the discharge of untreated waste containing pathogens, nutrients, and organic matter into the environment. Recently, especially in light of the COVID-19 pandemic, there is also increasing interest in utilizing WWTPs as a resource for surveilling communities and their health status by examining components and characteristics of raw sewage coming in to the WWTP (i.e., influent). The urban water cycle is strongly influenced by humans. One perspective of this cycle may focus on people contributing waste to sewer systems, this waste flows to WWTPs where it is treated and then discharged into the environment. People interact with this treated water through activities, including recreation (e.g., swimming, surfing), or as resources, when the water is used for irrigation or drinking water. This dissertation examines WWTPs from various aspects of this urban cycle through the lens of their role as a monitoring and mitigation point for antibiotic resistance.

1.1.1 Antibiotic Resistance

The discovery of antibiotics provided a revolutionary weapon against bacterial infections and, along with improved sanitation and hygiene, drastically increased life expectancy and decreased mortality rates from infectious diseases. Correspondingly, there was a major shift in mortality from the 19th and 20th centuries primarily from infectious to chronic diseases (1,2). However, with the rise in antimicrobial resistance (AMR), it is estimated that AMR infections will cause almost 10 million deaths globally by 2050. This is greater than the estimated deaths expected to be caused by cancer (8.2 million) in 2050 (3). AMR arises when microorganisms are no longer affected by the antibiotic, antibacterial, antiparasitic, antiviral or antifungal drugs designed to kill them (4). Antibiotic resistance is a subset of AMR specific to bacteria, wherein antibiotic resistance genes (ARGs) enable bacteria to survive in the presence of antibiotics. This can be accomplished by various mechanisms at the cellular level, including efflux pumps, cell wall modifications, enzymatic destruction, and modification of the antibiotic structure (5). A microorganism may become resistant to antimicrobials through mutations or by acquiring foreign DNA through horizontal gene transfer (HGT). Selective pressure is a powerful ecological force that can compel bacteria to acquire or maintain ARGs. Misuse and overuse of antibiotics elevates selective pressure in affected environments, including the corresponding gut bacteria, sewage collection and treatment systems, and environments that receive treated wastewater. Other antimicrobials, including disinfectants and heavy metals, can also impose selective pressure for antibiotic resistance. Mobile genetic elements (MGEs) (e.g., plasmids, transposons, integrons) play a key role in transferring ARGs as they promote intracellular DNA mobility (e.g., from the chromosome to a plasmid or between plasmids) and intercellular DNA mobility (i.e., HGT) (6). HGT is a major mechanism through which ARGs are shared via transformation (uptake, integration, and functional expression of free extracellular DNA), conjugation (transfer of DNA from cell to cell via surface pili), or transduction (transfer of genes via bacteriophages) (7).

Human behavior surrounding antibiotics has a direct impact on AMR in clinical settings and in the environment. Of particular interest to the research described in this dissertation were the drastic worldwide changes in behavior instigated by the COVID-19 pandemic. It has been hypothesized that antibiotic resistance likely increased as a result of pandemic-related behavioral

changes (8–11), such as decreased stewardship programs in hospitals (12,13), and the overprescription, and misuse of antibiotics, especially for COVID-19 patients (13,14).

1.1.2 COVID-19 Pandemic and Wastewater-Based Surveillance

An outbreak of pneumonia-like infections occurred in December of 2019 in Wuhan, China. The virus that caused this outbreak was later classified as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), an enveloped, positive single-stranded RNA virus. Coronavirus disease-2019 (COVID-19) cases quickly spread and the World Health Organization (WHO) declared a global pandemic (15). SARS-CoV-2 primarily causes a respiratory illness; however, symptoms of infection may vary from mild to multi-organ failure and death (16). On the other hand, some individuals who test positive may not experience symptoms. A literature review found that at least one third of cases are asymptomatic (17). Additional studies have found that 17.6% of patients with COVID-19 had gastrointestinal symptoms (18). GI symptoms increase the chances of a person shedding virus in their stool (with GI symptoms 52.5% vs. without 25.9%). Overall, 43-48.1% of patients shed SARS-CoV-2 RNA in their stool (18,19). This phenomenon allows for the monitoring of infections through wastewater surveillance. Monitoring of COVID-19 levels in communities through the detection of SARS-CoV-2 RNA concentrations in wastewater greatly expanded the field of wastewater surveillance and epidemiology during the pandemic. There are inherent challenges to scaling up this practice, however. For example, not all infected individuals shed SARS-CoV-2 RNA in their feces and, if they do, the concentrations shed do not consistently relate to the stage or severity of illness.

Wastewater surveillance was first used as a means to examine drug use within localized populations in the early 2000s (20,21). This approach was highlighted as a way to elucidate drug use patterns, without relying on invasive procedures or self-reporting via surveys, which can be unreliable. It can also be more cost effective to monitor the health of a community via wastewater, as each individual does not need to be tested because the wastewater effectively provides a pooled sample (22). Linking pathogens and chemicals found in wastewater to population-level health has grown into the field known as wastewater-based epidemiology (WBE) (23). Most recently, wastewater-based surveillance (WBS) (i.e., testing wastewater for public health benefit) has been used to track patterns and outbreaks of COVID-19. While neither WBE nor WBS can replace clinical testing, they can be important supplemental tools to help inform about the health of a community. In an analogous fashion, wastewater analyses can also act as a tool for the surveillance of AMR across a community, complimenting clinical surveillance efforts. Clinical surveillance focuses specifically on ill individuals, a smaller proportion of the population (24), and thus misses asymptomatic carriers. Also, clinical surveillance is thought to only capture a fraction of true resistant infection rates because often treatment proceeds without effort to culture etiological agents. The efficiency of wastewater surveillance could also potentially better support efforts to collect multiple comparable data points and track resistance trends over time. Such data are urgently needed to help inform priorities (national and global), assess interventions, identify new forms of resistance, and identify and stem outbreaks of resistant pathogens (25). The research presented herein used WBS to determine changes in the resistome (i.e., the sum of all ARGs (26)) in local WWTP influent prior, during, and after the COVID-19 pandemic to determine if changes in human behavior influencing AMR during this time could be detected in the wastewater.

1.1.3 AMR and Biological Treatment

As described above, wastewater can be a valuable monitoring point to surveil what is occurring in a community. However, it is also important to consider how AMR determinants might be impacting the wastewater resistome during treatment. Not all AMR is created equally: ARGs that encode resistance to clinically-relevant antibiotics with established importance to public health are of particular concern. The WHO established a list of medically important antimicrobials ranked based on a number of criteria. The antimicrobials that were categorized as the highest priority “critically important” were: cephalosporins (3rd, 4th and 5th generation), glycopeptides, macrolides and ketolides, polymyxins, and quinolones (27). The criteria (C) or prioritization factors (P) used by the WHO to distinguish medically important antibiotics are: (C1) limited available therapies to treat serious bacterial infections, (C2) used to treat infections caused by bacteria that may be transmitted from non-human sources or bacteria with resistance genes from non-human sources, (P1) are used to treat a large number of people (P2) often in high risk groups that have limited effective antimicrobials and (P3) there is extensive evidence for the transmission of ARB or ARGs from non-human sources (**Table 1.1**). Therefore, particular attention should be given to ARBs and ARGs that confer resistance to these antibiotics.

Table 1.1 Table of high priority antimicrobial classes that fell under all five of the criterion and prioritization factors as classified by the WHO (27) (Compiled from WHO Critically important antimicrobials for human medicine: 6th revision, page 11 and 24 with permission from WHO)

WHO Highest priority critically important antimicrobials					
Antimicrobial Class	Cephalosporins (3rd, 4th and 5th generation)	Glycopeptides	Macrolides and ketolides	Polymyxins	Quinolones
WHO Criterion (C) / Prioritization Factor (P)					
C1	C2	P1	P2	P3	
Sole, or one of limited available therapies, to treat serious bacterial infections in people	Used to treat infections caused by bacteria (1) possibly transmitted from non-human sources, or (2) with resistance genes from non-human sources	Used to treat a large number of people with infections for which limited antimicrobials are available	Used with high frequency in human medicine or in certain high risk groups	Used to treat human infections for which an extensive evidence exists on the transmission of resistant-bacteria or genes from non-human sources	

Relative to municipal sewage, hospital wastewater is known to have a greater microbial load of pathogens, ARBs, ARGs as well as higher concentrations of pharmaceuticals, including non-metabolized pharmaceuticals (unused and disposed of down the drain or excreted), disinfectants, heavy metals, and x-ray contrast media (28,29). Therefore, as an important source of antibiotic resistance in wastewater, hospitals and how they impact the wastewater resistome are an important monitoring point to consider. Examining AMR during wastewater treatment is important because of the concern that biological wastewater treatment may promote the spread of ARGs and AMR through HGT due to the high concentrations of bacteria in close proximity that may be exposed to sub-inhibitory antibiotic concentrations (30,31). Determining important sources of AMR in wastewater treatment and what is occurring during biological treatment is important to

then determine if additional pretreatment of relevant sources (i.e., hospital wastewater) is required. The research presented in this dissertation aimed to help further this question by examining a higher percentage of hospital wastewater and how that impacted the wastewater resistome in a controlled environment.

1.1.4 WWTP and the Environmental Resistome

Following along the urban water cycle, the impact that WWTPs have on aquatic environments must be explored. Concentrations and types of ARGs in aquatic environments have been shown to be influenced by a number of factors including storm events (32,33), urbanization (34), and WWTPs. Hiller and colleagues (34) compiled a review of studies using qPCR and culturing and found overall an increased relative and absolute abundance of ARGs downstream of WWTPs. Similarly, studies using metagenomic sequencing have found increased ARG abundances, taxonomic diversity (35), and bacteria resistant to clinically-relevant antibiotics (36) downstream of WWTPs. While WWTPs are often seen as hotspots for the dissemination of AMR into the environment (30) other sources include combined sewer overflow, aging infrastructure resulting in leaky pipes, and agricultural runoff (34,37,38).

The levels of AMR in affected environments are important because of the ways that we as humans interact with such environments (swimming, recreational, irrigation of food crops etc.). While hospitals have been shown as clear sources of antibiotic resistant infections (i.e., nosocomial infections) (39), the environment should not be discounted as an additional possible source. Studies have found increased rates of cefotaxime-resistant *Escherichia coli* colonization of surfer's guts compared with non-surfers (40) and that freshwater swimming is a risk factor for community-acquired urinary tract infections caused by extended-spectrum β -lactamase *Enterobacteriaceae* (41). While these links need to be further investigated for their impact on public health, they demonstrate that the aquatic environment represents a potential source of ARBs. The fourth chapter of this dissertation examined the aquatic resistomes upstream and downstream of WWTP discharge points to determine the impacts of the effluent across countries with variable land use, urbanization, population density, and antibiotic use.

1.1.5 Methods for examining AMR in the environment

Antibiotic resistance in the environment is commonly studied through culture-based or molecular-based approaches. Molecular methods include polymerase chain reaction (PCR)-derived techniques (e.g., PCR, quantitative PCR (qPCR), droplet digital PCR (ddPCR)) and next generation sequencing (e.g., amplicon sequencing, metagenomics, long-read sequencing) (42,43).

Culture-based methods allow enumeration of the viable culturable bacteria present in a sample. Further analyses on isolates allow for antibiotic resistance testing, sequence-based typing, and whole genome sequencing (44). Culture-based targets are also used in human health risk assessments. Culturing environmental samples for clinically-relevant bacteria is challenging when using media typically used for human samples. Environmental samples contain a wide variety of bacteria that yield false positives, many of which are not clinically relevant, but are intrinsically resistant (i.e., are able to grow in the presence of antibiotics used in human medicine) and able to grow on the selective media optimized for human materials (45).

PCR-based methods allow one to determine which genes are present in an environmental sample and depending on the method selected, to quantify the concentration of the gene of interest. PCR based methods are more sensitive, allowing for the detection of genes at lower concentrations, however, for sensitive detection limits pre-amplification may be required (46). A drawback of

PCR-based methods is the required *a priori* knowledge of the gene(s) of interest. While high-throughput qPCR allows for the quantification of hundreds of genes at once, it does not allow for the optimization of individual assays (47). qPCR is known to be sensitive to inhibitors found in the environment; ddPCR is gaining traction as it is less sensitive to inhibition and due to its high specificity and lower detection limits (48).

Next generation sequencing is a culture-independent method that allows researchers to examine an environment for DNA or RNA segments, these genetic materials can be used for the identification of genomes, species, or the ARGs present within a given sample. Next generation sequencing strategies have historically focused on short-reads (~150-300 bp) (e.g., Illumina Sequencing), which provide accuracy and depth, allowing for efficient profiling of ARGs or MGEs in an environment. Long-read sequencing (>10,000 bp) technologies (e.g., Oxford Nanopore) are less accurate than short read sequencing technologies, however, longer read sequencing provides stronger evidence of the physical location of genes in relation to one another and thus is more suitable for constructing whole genomes. Hybrid assembly combines the best of both worlds by either assembling long-reads and using short reads for assembly polishing by correcting incorrectly called bases (49), or by assembling short reads and then filling in gaps using long reads (50). Regardless of the assembly method, hybrid reads allow for greater recovery of distinct genomes that allow determination of sequences of interest and the species from which they are derived.

Each of the methods detailed above have been used to study the fate and transport of ARGs in wastewater. Research has aimed to explain the impact that WWTPs have on the environment, the detection of novel ARGs and monitoring their spread across the world, and to examine population use of antibiotics and how they contribute to the ARGs detected in wastewater. The research described in this dissertation primarily focuses on metagenomic sequencing to examine AMR markers in wastewater in different countries and over time under varying conditions.

1.2 Research Objectives

The primary goal of this research was to assess antibiotic resistance in wastewater through metagenomic sequencing and to understand the role that WWTPs play in receiving, attenuating, and disseminating ARGs. To achieve this goal, the following objectives were pursued:

1. Examine the relationship between the numbers of clinical cases of COVID-19, SARS-CoV-2 RNA concentrations, ARGs, quaternary ammonium compound resistance genes, and the wastewater microbiome of two rural WWTPs.
2. Investigate how different percentages of hospital wastewater discharged to domestic wastewater streams impact the performance of activated sludge treatment, which forms the core of most conventional WWTPs, and corresponding abundance of ARGs and key taxonomic groups.
3. Assess the impacts of WWTP effluent on stream microbiome and resistomes spanning multiple countries.

1.3 Annotated Dissertation Outline

Chapter 1: Introduction. This chapter provides background for relevant information giving context to the motivation of the research discussed in this dissertation.

Chapter 2: Surveillance of antimicrobial resistance in two rural U.S. WWTPs during the COVID-19 pandemic. This manuscript addresses the first objective by examining the impact that the COVID-19 pandemic had on markers of antibiotic resistance in wastewater. Two rural U.S.

WWTPs were surveilled. The WWTP receiving a larger proportion of university wastewater was found to be more drastically impacted, with reduced diversities of ARGs and taxa in 2020-2021. Changes in ARGs were statistically correlated with taxa as a possible explanation of differences between WWTPs and over time.

Chapter 3: Impact of hospital wastewater on simulated activated sludge treatment and antibiotic resistome. This manuscript addressed the second objective through lab-scale sequencing batch reactors (SBRs) fed a mixture of domestic and hospital wastewater. After hospital wastewater dosing the performance of the SBRs and important genera remained stable; the resistomes shifted slightly with clinically relevant ARGs (specifically *vanA*) detected more frequently in the SBRs dosed with hospital effluent. This manuscript is in preparation for submission to *Environmental Science: Water Research and Technology*.

Chapter 4: International comparison of antibiotic resistance genes from wastewater treatment plant final effluents and their receiving environment. This manuscript sought to address the third objective by examining rivers impacted by WWTPs in North America, Europe, and Asia. Results differed depending on method with qPCR indicating an increase of ARGs downstream of WWTPS and metagenomics revealing similar resistomes more strongly influenced by country and taxa than WWTP.

Chapter 5: Conclusions. This chapter summarizes the contributions this research made to the field of environmental engineering in understanding antibiotic resistance throughout wastewater treatment and makes recommendations for future work based on the observations made in this dissertation.

1.4 References

1. Schlipkötter U, Flahault A. Communicable diseases: Achievements and challenges for public health. Vol. 32, Public Health Reviews. 2010.
2. Chokshi A, Sifri Z, Cennimo D, Horng H. Global Contributors to Antibiotic Resistance. J Glob Infect Dis [Internet]. 2019 Jan 1 [cited 2021 Sep 2];11(1):36. Available from: [/pmc/articles/PMC6380099/](https://pubmed.ncbi.nlm.nih.gov/3280099/)
3. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. 2016;
4. WHO. Antimicrobial Resistance: Global Report on Surveillance [Internet]. 2014 [cited 2017 Oct 3]. Available from: http://apps.who.int/iris/bitstream/10665/112642/1/9789241564748_eng.pdf?ua=1
5. Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool MH, et al. Antibiotic resistance: a rundown of a global crisis. Infect Drug Resist [Internet]. 2018 [cited 2021 Feb 4];11:1645–1658. Available from: <http://dx.doi.org/10.2147/IDR.S173867>
6. Partridge SR, Kwong SM, Firth N, Jensen SO. Mobile genetic elements associated with antimicrobial resistance. Clin Microbiol Rev [Internet]. 2018 Oct 1 [cited 2022 Feb 16];31(4). Available from: <https://journals.asm.org/doi/abs/10.1128/CMR.00088-17>
7. von Wintersdorff CJ, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Front Microbiol [Internet]. 2016;7:173. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/26925045>

8. Usman M, Farooq M, Hanna K. Environmental side effects of the injudicious use of antimicrobials in the era of COVID-19. *Sci Total Environ*. 2020 Nov 25;745:141053.
9. Rawson TM, Ming D, Ahmad R, Moore LSP, Holmes AH. Antimicrobial use, drug-resistant infections and COVID-19. *Nat Rev Microbiol*. 2020;18(8):409–10.
10. Lu J, Guo J. Disinfection spreads antimicrobial resistance. *Science* (80-). 2021;371(6528):474.
11. Rezasoltani S, Yadegar A, Hatami B, Aghdaei HA, Zali MR. Antimicrobial resistance as a hidden menace lurking behind the COVID-19 outbreak: the global impacts of too much hygiene on AMR. *Front Microbiol*. 2020;11.
12. Chibabhai V, Duse AG, Perovic O, Richards GA. Collateral damage of the COVID-19 pandemic: Exacerbation of antimicrobial resistance and disruptions to antimicrobial stewardship programmes? *South African Med J* [Internet]. 2020 Jun 2 [cited 2021 Jul 19];110(7). Available from: <https://doi.org/10.1086/591708>
13. Ansari S, Hays JP, Kemp A, Okechukwu R, Murugaiyan J, Ekwanzala MD, et al. The potential impact of the COVID-19 pandemic on global antimicrobial and biocide resistance: an AMR Insights global perspective. *JAC-Antimicrobial Resist*. 2021;3(2):dlab038.
14. Cong W, Poudel N, Alhusein N, Wang H, Yao G, Lambert H. Antimicrobial use in COVID-19 patients in the first phase of the SARS-CoV-2 pandemic: Rapid review and evidence synthesis. *medRxiv*. 2021;
15. Cucinotta D, Vanelli M. WHO Declares COVID-19 a Pandemic. *Acta Bio Medica Atenei Parm* [Internet]. 2020 [cited 2021 Jul 19];91(1):157. Available from: </pmc/articles/PMC7569573/>
16. Puoti MG, Rybak A, Kiparissi F, Gaynor E, Borrelli O. SARS-CoV-2 and the Gastrointestinal Tract in Children. *Front Pediatr*. 2021 Feb 22;9:67.
17. Oran DP, Topol EJ. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic. <https://doi.org/107326/M20-6976> [Internet]. 2021 Jan 22 [cited 2021 Jul 22];174(5):655–62. Available from: <https://www.acpjournals.org/doi/abs/10.7326/M20-6976>
18. Cheung KS, Hung IFN, Chan PPY, Lung KC, Tso E, Liu R, et al. Gastrointestinal Manifestations of SARS-CoV-2 Infection and Virus Load in Fecal Samples From a Hong Kong Cohort: Systematic Review and Meta-analysis. *Gastroenterology*. 2020 Jul 1;159(1):81–95.
19. Zhang Y, Cen M, Hu M, Du L, Hu W, Kim JJ, et al. Prevalence and Persistent Shedding of Fecal SARS-CoV-2 RNA in Patients With COVID-19 Infection: A Systematic Review and Meta-analysis. *Clin Transl Gastroenterol* [Internet]. 2021 Apr 9 [cited 2021 Jul 22];12(4):e00343. Available from: </pmc/articles/PMC8036078/>
20. Zuccato E, Chiabrando C, Castiglioni S, Calamari D, Bagnati R, Schiarea S, et al. Cocaine in surface waters: A new evidence-based tool to monitor community drug abuse. *Environ Heal A Glob Access Sci Source* [Internet]. 2005 Aug 5 [cited 2021 Jan 30];4:14. Available from: </pmc/articles/PMC1190203/?report=abstract>
21. Daughton CG. Illicit Drugs in Municipal Sewage Proposed New Nonintrusive Tool to Heighten Public Awareness of Societal Use of Illicit-Abused Drugs and Their Potential for Ecological Consequences. 2001 [cited 2021 Jan 30]; Available from: <http://www.mnmter.net/~pu>
22. Hart OE, Halden RU. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities

- and challenges. *Sci Total Environ.* 2020 Aug 15;730:138875.
23. Larsen DA, Green H, Collins MB, Kmush BL. Wastewater monitoring, surveillance and epidemiology: a review of terminology for a common understanding. *FEMS Microbes* [Internet]. 2022 Jan 13 [cited 2022 Feb 16];2:11. Available from: <https://academic.oup.com/femsmicrobes/article/doi/10.1093/femsmc/xtab011/6354777>
 24. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology: Monitoring infectious disease spread and resistance to the community level. Vol. 139, *Environment International*. Elsevier Ltd; 2020. p. 105689.
 25. Harrington A, Vo V, Papp K, Tillett RL, Chang C-L, Baker H, et al. Urban monitoring of antimicrobial resistance during a COVID-19 surge through wastewater surveillance. *Sci Total Environ* [Internet]. 2022 Jan 2;853:158577. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969722056765>
 26. Wright GD. The antibiotic resistome: The nexus of chemical and genetic diversity. Vol. 5, *Nature Reviews Microbiology*. 2007.
 27. World Health Organisation. Critically Important Antimicrobials for Human Medicine 6th Revision 2018 [Internet]. 2019. Available from: <https://apps.who.int/Iris/Bitstream/Handle/10665/312266/9789241515528-Eng.Pdf?Ua=1>
 28. Rodriguez-Mozaz S, Lucas D, Barceló D. Full-Scale Plants for Dedicated Treatment of Hospital Effluents. In: *The Handbook of Environmental Chemistry* [Internet]. 60th ed. Springer, Cham; 2017 [cited 2022 Nov 18]. p. 189–208. Available from: <http://www.springer.com/series/698>
 29. Emmanuel E, Perrodin Y, Keck G, Blanchard JM, Vermande P. Ecotoxicological risk assessment of hospital wastewater: A proposed framework for raw effluents discharging into urban sewer network. *J Hazard Mater.* 2005 Jan 14;117(1):1–11.
 30. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Sci Total Environ* [Internet]. 2013 [cited 2017 Oct 3];447:345–60. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2013.01.032>
 31. Karkman A, Do TT, Walsh F, Virta MPJ. Antibiotic-Resistance Genes in Waste Water. 2017 [cited 2022 Nov 22]; Available from: <https://doi.org/10.1016/j.tim.2017.09.005>
 32. Baral D, Dvorak BI, Admiraal D, Jia S, Zhang C, Li X. Tracking the Sources of Antibiotic Resistance Genes in an Urban Stream during Wet Weather using Shotgun Metagenomic Analyses. *Environ Sci Technol* [Internet]. 2018 Aug 21 [cited 2021 Feb 6];52(16):9033–44. Available from: <https://pubs.acs.org/sharingguidelines>
 33. Garner E, Benitez R, von Wagoner E, Sawyer R, Schaberg E, Hession WC, et al. Stormwater loadings of antibiotic resistance genes in an urban stream. *Water Res.* 2017 Oct 15;123:144–52.
 34. Hiller CX, Hübner U, Fajnorova S, Schwartz T, Drewes JE. Antibiotic microbial resistance (AMR) removal efficiencies by conventional and advanced wastewater treatment processes: A review. Vol. 685, *Science of the Total Environment*. 2019.
 35. Zieliński W, Hubeny J, Buta-Hubeny M, Rolbiecki D, Harnisz M, Paukszto Ł, et al. Metagenomics analysis of probable transmission of determinants of antibiotic resistance from wastewater to the environment – A case study. *Sci Total Environ.* 2022;827.
 36. Ju F, Lee J, Beck K, Zhang G, Gekenidis M-T, Hummerjohann J, et al. Phenotypic Metagenomics tracks Wastewater-Associated Clinically Important Beta-lactam Resistant Bacteria Invading River Habitats [Internet]. In Review; 2022 Jan. Available from:

- <https://www.researchsquare.com/article/rs-1589365/v1>
37. Nappier SP, Liguori K, Ichida AM, Stewart JR, Jones KR. Antibiotic resistance in recreational waters: State of the science. Vol. 17, *International Journal of Environmental Research and Public Health*. 2020.
 38. He Y, Yuan Q, Mathieu J, Stadler L, Senehi N, Sun R, et al. Antibiotic resistance genes from livestock waste: occurrence, dissemination, and treatment. Vol. 3, *npj Clean Water*. 2020.
 39. Serra-Burriel M, Keys M, Campillo-Artero C, Agodi A, Barchitta M, Gikas A, et al. Impact of multi-drug resistant bacteria on economic and clinical outcomes of healthcare-associated infections in adults: Systematic review and meta-analysis. *PLoS One*. 2020;15(1).
 40. Leonard AFC, Zhang L, Balfour AJ, Garside R, Hawkey PM, Murray AK, et al. Exposure to and colonisation by antibiotic-resistant *E. coli* in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). *Environ Int*. 2018;114.
 41. Søråas A, Sundsfjord A, Sandven I, Brunborg C, Jenum PA. Risk Factors for Community-Acquired Urinary Tract Infections Caused by ESBL-Producing Enterobacteriaceae -A Case-Control Study in a Low Prevalence Country. *PLoS One*. 2013;8(7).
 42. Zhang X-X, Zhang T, Fang HHP. Antibiotic resistance genes in water environment. *Appl Microbiol Biotechnol*. 2009;82(3):387–414.
 43. Crofts TS, Gasparrini AJ, Dantas G. Next-generation approaches to understand and combat the antibiotic resistome. *Nat Rev Microbiol* [Internet]. 2017;15(7):422–34. Available from: www.nature.com/nrmicro
 44. Liguori K, Keenum I, Davis BC, Calarco J, Milligan E, Harwood VJ, et al. Antimicrobial Resistance Monitoring of Water Environments: A Framework for Standardized Methods and Quality Control. *Environ Sci Technol* [Internet]. 2022 Jan 17;56(13):9149–60. Available from: <https://doi.org/10.1021/acs.est.1c08918>
 45. Schreiber C, Zacharias N, Essert SM, Wasser F, Müller H, Sib E, et al. Clinically relevant antibiotic-resistant bacteria in aquatic environments – An optimized culture-based approach. *Sci Total Environ*. 2021;750.
 46. Ishii S, Segawa T, Okabe S. Simultaneous quantification of multiple food- and waterborne pathogens by use of microfluidic quantitative PCR. *Appl Environ Microbiol*. 2013;79(9).
 47. Waseem H, Jameel S, Ali J, Ur Rehman HS, Tauseef I, Farooq U, et al. Contributions and challenges of high throughput qPCR for determining antimicrobial resistance in the environment: A critical review. Vol. 24, *Molecules*. 2019.
 48. Park S, Rana A, Sung W, Munir M. Competitiveness of Quantitative Polymerase Chain Reaction (qPCR) and Droplet Digital Polymerase Chain Reaction (ddPCR) Technologies, with a Particular Focus on Detection of Antibiotic Resistance Genes (ARGs). *Appl Microbiol*. 2021;1(3).
 49. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* [Internet]. 2020;17:103–1110. Available from: <https://doi.org/10.1038/s41592-020-00971-x>
 50. Bertrand D, Shaw J, Kalathiyappan M, Hui A, Ng Q, Senthil Kumar M, et al. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. *Nat Biotechnol* [Internet]. 2019 [cited 2022 Feb 13];37:937–944. Available from: <https://doi.org/10.1038/s41587-019-0191-2>

CHAPTER 2: SURVEILLANCE OF ANTIMICROBIAL RESISTANCE IN TWO RURAL U.S. WWTPS DURING THE COVID-19 PANDEMIC

Co-authors: Monjura Afrin Rumi, Nazifa Moumi, Liqing Zhang, Haniyyah Majeed, Carla Finkielstein, Alessandro Ceci, Raul Gonzalez, Amy Pruden, Peter Vikesland

2.1 Abstract

It has been hypothesized that increased use of disinfectants and antibiotics during the COVID-19 pandemic would contribute to an increase in antimicrobial resistance in affected human and environmental microbial communities. The objective of this research was to assess the relationship between concentrations of SARS-CoV-2 RNA, antibiotic resistance genes, and quaternary ammonium compound resistance genes in the sewage influent to two rural wastewater treatment plants (WWTPs) and the corresponding wastewater microbiomes and numbers of clinical cases of COVID-19 in the community. Interestingly, the resistome and microbiome of the WWTP receiving a comparably larger proportion of wastewater from a university was more drastically impacted by behavioral changes (e.g., remote work) during the pandemic. Trends in resistance genes were strongly correlated with observed shifts in the microbiome and appeared to be more greatly impacted by seasonal variability and wastewater characteristics than specific factors at play during peak behavioral changes associated with the COVID-19 pandemic.

2.2 Introduction

In the early 2000s, wastewater was identified as an informative monitoring point for indicators of public health. Initially wastewater surveillance focused on a means to better quantify the scope of illicit drug use within localized populations (1,2). Expanding beyond drug use, wastewater surveillance has also been studied as a means to detect other targets reflective of human behavior, including plastic disposal (3) and alcohol consumption (4). Polio stands as a key example where such wastewater-based surveillance has actually been formally implemented at an institutional level, e.g., as an early warning system of outbreaks in Finland and Israel (5). The onset of the COVID-19 pandemic greatly accelerated interest in wastewater based surveillance (WBS) research and application, as summarized in a recent review by Larsen et al. (6). WBS has now been applied at previously unprecedented scale for the tracking of SARS-CoV-2 RNA, which has been successfully detected in wastewater around the world, including rural and urban areas as well as in countries of various income levels (7–10). Many studies have demonstrated striking trends in the concentration of SARS-CoV-2 RNA in wastewater and how they can be predictive of trends of clinical COVID-19 cases in the corresponding community (11,12). In September of 2020, the US Centers for Disease Control and Prevention (CDC) launched the National Wastewater Surveillance System (NWSS), which focuses on tracking SARS-CoV-2 levels in wastewater (13). NWSS has plans to expand to other public health targets, including antimicrobial resistance (AMR) (14,15). While WBS cannot replace clinical testing of diseases or disease markers, it holds promise as an important supplemental tool that can help provide key information about relevant public health indicators.

AMR is considered one of the top ten threats facing humanity. AMR limits the effectiveness of antibiotics for killing bacterial pathogens, thus making infections more difficult to treat (16). Excessive use and misuse of antibiotics contributes to the problem of AMR by exerting selection pressure, resulting in elevated antibiotic resistant bacteria (ARB) and antibiotic

resistance genes (ARGs) in the human gut, excreta, and corresponding sewage. For this reason, there has been emphasis on improving antibiotic stewardship practices, which are aimed at optimizing appropriate use of antibiotics when and where needed, and thus reducing opportunities for ARB and ARGs to evolve, become enriched, and spread (17). During the COVID-19 pandemic the World Health Organization (WHO) emphasized the importance of continuation of antibiotic stewardship programs. However, healthcare systems were overwhelmed and antibiotic stewardship became a lower priority relative to fighting the pandemic at hand. Surveys indicate that ~65% of respondents reported stewardship activities were negatively impacted by the pandemic (18,19)

At the onset of the pandemic, there was uncertainty regarding effective treatments for COVID-19. Therefore, antibiotics were often prescribed in an attempt to prevent secondary bacterial infections that were feared to arise from respiratory distress and other co-morbidities. It was reported that in U.S. hospitals, over 70% of COVID-19 patients were administered at least one antibiotic (February–April 2020 and March–October 2020), with 81% of prescriptions commencing immediately upon admission (20). Such practices occurred even though coinfection rates were subsequently found to be much lower than expected, ranging from 3–16% (21–23). Areas that experienced higher COVID-19 burdens especially struggled to maintain antibiotic stewardship practices. U.S. hospital data (n=716 hospitals) indicate that, on average, there were greater numbers of antibiotic therapy days per 1000 patient days than the years preceding the pandemic for cefepime (class beta-lactam), azithromycin (class macrolide), and doxycycline (class tetracycline) (20). According to the CDC, the infection rates reportedly decreased for some bacteria during the pandemic; however, increasing rates of infections were observed from 2019–2020 due to carbapenem-resistant *Acinetobacter* and Enterobacterales, extended spectrum beta-lactamase- (ESBL-) producing Enterobacterales, vancomycin-resistant Enterococcus (VRE), and multidrug-resistant (MDR) *Pseudomonas aeruginosa* (15).

While hospital inpatients may have been administered antibiotics at a higher frequency during the COVID-19 pandemic, outpatient prescriptions in the U.S. tell a different story. The U.S. National Action Plan for Combatting Antibiotic-Resistant Bacteria was released in 2015, with one major goal being reducing inappropriate outpatient antibiotic prescriptions by 50% by 2020. From 2011–2018 outpatient prescriptions were reduced by 10%; while from 2019–2020 there was a 25% decrease, partially due to decreased outpatient visits and healthcare seeking as a result of the pandemic (17). Data from the IQVIA Xponent database for oral antibiotic prescriptions reveal a similar story. From 2019 to 2020 the rates of oral antibiotic prescriptions decreased (per 1000 persons); however, the rates increased in 2021, though not back to the same rates seen in 2019 (24–26).

There has been a great deal of speculation as to how exactly the pandemic impacted AMR. Many scientists have hypothesized that the pandemic would give rise to increased levels of antibiotic resistant infections observed among patients and the greater release of ARBs and ARGs into the environment (27–30). This concern arose due to reports of decreased implementation of stewardship programs (31,32), over prescription and misuse of antibiotics for COVID-19 patients (32,33), and increased use of disinfectants due to calls for increased hygiene vigilance to prevent the spread of COVID-19 (30,34).

Increased hygiene vigilance appeared a logical intervention to prevent the spread of COVID-19, especially given what little was known about transmission routes at the onset of the pandemic. Correspondingly, the CDC recommended disinfecting surfaces in community settings at least daily (35), using one of the EPA's List N of disinfectants (36). From the list of over 500

EPA-approved products for COVID-19 disinfection, the most common active ingredients were quaternary ammonium compounds (QACs), followed by hydrogen peroxide, sodium hypochlorite, and ethanol. These four disinfectant types constitute ~75% of the EPA List N (37). Unfortunately, not only has it since been demonstrated that surface disinfection likely does little to prevent the spread of COVID-19, as selective agents, overuse of disinfectants creates the risk of promoting AMR.

Concerns about unanticipated negative consequences of overuse of disinfectants in the context of AMR were raised early during the COVID-19 pandemic. Chen and colleagues (37) voiced concerns that the lack of diversity in disinfectant-type may promote bacterial resistance since bacteria are repeatedly exposed to the same biocidal agent. Various disinfectants have been demonstrated to be capable of promoting antibiotic resistance through imposing a selective advantage for bacteria carrying antibiotic resistance genes (ARGs) and increasing tendency of ARGs to be horizontally transferred (30). Global sales of disinfectants and sanitizers increased during the pandemic and remained higher than pre-pandemic sales at least through March 2021 (37,38). Further, during the pandemic, wet disinfecting wipes, which may contain QACs (39), were flushed at such quantities that blockages were reported in many sewer systems (40,41). While not all disinfecting products are disposed of and/or washed down the drain, collectively these actions increased the likelihood of increased concentrations of disinfecting agents reaching sewers.

QACs, such as those found in common household products such as Lysol™, are biocides frequently used in solvents, cleaning products (soap and cleaning wipes), personal care products, and biomedical materials (42–44). QACs were used widely prior to the pandemic (45). Due to the presence of a positively-charged quaternary amine group and one or two long alkyl chains within the QAC structure (42) it has been hypothesized that QACs will sorb to negatively charged solids such as sludge, sediment, or soil. QACs have been detected in these environments as well as in wastewater effluent and receiving environments (42,46,47).

Some Gram-negative bacteria, such as *Pseudomonas aeruginosa*, possess natural resistance to QACs due to their intrinsic carriage of efflux pumps. Other bacteria acquire QAC resistance by either increasing the expression of existing efflux pumps, acquiring efflux pumps that remove QACs from the membrane, reducing membrane permeability, or increasing membrane stability (46). The QAC resistance gene *qacE* is also embedded in the class 1 integron, which is thought to play a critical role in the dissemination of multi-antibiotic resistance, and has been recovered from clinical isolates and environmental samples (48). Prior research is suggestive of an increase in co-resistance and cross resistance between QACs and a range of clinically-important antibiotics and disinfectants. A number of efflux pumps are able to mediate cross-resistance as a result of their ability to export both QACs and other antimicrobial agents. For example, *qacH*, which has been detected in clinical isolates on an integron on a conjugative plasmid, also encodes resistance to aminoglycosides, chloramphenicol, tetracycline, trimethoprim, beta-lactams, and sulfonamide (46,49).

The present moment is a critical opportunity to retrospectively examine how drastic shifts in antibiotic, antimicrobial, and disinfectant use during the COVID-19 pandemic might have affected AMR. The objective of this study was to apply WBS towards tracking trends in ARG composition in the sewage of two wastewater treatment plants (WWTPs) serving neighboring rural towns with distinct demographics. Shotgun metagenomic sequencing was applied as a means to comprehensively profile resistomes, i.e., ARGs collectively carried across microbial communities (50). SARS-CoV-2 RNA, QAC resistance genes, and microbial taxa were additionally analyzed in the wastewater samples. ARGs were hypothesized to increase due to greater use of disinfectants

and antibiotics during the pandemic. The two WWTPs were compared to examine how resistomes might be impacted differently given that nearly a quarter of the sewage flow of one town emanated from a large university. It was hypothesized that the wastewater resistome would be more greatly impacted by the university and that increases in AMR would be observed due to changes in antibiotic prescription practices and increased disinfectant use in these communities. The findings can help to promote better stewardship of antibiotics and disinfectants as the pandemic continues, and in preparation for future pandemics.

2.3 Materials and Methods

2.3.1 Site Description and Sampling

Samples were collected from two WWTPs located <25 km apart in southwestern Virginia. WWTP1 serves approximately 52,500 people and has an average influent flow of 6.5 MGD. WWTP2 serves approximately 21,500 people and has an average influent flow of 3 MGD. Only WWTP1 receives wastewater from the local university, approximately 24% of the flow, with over 34,000 on-campus and 37,000 off-campus students. WWTP1 also received wastewater from a hospital with 146 beds. Both WWTPs receive treated industrial waste in small volumes (0.4% and 0.05%, respectively). Neither the WWTP1 nor WWTP2 sewer systems are combined with stormwater.

Twenty-four-hour composite influent samples (ISCO 5800) were collected after bar screening each week (7am Thursdays to 7am Fridays), starting Aug 28, 2020 (one week after the semester began) through May 7, 2021 (last week of spring semester classes). All composite samplers were refrigerated at 4°C during collection. One liter was collected from each WWTP in sterile polypropylene bottles and transported to the lab on ice. A subset from each sample was prepared for processing for either RNA or DNA extraction. Samples collected prior to 2020 were grab samples as previously described in Majeed et al. (51).

WWTP operational conditions (flow) and wastewater physiochemical parameters (pH, total suspended solids (TSS), biochemical oxygen demand (BOD)) were determined by the staff at each WWTP. Oil and grease, ammonia, and total Kjeldahl nitrogen (TKN) were measured at WWTP2. Wastewater parameters measured by the Virginia Tech (VT) lab were ammonia and chemical oxygen demand (COD) measured by HACH methods 8000 and 8155 (HACH, Loveland, CO), respectively (52). TSS and volatile suspended solids (VSS) were measured following method 2540 according to the Standard Methods for the Examination of Water and Wastewater (53).

Climate data were downloaded from the National Climatic Data Center's (NCDC) archive available through the National Oceanic and Atmospheric Administration's Climate Data Online (CDO) website (54). Rainfall and temperature daily summaries were included from weather stations USC00440766 for WWTP1 and US1VAMN0005 for WWTP2. The mean weekly rainfall and temperatures were calculated using data from the collection date and the six days prior.

2.3.2 SARS-CoV-2

Sample concentration for RNA extraction followed methods previously described (55,56). Briefly, MgCl₂ was added to a 150-200 mL sample to achieve a final concentration of 25 mM and the pH was adjusted to 3-4. After these adjustments, the samples were filtered through 0.45-µm mixed cellulose ester filters (Thermo Fisher Scientific, Waltham, MA) onto sterilized vacuum filters within a BSL-2 hood. The filters were then rolled and fragmented into 3-5-mm size pieces with flame sterilized tweezers and put into centrifuge tubes. The ripped filters were stored at -80 °C until they were transferred on ice to the Fralin Biomedical Research Institute for further

processing. RNA extraction followed a mechanical procedure to release particles from the filters after which an aliquot of eluate (300 μ L) was subjected to TRIzol/EtOH-based nucleic acid extraction following standard extraction/washing procedures (57). Samples were neutralized in TRIzol supplemented with ethanol, then purified in a column-based system (Zymo Research, Irvine, CA). Lastly, the samples were eluted in water and quantified using a Nanodrop system before being stored at -80 °C.

SARS-CoV-2 genes (*N* and *E*) as well as pepper mild mottle virus (PMMoV) were quantified in RNA extracts by staff at the Hampton Roads Sanitation District (HRSD). Analysis was done using a one-step ddPCR multiplex protocol that included no template controls, field blanks, and Twist Synthetic SARS-CoV-2 RNA Control 4 positive standards (Twist Bioscience, San Francisco, CA, USA), following methods described in Gonzalez et al. (58).

Clinical case data were shared by the Virginia Department of Health (VDH), which reported the number of cases based on zip code (relating to WWTP1 or WWTP2) on a given date as either the date that the case was reported to VDH, the date of illness onset, or the date that a test was conducted. The clinical cases were summarized based on symptom onset date; therefore, the date of illness onset was first used to count the number of cases. If these data were not available, then the test date was used followed by the date reported. The date reported was not used alone as this date might differ from when an individual was actually positive.

2.3.3 Sequencing and Bioinformatics

Sample concentration procedures for DNA extraction followed methods similar to those which were applied for RNA extraction (e.g., filtration). However, neither cation addition nor pH adjustment were made to the samples prior to filtering them through 0.22- μ m filters (Thermo Fisher Scientific, Waltham, MA). Filtration was carried out until the filters clogged and the corresponding volume filtered was recorded (24-180 mL). DNA was extracted using the FastDNA™ SPIN Kit for Soil (MP Biomedicals, Solon, OH). This extraction kit was chosen to ensure comparability with previous samples collected at WWTP2. These prior samples include those collected bi-monthly over the course of 18 months prior to the pandemic (2017-2018) (51), the year prior (2019), and at the start of the pandemic (3/24/2020, 3/27/2020, 4/7/2020) (see Chapter 3). Filters and extracts were stored at -20 °C. DNA extracts were subject to library preparation and sequencing at the Duke University Center for Genomic and Computational Biology. DNA extracts were sequenced using an Illumina NovaSeq 6000 platform using 2 $\times$ 150 paired-end reads targeting approximately 36 Gbp per sample. A total of 80 DNA extracts were sent for short-read sequencing; 34 from WWTP1 and 46 from WWTP2 (which included three sampling date replicates). An additional 8 metagenomic sequencing libraries from WWTP2 were reprocessed, but not resequenced resulting in a total of 88 metagenomic sequencing libraries. Six WWTP2 samples were collected prior to the start of sampling at WWTP1 in August and September, 2020. One WWTP1 sample did not have a corresponding WWTP2 sampling date for direct comparison (4/9/2021).

Short reads from the Illumina platform and samples previously sequenced were processed using fastp (59) to trim all adapters and filter reads. Reads that mapped to cow, masked dog, and human reference genomes were removed using WWTP1duk as suggested by Clum et al. (60). Sequences were annotated against the Comprehensive Antibiotic Resistance Database (CARD v3.1.4) ($\geq 80\%$ identity and $E < 10^{10}$) (61) and using Kraken2 (62) against the Genome Taxonomy Database (GTDB v202 released April 27, 2021) (63,64).

To identify the known hosts of QAC resistance genes the gene sequence for *qacE*, *qacEdelta1*, *qacL*, and *qacH* as reported in CARD was searched using BLASTN (megablast (65)) against National Center for Biotechnology Information's (NCBI) nucleotide collection (non-redundant protein (NR) and nucleotide sequences (NT)) database (searched on December, 20th 2022) (66). Results were filtered and matches with at least 90% identity and 90% coverage were considered hits as previously described (67).

2.3.4 PPCPs and QACs

Grab samples for antibiotic analysis were collected from WWTP2 prior to the start of the pandemic for semi-quantitative non-target screening of pharmaceuticals and personal care products (PPCPs) (7/23/2019, 8/16/2019, 2/4/2020). Samples were also collected approximately monthly during the pandemic for quantitative analysis of targeted compounds (10/5/2020, 12/11/2020, 1/18/2021, 3/1/2021, 3/29/21, 4/2/2021). Samples were collected in glass amber bottles that were acid washed and rinsed with methanol prior to use. Samples were transported to the lab on ice then stored at -20 °C until analysis. Additional influent samples were collected from both WWTP2 and WWTP1 for quantification of quaternary ammonium compounds: octyl decyl dimethyl ammonium chloride, dioctyl dimethyl ammonium chloride, dodecyl dimethyl ammonium chloride, alkyl dimethyl benzyl ammonium chloride (9/2/2020, 9/4/2020, 9/7/2020, 9/8/2020, 9/16/2020, 9/24/2020, 11/29/2020, 12/6/2020). Antibiotics and QACs were analyzed in the laboratory of Dr. Kang Xia at Virginia Tech.

Ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) was used to measure PPCPs and their metabolites via both targeted and non-targeted analysis. Non-target analysis enables screening of over 140 compounds in a single run and allows for semi-quantitative comparison of results (i.e., a comparison of peak areas) following solid phase extraction (SPE) conducted by following EPA method 1694 (68). Non-target analysis limits comparison across time (as peak areas may not be compared across runs), therefore target analysis was chosen to allow better comparability of antibiotics over time, antibiotics were chosen for targeted analysis based on antibiotics previously detected at WWTP2 using non-targeted analysis; targeted analysis will quantify the antibiotics N¹-acetylsulfamethoxazole, sulfamethoxazole, sulfamethoxazole, sulfadiazine, anhydroerythromycin A, azithromycin, clarithromycin, and ciprofloxacin.

2.3.5 Statistics

All statistical tests were conducted in R (v 4.1.3) (69). An α of 0.05 (p-value < 0.05) was used as the level of significance, unless otherwise stated, with Bonferroni-Holm corrections for multiple comparisons (70). Metagenomic samples were found to be non-normally distributed by the Anderson-Darling test (71). Physiochemical parameters were tested for normality using Shapiro-Wilks (72). Normally distributed data were compared using a t-test of log transformed data and non-normally distributed data was compared using Wilcoxon signed-rank test using the rstatix package (73). Spearman's rank correlation coefficients were calculated to determine correlations between clinical cases and SARS-CoV-2 RNA and between QAC resistance genes and bacterial families.

Alpha diversity was determined using the Shannon diversity index, which describes the diversity within a sample after taking proportional abundance into account. Beta diversity, measuring differences between samples, was determined via Bray-Curtis dissimilarity matrices and visualized using non-metric multidimensional scaling (NMDS); significant differences between groups were calculated using Permutational Multivariate Analysis of Variance

(PERMANOVA). All ordination analyses were done using the R package *vegan* (74). The *vegan* function *envfit* was used to calculate regressions and determine significant intrinsic and extrinsic factors; extrinsic factors were considered significant given a $p < 0.001$, the significance of extrinsic factors were selected given a $p \leq 0.001$ as this controlled for a false discovery rate (FDR) of 0.1 (following (75)) using the Benjamini-Hochberg procedure (76).

All figures were generated in R using *ggplot2* (77), *cowplot* (78), and *GGally* (79); color scales were generated using *viridis* (80) and *pals* (81) packages.

2.4 Results

2.4.1 WWTP Physiochemical Parameters

The average WWTP1 flow was twice that of WWTP2 over the duration of the study period. Conversely, all of the physiochemical parameters measured by both WWTPs were higher in WWTP2 and were statistically different from WWTP1 (BOD, pH, and TSS: Wilcox, $p < 0.001$ **Figure 2.1**). Ammonia and COD were found to be similar between the two plants (**Table A.1**).

A subset of WWTP parameters were compared for 2021 and 2022 (March-April) to account for temporal changes such as those resulting from seasonal variability. Water quality measures at WWTP2 were found to be more variable than at WWTP1 during this time period, with greater BOD and TSS (BOD: 273 ± 68 and 217 ± 64 mg/L, TSS: 259 ± 65 and 217 ± 65 mg/L, respectively, Wilcox $p < 0.001$) and lower flow and pH (2.7 ± 0.4 and 3.5 ± 1.0 MGD; pH 7.69 ± 0.08 and 7.86 ± 0.08 , Wilcox $p < 0.0001$). For WWTP1, only TSS was greater in 2022 than 2021 (166 ± 45.3 and 142 ± 41.3 mg/L, respectively, Wilcox $p = 0.021$).

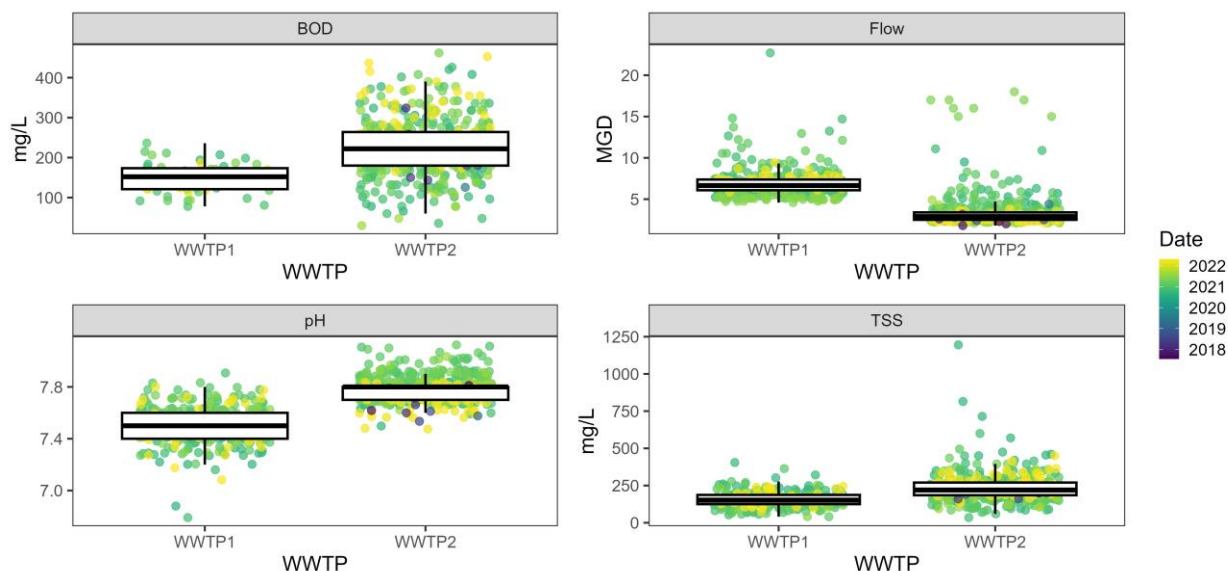


Figure 2.1 BOD, flow rate, pH, and TSS measured at each WWTP, each physiochemical parameter was significantly different between the WWTPs (Wilcox, $p < 0.001$). Outlier high flow events correspond to efforts by staff to address operational issues occurring at the WWTPs.

2.4.2 SARS-CoV-2

The mean and statistical analyses for COVID-19 clinical cases were calculated using the sum of daily clinical cases reported each week (8/9/2020-5/16/2021). Comparisons between

SARS-CoV-2 RNA and clinical cases were conducted using the sum of clinical cases reported on Thursdays and Fridays corresponding to the WWTP composite sample collection days. The number of clinical cases were almost 3× greater in the town served by WWTP1 than the town served by WWTP2 for both 2020 and 2021 (t-test log10 transformed data, $p < 0.01$ and $p < 0.001$, respectively) (**Table A.2**). The number of cases reported in the 2020 vs. the 2021 study time periods were very similar for WWTP1 (41 ± 32 cases/week on average), but more variable for WWTP2 (15 ± 10 cases/week on average, t-test (log10), $p = 0.025$). WWTP1 had two spikes in case numbers in August-September, 2020 (57 ± 40 cases/week on average) and January-February 2021 (69 ± 16 cases/week on average). WWTP2 only had only one spike in cases with fewer cases/week than WWTP1, however, this spike spanned three months: November 2020-January 2021 (28 ± 7 cases/week on average). The spike in cases for WWTP1 corresponded with the university semester start dates since undergraduate classes resumed August 24th, 2020 and again for the spring semester on January 19th, 2021, but were held virtually until January 25th, 2021. Classes were primarily virtual during the fall 2020 semester and were only virtual after the Thanksgiving holiday (beginning November 21, 2020).

SARS-CoV-2 RNA genes were normalized both to the volume of sample filtered (per 100 mL) (**Figure 2.2 A**) and to the concentration of PMMoV (**Figure 2.2 B**). *N* and *E* genes normalized by either approach were similar enough in both WWTPs and across 2020 and 2021 to not be statistically different. When comparing WWTP1 and WWTP2 month-to-month there were some discernible differences from December 2020-March 2021. Genes greater in WWTP2 than WWTP1 were *N* and *E* in December, 2020 and *E*/PMMoV in January, 2021. Genes greater in WWTP1 than WWTP2 were *E* and *E*/PMMoV in February, 2021 and *N* and *E* in March, 2021 (**Table 2.1**). Concentrations of genes were greatest in February for WWTP1 (*E*: 3866 ± 2534 and *N*: 1937 ± 1212 gene copies/100mL; *N*/PMMoV: 0.00081 ± 0.0006 and *E*/PMMoV: 0.0016 ± 0.0013 gene copies/PMMoV). For WWTP2 both *E* and *N* were highest in January (3431 ± 2892 and 1382 ± 1187 gene copies/100mL, respectively) (samples collected in May were not compared as they represented only one date). When normalized to PMMoV *E* was greatest in September (0.0011 ± 0.0018 gene/PMMoV) and *N* greatest in December (0.0008 ± 0.0004 gene/PMMoV) (*E* was also second highest in December 0.0010 ± 0.0008 gene/PMMoV).

The total clinical cases (sum of each reporting type from Thursdays and Fridays) were compared against the SARS-CoV-2 RNA wastewater signal. When normalized by PMMoV both *N* and *E* genes positively correlated with the number of clinical cases. WWTP1 was more strongly correlated than WWTP2, and for both WWTPs, *E* was more strongly correlated to clinical cases than *N* (WWTP1 *E*: $p < 0.0001$, WWTP1 *N*: $p < 0.001$, WWTP2 *E*: $p = 0.03$, WWTP2 *N*: $p = 0.03$) (**Figure 2.3**). The *N* gene correlated with the clinical cases for WWTP2 and WWTP1 when normalized by volume, *E*/100mL was only correlated with clinical cases for WWTP1 (**Figure A.1**). However, these correlations were not as strong for genes normalized by PMMoV.

Table 2.1 Concentrations of SARS-CoV-2 gene copy (gc) numbers that were significantly different between WWTP1 and WWTP2 (Wilcox test, $p < 0.05$), averaged by month

Month	Gene	WWTP1 (gc/100mL or gc/gc) (mean $\pm$ SD)	WWTP2 (gc/100mL or gc/gc) (mean $\pm$ SD)
Dec, 2020	<i>N</i>	397 $\pm$ 181	1071 $\pm$ 344
	<i>E</i>	464 $\pm$ 198	1230 $\pm$ 425
Jan, 2021	<i>E</i> /PMMoV	0.00039 $\pm$ 0.00014	0.00065 $\pm$ 0.00013
Feb, 2021	<i>E</i>	3866 $\pm$ 2534	639 $\pm$ 373
	<i>E</i> /PMMoV	0.0016 $\pm$ 0.0013	0.0003 $\pm$ 0.0001
Mar, 2021	<i>N</i>	852 $\pm$ 177	572 $\pm$ 79;
	<i>E</i>	1516 $\pm$ 379	573 $\pm$ 330

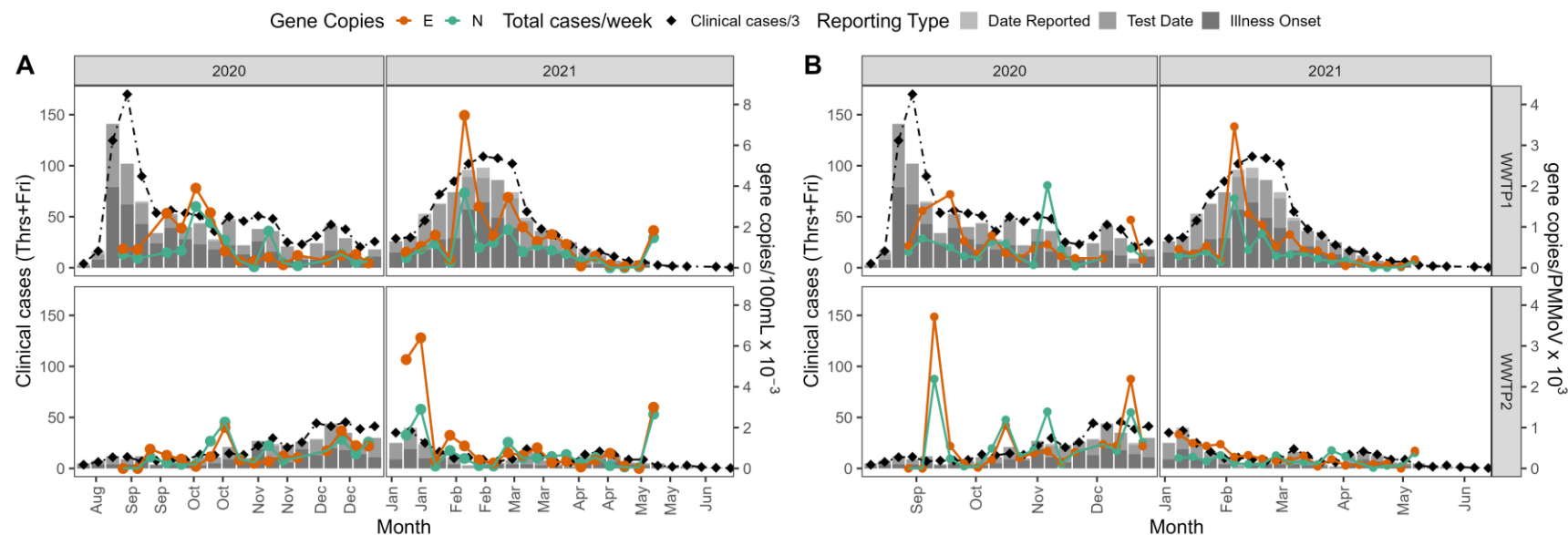


Figure 2.2 SARS-CoV-2 measurements (*E* and *N* genes) for WWTP1 and WWTP2 influents overlaid over case data obtained from the Virginia Department of Health (VDH). Reporting type represents the weekly summary of daily cases based on dates selected based on the illness onset date (if available). If illness onset date was not available, then the test date was selected. Finally, the date the case was reported to VDH was used if no other date was available. Total cases/week were scaled by dividing by three.

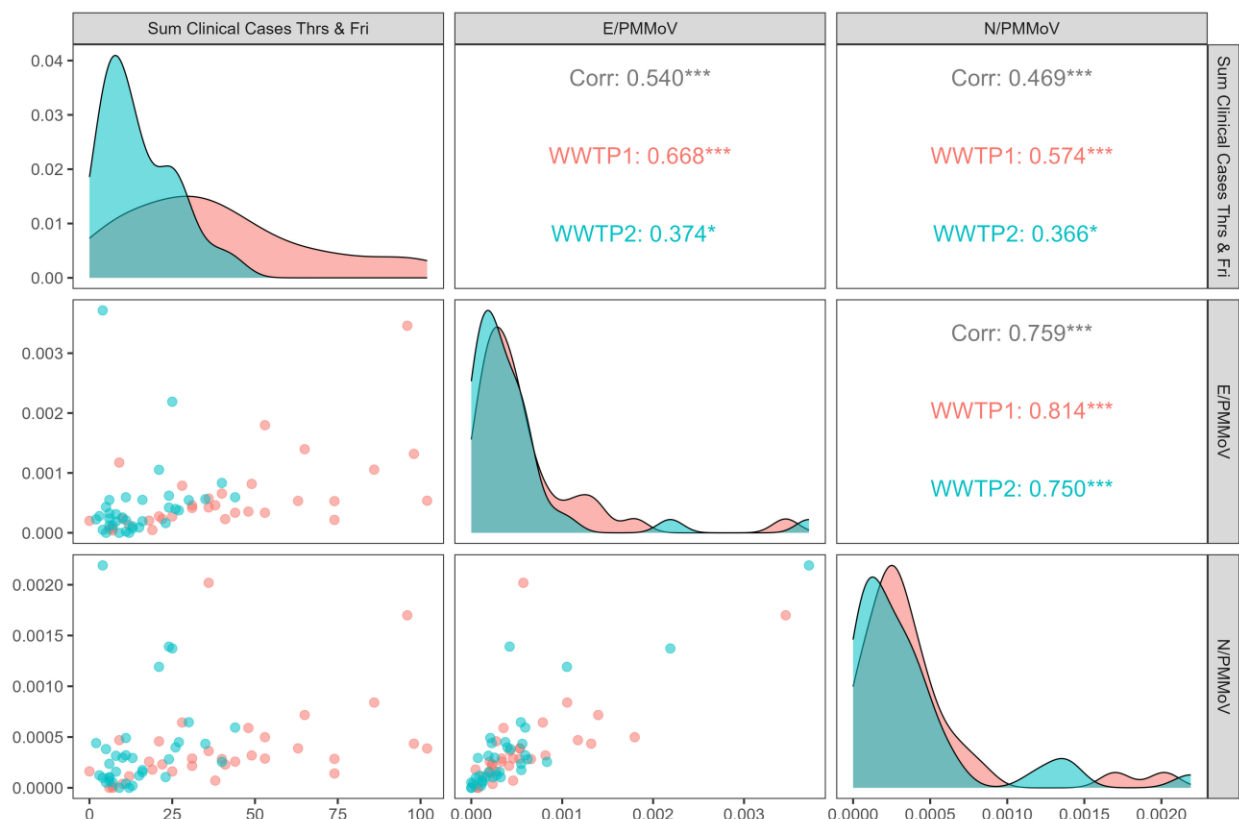


Figure 2.3 Histograms (displayed diagonally) for the clinical cases of COVID-19 within each WWTP's sewershed (top row), or concentrations of *E*/PMMoV (middle row) and *N*/PMMoV (bottom row). The bottom left panels provide scatter plots of these data and the top right panels are the corresponding correlations (Spearman). Corr indicates the Spearman's Rho for both WWTPs together, below which are the correlations for WWTP1 and WWTP2 when compared separately. The units are total counts of clinical COVID cases or *N* or *E* gene copies per PMMoV. The graph was generated in R generated using GGally.

2.4.3 QAC Resistance Genes

QAC resistance genes *qacE*, *qacEdelta1*, *qacL*, and *qacH* were detected in both WWTP2 and WWTP1 samples. *qacL* (on average 0.015 ± 0.005 ARGs/16S rRNA) had the greatest relative abundance in both WWTPs, followed by *qacEdelta1* (0.004 ± 0.001), *qacE* (0.0010 ± 0.0004), and *qacH* (0.00002 ± 0.00001). The relative abundances of these genes varied over time with peaks at different time points for both WWTPs (**Figure 2.4 A**). *qacE* was greater in WWTP2 than WWTP1 in 2021 (Wilcox, $p=0.007$), *qacEdelta1* was greater in WWTP2 than WWTP1 in both 2020 and 2021 (Wilcox, $p=0.0016$ and $p=0.026$, respectively), and *qacL* was greater in WWTP1 than WWTP2 in 2020 and 2021 (Wilcox, $p=0.001$ and $p=0.007$ respectively). For WWTP1, *qacE* and *qacEdelta1* decreased from 2020 to 2021 (Wilcox, Holm adjusted $p=0.012$ and $p=0.03$, respectively), whereas *qacL* increased from 2021 to 2022 (Wilcox, Holm adjusted, $p=0.012$). For WWTP2, *qacL* increased from 2020 to 2021 (Wilcox, Holm adjusted $p=0.002$) and *qacEdelta1* decreased from 2020 to 2021 (Wilcox, Holm adjusted $p<0.0001$). In WWTP1, the total QAC resistance genes were greater prior to an increase in clinical cases, however, the opposite trend was observed for WWTP2 (**Figure 2.4 B**).

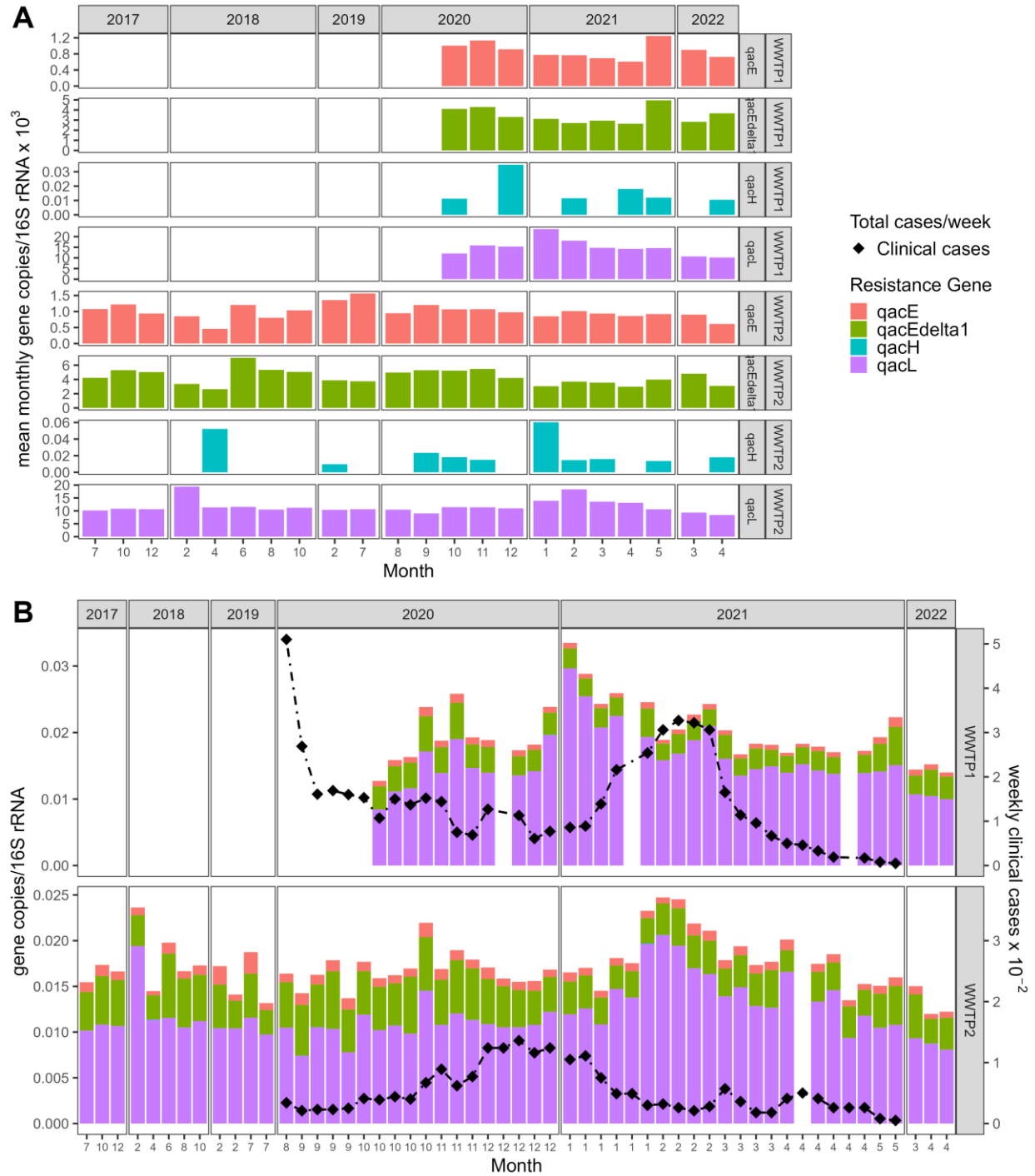


Figure 2.4 (A) Monthly mean of QAC resistance genes detected at both WWTPs over time and (B) QAC resistance genes with the sum of clinical cases overlaid. Missing bars in WWTP1 2020-2021 are due to replicate samples from WWTP2. There was one missing sample in WWTP2 that did not have a comparable date to WWTP1. Samples were not collected at WWTP1 before 2020.

Bacterial families previously shown to harbor QAC resistance genes in NCBI were compared with the detected QAC resistance genes (*qacE*, *qacEdelta1*, *qacL*, and *qacH*) via linear correlation. The QAC resistance gene *qacE* was associated with the most families followed by *qacEdelta1* (**Table A.3**); *qacEdelta1* was associated with the same taxa as *qacE* with the exception of Pasteurellaceae, which was not associated with *qacE*. *qacL* was hosted by many of the same taxa as *qacE* and *qacEdelta1* (Comamonadaceae, Aeromonadaceae, Enterobacteriaceae, Moraxellaceae, and Xanthomonadaceae). *qacH* was only known to be hosted by Staphylococcaceae and Enterococcaceae. The families that were best correlated varied for different QAC resistance genes and between the WWTPs.

qacL was positively correlated with a number of taxa (Burkholderiaceae, Rhizobiaceae, Microbacteriaceae, Micrococcaceae) in both WWTPs ($p < 0.05$). In WWTP1, *qacL* was most strongly correlated with Burkholderiaceae and Rhizobiaceae ($p < 0.0001$). In WWTP2, *qacL* was correlated with these same taxa, however with a lower Rho (**Figure A.2**). *qacE* and *qacEdelta1* were correlated with similar taxa for both WWTP1 and WWTP2 (**Table 2.2**). *qacH* was only significantly correlated with Rhizobiaceae in WWTP1.

The relative abundance of *qacL* was graphed over time for relationships that showed the strongest correlations (Burkholderiaceae, Rhizobiaceae, and Microbacteriaceae) (**Figure 2.5**) (additional comparison **Figure A.3**). *qacE* and *qacEdelta1* genes were positively correlated with each other in WWTP1 and WWTP2 ($p < 0.0001$); however, the taxa that were correlated with these genes were not always the same for both WWTPs. In WWTP1, Pseudomonadaceae and Staphylococcaceae were significantly correlated with *qacEdelta1*, but not with *qacE*. whereas, in WWTP2 Enterobacteriaceae and Pasteurellaceae correlated with *qacEdelta1*, but not with *qacE*.

Table 2.2 Bacterial families that significantly and positively correlated with specific QAC resistance genes (Rho ≥ 0.35)

WWTP	Gene	Family	Spearman's Rho
WWTP1	<i>qacE</i>	Pasteurellaceae, Rhodocyclaceae, Vibronaceae, Rhizobiaceae_A, Enterobacteriaceae	0.52****, 0.50**, 0.48**, 0.48**, 0.44*
	<i>qacEdelta1</i>	Rhizobiaceae_A, Vibronaceae, Pasteurellaceae, Rhodocyclaceae, Enterobacteriaceae, Staphylococcaceae, Pseudomonadaceae	0.66****, 0.65***, 0.64****, 0.62**, 0.59**, 0.43**, 0.34*,
	<i>qacL</i>	Burkholderiaceae, Rhizobiaceae, Microbacteriaceae, Xanthomonadaceae, Micrococcaceae, Actinomycetaceae, Mycobacteriaceae	0.82****, 0.73****, 0.60***, 0.58***, 0.55***, 0.54**, 0.45**
	<i>qacH</i>	Rhizobiaceae	0.79*
WWTP2	<i>qacE</i>	Vibronaceae, Shewanellaceae, Rhodocyclaceae	0.61****, 0.38**, 0.32*
	<i>qacEdelta1</i>	Vibronaceae, Enterobacteriaceae, Rhodocyclaceae, Shewanellaceae, Pasteurellaceae, Aeromonadaceae	0.77****, 0.70****, 0.63****, 0.60****, 0.60****, 0.32*
	<i>qacL</i>	Burkholderiaceae, Rhizobiaceae, Microbacteriaceae, Xanthomonadaceae, Pseudomonadaceae, Micrococcaceae	0.58****, 0.55**, 0.55**, 0.44**, 0.35**, 0.33*
* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$			

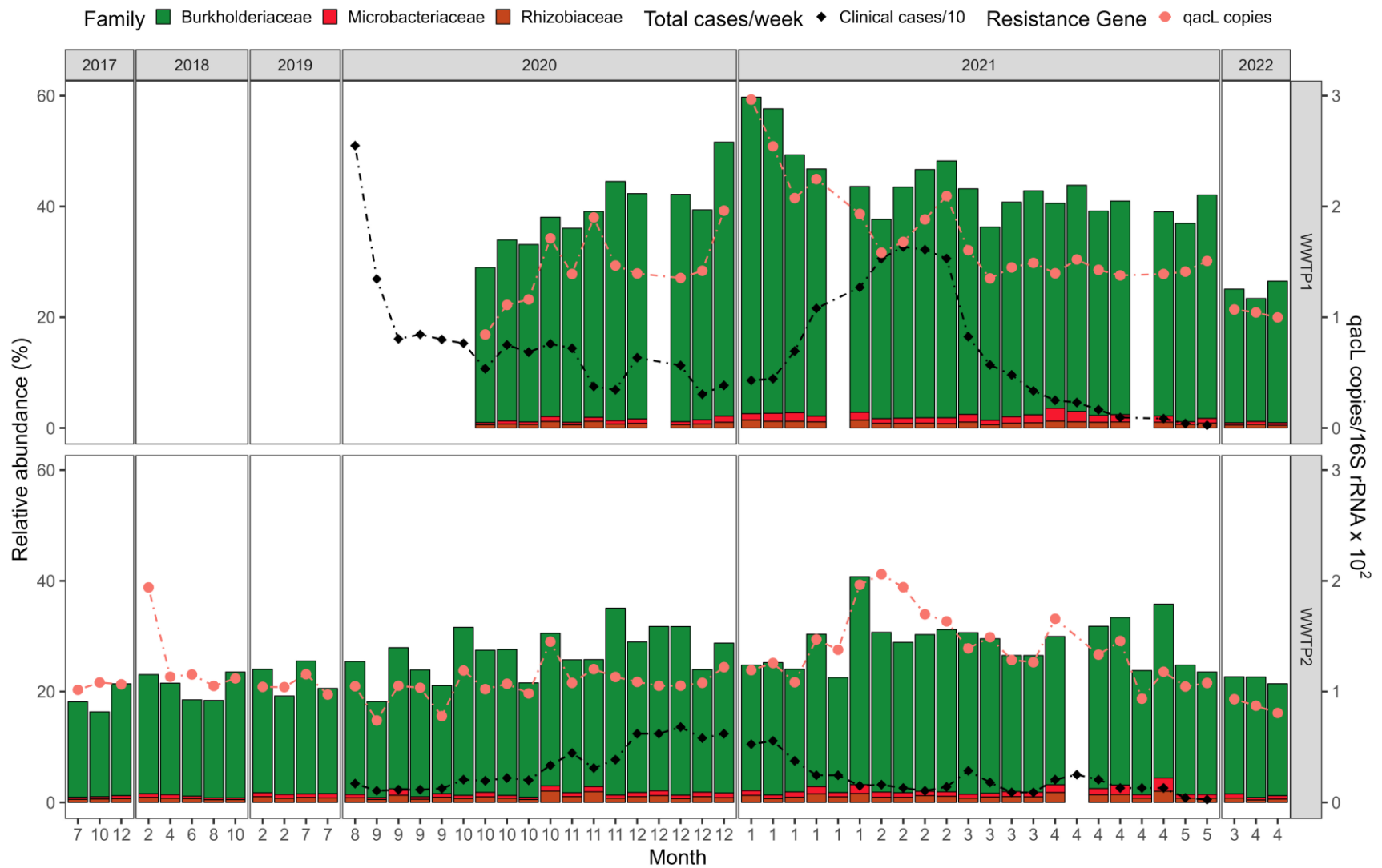


Figure 2.5 The three families most correlated with *qacL* were plotted together to visually show how they were correlated or changed over time. The clinical cases are the weekly sum which were scaled down by 10 $\times$.

2.4.4 Resistome Comparison

The most abundant ARGs for both WWTPs were: *msrE* (MLS resistance), *mphE* (MLS resistance), *rsmA* (multidrug resistance), *qacL* (other/disinfecting agents and antiseptics), and *tet(39)* (tetracycline resistance). These ARGs were also the most abundant in the shared resistome (**Figure A.4 A**). Additional QAC resistance genes, within the most abundant ARGs, were *qacEdelta1* (19th most abundant in WWTP2 20th in WWTP1) and *qacE* (79th most abundant in WWTP1 and WWTP2 84th in WWTP2) (**Table A.4**). Core resistance genes were similar across both WWTPs, with the top five most abundant resistance classes being multidrug, MLS, beta-lactam, tetracycline, and aminoglycoside. Mean relative abundances were similar between WWTPs, however, multidrug was slightly higher in WWTP2 (WWTP1: 0.104 ± 0.013 and WWTP2: 0.116 ± 0.019 ARG/16S rRNA, Wilcox $p < 0.00001$) and MLS slightly higher in WWTP1 (MLS WWTP1: 0.104 ± 0.017 and WWTP2: 0.084 ± 0.019 ARG/16S rRNA, Wilcox $p = 0.015$). The total relative ARG abundances were similar between WWTPs; however, samples sequenced prior to 2019 indicated a higher total abundance than 2020-2022 samples (0.50 ± 0.06 and 0.37 ± 0.02 , respectively) (**Figure A.4 B**). The differences in total abundance could potentially reflect differences in sequencing depth, as the samples sequenced prior to 2019 had a much lower sequencing depth (seven Gbp per sample as opposed to 36 Gbp per sample).

Differences between WWTP resistomes (beta diversity) were determined using Permutational Multivariate Analysis of Variance (PERMANOVA) and Bray-Curtis distances were visualized in non-metric multidimensional scaling (NMDS) plots. WWTP1 and WWTP2 groupings were distinct from each other (PERMANOVA: R^2 0.28, p -value 0.001). Interestingly, however, WWTP1 samples collected from three separate dates were more similar to WWTP2 samples than to other WWTP1 samples (10/9/2020, 4/1/2022, and 4/15/2022, notably the third 2022 sample was between WWTP1 and WWTP2 groups). The year sampled played a small but significant role in differentiating both WWTP1 and WWTP2 when compared separately (PERMANOVA $R^2=0.10$ $p=0.008$ and $R^2=0.18$ $p=0.001$, respectively)

The Shannon (alpha) diversity of ARGs was different between WWTP1 and WWTP2 (Wilcox, $p < 0.0001$). WWTP2 samples had greater diversity than WWTP1 (4.84 ± 0.08 and 4.56 ± 0.13 , respectively). Similar to the beta-diversity results, the alpha diversity for WWTP1 samples collected in 2022 were more similar (i.e., higher) to WWTP2 (**Figure 2.6**). Alpha diversity increased in WWTP2 from 2018-2020 ($p=0.011$) and decreased from 2020-2021 (Wilcox, Holm adjusted $p=0.042$) and for WWTP1 increased from 2021-2022 (Wilcox, Holm adjusted $p=0.003$). The extrinsic factors that best explained the differences in the WWTP2 and WWTP1 resistomes were characteristics of the WWTPs (flow) and wastewater physiochemical parameters (pH, TSS, and BOD) (**Figure A.5**). Additional factors that contributed to the differences in samples were year and the maximum and minimum temperatures (**Table A.5**). Furthermore, the sum of clinical cases played a small, but significant role ($R^2=0.15$, $p=0.022$)

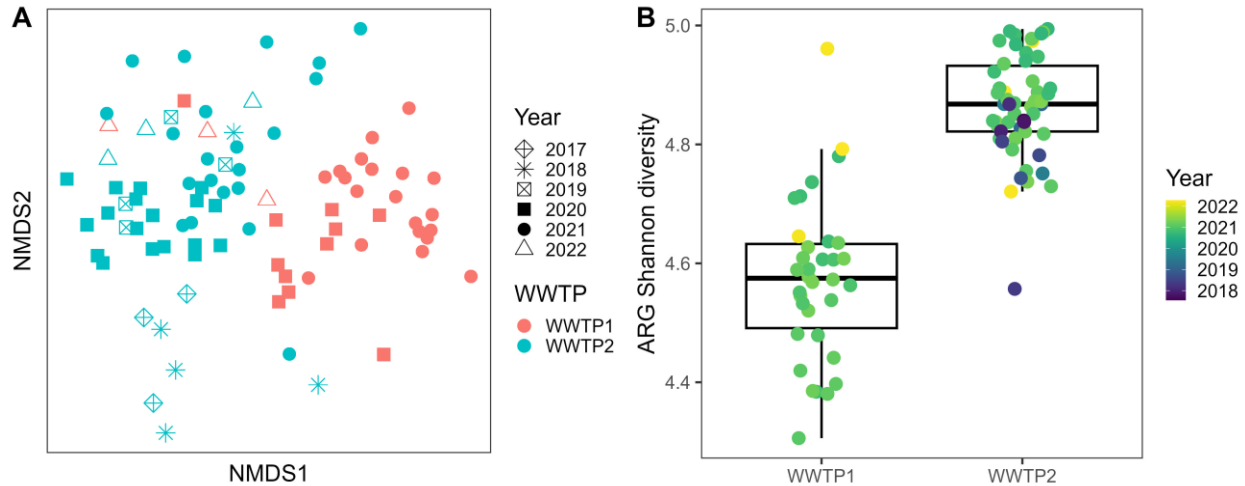


Figure 2.6 (A) Beta diversity of ARGs via NMDS (stress = 0.1432995) clustered by WWTP (PERMANOVA: R^2 0.28, p-value 0.001) and (B) ARG alpha diversity (Shannon) which were different for WWTP1 and WWTP2 influent (Wilcox, $p < 0.0001$).

Both WWTPs shared 76% ($n=1409$) of the same ARGs across all sampling dates. WWTP2 had a greater number of discriminatory ARGs than WWTP1 (313 and 131, respectively) across all sampling dates. The unique ARGs from both WWTPs most commonly contained genes conferring resistance to beta-lactam and aminoglycoside antibiotics (these occurred at the greatest relative abundances and had the greatest number of occurrences (counts) across all dates) (**Figure A.6**). WWTP1 and WWTP2 samples collected in 2022 shared a similar percentage of core ARGs (74%, $n=900$). WWTP1 samples collected in 2022 only shared 65% ($n=994$) of genes with WWTP1 samples collected from 2020-2021, highlighting that WWTP1 in 2022 was more similar to WWTP2 than to itself in the prior two years (**Figure 2.7 A**) Thirty-five genes were unique to WWTP1 in 2022 (24 of those genes were also detected in WWTP2 prior to 2022). Most of the ARGs unique to WWTP1 in 2022 conferred resistance to beta-lactams; however, genes conferring resistance to aminoglycoside, MLS, and tetracycline were also detected. Genes detected in both WWTP1 and WWTP2 in 2022 (also detected prior to 2022 for WWTP2, but not for WWTP1) frequently contained beta-lactam ARGs (five unique ARGs detected across WWTP1 and WWTP2 in 2022); however, ARGs conferring resistance to multiple antibiotics and trimethoprim were also detected (*otrC*: multidrug resistance, detected 2 $\times$ in WWTP1 and 21 $\times$ in WWTP2; *drfA22* and *drfB1*: trimethoprim resistance, detected 12 $\times$ in WWTP2, 2 $\times$ in WWTP1) (**Figure 2.7 B**).



Figure 2.7 ARGs unique to WWTP1 shown as (A) Venn diagrams of ARGs unique to WWTP2 and WWTP1 grouped by different years (B) ARGs that appeared in WWTP1 in 2022 but not in WWTP2 in 2022 and ARGs that were detected in WWTP1 and WWTP2 in 2022 but not in WWTP1 in 2020-2021. Relative abundance is represented by the darkness/lightness of the color which is based on the mean ARG relative abundance from each year, the number in each tile represents the number of times that gene was detected in a given year.

Clinically-relevant ARGs conferring resistance to beta-lactam, peptide, glycopeptide, and quinolone antibiotics were detected in both WWTPs. Across all sampling dates, the relative abundances of these ARGs were greater for WWTP2 relative to WWTP1 for all resistance classes (Wilcox, $p < 0.05$). Glycopeptide abundances were most similar between the WWTPs, however, they also had the least number of detections (WWTP1: 0.0004 ± 0.0002 , detected $8\times$ and WWTP2: 0.0006 ± 0.0003 ARGs/16S rRNA, detected $17\times$). Beta-lactam ARGs had similar relative abundances between WWTP1 and WWTP2 (0.025 ± 0.005 and 0.030 ± 0.008 ARGs/16S rRNA, respectively); quinolone abundances were most different between WWTP1 and WWTP2 (0.0005 ± 0.0003 and 0.0019 ± 0.0008 ARGs/16S rRNA, respectively) (**Figure A.7**). Clinically relevant ARGs were different between WWTP1 and WWTP2 when grouped by 2020 and 2021 (Wilcox, $p < 0.05$ for all but glycopeptide which did not have enough detections to make a comparison). Glycopeptide resistance genes were not detected in WWTP1 in 2020 or 2022. The mean relative abundances for the different resistance classes were greater in WWTP2 than WWTP1 for 2020 and 2021; however, in 2022 beta-lactam and peptide ARGs were more abundant in WWTP1 than in WWTP2 (beta-lactam: 0.024 ± 0.001 and 0.023 ± 0.002 (3% greater) and peptide: 0.0008 ± 0.0001 and 0.0006 ± 0.0001 (20% greater) ARG/16S rRNA, respectively). Clinically relevant quinolone resistance genes were greater in WWTP1 in 2020 than in 2021 (Wilcox, Holm adjusted $p = 0.045$) and even higher in 2022 (though this was not statistically significant, Wilcox, Holm adjusted $p = 0.3$). Clinically-relevant beta-lactam ARGs in WWTP2 decreased over the years, with the lowest abundance in 2022 compared with 2020 and 2021 (Wilcox, Holm adjusted $p < 0.001$).

2.4.5 Intrinsic and Extrinsic Factors

Taxa at the family level showed similar trends to ARGs with distinct clustering between WWTPs (PERMANOVA, $R^2 = 0.31$, $p = 0.001$). The trend evident for ARGs was also observed for taxa with 2022 WWTP1 samples clustering more closely with WWTP2. Alpha diversity was greater in WWTP2 than WWTP1 (3.40 ± 0.21 and 2.84 ± 0.23 , respectively; Wilcox, $p < 0.0001$). The alpha diversity for WWTP1 2022 samples was higher and closer to WWTP2 samples (**Figure A.8**). Differences were less evident by year (PERMANOVA $R^2 = 0.12$, $p = 0.001$), though year was a better explanation for differences between samples for WWTP2 than WWTP1 ($R^2 = 0.17$ $p = 0.001$ and $R^2 = 0.08$, $p = 0.058$, respectively)

Extrinsic and intrinsic factors were assessed using multivariate regression analyses to elucidate factors contributing to sample dissimilarity. The extrinsic factors that best explained the distribution of points in the family NMDS were WWTP characteristics (flow) or the physiochemical parameters (pH, TSS, and BOD). Temperature, date (year and month), as well as the clinical COVID-19 cases and concentration of *N*/100 mL were also contributing factors (**Figure 2.8 A**). When comparing each WWTP separately, many of the same extrinsic factors explained the sample spread; for WWTP2 the greatest impacts were driven by temperature whereas for WWTP1 the factor with the greatest correlation was pH (**Table 2.3**).

Intrinsic factors that best explain the differences between samples were first filtered to remove rare taxa (family $\leq 0.5\%$ abundance). This processing resulted in 27 taxa for WWTP2 and WWTP1 (**Figure 2.8 B**). When examining the plants separately, there were 28 and 22 significant intrinsic factors for WWTP2 and WWTP1, respectively. The significant families varied slightly between WWTP2 and WWTP1 over time, however, most of the taxa remained the same. Intrinsic factors were graphed if they had a $p < 0.001$ (controlled for a false discovery rate of 0.1). Bacterial family abundances in 2022 varied from 2020-2021; in WWTP1 Aeromonadaceae were $2.7\times$ greater in 2022 and Microtrichaceae went from being undetected to having similar relative

abundances as WWTP2 in 2022. Burkholderiaceae were greater in WWTP1 than WWTP2 across all years, however, they made up a smaller proportion in 2022 (from 1.5× greater in 2020-2021 to 1.1× in 2022). Arcobacteraceae and Bacteroidaceae were greater in WWTP2 across all years (67 and 57% greater, respectively); however, in 2022 the percentage in WWTP1 was more similar to WWTP2 (18 and 29% greater, respectively) (**Figure 2.8 C**).

Table 2.3 Extrinsic factors influencing the spread of taxa and their associated R^2 and p-value.

Extrinsic Factor	WWTP1 & 2		WWTP2		WWTP1	
	R^2	p-value	R^2	p-value	R^2	p-value
Flow	0.339	0.001			0.290	0.048
pH	0.293	0.001	0.353	0.001	0.626	0.002
TSS	0.226	0.003			0.487	0.009
Daily minimum temperature	0.172	0.007	0.329	0.003		
Year	0.167	0.01	0.401	0.001	0.537	0.006
Clinical cases/week	0.148	0.016				
Weekly minimum temp	0.144	0.02	0.406	0.001	0.349	0.031
BOD	0.150	0.02			0.326	0.046
Month	0.128	0.026	0.303	0.004	0.522	0.009
N copies /100mL	0.124	0.041			0.450	0.009
Daily maximum temperature	0.113	0.05	0.251	0.007	0.406	0.017
Weekly maximum temperature			0.406	0.001	0.492	0.009

2.5 Discussion

2.5.1 Pandemic/University Impact

The two WWTPs examined in this study had distinct physiochemical parameters. WWTP1 had greater flow, while WWTP2 had greater TSS, BOD, and pH. Samples from both WWTPs were collected as 24 hour composites after bar screening. Accordingly, the differences between samples are expected to be due to the community and underlying sewer systems rather than operational parameters at either WWTP. Sewer systems can be thought of as physio-biochemical reactors since processes occurring in the sewer such as dilution, deposition of solids, and biochemical transformations impact the concentrations and nature of the physiochemical properties of the wastewater (82). Similarly, sewer systems are not just conveyance systems for microbes, but may be viewed as a habitat for microbial propagation (83). The composition of sewer microbial communities is more closely related to human fecal microbes (~35%) at the neighborhood level, however, once at the WWTP only ~15% of the microbial community reflects human fecal microbial communities (83–85). Sewer microbial communities are influenced by the sewer design and may shift due to factors such as the availability of growth substrates, the presence of terminal electron acceptors, and the fraction of active biomass (82,86). During transport, environmental microbes, introduced through rainwater or stormwater infiltration, change the community composition (82). Seasonal variability (e.g., temperature) and dilution by rainfall have been shown to impact wastewater characteristics as well as the decay kinetics of some sewer system pathogens (87).

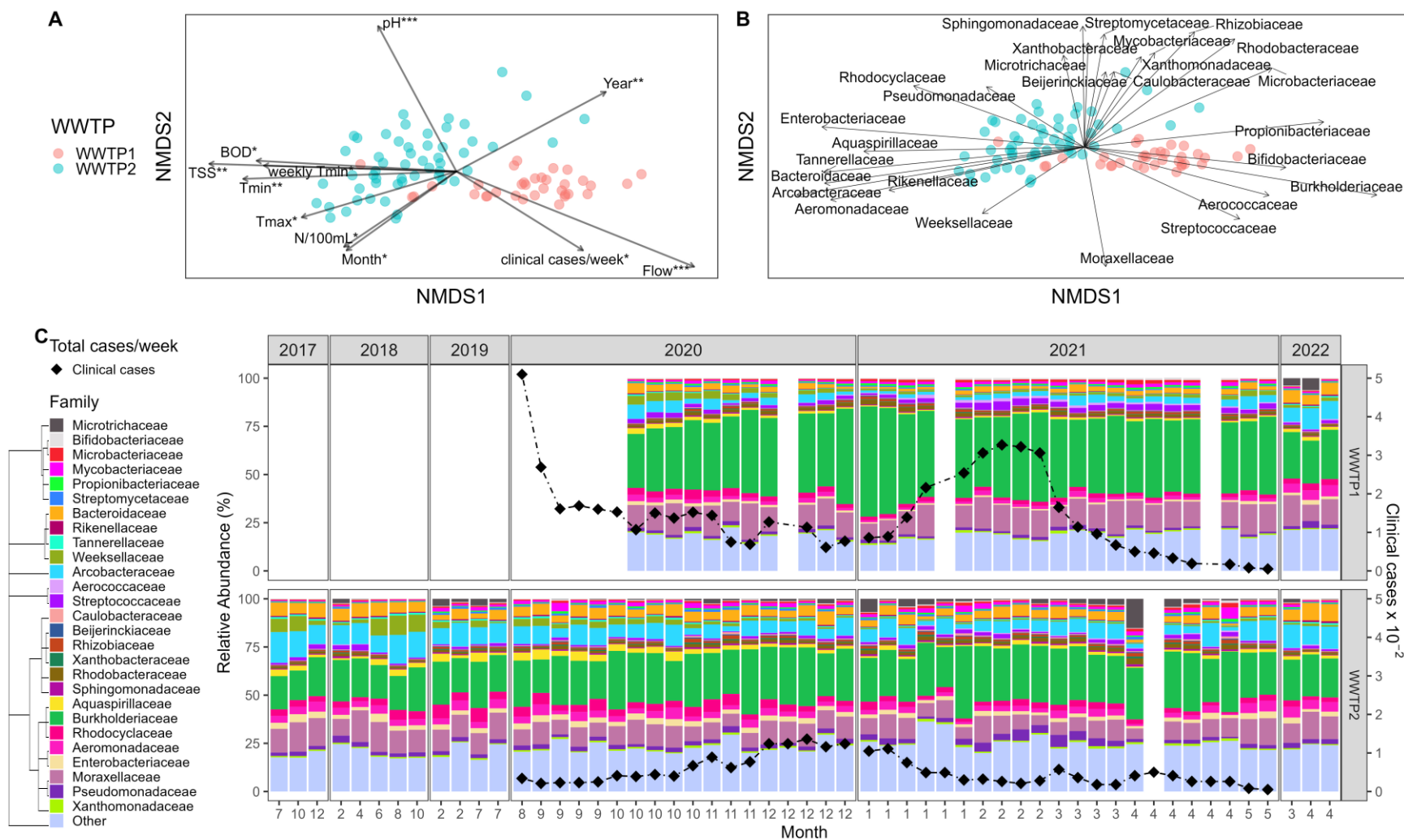


Figure 2.8 (A) Extrinsic factors impacting spread of points, graphed factors with $p < 0.05$, stars indicate $*p \leq 0.05$, $**p \leq 0.01$, and $***p \leq 0.001$. (B) Taxa with $p = 0.001$ which accounts for significance at FDR of 0.1. (C) Relative abundances of the taxa that were determined to be impacting the differences between samples (all other taxa in group “other”, not including unclassified taxa; mean unclassified in WWTP1 $26.3 \pm 4.41\%$ and WWTP2 $31.3 \pm 5.70\%$).

The two WWTPs had distinct resistomes and microbiomes when examined via metagenomic sequencing. While remaining distinct from each other, both WWTP1 and WWTP2 showed variability over time with regards to physiochemical parameters and year-to-year diversity. Some of the microbial variability may reflect seasonal changes since microbial communities have been shown to vary based on season/ambient temperature (85) as well as pH (88). Interestingly, the WWTP1 samples were more similar to WWTP2 samples in 2022. There were no samples from WWTP1 collected prior to the start of the pandemic, however, it is possible that the dissimilarity between 2020-2021 and the 2022 samples reflects changes within the local community and a return to a “pre-pandemic normal.” It is reasonable to expect that the WWTP1 and WWTP2 influents would have similar microbial communities as the two systems are geographically close and are influenced by highly similar seasonality and weather patterns. Further, while the two towns are separate entities, there is overlap between them with respect to work location, residence, recreation, and commerce and therefore it is expected that people contribute to one or both WWTPs. VandeWalle and colleagues’ (84) previously found that two separate WWTPs serving Milwaukee, WI had similar microbial communities even though the systems were physically separate and our result is thus consistent with their prior study. The shift in the microbial community in 2022 may be reflect the later days of the pandemic and changes in behavior and regulations due to alterations in COVID-19 safety precautions.

Water use changes leading to differences in flow may have occurred during the pandemic potentially shifting the sewer system microbial communities. A study examining wastewater flow in an urban metropolitan city in Texas found that social distancing policies put in place in 2020 changed the peak wastewater flow from what had been observed in 2019. The peak flow in 2020 was greater in residential areas and lower in non-residential (e.g., business) areas presumably since more individuals were working from home and/or increased their hygiene practices in response to the pandemic. Reopening saw another shift in the peak flow somewhat returning back to pre-pandemic conditions, however, the residential peak flow remained higher than before likely due to people continuing to work from home (89). Per university policy, many university students and employees worked remotely in 2020-2021 which might have led to them contribute to the sewer differently or to even contribute to a different sewer system. The university contributes approximately 24% of the total flow to WWTP1, therefore, while the total flow was not different for months sampled in 2022, changes in water use at the university could have impacted on the wastewater quality and microbial communities.

One of the objectives of this study was to determine if the pandemic negatively impacted AMR given concerns over potential increases in antibiotic prescriptions and decreasing stewardship practices due to concerns over co-infections and shifts in resources during the pandemic. Conversely, it was considered that the pandemic might act as a preventative measure towards AMR due to behavioral changes such as social distancing, telemedicine and more. The latter hypothesis may be true as it was found that a majority of the mean monthly outpatient prescriptions for common antibiotics in the US were lower in 2020 than expected, this was speculated to be due to changes in behavior such as masking and other preventative measures that reduced infections leading to fewer antibiotic prescriptions (90). Early in the pandemic (January-May 2020) it was found that antibiotic outpatient prescriptions decreased in the US, especially for agents typically prescribed for non-COVID-19 respiratory infections (e.g. common cold, flu) (91). The general public’s behaviors regarding medical care may have shifted during the pandemic, the US Centers for Disease Control estimated that approximately 40% of adults delayed medical care

from April-July 2020, this number decreased to ~33% for the latter half of 2020 and at the conclusion of the survey period was ~15% in July, 2021 (92). Because of the large proportion of the population in WWTP1 being university students, changes in their behavior may have had a greater impact on the community than was typical for other areas. Antibiotics typically prescribed to young adults might also have shifted during the pandemic, for example, from 2010-2018 influenza associated antibiotic prescribing (e.g., misuse) was highest in a cohort study for adults aged 18-19 (compared with adults aged 18-59) (93), therefore, if prescriptions relating to respiratory illnesses decreased overall in 2020, then perhaps the university students had a larger influence on antibiotic use in WWTP1. There were a greater number of unique resistance genes in WWTP1 2022, most often related to beta-lactam resistance, therefore, it may be that more university students were prescribed beta-lactams during 2022, though the relative abundance of clinically relevant beta-lactam ARGs was similar to years prior. Beta-lactam prescriptions in physician offices declined from 2009-2016, however, they still remained among the top three most prescribed oral antibiotics along with macrolides and fluoroquinolones (94) and were higher in 2021 than 2020 (25,26).

2.5.2 QAC Resistance Trends

It was hypothesized that abundances of QAC resistance genes detected in wastewater would change due to increased usage of disinfectants containing QAC during the pandemic. However, the data hinted that the trends (i.e., increases and decreases) seen in the wastewater were more strongly correlated with bacterial families such as those that likely house these QAC resistance genes. When compared with clinical COVID-19 cases (as a proxy for the pandemic) there were no clear correlations between infections and QAC resistance gene concentrations across both WWTPs. However, as has been established, the two WWTPs were distinct, therefore, it may be that the increases in QAC resistance genes seen after a spike in cases in WWTP1 and before a spike in WWTP2 are indicative of different community behaviors regarding disinfectant use. While the linear correlations between QAC resistance genes and their potential bacterial hosts cannot prove causation, they provide an idea as to why QAC resistance may have changed over time if not due to the pandemic. These findings also highlight the need to not only examine ARGs in an environment, but also the importance of examining the microbial taxa that may be harboring these genes leading to shifts over time or differences between environments.

2.5.3 Wastewater Based Surveillance

WBS has been used extensively to examine trends of infections in the community, especially SARS-CoV-2 during the pandemic. Studies have found that SARS-CoV-2 RNA in wastewater can be predictive of the trends in the community from two days to three weeks (11,12,95). This is especially important when clinical testing is not available and/or practical. This study did not find a similar predictive power for wastewater; however, the SARS-CoV-2 RNA concentrations in wastewater were positively correlated with clinical cases as reported by VDH when broken down by zip code. The trends in SARS-CoV-2 RNA in WWTP1 wastewater were more similarly correlated with clinical cases; this may be indicative that the population contributing to WWTP1 is more similar to the people that reside in the corresponding zip code for that town. WWTP2 was more variable, this could be due to differences in the sewer systems, populations contributing unequally to the WWTP/transient populations, or perhaps potential discrepancies with zip-code delineations and the sewershed.

Increases in clinical cases in WWTP1 occurred at the start of the fall and spring semesters, highlighting how the community is influenced by the large student population as the trends in clinical cases coincided more closely with their return than trends seen cross the nation. On the other hand, WWTP2 saw the greatest increase in cases in November 2020-January 2021 which corresponded with a similar observed increase in nationwide cases indicating that WWTP2 may be more indicative of trends seen across the US (96). WWTP2 is a rural community, therefore, the limited impact of the pandemic on the wastewater resistome/microbiome may be indicative of differences in responses to the pandemic seen in different communities, for example, mask use has been observed to be lower in rural populations than urban in the US (97). While clinical cases were used in this study as a proxy for the pandemic severity and a marker of potential behavioral changes in the community, there may be other markers that are more influential to the wastewater resistome and microbiome such as changes in water usage, antibiotic prescriptions, and transient populations/who is contributing to the WWTP.

2.5.4 Next Steps

Additional wastewater samples collected from WWTP1 and WWTP2 in 2020 and 2022 may allow for a better understanding of differences between WWTPs and the shifts that were seen in WWTP1 in 2022. These samples will be analyzed for antibiotic and QAC concentrations to be able to better determine if antibiotic and/or QAC concentrations changed during the course of the pandemic and how these concentrations differed between WWTPs. It is expected that the concentrations in wastewater would mirror the trends seen for both communities with concentrations greater in WWTP2 in 2020 and more similar to WWTP1 in 2022. If this trend is not observed, it would infer that there remain additional unexplained parameters as to why the WWTP1 microbiome and resistome shifted in 2022. Data have been requested from both the regional hospital (IRB pending) and the university health center regarding antibiotic prescription information and QAC use prior and during the course of the pandemic. If these data are obtained they would help further explain changes in behavior and antibiotic usage that may have led to the shifts seen in the resistome over time and between the two WWTPs.

2.6 Conclusions

This study showed that the resistomes and microbiomes from two rural WWTP were impacted differently during the pandemic, with WWTP1 receiving university wastewater being more drastically impacted during the pandemic (2020-2021) vs later/after (2022). Trends in SARS-CoV-2 RNA detected in the wastewater were positively correlated with the clinical cases reported for the community; both of these markers of the pandemic also varied between the two WWTPs. The trends in clinical cases for the university town WWTP1 increased with the return of undergraduate students at the start of each semester, whereas WWTP2 was more similar to nationwide case levels. QAC resistance genes in wastewater did not increase over time nor did they show consistent patterns across WWTPs, they did however strongly correlate with bacterial families known to be hosts of these genes. While this study did not detect overall increases in AMR in wastewater, additional indicators of resistance must be considered when considering the impact that the pandemic had on AMR around the world. Furthermore, it must be acknowledged that the trends revealed here may be specific to the rural community surveilled and changes in AMR may differ for areas that had different responses to the pandemic.

2.7 Acknowledgements

This study was supported by an internal award from Virginia Tech to study the sewage during the pandemic and also NSF CSSI Award 2004751, NSF NRT Award 2125798

The authors would like to extend their thanks to the WWTP operators who provided samples as well as data over the years. Thank you to Monica Gurung, Callen Barlik, and Ana Haines for all of your help in the lab. Thank you to Jason Deese from VDH for providing local clinical case data.

2.8 References

1. Zuccato E, Chiabrando C, Castiglioni S, Calamari D, Bagnati R, Schiarea S, et al. Cocaine in surface waters: A new evidence-based tool to monitor community drug abuse. *Environ Heal A Glob Access Sci Source* [Internet]. 2005 Aug 5 [cited 2021 Jan 30];4:14. Available from: /pmc/articles/PMC1190203/?report=abstract
2. Daughton CG. Illicit Drugs in Municipal Sewage Proposed New Nonintrusive Tool to Heighten Public Awareness of Societal Use of Illicit-Abused Drugs and Their Potential for Ecological Consequences. 2001 [cited 2021 Jan 30]; Available from: <http://www.mnmter.net/~pu>
3. Okoffo ED, Tschärke BJ, O'Brien JW, O'Brien S, Ribeiro F, Burrows SD, et al. Release of Plastics to Australian Land from Biosolids End-Use. *Environ Sci Technol* [Internet]. 2020 Dec 1 [cited 2021 Jul 22];54(23):15132–41. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.est.0c05867>
4. Thai PK, Tschärke BJ, O'Brien JW, Zheng Q, Thomas K V., Mueller JF. Estimating Alcohol Consumption by Wastewater-Based Epidemiology: An Assessment of the Correction Factor for Ethyl Sulfate Using Large-Scale National Monitoring Data. *Environ Sci Technol Lett* [Internet]. 2021 Apr 13 [cited 2021 Jul 22];8(4):333–8. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.estlett.1c00065>
5. Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Heal* [Internet]. 2020 [cited 2021 Oct 15];17:1–7. Available from: <https://doi.org/10.1016/j.coesh.2020.04.006>
6. Larsen DA, Green H, Collins MB, Kmush BL. Wastewater monitoring, surveillance and epidemiology: a review of terminology for a common understanding. *FEMS Microbes* [Internet]. 2022 Jan 13 [cited 2022 Feb 16];2:11. Available from: <https://academic.oup.com/femsmicrobes/article/doi/10.1093/femsmc/xtab011/6354777>
7. Sharif S, Ikram A, Khurshid A, Salman M, Mehmood N, Arshad Y, et al. Detection of SARS-Coronavirus-2 in wastewater, using the existing environmental surveillance network: An epidemiological gateway to an early warning for COVID-19 in communities. *medRxiv*. 2020 Jun 7;2020.06.03.20121426.
8. Rosa G La, Iaconelli M, Mancini P, Ferraro GB, Veneri C, Bonadonna L, et al. First Detection of Sars-Cov-2 in Untreated Wastewaters in Italy. *medRxiv*. 2020 May 7;2020.04.25.20079830.
9. Jahn K, Dreifuss D, Topolsky I, Kull A, Ganesanandamoorthy P, Fernandez-Cassi X, et al. Detection of SARS-CoV-2 variants in Switzerland by genomic analysis of wastewater samples. *medRxiv* [Internet]. 2021 [cited 2021 Feb 26]; Available from: <https://doi.org/10.1101/2021.01.08.21249379>
10. Ahmed W, Tschärke B, Bertsch PM, Bibby K, Bivins A, Choi P, et al. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission

- in the community: A temporal case study. *Sci Total Environ*. 2021 Mar 20;761:144216.
11. Medema G, Heijnen L, Elsinga G, Italiaander R, Brouwer A. Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environ Sci Technol Lett* [Internet]. 2020 [cited 2020 Jun 9];7:511–6. Available from: <https://pubs-acsc-org.sire.ub.edu/doi/pdf/10.1021/acs.estlett.0c00357>
 12. Peccia J, Zulli A, Brackney DE, Grubaugh ND, Kaplan EH, Casanovas-Massana A, et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat Biotechnol* 2020 3810 [Internet]. 2020 Sep 18 [cited 2022 Feb 13];38(10):1164–7. Available from: <https://www.nature.com/articles/s41587-020-0684-z>
 13. CDC. National Wastewater Surveillance System [Internet]. Centers for Disease Control and Prevention. 2022. Available from: <https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html>
 14. Harrington A, Vo V, Papp K, Tillett RL, Chang C-L, Baker H, et al. Urban monitoring of antimicrobial resistance during a COVID-19 surge through wastewater surveillance. *Sci Total Environ* [Internet]. 2022 Jan 2;853:158577. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969722056765>
 15. CDC. COVID-19: U.S. Impact on Antimicrobial Resistance, Special Report 2022 [Internet]. National Center for Emerging and Zoonotic Infectious Diseases; 2022 Jan. Available from: <https://stacks.cdc.gov/view/cdc/117915>
 16. WHO. Antimicrobial resistance [Internet]. 2021. Available from: <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>
 17. CDC. Antibiotic Use in the United States, 2021 Update: Progress and Opportunities [Internet]. US Department of Health and Human Services; 2021. Available from: <https://www.cdc.gov/antibiotic-use/stewardship-report/current.html>
 18. Ashiru-Oredope D, Kerr F, Hughes S, Urch J, Lanzman M, Yau T, et al. Assessing the Impact of COVID-19 on Antimicrobial Stewardship Activities/Programs in the United Kingdom. *Antibiotics* [Internet]. 2021;10(2):110. Available from: <https://doi.org/10.3390/antibiotics>
 19. Tomczyk S, Taylor A, Brown A, De Kraker MEA, El-Saed A, Alshamrani M, et al. Impact of the COVID-19 pandemic on the surveillance, prevention and control of antimicrobial resistance: A global survey. Vol. 76, *Journal of Antimicrobial Chemotherapy*. 2021.
 20. Rose AN, Baggs J, Wolford H, Neuhauser MM, Srinivasan A, Gundlapalli A V, et al. Trends in Antibiotic Use in United States Hospitals During the Coronavirus Disease 2019 Pandemic. *Open Forum Infect Dis* [Internet]. 2021 Jan 1;8(6):ofab236. Available from: <https://doi.org/10.1093/ofid/ofab236>
 21. Kubin CJ, Calfee DP, Singh HK, Greendyke WG, Simon MS, Yoko Furuya E. Antimicrobial stewardship perspectives from a New York City hospital during the COVID-19 pandemic: Challenges and opportunities. *Am Soc Heal Pharm* 2021 [Internet]. 2021 [cited 2021 Dec 22];78(8). Available from: <https://academic.oup.com/ajhp/article/78/8/743/6128902>
 22. Musuuza JS, Watson L, Parmasad V, Putman-Buehler N, Christensen L, Safdar N. Prevalence and outcomes of co-infection and superinfection with SARS-CoV-2 and other pathogens: A systematic review and meta-analysis. *PLoS One* [Internet].

- 2021;16(5):e0251170. Available from: <https://doi.org/10.1371/journal.pone.0251170>
23. Alhumaid S, Al Mutair A, Al Alawi Z, Alshawhi AM, Alomran SA, Almuhanha MS, et al. Coinfections with Bacteria, Fungi, and Respiratory Viruses in Patients with SARS-CoV-2: A Systematic Review and Meta-Analysis. *Pathogens* [Internet]. 2021;10(7):809. Available from: <https://doi.org/10.3390/pathogens10070809>
24. CDC. Outpatient Antibiotic Prescriptions — United States, 2019 [Internet]. 2019 [cited 2022 Feb 7]. Available from: <https://www.cdc.gov/antibiotic-use/data/report-2019.html>
25. CDC. Outpatient Antibiotic Prescriptions — United States, 2020 [Internet]. 2020 [cited 2021 Dec 24]. Available from: <https://www.cdc.gov/antibiotic-use/data/report-2020.html>
26. CDC. Outpatient antibiotic prescriptions-United States, 2021 [Internet]. 2021. Available from: <https://www.cdc.gov/antibiotic-use/pdfs/Annual-Report-2021-H.pdf>
27. Usman M, Farooq M, Hanna K. Environmental side effects of the injudicious use of antimicrobials in the era of COVID-19. *Sci Total Environ*. 2020 Nov 25;745:141053.
28. Rawson TM, Ming D, Ahmad R, Moore LSP, Holmes AH. Antimicrobial use, drug-resistant infections and COVID-19. *Nat Rev Microbiol*. 2020;18(8):409–10.
29. Lu J, Guo J. Disinfection spreads antimicrobial resistance. *Science* (80-). 2021;371(6528):474.
30. Rezasoltani S, Yadegar A, Hatami B, Aghdaei HA, Zali MR. Antimicrobial resistance as a hidden menace lurking behind the COVID-19 outbreak: the global impacts of too much hygiene on AMR. *Front Microbiol*. 2020;11.
31. Chibabhai V, Duse AG, Perovic O, Richards GA. Collateral damage of the COVID-19 pandemic: Exacerbation of antimicrobial resistance and disruptions to antimicrobial stewardship programmes? *South African Med J* [Internet]. 2020 Jun 2 [cited 2021 Jul 19];110(7). Available from: <https://doi.org/10.1086/591708>
32. Ansari S, Hays JP, Kemp A, Okechukwu R, Murugaiyan J, Ekwanzala MD, et al. The potential impact of the COVID-19 pandemic on global antimicrobial and biocide resistance: an AMR Insights global perspective. *JAC-Antimicrobial Resist*. 2021;3(2):dlab038.
33. Cong W, Poudel N, Alhusein N, Wang H, Yao G, Lambert H. Antimicrobial use in COVID-19 patients in the first phase of the SARS-CoV-2 pandemic: Rapid review and evidence synthesis. *medRxiv*. 2021;
34. Nayak B, Parshionikar S. Beware of Superbugs in a Post-COVID World. *J Am Water Works Assoc* [Internet]. 2021 May 1 [cited 2021 Jul 19];113(4):82. Available from: <https://pmc/articles/PMC8206946/>
35. CDC. Cleaning and Disinfecting Your Facility [Internet]. Centers for Disease Control and Prevention2. 2021 [cited 2021 Sep 16]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/community/disinfecting-building-facility.html>
36. EPA. EPA List N Advanced Search Page: Disinfectants for Coronavirus (COVID-19) [Internet]. 2021 [cited 2021 Jul 13]. Available from: <https://cfpub.epa.gov/wizards/disinfectants/>
37. Chen B, Han J, Dai H, Jia P. Biocide-tolerance and antibiotic-resistance in community environments and risk of direct transfers to humans: Unintended consequences of community-wide surface disinfecting during COVID-19? *Environ Pollut*. 2021 Aug 15;283:117074.
38. Lewis D. COVID-19 rarely spreads through surfaces. So why are we still deep cleaning?

- Nature [Internet]. 2021 [cited 2021 Sep 16];590(7844):26–8. Available from: https://www.nature.com/articles/d41586-021-00251-4?utm_source=Nature+Briefing&utm_campaign=65657f8ef0-briefing-dy-20210201&utm_medium
39. Dear K, Palmer A, Nixon R. Contact allergy and allergic contact dermatitis from benzalkonium chloride in a tertiary dermatology center in Melbourne, Australia. *Contact Dermatitis*. 2021;85(2).
 40. Alda-Vidal C, Browne AL, Hoolohan C. “Unflushables”: Establishing a global agenda for action on everyday practices associated with sewer blockages, water quality, and plastic pollution. *Wiley Interdiscip Rev Water* [Internet]. 2020 Jul 1 [cited 2022 Jan 31];7(4):e1452. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/wat2.1452>
 41. Harmeet Kaur. Disinfecting wipes are being flushed down toilets and causing major pipe problems [Internet]. CNN. 2020 [cited 2022 Jan 31]. Available from: <https://www.cnn.com/2020/03/20/us/wipes-flush-toilet-pipes-wellness-trnd/index.html>
 42. Pati SG, Arnold WA. Comprehensive screening of quaternary ammonium surfactants and ionic liquids in wastewater effluents and lake sediments. *Environ Sci Process Impacts* [Internet]. 2020 Feb 26 [cited 2022 Jan 29];22(2):430–41. Available from: <https://pubs.rsc.org/en/content/articlehtml/2020/em/c9em00554d>
 43. Schrank CL, Minbiole KPC, Wuest WM. Are Quaternary Ammonium Compounds, the Workhorse Disinfectants, Effective against Severe Acute Respiratory Syndrome-Coronavirus-2? *ACS Infect Dis* [Internet]. 2020 Jul 10 [cited 2022 Jan 29];6(7):1553–7. Available from: <https://pubs.acs.org/doi/abs/10.1021/acsinfecdis.0c00265>
 44. Zheng G, Filippelli GM, Salamova A. Increased Indoor Exposure to Commonly Used Disinfectants during the COVID-19 Pandemic. *Environ Sci Technol Lett* [Internet]. 2020 Oct 13 [cited 2022 Jan 29];7(10):760–5. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.estlett.0c00587>
 45. Hora PI, Pati SG, McNamara PJ, Arnold WA. Increased Use of Quaternary Ammonium Compounds during the SARS-CoV-2 Pandemic and Beyond: Consideration of Environmental Implications. *Environ Sci Technol Lett* [Internet]. 2020 Sep 8 [cited 2022 Jan 29];7(9):622–31. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.estlett.0c00437>
 46. Hegstad K, Langsrud S, Lunestad BT, Scheie AA, Sunde M, Yazdankhah SP. Does the Wide Use of Quaternary Ammonium Compounds Enhance the Selection and Spread of Antimicrobial Resistance and Thus Threaten Our Health? *Microb drug Resist*. 2010;16(2):91–104.
 47. Zheng G, Webster TF, Salamova A. Quaternary ammonium compounds: Bioaccumulation potentials in humans and levels in blood before and during the covid-19 pandemic. *Environ Sci Technol* [Internet]. 2021 Nov 2 [cited 2022 Jan 29];55(21):14689–98. Available from: <https://pubs.acs.org/doi/abs/10.1021/acs.est.1c01654>
 48. Gillings MR, Xuejun D, Hardwick SA, Holley MP, Stokes HW. Gene cassettes encoding resistance to quaternary ammonium compounds: A role in the origin of clinical class 1 integrons? *ISME J*. 2009;3(2).
 49. Ceccarelli D, Salvia AM, Sami J, Cappuccinelli P, Colombo MM. New cluster of plasmid-located class 1 integrons in *Vibrio cholerae* O1 and a *dfrA15* cassette-containing integron in *Vibrio parahaemolyticus* isolated in Angola. *Antimicrob Agents Chemother* [Internet].

- 2006 Jul [cited 2022 Jan 29];50(7):2493–9. Available from: <https://journals.asm.org/doi/abs/10.1128/AAC.01310-05>
50. Wright GD. The antibiotic resistome: The nexus of chemical and genetic diversity. Vol. 5, *Nature Reviews Microbiology*. 2007.
 51. Majeed HJ, Riquelme M V., Davis BC, Gupta S, Angeles L, Aga DS, et al. Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Front Microbiol* [Internet]. 2021 May 13 [cited 2021 Jun 8];12:1048. Available from: www.frontiersin.org
 52. Jirka AM, Carter MJ. Micro–semi-automated analysis of surface and wastewaters for chemical oxygen demand. *Anal Chem* [Internet]. 1975 [cited 2022 Jun 3];47(8):1397–402. Available from: <https://pubs.acs.org/sharingguidelines>
 53. APHA-AWWA-WEF. Standard methods for the examination of water and wastewater. 23rd ed. Baird R, Bridgewater L, editors. Washington, D.C.: American Public Health Association; 2017.
 54. Climate Data Online (CDO) [Internet]. 2022. Available from: <https://www.ncei.noaa.gov/cdo-web/>
 55. Ahmed W, Bertsch PM, Bivins A, Bibby K, Farkas K, Gathercole A, et al. Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Sci Total Environ* [Internet]. 2020 Oct 15 [cited 2020 Jun 15];739:139960. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S004896972033480X>
 56. Cohen A, Maile-Moskowitz A, Grubb C, Gonzalez RA, Ceci A, Darling A, et al. Subsewershed SARS-CoV-2 Wastewater Surveillance and COVID-19 Epidemiology Using Building-Specific Occupancy and Case Data. *ACS ES&T Water* [Internet]. 2022 Dec 29;2(11):2047–59. Available from: <https://doi.org/10.1021/acsestwater.2c00059>
 57. Ceci A, Muñoz-Ballester C, Tegge AN, Brown KL, Umans RA, Michel FM, et al. Development and implementation of a scalable and versatile test for COVID-19 diagnostics in rural communities. *Nat Commun* 2021 121 [Internet]. 2021 Jul 20 [cited 2021 Oct 28];12(1):1–14. Available from: <https://www.nature.com/articles/s41467-021-24552-4>
 58. Gonzalez R, Larson A, Thompson H, Carter E, Cassi XF. Redesigning SARS-CoV-2 clinical RT-qPCR assays for wastewater RT-ddPCR. *medRxiv* [Internet]. 2021 Jan 13 [cited 2022 Feb 10];2021.03.02.21252754. Available from: <https://www.medrxiv.org/content/10.1101/2021.03.02.21252754v2>
 59. Chen S, Zhou Y, Chen Y, Gu J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. In: *Bioinformatics*. 2018.
 60. Clum A, Huntemann M, Bushnell B, Foster B, Foster B, Roux S, et al. DOE JGI Metagenome Workflow. *mSystems* [Internet]. 2021 Jun 29 [cited 2022 Feb 13];6(3). Available from: <https://pubmed.ncbi.nlm.nih.gov/34006627/>
 61. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2020;48(D1).
 62. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20(1).
 63. Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol*. 2020;38(9).

64. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarszewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36(10).
65. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. Database indexing for production MegaBLAST searches. In: *Bioinformatics.* 2008.
66. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2022;50(D1).
67. Hutinel M, Larsson DGJ, Flach CF. Antibiotic resistance genes of emerging concern in municipal and hospital wastewater from a major Swedish city. *Sci Total Environ.* 2022;812.
68. EPA. Method 1694: Pharmaceuticals and Personal Care Products in Water, Soil, Sediment, and Biosolids by HPLC/MS/MS. 2007 [cited 2022 Feb 13]; Available from: <http://www.epa.gov/waterscience>
69. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
70. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat [Internet].* 1979 Jan 2;6(2):65–70. Available from: <https://www.jstor.org/stable/4615733>
71. Gross J, Ligges U. nortest: Tests for Normality. R package version 1.0-4; 2015.
72. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika.* 1965;52(3–4).
73. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests [Internet]. R package. 2021. Available from: <https://cran.r-project.org/package=rstatix>
74. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. The vegan Package Title Community Ecology Package [Internet]. R package. 2022 [cited 2022 Jun 2]. Available from: <https://cran.r-project.org/package=vegan>
75. Smith AJH, Potvin LR, Lilleskov EA. Fertility-dependent effects of ectomycorrhizal fungal communities on white spruce seedling nutrition. *Mycorrhiza.* 2015;25(8).
76. McDonald JH. Handbook of Biological Statistics. 2009; Available from: <http://biostathandbook.com/HandbookBioStatSecond.pdf>
77. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
78. Wilke CO. cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2” R package version 1.1.1. 2018.
79. Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, et al. Ggally: Extension to ggplot2 R package version 2.1.2. 2020.
80. Garnier S, Ross N, Rudis R, Camargo AP, Sciaini M, Scherer C. Rvision - Colorblind-Friendly Color Maps for R. R package version 0.6.1. 2021;
81. Wright K. pals: Color Palettes, Colormaps, and Tools to Evaluate Them. R package version 1.7. 2021; Available from: <https://cran.r-project.org/package=pals>
82. Gao Y, Shi X, Jin X, Wang XC, Jin P. A critical review of wastewater quality variation and in-sewer processes during conveyance in sewer systems. *Water Res [Internet].* 2023 Dec 28;228:119398. Available from: <https://www.sciencedirect.com/science/article/pii/S0043135422013434>
83. McLellan SL, Roguet A. The unexpected habitat in sewer pipes for the propagation of microbial communities and their imprint on urban waters. Vol. 57, *Current Opinion in*

- Biotechnology. 2019.
84. VandeWalle JL, Goetz GW, Huse SM, Morrison HG, Sogin ML, Hoffmann RG, et al. *Acinetobacter*, *Aeromonas* and *Trichococcus* populations dominate the microbial community within urban sewer infrastructure. *Environ Microbiol*. 2012;14(9).
 85. LaMartina E Lou, Mohaimani AA, Newton RJ. Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome*. 2021;9(1).
 86. Ramin P, Brock AL, Polesel F, Causanilles A, Emke E, De Voogt P, et al. Transformation and sorption of illicit drug biomarkers in sewer systems: Understanding the role of suspended solids in raw wastewater. *Environ Sci Technol*. 2016;50(24).
 87. Guo Y, Sivakumar M, Jiang G. Decay of four enteric pathogens and implications to wastewater-based epidemiology: Effects of temperature and wastewater dilutions. Vol. 819, *Science of the Total Environment*. 2022.
 88. Fierer N, Holland-Moritz H, Alexiev A, Batther H, Dragone NB, Friar L, et al. A Metagenomic Investigation of Spatial and Temporal Changes in Sewage Microbiomes across a University Campus. *mSystems* [Internet]. 2022 Dec 28;7(5):e00651-22. Available from: <https://journals.asm.org/doi/full/10.1128/msystems.00651-22>
 89. Bakchan A, Roy A, Faust KM. Leveraging water-wastewater data interdependencies to understand infrastructure systems' behaviors during COVID-19 pandemic. *J Clean Prod* [Internet]. 2022 Dec 29;367:132962. Available from: <https://www.sciencedirect.com/science/article/pii/S0959652622025549>
 90. Buehrle DJ, Wagener MM, Nguyen MH, Clancy CJ. Trends in Outpatient Antibiotic Prescriptions in the United States during the COVID-19 Pandemic in 2020. *JAMA Netw Open*. 2021;4(9).
 91. King LM, Lovegrove MC, Shehab N, Tsay S, Budnitz DS, Geller AI, et al. Trends in US Outpatient Antibiotic Prescriptions during the Coronavirus Disease 2019 Pandemic. *Clin Infect Dis*. 2021;73(3).
 92. CDC. Reduced Access to Care - Household Pulse Survey - COVID-19 [Internet]. 2022. Available from: <https://www.cdc.gov/nchs/covid19/pulse/reduced-access-to-care.htm>
 93. Goldstein E, Fireman BH, Klein NP, Lipsitch M, Ray GT. Antibiotic prescribing across age groups in the Kaiser Permanente Northern California population in association with different diagnoses, and with influenza incidence, 2010-2018. *Epidemiol Infect*. 2022;150.
 94. Young EH, Panchal RM, Yap AG, Reveles KR. National Trends in Oral Antibiotic Prescribing in United States Physician Offices from 2009 to 2016. *Pharmacotherapy*. 2020;40(10).
 95. Karthikeyan S, Ronquillo N, Belda-Ferre P, Alvarado D, Javidi T, Longhurst CA, et al. High-Throughput Wastewater SARS-CoV-2 Detection Enables Forecasting of Community Infection Dynamics in San Diego County. *mSystems*. 2021;6(2).
 96. CDC. COVID Data Tracker [Internet]. Atlanta, GA; 2023. Available from: <https://covid.cdc.gov/covid-data-tracker%0A>
 97. Haischer MH, Beilfuss R, Hart MR, Opielinski L, Wrucke D, Zirgaitis G, et al. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. *PLoS One*. 2020;15(10 October).

APPENDIX A

A.1 Supporting Tables

Table A.1 Table of mean, standard deviation, and statistics results for parameters measured by each WWTP and for parameters measured in the VT lab

	Parameter	WWTP1		WWTP2		Wilcox, p-value
		Mean	SD	Mean	SD	
Measured by WWTP	BOD (mg/L)	149	37.2	226	72.1	1.51e- 16 ****
	Flow (MGD)	6.93	1.67	3.43	2.09	1.78e-116 ****
	pH	7.50	0.157	7.79	0.0989	1.62e- 89 ****
	TSS (mg/L)	157	54.1	235	96.5	4e-40 ****
	Ammonia (mg/L)	--	--	24.9	9.29	--
	Oil and grease (mg/L)	--	--	13.6	5.90	--
	TKN (mg/L)	--	--	36.5	18.0	--
Measured in the VT lab	Ammonia	21.6	15.6	20.9	12.8	0.85
	COD	226	75.8	276	124	0.412
	TSS	139	45.3	181	65.2	0.0105 *
	VSS	129	49.0	173	64.7	0.00826 *

* p≤0.05, **p≤0.01, ***p≤0.001, ****p≤0.0001

Table A.2 Table of mean, standard deviation for weekly COVID-19 cases broken down by month

WWTP	Year	Month	Mean cases/week	SD cases/week
#1	2020	8	66	67
		9	48	14
		10	36	13
		11	26	12
		12	24	14
	2021	1	49	21
		2	89	11
		3	35	10
		4	11	6
		5	5	3
#2	2020	8	9	3
		9	7	3
		10	13	5
		11	26	2
		12	34	8
	2021	1	23	11
		2	6	2
		3	8	6
		4	12	4
		5	4	2

Table A.3 Taxonomy classifications and the number of organisms associated with QAC resistance genes in NCBI (BLASTN (megablast)) filtered for 90-100 query coverage and 90-100 percent identity.

GTDB Order	GTDB Family	NCBI Family	Number of Organisms (family)			
			<i>qacE</i>	<i>qacEdelta1</i>	<i>qacL</i>	<i>qacH</i>
Actinomycetales	Actinomycetaceae	Actinomycetaceae	2			
	Microbacteriaceae	Microbacteriaceae	2			
	Micrococcaceae	Micrococcaceae	4			
Mycobacteriales	Mycobacteriaceae	Nocardiaceae	2			
Lactobacillales	Enterococcaceae	Enterococcaceae				3 (genera <i>Enterococcus</i>)
Staphylococcales	Staphylococcaceae	Staphylococcaceae				7
Rhizobiales	Rhizobiaceae	Rhizobiaceae	3			
Rhizobiales_A	Rhizobiaceae_A					
Burkholderiales	Burkholderiaceae	Alcaligenaceae	11	11		
		Comamonadaceae	12	12	4	
	Rhodocyclaceae	Zoogloeaceae	2 (genera <i>Thauera</i>)	2 (genera <i>Thauera</i>)		
Enterobacterales	Aeromonadaceae	Aeromonadaceae	21 (genera <i>Aeromonas</i>)	21 (genera <i>Aeromonas</i>)	3 (genera <i>Aeromonas</i>)	
	Enterobacteriaceae	Enterobacteriaceae	184	192	74	
		Morganellaceae	18	18	6	
		Yersiniaceae	2	2		
		Edwardsiella	3	3		
		Pantoea	3	3		
	Pasteurellaceae	Pasteurellaceae		2		
	Shewanellaceae	Shewanellaceae	9 (genera <i>Shewanella</i>)	8 (genera <i>Shewanella</i>)		
	Vibrionaceae	Vibrionaceae	10	9		
Pseudomonadales	Moraxellaceae	Moraxellaceae	28	28	4	
	Pseudomonadaceae	Pseudomonadaceae	29	27		
Xanthomonadales	Xanthomonadaceae	Xanthomonadaceae	4	4	3	

Table A.4 Top ARGs and their resistance classes all of which are also core ARGs
 qacEdelta1 is also within the top 20 most abundant ARGs
 qacE is within top 90 resistance genes for WWTP1 (#79) and WWTP2 (#84)

Gene	Resistance Class	WWTP1		WWTP2	
		Mean	SD	Mean	SD
msrE	MLS	0.038	0.008	0.029	0.007
mphE	MLS	0.034	0.007	0.028	0.007
rsmA	Multidrug	0.016	0.003	0.017	0.004
qacL	Other	0.016	0.004	0.012	0.003
tet(39)	Tetracycline	0.016	0.004	0.012	0.004
MLS indicates resistance to macrolide, lincosamide, and streptogramin antibiotics.					

Table A.5 Parameters that were significantly impactful to the spread of samples from ARG NMDS (Figure A.1), listed in p-value decreasing order for both WWTP1 and WWTP2 together

Environmental Parameter	R ²	p-value	
year	0.4391	0.001	***
Flow	0.3483	0.001	***
Weekly mean minimum temperature	0.2538	0.002	**
Daily minimum temperature	0.2485	0.002	**
pH	0.2401	0.002	**
TSS	0.1944	0.005	**
Daily maximum temperature	0.2056	0.007	**
BOD	0.1732	0.013	*
Weekly mean maximum temperature	0.158	0.018	*
Weekly sum of clinical COVID-19 cases	0.149	0.022	*
* p≤0.05, **p≤0.01, ***p≤0.001, ****p≤0.0001			

A.2 Supplementary Figures

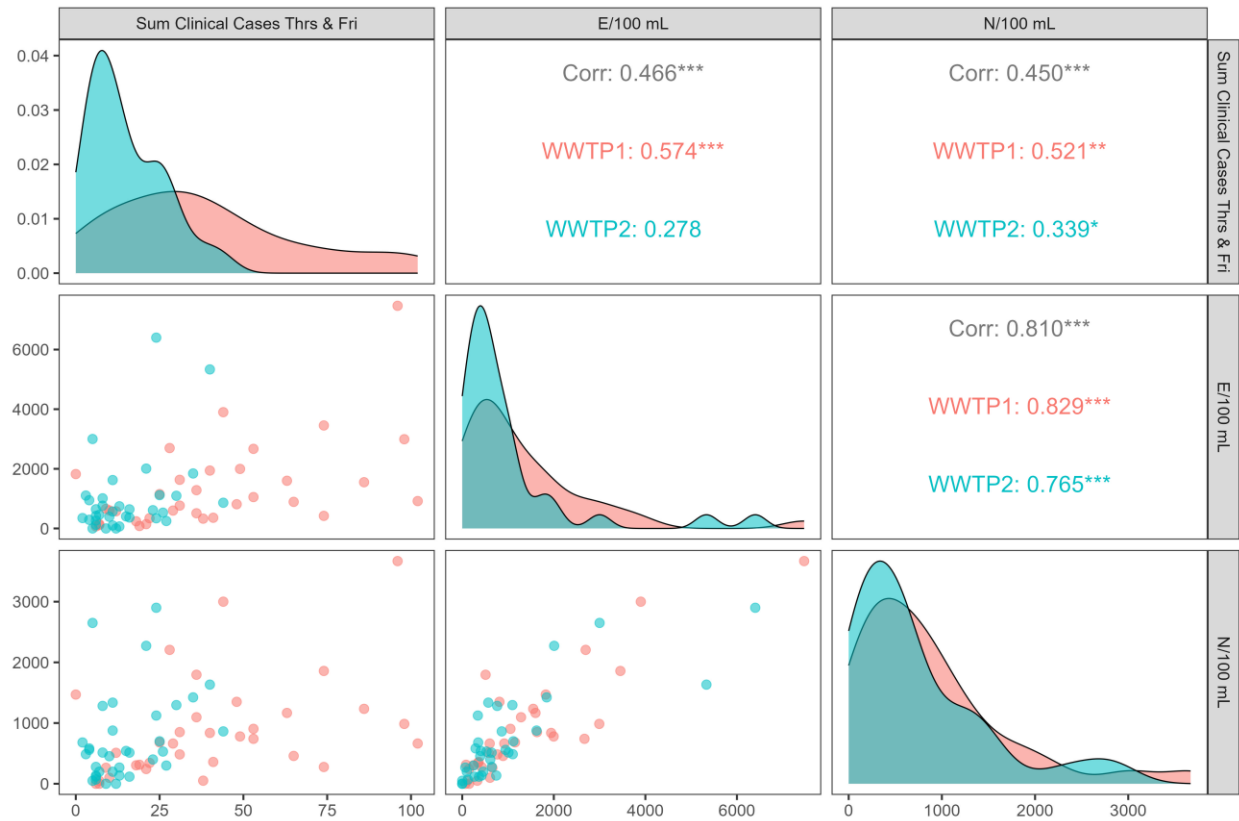


Figure A.1 Histograms (displayed diagonally) for the clinical cases of COVID-19 within each WWTP's sewershed (top row), or concentrations of *E*/100 mL (middle row) and *N*/100 mL (bottom row). The bottom left panels provide scatter plots of these data and the top right panels are the corresponding correlations (Spearman). Corr indicates the Spearman's Rho for both WWTPs together, below which are the correlations for WWTP1 and WWTP2 when compared separately. The units are total counts of clinical COVID cases or *N* or *E* gene copies per 100 mL. The graph was generated in R generated using GGally.

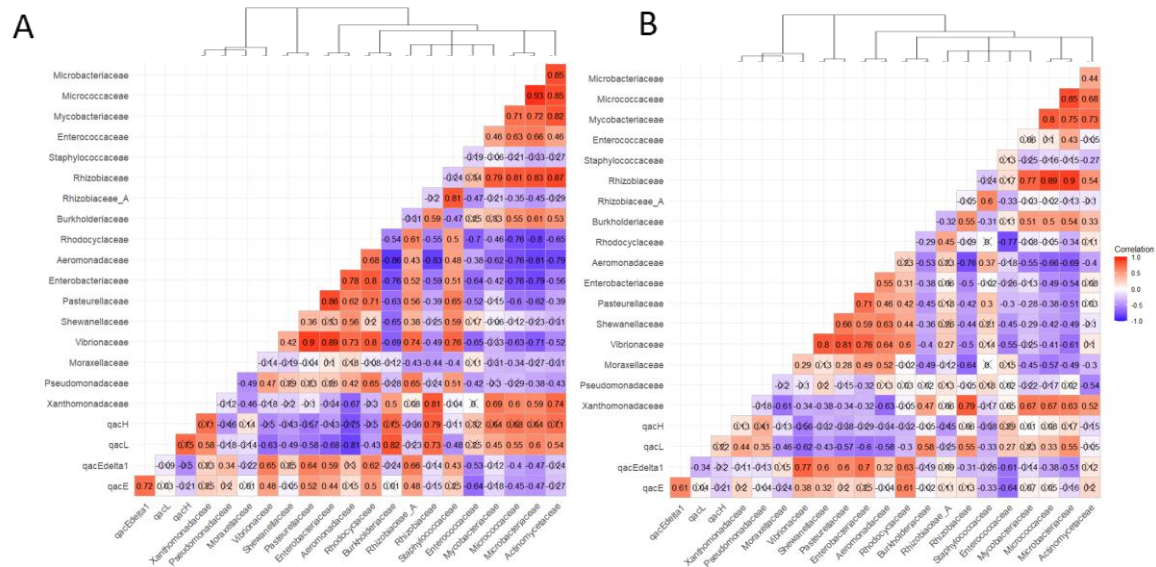


Figure A.2 Spearman correlation matrices (values indicate Spearman's Rho) showing the correlations between QAC resistance genes and their potential hosts for (A) WWTP1 and (B) WWTP2. The correlation extends and also compares QAC resistance genes and the host families. The red color is more positively correlated; the purple color is more negatively correlated. The correlation (R) is written in the matrix square/box, the p-value is not written in the figure, however, square/boxes with an X indicate that the $p > 0.05$ and therefore the correlation is not significant. Dendrogram shows the relatedness of the taxonomy.

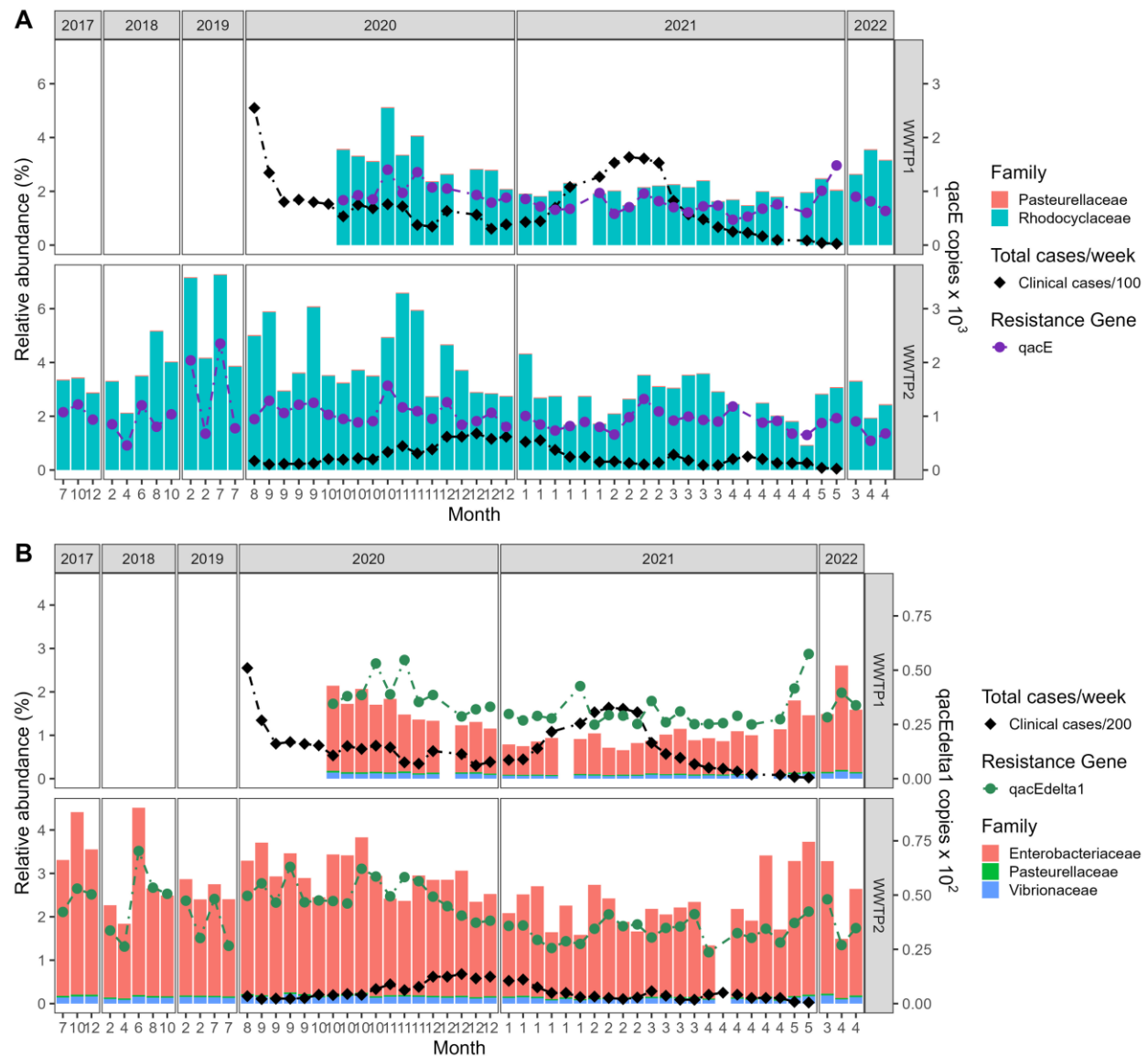


Figure A.3 (A) *qacE* and the most strongly correlated taxa for WWTP1 (Pasteurellaceae and Rhodocyclaceae) and WWTP2 (Rhodocyclaceae) and (B) *qacEdelta1* and the most strongly correlated taxa for both WWTP1 and WWTP2 (Pasteurellaceae, Enterobacteriaceae, and Vibrionaceae).

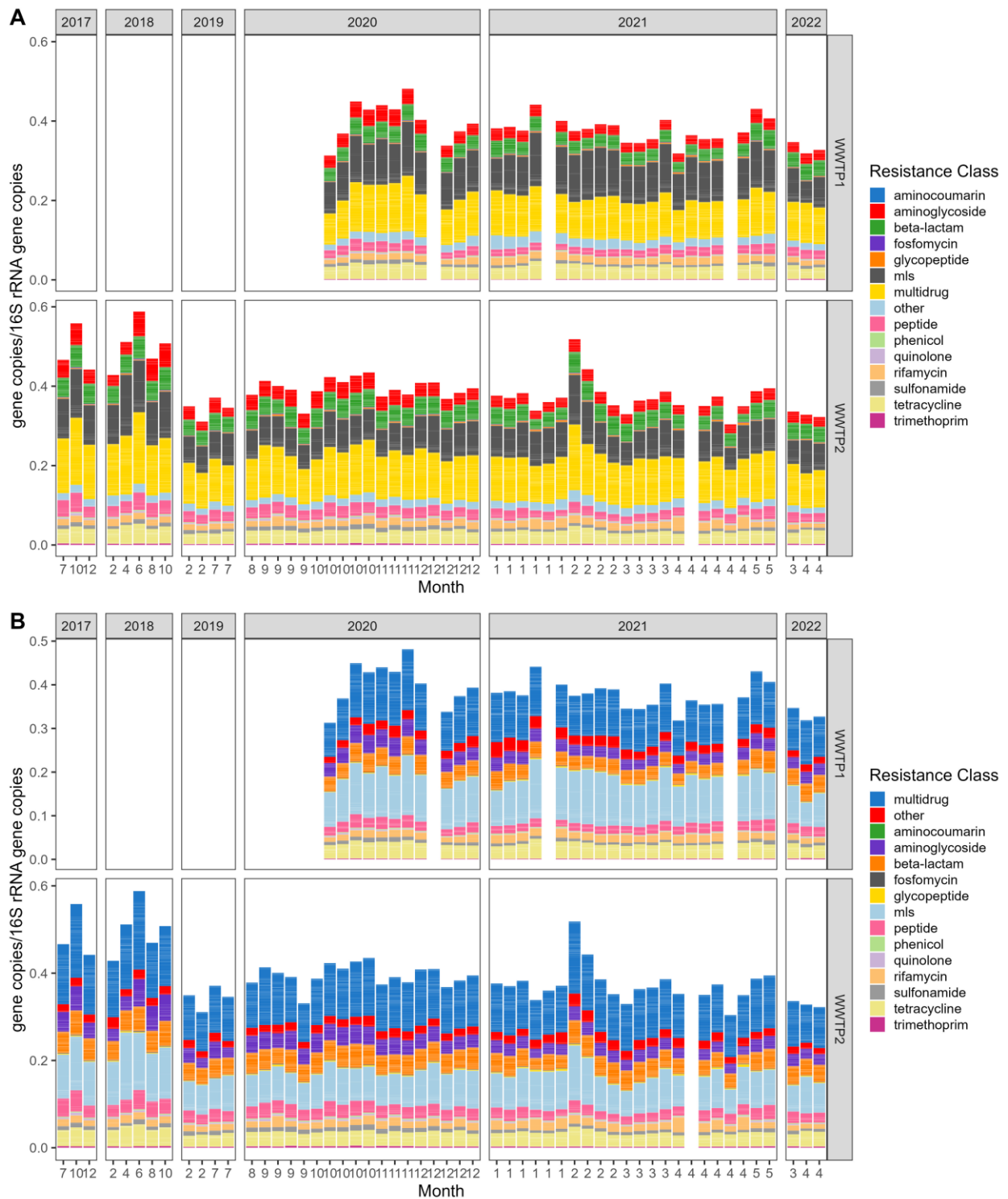


Figure A.4 (A) Core resistance genes and (B) all resistance genes detected in all WWTP1 and WWTP2 samples across all dates.

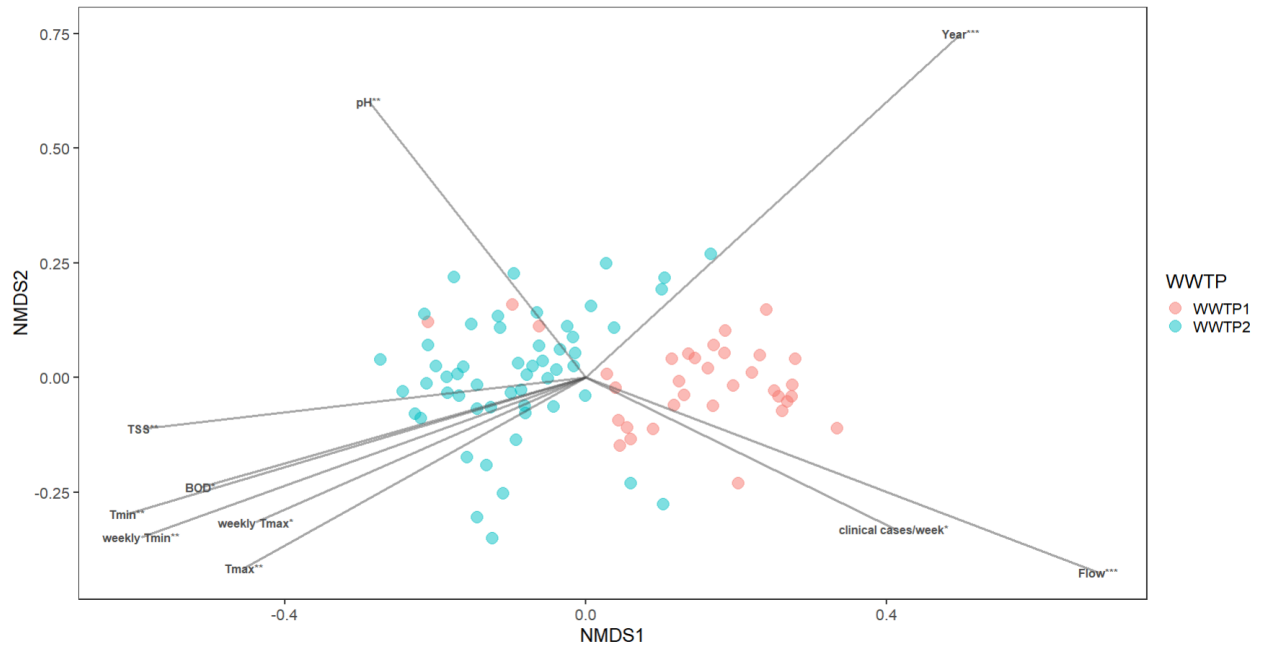


Figure A.5 ARG NMDS with extrinsic factors plotted over top (shows all of the environmental parameters that were considered not just ones that were statistically different)

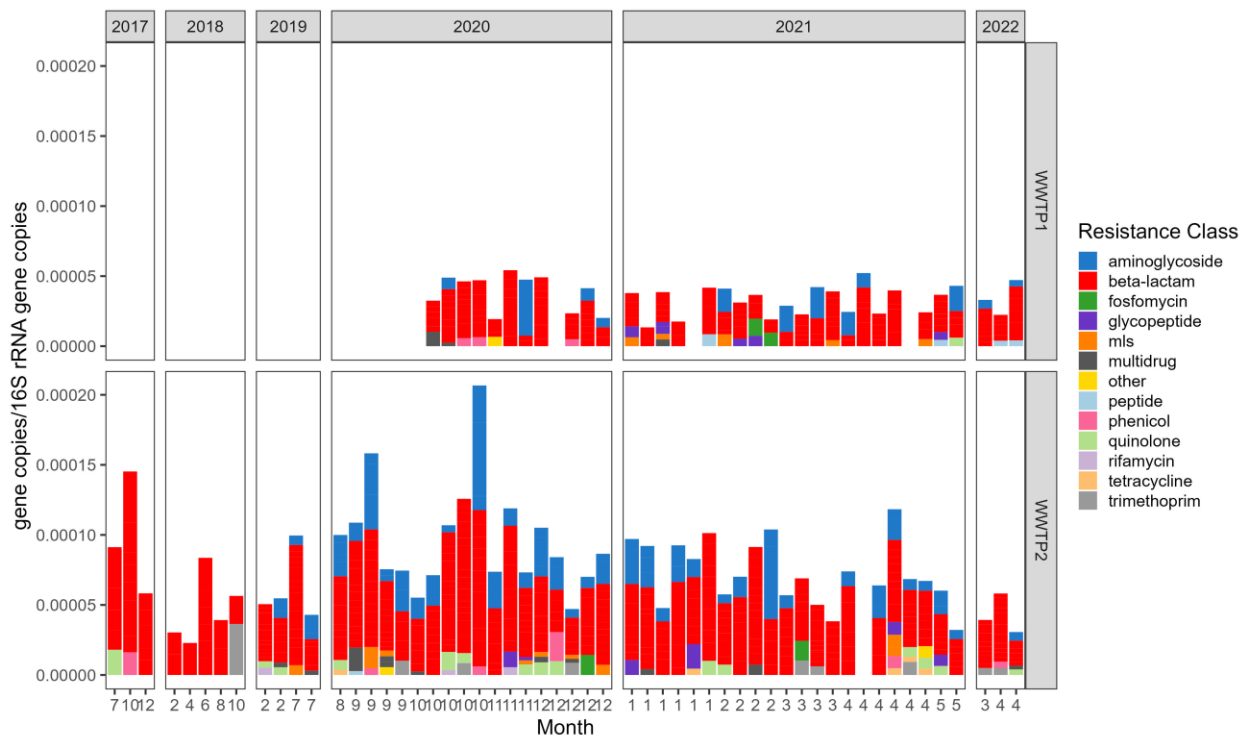


Figure A.6 Discriminatory ARGs for all WWTP1 and WWTP2 samples (when compared across all dates).

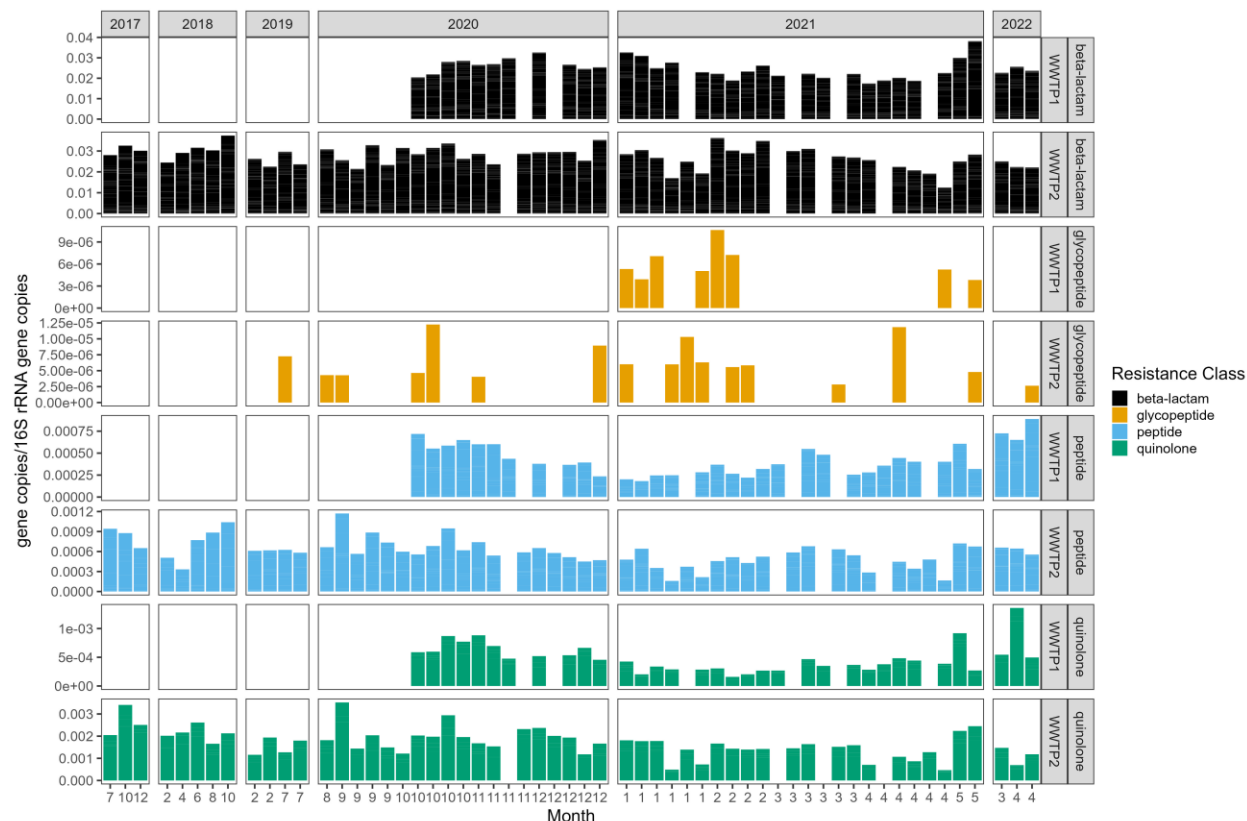


Figure A.7 Clinically relevant ARGs conferring resistance to beta-lactam, glycopeptide, peptide, and quinolone antibiotics detected across WWTPs.

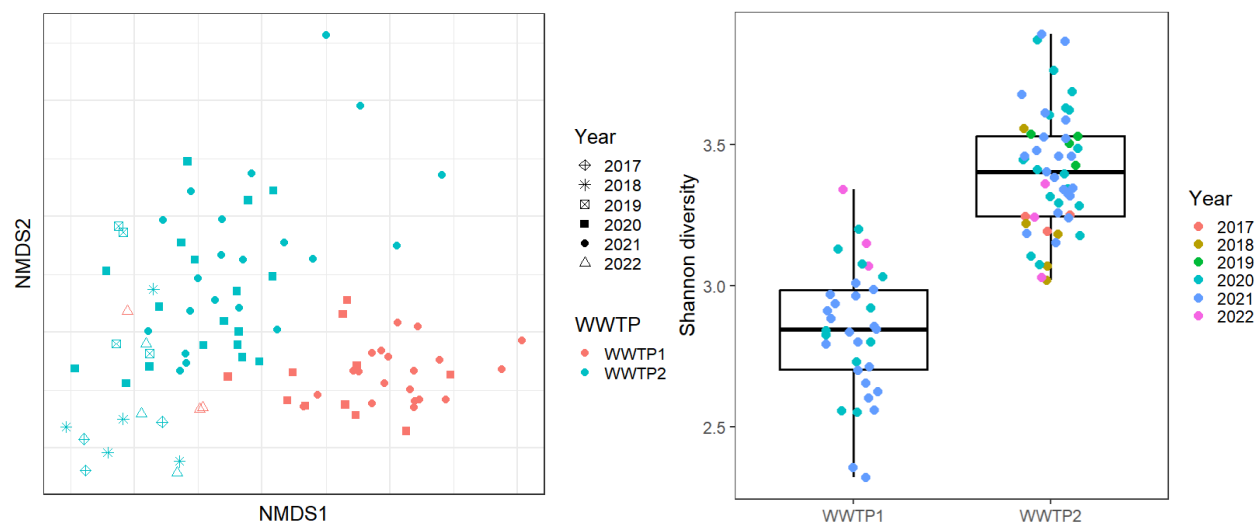


Figure A.8 (A) NMDS of family colored by WWTP with shapes indicating different years (stress = 0.1136702). Groups are distinct between WWTP2 and WWTP1 when grouped by WWTP (PERMANOVA, $R^2=0.30607$, $p=0.001$). Though less explanatory, there are also differences when grouping by year (PERMANOVA, $R^2=0.11821$, $p=0.001$). (B) Shannon diversity of family with colors indicating different years. WWTP1 samples collected in 2022 are highest/most similar to WWTP2.

CHAPTER 3: IMPACT OF HOSPITAL WASTEWATER ON SIMULATED ACTIVATED SLUDGE TREATMENT AND ANTIBIOTIC RESISTOME

Coauthors: Connor Brown, Latania Logan, Kang Xia, Amy Pruden, Peter Vikesland

3.1 Abstract

Wastewater treatment plants (WWTPs) are becoming focal points for monitoring both chemical and/or biological constituents released into wastewater by the community and their potential to escape treatment and be released into the environment. Release of hospital wastewater (HWW) to the municipal sewage stream is of special concern, especially with respect to antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARGs). The objective of this study was to assess whether blending of HWW with domestic wastewater can impact activated sludge wastewater treatment and overall ARB/ARG profiles. Sequencing batch reactors (SBRs) were operated as a controlled simulation of activated sludge treatment with varying inputs of HWW (0 or 10%). Metagenomic sequencing was applied to yield high resolution profiles of the relative abundances of ARGs and corresponding genera carrying them. It was found that the HWW contained a greater abundance of ARGs relative to the background domestic wastewater feed. However, these differences did not negatively impact SBR performance or taxonomic composition. The addition of HWW to the SBRs did result in some differences in ARG abundances, notably glycopeptide resistance gene *vanA* was more frequently detected in the feed and SBRs amended with 10% versus 0% HWW. However, SBR resistomes were not significantly separated according to broad-scale nonmetric multidimensional scaling analysis and total relative ARG abundances. While broad-scale effects of HWW in the feed on activated sludge performance or resistome composition were not apparent from this study, it is important to examine further specific ARGs of HWW origin that may escape treatment, especially those with clinical relevance.

3.2 Water Impact

Hospital wastewater is gaining increasing attention as an important point for monitoring and mitigating antimicrobial resistance. Here we fed laboratory-scale wastewater reactors influent with varying proportions of hospital sewage and did not find evidence of broad-scale impact to the microbial community or antibiotic resistance gene (ARG) profiles, pointing to the value of monitoring individual ARGs and resistant organisms of concern.

3.3 Introduction

Antimicrobial resistance (AMR) is an ongoing threat to human health because it undermines the efficacy of antibiotics to cure diseases. AMR is exacerbated by overuse of antibiotics and other antimicrobials (1). Wastewater treatment plants (WWTPs) are gaining increasing attention as a useful surveillance point for profiling forms of antibiotic resistance circulating in a community, but also as a potentially important point for mitigating further spread of antibiotic resistance into aquatic receiving environments. WWTPs are not intentionally designed to attenuate antibiotic resistant bacteria (ARB) or their carriage of antibiotic resistance genes (ARGs). Further, there is concern that biological wastewater treatment might even promote the spread of ARGs through horizontal gene transfer facilitated by high concentrations of bacteria. Selective agents in the wastewater, such as antibiotics and heavy metals, can potentially induce horizontal gene transfer, even when present at sub-inhibitory concentrations (2,3). A number of studies have recently been published wherein WWTP influent is profiled for pathogens,

antibiotics, and ARGs as an indicator of the antibiotic use and resistance status of the corresponding human population (4). Treated wastewater can be monitored to determine what characteristics (e.g., antibiotics, ARGs, pathogens) are released and can potentially impact the environmental resistome (5).

Hospital wastewater (HWW) might be of even greater concern than municipal wastewater from an AMR standpoint because it often contains higher concentrations of pharmaceuticals; including non-metabolized pharmaceuticals (unused and disposed of down the drain or excreted), disinfectants, heavy metals, x-ray contrast media, and a greater microbial load that may contain pathogens, especially ARB (6,7). There are reference standards and treatment methods for industrial effluents that are released to WWTPs; however, there exists little in the way of analogous recommendations for HWW. HWW is often considered to be essentially equivalent to domestic wastewater in terms of its physiochemical parameters (8,9). The World Health Organization (WHO) has guidelines on the safe management of waste from health care activities (10), but the US Environmental Protection Agency (EPA) allows entities (e.g., states, WWTPs) to determine if HWW must be treated prior to its discharge into the sewer (11). Few countries have specific guidelines for pretreating HWW, however, some hospitals in Denmark and France do utilize pre-treatment processes to remove contaminants before sewage discharge (12,13). Hospital wastewater typically contributes <3% (**Table B.1**) of the total volume that WWTPs receive, therefore some have argued that the pre-treatment of HWW is a waste of effort and resources (14).

The impact of ARB and ARGs from HWW on domestic wastewater treatment remains unclear. A prior study of two Dutch WWTPs revealed that while HWW typically contains greater abundances of human-associated bacteria and ARGs, there was no obvious difference between the composition of influent microbiota and levels of ARGs as a function of their content of HWW (14). Conversely, a study of two WWTPs in Poland showed that there was a greater abundance of beta-lactamase-producing ARGs in wastewater receiving greater percentages of HWW (15). If selective agents play a significant role at hospitals, then one might expect that domestic wastewater containing HWW would be enriched in antibiotic resistance determinants that encode resistance to antibiotics administered in corresponding hospital settings. For example, the route of administration for glycopeptides (typically vancomycin) is typically parenteral (i.e., intravenously), though it may also be administered orally (16), and therefore it is commonly administered only in hospital settings (17); correspondingly, hospital wastewater has been shown to be a source of vancomycin-resistant *Enterococcus* (VRE) (18). Glycopeptides are considered critically important antimicrobials according to the WHO meaning there are limited available therapies to treat serious bacterial infections in people and they are used to treat infections caused by bacteria that may be transmitted from non-human sources or bacteria with resistance genes from non-human sources (19), therefore, it is of particular concern to limit the spread of vancomycin resistance.

A number of methods can be used for AMR monitoring of wastewater. Culture-based approaches allow determination of viable organisms; however, viable but non-culturable cells are not easily grown in laboratory settings and are often missed (20). Targeted molecular techniques (e.g., quantitative or digital polymerase chain reaction) can provide precise quantification of individual genes of interest, such as ARGs; however, they require prior delineation of the target organism or gene of interest. Next generation sequencing (NGS) allows determination of the microbial community's composition as well as their functional potential without *a priori* selection of gene targets. In particular, shot-gun metagenomic sequencing can be applied to capture the full suite of ARGs carried by a microbial community, or the resistome (i.e., the sum of all ARGs (21)).

Metagenomic analysis thus could be applied towards shedding light on the precise effect of HWW on affected WWTP resistomes.

The objective of this study was to assess effects of HWW on the performance of activated sludge wastewater treatment and corresponding resistomes. Lab-scale sequencing batch reactors (SBRs) were applied as controlled, replicable simulations of activated sludge treatment. SBRs provide a small-footprint simulation of an activated sludge WWTP, cycling through aeration, solids wasting, and settling in a single vessel (22). In this study, triplicate SBRs were operated in parallel with 0%, 1%, and 10% HWW blended with local municipal sewage as influent. Shotgun metagenomic sequencing was applied towards obtaining a high resolution comparison of the effects of the various blends of the hospital sewage on the microbial community and resistome composition. The findings of this study help to shed light on whether reducing inputs of HWW to municipal wastewater streams would have a measurable impact on wastewater resistomes and specific ARGs that escape treatment.

3.4 Materials and Methods

3.4.1 SBR Setup and Operation

SBRs (n=9) were maintained for 17 days post HWW addition in a 20 °C temperature-controlled room in three-liter glass beakers with an active volume of two liters. The SBRs were operated on a 12-hour cycle with a two-day hydraulic retention time (HRT) and five-day solids retention time (SRT). Each cycle consisted of a 646.8 minute (10.78 h) aeration/react period, including a 60-minute slow feed, followed by 8 minutes of solids wasting (decanting of ML), 53 minutes of solids settling, and a 12-minute effluent decant. During the aeration/react period the SBRs were aerated using Top Fin® Aquarium Air Pumps and mixed using stir plates. SBRs were fed and decanted using three- or four-roller peristaltic Masterflex® EasyLoad pump heads controlled by Masterflex® pump drives (Model 7553-80). HWW dosing began after the SBRs reached stable mixed liquor suspended solids (MLSS, 430 ± 65 mg/L), mixed liquor volatile suspended solids (MLVSS, 382 ± 61 mg/L), sludge volume index (SVI, 134 ± 41 mg/L), pH (average 6.5 ± 0.5), and effluent soluble chemical oxygen demand (sCOD, 27.4 ± 10.6 mg/L).

The SBRs were fed influent collected prior to primary sedimentation twice weekly from a local WWTP plant in Virginia receiving primarily domestic waste (0.4% pre-treated industrial waste and no hospital waste) with an average flow of 3 MGD. HWW dosing was compared in triplicate with SBRs receiving only WWTP influent (0%), 1% HWW, or 10% HWW (by volume). HWW was collected from an urban medical center located in Chicago, Illinois serving approximately 1.5 million people. The medical center consists of a children's hospital (120 beds), a health center for the elderly (79 beds), and an urban hospital (676 operating beds, including 12 adult critical care beds). The untreated HWW was collected from manholes over a 24-hour period and shipped to the Virginia Tech lab on ice. Upon arrival, subsets of samples were processed for further analysis. HWW was stored at 4 °C for 56 days prior to commencing the experiment. Further shipments of HWW were not possible because the experiment was initiated in February 2020, just before global COVID-19 stay-at-home orders were emplaced.

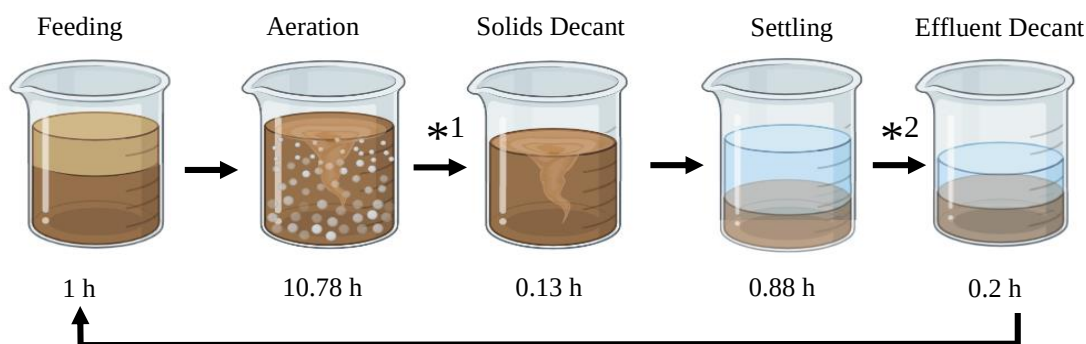


Figure 3.1 Sequencing batch reactor (SBR) 12-hour cycle. Each condition 0, 1, 10% HWW was operated in triplicate. The 1 hour feed time is included within the 10.78 hour aeration time. * indicates sampling of mixed liquor (1) and final effluent (2). Image generated using BioRender.

3.4.2 Sample Collection and DNA Extraction

Once HWW dosing commenced, the pH, sludge volume index (SVI), effluent soluble chemical oxygen demand (sCOD), mixed liquor suspended solids (MLSS), and mixed liquor volatile suspended solids (MLVSS) were tested every 2-5 days and the effluent total suspended solids (TSS) were tested every 1-3 days (**Table B.2**). The HACH reactor digestion method (method 8000, HACH, Loveland, CO) (23) was used for COD and sCOD. Samples for sCOD analysis were filtered through 0.45- μm filter prior to digestion. Ammonia was tested using the HACH salicylate method (Method 8155 HACH, Loveland, CO). Alkalinity, TSS, volatile suspended solids (VSS), and SVI were tested according to Standard Methods for the Examination of Water and Wastewater (methods 2320, 2540, and 213 respectively) (24).

Soluble elements, typically metals, were collected from the mixed liquor fraction of the SBRs. Samples were filtered through a 0.45- μm filter, acidified with nitric acid (2% v/v), and then quantified using inductively coupled plasma-mass spectrometry (ICP-MS) following method 3125 B (24). Pharmaceuticals and personal care products (PPCPs) as well as their transformation products were measured from combined mixed liquor and final effluent using ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS). This non-target analysis enabled screening of over 140 compounds in one run and semi-quantitative comparison of results (i.e., a comparison of peak areas obtained in one run) following solid phase extraction (SPE) conducted by following EPA method 1694 (25).

Samples were filter-concentrated for DNA extraction for both short- and long-read sequencing. Briefly, samples were filtered through 0.22- μm filters (Millipore, Billerica, MA, United States) until clogging, after which they were ripped into small (3-5 mm) pieces with flame sterilized tweezers and stored in microcentrifuge tubes at -20 °C until extraction. DNA was extracted using the FastDNA™ SPIN Kit for Soil (MP Biomedicals, Solon, OH, United States) following kit instructions and then stored at -20 °C.

Climate data were downloaded from the National Climatic Data Center's (NCDC) archive available through the National Oceanic and Atmospheric Administration's Climate Data Online (CDO) website (26). Rainfall and temperature daily summaries were averaged across weather stations near the WWTP (US1VAMN0040, US1VAMN0004, US1VAMN0036, USC00441692, US1VAMN0005, and US1VAMN0038).

3.4.3 Shotgun Metagenomic Sequencing Analysis

Fifty-four mixed liquor (ML) DNA extracts from representing all triplicate SBR across nine sampling dates were subject to sequencing, including forty-five final effluent (FE) DNA extracts and twelve feed samples (**Table B.3**). DNA extracts were subject to library preparation and sequencing at the Duke University Center for Genomic and Computational Biology. Extracts were sequenced using an Illumina NovaSeq 6000 platform using 2×150 paired-end reads targeting approximately 5 Gbp per sample.

DNA extracts were also subject to long-read metagenomic sequencing using a minION sequencer (Oxford Nanopore). The ligation sequencing (SQK-109) and native genomic DNA barcoding (NBD-104) kits were used following manufacturer's recommendations (protocol version NBE_9065_v109_revAD_14Aug2019). Cumulatively, Illumina short-read and nanopore long-reads sequencing yielded 6.5×10^{11} basepairs of metagenome sequence data.

Annotation of short reads for taxonomy was carried out using mmseqs (27) against GTDB (28,29)) and for ARGs using CARD 3.0.7 (30). Diamond blastx was applied for alignment, with cutoffs of $\geq 80\%$ identity and $E < 10^{-10}$. Co-assembly of short and long reads was carried out using OPERA-MS (31) (using MEGAHIT (32) for initial contig generation) as well as HybridSPAdes (33). Multiple coassemblies were generated using OPERA-MS, including: (1) all samples; (2) samples representing specific experimental conditions (e.g., all of 10%); (3) samples representing specific stages of treatment (e.g., all mixed liquor samples); (4) all-condition samples (e.g., all of the 10% samples across all stages of treatment). The merged all-sample coassembly was mapped using both minimap2 (34) and bbmap (35), while the rest of the assemblies were mapped to using only minimap2. From contigs, open reading frames were predicted using prodigal (-p meta) and the resulting sequences were annotated against CARD ($\geq 90\%$ identity and $E < 10^{-10}$). Taxonomy of contigs were assigned using mmseqs2 (--tax-lineage 1 --majority 0.7 --vote-mode 1 --lca-mode 3 --orf-filter 1) against GTDB (v202 released April 27, 2021).

Clinically-relevant ARGs were selected from a list of 953 genes (**Table B.4**), revised from (36)). Genera of bacteria known to be important to WWTP processes (**Table B.5**) annotated to GTDB were selected from a combination of sources (37–41); the classification guide of functional guilds was used from the Microbial Database for Activated Sludge (MiDAS) (38). Genera that may contain pathogenic species (i.e., potential pathogens) were selected from a previously curated list (42) (**Table B.6**).

3.4.4 Statistical Analyses

All statistical tests were conducted in R (v 4.1.3) (43) using the rstatix package (44) and an α of 0.05 (p-value < 0.05) as the level of significance, unless otherwise specified. Metagenomic sequencing data were found to be non-normally distributed by the Anderson-Darling test (45). Physiochemical parameters were tested for normality using Shapiro–Wilks (46). Normally-distributed data were analyzed using ANOVA (47) *post-hoc* Tukey HSD test (48,49) and non-normally distributed data using Kruskal-Wallis *post hoc* Dunn's test (50). Samples compared over time were considered dependent if the sampling dates were within five days of each other (given the 5-day SRT, **Table B.7**). For the same SBR, the mixed liquor and final effluent samples were considered dependent because the final effluent is a subfraction of the mixed liquor (i.e., the portion that did not settle).

Alpha diversity was determined using the Shannon diversity index, which describes the diversity within a sample by taking proportional abundance into account. Significant differences between Shannon diversity index scores were determined using Wilcoxon rank sum test

(Bonferroni-Holm corrected). Beta diversity, measuring differences between samples, was determined via Bray-Curtis dissimilarity matrices and visualized using non-metric multidimensional scaling (NMDS); significant differences between groups were calculated using Permutational Multivariate Analysis of Variance (PERMANOVA). All ordination analyses were done using the R package *vegan* (51).

A Wilcoxon rank sum was used to determine which genes significantly increased or decreased relative to the three sampling dates prior to dosing HWW (-5, -3, -1 day(s) pre dosing) to the final three sampling days (9, 13, and 16 days post dosing). The 10% condition was compared with control (0%) to account for variability associated with the influent sewage from the WWTP that was applied to the SBRs. The (graphed) magnitude of change was based on absolute difference and the change was calculated based on difference between the average three prior and three final sampling dates. All figures were generated in R using *ggplot2* (52) and *circlize* (53).

3.5 Results

3.5.1 Feed

Alkalinity, COD, and sCOD were not different between the 0, 1, and 10% SBR feeds. The only physiochemical parameter that was statistically different was ammonia, which was higher in the 10% vs 0% feed (ANOVA *post hoc* Tukey, $p = 0.0396$). Ammonia was also higher in the 100% HWW when compared with municipal wastewater (115 mg/L vs 35 mg/L on day 6). Variability in the HWW and municipal sewage stock feeds to the SBRs was also tracked with time. COD and sCOD in the HWW was noted to decrease from 1/30/20 to 3/31/20 were 15.5% and 45.4%, respectively. Seasonal variability (e.g., temperature) and dilution by rainfall were monitored since they can impact wastewater characteristics as well as the decay kinetics of some pathogens found in sewer systems (54). There was some variability in physicochemical parameters with time consistent with expectations for seasonable variability (e.g., temperature and rainfall), but in the weeks before and after dosing they were fairly constant and did not appear to impact the feed (**Figure 3.2**).

Metagenomic sequencing of the feed revealed that total ARGs were similar in relative abundance for 10% and 0% SBRs (0.51 ± 0.03 and 0.54 ± 0.06 gene copies/16S rRNA gene copies, respectively). As expected, the fresh HWW had a greater abundance of total ARGs (0.83 ± 0.03). The five most abundant resistance classes across all feeds were macrolide/lincosamide/streptogramin (MLS), multidrug, beta-lactam, tetracycline, and aminoglycoside. The greater abundance of ARGs in the HWW was examined further for day six which revealed a slightly greater total abundance of shared genes (i.e., genes present in both 0% and 10% feed) and a greater number of unique genes present in the 10% versus 0% (259 versus 21, respectively). The expected abundance of shared ARGs was calculated using the 100% HWW sequencing results and therefore does not account for the potential degradation of genes during storage. The 10% feed calculations were done on a gene-by-gene basis then graphed based on resistance class. The calculated 10% feed had similar abundances of total ARGs and their respective resistance classes, however, individual genes varied up to 180% greater or less than the measured abundance.

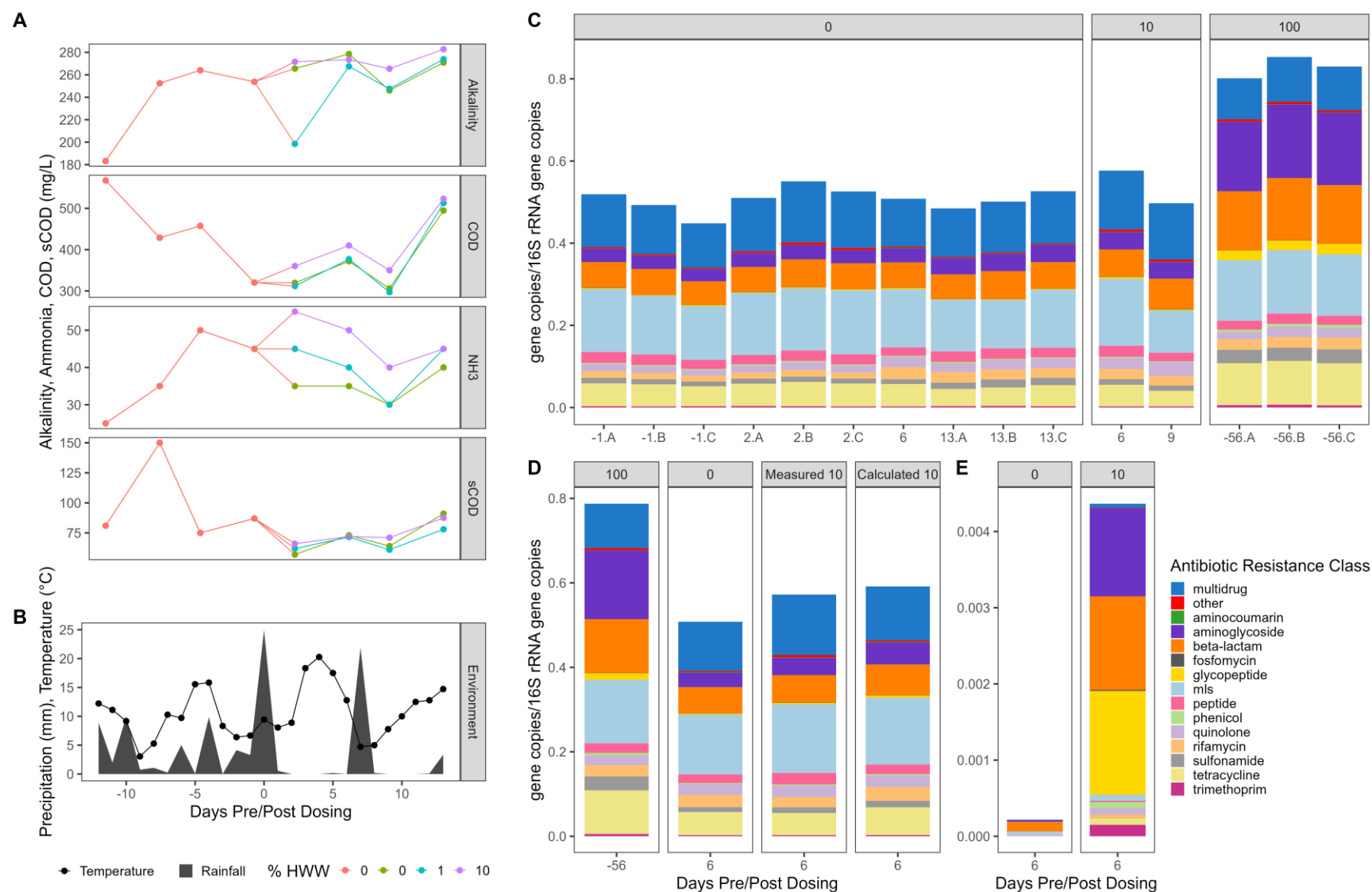


Figure 3.2 A) SBR feed physiochemical parameters, B) climate data, C) relative abundances of total annotated ARGs in the feeds (x-axis A, B, and C reflect extract replicates), D) measured and calculated shared ARGs for 10% HWW along with 0 and 100% HWW E) unique ARGs on day six for 0 and 10% HWW feed.

Taxonomic composition at the genus level was found to be distinct between the feeds themselves and within the SBRs. First, the percentages of genera that were unclassified were lower in the feeds, especially the 100% HWW ($12.1 \pm 0.22\%$), in comparison to mixed liquor and final effluent (56.6 ± 3.6 and $66.1 \pm 4.3\%$, respectively) (unclassified in 0% feed: $33.0 \pm 5.4\%$, 10% feed: $43.1 \pm 3.9\%$). The most abundant taxa in both 0% and 10% were *Acidovorax*, *Acinetobacter*, *Aeromonas*, *Aliarcobacter*, *Bacteroides*, *Flavobacterium*, and *Microthrix*. The 100% HWW had similarly greater abundances of *Acidovorax*, *Acinetobacter* and *Bacteroides*; however, the other top most abundant genera were *Agathobacter*, *Alistipes*, *Blautia*, *Faecalibacterium*, and *Prevotella*.

3.5.2 SBR Performance

pH, SVI (mg/L), MLSS (mg/L), MLVSS (mg/L), effluent sCOD (mg/L) and effluent TSS (mg/L) measurements are reported in **Figure B.1**. Neither Kruskal-Wallis nor ANOVA tests revealed that any of the SBRs performed differently. The only apparent difference was that, on day seven, the 10% SBR did begin to diverge from the other conditions with a slightly lower pH and higher MLSS and MLVSS. On four dates (two prior to dosing HWW and two after) the SBRs were tested for ammonia, sCOD, and alkalinity (**Figure B.2**) during the react period. Each of the parameters decreased on average from the start of feeding compared to ten hours later. The greatest decline was measured for ammonia ($93.7 \pm 4.0\%$) followed by alkalinity ($67.7 \pm 13.3\%$). sCOD was more variable, but on average still decreased ($26.9 \pm 30.1\%$). None of the parameters measured were different across SBRs for the four test dates. However, ammonia had a higher starting concentration in the 10% SBRs on the first date after starting HWW dosing, though this was not statistically different (62 mg/L in 0%, 97 mg/L in 1% and 165 mg/L in 10%).

The most abundant genera in the mixed liquor (*Microthrix*, *Streptomyces*, *Rhodococcus*, and *Rubrivivax*) and final effluent (*Rhodococcus*, *UBA11063*, *Microthrix*, *Streptomyces*, and *Mycolicibacterium*) were the same for the 10% and 0% conditions. Of interest was whether HWW contributed any specific genera of bacteria that could be problematic to SBR performance, or otherwise negatively impact important functional groups of bacteria. Accordingly, short read sequence annotations were examined for putative ammonia-oxidizing bacteria (AOB), nitrite-oxidizing bacteria (NOB), and filamentous bacteria (**Figure 3.3**). The 100% HWW was found to contain lower mean abundances of all of these taxa ($2.69 \pm 4.84\%$). The abundances in the 0% feed were variable over time. However, this variability was not due to the addition of HWW, as 0% and 10% feeds were similar (5.34 ± 8.53 and $6.72 \pm 10.7\%$, respectively). As expected, the mixed liquor contained greater mean abundances of these organisms than the final effluent (5.16 ± 6.88 and $4.32 \pm 5.70\%$, respectively). The seven most abundant genera were the same for mixed liquor and final effluent (0% and 10%): *Gordonia*, *Microthrix*, *Mycobacterium*, *Nitrosomonas*, *Nitrospira_A*, *Nocardia*, and *Rhodococcus*. These taxa were similarly dominant in feed samples except for *Gordonia*. In the 100% HWW *Nitrospira* and *Microthrix* were not within the top most abundant genera.

3.5.3 Comparison of Mixed Liquors Across the SBR Conditions

Of the 149 PPCPs (**Table B.8**) and metabolites measured in combined mixed liquor and final effluent, three were different across SBRs based upon Kruskal-Wallis *post hoc* Dunn's test: diclofenac 4-hydroxy ($1\% < 10\%$ $p = 0.043$), diclofenac ($0\% < 10\%$ $p = 0.034$), and norfentanyl ($1\% < 10\%$ $p = 0.046$). Across each SBRs and all dates, the four most abundant PPCPs were: desmethylvenlafaxine, lidocaine, tramadol, and venlafaxine. No antibiotics were significantly

different across SBRs. The top five most abundant antibiotics across most dates and SBRs were azithromycin, ciprofloxacin, clarithromycin, thiabendazole, and trimethoprim (**Figure B.3**).

Of the 29 soluble elements measured in mixed liquor, four were significantly different across SBRs after HWW dosing began. Compared using Kruskal-Wallis *post hoc* Dunn's test differences were seen in barium (0% greater than 10%, $p = 0.0007$, 1% greater than 10%, $p = 0.002$), sulfur (10% greater than 0%, $p = 0.002$, 10% greater than 1%, $p = 0.007$), silicon (0% greater than 10%, $p = 0.001$, 1% greater than 10%, $p = 0.044$), and uranium (1% greater than 10%, $p = 0.001$). The five most abundant elements found across all mixed liquor SBRs were: calcium, chlorine, potassium, magnesium, and sodium. Elements that were more abundant in the 10% feed than 0% or 1% were silver, aluminum, chromium, copper, iron, lead, and tin (**Figure B.4**).

3.5.4 Taxonomic Groups Containing Pathogens

Genera of bacteria commonly found in sewage and known to contain human pathogens were selected from a list of 140, of those 134 (95.7 %) were found with an abundance of $>0.01\%$ of the microbial community in the samples using the short read data (**Figure B.5**). The abundance of these genera in the feed ($22.3 \pm 4.28\%$) was greater than that in the mixed liquor or final effluent (10.2 ± 1.02 and $14.7 \pm 3.20\%$, respectively). *Bacteroides* and *Prevotella* were greater in the 100% HWW (8.9 ± 0.13 and $6.0 \pm 0.08\%$, respectively) than the 0% and 10% feeds (on average 2.1 ± 0.31 and $0.9 \pm 0.28\%$, respectively). On the other hand, abundances of *Acinetobacter* and *Aeromonas* were higher in the 0% and 10% feeds (**Table 3.1**). The top five most abundant pathogen associated genera in the feed were: *Acinetobacter*, *Bacteroides*, *Aeromonas*, and *Prevotella*. The top five pathogen-associated genera were the same in the mixed liquor and final effluent: *Rhodococcus*, *Acinetobacter*, *Burkholderia*, *Mycobacterium*, and *Nocardia*. The abundances of *Rhodococcus* and *Acinetobacter* were greater in the final effluent than in the mixed liquor (*Rhodococcus* ML: 2.3 ± 0.72 , FE: $3.9 \pm 2.2\%$; *Acinetobacter* ML: 0.78 ± 0.42 , FE: $2.3 \pm 2.1\%$).

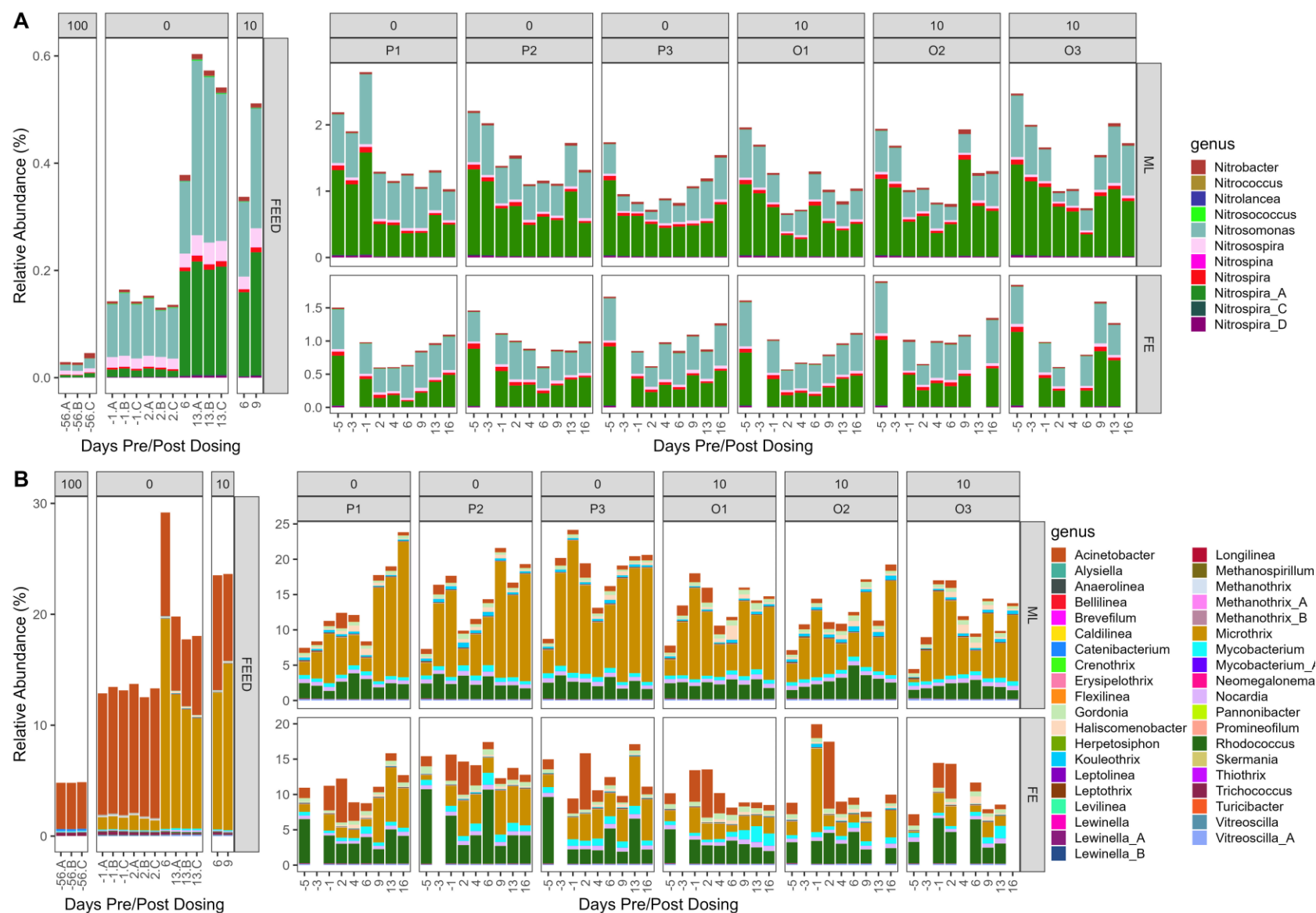


Figure 3.3 Relative abundances of genera that have the potential to (A) nitrify and (B) form flocs. Triplicate SBR denoted as O1-3 for 10% and P1-3 for 0% HWW fed SBRs.

Table 3.1 Top ten most abundant potentially pathogenic genera in the feeds and HWW. Abundance is the mean of the percent relative abundance across all samples for the given condition (dates or replicates). The first ten genera listed are the most abundant potential pathogens in each of the feeds. Below the dashed line are the genera that were not in the top 10 for that sample/group but do correspond to the top 10 for one/both of the other samples.

0% HWW			10% HWW			100% HWW		
Genus	Mean	SD	Genus	Mean	SD	Genus	Mean	SD
<i>Acinetobacter</i>	9.68	2.23	<i>Acinetobacter</i>	9.09	1.78	<i>Bacteroides</i>	8.86	0.13
<i>Aeromonas</i>	3.91	1.00	<i>Aeromonas</i>	2.90	0.41	<i>Prevotella</i>	6.00	0.08
<i>Bacteroides</i>	2.53	0.58	<i>Bacteroides</i>	1.73	0.03	<i>Acinetobacter</i>	4.18	0.02
<i>Prevotella</i>	1.16	0.45	<i>Prevotella</i>	0.69	0.10	<i>Aeromonas</i>	1.28	0.04
<i>Stenotrophomonas</i>	0.28	0.03	<i>Comamonas</i>	0.56	0.03	<i>Bifidobacterium</i>	1.03	0.05
<i>Brevundimonas</i>	0.27	0.10	<i>Stenotrophomonas</i>	0.34	0.07	<i>Streptococcus</i>	1.01	0.06
<i>Bifidobacterium</i>	0.23	0.06	<i>Brevundimonas</i>	0.29	0.01	<i>Escherichia</i>	0.95	0.02
<i>Burkholderia</i>	0.20	0.08	<i>Bifidobacterium</i>	0.27	0.09	<i>Stenotrophomonas</i>	0.64	0.03
<i>Citrobacter</i>	0.19	0.03	<i>Burkholderia</i>	0.22	0.05	<i>Collinsella</i>	0.54	0.02
<i>Escherichia</i>	0.17	0.07	<i>Achromobacter</i>	0.17	0.03	<i>Ochrobactrum</i>	0.49	0.05
<i>Achromobacter</i>	0.14	0.05	<i>Citrobacter</i>	0.17	0.01	<i>Comamonas</i>	0.26	0.01
<i>Collinsella</i>	0.14	0.04	<i>Collinsella</i>	0.14	0.05	<i>Citrobacter</i>	0.23	0.01
<i>Comamonas</i>	0.11	0.01	<i>Streptococcus</i>	0.11	0.03	<i>Achromobacter</i>	0.09	0.00
<i>Streptococcus</i>	0.10	0.02	<i>Escherichia</i>	0.09	0.01	<i>Brevundimonas</i>	0.08	0.01
<i>Ochrobactrum</i>	0.02	0.00	<i>Ochrobactrum</i>	0.05	0.01	<i>Burkholderia</i>	0.07	0.00

3.5.5 Antibiotic Resistance Genes

ARG Shannon diversity was highest for fresh HWW (4.8 ± 0.02). Both the 0% and 10% feeds had similar diversities (4.5 ± 0.06 and 4.6 ± 0.05 , respectively). Shannon diversity of the feeds was greater than that for either mixed liquor or final effluent (**Figure 3.4 A**). The mean alpha diversity was different between mixed liquor and final effluent for both 0% and 10% SBRs (ML: 4.3 ± 0.07 and FE: 4.1 ± 0.20 , respectively, Wilcox $p < 0.05$). ARG alpha diversity of mixed liquor was slightly greater and less variable over time. No difference in alpha diversity was found between the 0% and 10% conditions for either the mixed liquor or final effluent (Wilcox, $p > 0.05$) (**Table B.9**).

Short read metagenomic sequencing revealed the most abundant antibiotic resistance classes for mixed liquor and final effluent were multidrug, rifamycin, macrolide/lincosamide/streptogramin resistance (MLS), aminoglycoside, and sulfonamide (**Figure 3.4 B**). Trends in resistance classes were similar between 0% and 10% and between mixed liquor and final effluent. However, NMDS analysis revealed some differences among the Bray-Curtis dissimilarity matrices. Differences between groups were determined by PERMANOVA ($R^2 = 0.56932$ with feed or $R^2 = 0.32107$ without feed, p -value = 0.001) indicating that approximately 32% of the variance can be explained when examining samples grouped by before/initial/after, HWW percentage, and by mixed liquor or final effluent (~57% when the feed is included) (**Figure B.6**).

While the 0% and 10% conditions were very similar according to the broad-level comparison provided by NMDS analysis, there were differences between certain ARGs. Using unadjusted p -values, there were 53 ARGs with $p < 0.05$ across mixed liquor and/or final effluent,

to decrease the rate of false positives the p-value was adjusted (55) to account for multiple comparisons within type (mixed liquor and final effluent) and category (before and after) this increased the p-value and reduced the number of ARGs to 21 (**Figure 3.4 C**). ARGs that varied ($p_{adj} < 0.05$) belonged to aminoglycoside, multidrug, rifamycin, sulfonamide, tetracycline, peptide, fosfomycin, MLS, and other resistance classes. When applying a stricter cut-off, as has been advised by some to better control for false discovery rates that may arise due to the complexity of metagenomic data and the high number of unique genes (56), only three ARGs were identified that were different between the 0% and 10% mixed liquor: *tet(V)*, *mexQ*, and *TriC*, conferring resistance to tetracycline, multidrug, and other (specifically disinfecting agents and antiseptics). Regardless of the p-value cutoff chosen, the same trend is evident that there is some variability between 0% and 10% SBRs prior to dosing (when both were receiving 0%) and that this variability increased after dosing with more genes increasing in the mixed liquor of the 10% HWW condition (**Figure B.7**). Because the Wilcox test excluded genes that were not detected, this test compared only the shared ARGs. Six of the ARGs that were found to be different between 0% and 10% also appeared within the top 30 most abundant shared ARGs. ARGs that were highly abundant in 0% and 10%: *msrE*, *Bifidobacterium adolescentis rpoB* mutants conferring resistance to rifampicin, *sul1*, *ugd*, and *sul2*. The sixth gene was only present in the 30 most abundant ARGs for the 10% SBRs: *MexF* (57th most abundant in the 0%).

The higher abundances of *mexQ* and *TriC* may be due to their greater abundances in the 10% feed. All of the feeds contained lower abundances of *tet(V)* which was more abundant in the SBR mixed liquors, especially the 10% (**Figure B.8**), however, tetracycline was rarely detected in the SBRs (**Figure B.3 D**). Contigs with these ARGs revealed that *TriC* was associated with *Rhodobacteraceae* and *Pseudomonadaceae*; *mexQ* was associated with more taxa including a good deal of unclassified taxa and *Pseudomonadaceae* but also *Burkholderiaceae*, *Caulobacteraceae*, and *Rhodobacteraceae*. There were no assembled contigs with *tet(V)*.

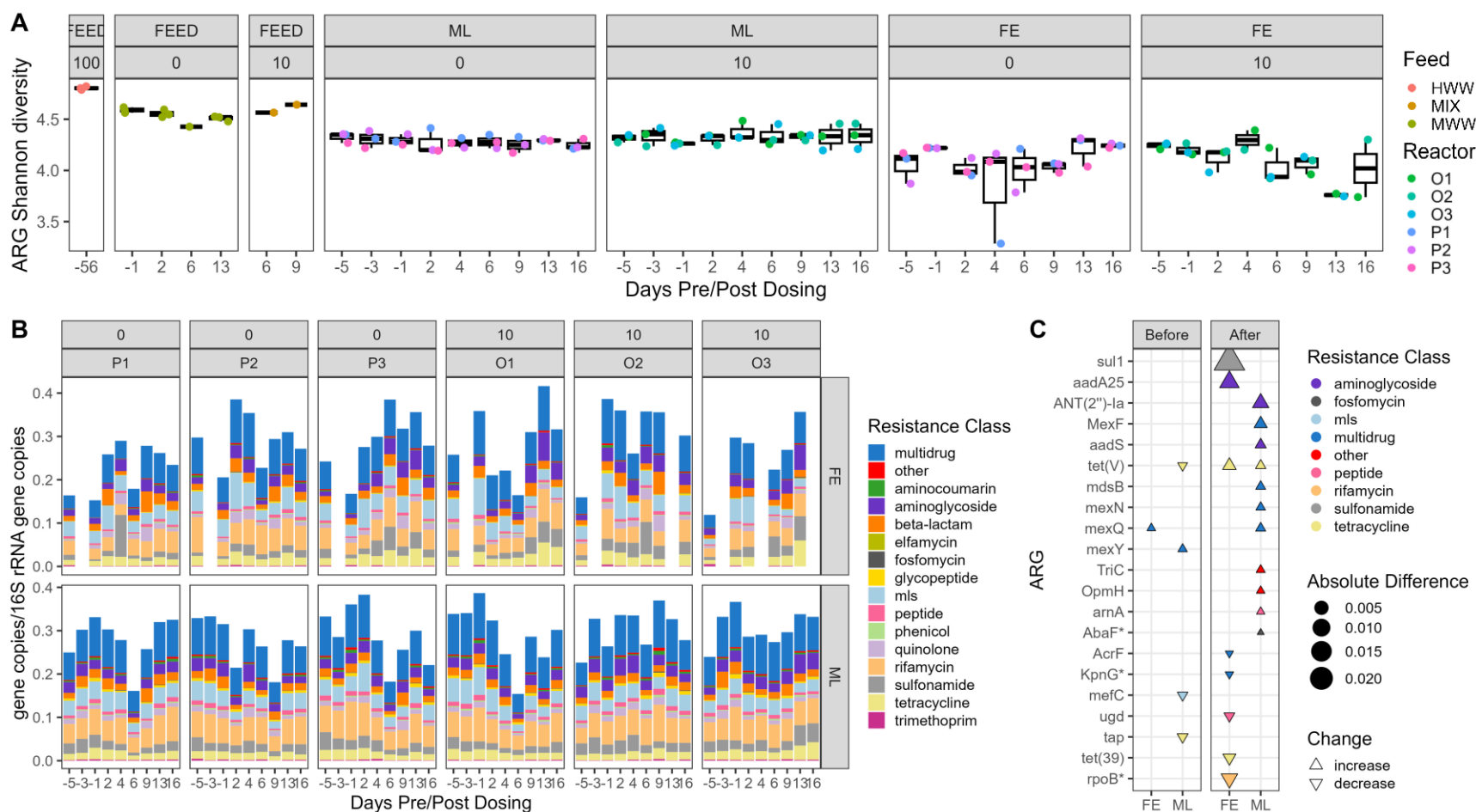


Figure 3.4 S Shannon diversity and total relative abundances of ARGs and corresponding differences observed between the 0% and 10% SBRs. (A) Shannon diversity of ARGs, reactors: MIX = 10% feed, MWW = 0% feed, O1-3 = 10% SBR, P1-3 = 0% SBR (B) Total ARG (i.e., all ARGs detected) relative abundance (ARG copies per copies of 16S rRNA normalized as previously described Li and colleagues (57)) grouped by antibiotic resistance class. MLS indicates resistance to macrolides, lincosamides, and streptogramins and (C) absolute abundance of ARGs that were different between 0% and 10% SBRs (Wilcox p.adj<0.05) before and after HWW dosing. *ARG names were shortened from: *Acinetobacter baumannii* AbaF, *Klebsiella pneumoniae* KpnG, and *Bifidobacterium adolescentis* rpoB mutants conferring resistance to rifampicin.

3.5.6 Clinically-Relevant ARGs and Associated Taxa

When examining the clinically-relevant ARGs there were 225 genes detected in the feed, ML, and FE. Most of the clinically-relevant ARGs that were detected were found to convey resistance to beta-lactams followed by glycopeptide, peptide, and quinolone resistance (**Figure B.9**). ARGs encoding resistance to glycopeptide antibiotics (specifically vancomycin) were prominent in the 100% HWW and not found in the 0% feed, indicating strong association of this form of resistance with HWW. Glycopeptide resistance was determined through detection of *vanA*, which was found in the 0% mixed liquor and final effluent (even though it was not detected in the 0% feed) though less frequently than in the 10% (for mixed liquor) (**Figure 3.5**).

Hybrid assembled contigs were used to determine which taxa harbored clinically-relevant ARGs. Resistance to beta-lactams (specifically cephalosporin and carbapenams) was found to be associated with a variety of taxa primarily belonging to the class Gammaproteobacteria. ARGs associated with fluoroquinolone antibiotics appeared in *Alteromonadaceae*, *Enterobacteriaceae*, and *Vibrionaceae*. ARGs encoding resistance to peptide antibiotics were found in *Aeromonadaceae*, *Burkholderiaceae*, and *Rhodocyclaceae*. ARGs encoding resistance to glycopeptide antibiotics were only found in *Enterococcaceae*.

The contigs containing *vanA* were most frequently encountered in the 100% HWW. There were two contigs containing *vanA* that were only found in the 10% SBRs (mixed liquor and/or final effluent) and another two were found in mixed liquor or to coassembly (which can contain all samples). Because neither final effluent nor 0% appeared as separate contigs, it might indicate that these samples did not assemble separately into contigs (though they may be within the coassembly or mixed liquor contigs) indicating that perhaps the HWW and 10% are more distinct reservoirs of *vanA* (**Figure 3.5**).

Vancomycin antibiotic was detected more frequently in the 1% and 10% SBRs after HWW dosing and on average at higher concentrations in the 10% SBRs (**Figure B.3 C**). HWW and WWTP influent samples collected from the same hospital and WWTP prior to starting the SBR experiment revealed that vancomycin was detected at each sampling date in the HWW, and was on average over 500× lower in the municipal wastewater that also received hospital wastewater and almost 3000× lower in the WWTP that did not receive HWW indicating that while vancomycin was coming from the hospital, it was very diluted by the time it came to the WWTP influent (**Table B.10**).

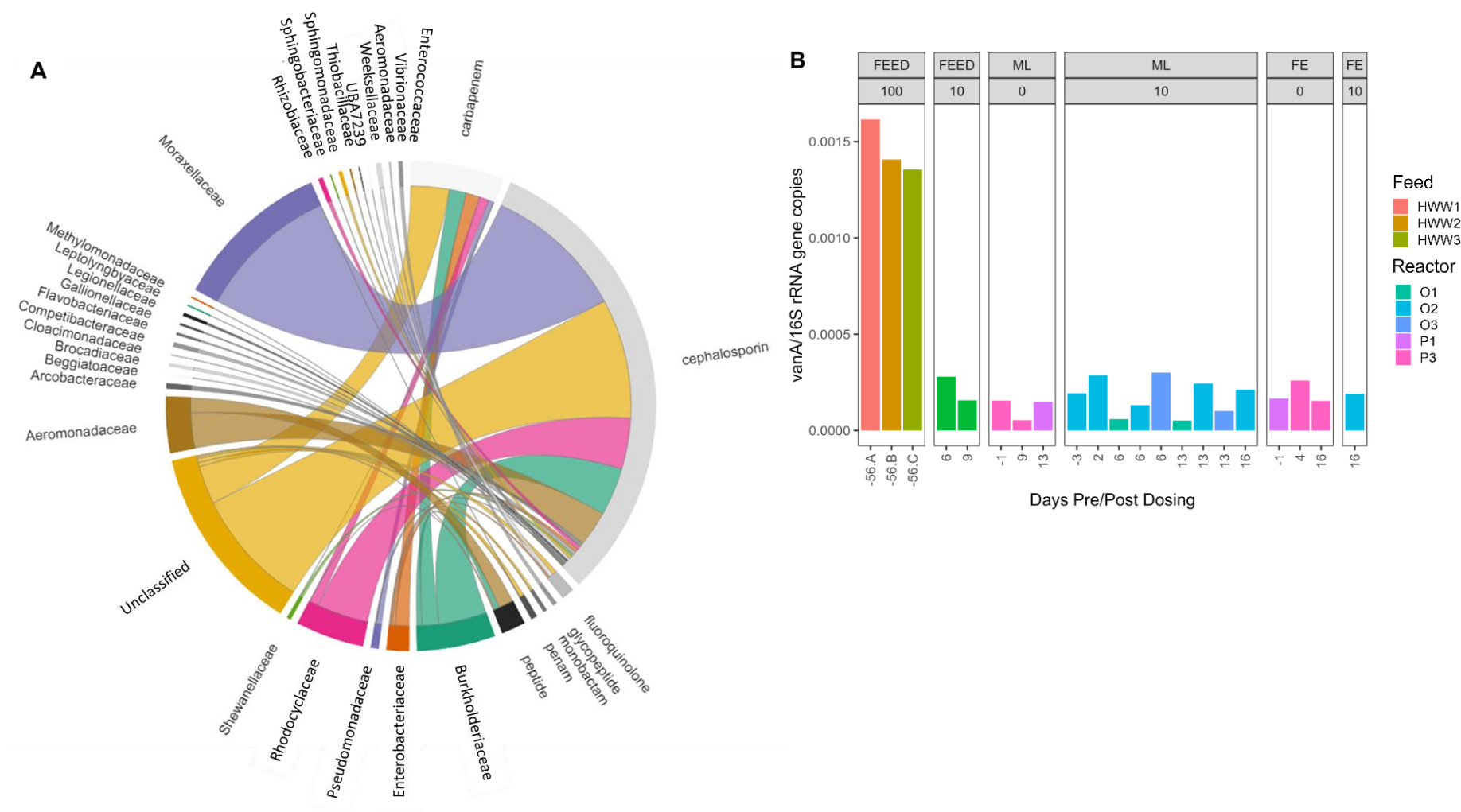


Figure 3.5 (A) Contigs found to contain clinically-relevant ARGs and their corresponding taxonomic classifications. There were a number of contigs that could not be assigned a taxonomic classification (Unclassified) and (B) all *vanA* detected in samples. *vanA* was detected in both 10% and 100% HWW. Although *vanA* was not detected in the 0% feed it was detected in two of the 0% SBR (P1, P3) in both mixed liquor and final effluent though less frequently than in the 10% (O1-O3) where it tended to be in the mixed liquor.

3.6 Discussion

SBRs have been used extensively to treat various types of wastewater (i.e., industrial, municipal) (58,59) across a range of sizes (lab-scale experiments to full-scale wastewater treatment plants) (60,61). Bench-scale SBRs have been used previously to examine the effects of different inocula (hospital, meat processing, or municipal WWTP waste) (62) as well as the effects of antibiotics on microbiota and ARGs (63).

3.6.1 Differences in HWW and SBR Feed

The physiochemical parameters (e.g., ammonia, alkalinity, COD, sCOD) indicated that overall the HWW was of higher strength than the domestic wastewater that served as the background feed to the SBRs. Lower ammonia and COD in the domestic wastewater could reflect degradation and/or dilution as it traveled through the sewers to reach the collection point at the WWTP. Higher ammonia concentration of HWW was also found in other studies. For example, Carraro and colleagues (8) found from a review of the published literature that concentrations of ammonium ion (NH_4^+) were generally higher in HWW than the corresponding urban wastewater inlet.

The relative abundances of genera were different for the feeds, especially the 100% HWW. The 100% HWW had lower abundances of filamentous bacteria and higher abundances of potential pathogens that tend to be found in the human gut microbiome. This finding may be indicative of shifts in the environment as aerobic bacteria thrive over the anaerobic bacteria more commonly found in fecal (64) or HWW (65) samples. The 0% and 10% feed had similarly abundant genera suggesting that the HWW did not contribute substantially to the diversity of these taxa. This result has also been shown from HWW and municipal wastewater samples collected in the field (14).

The feed contained greater abundances of ARGs, especially for the 100% HWW, as was expected based on previous studies. A literature review showed in 81% of studies ($n=37$) that hospital wastewater contained higher frequencies of antibiotic resistance determinants (from a list of 15 ARGs) when compared with the corresponding community wastewater (66). Additional studies have found that HWW contains greater abundances of ARGs conveying resistance to carbapenems, sulfonamides, and tetracyclines and that these abundances vary by country (67).

3.6.2 Activated Sludge is a Distinct Ecosystem that was Only Subtly Affected by the HWW

The 10% feed appeared to be impacted by the HWW as evidenced by higher concentrations of ammonia and certain elements (e.g., silver, aluminum, chromium, copper, iron, lead, and tin), as well as a greater number of unique ARGs. These differences were also reflected in the SBRs, as the corresponding parameters increased in the 10% SBRs as well. However, none of these changes resulted in a disruption of efficacy of wastewater treatment for removal of organic carbon and nitrogen. Similarly, there was greater relative abundances of potential pathogens in the feeds, but these same pathogens were not dominant in the mixed liquor or final effluent. The activated sludge process is recognized for its capacity to reduce pathogen concentrations through degradation and predation (68).

The robustness of the mixed liquor for attenuating markers of antibiotic resistance is further evident through the stability of the Shannon diversity for both ARGs and taxa over time. The relative abundances of ARGs and genera were also very stable over time (see stacked bar charts). Another study found that the shared ARGs in activated sludge were highly similar even though the different WWTPs studied received wastewater from different sources (domestic, landfill leachate, and/or industrial waste) (69). Additional studies have found that WWTP activated sludge

was stable over time and for different regions of the world (Asia and North America) at different seasons (i.e., temperatures) with the same core functional bacteria being present in activated sludge treating municipal and industrial wastewater (70).

3.6.3 Impact of Hospital Wastewater

The ARGs that were different between SBRs were higher in the 10% mixed liquor after HWW dosing. These increases may correspond to their increased abundance in the feed dosed with HWW (*TriC* and *mexQ*). However, *tet(V)* did not follow this pattern as the ARG was lower in all of the feeds than the SBRs, this may be due to the greater abundance of *Mycolicibacterium*, which harbors *tet(V)* as part of its intrinsic resistome (71), in the SBRs compared to the feed. The relative abundance of *Mycolicibacterium* was not different between the 0% and 10% ML, however, indicating there may have been additional selection pressure resulting in the increase in *tet(V)* (though tetracycline was rarely detected in the SBRs and was not higher in the HWW when compared with the WWTP that receives that waste).

Vancomycin and vancomycin ARGs were detected in the HWW at a greater abundance than in domestic wastewater. Notably, the assembled data confirmed association of *vanA* with *Enterococcus*, which are generally known to be a carrier of this gene (72). The CDC reported data from the National Healthcare Safety Network that from isolates collected from patients 2011-2020 3.5% of *Enterococcus faecalis* (73) and 65.7% of *Enterococcus faecium* (74) isolates were resistant to vancomycin. Varela and colleagues (75) similarly found that HWW had 9× more culturable vancomycin resistant Enterococci (predominantly *E. faecalis*) than the wastewater treatment plant influent. Additional studies have found some ambiguity whether vancomycin resistance was higher due to HWW in wastewater, however, the authors speculate this may be due to differences in sampling or culturing methods (66). Furthermore, variability in vancomycin usage in a hospital and/or community will impact its detection in wastewater. This study supports using *vanA* as a monitoring target for clinically-relevant ARGs (76) as it is often found to be associated with hospitals (77). *vanA* was detected less frequently in final effluent samples, however, the presence of clinically-relevant ARGs has been found in other studies in treated municipal wastewater highlighting that they may still be released into the environment (65).

Statistical analyses revealed more variability between the 0% and 10% SBRs after HWW dosing across all sequencing data (ARGs, taxa, and not included metabolic pathways, and metal and biocide resistance genes). While some of these changes may be attributable to HWW (discussed below), this variability may also be due to regional differences (i.e., rural vs urban, Midwest vs South). Regarding differences in ARGs, the rates of prescriptions per 1,000 persons are similar for Virginia and Illinois (591 and 606, respectively) (78). However, studies have found that there may be differences in urban and rural prescription practices (79). This is reflected in the slightly greater rates of antibiotic prescriptions in Roanoke than Chicago (80.7 and 74.7 per 100 members, respectively) though broad spectrum antibiotic prescriptions were similar for both cities (40.2 and 39.7, respectively) (80). Regional differences must also be considered as the Southern US reportedly has higher antibiotic prescription rates than other areas in the US (81,82). Therefore, while the measured differences observed may reflect the addition of HWW, it cannot be ruled out that these differences were driven by variability in the underlying resistomes.

3.6.4 Limitations

The SBRs were operated until they reached stable conditions before dosing in HWW. This was necessary to ensure that any changes detected were due to the addition of HWW. The time

required to reach steady state; however, resulted in the HWW being stored for two months prior to use. While it would have been ideal to utilize fresh hospital waste this was not possible since at the start of these experiments the hospital was already treating COVID-19 patients and it was not possible for them to re-sample the wastewater due to safety concerns. The pandemic also shortened the run time of the SBRs. Ideally, they would have been run until they reached a new steady state, however, due to limited personnel and the need to shut down the university laboratories for non-essential tasks it was necessary to stop the experiment after gathering just over two weeks' worth of data.

The experimental design does not take into account the mixing and degradation that would normally take place in a sewer. Alteration of the HWW during transportation or storage may have resulted in changes in the microbial community that would not occur in a normal sewer thereby changing which ARGs or trends were detected.

This study was designed with collaborators to try and determine if hospital/wastewater treatment plant effluent was contributing to antibiotic resistance levels observed in the community. Therefore, hospital wastewater was collected from a Midwestern urban hospital, however, to be able to more clearly determine the impacts of HWW without having to account for differences in the underlying community resistome future experiments should use HWW from a local hospital. There are also changes in prescription practices and wastewater characteristics over time, therefore, it would be ideal to conduct this experiment for longer to be able to determine the impacts of environment and season.

3.7 Conclusions

This study showed that the HWW examined in this study was more different from domestic wastewater with greater concentrations of total ARGs and relative abundances of *Bacteroides* and *Prevotella*. While the composition of HWW is more different according to NMDS analysis, it is greatly diluted when mixed with domestic wastewater, even at a realistically uncommon high 10%. Similar to what other studies have shown regarding activated sludge, mixed liquor was resistant to change as the SBRs dosed with HWW remained similar to their control counterparts with respect to performance, ARGs, and relative abundances of genera. There was some elevation in the relative abundance of certain ARGs in the mixed liquor of the 10% SBR, notably resistance to vancomycin via *vanA*, which is known to be connected to clinical settings. Because WWTPs are not typically designed to remove ARB or ARGs, continued attention should be given to interactions between WWTP and hospitals to prevent the dissemination of resistance to clinically relevant antibiotics into the environment.

3.8 Acknowledgements

This project was partially supported by the National Science Foundation (NSF) Halting Environmental Antimicrobial Resistance Dissemination Partnership for International Research and Education (#1545756), an NSF Graduate Research Fellowship, and the NSF Combating Antimicrobial Resistance Research Traineeship (#2125798). The authors would also like to extend their sincere gratitude to Abe Cullom, Delaney Snead, Eldridge Hager, Ishi Keenum, Joyce Ni Zhu, Mariah Gnegy-Spencer, Storme Spencer, and Yang Song for helping process samples at the start of a pandemic. Thank you to Dr. Jeffrey Parks for his technical and laboratory assistance. Further thanks are extended to the WWTP operators for allowing the collection of influent each week, and for the hospital wastewater operators for collecting and shipping sewage.

3.9 References

1. Michael CA, Dominey-Howes D, Labbate M, Maria C, Elisabeth J, Pitout J, et al. The antimicrobial resistance crisis: causes, consequences, and management. *Front Public Heal* [Internet]. 2014 [cited 2022 Nov 21];2(145). Available from: www.frontiersin.org
2. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Sci Total Environ* [Internet]. 2013 [cited 2017 Oct 3];447:345–60. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2013.01.032>
3. Karkman A, Do TT, Walsh F, Virta MPJ. Antibiotic-Resistance Genes in Waste Water. 2017 [cited 2022 Nov 22]; Available from: <https://doi.org/10.1016/j.tim.2017.09.005>
4. Prieto Riquelme MV, Garner E, Gupta S, Metch J, Zhu N, Blair MF, et al. Demonstrating a Comprehensive Wastewater-Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environ Sci Technol* [Internet]. 2022 Dec 19;56(21):14982–93. Available from: <https://doi.org/10.1021/acs.est.1c08673>
5. Thornton CN, Tanner WD, Vanderslice JA, Brazelton WJ. Localized effect of treated wastewater effluent on the resistome of an urban watershed. *Gigascience*. 2020;9(11).
6. Rodriguez-Mozaz S, Lucas D, Barceló D. Full-Scale Plants for Dedicated Treatment of Hospital Effluents. In: *The Handbook of Environmental Chemistry* [Internet]. 60th ed. Springer, Cham; 2017 [cited 2022 Nov 18]. p. 189–208. Available from: <http://www.springer.com/series/698>
7. Emmanuel E, Perrodin Y, Keck G, Blanchard JM, Vermande P. Ecotoxicological risk assessment of hospital wastewater: A proposed framework for raw effluents discharging into urban sewer network. *J Hazard Mater*. 2005 Jan 14;117(1):1–11.
8. Carraro E, Bonetta S, Bertino C, Lorenzi E, Bonetta S, Gilli G. Hospital effluents management: Chemical, physical, microbiological risks and legislation in different countries. *J Environ Manage* [Internet]. 2016 [cited 2018 Dec 5];168:185–99. Available from: <http://dx.doi.org/10.1016/j.jenvman.2015.11.021>
9. Verlicchi P, Al Aukidy M, Zambello E. What have we learned from worldwide experiences on the management and treatment of hospital effluent? — An overview and a discussion on perspectives. *Sci Total Environ* [Internet]. 2015 May 5 [cited 2022 Nov 18];514:467. Available from: [/pmc/articles/PMC7112026/](http://pmc/articles/PMC7112026/)
10. WHO. Safe management of wastes from health-care activities [Internet]. 2014. Available from: <https://www.who.int/publications/i/item/9789241548564>
11. EPA. Local Limits Development Guidance [Internet]. 2004. Available from: https://www3.epa.gov/npdes/pubs/final_local_limits_guidance.pdf
12. DHI. Full scale advanced wastewater treatment at Herlev Hospital Treatment performance and evaluation. Hørsholm; 2016.
13. Wiest L, Chonova T, Bergé A, Baudot R, Bessueille-Barbier F, Ayouni-Derouiche L, et al. Two-year survey of specific hospital wastewater treatment and its impact on pharmaceutical discharges. *Environ Sci Pollut Res* [Internet]. 2018 Apr 17 [cited 2019 Oct 9];25(10):9207–18. Available from: <http://link.springer.com/10.1007/s11356-017-9662-5>
14. Buelow E, Bayjanov JR, Majoor E, Willems RJL, Bonten MJM, Schmitt H, et al. Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. *FEMS Microbiol Ecol* [Internet]. 2018 [cited 2022 Nov 22];94:87. Available from: <https://academic.oup.com/femsec/article/94/7/fiy087/4995906>

15. Hubeny J, Ciesielski S, Harnisz M, Korzeniewska E, Dulski T, Jałowiecki Ł, et al. Impact of Hospital Wastewater on the Occurrence and Diversity of Beta-Lactamase Genes During Wastewater Treatment with an Emphasis on Carbapenemase Genes: A Metagenomic Approach. *Front Environ Sci*. 2021 Oct 22;9:479.
16. Goodman KE, Cosgrove SE, Pineles L, Magder LS, Anderson DJ, Dodds Ashley E, et al. Significant Regional Differences in Antibiotic Use across 576 US Hospitals and 11 701 326 Adult Admissions, 2016-2017. *Clin Infect Dis*. 2021;73(2).
17. Kümmerer K. Promoting resistance by the emission of antibiotics from hospitals and households into effluent. *Clin Microbiol Infect*. 2003;9(12).
18. Nuñez L, Tornello C, Puentes N, Espigares E, Moreno E, Espigares M, et al. Hospital effluent constitutes a source of vancomycin-resistant enterococci. *Ars Pharm*. 2016;57(3).
19. World Health Organisation. Critically Important Antimicrobials for Human Medicine 6th Revision 2018 [Internet]. 2019. Available from: <https://apps.who.int/Iris/Bitstream/Handle/10665/312266/9789241515528-Eng.Pdf?Ua=1>
20. Zhang Y, Zhang T. Culturing the uncultured microbial majority in activated sludge: A critical review. *Crit Rev Environ Sci Technol* [Internet]. 2022;0(0):1–24. Available from: <https://doi.org/10.1080/10643389.2022.2077063>
21. Wright GD. The antibiotic resistome: The nexus of chemical and genetic diversity. Vol. 5, *Nature Reviews Microbiology*. 2007.
22. Irvine RL, Ketchum LH. Sequencing batch reactors for biological wastewater treatment. *Crit Rev Environ Control*. 1989;18(4).
23. Jirka AM, Carter MJ. Micro–semi-automated analysis of surface and wastewaters for chemical oxygen demand. *Anal Chem* [Internet]. 1975 [cited 2022 Jun 3];47(8):1397–402. Available from: <https://pubs.acs.org/sharingguidelines>
24. APHA-AWWA-WEF. Standard methods for the examination of water and wastewater. 23rd ed. Baird R, Bridgewater L, editors. Washington, D.C.: American Public Health Association; 2017.
25. EPA. Method 1694: Pharmaceuticals and Personal Care Products in Water, Soil, Sediment, and Biosolids by HPLC/MS/MS. 2007 [cited 2022 Feb 13]; Available from: <http://www.epa.gov/waterscience>
26. Climate Data Online (CDO) [Internet]. 2022. Available from: <https://www.ncei.noaa.gov/cdo-web/>
27. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Vol. 35, *Nature Biotechnology*. 2017.
28. Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol*. 2020;38(9).
29. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarszewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. 2018;36(10).
30. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2017;45(D1):D566–73.
31. Bertrand D, Shaw J, Kalathiyappan M, Hui A, Ng Q, Senthil Kumar M, et al. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. *Nat Biotechnol* [Internet]. 2019 [cited 2022 Feb 13];37:937–944. Available from: <https://doi.org/10.1038/s41587-019-0191-2>

32. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31(10).
33. Antipov D, Korobeynikov A, McLean JS, Pevzner PA. HybridSPAdes: An algorithm for hybrid assembly of short and long reads. *Bioinformatics*. 2016;32(7).
34. Li H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18).
35. Bushnell B. BBMap: A Fast, Accurate, Splice-Aware Aligner. In: United States. 2014.
36. Keenum I, Williams RK, Ray P, Garner ED, Knowlton KF, Pruden A. Combined effects of composting and antibiotic administration on cattle manure–borne antibiotic resistance genes. *Microbiome*. 2021;9(1).
37. Nierychlo M, McIlroy SJ, Kucheryavskiy S, Jiang C, Ziegler AS, Kondrotaitė Z, et al. *Candidatus Amarolinea* and *Candidatus Microthrix* Are Mainly Responsible for Filamentous Bulking in Danish Municipal Wastewater Treatment Plants. *Front Microbiol*. 2020;11.
38. Dueholm MKD, Nierychlo M, Andersen KS, Rudkjøbing V, Knutsson S, Arriaga S, et al. MiDAS 4: A global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. *Nat Commun*. 2022;13(1).
39. Li BB, Peng ZY, Zhi LL, Li HB, Zheng KK, Li J. Distribution and diversity of filamentous bacteria in wastewater treatment plants exhibiting foaming of Taihu Lake Basin, China. *Environ Pollut*. 2020;267.
40. You J, Das A, Dolan EM, Hu Z. Ammonia-oxidizing archaea involved in nitrogen removal. Vol. 43, *Water Research*. 2009.
41. Daims H, Lückner S, Wagner M. A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. Vol. 24, *Trends in Microbiology*. 2016.
42. Cai L, Ju F, Zhang T. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Appl Microbiol Biotechnol*. 2014;98(7).
43. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
44. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests [Internet]. R package. 2021. Available from: <https://cran.r-project.org/package=rstatix>
45. Gross J, Ligges U. nortest: Tests for Normality. R package version 1.0-4; 2015.
46. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3–4).
47. Chambers JM, Freeny A, Heiberger RM. Analysis of variance; designed experiments. In: John M. Chambers, Trevor Hastie TJH, editor. *Statistical Models in S*. Wadsworth & Brooks/Cole; 1992.
48. Miller RG. *Simultaneous Statistical Inference*. Springer. 1981.
49. Yandell BS. *Practical Data Analysis for Designed Experiments*. In Chapman & Hall; 1997.
50. Dunn OJ. Multiple comparisons using rank sums. *Technometrics*. 1964;6:241–252.
51. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. The vegan Package Title Community Ecology Package [Internet]. R package. 2022 [cited 2022 Jun 2]. Available from: <https://cran.r-project.org/package=vegan>
52. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag;

- 2016.
53. Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular visualization in R. *Bioinformatics*. 2014;30(19).
54. Guo Y, Sivakumar M, Jiang G. Decay of four enteric pathogens and implications to wastewater-based epidemiology: Effects of temperature and wastewater dilutions. Vol. 819, *Science of the Total Environment*. 2022.
55. Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics*. 2010;26(6).
56. Kaiser T, Jahansouza C, Staley C. Network-based approaches for the investigation of microbial community structure and function using. *Future Microbiol*. 2022;17(8):621–31.
57. Li B, Yang Y, Ma L, Ju F, Guo F, Tiedje JM, et al. Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J* [Internet]. 2015 [cited 2018 Aug 27];9:2490–502. Available from: <http://www.cdc.gov/narms/>
58. Keller J, Subramaniam K, Gösswein J, Greenfield PF. Nutrient removal from industrial wastewater using single tank sequencing batch reactors. *Water Sci Technol*. 1997;35(6).
59. Neczaj E, Kacprzak M, Kamizela T, Lach J, Okoniewska E. Sequencing batch reactor system for the co-treatment of landfill leachate and dairy wastewater. *Desalination*. 2008;222(1–3).
60. Pronk M, de Kreuk MK, de Bruin B, Kamminga P, Kleerebezem R, van Loosdrecht MCM. Full scale performance of the aerobic granular sludge process for sewage treatment. *Water Res*. 2015;84.
61. Wu L, Wang M, Rong L, Wang W, Chen L, Wu Q, et al. Structural effects of sulfonamides on the proliferation dynamics of sulfonamide resistance genes in the sequencing batch reactors and the mechanism. *J Environ Sci* [Internet]. 2022 Dec 19; Available from: <https://www.sciencedirect.com/science/article/pii/S1001074222006003>
62. Wittebolle L, Verstraete W, Boon N. The inoculum effect on the ammonia-oxidizing bacterial communities in parallel sequential batch reactors. *Water Res*. 2009;43(17).
63. Zhao R, Feng J, Liu J, Fu W, Li X, Li B. Deciphering of microbial community and antibiotic resistance genes in activated sludge reactors under high selective pressure of different antibiotics. *Water Res*. 2019;151.
64. Bengtsson-Palme J, Hammaren R, Pal C, Ostman M, Bjorlenius B, Flach C-F, et al. Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. *Sci Total Environ*. 2016 Dec;572:697–712.
65. Hutinel M, Larsson DGJ, Flach CF. Antibiotic resistance genes of emerging concern in municipal and hospital wastewater from a major Swedish city. *Sci Total Environ*. 2022;812.
66. Hassoun-Kheir N, Stabholz Y, Kreft JU, de la Cruz R, Romalde JL, Nesme J, et al. Comparison of antibiotic-resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: A systematic review. Vol. 743, *Science of the Total Environment*. 2020.
67. Zhang S, Huang J, Zhao Z, Cao Y, Li B. Hospital Wastewater as a Reservoir for Antibiotic Resistance Genes: A Meta-Analysis. Vol. 8, *Frontiers in Public Health*. 2020.
68. Kim TD, Unno H. The roles of microbes in the removal and inactivation of viruses in a biological wastewater treatment system. In: *Water Science and Technology*. 1996.
69. Liu Z, Klümper U, Liu Y, Yang Y, Wei Q, Lin JG, et al. Metagenomic and

- metatranscriptomic analyses reveal activity and hosts of antibiotic resistance genes in activated sludge. *Environ Int.* 2019;129.
70. Ju F, Xia Y, Guo F, Wang Z, Zhang T. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. *Environ Microbiol.* 2014;16(8).
 71. Kyselková M, Chroňáková A, Volná L, Němec J, Ulmann V, Scharfen J, et al. Tetracycline resistance and presence of tetracycline resistance determinants tet(V) and tap in rapidly growing mycobacteria from agricultural soils and clinical isolates. *Microbes Environ.* 2012;27(4).
 72. Davis BC, Keenum I, Calarco J, Liguori K, Milligan E, Pruden A, et al. Towards the standardization of *Enterococcus* culture methods for waterborne antibiotic resistance monitoring: A critical review of trends across studies. *Water Res X* [Internet]. 2022 Jan 23;17:100161. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9712764/>
 73. CDC. Vancomycin-resistant *Enterococcus faecalis* [Internet]. National Healthcare Safety Network. 2020. Available from: <https://arpsp.cdc.gov/profile/antibiotic-resistance/vancomycin-resistant-enterococcus-faecalis>
 74. CDC. Vancomycin-resistant *Enterococcus faecium* [Internet]. National Healthcare Safety Network. 2020. Available from: <https://arpsp.cdc.gov/profile/antibiotic-resistance/vancomycin-resistant-enterococcus-faecium#healthcare-facility-type>
 75. Varela AR, Ferro G, Vredenburg J, Yanik M, Vieira L, Rizzo L, et al. Vancomycin resistant enterococci: From the hospital effluent to the urban wastewater treatment plant. *Sci Total Environ.* 2013;450–451.
 76. Keenum I, Liguori K, Calarco J, Davis BC, Milligan E, Harwood VJ, et al. A framework for standardized qPCR-targets and protocols for quantifying antibiotic resistance in surface water, recycled water and wastewater. Vol. 52, *Critical Reviews in Environmental Science and Technology*. 2022.
 77. Paulus GK, Hornstra LM, Alygizakis N, Slobodnik J, Thomaidis N, Medema G. The impact of on-site hospital wastewater treatment on the downstream communal wastewater system in terms of antibiotics and antibiotic resistance genes. *Int J Hyg Environ Health* [Internet]. 2019 May 1 [cited 2019 Jun 18];222(4):635–44. Available from: <https://www.sciencedirect.com/science/article/pii/S1438463918308289?via%3Dihub>
 78. CDC. Outpatient Antibiotic Prescriptions — United States, 2020 [Internet]. 2020 [cited 2021 Dec 24]. Available from: <https://www.cdc.gov/antibiotic-use/data/report-2020.html>
 79. Clark AW, Durkin MJ, Olsen MA, Keller M, Ma Y, O’Neil CA, et al. Rural–urban differences in antibiotic prescribing for uncomplicated urinary tract infection. *Infect Control Hosp Epidemiol.* 2021;42(12).
 80. Blue Cross Blue Shield. Antibiotic prescription fill rates declining in the US [Internet]. 2017. Available from: <https://www.bcbs.com/sites/default/files/file-attachments/health-of-america-report/HoA.Antibiotics.Report.pdf>
 81. Gouin KA, Fleming-Dutra KE, Tsay S, Bizune D, Hicks LA, Kabbani S. Identifying Higher-Volume Antibiotic Outpatient Prescribers Using Publicly Available Medicare Part D Data — United States, 2019. *MMWR Recomm Reports.* 2022;71(6).
 82. CDC. Outpatient antibiotic prescriptions-United States, 2021 [Internet]. 2021. Available from: <https://www.cdc.gov/antibiotic-use/pdfs/Annual-Report-2021-H.pdf>

APPENDIX B

B.1 Supplementary Tables

Table B.1 Non-systematic literature review of the proportion of hospital wastewater (HWW) found in wastewater treatment plant influent based.

Country	Approximate HWW %	# hospitals	# beds in the hospital	WWTP average flow	Reference
Portugal	0.1%	1	1120 beds	1.1×10^6 m ³ /day	(1)
Portugal	0.1%	1	1120 beds	974000 m ³ /d	(2)
Italy	2%	1	Hospital B: 900 beds	28000 m ³ /d	(3)
Japan	0.34%	1	477	134000 m ³ /d	(4)
Poland	0.17%	Not given	Not given	96000 m ³ /d	(5)
France	0.56%	1	87 beds	4300 m ³ /d	(6)
Turkey	0.011– 3.57% in summer and 0.003–11.4% in winter	16	16 hospitals: university (903 - 1298 beds), pediatric (363 beds), general (medium 194 – 600 and small 27-103)	159,800 m ³ /day (on sampling day not average)	(7)
Portugal	1.25%	4	1456 (university hospital), 350 (general hospital), 110 (pediatric hospital), 96 (maternity hospital)	30,564 m ³ /d	(8)
Spain	0.50%	1	1000 beds		(9)
Portugal	15%	1	1,200 beds	32,500 m ³ /d	(10)
Ireland	2%	1			(11)
UK, Denmark, Spain	3%	1	450 beds UK, 1000 beds Denmark, 1395 beds, Spain	Don't think it was given	(12)
Spain	2%	2	1300 beds	54,560 m ³ /d	(13)

Table B.2 The different samples collected and how they correlate with each other (by date). Yellow shading indicates day started to dose hospital effluent, green is 2 day HRT (from dose day) and blue is 5 day HRT (from dose day). *Hourly sampling was for sCOD, ammonia, alkalinity, and DNA extraction during aeration (10am, 12pm, 4pm, 8pm).

Sample Processing	Sample Type	Dose Day																				
		-5	-4	-3	-2	-1	0	2	3	4	5	6	7	8	9	10	12	13	14	15	16	17
COD/sCOD	Feed	x				x		x				x			x			x				
	FE	x			x				x		x			x			x			x		
TSS/VSS	ML	x			x				x		x			x			x			x		
	FE	x	x		x		x		x		x		x	x		x	x		x	x		
SVI	ML	x			x				x		x						x			x		
Filtering for DNA Extraction	Feed	x				x		x				x			x			x				
	ML	x		x		x		x		x		x			x			x			x	
	FE	x		x		x		x		x		x			x			x			x	
Culturing (5 media types: data not included in this study)	Feed					x		x				x							x			
	ML					x		x				x							x			
	FE					x		x				x							x			
PPCPs	FE + ML				x						x											x
Metals	ML					x		x				x			x						x	
Ammonia	Feed	x				x		x				x			x			x				
Alkalinity	Feed	x				x		x				x			x			x				
Sample ~hourly*	ML			x										x				x				

Table B.3 Samples sequenced for shotgun metagenomic sequencing

Sample Type	Percent Hospital	SBR Replicate	Dose Day									
			-56	-5	-3	-1	2	4	6	9	13	16
ML	10%	1		x	x	x	x	x	x	x	x	x
ML	10%	2		x	x	x	x	x	x	x	x	x
ML	10%	3		x	x	x	x	x	x	x	x	x
ML	0%	1		x	x	x	x	x	x	x	x	x
ML	0%	2		x	x	x	x	x	x	x	x	x
ML	0%	3		x	x	x	x	x	x	x	x	x
FE	10%	1		x		x	x	x	x	x	x	x
FE	10%	2		x		x	x	x	x			x
FE	10%	3		x		x	x		x	x	x	
FE	0%	1		x		x	x	x	x	x	x	x
FE	0%	2		x		x	x	x	x	x	x	x
FE	0%	3		x		x	x	x	x	x	x	x
FEED		WWTP				x	x				x	
						x	x				x	
						x	x		x		x	
FEED		10% Hospital							x	x		
FEED		Fresh HWW	x									
			x									
			x									

Table B.4 Clinically relevant ARGs

CARB-1	CTX-M-54	MCR-7.1	OXA-204	OXA-355	OXA-92	SHV-3	TEM-145	TEM-88
CARB-10	CTX-M-55	MCR-8.1	OXA-205	OXA-356	OXA-93	SHV-30	TEM-146	TEM-89
CARB-12	CTX-M-56	MCR-9.1	OXA-206	OXA-357	OXA-94	SHV-31	TEM-147	TEM-90
CARB-14	CTX-M-58	mecA	OXA-207	OXA-358	OXA-95	SHV-32	TEM-148	TEM-91
CARB-16	CTX-M-59	mecB	OXA-208	OXA-359	OXA-96	SHV-33	TEM-149	TEM-92
CARB-17	CTX-M-6	mecC	OXA-209	OXA-36	OXA-97	SHV-34	TEM-15	TEM-93
CARB-18	CTX-M-60	mecD	OXA-21	OXA-360	OXA-98	SHV-35	TEM-150	TEM-94
CARB-19	CTX-M-61	mecI	OXA-210	OXA-361	OXA-99	SHV-36	TEM-151	TEM-95
CARB-2	CTX-M-62	mecR1	OXA-211	OXA-362	QnrA1	SHV-37	TEM-152	TEM-96
CARB-20	CTX-M-63	NDM-1	OXA-212	OXA-363	QnrA2	SHV-38	TEM-153	vanA
CARB-21	CTX-M-64	NDM-10	OXA-213	OXA-364	QnrA3	SHV-39	TEM-154	VIM-1
CARB-22	CTX-M-65	NDM-12	OXA-214	OXA-365	QnrA4	SHV-40	TEM-155	VIM-10
CARB-23	CTX-M-66	NDM-13	OXA-215	OXA-366	QnrA5	SHV-41	TEM-156	VIM-11
CARB-3	CTX-M-67	NDM-14	OXA-216	OXA-368	QnrA6	SHV-42	TEM-157	VIM-12
CARB-4	CTX-M-68	NDM-17	OXA-217	OXA-37	QnrA7	SHV-43	TEM-158	VIM-13
CARB-5	CTX-M-69	NDM-2	OXA-219	OXA-370	QnrS1	SHV-44	TEM-159	VIM-14
CARB-6	CTX-M-7	NDM-3	OXA-22	OXA-371	QNRS11	SHV-45	TEM-16	VIM-15
CARB-7	CTX-M-71	NDM-4	OXA-223	OXA-372	QnrS2	SHV-46	TEM-160	VIM-16
CARB-8	CTX-M-72	NDM-5	OXA-224	OXA-373	QnrS3	SHV-48	TEM-162	VIM-17
CARB-9	CTX-M-74	NDM-6	OXA-225	OXA-374	QnrS4	SHV-49	TEM-163	VIM-18
CTX-M-1	CTX-M-75	NDM-7	OXA-226	OXA-375	QnrS5	SHV-5	TEM-164	VIM-19
CTX-M-10	CTX-M-76	NDM-8	OXA-228	OXA-376	QnrS6	SHV-50	TEM-166	VIM-2
CTX-M-100	CTX-M-77	NDM-9	OXA-229	OXA-377	QnrS7	SHV-51	TEM-167	VIM-20
CTX-M-101	CTX-M-78	OXA-1	OXA-23	OXA-378	QnrS8	SHV-52	TEM-168	VIM-23
CTX-M-102	CTX-M-79	OXA-10	OXA-230	OXA-379	QnrS9	SHV-53	TEM-169	VIM-24
CTX-M-103	CTX-M-8	OXA-100	OXA-231	OXA-380	SHV-1	SHV-55	TEM-17	VIM-25
CTX-M-104	CTX-M-80	OXA-101	OXA-232	OXA-381	SHV-100	SHV-56	TEM-171	VIM-26

CTX-M-105	CTX-M-81	OXA-104	OXA-233	OXA-382	SHV-101	SHV-57	TEM-176	VIM-27
CTX-M-106	CTX-M-82	OXA-106	OXA-235	OXA-383	SHV-102	SHV-59	TEM-177	VIM-28
CTX-M-107	CTX-M-83	OXA-107	OXA-236	OXA-384	SHV-103	SHV-6	TEM-178	VIM-29
CTX-M-108	CTX-M-84	OXA-108	OXA-237	OXA-385	SHV-104	SHV-60	TEM-182	VIM-3
CTX-M-109	CTX-M-85	OXA-109	OXA-239	OXA-386	SHV-105	SHV-61	TEM-183	VIM-30
CTX-M-11	CTX-M-86	OXA-11	OXA-24	OXA-387	SHV-106	SHV-62	TEM-184	VIM-31
CTX-M-110	CTX-M-87	OXA-110	OXA-240	OXA-388	SHV-107	SHV-63	TEM-185	VIM-32
CTX-M-111	CTX-M-88	OXA-111	OXA-241	OXA-389	SHV-108	SHV-64	TEM-186	VIM-33
CTX-M-112	CTX-M-89	OXA-112	OXA-242	OXA-390	SHV-109	SHV-65	TEM-187	VIM-34
CTX-M-113	CTX-M-9	OXA-113	OXA-243	OXA-391	SHV-11	SHV-66	TEM-188	VIM-35
CTX-M-114	CTX-M-90	OXA-114a	OXA-244	OXA-392	SHV-110	SHV-67	TEM-189	VIM-36
CTX-M-115	CTX-M-91	OXA-115	OXA-245	OXA-397	SHV-111	SHV-69	TEM-19	VIM-37
CTX-M-116	CTX-M-92	OXA-116	OXA-246	OXA-398	SHV-112	SHV-7	TEM-190	VIM-38
CTX-M-117	CTX-M-93	OXA-117	OXA-247	OXA-4	SHV-119	SHV-70	TEM-191	VIM-39
CTX-M-12	CTX-M-94	OXA-118	OXA-248	OXA-407	SHV-12	SHV-71	TEM-192	VIM-4
CTX-M-121	CTX-M-95	OXA-119	OXA-249	OXA-415	SHV-120	SHV-72	TEM-193	VIM-42
CTX-M-122	CTX-M-96	OXA-12	OXA-25	OXA-417	SHV-121	SHV-73	TEM-194	VIM-43
CTX-M-123	CTX-M-98	OXA-120	OXA-250	OXA-418	SHV-122	SHV-74	TEM-195	VIM-5
CTX-M-124	CTX-M-99	OXA-121	OXA-251	OXA-42	SHV-123	SHV-75	TEM-196	VIM-6
CTX-M-125	GES-1	OXA-125	OXA-253	OXA-420	SHV-124	SHV-76	TEM-197	VIM-7
CTX-M-126	GES-10	OXA-128	OXA-254	OXA-421	SHV-125	SHV-77	TEM-198	VIM-8
CTX-M-129	GES-11	OXA-129	OXA-255	OXA-422	SHV-126	SHV-78	TEM-199	VIM-9
CTX-M-13	GES-12	OXA-13	OXA-256	OXA-423	SHV-127	SHV-79	TEM-2	
CTX-M-130	GES-13	OXA-130	OXA-257	OXA-424	SHV-128	SHV-8	TEM-20	
CTX-M-131	GES-14	OXA-131	OXA-258	OXA-425	SHV-129	SHV-80	TEM-201	
CTX-M-132	GES-15	OXA-132	OXA-26	OXA-426	SHV-13	SHV-81	TEM-205	
CTX-M-134	GES-16	OXA-133	OXA-266	OXA-427	SHV-133	SHV-82	TEM-206	
CTX-M-136	GES-17	OXA-134	OXA-27	OXA-43	SHV-134	SHV-83	TEM-207	
CTX-M-137	GES-18	OXA-136	OXA-270	OXA-435	SHV-135	SHV-84	TEM-208	
CTX-M-139	GES-19	OXA-137	OXA-272	OXA-436	SHV-137	SHV-85	TEM-209	
CTX-M-14	GES-2	OXA-138	OXA-274	OXA-443	SHV-14	SHV-86	TEM-21	
CTX-M-141	GES-20	OXA-139	OXA-275	OXA-45	SHV-140	SHV-89	TEM-211	
CTX-M-142	GES-21	OXA-14	OXA-278	OXA-454	SHV-141	SHV-9	TEM-213	
CTX-M-144	GES-22	OXA-141	OXA-28	OXA-46	SHV-142	SHV-92	TEM-214	
CTX-M-147	GES-23	OXA-142	OXA-280	OXA-464	SHV-143	SHV-93	TEM-215	
CTX-M-148	GES-24	OXA-143	OXA-281	OXA-47	SHV-144	SHV-94	TEM-216	
CTX-M-15	GES-26	OXA-144	OXA-282	OXA-48	SHV-145	SHV-95	TEM-217	
CTX-M-151	GES-3	OXA-145	OXA-287	OXA-486	SHV-147	SHV-96	TEM-219	
CTX-M-152	GES-4	OXA-146	OXA-288	OXA-49	SHV-148	SHV-97	TEM-22	
CTX-M-155	GES-5	OXA-147	OXA-29	OXA-5	SHV-149	SHV-98	TEM-220	
CTX-M-156	GES-6	OXA-148	OXA-294	OXA-50	SHV-15	SHV-99	TEM-24	
CTX-M-157	GES-7	OXA-149	OXA-296	OXA-51	SHV-150	TEM-1	TEM-26	
CTX-M-158	GES-8	OXA-15	OXA-297	OXA-53	SHV-151	TEM-10	TEM-28	
CTX-M-159	GES-9	OXA-150	OXA-298	OXA-535	SHV-152	TEM-101	TEM-29	
CTX-M-16	KPC-1	OXA-16	OXA-299	OXA-54	SHV-153	TEM-102	TEM-3	
CTX-M-160	KPC-10	OXA-160	OXA-3	OXA-55	SHV-154	TEM-104	TEM-30	
CTX-M-17	KPC-11	OXA-161	OXA-300	OXA-56	SHV-155	TEM-105	TEM-33	
CTX-M-19	KPC-12	OXA-162	OXA-304	OXA-57	SHV-156	TEM-106	TEM-34	
CTX-M-2	KPC-13	OXA-163	OXA-306	OXA-58	SHV-157	TEM-107	TEM-4	
CTX-M-20	KPC-14	OXA-164	OXA-308	OXA-59	SHV-158	TEM-108	TEM-40	
CTX-M-21	KPC-15	OXA-165	OXA-309	OXA-60	SHV-159	TEM-109	TEM-42	
CTX-M-22	KPC-16	OXA-166	OXA-31	OXA-61	SHV-16	TEM-11	TEM-43	
CTX-M-23	KPC-17	OXA-167	OXA-312	OXA-62	SHV-160	TEM-110	TEM-45	

CTX-M-24	KPC-19	OXA-168	OXA-313	OXA-63	SHV-161	TEM-111	TEM-47	
CTX-M-25	KPC-22	OXA-169	OXA-314	OXA-64	SHV-162	TEM-112	TEM-48	
CTX-M-26	KPC-3	OXA-17	OXA-315	OXA-65	SHV-163	TEM-113	TEM-49	
CTX-M-27	KPC-4	OXA-170	OXA-316	OXA-66	SHV-164	TEM-114	TEM-52	
CTX-M-28	KPC-5	OXA-171	OXA-317	OXA-664	SHV-165	TEM-115	TEM-53	
CTX-M-29	KPC-6	OXA-172	OXA-32	OXA-665	SHV-167	TEM-116	TEM-54	
CTX-M-3	KPC-7	OXA-173	OXA-320	OXA-67	SHV-168	TEM-117	TEM-55	
CTX-M-30	KPC-8	OXA-174	OXA-322	OXA-68	SHV-172	TEM-118	TEM-57	
CTX-M-31	KPC-9	OXA-175	OXA-323	OXA-69	SHV-173	TEM-12	TEM-59	
CTX-M-32	MCR-1	OXA-176	OXA-324	OXA-7	SHV-178	TEM-120	TEM-6	
CTX-M-33	MCR-1.1	OXA-177	OXA-325	OXA-70	SHV-179	TEM-121	TEM-60	
CTX-M-34	MCR-1.10	OXA-178	OXA-326	OXA-71	SHV-18	TEM-122	TEM-63	
CTX-M-35	MCR-1.2	OXA-179	OXA-327	OXA-72	SHV-180	TEM-123	TEM-67	
CTX-M-36	MCR-1.5	OXA-18	OXA-328	OXA-724	SHV-182	TEM-124	TEM-68	
CTX-M-37	MCR-2	OXA-180	OXA-329	OXA-73	SHV-183	TEM-125	TEM-7	
CTX-M-38	MCR-2.1	OXA-181	OXA-33	OXA-74	SHV-185	TEM-126	TEM-70	
CTX-M-39	MCR-3	OXA-182	OXA-330	OXA-75	SHV-186	TEM-127	TEM-71	
CTX-M-4	MCR-3.1	OXA-183	OXA-331	OXA-76	SHV-187	TEM-128	TEM-72	
CTX-M-40	MCR-3.10	OXA-184	OXA-332	OXA-77	SHV-188	TEM-129	TEM-73	
CTX-M-41	MCR-3.12	OXA-19	OXA-333	OXA-78	SHV-189	TEM-130	TEM-75	
CTX-M-42	MCR-3.3	OXA-192	OXA-334	OXA-79	SHV-19	TEM-131	TEM-76	
CTX-M-43	MCR-3.4	OXA-194	OXA-335	OXA-80	SHV-2	TEM-132	TEM-77	
CTX-M-44	MCR-3.5	OXA-195	OXA-338	OXA-82	SHV-20	TEM-133	TEM-78	
CTX-M-45	MCR-3.6	OXA-196	OXA-34	OXA-83	SHV-21	TEM-134	TEM-79	
CTX-M-46	MCR-3.7	OXA-197	OXA-347	OXA-84	SHV-22	TEM-135	TEM-8	
CTX-M-47	MCR-3.9	OXA-198	OXA-348	OXA-85	SHV-23	TEM-136	TEM-80	
CTX-M-48	MCR-4	OXA-199	OXA-349	OXA-86	SHV-24	TEM-137	TEM-81	
CTX-M-49	MCR-4.2	OXA-2	OXA-35	OXA-87	SHV-25	TEM-138	TEM-82	
CTX-M-5	MCR-4.3	OXA-20	OXA-350	OXA-88	SHV-26	TEM-139	TEM-83	
CTX-M-50	MCR-5	OXA-200	OXA-351	OXA-89	SHV-27	TEM-141	TEM-84	
CTX-M-51	MCR-5.1	OXA-201	OXA-352	OXA-9	SHV-28	TEM-142	TEM-85	
CTX-M-52	MCR-5.2	OXA-202	OXA-353	OXA-90	SHV-29	TEM-143	TEM-86	
CTX-M-53	MCR-6.1	OXA-203	OXA-354	OXA-91	SHV-2A	TEM-144	TEM-87	

Table B.5 List of genera that are functionally important to wastewater treatment falling under the functional guilds: filamentous, AOB, NOB, nitrite reduction, and anammox (the list was generated from a number of sources but most can be found in (14)).

<i>Acidovorax</i>	<i>Ca Solibacter</i>	<i>Leptothrix</i>	<i>Nitrotoga</i>
<i>Alysiella</i>	<i>Ca Villigracilis</i>	<i>Levilinea</i>	<i>Nocardia</i>
<i>Anaerolinea</i>	<i>Catenibacterium</i>	<i>Lewinella</i>	<i>Pannonibacter</i>
<i>Aquicella</i>	<i>Corynebacterium</i>	<i>Longilinea</i>	<i>Pseudomonas</i>
<i>Azoarcus</i>	<i>Crenothrix</i>	<i>Luteimonas</i>	<i>Rhodococcus</i>
<i>Azospira</i>	<i>Dechloromonas</i>	<i>Methanospirillum</i>	<i>Rhodoferax</i>
<i>Bellilinea</i>	<i>Denitratisoma</i>	<i>Methanothrix</i>	<i>Rhodoplanes</i>
<i>Bradyrhizobium</i>	<i>Diaphorobacter</i>	<i>Mycobacterium</i>	<i>Simplicispira</i>
<i>Brocadia</i>	<i>Dokdonella</i>	<i>Neomegalonema</i>	<i>Skermania</i>
<i>Ca Accumulibacter</i>	<i>Erysipelothrix</i>	<i>Nitrobacter</i>	<i>Steroidobacter</i>
<i>Ca Alysiosphaera</i>	<i>Flexilinea</i>	<i>Nitrococcus</i>	<i>Sterolibacterium</i>
<i>Ca Amarolinea</i>	<i>Gordonia</i>	<i>Nitrolancea</i>	<i>Sulfurimonas</i>
<i>Ca Brevefilum</i>	<i>Haliangium</i>	<i>Nitromaritima</i>	<i>Sulfuritalea</i>

<i>Ca Brocadia</i>	<i>Halioglobus</i>	<i>Nitrosococcus</i>	<i>Thauera</i>
<i>Ca Competibacter</i>	<i>Haliscomenobacter</i>	<i>Nitrosolobus</i>	<i>Thermogutta</i>
<i>Ca Microthrix</i>	<i>Herpetosiphon</i>	<i>Nitrosomonas</i>	<i>Thiothrix</i>
<i>Ca Nostocoida</i>	<i>Hyphomonas</i>	<i>Nitrospira</i>	<i>Turicibacter</i>
<i>Ca Phosphitivorax</i>	<i>Iamia</i>	<i>Nitrosovibrio</i>	<i>Uruburuella</i>
<i>Ca Promineofilum</i>	<i>Kouleothrix</i>	<i>Nitrospina</i>	<i>Vitreoscilla</i>
<i>Ca Sarcinithrix</i>	<i>Leptolinea</i>	<i>Nitrospira</i>	<i>Zoogloea</i>

Table B.6 List of genera that may contain pathogenic species (15) (i.e., potential pathogens)

<i>Abiotrophia</i>	<i>Chryseobacterium</i>	<i>Kluyvera</i>	<i>Providencia</i>
<i>Achromobacter</i>	<i>Citrobacter</i>	<i>Lactobacillus</i>	<i>Pseudomonas</i>
<i>Acidaminococcus</i>	<i>Clostridium</i>	<i>Legionella</i>	<i>Pseudonocardia</i>
<i>Acinetobacter</i>	<i>Collinsella</i>	<i>Leifsonia</i>	<i>Pseudoramibacter</i>
<i>Actinobacillus</i>	<i>Comamonas</i>	<i>Leptospira</i>	<i>Psychrobacter</i>
<i>Actinomyces</i>	<i>Corynebacterium</i>	<i>Leptotrichia</i>	<i>Rahnella</i>
<i>Aerococcus</i>	<i>Coxiella</i>	<i>Listeria</i>	<i>Ralstonia</i>
<i>Aeromonas</i>	<i>Delftia</i>	<i>Mannheimia</i>	<i>Rhodococcus</i>
<i>Alcaligenes</i>	<i>Dermatophilus</i>	<i>Megamonas</i>	<i>Rickettsia</i>
<i>Amycolatopsis</i>	<i>Dichelobacter</i>	<i>Megasphaera</i>	<i>Rothia</i>
<i>Anaerococcus</i>	<i>Edwardsiella</i>	<i>Micromonas</i>	<i>Ruminococcus</i>
<i>Anaplasma</i>	<i>Eggerthella</i>	<i>Mogibacterium</i>	<i>Saccharomonospora</i>
<i>Arcanobacterium</i>	<i>Ehrlichia</i>	<i>Moraxella</i>	<i>Saccharopolyspora</i>
<i>Arcobacter</i>	<i>Eikenella</i>	<i>Morganella</i>	<i>Salmonella</i>
<i>Bacillus</i>	<i>Enterobacter</i>	<i>Mycobacterium</i>	<i>Sealdella</i>
<i>Bacteroides</i>	<i>Enterococcus</i>	<i>Mycoplasma</i>	<i>Selenomonas</i>
<i>Bartonella</i>	<i>Erysipelothrix</i>	<i>Myroides</i>	<i>Serratia</i>
<i>Bergeyella</i>	<i>Escherichia</i>	<i>Neisseria</i>	<i>Shigella</i>
<i>Bifidobacterium</i>	<i>Eubacterium</i>	<i>Neorickettsia</i>	<i>Sphingomonas</i>
<i>Bilophila</i>	<i>Ewingella</i>	<i>Nocardia</i>	<i>Spirillum</i>
<i>Bordetella</i>	<i>Fibrobacter</i>	<i>Ochrobactrum</i>	<i>Staphylococcus</i>
<i>Borrelia</i>	<i>Filifactor</i>	<i>Oligella</i>	<i>Stenotrophomonas</i>
<i>Brevibacillus</i>	<i>Finegoldia</i>	<i>Orientia</i>	<i>Streptobacillus</i>
<i>Brevundimonas</i>	<i>Fluoribacter</i>	<i>Paenibacillus</i>	<i>Streptococcus</i>
<i>Brucella</i>	<i>Francisella</i>	<i>Pantoea</i>	<i>Suttonella</i>
<i>Burkholderia</i>	<i>Fusobacterium</i>	<i>Pasteurella</i>	<i>Tatlockia</i>
<i>Campylobacter</i>	<i>Gardnerella</i>	<i>Peptinophilus</i>	<i>Tatumella</i>
<i>Capnocytophaga</i>	<i>Gemella</i>	<i>Peptococcus</i>	<i>Treponema</i>
<i>Cardiobacterium</i>	<i>Gordonia</i>	<i>Peptostreptococcus</i>	<i>Tropheryma</i>
<i>Cedecea</i>	<i>Granulicatella</i>	<i>Photobacterium</i>	<i>Tsukamurella</i>
<i>Cellulomonas</i>	<i>Haemophilus</i>	<i>Plesiomonas</i>	<i>Ureaplasma</i>
<i>Centipeda</i>	<i>Hafnia</i>	<i>Porphyromonas</i>	<i>Veillonella</i>
<i>Chlamydia</i>	<i>Helicobacter</i>	<i>Prevotella</i>	<i>Vibrio</i>
<i>Chlamydophila</i>	<i>Kingella</i>	<i>Propionibacterium</i>	<i>Wolinella</i>
<i>Chromobacterium</i>	<i>Klebsiella</i>	<i>Proteus</i>	<i>Yersinia</i>

Table B.7 Overview of samples collected for metagenomic sequencing with each new WWTP feed highlighted with a different color in the first column. The five day SRT is indicated by gray bars with the two day HRT indicated by lined cells. This table illustrates that the samples are dependent when considering dates close to each other as there may be overlapping feed, HRT, and/or SRT. Therefore, samples were only considered independent if the sampling dates were far enough apart.

[illegible]

Table B.8 Pharmaceuticals and personal care products (PPCP) and their metabolites grouped into different drug classes.

Compound	Class	Compound	Class	Compound	Class	Compound	Class
(±)11-nor-9-carboxy-delta-THC	Psychoactive	Diclofenac	Analgesic	Meprobamate	Other	Phenytoin	Anticonvulsant
10,11-dihydro-10-hydroxycarbamazepine	Anticonvulsant	Diclofenac 4-hydroxy	Analgesic	Metformin	Antidiabetic	Pioglitazone	Antidiabetic
6-Acetylmorphine	Narcotic	Diltiazem	Antihypertensive	Methadone	Narcotic	Pravastatin	Statin
Acebutolol	Beta Blocker	Diphenhydramine	Antihistamine	Methamphetamine	Stimulant	Pregabalin	Anticonvulsant
Acetaminophen	Analgesic	Disopyramide	Antiarrhythmic	Methotrexate	Other	Primidone	Anticonvulsant
Albuterol	Bronchodilators	Donepezil	Other	Methylparaben	Antimicrobial	Propranolol	Beta Blocker
Amitriptyline	Antidepressant	Duloxetine	Antidepressant	Methylphenidate	Stimulant	Pseudoephedrine	Other
Amitriptyline metabolite	Antidepressant	Ecgonine methyl ester	Stimulant	Metoprolol	Beta Blocker	Quetiapine	Antipsychotic
Amphetamine	Stimulant	EDDP	Narcotic	Mevastatin	Statin	Ritalinic acid	Stimulant
Aripiprazole	Antipsychotic	Erythromycin	Antimicrobial	m-Hydroxybenzoylecgonine	Stimulant	Sertraline	Antidepressant
Atenolol	Beta Blocker	Erythromycin-anhydro	Antimicrobial	Modafinil	Stimulant	Sildenafil	Other
Atorvastatin	Statin	Escitalopram	Antidepressant	Modafinil acid	Stimulant	Simvastatin	Statin
Atrazine	Herbicide	Famotidine	Antihistamine	Monoethylglycinexylidide	Antiarrhythmic	Sotalol	Antiarrhythmic
Azithromycin	Antimicrobial	Fenbufen	Analgesic	Montelukast	Other	Sulfadimethoxin	Antimicrobial
Benzoyllecgonine	Stimulant	Fentanyl	Narcotic	Morphine	Narcotic	Sulfamethazine	Antimicrobial
Bezafibrate	Other	Fluoxetine	Antidepressant	Naproxen	Analgesic	Sulfamethoxazole	Antimicrobial
Buprenorphine	Narcotic	Fluticasone propionate	Other	n-Butylparaben	Antimicrobial	Sumatriptan	Other
Bupropion	Antidepressant	Furosemide	Diuretic	Nifedipine	Antihypertensive	Tadalafil	Other
Caffeine	Stimulant	Gabapentin	Anticonvulsant	Nifedipine oxidized	Antihypertensive	Temazepam	Sedative
Carbamazepine	Anticonvulsant	Gemfibrozil	Other	Norfentanyl	Narcotic	Tetracycline	Antimicrobial
Carbamazepine 10,11 epoxide	Anticonvulsant	Glyburide	Antidiabetic	Norfloxacin	Antimicrobial	Thiabendazole	Antimicrobial
Carisoprodol	Other	Hydrochlorothiazide	Diuretic	Norfluoxetine	Antidepressant	Tramadol	Antidepressant
Celecoxib	Analgesic	Hydrocodone	Narcotic	Normeperidine	Narcotic	Trazadone	Antidepressant
Chloramphenicol	Antimicrobial	Hydromorphone	Narcotic	Norquetiapine	Antidepressant	Triamterene	Diuretic
Chlorotetracycline	Antimicrobial	Hydroxybupropion	Antidepressant	Norsertraline	Antidepressant	Triclocarban	Antimicrobial
Chlorpheniramine	Antihistamine	Ibuprofen	Analgesic	Norverapamil	Antihypertensive	Triclosan	Antimicrobial
Ciprofloxacin	Antimicrobial	Imipenem	Antimicrobial	Omeprazole	Other	Trimethoprim	Antimicrobial
Clarithromycin	Antimicrobial	Ketoprofen	Analgesic	Oxazepam	Sedative	Tylosin	Antimicrobial
Clenbuterol	Other	Lamotrigine	Anticonvulsant	Oxcarbazepine	Anticonvulsant	Valsartan	Antihypertensive
Clopidogrel carboxylic acid	Anticoagulant	Levorphanol	Narcotic	Oxolinic acid	Antimicrobial	Vancomycin	Antimicrobial
Cocaethylene	Stimulant	Lidocaine	Anesthetic	Oxycodone	Narcotic	Venlafaxine	Antidepressant
Cocaine	Stimulant	Loratadine	Antihistamine	Oxymorphone	Narcotic	Verapamil	Antihypertensive

Codeine	Narcotic	Lorazepam	Sedative	Oxymorphone glucuronide	Narcotic	Warfarin	Anticoagulant
Cotinine	Other	MDA	Stimulant	Paroxetine	Antidepressant	Zolpidem	Sedative
Dehydroaripiprazole	Antipsychotic	MDEA	Stimulant	Phenmetrazine	Stimulant	Zolpidem phenyl-4-carboxylic acid	Sedative
Desmethylocitalopram	Antidepressant	MDMA	Stimulant	Phenobarbital	Anticonvulsant		
Desmethylenlafaxine	Antidepressant	Mefenamic acid	Analgesic	Phentermine	Stimulant		
Dextromethorphan	Other	Meperidine	Narcotic	Phenylpropanolamine	Stimulant		

Table B.9 Shannon diversity statistics comparison

Database	Groups Compared	p-value (Wilcox, Holm adjusted)
ARGs	ML vs FE	1.98e-11 ***
	0% ML vs FE	0.000000634 ***
	10% ML vs FE	0.000000554 ***
	FE 0% vs 10%	0.866
	ML 0% vs 10%	0.123
	ML vs FEED	6.18e- 9 ****
	FE vs FEED	1.13e-13 ****

* p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

Table B.10 Vancomycin in municipal WWTPs (MWW) with and without hospital wastewater (HWW). Samples were run on separate dates; therefore, the peak areas may only be compared across one date.

PPCP	Date	Location	Peak Area
Vancomycin	July 2019	MWW without HWW	22
		HWW	69794
		MWW with HWW	163
	Sept 2019	MWW without HWW	25
		HWW	67380
		MWW with HWW	102
	Jan 2020	MWW without HWW	0
		HWW	17426
		MWW with HWW	0

B.2 Supplementary Figures

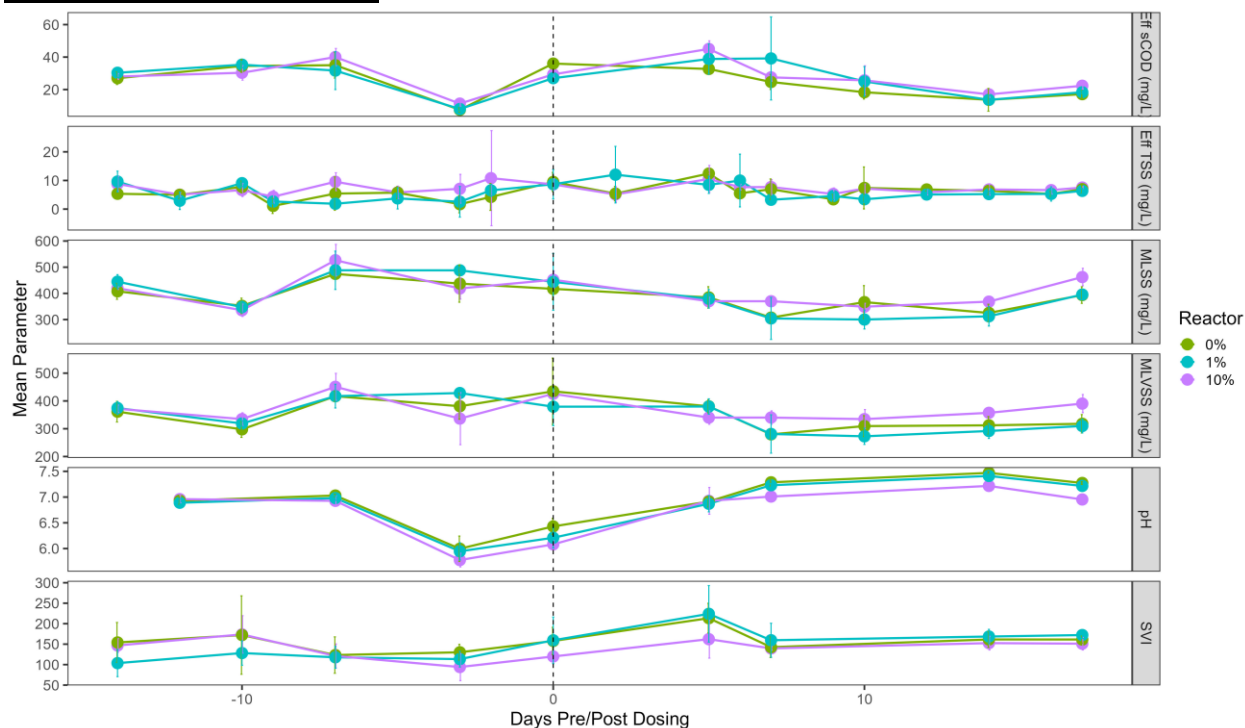


Figure B.1 Mean SBR performance parameters. Parameters were not different between reactors 0-17 days (ANOVA or Kruskal-Wallis, $p > 0.05$). The mean effluent TSS across all SBRs was 6.2 ± 2.8 mg/L and mean effluent sCOD across all SBR was 27.0 ± 4.3 mg/L thereby the SBR were performing within NPDES limits of 30 mg/L 30-day average for BOD and TSS.

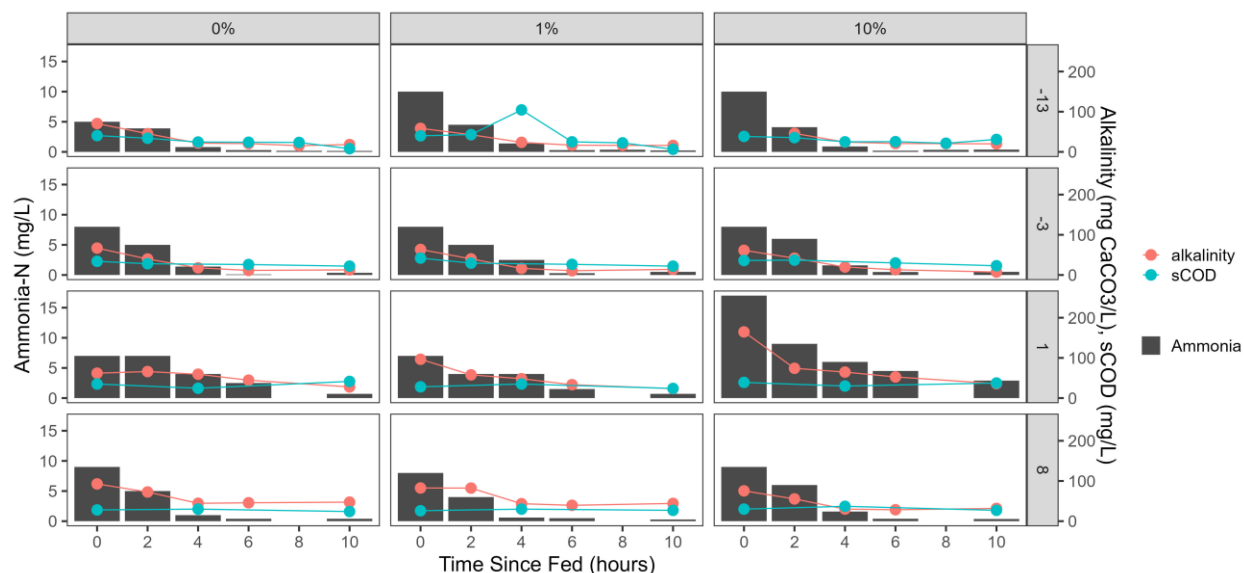


Figure B.2 Ammonia, alkalinity and sCOD reduction in the SBR during aeration. Because nitrification can consume alkalinity this parameter was monitored to check if alkalinity needed to be added to the feed.

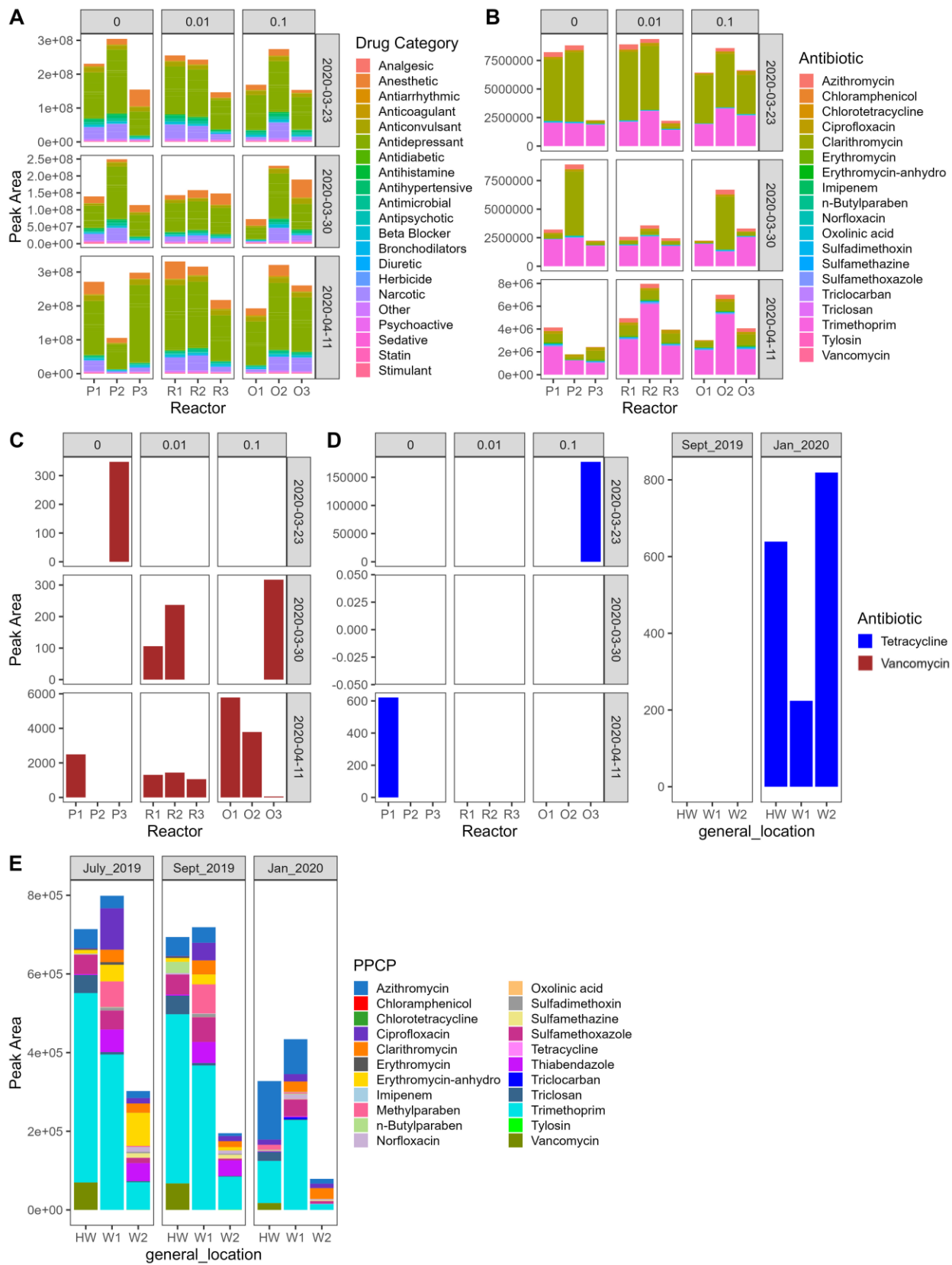


Figure B.3 Pharmaceuticals and personal care products (PPCPs) and their metabolites detected by UPLC-MS/MS (A) 149 total compounds grouped into drug classes (B) PPCP antibiotics (C) Vancomycin detected in the SBR with rows representing different dates (top is prior to HWW dosing) (D) (left) Similar to C except the antibiotic tetracycline (right) pulled out tetracycline from the earlier samples looking at PPCPs in WWTP without HWW (W1), HWW (W1), and WWTP receiving the HWW (W2). Tetracycline was not screened for in July 2019, not detected in September 2019, detected in all January 2020 samples (E) Similar to (D, right), however this graph is showing all antibiotics. Note that these data were run on different days, so while they are graphed next to each other the abundances cannot be compared across sampling dates. Nevertheless, vancomycin (green bar on the bottom) is clearly visible in the HWW at a much greater abundance than W1 or W2 WWTPs. Vancomycin was not detected

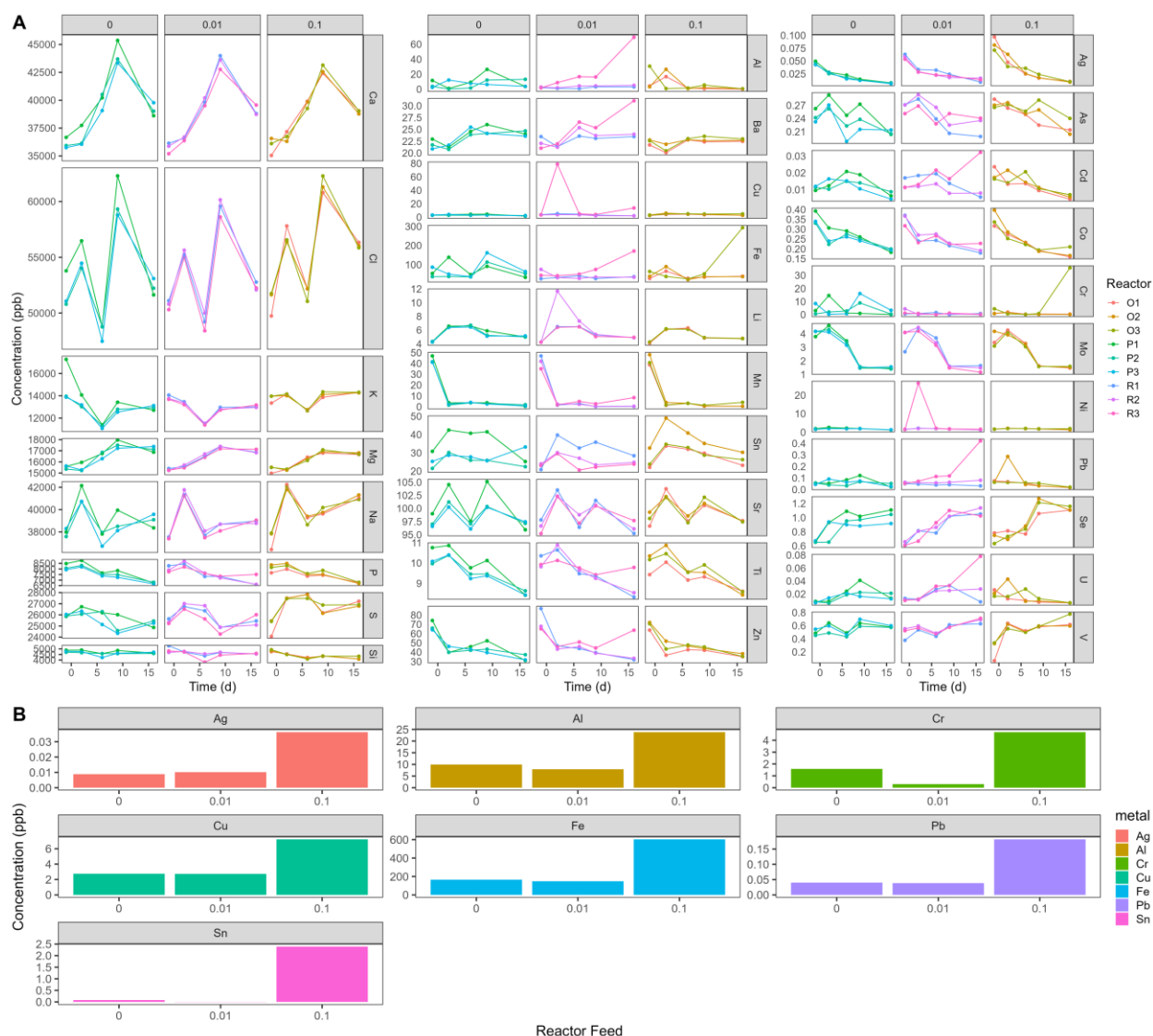


Figure B.4 29 soluble elements were analyzed from mixed liquor from each SBR (10% = O1-3, 1% = R1-3, 0% = P1-3) and (B) Elements that were more abundant in 10% feed than 0% or 1%: silver, aluminum, chromium, copper, iron, lead, tin. Anything with greater than a 0.3 log change between 0 and 10% and/or 1 and 10% feed

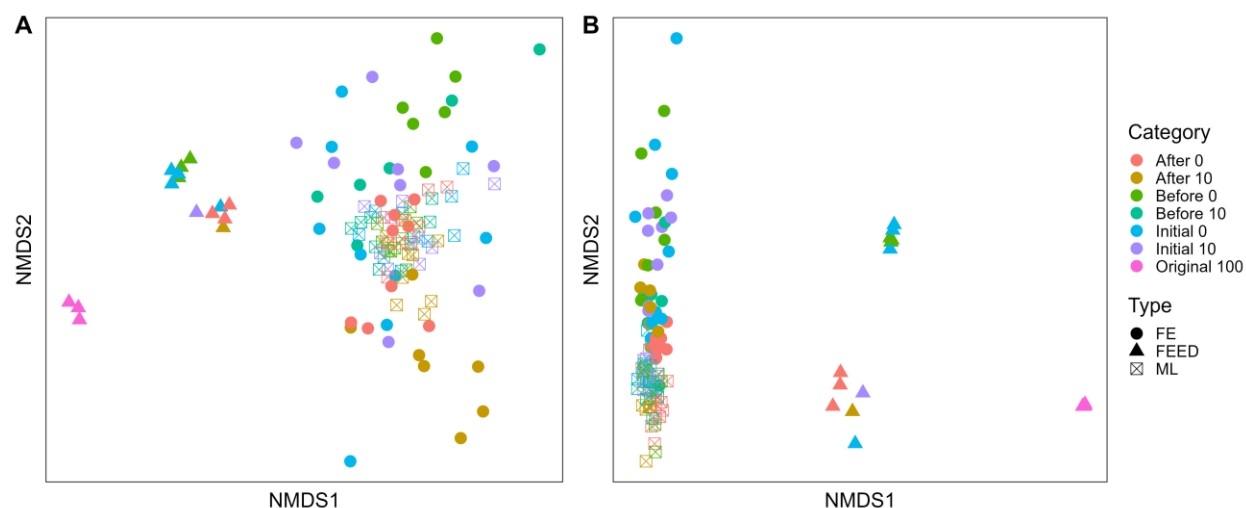


Figure B.6 (A) NMDS of ARGs based on Bray Curtis dissimilarity (stress = 0.1258261, with feed). PERMANOVA: $R^2=0.56932$, $p=0.001$ (with feed) $R^2: 0.32107$, $p=0.001$ (without feed). (B) NMDS of genera based on Bray Curtis dissimilarity (stress 0.04589027). PERMANOVA: $R^2 0.85052$, $p=0.001$ (with feed) $R^2 = 0.51385$ $p=0.001$ (without feed).

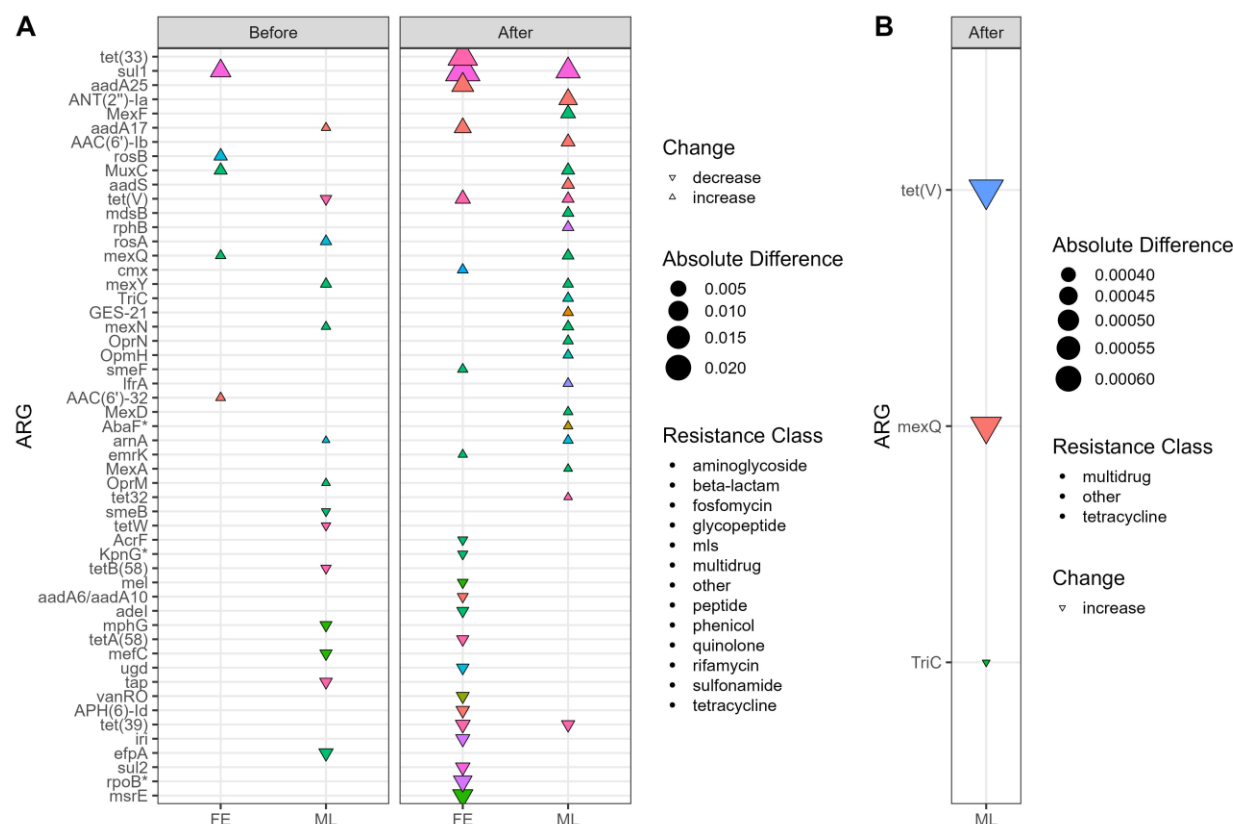


Figure B.7 Wilcox test comparing 10% and 0% ARGs for different iterations of p-value cutoffs (A) $p\text{-value} < 0.05$ and (B) $p\text{-value} < 0.001$



Figure B.8 Showing ARGs for the three genes that had $p < 0.001$ for (A) all samples and (B) mean abundances only for “after” dates in the SBR (mixed liquor and final effluent) and for 0%, 10%, 100% feed, and calculated 10% feed.

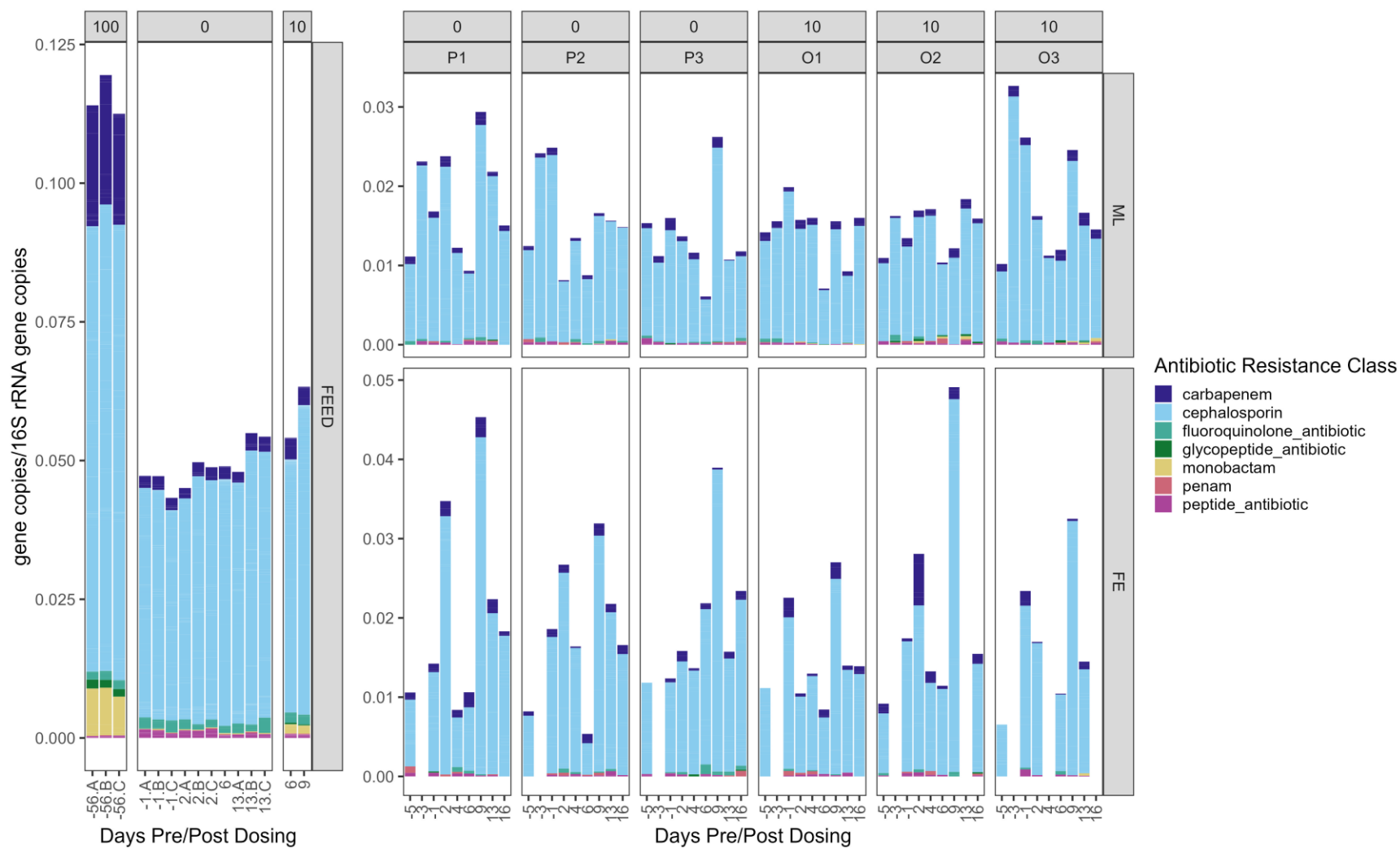


Figure B.9 All clinically relevant ARGs grouped by antibiotic resistance class.

CHAPTER 4: INTERNATIONAL COMPARISON OF ANTIBIOTIC RESISTANCE GENES FROM WASTEWATER TREATMENT PLANT FINAL EFFLUENTS AND THEIR RECEIVING ENVIRONMENT

Coauthors: Maria Virginia Prieto Riquelme, Emily Garner, Jake Metch, Marjorie Willner, Matthew Blair (Virginia Tech), D.G. Joakim Larsson (Department of Infectious Diseases, University of Gothenburg), Helmut Bürgmann (Eawag: Swiss Federal Institute of Aquatic Science and Technology), Tong Zhang (Department of Civil Engineering, The University of Hong Kong), Indumathi M. Nambi (Department of Civil Engineering, Indian Institute of Technology), Diana Aga (Department of Chemistry, University at Buffalo), Amy Pruden (Virginia Tech), Peter Vikesland (Virginia Tech)

4.1 Abstract

Antimicrobial resistance is an important public health concern that spans clinical settings and environments across the world. One anthropogenic source for antibiotic resistance genes (ARGs) and antibiotic resistant bacteria in aquatic environments are wastewater treatment plant (WWTP) effluents. This study sought to further our understanding of the impacts of WWTP effluents on receiving environments. Metagenomic sequencing of globally sourced final effluents as well as corresponding upstream and downstream receiving environment samples was conducted. Analyses using qPCR found results typical of the literature (an increase in ARGs downstream); however, metagenomic sequencing revealed similar upstream and downstream resistomes. Aquatic resistomes were distinct between countries a finding which reflects differences in antibiotic burden, local pollution, and river microbiomes. Additional studies are needed to more thoroughly explain the differences in river resistomes around the world and their anthropogenic inputs.

4.2 Introduction

The discovery of antibiotics enabled major advancements in medicine that led to significant overall improvements in people's lives. However, due to the rise of antibiotic resistance, the societal benefits that antibiotics provide may lessen and ultimately disappear (1). In addition to losses in medicinal benefits, antibiotic resistance burdens the economy and healthcare systems with longer hospital stays and the potential need for supplemental medicines for treatment of antibiotic resistant infections (2). The development of new antibiotics is not economically profitable; therefore, it has been suggested that, in tandem with combating inappropriate antibiotic use, changes need to be made in the market that incentivize antibiotic production (3). A coordinated action-based plan is necessary to fight antibiotic resistance; with this though in mind the World Health Organization (WHO) developed a global action plan to combat antibiotic resistance that includes improved awareness, surveillance, research, sanitation measures, improved antimicrobial use, and sustainable economic investments (4).

Antibiotic resistance is a global problem that is not constrained by state or country boundaries; however, there are differences in key contributors between high-income (HIC) and low-middle-income (LMIC) countries. Chokshi and colleagues (5) summarized key contributors in LMICs as the lack of surveillance of resistance development, poor control of available antibiotics, clinical misuse, and ease of availability. On the other hand, key contributions in HICs related to poor hospital-level regulation and overuse of antibiotics for maintenance of food-producing animals. Human behavior plays an important role in increasing or disseminating antibiotic resistance. Travel to regions with high rates of antibiotic resistance may result in the

transfer of resistance to regions with lower or different types of resistance (2). Antibiotics may be misused when they are prescribed without prescriptions and/or prescribed when unnecessary (6). Changes to these behaviors require global political action and cooperation.

Antibiotic resistance is a natural and ancient phenomena involving communication and antagonism between bacteria (7). As expected, ARGs have therefore been found in numerous pristine environments (7). Despite its natural occurrence, the spread and evolution of antibiotic resistance is impacted by human activities as ARBs and ARGs have been linked to anthropogenic waste and agricultural practices (8). While wastewater treatment plants (WWTPs) are often seen as hotspots for the dissemination of antimicrobial resistance into the environment (9) additional sources include combined sewer overflow, aging infrastructure resulting in leaky pipes, and agricultural runoff (10–12). WWTPs are nodes where the consequences of antibiotic use and other human behaviors converge. Antibiotics that have not been fully metabolized are excreted along with antibiotic resistance genes (ARGs) and antibiotic resistant bacteria (ARB). WWTPs are not explicitly designed to remove antibiotics or to treat their resistance determinants (i.e., ARGs) therefore these may pass through the WWTP and out into the receiving environment where treated water is discharged. The high microbial loads present during wastewater treatment, along with the presence of antibiotics and metals that may provide selective pressure promoting resistance, suggest that WWTPs may be ‘hotspots’ for the dissemination of antibiotic resistance (9).

The objective of this study was to assess the impact of WWTPs on antibiotic resistance in a receiving aquatic environment across five countries spanning North America, Europe, and Asia. Samples collected were compared using both quantitative polymerase chain reaction (qPCR) and metagenomic sequencing. The data from qPCR allow for quantitative measure of change, whereas metagenomic data provide a broader perspective of the environment’s resistome (i.e., a measure of total ARGs). It was hypothesized that the total relative abundance of ARGs (i.e., resistome) would increase downstream of the WWTPs and that samples collected in urban environments would have greater loads of ARGs than those collected from rural streams.

4.3 Materials and Methods

4.3.1 Site Description

Stream samples were collected upstream and downstream of wastewater treatment plants in each of the following countries: Switzerland (CHE), India (IND), the Philippines (PHL), Sweden (SWE) and the United States (USA) between March 2016 – July 2017. Detailed information on the WWTPs may be found in Prieto Riquelme et al. (13). Stream samples were collected from two streams (**Table C.1**) in all countries except for PHL where only one upstream sample was collected. The streams in IND were sampled once in 2016 and again in 2017, all other samples were collected in 2016. Herein, samples are denoted by their country of origin, sampling trip (#1 for all samples except for IND which included a second trip), WWTP (P1 or P2) and stream or WWTP sampling location upstream (UP), final effluent (FE), or downstream (DN). Distances were measured using Google maps between the upstream and downstream sample locations following the path of the river. Distances from a given sampling location to the nearest hospital were measured as a direct point-to-point distance. Land use was determined using an ArcGIS world land cover map (14) with local land use based on a 1 km circular distance around the sampling location.

4.3.2 Sample Collection and Processing

Sampling locations varied in distance from the WWTP effluent discharge point due to stream accessibility. Distances between upstream and downstream locations ranged from ~200 m to almost 9 km (**Table C.1**). Grab stream samples were collected in sterile (autoclaved) polypropylene bottles for DNA and in acid-washed, pre-combusted (500 °C) glass amber bottles as described previously for antibiotics (13,15,16). Briefly, samples were transferred back to the local laboratory where within 12-hours they were filtered in triplicate for DNA through 0.22- μ m mixed cellulose ester membranes (MilliporeSigma Darmstadt, Germany) until clogging. The filters were then preserved in 50% molecular grade ethanol and stored at -20 °C until shipment. Samples were shipped back to the Virginia Tech laboratory and stored at -20 °C until further processing. Samples for DNA were torn into small pieces (~3 mm²) using flame sterilized forceps and extracted using the FastDNA® SPIN Kit for Soil (MP Biomedicals, Solon, Ohio) following the manufacturer's instructions. Samples to be analyzed for antibiotic concentrations were spiked with Na₂EDTA (2 mL, 5% v/v in water) and isotopically labeled surrogate standards (50 μ L of 1000 μ g/L surrogate mix solution) and filtered using 0.45- μ m glass microfiber filters. These samples were shipped to University at Buffalo for analysis using high performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) as described previously (17). Analyses of the following antibiotics was conducted: anhydro-erythromycin, azithromycin, clarithromycin, roxithromycin, ciprofloxacin, enrofloxacin, norfloxacin, oxolinic acid, sarafloxacin, sulfamethoxazole, sulfadiazine, acetylsulfamethoxazole, and sulfamethoxydiazine.

4.3.3 Molecular Analysis and High Throughput Sequencing

Ten upstream and nine downstream samples were quantified for target ARGs via qPCR using a CFX96 Real Time System (BioRad, Hercules, CA, United States) using previously published protocols for the following genes: *bla*_{TEM} (18), *ermB* (19), *qnrS* (20), *sul1* (21), *tet(X)* (22), *vanA* (23), *intI1* (24), and 16S rRNA (25) as described previously (15,26). Genes were analyzed in triplicate using a dilution factor of 1:10 along with standard curves of 10 $\times$ serially diluted standards and triplicate negative controls. Primer sequences, R², and efficiency values are reported in **Table C.2**.

Eight upstream, seven downstream, and ten final effluent (including replicates from two WWTPs) samples were selected for shotgun metagenomic sequencing. Triplicate DNA extracts were pooled by equal mass and prepared for sequencing using TrueSeq library preparation (Illumina, San Diego, California). Sequencing was performed at the Virginia Tech Biocomplexity Institute Genomic Sequencing Center (Blacksburg, VA) on an Illumina HiSeq 2500 using 2 $\times$ 100 paired-end reads. Sequencing depth ranged from 1.8-18 million reads per sample after quality trimming and filtering (on average 10 million reads per sample; **Table C.3**). Samples were uploaded to the MetaStorm pipeline (27) where quality filtering was performed and reads were annotated against the Comprehensive Antibiotic Resistance Database (CARD v 2.0.1) to identify ARGs (28), the Silva rRNA database for 16S rRNA genes (29), and Metagenomic Phylogenetic Analysis 2 (MetaPhlAn2) (30).

4.3.4 Statistical Analyses

All statistical tests were conducted in R (v 4.1.3) (31) using the rstatix package (32) using an α of 0.05 (p-value < 0.05) as the level of significance. Metagenomic samples were found to be non-normally distributed by the Anderson-Darling test (33). qPCR and individual ARGs from metagenomics data were found to be normally distributed via Shapiro–Wilks (34) when log

transformed. Comparisons between stream samples were conducted using a Wilcoxon rank sum. Alpha diversity was determined using Shannon diversity index (35) and Pielou's evenness index (36). Beta-diversity was determined via Bray-Curtis dissimilarity matrices and visualized using non-metric multidimensional scaling (NMDS); significant differences between groups were calculated using Permutational Multivariate Analysis of Variance (PERMANOVA). All ordination analyses were done using the R package *vegan* (37). All figures were generated in R using *ggplot2* (38).

4.4 Results

4.4.1 qPCR and Metagenomics

ARGs quantified via qPCR were typically higher in DN samples than UP, the exceptions being CHE P2, IND2 P2, and USA P1 which had similar medians for all genes for both upstream and downstream samples (**Figure 4.1 A**). With the exception of SWE P2, all locations had elevated concentrations of integrase gene *intI1* downstream (**Figure 4.1 B**). We note that the ARGs that were quantifiable via qPCR were not always detectable via metagenomics; however, genes representing *ermB*, *qnrS* (*qnrS2*, *qnrS4*, and *qnrS5* which are all captured by the qPCR target gene *qnrS*), and *sul1* were detected in a subset of stream metagenomics samples. Patterns were harder to discern using metagenomics data as not all the genes were detected in both upstream and downstream samples; the two samples that had detections of these ARGs in both upstream and downstream samples were IND P1 and SWE P2. Metagenomics data for IND P1 *sul1* was elevated downstream (this matched the qPCR results), however, *qnrS2* and *qnrS4* had lower relative abundances downstream (this was opposite of the qPCR results). SWE P2 metagenomics data had a lower relative abundance of *sul1* and an elevated relative abundance of *ermB* downstream (both of these matched with the qPCR results) (**Figure 4.1 C**). To further determine how well the qPCR and metagenomics results were aligned, the ARG that was most frequently detected for both qPCR and metagenomics was compared; *sul1* relative abundances were normally distributed when log transformed and therefore the correlation was assessed via Pearson's correlation coefficient. *sul1* was found to be positively correlated between both qPCR and metagenomics data ($R=0.83$ and $R^2=0.68$, $p<0.0001$) (**Figure 4.1 D**), therefore, it is likely that metagenomics does well at truly representing the relative abundances of frequently detected genes.

4.4.2 Resistome

The total ARG relative abundance was typically higher in final effluent than the corresponding stream samples. While the final effluent samples from most countries had greater total relative abundance than all other countries' stream resistomes, USA P2 had a lower final effluent total ARG relative abundance (though this was still higher than USA P2 stream samples) (**Figure 4.2 A**). When examining upstreams and downstreams across countries and WWTPs the resistomes tell a different story from qPCR. Unlike the trends seen via qPCR, the resistomes tended to be higher (i.e., have a greater total relative abundance of ARGs) upstream than downstream with the exception of USA P2 which was slightly higher downstream and for IND P1 and USA P1 which had similar levels of total ARGs in both upstreams and downstreams samples (**Figure 4.2 B**).

Comparing total resistance classes and their change between upstream and downstream reveals more nuanced story than that for total ARGs. The WWTPs that had the most antibiotic resistance classes increase (i.e., the sum ARG relative abundance for that class) downstream were USA P2 and CHE P2 (77% and 71% of resistance classes increased, respectively). IND P1 fell in

the middle with 40% of resistance classes increasing downstream. The remaining WWTPs (CHE P1, SWE P1, SWE P2, and USA P1) all had lower percentages of resistance genes increase downstream (i.e., most resistance classes had greater relative abundances upstream) (on average $28 \pm 2\%$ of resistance classes increased downstream) (**Figure 4.3**).

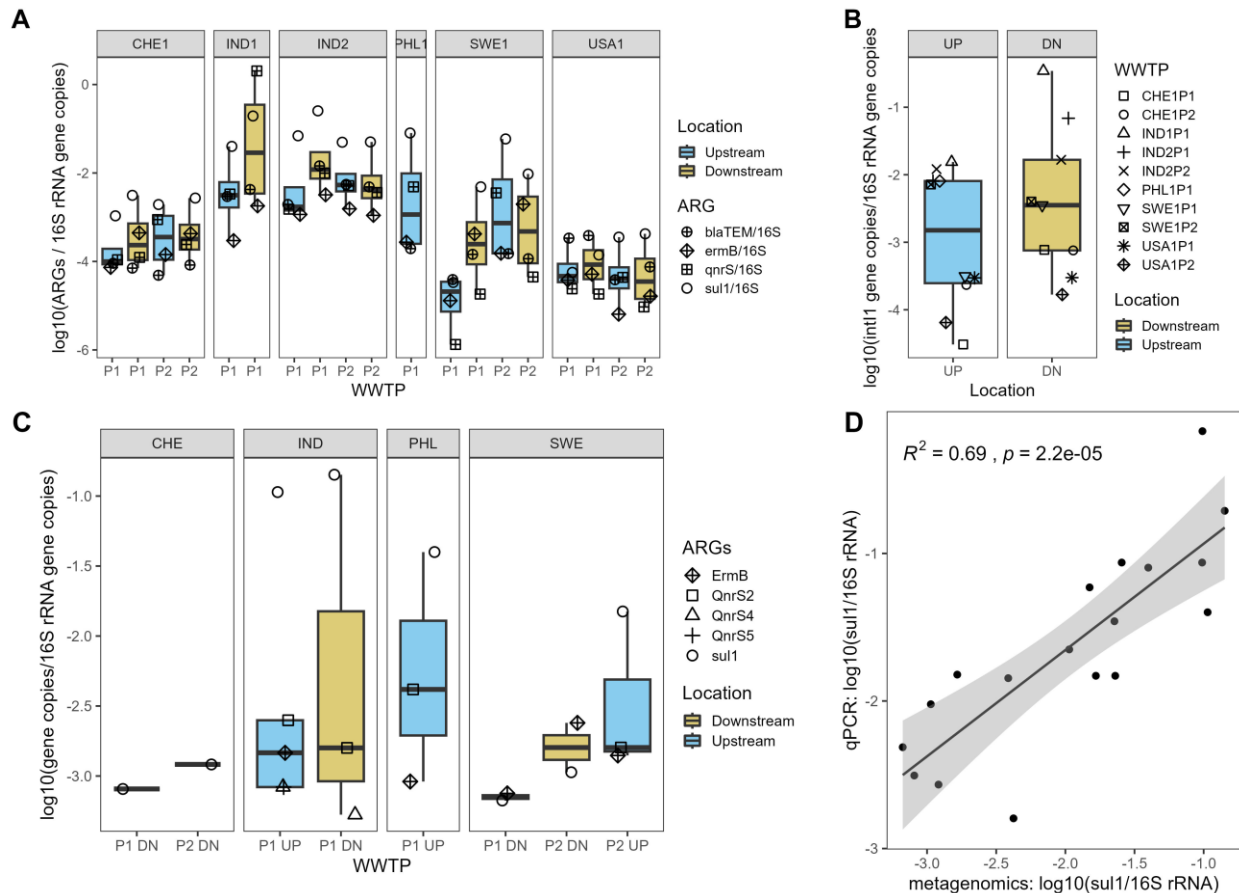


Figure 4.1 qPCR results for (A) ARGs and (B) *intI1* for upstream and downstream samples, (C) ARGs from metagenomic sequencing that matched qPCR and (C) correlation between qPCR and metagenomics for *sul1* (log transformed, Pearson correlation), data included *sul1* from UP DN, and FE samples.

The resistance classes that tended to increase or decrease most across WWTPs were aminoglycoside, macrolide/lincosamide/streptogramin (MLS), quinolone, multidrug and other (i.e., non-antibiotic classes including disinfectants, antifungals). Aminoglycoside resistance was greater upstream across all WWTPs where ARGs were detected (in four WWTPs). MLS, multidrug, and other resistance classes tended to decrease downstream (n=6 decreased and n=1 increased for all WWTPs). Quinolone was the only resistance class that tended to increase downstream (n=5 WWTPs increased and n=2 decreased). Multidrug resistant ARGs were detected most frequently in upstream and downstream samples across all WWTPs (68 total multidrug ARGs were detected, not including housekeeping genes). There were 12 ARGs (18% of all multidrug ARGs) that decreased downstream in all samples (in at least two sampling locations) (**Figure C.1**). The ARGs that consistently decreased were *PmpM*, *oqx*B, *OprM*, *oprJ*, *mexY*, *mexJ*, *mexB*, *efpA*, *ceoB*, *acrB*, and *abeM* (**Figure 4.4**). Statistical comparisons between upstream and downstream

were limited due to the low numbers of ARGs detected in both upstream and downstream samples; however, differences were found to be significant in SWE P1 for aminoglycoside, multidrug, and tetracycline ($p < 0.05$). Multidrug resistance was also significantly different for CHE P2 ($p < 0.00001$) and beta-lactam resistance was different in IND P1 ($p < 0.05$) (**Table 4.1**).

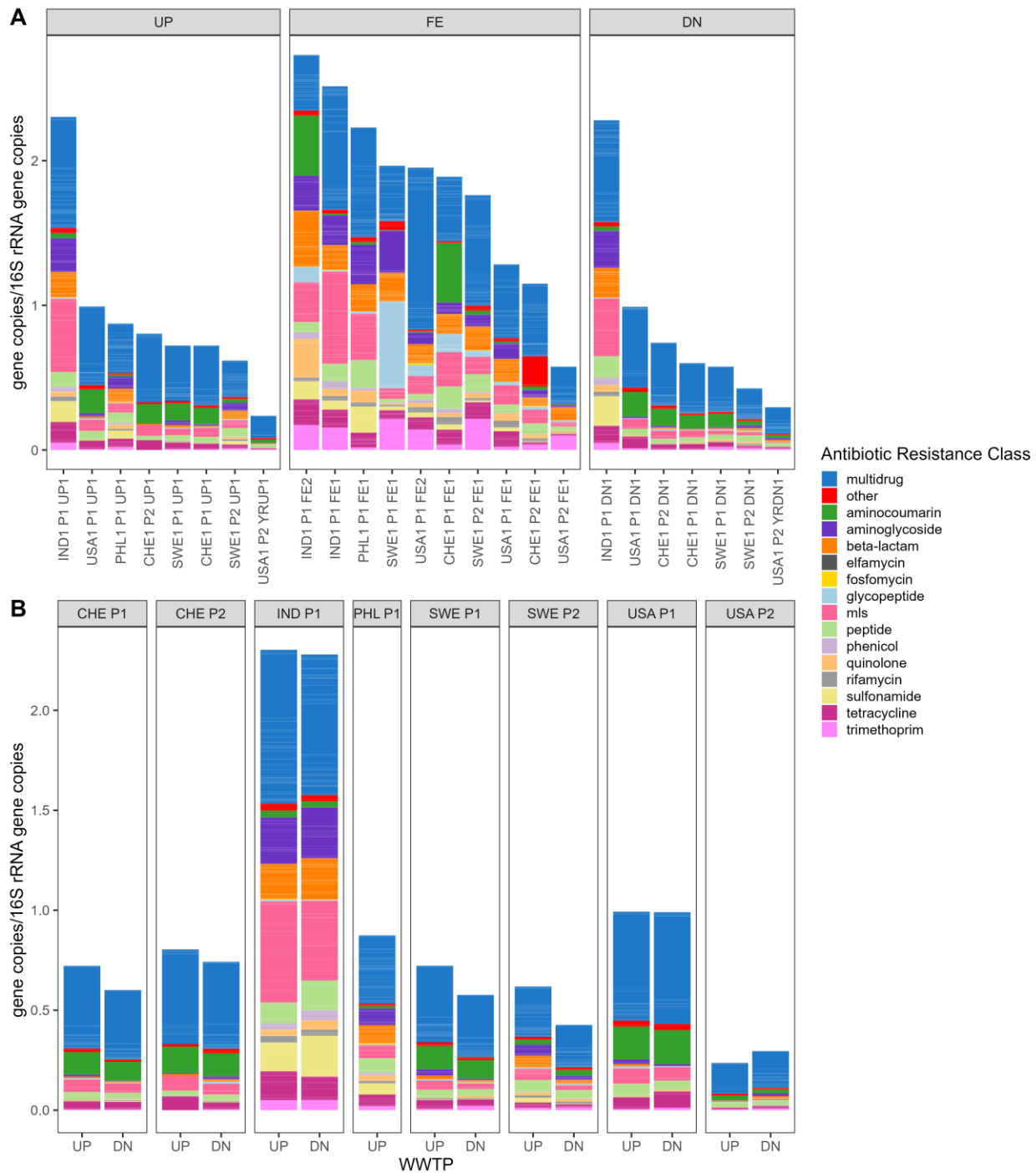


Figure 4.2 Total ARGs in (A) UP, FE, and DN samples and (B) only UP and DN samples. MLS indicates resistance to macrolide, lincosamide, and streptogramin antibiotics

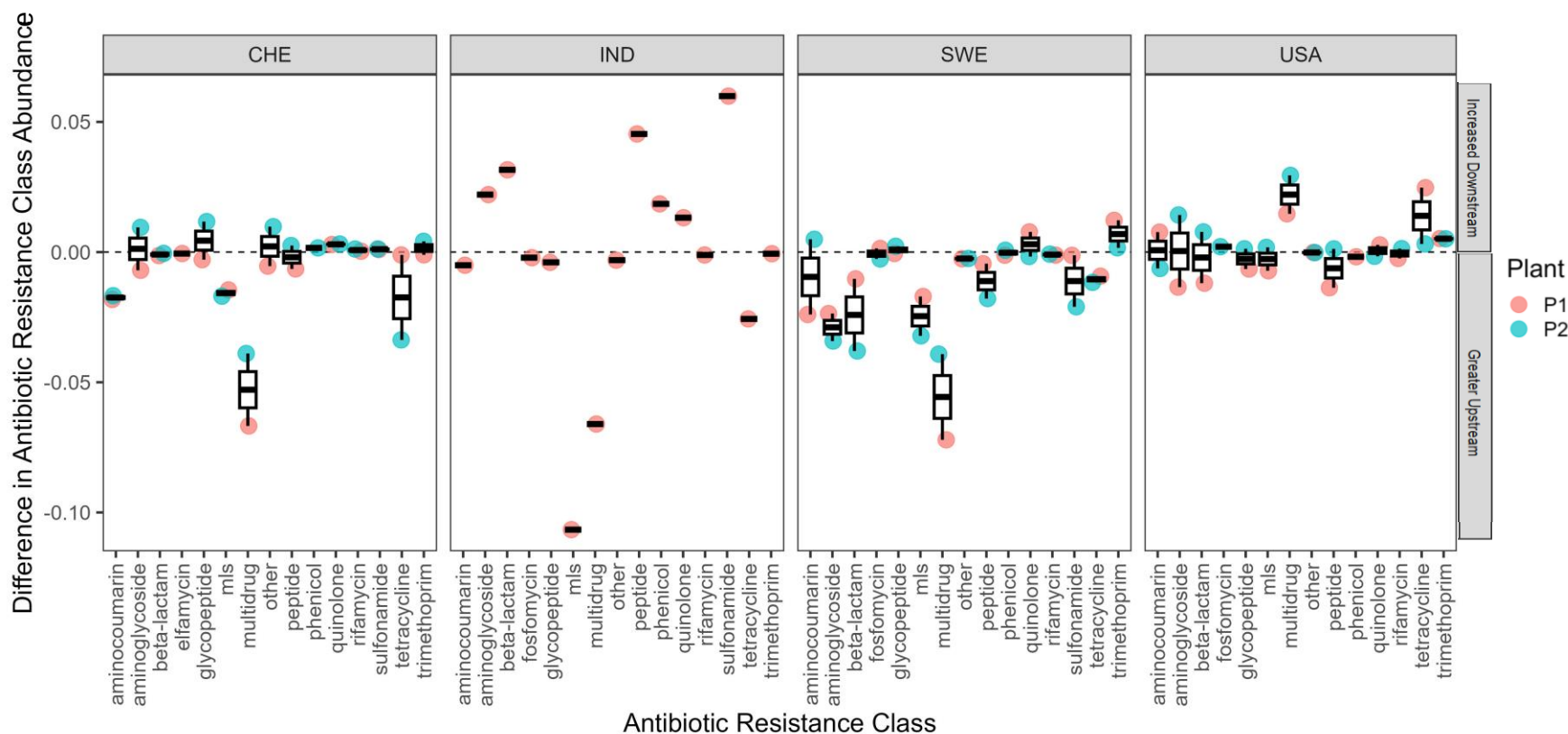


Figure 4.3 Differences in relative abundance for resistance classes UP vs DN. Box-plots display median between WWTPs from the same country. y-axis is relative abundance of total resistance class in downstream minus abundance of total resistance class in upstream for the same WWTP. Housekeeping genes for beta-lactam and multidrug resistance classes were removed.

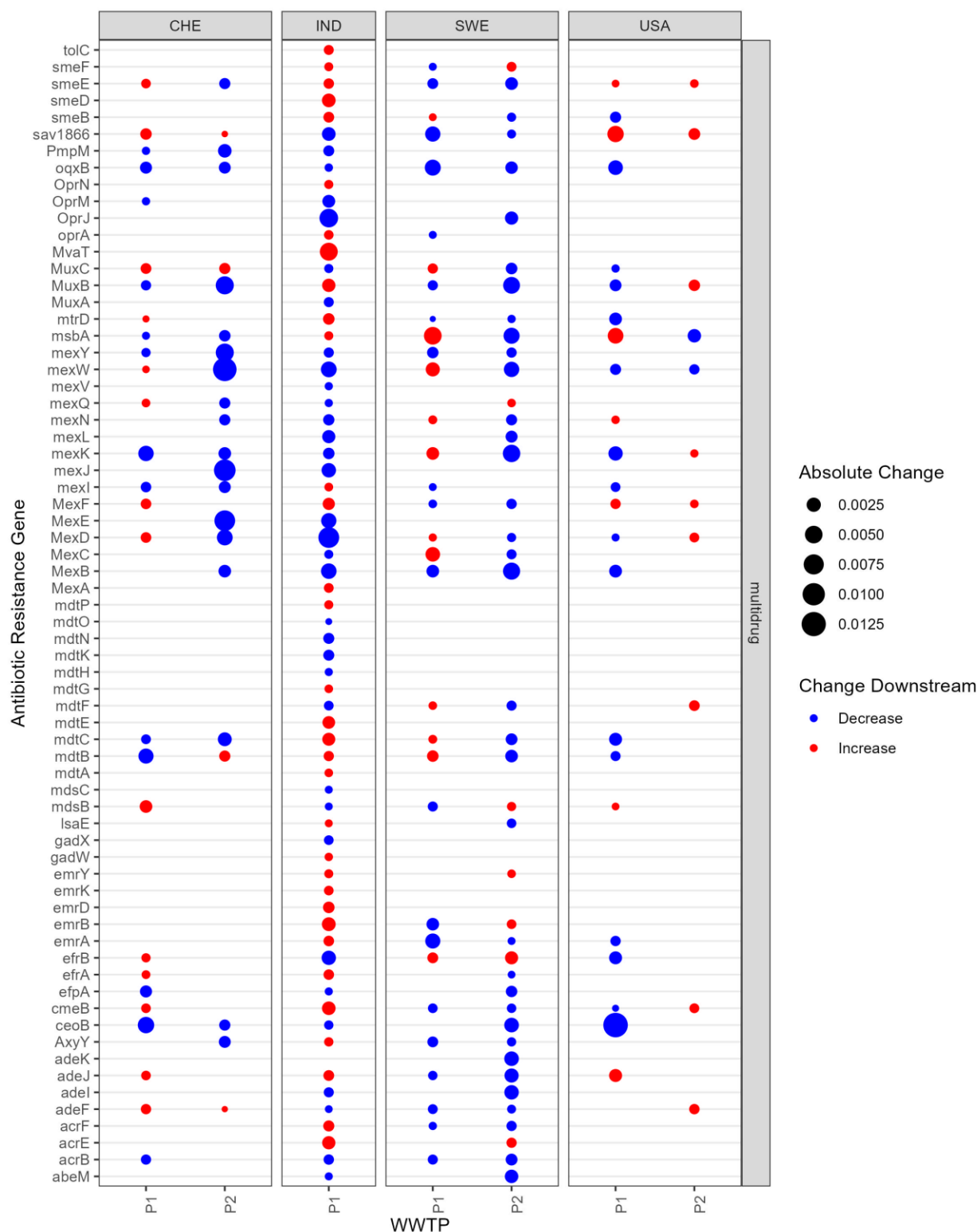


Figure 4.4 Multidrug resistance genes that increased or decreased between upstream and downstream. Removed housekeeping genes (listed **Figure C.2**). This figure does not include genes that were detected only in upstream or downstream. Statistical comparisons were not performed as there were not enough samples to compare upstream and downstream locations.

Table 4.1 Comparing upstream and downstream resistance classes for all upstream and downstream locations by plant and country. # indicates the number of ARGs detected in that resistance class for each WWTP (elfamycin was not included because resistance was only detected once in CHE P1 upstream). The p-value is a result of a Wilcox test comparing upstream and downstream ARGs grouped by resistance class with bolded values indicating significant differences ($p < 0.05$); blank cells indicate that there were not enough ARGs in both groups (upstream and downstream) to make a comparison.

Resistance Class	USA P1		USA P2		SWE P1		SWE P2		CHE P1		CHE P2		IND P1	
	#	p-value	#	p-value	#	p-value	#	p-value	#	p-value	#	p-value	#	p-value
aminocoumarin	4	0.667	4	1	4	0.667	4	0.667	4	1	4	1	4	0.667
aminoglycoside	9	0.883	6	0.333	18	0.0234	35	0.669	10	0.476	9	0.173	59	0.91
beta-lactam	2		6	1	20	0.0777	59	0.447	11	0.0560	12	0.182	136	0.0441
fosfomycin			1		1		2						4	1
glycopeptide	9	0.697	1		9	0.413	15	0.183	8	0.549	5		16	0.0659
mls	11	0.931	6	1	12	0.0732	33	0.538	13	0.836	14	0.142	61	0.67
multidrug	70	0.265	55	0.0407	98	0.0138	137	0.0665	83	0.9820	72	1.3×10^{-6}	190	0.9440
other	11	0.527	10	0.352	17	0.601	17	0.673	13	1	9	0.667	18	0.863
peptide	16	0.263	15	0.524	19	0.487	25	0.935	16	0.382	13	0.0769	28	0.631
phenicol	1				1		6	0.1			1		29	0.844
quinolone	6	0.658	3	0.667	5	0.8	11	0.26	7	0.558	1		25	0.428
rifamycin	2	1	2		8	0.686	11	0.931	6	1	4	1	14	0.798
sulfonamide					3	0.667	5	0.2	3	1	1		8	0.686
tetracycline	13	0.731	17	0.311	20	0.0185	37	0.722	24	0.154	13	0.0513	55	0.846
trimethoprim	4	0.667	4	0.346	6	0.667	11	0.104	3	0.667	2		25	0.395

The resistomes were different across samples as revealed by NMDS when grouped by country both with (PERMANOVA $R^2=0.29$, $p=0.001$) and without ($R^2=0.58$, $p=0.001$) considering final effluent. When grouped by sampling location (i.e., upstream and downstream) the resistomes were different when compared with final effluent ($R^2=0.17$, $p=0.003$) but not when only comparing upstream and downstream ($R^2=0.03$, $p=0.837$) (**Figure 4.5 A and B**). When comparing the ARG alpha diversity, final effluent had a higher Shannon diversity (4.25 ± 0.421), though this was only statistically different from downstream samples (Wilcox, $p=0.018$). Upstream and downstream samples were similarly diverse (3.37 ± 1.11 and 3.21 ± 0.802 Shannon diversity, respectively, Wilcox $p=0.955$) (**Figure 4.5 C**). The average ARG Shannon diversity was similar for both CHE streams (P1: 2.60 ± 0.08 and P2: 2.56 ± 0.29), IND and PHL had similar and higher Shannon diversities (4.52 ± 0.066 and 4.90 , respectively). The USA streams had similar Shannon diversities (P1: 2.54 ± 0.28 and P2: 2.77 ± 0.78). The SWE streams were most different with P2 having a greater diversity than P1 (4.31 ± 0.52 and 2.98 ± 0.03 , respectively) (**Figure 4.5 D**). The Wilcox-test was not able to find differences between upstream and downstream samples when considering all countries together. Furthermore, none of the stream Shannon diversities were different when compared by country (both when including upstream and downstream grouped together and when comparing the upstream and downstream separately).

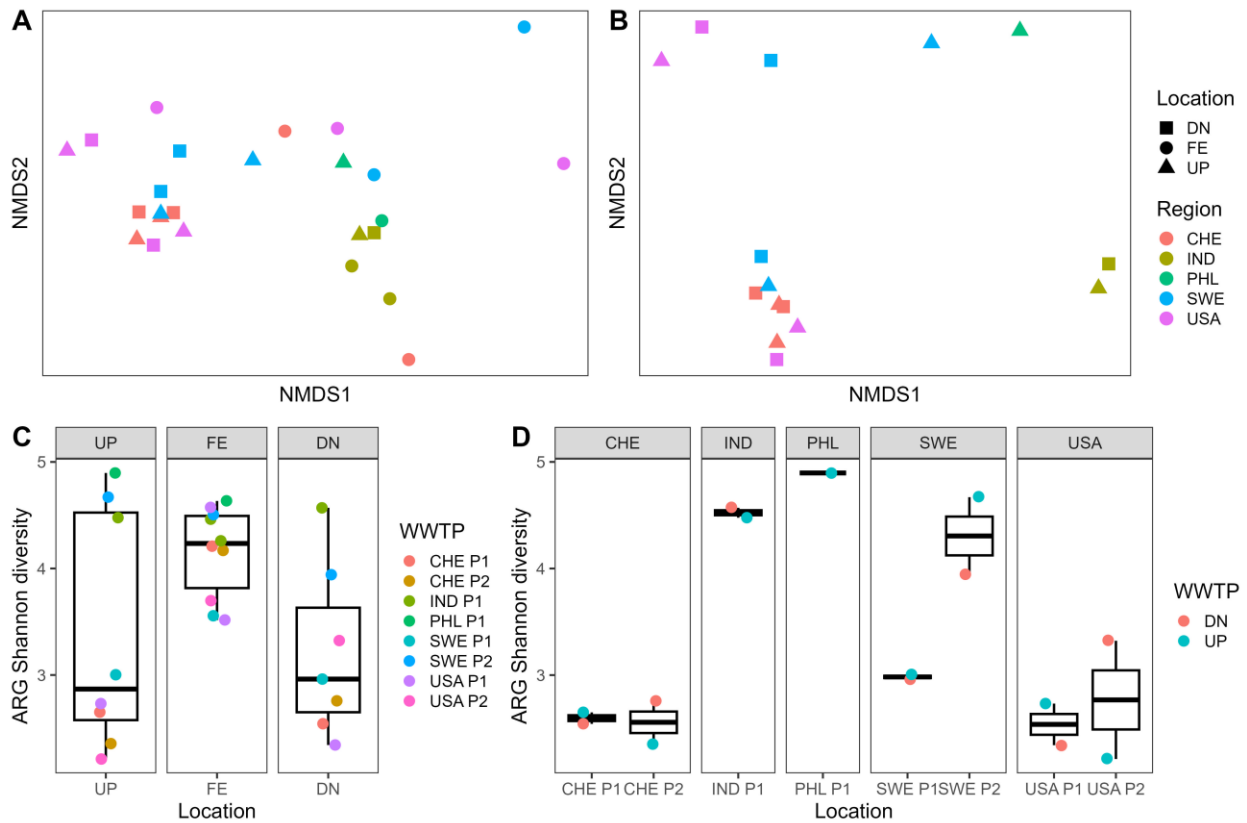


Figure 4.5 Beta-diversity (A) with FE (stress = 0.1369822) and (B) without FE (stress= 0.03223323). (C) Shannon diversities of ARGs for final effluent (FE), upstream (UP), and downstream (DN) and (D) only for UP and DN

Across all samples, the mean total relative abundance for clinically relevant beta-lactam ARGs was similar for upstream and downstream (on average 0.014 ± 0.024 ARG/16S rRNA) and higher for final effluent (0.039 ± 0.062). The mean total relative abundance was higher for all resistance classes in final effluent than upstream and downstream; upstream samples also had slightly higher relative abundances than downstream for clinically relevant beta-lactam, peptide, and quinolone resistance genes. Clinically relevant glycopeptide ARGs were only detected in final effluent in CHE P2, IND P1, and SWE P1 (**Figure 4.6 A**). Clinically relevant beta-lactam ARGs were detected most frequently and at the greatest relative abundance when compared with other clinically relevant resistance classes ($n=409$ unique clinically relevant beta-lactam resistance genes detected across all samples, followed by $n=55$ for peptide, $n=19$ for quinolone and lastly, $n=3$ for glycopeptide). Final effluent samples typically had a greater number of unique clinically relevant ARGs (on average 31 ARGs/sample). The top three final effluent samples having the greatest number of clinically relevant ARGs were USA P1, SWE P1, and IND P1. Upstream samples were split into two groups either having a higher (28 ± 4 clinically relevant ARGs/sample for IND P1, PHL, and SWE P2) or lower (3 ± 0.5 clinically relevant ARGs/sample for USA P1, USA P2 and SWE P1) relative ARG count. With the exception of IND P1 (40 ARGs/sample) all downstream samples had few ARGs (5 ± 1 for SWE P1; SWE P2; CHE P2; USA P2) (**Figure 4.6 B**). The greatest distance between a sampling location and hospital was 7 miles (~ 11 km) (straight distance), however, most of the sampling locations had a hospital within 3 miles (~ 5 km) of the sampling site (**Table C.1**).

4.4.3 Taxonomy

When comparing genera from final effluent, upstream, and downstream the variability between samples were best explained by country (PERMANOVA $R^2=0.39$, $p=0.001$), though there were also significant difference when grouped by sampling location (i.e., upstream, downstream, final effluent) (PERMANOVA, $R^2=0.18$ $p=0.01$). When only comparing upstream and downstream the samples were different when grouped by country (PERMANOVA $R^2=0.57$, $p=0.002$), but not when grouped by stream sampling location (PERMANOVA $R^2=0.035$, $p=0.809$) (**Figure 4.7 A**). When compared together, the mean Shannon diversity indices of genera for upstream and downstream were all very similar (range 1.57–2.49). PHL (UP: 2.49) and SWE P2 had the greatest diversities (UP: 2.55 and DN: 2.44). IND P1 and USA P2 had the lowest mean Shannon diversity indices (mean of UP and DN: 1.58 ± 0.21 and 1.57 ± 0.03 , respectively). None of the upstream and downstream samples were different from each other (Wilcox $p>0.05$), however, these values were compared for all countries grouped together since there were not enough samples to compare upstream and downstream for each location (**Figure 4.7 B**). The mean genera Shannon diversity was highest for final effluent; this is likely due to the biological treatment at the WWTPs which promote bacterial growth to break down organic matter. When comparing upstream and downstream for each WWTP five WWTPs had a greater Shannon diversity downstream than upstream, in comparison only two had a greater upstream diversity. This slight increase in the diversity downstream may be indicative of some of the taxa from final effluent persisting downstream.

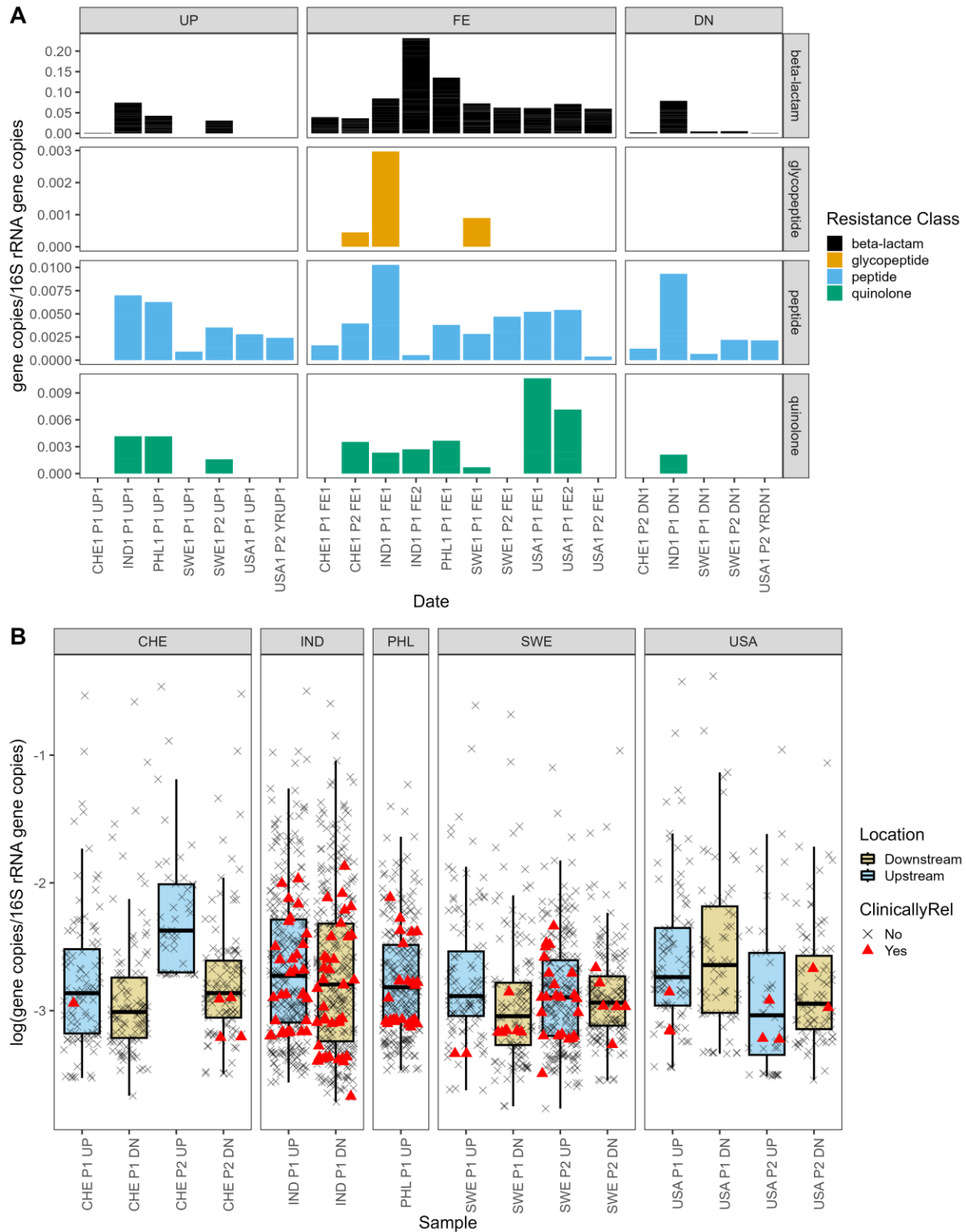


Figure 4.6 Clinically relevant ARGs (A) showing the resistance class and relative abundance of those clinically relevant ARGs for all samples and (B) clinically relevant ARGs for upstream (UP) and downstream (DN) only, the red triangles indicate a clinically relevant ARGs within a sample

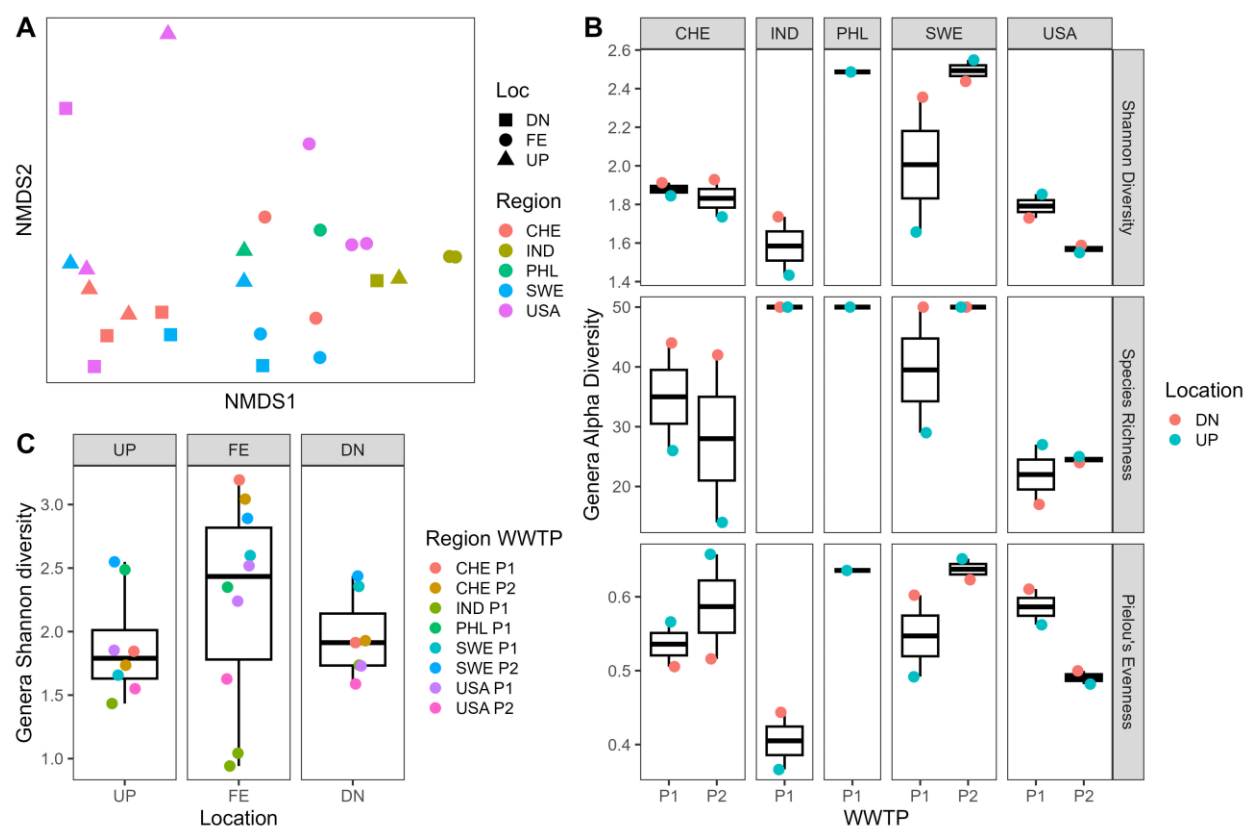


Figure 4.7 (A) NMDS of genera (left) with final effluent (FE) (stress=0.1479573) and alpha diversity of genera (Shannon) (B) with final effluent and (C) for only upstream (UP) and downstream (DN).

The lower Shannon diversity in IND was driven by the uneven distribution of taxa as seen by the similarly low Pielou's evenness index. Species richness was high in IND indicating that there are many unique genera, however, the large proportion of *Thauera* appear to have skewed the diversity, as Shannon diversity takes the proportion of individuals into account. Trends for final effluent were very similar though the Shannon diversity for CHE was especially higher than its upstream and downstream counterparts. Species richness was high across final effluent samples except for USA P2 which was only slightly higher than upstream and downstream (**Figure 4.7 C**).

CHE P1 stream samples had similar relative abundances of unnamed genera of the phylum Actinobacteria and class Alphaproteobacteria dominating both upstream and downstream samples. *Burkholderia* were also in greater relative abundance upstream. The final effluent sample was more different from upstream and downstream samples with abundant genera including: *Acinetobacter*, *Bacteroides*, *Thiomonas*, *Nitrospira*, and *Pseudomonas*. CHE P2 upstream and downstream samples were very similar to each other with the top most abundant genera being *Pseudomonas*, *Limnohabitans* and genera belonging to the phylum Actinobacteria and class Alphaproteobacteria. Final effluent also had high relative abundances of *Pseudomonas*, however the other most abundant genera differed from the stream samples (*Brevundimonas*, *Thiomonas*, *Nitrospira*).

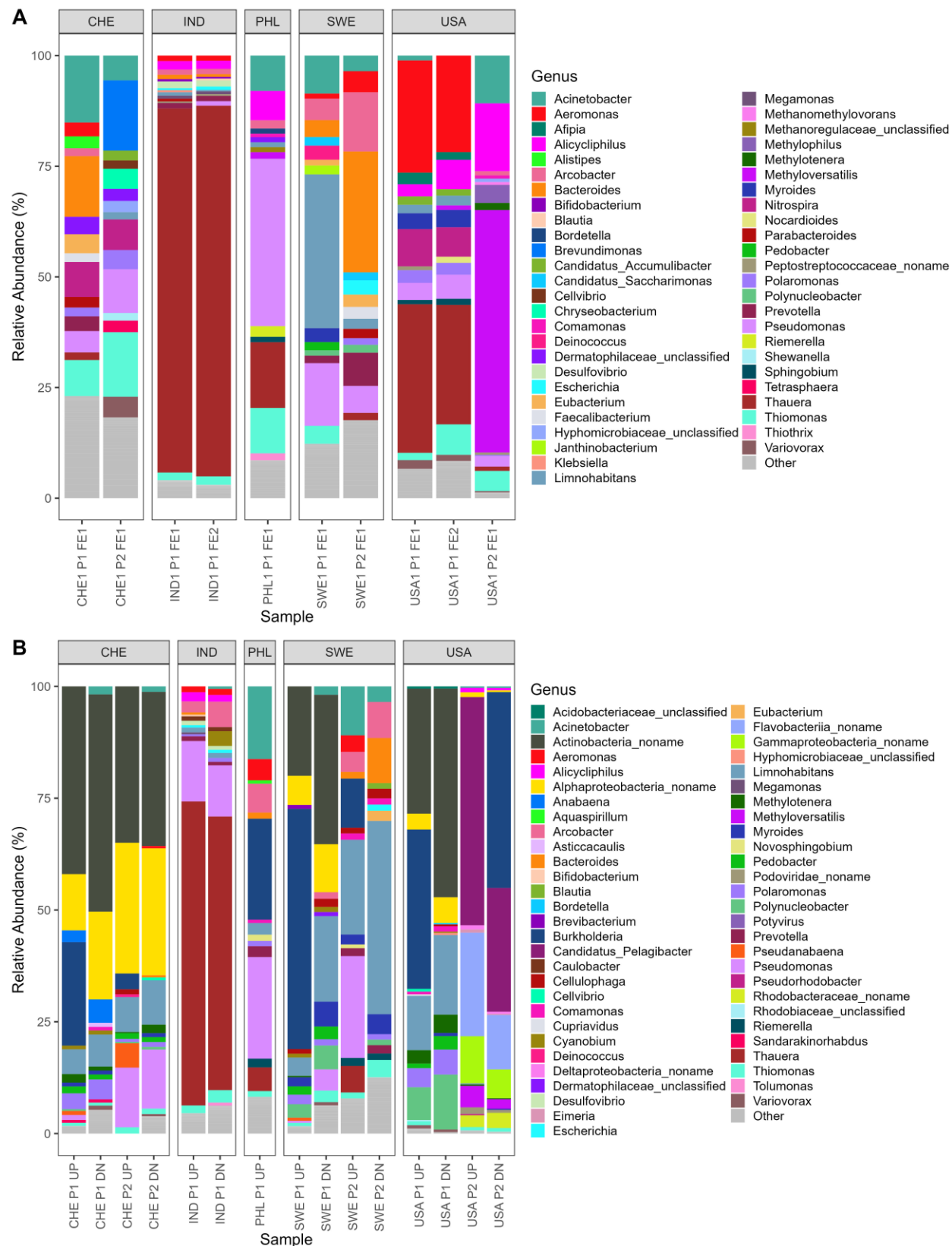
IND stream samples were very similar to final effluent with the most dominant genus across all samples being *Thauera*; *Thiomonas* was also present in all IND samples but at a lower relative abundance. *Pseudomonas* and *Arcobacter* were more highly abundant in IND stream samples whereas *Desulfovibrio* and *Prevotella* were more abundant in the final effluent. *Pseudomonas*, *Thauera*, *Thiomonas*, and *Acinetobacter* genera were dominant in PHL upstream and final effluent samples. *Thiomonas* and *Alicyclophilus* were more dominant in PHL final effluent and *Burkholderia* and *Arcobacter* were more dominant upstream.

SWE P1 upstream was more different from its downstream counterpart with a greater relative abundance of *Burkholderia*. Final effluent and downstream both had greater relative abundances of *Limnohabitans*. Upstream and downstream SWE P1 samples were similar in that they had greater relative abundances of genera belonging to phylum Actinobacteria and class Alphaproteobacteria. SWE P2 also had greater differences between upstream and downstream with upstream having more *Pseudomonas* and *Thauera*. SWE P2 final effluent had greater relative abundances of *Bacteroides* and *Arcobacter* which were also in greater relative abundance downstream.

USA P1 upstream and downstream were also different with the upstream having greater relative abundance of *Burkholderia* (similar to SWE P2 upstream). Both USA P1 upstream and downstream had high abundances of *Limnohabitans*, *Polynucleobacter* and phylum Actinobacteria. Final effluent replicates were similar to each other with greater relative abundances of *Thauera*, *Aeromonas*, and *Nitrospira*. USA P2 upstream and downstream both had high relative abundances of *Candidatus Pelagibacter*, and genera belonging to classes Flavobacteriia, and Gammaproteobacteria. USA P2 downstream also had higher relative abundances of *Burkholderia*. The final effluent was highly dominant in *Methyloversatilis* followed by *Alicyclophilus* and *Acinetobacter* (**Figure 4.8**).

4.4.4 Antibiotics

Concentrations of MLS antibiotics were highest in IND P2 when compared with other sampling locations. MLS antibiotics in IND P2 were higher upstream than downstream (6293 ± 3539 and 2907 ± 2322 ng/L, respectively). The only other location where MLS antibiotics were detected in both upstream and downstream samples was CHE P1, here the concentrations were slightly higher downstream than upstream (8 ± 6 and 4 ± 0 ng/L, respectively). Quinolone antibiotics were only detected in IND P1, IND P2, and SWE P2. Among these locations concentrations of quinolone antibiotics were highest for IND2 and lowest for SWE P2. IND P2 had higher concentrations of quinolone antibiotics upstream than downstream (434 ± 201 and 331 ± 150 ng/L, respectively). IND P1 had higher concentrations of quinolone antibiotics downstream than upstream (220 ± 55 and 62 ± 0 ng/L) and SWE P2 had similar quinolone concentrations upstream and downstream (8 ± 0 and 9 ± 0 ng/L). Sulfonamide antibiotics were only detected in USA, CHE, and SWE samples. The highest concentrations of sulfonamide antibiotics detected were the same for CHE P1 upstream and downstream (53 ng/L); the location with the second highest concentration of sulfonamide antibiotics was USA P2 (18 ± 3 ng/L), all other locations had lower concentrations (on average 4 ± 2 ng/L) (**Figure 4.9**).



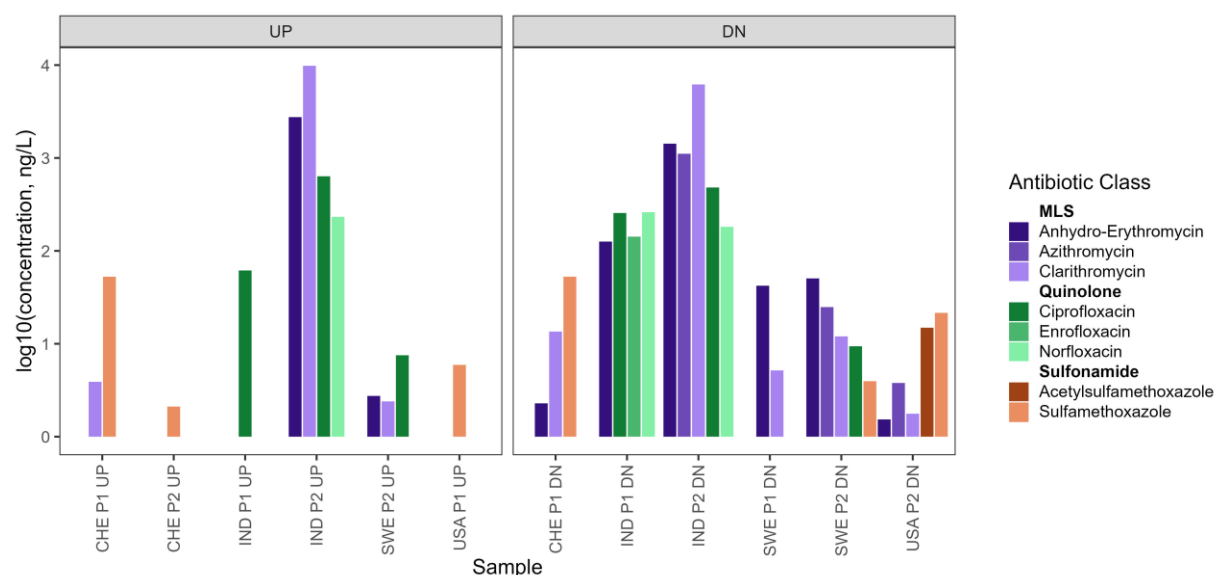


Figure 4.9 Concentrations of antibiotics ($\log_{10}(\text{ng/L})$) for upstream (UP) and downstream (DN) samples. IND samples were collected the year after qPCR and metagenomics samples were collected. IND P2 was included because, while there are no metagenomics data for this date, there are qPCR results.

4.5 Discussion

The qPCR results presented here align with what other studies have reported: an increased relative abundance of ARGs in urban surface water and downstream of WWTPs (11,39,40). Concentrations of ARGs will vary as they typically decrease with increasing distances from the WWTP effluent discharge point (7,41). Anthropogenic sources of ARGs may also be more complex than arising just from WWTP effluents. Damashek and colleagues (42) investigated a North American watershed and found that both human fecal markers and ARGs were strongly correlated with sewage infrastructure, the authors hypothesized that this was due to sewer failings associated with aging infrastructure. Therefore, the origins of ARGs in watersheds are complex and many sources should be considered.

ARGs associated with multidrug resistance frequently decreased downstream of WWTPs, influencing the decrease in total relative abundance. Multidrug ARGs have also been shown to influence lake resistomes where the total relative abundances of ARGs were inversely correlated with the human watershed impact, this relationship was largely driven by bacitracin and multidrug ARG concentrations in Canadian lakes (7). Multidrug ARGs have previously been shown to be highly abundant in environmental samples where resistance is typically conferred through efflux pumps (43). Because efflux pumps are often intrinsic and may be present in bacteria that are both resistant and susceptible to antibiotics, they may not be an ideal measure of anthropogenic impacts on environmental samples (43,44). While the results from this study did not find a clear increase in ARGs due to WWTP discharge a metagenomics study have found increased diversity of taxa and relative abundances of ARGs downstream of WWTPs in Poland (45).

It was expected that WWTPs would contribute to the clinically relevant ARGs in streams since phenotypic metagenomics have revealed increased relative abundances of clinically relevant carbapenem resistant and cephalosporin-resistant bacteria downstream of WWTPs in Swiss rivers (46). The distance to the nearest hospital did not appear to influence the relative abundance of

clinically relevant ARGs detected in the streams. The higher relative abundances of clinically relevant ARGs in the SWE P2 stream may be due to the fact that it is an urban stream that passes through an underground culvert. The higher relative abundances of clinically relevant ARGs in IND and PHL may relate back to the higher rate of urbanization and large populations around the streams (**Figure C.3**). The distance from sampling location to the WWTP outfall must also be considered, while the effluent discharge point was not recorded, the SWE samples were close to the WWTP, whereas most other samples were further removed (with the exception being CHE P2 which was also close to the WWTP).

While the impact of the WWTPs remains unclear from our metagenomics data, the final effluent and stream clustered separately when examined via NMDS which is consistent with the literature. A comparative study of influent around the world found distinct groupings by regions with the resistomes of different countries being strongly linked with socioeconomic factors and best described by parameters associated with sanitation and health measures reported for different countries (for example, investment in water and sanitation was negatively correlated with antimicrobial resistance levels, diarrhea treatment was positively correlated etc.) (47).

There were some downstream samples that had dominant genera more similar with their final effluent counterparts than to their upstream (especially SWE P1 and somewhat SWE P2). Samples typically had a greater relative abundance of *Burkholderia* upstream rather than downstream (except for USA P2 which showed the opposite). Notably IND stream samples were very similar to the corresponding final effluent with a very high relative abundance of *Thauera*. Reportedly the Adyar River has high levels of domestic developments along its banks and untreated human waste makes up a large proportion of the flow during summer months. The anthropogenic influences are evident through high concentrations of dissolved organic nitrogen and due to nitrification and denitrification occurring in the river (48). *Thauera* have been isolated from a number of environments including soils, sediments and WWTPs (e.g., in activated sludge) (49,50). A number of *Thauera* species are known to denitrify and degrade aromatic compounds (49) and have been found in greater relative abundances in Chinese WWTPs treating industrial waste (primarily textile dyeing and pharmaceutical) (51). The similarity of the genera between the IND stream and final effluent samples may be indicative of the higher levels of urbanization and pollution seen around/in the Adyar River.

Burkholderia are common in aquatic environments though their relative abundance has been shown to vary by season (52). *Burkholderia* have been found with greater relative abundances in streams influenced by agriculture with higher levels of dissolved organic carbon and nitrate (53). *Burkholderia* have also been shown to be associated with the occurrence of ARGs in an Indian lake (54). The relative abundances of *Burkholderia* are clearly different between the upstream and downstream samples, these differences may offer additional explanations as to the differences in ARGs found between stream samples. Multidrug resistance genes that decreased downstream in multiple samples included *OprM*, *mexB*, and *mexY*, which are part of a complex of efflux pumps, notably *amrAB-OprM* which is intrinsic to *Burkholderia* (55) and *MexAB-OprM* and *MexXY-OprM* which are intrinsic to *Pseudomonas aeruginosa* (56). The genera *Pseudomonas* also decreased across multiple samples (**Figure C.4**) further lending credence to the tie between genera and resistance genes. The similarity of taxa between upstream and downstream samples has been shown previously as no significant changes in phyla were shown downstream of two U.S. WWTPs (57).

Quinolone resistance genes were detected in all samples, however, clinically relevant quinolone ARGs and antibiotic concentrations revealed matching trends as the genes and

antibiotics were only detected in IND and SWE P2 samples. Antibiotic concentrations were similar between upstream and downstream samples, however, there were more unique antibiotics detected downstream; this trend matches more closely with the qPCR results which revealed that there was an increase in resistance downstream of the WWTPs. The antibiotic, metagenomics, and qPCR results furthermore all agree that there are greater relative abundances of antibiotics and ARGs in IND in comparison to USA, CHE, and SWE.

Klein and colleagues (58) calculated the Drug Resistance Index (DRI) (an aggregated measure of drug resistance combining antibiotic consumption and antibiotic resistant pathogens) for 41 countries between 2012-2015 and found the DRI to be highest for India, indicating lower effectiveness of antibiotic therapies; Sweden had the lowest DRI and the U.S. fell in the middle. Antibiotic consumption rates and the classes of antibiotics most frequently consumed vary across countries. Of the countries in this study IQVIA data from 2015 showed the highest total antibiotic consumption to be in the U.S. (10,298 defined daily doses (DDD)/1000 people), followed by Switzerland (5128 DDD/1000 people). Antibiotic use in Sweden and India were similar (4722 and 4950 DDD/1000 people, respectively). The Philippines had the lowest antibiotic use of the countries in this study at 2600 DDD/1000 people. Beta-lactam antibiotics were most used in these countries though the specific antibiotic (broad and narrow spectrum penicillin and cephalosporin) varied depending by country; overall broad spectrum penicillin antibiotics were most frequently used (59). Interpreting associations is complex as antibiotic concentrations have not shown correlations with ARGs in many studies (13,15,47,60) however, links have been found between the use of specific antibiotic classes (rather than the total use of antibiotics) and ARGs in wastewater influent across countries (47).

Regardless of the impacts or lack therefore seen from WWTPs, examining aquatic environments for ARGs, especially those that are clinically relevant, is an important aspect of monitoring for antimicrobial resistance since people may interact with these aquatic environments. Leonard et al. (61) found that U.K. surfers (as an example of those who are frequently exposed to water that may contain ARB and who might ingest that water) were more likely to have their gut colonized with cefotaxime-resistant *Escherichia coli* (9.1% for surfers' vs 3.1% for non-surfers; risk ratio 2.95, 95% CI= 1.05 to 8.32). Søråas and colleagues (62) found that the strongest risk factor for community-acquired urinary tract (CA-UTI) infections caused by extended-spectrum β -lactamase (ESBL)-*Enterobacteriaceae* was international travel followed by antibiotic use, however, freshwater swimming in the past year (odds ratio=2.1; 95% CI: 1.0–4.0) was also contributing factor of ESBL positive CA-UTI risk in Norwegian populations (62). Further research is necessary to understand the interactions between water and ARB infections/gut microbiome colonization and their impact on public health, however, these studies exemplify the importance of the environment as a potential source of antimicrobial resistance in the public.

4.6 Conclusions

Stream samples saw increases in the relative abundance of certain ARGs and integrase 1 genes downstream of WWTPs, this was especially evident when examined through qPCR. Metagenomics sequencing revealed that the overall resistome of the streams did not appear to be strongly impacted by the WWTPs as determined by broad-scale metagenomic ARG profiling. The differences between stream samples were best explained by country rather than stream sampling location (upstream or downstream); this was to be expected given differences in land use around the streams (e.g., urban, agriculture, forest) and the different communities and behavioral practices

surrounding antibiotic use in each of the countries studied. Additional studies are needed to fully understand the contributions of antimicrobial resistance in aquatic environments.

4.7 Acknowledgements

This manuscript was supported by the U.S. National Science Foundation (NSF) Award OISE: 1545756. Additional funding to D.G.J.L. was provided by the Swedish research council VR (2018-05771) and the Region Västra Götaland under the ALF agreement (grant number ALFGBG-717901). Funding for H.B. was provided by the Swiss National Science Foundation under the NRP72 “Antimicrobial resistance”, grant 167116

4.8 References

1. Laxminarayan R, Duse A, Wattal C, M Zaidi AK, L Wertheim HF, Sumpradit N, et al. Antibiotic resistance-the need for global solutions. *Lancet Infect Dis* [Internet]. 2013 [cited 2018 Oct 24];13:1057–98. Available from: <http://dx.doi.org/10.1016/>
2. Frost I, Van Boeckel T, Pires J, Craig J, Laxminarayan R. Global geographic trends in antimicrobial resistance: the role of international travel. *J Travel Med* [Internet]. 2019 [cited 2021 Feb 4];1–13. Available from: <http://orcid.org/0000-0002-0095-0845>
3. Roope LSJ, Smith RD, Pouwels KB, Buchanan J, Abel L, Eibich P, et al. The challenge of antimicrobial resistance: What economics can contribute [Internet]. Vol. 364, *Science*. American Association for the Advancement of Science; 2019 [cited 2021 Feb 4]. Available from: <https://science.sciencemag.org/content/364/6435/eaau4679>
4. WHO. Global Action Plan on Antimicrobial Resistance [Internet]. Geneva, Switzerland: World Health Organization; 2015. Available from: <https://www.who.int/antimicrobial-resistance/publications/global-action-plan/en/>
5. Chokshi A, Sifri Z, Cennimo D, Horng H. Global Contributors to Antibiotic Resistance. *J Glob Infect Dis* [Internet]. 2019 Jan 1 [cited 2021 Sep 2];11(1):36. Available from: </pmc/articles/PMC6380099/>
6. Auta A, Hadi MA, Oga E, Adewuyi EO, Abdu-Aguye SN, Adeloye D, et al. Global access to antibiotics without prescription in community pharmacies: A systematic review and meta-analysis. *J Infect* [Internet]. 2019 [cited 2021 Feb 4];78:8–18. Available from: <https://doi.org/10.1016/j.jinf.2018.07.001>
7. Kraemer SA, Barbosa da Costa N, Oliva A, Huot Y, Walsh DA. A resistome survey across hundreds of freshwater bacterial communities reveals the impacts of veterinary and human antibiotics use. *Front Microbiol* [Internet]. 2022 Jan 12;13. Available from: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.995418>
8. Scott LC, Aubee A, Wilson MJ, Esser S, Descamps D, Lee N, et al. Leave No Trace? Ecological and anthropogenic determinants of antibiotic resistant bacteria in a recreational alpine environment. *Environ Res* [Internet]. 2023;216(Pt 2):114617. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/36273598>
9. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Sci Total Environ* [Internet]. 2013 [cited 2017 Oct 3];447:345–60. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2013.01.032>
10. Nappier SP, Liguori K, Ichida AM, Stewart JR, Jones KR. Antibiotic resistance in recreational waters: State of the science. Vol. 17, *International Journal of Environmental Research and Public Health*. 2020.

11. Hiller CX, Hübner U, Fajnorova S, Schwartz T, Drewes JE. Antibiotic microbial resistance (AMR) removal efficiencies by conventional and advanced wastewater treatment processes: A review. Vol. 685, *Science of the Total Environment*. 2019.
12. He Y, Yuan Q, Mathieu J, Stadler L, Senéhi N, Sun R, et al. Antibiotic resistance genes from livestock waste: occurrence, dissemination, and treatment. Vol. 3, *npj Clean Water*. 2020.
13. Prieto Riquelme MV, Garner E, Gupta S, Metch J, Zhu N, Blair MF, et al. Demonstrating a Comprehensive Wastewater-Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environ Sci Technol* [Internet]. 2022 Dec 19;56(21):14982–93. Available from: <https://doi.org/10.1021/acs.est.1c08673>
14. ESRI. World Land Cover 30m BaseVue 2013 [Internet]. 2021. Available from: <https://www.arcgis.com/home/item.html?id=1770449f11df418db482a14df4ac26eb>
15. Majeed HJ, Riquelme M V., Davis BC, Gupta S, Angeles L, Aga DS, et al. Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Front Microbiol* [Internet]. 2021 May 13 [cited 2021 Jun 8];12:1048. Available from: www.frontiersin.org
16. Li A-D, Metch JW, Wang Y, Garner E, Zhang AN, Riquelme M V, et al. Effects of sample preservation and DNA extraction on enumeration of antibiotic resistance genes in wastewater. *FEMS Microbiol Ecol*. 2018 Feb;94(2).
17. Singh RR, Angeles LF, Butryn DM, Metch JW, Garner E, Vikesland PJ, et al. Towards a harmonized method for the global reconnaissance of multi-class antimicrobials and other pharmaceuticals in wastewater and receiving surface waters. *Environ Int*. 2019;124.
18. Bibbal D, Dupouy V, Ferré JP, Toutain PL, Fayet O, Prère MF, et al. Impact of three ampicillin dosage regimens on selection of ampicillin resistance in Enterobacteriaceae and excretion of blaTEM genes in swine feces. *Appl Environ Microbiol*. 2007;73(15).
19. Chen J, Yu Z, Michel FC, Wittum T, Morrison M. Development and application of real-time PCR assays for quantification of erm genes conferring resistance to macrolides-lincosamides-streptogramin B in livestock manure and manure management systems. *Appl Environ Microbiol*. 2007;73(14).
20. Colomer-Lluch M, Jofre J, Muniesa M. Quinolone resistance genes (qnrA and qnrS) in bacteriophage particles from wastewater samples and the effect of inducing agents on packaged antibiotic resistance genes. *J Antimicrob Chemother*. 2014;69(5).
21. Pei R, Kim SC, Carlson KH, Pruden A. Effect of River Landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Res*. 2006;40(12).
22. Ng L-K, Martin I, Alfa M, Mulvey M. Multiplex PCR for the detection of tetracycline resistant genes. *Mol Cell Probes* [Internet]. 2001 [cited 2018 Jun 1];15:209–15. Available from: <http://www.idealibrary.com>
23. Dutka-Malen S, Evers S, Courvalin P. Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant enterococci by PCR. *J Clin Microbiol*. 1995;33(1).
24. Hardwick SA, Stokes HW, Findlay S, Taylor M, Gillings MR. Quantification of class 1 integron abundance in natural environments using real-time quantitative PCR. *FEMS Microbiol Lett*. 2008;278(2).
25. Suzuki MT, Taylor LT, DeLong EF. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol*.

- 2000;66(11).
26. Garner E, Chen C, Xia K, Bowers J, Engelthaler DM, McLain J, et al. Metagenomic Characterization of Antibiotic Resistance Genes in Full-Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. *Environ Sci Technol*. 2018 Jun;52(11):6113–25.
 27. Arango-Argoty G, Singh G, Heath LS, Pruden A, Xiao W, Zhang L. MetaStorm: A public resource for customizable metagenomics annotation. *PLoS One*. 2016;11(9).
 28. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2017;45(D1):D566–73.
 29. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(D1).
 30. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Vol. 12, *Nature Methods*. 2015.
 31. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
 32. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests [Internet]. R package. 2021. Available from: <https://cran.r-project.org/package=rstatix>
 33. Gross J, Ligges U. nortest: Tests for Normality. R package version 1.0-4; 2015.
 34. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3–4).
 35. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J*. 1948;27(3).
 36. Pielou EC. The measurement of diversity in different types of biological collections. *J Theor Biol*. 1966;13(C).
 37. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. The vegan Package Title Community Ecology Package [Internet]. R package. 2022 [cited 2022 Jun 2]. Available from: <https://cran.r-project.org/package=vegan>
 38. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
 39. Roberto AA, Van Gray JB, Engohang-Ndong J, Leff LG. Distribution and co-occurrence of antibiotic and metal resistance genes in biofilms of an anthropogenically impacted stream. 2019 [cited 2021 Feb 6]; Available from: <https://doi.org/10.1016/j.scitotenv.2019.06.053>
 40. Kneis D, Berendonk TU, Forslund SK, Hess S. Antibiotic Resistance Genes in River Biofilms: A Metagenomic Approach toward the Identification of Sources and Candidate Hosts. *Environ Sci Technol* [Internet]. 2022 Jan 12;56(21):14913–22. Available from: <https://doi.org/10.1021/acs.est.2c00370>
 41. Lee J, Ju F, Maile-Moskowitz A, Beck K, Maccagnan A, McArdell CS, et al. Unraveling the riverine antibiotic resistome: The downstream fate of anthropogenic inputs. *Water Res*. 2021;197.
 42. Damashek J, Westrich JR, McDonald JMB, Teachey ME, Jackson CR, Frye JG, et al. Non-point source fecal contamination from aging wastewater infrastructure is a primary driver of antibiotic resistance in surface waters. *Water Res* [Internet]. 2022 Jan 12;222:118853. Available from: <https://www.sciencedirect.com/science/article/pii/S0043135422008004>

43. Nesme J, Cécillon S, Delmont TO, Monier JM, Vogel TM, Simonet P. Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr Biol*. 2014;24(10).
44. Sharma A, Gupta VK, Pathania R. Efflux pump inhibitors for bacterial pathogens: From bench to bedside. Vol. 149, *Indian Journal of Medical Research*. 2019.
45. Zieliński W, Hubeny J, Buta-Hubeny M, Rolbiecki D, Harnisz M, Pauksztó Ł, et al. Metagenomics analysis of probable transmission of determinants of antibiotic resistance from wastewater to the environment – A case study. *Sci Total Environ*. 2022;827.
46. Ju F, Lee J, Beck K, Zhang G, Gekenidis M-T, Hummerjohann J, et al. Phenotypic Metagenomics tracks Wastewater-Associated Clinically Important Beta-lactam Resistant Bacteria Invading River Habitats [Internet]. In Review; 2022 Jan. Available from: <https://www.researchsquare.com/article/rs-1589365/v1>
47. Hendriksen RS, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat Commun*. 2019 Mar;10.
48. Nirmal Rajkumar A, Barnes J, Ramesh R, Purvaja R, Upstill-Goddard RC. Methane and nitrous oxide fluxes in the polluted Adyar River and estuary, SE India. *Mar Pollut Bull*. 2008;56(12).
49. Liu B, Mao Y, Bergaust L, Bakken LR, Frostegård Å. Strains in the genus *Thauera* exhibit remarkably different denitrification regulatory phenotypes. *Environ Microbiol*. 2013;15(10).
50. Morgan-Sagastume F, Nielsen JL, Nielsen PH. Substrate-dependent denitrification of abundant probe-defined denitrifying bacteria in activated sludge. *FEMS Microbiol Ecol*. 2008;66(2).
51. Zhang B, Xu X, Zhu L. Structure and function of the microbial consortia of activated sludge in typical municipal wastewater treatment plants in winter. *Sci Rep*. 2017;7(1).
52. Leff LG, Leff AA, Lemke MJ. Seasonal changes in planktonic bacterial assemblages of two Ohio streams. *Freshw Biol*. 1998;39(1).
53. Olapade OA, Gao X, Leff LG. Abundance of three bacterial populations in selected streams. *Microb Ecol*. 2005;49(3).
54. Chakraborty J, Sapkale V, Rajput V, Shah M, Kamble S, Dharne M. Shotgun metagenome guided exploration of anthropogenically driven resistomic hotspots within Lonar soda lake of India. *Ecotoxicol Environ Saf*. 2020 May 1;194:110443.
55. Podnecky NL, Rhodes KA, Schweizer HP. Efflux pump-mediated drug resistance in *Burkholderia*. *Front Microbiol*. 2015;6(MAR).
56. Masuda N, Sakagawa E, Ohya S, Gotoh N, Tsujimoto H, Nishino T. Contribution of the MexX-MexY-OprM efflux system to intrinsic resistance in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother*. 2000;44(9).
57. Lambirth K, Tsilimigras M, Lulla A, Johnson J, Al-Shaer A, Wynblatt O, et al. Microbial Community Composition and Antibiotic Resistance Genes within a North Carolina Urban Water System. *WATER*. 2018 Nov;10(11).
58. Klein EY, Tseng KK, Pant S, Laxminarayan R. Tracking global trends in the effectiveness of antibiotic therapy using the Drug Resistance Index. *BMJ Glob Heal*. 2019;4(2).
59. OneHealthTrust. ResistanceMap: Antibiotic resistance [Internet]. Center for Disease Dynamics, Economics & Policy; 2023 [cited 2023 Jul 1]. Available from: <https://resistancemap.cddep.org/AntibioticResistance.php>

60. Pärnänen KMM, Narciso-Da-Rocha C, Kneis D, Berendonk TU, Cacace D, Do TT, et al. Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Sci Adv.* 2019;5(3).
61. Leonard AFC, Zhang L, Balfour AJ, Garside R, Hawkey PM, Murray AK, et al. Exposure to and colonisation by antibiotic-resistant *E. coli* in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). *Environ Int.* 2018;114.
62. Søråas A, Sundsfjord A, Sandven I, Brunborg C, Jenum PA. Risk Factors for Community-Acquired Urinary Tract Infections Caused by ESBL-Producing Enterobacteriaceae -A Case-Control Study in a Low Prevalence Country. *PLoS One.* 2013;8(7).

APPENDIX C

C.1 Supplementary Figures

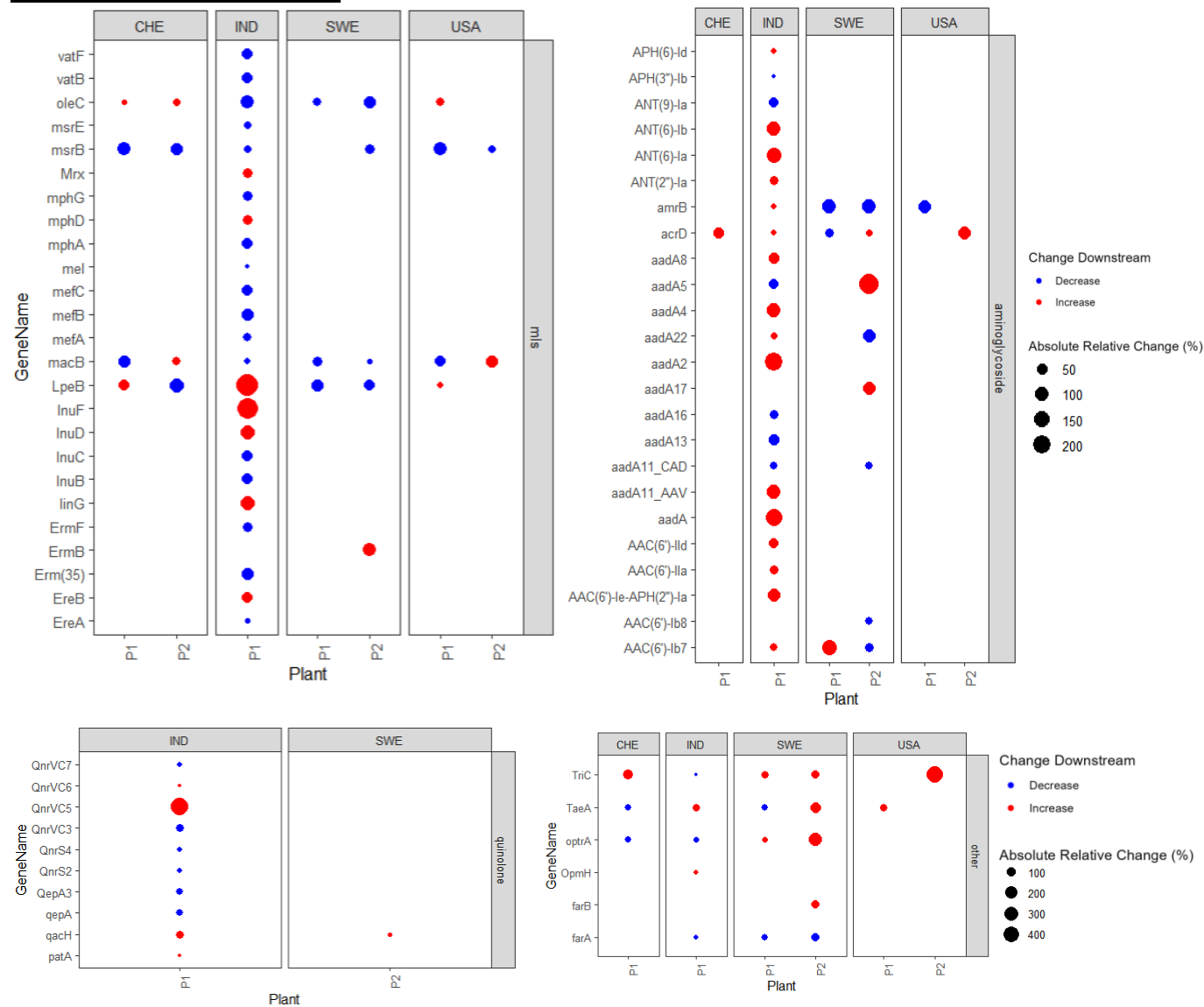


Figure C.1 Changes in ARGs for resistance classes MLS, aminoglycoside, quinolone, and other. Housekeeping genes that removed: multidrug (mupA, mupB, and Bifidobacteria-intrinsic-ileS-conferring-resistance-to-mupirocin), other (LlmA-23S-ribosomal-RNA-methyltransferase), aminoglycoside (kdpE), quinolone (arlR – response regulator, arlS – promoter)

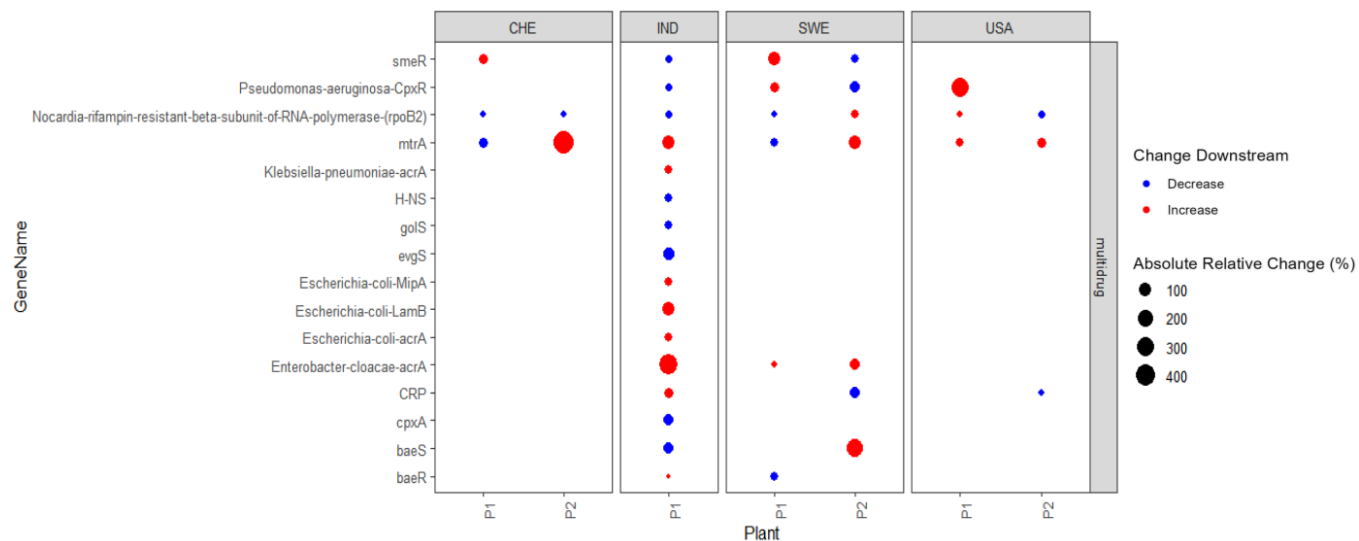


Figure C.2 (A) Multidrug housekeeping genes that removed from Figure 4.4.

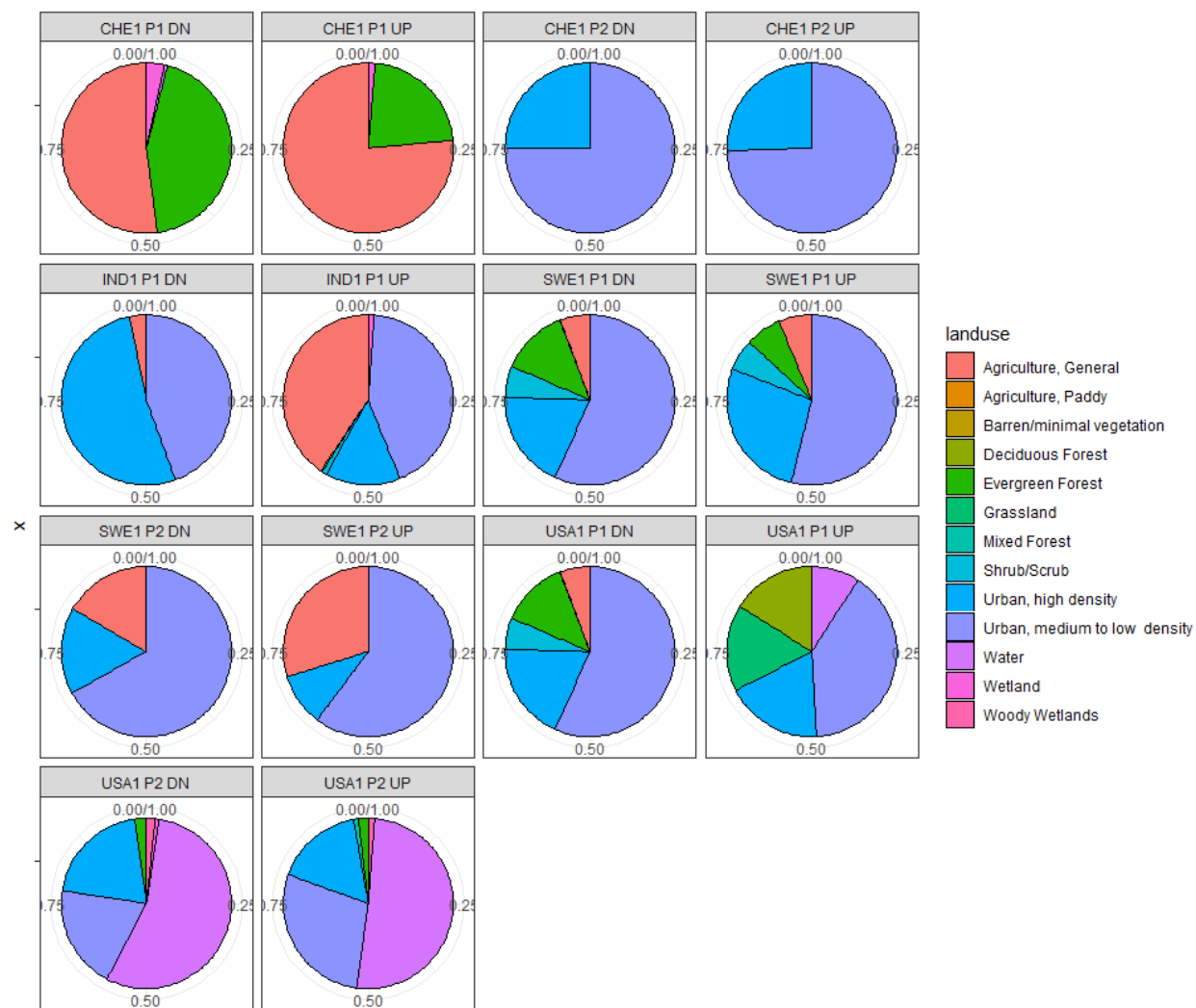


Figure C.3 Land use from ArcGIS. Greatest percentages of land use fell into the following categories: agriculture (general), urban (high density), urban (medium- low density), water (USA P2, York River), forest

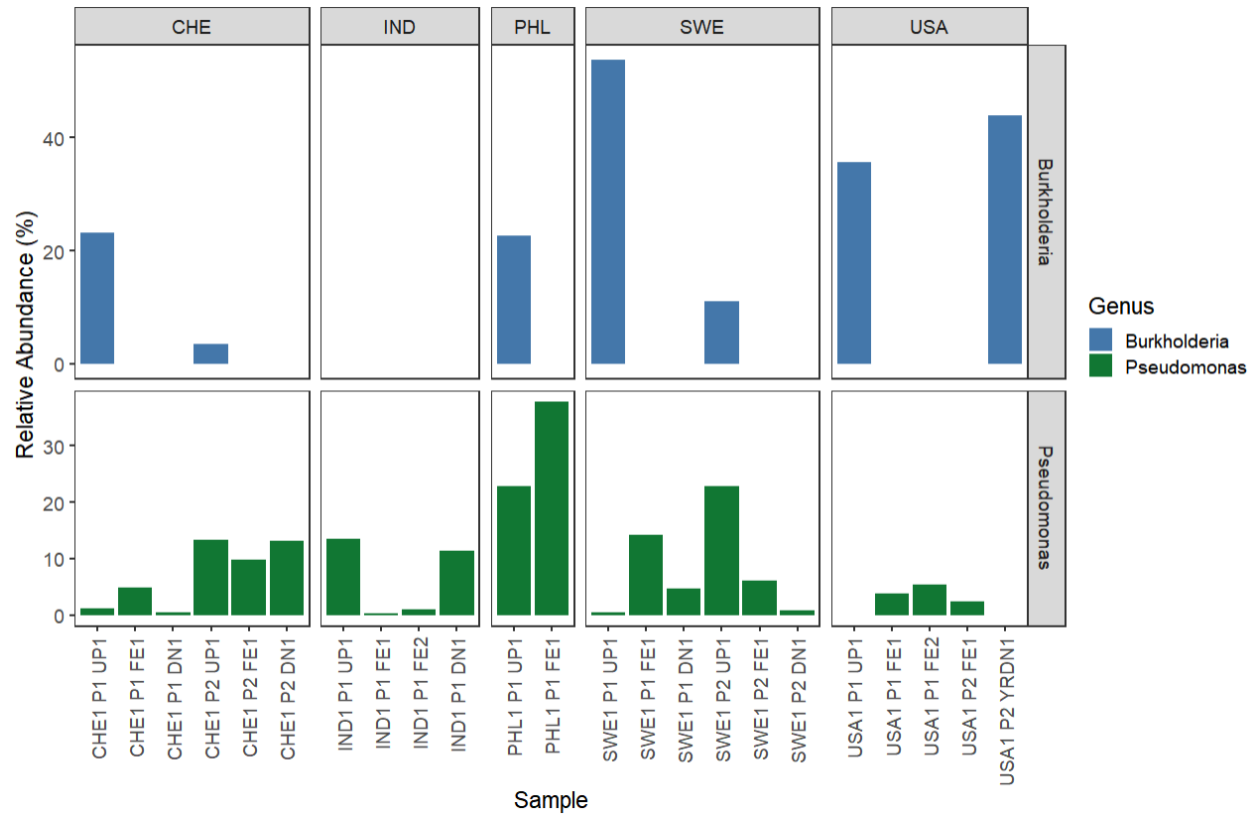


Figure C.4 Showing relative abundances of *Burkholderia* and *Pseudomonas* genera detected across all samples

C.2 Supporting Tables

Table C.1 Site description, Google Earth coordinates can be found at https://earth.google.com/earth/d/1-TGg63PV5zGr2uEyDz6MgYWQD0g_jjcR?usp=sharing

Country	Sample	Date Sampled	Latitude	Longitude	Approx. Distance UP vs DN (m)	River	WWTP contribution to stream flow	~City Population	Straight Distance to Nearest Hospital
Switzerland	P1 UP	5/17/2016	47.0906	8.326701	920	Reuss River	5.9% at low tide	30,926	2.60 mi (4.18 km)
Switzerland	P1 DN		47.0848	8.318091					2.05 mi (3.29 km)
Switzerland	P2 UP	5/18/2016	47.4063	8.595474	230	Glatt River	11% for dry weather (mean 6.08%)	28,667	2.98 mi (4.80 km)
Switzerland	P2 DN		47.4059	8.598409					3.07 mi (4.93 km)
India	P1 UP	3/10/2016	13.0206	80.18794	2600	Adyar River		10,455,606	1.18 mi (1.89 km)
India	P1 DN		13.0271	80.2078					1.30 mi (2.09 km)
Philippines	P1 UP	12/2/2016	14.2586	121.0563	No DN	Silang-Santa Rosa River		3,382,193	1.01 mi (1.63 km)
Sweden	P1 UP	6/8/2016	57.7062	12.92857	400	Viskan River		66,273	1.74 mi (2.79 km)
Sweden	P1 DN		57.7084	12.92663					1.68 mi (2.71 km)
Sweden	P2 UP	6/9/2016	58.3919	13.87852	min ~325	underground culvert		34,466	2.57 mi (4.14 km)
Sweden	P2 DN		58.3893	13.880684					2.79 mi (4.49 km)
USA	P1 UP	11/1/2016	37.1612	-80.5524	8700	New River		22,615	5.56 mi (8.95 km)
USA	P1 DN		37.1431	-80.5628					4.88 mi (7.86 km)
USA	P2 YRUP	1/19/2017	37.2163	-76.4617	770	York River (off Chesapeake Bay)		1.7 million (Hampton Roads) or 184,587 (Newport News)	6.94 mi (11.17 km)
USA	P2 YRUP		37.2185	-76.4551					5.43 mi (8.73 km)

Table C.2 Primers and qualitative control metrics (standard curve R^2 and efficiency) for qPCR

Assay	Primer Sequence (5'-3')	R^2	Efficiency*	Reference
16S rRNA	F: CCGTGAATACGTTTCYCGG R: GGWTACCTTGTTACGACTT	90.37 ± 2.70	0.995 ± 0.005	(Suzuki et al., 2000)
<i>bla_{TEM}</i>	F: TTCCTGTTTTTGCTCACCCAG R: CTCAAGGATCTTACCGCTGTTG	94.93 ± 0.94	0.975 ± 0.011	(Bibbal et al., 2007)
<i>ermB</i>	F: GATACCGTTTACGAAATTGG R: GAATCGAGACTTGAGTGTGC	83.50 ± 3.57	0.996 ± 0.002	(Chen et al., 2007)
<i>intI1</i>	F: CTGGATTTTCGATCACGGCACG R: ACATGCGTGTAATCATCGTCG	80.25 ± 4.00	0.993 ± 0.0004	(Hardwick et al., 2008)
<i>sul1</i>	F: CGCACCGGAAACATCGCTGCAC R: TGAAGTTCCGCCGCAAGGCTCG	90.99 ± 9.63	0.992 ± 0.006	(Pei et al., 2006)
<i>vanA</i>	F: GGGAAAACGACAATTGC R: GTACAATGCGGCCGTTA	93.80 ± 5.45	0.992 ± 0.003	(Dutka-Malen et al., 1995),
<i>qnrS</i>	F: CGACGTGCTAACTTGCGTGA R: ATGCAAGTTTCCAACAATGCC	73.08 ± 4.23	0.992 ± 0.009	(Colomer-Lluch et al., 2014)
<i>tet(X)</i>	F: CAATAATTGGTGGTGGACCC R: TTCTTACCTTGACATCCCG	100.75 ± 3.37	0.988 ± 0.004	(Ng et al., 2001)

*efficiency = $10^{(-1/\text{slope})-1}$

Table C.3 Raw reads and quality control (QC) sequence metrics performed by MetaStorm

Country	WWTP	Location (Replicate)	Raw Read Pairs	Read Pairs after QC	% Passing QC
CHE	P1	UP	9004487	8691210	96.5
CHE	P2	UP	1868000	1819278	97.4
IND	P1	UP	12640514	12141491	96.1
PHL	P1	UP	8959862	8700123	97.1
SWE	P1	UP	8772912	8501184	96.9
SWE	P2	UP	9589213	9316428	97.2
USA	P1	UP	7786876	7537489	96.8
USA	P2	UP	7886940	7633592	96.8
CHE	P1	FE	10683841	10489905	98.2
CHE	P2	FE	15187655	14849493	97.8
IND	P1	FE1	13738688	13451894	97.9
IND	P1	FE2	7693535	7566499	98.3
IND	P1	FE	8126881	7861533	96.7
IND	P2	FE	14989241	14368769	95.9
IND	P2	FE	8797468	8519103	96.8
PHL	P1	FE	15641018	15018872	96.0
SWE	P1	FE	17807714	17467346	98.1
SWE	P2	FE	15280977	15019026	98.3
USA	P1	FE1	10321000	9948362	96.4
USA	P1	FE2	14017561	13529393	96.5
USA	P2	FE	12906961	12690702	98.3
CHE	P1	DN	12646625	12114752	95.8

CHE	P2	DN	10315780	10020992	97.1
IND	P1	DN	15189431	14872349	97.9
SWE	P1	DN	13183487	12777027	96.9
SWE	P2	DN	12582924	12200283	97.0
USA	P1	DN	6524827	6303807	96.6
USA	P2	DN	7818468	7538990	96.4

CHAPTER 5: CONCLUSIONS

Antibiotic resistance is a major threat to public health. Sampling sewage influents at WWTPs provide capacity for surveillance of antibiotic resistance at the community scale through the examination of resistance determinants (e.g., ARGs, ARBs, antibiotics). WWTPs act as barriers preventing the dissemination of pollutants into the environment and surrounding communities, however, they are not typically designed to remove antibiotics or ARGs. Accordingly, it is important to monitor such determinants during and after treatment to see what is being released into the environment. This research effort sought to characterize the resistome of wastewater in three different stages: 1) ARGs entering WWTPs under the lens of the COVID-19 pandemic to reveal how changes in human behavior during the pandemic influenced the resistome. 2) What happens to the resistome during biological treatment when hospital wastewater, known to harbor high quantities of resistant organisms and ARGs, was mixed with domestic wastewater prior to treatment. And finally, 3) the impact that WWTP effluent had on receiving stream resistomes in North America, Europe, and Asia.

The research presented in Chapter Two showed that WWTP resistomes and microbiomes were variably impacted by the pandemic. The WWTP receiving a larger percentage of university wastewater had lower diversity and a distinct microbiome when university classes were primarily held virtually. This research demonstrated the value of using wastewater as a monitoring tool as SARS-CoV-2 RNA concentrations detected in the wastewater were positively correlated with COVID-19 clinical cases reported for the areas studied. This research focused on rural WWTPs and therefore additional research is needed to determine how the pandemic influenced ARGs and taxa of wastewater in urban communities and in countries with different responses to the pandemic.

In Chapter Three domestic wastewater was dosed with 0, 1, and 10% hospital wastewater which was then fed to sequencing batch reactors (SBRs). The overall resistome, genera important in biological treatment, and genera that may contain pathogenic species remained stable despite the addition of hospital wastewater; this research demonstrated the robustness of activated sludge treatment. Notably, the clinically relevant ARG *vanA* was detected at greater frequencies in the hospital wastewater dosed SBRs, indicating that while the resistome may not have changed as a whole, hospital wastewater may still be contributing important ARGs to WWTPs that warrant further monitoring.

Chapter Four examined the impact of WWTP effluents on stream resistomes and found that when examining specific target ARGs through qPCR there was an increase in relative abundance of these genes downstream of the WWTPs. Metagenomic sequencing revealed that the resistomes were similar both upstream and downstream of the WWTPs. The stream resistomes and their genera varied more greatly between countries and sampling sites than between their upstream and downstream counterparts thus highlighting that environmental resistomes may be more strongly influenced by the surrounding community than by the WWTP effluents.

Overall, the research presented here demonstrates the complexity of wastewater resistomes and the many ways that metagenomic sequencing of wastewater may be used to quantify antibiotic resistance in a community, within WWTPs, and in the receiving environment. This research was focused either on rural wastewater treatment plants (Chapter Two) or sequencing batch reactors (Chapter Three) and was limited by a low number of sampling replicates collected as grab samples (for Chapter Four). Additional research is needed to continue to understand the wastewater resistome and how it is impacted by the community and how it in turn impacts the community once discharged into the environment.

Research on the impact of the COVID-19 pandemic is ongoing with many studies looking at antibiotic use and antibiotic resistant nosocomial infections. Additional wastewater-focused research should accommodate a wider range of communities to be able to distinguish how sewage resistomes were impacted in urban settings and under varying COVID-19 burdens. Future research efforts would benefit by the establishment of stronger connections and communications between public health officials and agencies. These connections would require translation of concentrations and trends measured in wastewater into a metric that public health officials can easily understand. These connections would also allow for public health officials to directly voice their concerns with regards to AMR to researchers and would allow scientists and engineers to target ARGs that are of particular concern within local communities.

The impact that hospital wastewater has on the wastewater resistome was explored in a controlled setting, however, restrictions due to the pandemic limited conclusions that could be drawn from this study. Future research is needed to better quantify the impact of hospital wastewater and how this translates to real world settings. All future studies using metagenomic sequencing should consider deeper sequencing depths along with long-read sequencing where possible to be able to better determine not only the differences or similarities within a community but also to if shifts in ARGs are due to changes in the microbial community harboring these genes or if more/different bacteria having acquired ARGs.

APPENDIX D

D.1 Sub-sewershed SARS-CoV-2 Wastewater Surveillance and COVID-19 Epidemiology Using Building-specific Occupancy and Case Data

This manuscript has been published:

Reprinted with permission from:

Cohen, A., Maile-Moskowitz, A., Grubb, C., Gonzalez, R.A., Ceci, A., Darling, A., Hungerford, L., Fricker Jr, R.D., Finkelstein, C.V., Pruden, A. and Vikesland, P.J. (2022). Subsewershed SARS-CoV-2 Wastewater Surveillance and COVID-19 Epidemiology Using Building-Specific Occupancy and Case Data. *ACS ES&T Water*. 2(11):2047–2059. Copyright (2022) American Chemical Society

Abstract

To evaluate the use of wastewater-based surveillance and epidemiology to monitor and predict SARS-CoV-2 virus trends, over the 2020-2021 academic year we collected wastewater samples twice weekly from 17 manholes across Virginia Tech's main campus. We used data from external door swipe card readers and student isolation/quarantine status to estimate building-specific occupancy and COVID-19 case counts at a daily resolution. After analyzing 673 wastewater samples using RT-qPCR, we re-analyzed 329 samples from isolation- and non-isolation dormitories and the campus sewage outflow using RT-ddPCR. Population-adjusted viral copy means from isolation dormitory wastewater were 48% and 66% higher than unadjusted viral copy means for *N* and *E* genes (1,846 to 2,733/100mL/100people, and 2,312 to 3,828/100mL/100people, respectively; n=46). Pre-specified analyses with random-effects Poisson regression and dormitory/cluster-robust standard errors showed that detection of *N* and *E* genes were associated with increases of 85% and 99% in the likelihood of COVID-19 cases eight days later (IRR=1.845, p=0.013, and IRR=1.994, p=0.007, respectively; n=215), and one-log increases in swipe card normalized viral copies (copies/100mL/100people) for *N* and *E* were associated with increases of 21% and 27% in the likelihood of observing COVID-19 cases eight days following sample collection (IRR=1.206, p<0.001, n=211 for *N*; and IRR=1.265, p<0.001, n=211 for *E*). One-log increases in swipe normalized copies were also associated with 40% and 43% increases in the likelihood of observing COVID-19 cases five days after sample collection (IRR=1.403, p=0.002, n=212 for *N*; and IRR=1.426, p<0.001, n=212 for *E*). Our findings highlight the use of building-specific occupancy data and add to evidence on the potential of wastewater-based epidemiology for predicting COVID-19 trends at sub-sewershed scales.

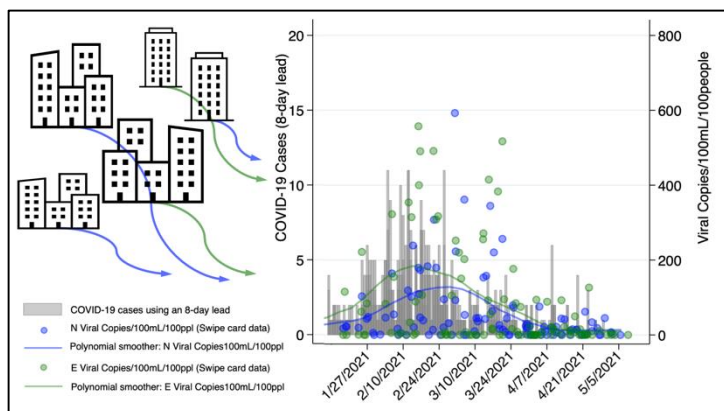
Keywords

Wastewater-based surveillance, wastewater-based epidemiology, sub-sewershed, COVID-19, SARS-CoV-2, digital droplet PCR

Synopsis

Wastewater samples, population estimates, and case outcome data from individual buildings can be used to monitor and predict COVID-19 trends

Table of Contents Art



Introduction

Wastewater-based surveillance (WBS) has long been used as a supplement to clinical surveillance for diseases such as polio (since the 1940s),¹ as well as for other purposes such as monitoring illicit drug use,^{2,3} and supporting vaccination campaigns aimed at eradicating poliomyelitis and vaccine-derived poliovirus outbreaks.⁴ During the early stages of the current coronavirus disease 2019 (COVID-19) pandemic, it was established that SARS-CoV-2, the virus that causes COVID-19, was shed in feces.^{5,6} Subsequently, globally distributed teams adopted WBS to monitor and assess infection trends, particularly in regions with limited clinical testing.

SARS-CoV-2 may be quantified in clinical or wastewater samples through the detection of viral RNA; this requires targeting viral genes that are unique to the virus. Early in the pandemic, national health agencies and private companies worked to develop diagnostic kits primarily using reverse transcription quantitative polymerase chain reaction (RT-qPCR).⁷ The World Health Organization (WHO) released guidelines for diagnosing COVID-19 in which they recommended nucleic acid amplification methods (e.g., RT-qPCR) that target regions on the N, E, and S genes of the virus.⁸ The U.S. Centers for Disease Control and Prevention (CDC) developed its own primers targeting three regions of the nucleocapsid gene (N1, N2, N3),⁹ whereas others developed primers targeting the envelope (E) and spike protein (S) genes.¹⁰ RT-qPCR results yield cycle threshold (Ct) values that indicate the number of cycles needed for a sample to amplify above a specified threshold level, and these Ct values are expected to be inversely correlated with viral concentrations.¹¹

Once it was shown that viral RNA was detectable in wastewater by RT-qPCR and could be used to predict trends in clinical cases,¹² other groups adopted WBS as a means to attempt to monitor and predict COVID-19 infection trends and outbreaks.^{13–16} Building on this work, in mid 2020 Virginia Tech (VT) - a large university that comprises the majority of the population of rural Blacksburg, VA - added wastewater surveillance to the university's COVID-19 operational plan^{17,18} with the goal of providing additional data to support campus-wide public health decision-making and the allocation of testing resources for the university's on-campus residential facilities (dormitories). Many other universities and colleges also initiated COVID-19-focused WBS programs for their campuses as well as surrounding communities.^{19–21} While WBS can be used to support disease monitoring and mitigation efforts, there are many challenges associated with the use of WBS for the detection and monitoring of viruses such as SARS-CoV-2. For one, viral shedding timing and concentrations vary by pathogen, pathogen variant, and infected individual, meaning that depending on the size of a given population and case incidence, WBS

results may or may not reliably reflect infection trends.^{22–24} Constituents that are common in wastewater may also inhibit the amplification of genes, resulting in false negatives.²³ The manner in which samples are collected also plays an important role, as does the method of analysis.²⁵ Composite and grab sampling techniques must be considered along with heterogeneity in the volume, flow, and nature of the wastewater sampled.²⁶ With regard to extending WBS to wastewater-based epidemiology (WBE), in addition to the challenges associated with measuring and quantifying the signal in wastewater samples, sufficiently reliable and linked outcome data are needed, as are data on the size and nature of the population contributing to the wastewater being sampled. Estimation bias associated with these interrelated issues varies spatially and temporally, as well as by scale of sampling and analysis - from sewershed, to sub-sewershed, to building-specific WBS.

In the process of assisting the VT administration with the establishment of a COVID-19 WBS program that could effectively generate data to support efforts to monitor and control SARS-CoV-2 infections on campus, we identified knowledge gaps that needed to be addressed to advance application of WBS and WBE of COVID-19 more broadly. The overarching objective of this research was to improve understanding of the relationships between SARS-CoV-2 detected in sewage and COVID-19 case data at building-scale, and to improve guidance on sub-sewershed sample collection and analysis strategies. To achieve this, we collected and analyzed wastewater samples collected from a cross-section of buildings across campus, with an emphasis on dormitories, to quantify SARS-CoV-2. To extend from WBS to WBE, we coupled these data with building-specific high-resolution occupancy data and student COVID-19 case status data. With regard to specific research objectives, we sought to evaluate whether analysis of unadjusted and population-adjusted SARS-CoV-2 viral concentrations in wastewater samples could be used to reliably predict cases of COVID-19 at building-specific scales using pre-specified lead times from sample collection to case identification. We anticipate that findings from this research can be used to help improve WBS and WBE application at sub-sewershed scales.

Methods

Our analyses were based on a number of data sources described in more detail in the sections that follow. Briefly, wastewater samples were collected from multiple sites/buildings, as well as the university's primary wastewater outflow, twice weekly, starting September 2020 through May 2021 (the VT fall 2020 semester started in August and concluded in December 2020, and the spring 2021 semester started in January and concluded in May 2021). Building occupancy estimates were derived from two data sources: swipe card data for those buildings with exterior doors operated via card readers, and weekly occupancy counts from university administration. Building-specific estimated COVID-19 cases were derived from data on student-specific assignments to isolation and quarantine. University-wide COVID-19 case data were obtained from the university's health center and COVID-19 dashboard management team.

Our study was approved by VT's Institutional Review Board (#21-110). Prior to data analysis we prepared a pre-specified statistical analysis plan (uploaded to Open Science Framework).²⁷

Wastewater sampling sites & protocols

With the broader goal of sampling widely across the campus, sampling sites were selected based on practicality (layout and accessibility of the sewer) and with a focus on sites from which we could collect outflow from specific dormitories/buildings, rather than those sampling sites with mixed wastewater flows from multiple residential and non-residential buildings (though we did also collect samples from several multi-building collection points). Wastewater samples were collected twice weekly (typically Mondays and Thursdays) from 17 manholes across the VT main campus in Blacksburg, VA from September 17, 2020 to May 6, 2021.

During the fall 2020 semester, grab samples were collected at all locations until composite samplers became available in October 2020. Depending on manhole access and depth, as well as other logistical considerations, samples were collected as 24-hour composite samples where possible, and as grab samples otherwise. Composite samples were collected using either ISCO 6712 or ISCO 3700 (Teledyne Inc., Louisville, KY) automatic samplers every 30 minutes for 24 hours (7am-7am) on weekends and during the week. Some sampling dates were missed due to extreme weather, university holidays (e.g., Thanksgiving, November 21-29, 2020), and the winter break between the fall and spring semesters (December 18, 2020 to January 19, 2021). Data on selected sampling sites and methods by building and semester are summarized in the Supporting Information (SI) Table S1.

With regard to protocols, one liter was collected from each site and stored on ice until transported to the laboratory for analysis (typically within 1.5-2 hours of collection). A subset from each sample was prepared for further processing following the methods described in Ahmed et al. (2020).²⁸ Briefly, $MgCl_2$ was added to 150-200mL samples for a final concentration of 25mM and the pH was adjusted to 3-4. To quantify loss through filtering and extraction, bovine coronavirus (CALF-GUARD; Zoetis, Parsippany, NJ) was spiked into the samples at 1 μ L BCoV/1mL of sample (between 1.43E7 and 1.90E7 copies). Following these adjustments, samples were filtered through 0.45 μ m mixed cellulose ester filters (ThermoFisher Scientific, Waltham, MA) using sterilized vacuum filter holders within a BSL-2 hood. Filters were then rolled and torn into 3-5mm sized pieces using flame sterilized tweezers, and then transferred into centrifuge tubes. Torn filters were stored at -80°C until they could be transported on ice to the Molecular Diagnostics Laboratory at the VT Fralin Biomedical Research Institute where RNA extraction and RT-qPCR was performed (described below).

Extraction & PCR

To determine the viral concentrations in wastewater samples we used two PCR approaches. Samples were labeled using coded identifiers so those conducting the RT-qPCR and RT-ddPCR analyses would not know which sites samples had been collected from. Initially, we used an RT-qPCR assay developed by Ceci et al (2021),²⁹ who had received Emergency Use Authorization (EUA #200383) from the U.S. Food & Drug Administration (FDA). We opted to use this approach because it was FDA authorized, used for clinical testing for the VT campus (as well as for samples from Health Districts across southwest VA), and because the Molecular Diagnostics Laboratory at the VT Fralin Biomedical Research Institute had the capacity and staff to analyze wastewater samples in a timely fashion (typically providing results the same day processed samples were received).

Specimens from filter membranes were eluted in a DNA/RNA stabilizer buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 2 mM KH₂PO₄, 50 mM KCl, 8 mM DTT, 40 Units/mL of RNasin, pH7.6) following a mechanical procedure to facilitate the release of particles from the membrane. An aliquot of the eluate (300µl) was subjected to a TRIzol/EtOH-based nucleic acid extraction protocol in a 96-well format column system (ZYMO RESEARCH) following standard extraction/washing procedures, as described in Ceci et al. (2021).²⁹ Total RNA was eluted in water and quantified using a nanodrop system (Nanodrop 200c [Thermofisher], RNA concentrations varied from 5ng/µL to 280ng/µL depending on the original sample). Five ng of RNA were used as template in a *Power SYBR[®] Green RNA-to CT[™] 1-Step* (Thermofisher) RT-qPCR reaction, in a formulation that contains RNase inhibitor and additives to reduce the primer's secondary structure formation. Under these conditions, false positives resulting from primer dimers were less likely to occur.

Three sets of SARS-CoV-2 genes, *N*, *E*, and *S*, along with a housekeeping gene (hRPP30), were evaluated for amplification in independent reaction mixtures (standard curves were included in each plate and for each run, and curves were considered acceptable when R² values were >0.98). This approach allowed for higher specificity (and helped reduce the would-be number of false negative results among clinical samples). Serial dilutions (10,000 to 10 copies per reaction) of 2019-nCoV_N_Positive control (200,000 copies/ul, IDT) and 2019-nCoV_N_Hs_RPP30_Positive control (200,000 copies/µL, IDT) plasmids were used for each standard curve. Because the primers were designed to anneal in the 5' and 3' flanking regions of the target control gene there was no need to linearize either plasmid control. Cut-off Ct and 95% confidence interval (CI) values were experimentally determined based on standard curves. The limit of detection (LoD) was 10 copies per reaction and determined as previously described.²⁹ Negative controls were used to monitor cross-contamination during RNA extraction and sample set up (no template RT-qPCR reaction for each set of primers tested in each plate). Heat-inactivated SARS-CoV-2 (ATCC cat# VR-1986HK) was used as a positive amplification control for *N*, *E*, and *S* sets of primers.

To explore the potential for improved detection sensitivity,²⁵ following analysis of the RT-qPCR results from all collected wastewater samples (described below), in November 2021, RNA extracts from sampling sites matched to dormitories for which we had relatively complete occupancy and COVID-19 case data (n=280 samples), as well as RNA extracts from the campus outflow site samples (n=49), were sent on dry ice to the Hampton Roads Sanitation District (HRSD), Virginia Beach, VA, for independent analysis using reverse transcription digital droplet polymerase chain reaction (RT-ddPCR). For this subset of samples, a one-step RT-ddPCR multiplex was used to quantify *N* and *E* genes in samples, using no template controls, field blanks, and Twist Synthetic SARS-CoV-2 RNA Control 4 (Twist Bioscience, San Francisco, CA, USA) positive standards, according to methods described in Gonzalez et al. (2021).³⁰ LoD, irrespective of sample volume, was 2.2 and 16 copies per reaction for the *N* and *E* gene assays, respectively. RT-ddPCR total process percent recoveries using spiked bovine coronavirus had a mean of 4.12, and a standard deviation (SD) of 5.22.

COVID-19 Case Data

We used dormitory-specific student case status data as the primary outcome variable for our WBE analyses. During the 2020/21 academic year, VT students living in on-campus housing who tested

positive for COVID-19, or who were in close-contact to someone who tested positive (as identified via contact tracing), were assigned by the administration to isolation (for diagnosed cases) or quarantine (for those likely exposed to diagnosed cases). These data were received in the form of daily reports (similar to that of hotel bookings) with information on current assignment status (isolation or quarantine), status start and end dates, off- or on-campus housing status, and the names of the dormitories where students were assigned (for isolation) or already residing (for quarantine). See Figure S1 for aggregate assignment data by day of the week and semester.

For our analyses and reporting of results we used codes in place of names for university dormitories, other buildings, and wastewater sampling sites. The university used one dormitory for housing students assigned to COVID-19 isolation (referred to here as the isolation dormitory, or building code 25), as well as a second dormitory used when the primary isolation dormitory was at capacity (building code 19). We used building code 99 to represent the sampling point for wastewater outflow from most of the university main campus.

We used a silo-based approach for data management, in that only two members of the team interacted with the full and identified data provided by the university administration. Identified data were stored on a secure server and we used automated code/scripts to remove personal identifiers and extract and collate daily isolation and quarantine counts per building per day for students living in on-campus housing (i.e., the number of new students assigned to isolation or quarantine by building and by day, as well as running tallies of the total number of students by assignment status and building).

In addition to these building-specific data, for analysis of wastewater samples from the campus outflow site we also used data from the university's online COVID-19 dashboard, including the daily number of student positive cases and the daily positive testing rate (these data included results from students residing in on- and off-campus housing).

Building-specific Occupancy Data

We used data from two sources to estimate daily and weekly occupancy rates by building. Daily estimates were derived from swipe card data from VT buildings with swipe card readers on all external doors. For our primary WBS and WBE analyses (described below), we focused on those dormitories that required the use of a card swipe for entry 24 hours/day, and for which we had building-specific wastewater sample collection. We also considered analysis of swipe card data from other mixed-use buildings that included student housing as well as other facilities and offices, but because such buildings were open to the public during regular operating/business hours we were unable to derive sufficiently accurate occupancy estimates. Disaggregated swipe card data for the five VT dormitories we focused on in our study were provided in an automated fashion on a daily basis, and included counts for all student and non-student unique (i.e., card-specific) and total swipes per day.

In addition to the swipe card data, we also received weekly counts of the number of people assigned to all residential buildings on VT's main campus (provided every Monday for the previous week's data). As part of our analysis, we evaluated associations between these official weekly counts and the swipe card data. Because the swipe data evidenced substantial variance across the week, and better accounted for occupancy during weekends and university holidays or breaks, for our primary

analyses we used the swipe card data to estimate daily occupancy, based on the total number of swipes per day by students and non-students.

Statistical Analyses & Modeling

For the RT-qPCR results, mean Ct values were calculated for each *N*, *E*, and *S* gene measurement by taking the arithmetic mean of the two values for each site by sample, ignoring Ct values of ≥ 45 (i.e., non-detects, based on the limit of detection for the assay used)²⁹ and missing values; in instances where only one value was available and the other was a non-detect, the available value was used for the mean (“*CTmean_g*”). We also created variables for the mean of the Ct means in two ways: when values for all three sites (*N*, *E*, and *S*) were available (“All NES Ct Mean”), and a mean of the Ct means derived from all or any available Ct mean data (“Any NES Ct Mean”). In addition, we generated binary variables for each *N*, *E*, *S* gene mean (detect, non-detect) as well as a binary variable (detect, non-detect) derived from the mean of Ct means (“Any NES Ct Mean as Binary”).

We used plate-specific standard curves to calculate estimated viral copies per 100mL for *N*, *E*, and *S* genes individually (“*ViralCopies_g*”), as described in Equation 1.

$$ViralCopies_g = 20 * e^{\left(\frac{CTmean_g - \beta_p}{\alpha_p}\right)}$$

Equation 1: Equation for calculating viral copies from RT-qPCR Ct data

Where *CTmean_g* is the arithmetic mean for genes *g*, *g* = *N*, *E*, or *S*; β_p and α_p are the beta and y-intercept coefficients from the standard curve from each plate *p*; 20 is a constant to adjust for the qPCR dilution factor; and non-detect values (Ct ≥ 45) were coded as zeros. After calculating estimated viral copies for *N*, *E*, and *S* genes individually, we created a variable for mean viral copies (“Any NES Mean Viral Copies/100mL”), calculated in the same fashion as described above for the mean of mean Ct values.

For RT-ddPCR runs, samples were considered positive/quantifiable if there were at least three positive droplets per reaction. A reaction was rerun if less than 10,000 total droplets were generated for a PCR reaction. Only those RT-ddPCR *N* and *E* viral copy data that were greater than the observation-specific LoD were used for analyses, treating all values below the LoD as zero.

We visually inspected data distributions using histograms as well as QQ-plots to assess normality. We also pre-specified that no imputation would take place for missing data (but that we would conduct some missing data/sensitivity analysis for those variables missing >10% of expected observations).

To adjust Ct and viral copy data by building-specific populations, we used the total unique daily swipe counts per building, for students and non-students (since non-students would also be contributing to the wastewater outflows from the buildings), as well as the weekly occupancy data, to calculate mean Ct and estimated viral copies per 100mL per 100 people by building. To assess potential associations using SARS-CoV-2 signals from the campus outflow sampling site, we used historic data from the university COVID-19 dashboard on the total number of positive tests for students university-wide (as well as the daily positivity rate).

With respect to analyzing associations between Ct and viral copy data and our primary outcome (proxy data for COVID-19 cases by building), we pre-specified that - data allowing - we would assess three different approaches to weighting the *N*, *E*, and *S* site data: Equal weighting; 50/25/25 (upweighting *N*); and learning optimal weights via a machine-learning based approach (which we did not pursue given the relatively low proportion of detects from the RT-qPCR results). While our overarching objective was to analyze the signal from the wastewater with the highest resolution possible, we also planned to analyze the data by discretizing the wastewater results/signal into either trinary (low, medium, high) or binary (detect, non-detect) variables.

For our primary analyses and statistical modeling of the potential association between SARS-CoV-2 signals in the wastewater samples and COVID-19 cases for the five dormitories for which we had high-resolution occupancy and case data, we planned to analyze associations between wastewater samples on day *X* and corresponding case data on day *X*+8 (i.e., using an eight-day lead for the case outcome data). The choice to focus our analyses using an eight-day case-lead was informed by an interest in better understanding the potential of WBS as an early warning system, our understanding of COVID-19 research literature on incubation periods at the time we prepared our statistical analysis plan (finalized July, 2021), and our desire to limit the total number of comparisons (models) conducted. Although we pre-specified analyzing case data 8 ± 2 days following wastewater sample collection, given advancements in the understanding of SARS-CoV-2 incubation and shedding dynamics, 24,31–33 we modified this to analyze a range of four to nine days after wastewater sampling, as well as 3-day bins of cases centered on eight days. With regard to normalization, we pre-specified the use of three normalization schemes (weekly occupancy data, daily swipe card data, and a combination of the two), but based on the nature of the occupancy data we primarily used swipe card data for our analyses.

To assess building-specific associations between SARS-CoV-2 detection in wastewater samples and associated COVID-19 cases we used random-effects Poisson regression with cluster-robust standard errors (SE), treating each of the five dormitories in our primary analysis as a cluster (to control for variance within and between clusters). Prior to running the models, data were organized with buildings/clusters serving as panel variables and the date (daily) as the time variable; unadjusted and population-adjusted viral copy data were log₁₀ transformed, coding non-detects and values <LoD as zero. We exponentiated model beta coefficients to yield incident-rate ratios (IRR), which - given the underlying structure of our data - may be interpreted as a percentage change in likelihood. SEs and associated p-values were calculated using a standard confidence level of 95%.

All statistical analyses were conducted using Stata (v16.1; by AC), and primary analyses were replicated using R (4.0.2; by CB).

Results & Discussion

Wastewater Sample Collection & Assays for SARS-CoV-2 Detection

Over the 2020-2021 academic year, we analyzed 673 wastewater samples collected across the VT campus for SARS-CoV-2 genes *N*, *E*, and *S* using RT-qPCR (Figure 1). As shown in Table 1Error! Reference source not found., across all 673 wastewater samples analyzed for SARS-

CoV-2 using RT-qPCR, *N*, *E*, and *S* genes were detected ($C_t < 45$) in 7.7% ($n=52$), 12.5% ($n=84$), and 21.6% ($n=145$) of samples, respectively.

As discussed above, following initial analyses of all RT-qPCR results, 329 samples were independently re-analyzed using RT-ddPCR with the goal of obtaining higher resolution quantification of SARS-CoV-2 genes *N* and *E*. Across this subset, *N* and *E* genes were detected in 61.1% ($n=201$) and 35.0% ($n=115$) of samples, respectively. This was a substantial improvement in sensitivity relative to detection by RT-qPCR, where *N* and *E* genes had been detected in 7.0% ($n=23$) and 12.2%, respectively ($n=40$), of this subset of samples. RT-ddPCR LoD data as well as viral copies/100mL for *N* and *E* genes, with and without adjustment for LoD (i.e., viral copies if $> \text{LoD}$ or if > 0), are summarized in Table S2. Comparisons of the percentages of samples with SARS-CoV-2 detection by building code (and for the outflow site) based on RT-qPCR and RT-ddPCR results are provided in Table S3.



Figure 1: A composite sampler surrounded by safety fencing in August of the 2020 fall semester (a), another composite sampler surrounded by safety fencing in January of the 2021 spring semester (b), and a submerged strainer (connected to tubing affixed to a PVC pipe) inside a manhole connected to a composite sampler (photo credits: A. Maile-Moskowitz).

Table 1: Wastewater Samples & SARs-CoV-2 Detection by Gene & Building Codes: RT-qPCR and RT-ddPCR Comparison

	Gene	SARs-CoV-2	Building Codes for Dormitories with Site-specific Sampling Points							Building Codes for Grouped/Clustered Buildings and Sampling Points							Campus Outflow	Total
			B8	B13	B15	B17	B19	B22	B25	G1	G2	G3	G4	G5	G6	G7	B99	
RT-qPCR	<i>N</i>	Not Detected	46	41	47	19	16	47	42	46	50	42	41	47	43	46	48	621
		Detected	1	6	5	0	0	2	6	3	2	7	6	3	6	2	3	52
		Total	47	47	52	19	16	49	48	49	52	49	47	50	49	48	51	673
	<i>E</i>	Not Detected	44	44	43	19	14	40	39	43	47	40	40	40	45	45	46	589
		Detected	3	3	9	0	2	9	9	6	5	9	7	10	4	3	5	84
		Total	47	47	52	19	16	49	48	49	52	49	47	50	49	48	51	673
	<i>S</i>	Not Detected	38	42	37	17	14	37	35	34	37	40	39	37	43	44	34	528
		Detected	9	5	15	2	2	12	13	15	15	9	8	13	6	4	17	145
		Total	47	47	52	19	16	49	48	49	52	49	47	50	49	48	51	673
RT-ddPCR	<i>N</i>	Not Detected	26	23	19	10	11	15	15								9	128
		Detected	23	26	32	9	6	33	32								40	201
		Total	49	49	51	19	17	48	47								49	329
	<i>E</i>	Not Detected	36	36	31	15	15	26	24								31	214
		Detected	13	13	20	4	2	22	23								18	115
		Total	49	49	51	19	17	48	47								49	329

Note: RT-qPCR and RT-ddPCR data are from samples collected in the fall 2020 and spring 2021 semesters.

Visual inspection of histograms and QQ-plots for RT-ddPCR viral copy results for *N* and *E* showed right-skewed distributions prior to log-10 transformation (see Figures S2 and S3), as was also the case for RT-qPCR results. RT-ddPCR viral copy results for *N* and *E* genes were highly correlated ($r=0.967$), with slightly higher viral copy concentrations for *E* than for *N* overall (Figure S4). These RT-ddPCR results, and higher average LoD across samples for the *E* gene (mean=244, SD=64) compared with the *N* gene (mean=33, SD=9; see Table S2), were likely due to differences in assay design and target region characteristics. A comparison of viral copy concentrations for RT-ddPCR and RT-qPCR results for *N* and *S* (for those samples for which data were available for pairwise comparison) is provided in Figure S5. Unless otherwise noted, all subsequent results and analyses reported herein are based on the RT-ddPCR data (using only observations >LoD).

Building-specific COVID-19 Case Data & Occupancy Estimation

The cumulative weekly numbers of students from five non-isolation dormitories (building codes 8, 13, 15, 17, and 22) assigned to isolation and quarantine for the fall 2020 and spring 2021 semesters are shown in Figure 2 (assignment to isolation and quarantine by day of the week and semester shown in Figure S1).

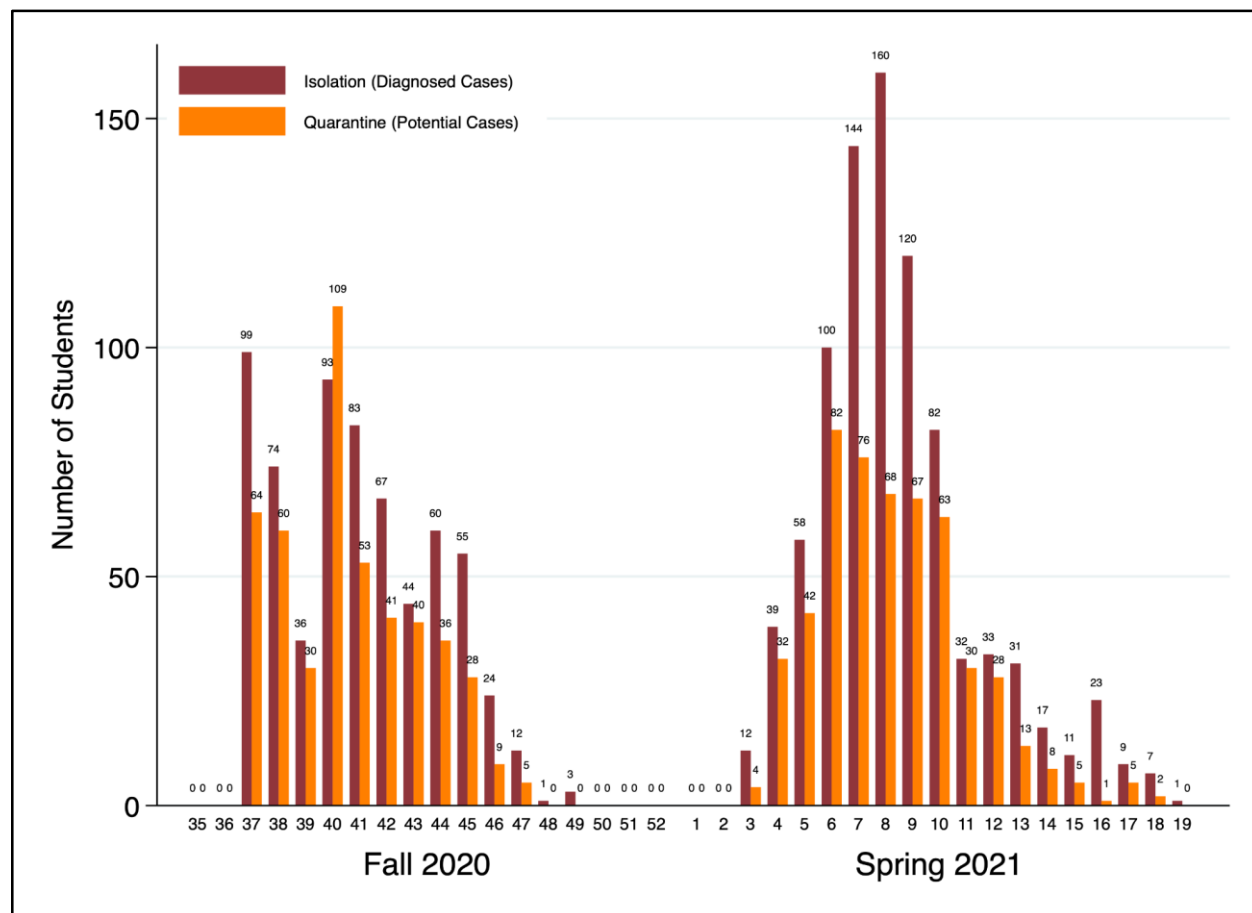


Figure 2: Number of students assigned to COVID-19 isolation or quarantine status by calendar week of the year for five dormitories (building codes B8, B13, B15, B17, and B22).

During the fall 2020 semester at VT, in response to the COVID-19 pandemic courses were scheduled to resume in an online format following the November Thanksgiving holiday.

Consequently, most students residing on campus left before the holiday break began on November 21, 2020; this shift in on-campus residency is reflected in the isolation and quarantine data in Figure 2 for the 47th and 48th weeks of the calendar year. Aggregated occupancy estimates by building code derived from daily swipe card data for students and others, students only, and based on weekly counts, are summarized for the fall and spring semesters in Table 2 based on occupancy data from September 1, 2020 through November 25, 2020, and from January 15, 2021 through May 13, 2021. The large SDs observed (relative to the means) for the swipe card data are due primarily to inter- and intra-weekly variation, as well as relative decreases in student occupancy in the weeks before and after the winter break between semesters, as shown in Figure S6 for the five non-isolation dormitories.

Table 2: Building Occupancy Estimates by Building/Site Code

Building Code	Daily Swipe Card Data for Students & Others				Daily Swipe Card Data for Students Only				Weekly Occupancy Data			
	Median	Mean	SD	n	Median	Mean	SD	n	Median	Mean	SD	n
B8	424	396.3	94.0	204	412	386.6	91.6	204	396	405.1	32.9	191
B13	418	394.1	90.3	204	402	380.8	88.0	204	413	420.7	28.0	191
B15	443	423.2	74.5	193	421	404.7	70.7	193	517	530.6	20.3	191
B17	63	66.3	17.6	204	57	61.5	17.0	204	58	69.7	13.7	191
B22	727	695.7	145.4	204	706	675.6	143.7	204	881	911.8	59.5	191
B19 (I-Dorm2)	32	30.5	18.9	159	20	20.0	16.8	159	0	5.0	18.5	190
B25 (I-Dorm)	44	45.3	20.9	193	27	29.9	16.5	193	16	21.7	17.1	60

Notes: Data from the fall 2020 and spring 2021 semesters (excludes data from the 2020/21 winter break). Building code B15 represents data from a cluster of special-purpose on-campus housing with a shared wastewater outflow sampling site/manhole. Due to low occupancy in B17 relative to the other building-sampling sites, wastewater sampling for B17 was discontinued after the fall 2020 semester. I-Dorm = isolation dormitory.

Population-Adjusted SARS-CoV-2 Data

As shown in Table 3, among the subset of sampling sites analyzed using RT-ddPCR, mean SARS-CoV-2 viral copies were highest for both *N* and *E* genes in samples collected from dormitory code B13 (17,220 and 3,934 copies/100mL for *N* and *E*, respectively) as well as the primary isolation dormitory (1,846 and 2,312 copies/100mL for *N* and *E*, respectively, at site B25). After adjusting copy numbers based on the estimated populations contributing to wastewater outflows from these buildings, mean viral copies/100mL/100people were still highest for dormitory B13 for *N* (3,901 copies/100mL/100people), but not for *E* (900 copies/100mL/100people), followed by the population-adjusted means for the isolation dormitory, which were higher than the unadjusted copy numbers (2,733 and 3,828 copies/100mL/100people for *N* and *E*, respectively). Similar trends were observed for population-adjusted means using the weekly housing count data.

One would expect mean SARS-CoV-2 viral copy numbers to be higher on a per capita basis for samples collected from COVID-19 isolation dormitories given the higher rates of infection and viral shedding. Indeed, when comparing unadjusted and population adjusted viral copy means across all sites (Table 3), aside from a slight increase in the population-adjusted mean for B17 (33%, from 110 to 146, for *N*; and 26%, 68 to 86, for *E*; though samples were only collected for B17 during the fall semester), population-adjusted means were lower than the unadjusted means for all non-isolation dormitories. Conversely, for the primary isolation dormitory, population-adjusted viral copy means were 48% and 66% higher than unadjusted viral copy means for *N* and *E* (from 1,846 to 2,733, and from 2,312 to 3,828, respectively), and 116% and 131% higher for the secondary isolation dormitory for the same comparison of viral copy means for *N* and *E* (from 255 to 550, and from 746 to 1,724, respectively). These same trends were observed for viral copy means when adjusted (data allowing) using the weekly housing data.

Table 3: Unadjusted & Population-adjusted RT-ddPCR Viral Copies/100mL for N & E Genes by Building

	Building Code	Viral Copies/100mL				Viral Copies/100mL per 100 People: Swipe Card Data (S&O)				Viral Copies/100mL per 100 People: Housing Data (S)			
		Median	Mean	SD	n	Median	Mean	SD	n	Median	Mean	SD	n
N	B8	0	331	1,061	49	0	83	276	49	0	88	277	46
	B13	34	17,220	11,9018	49	7	3,901	26,927	49	9	4,431	29,670	46
	B15	73	336	628	51	25	80	142	47	20	67	118	47
	B17	0	110	181	19	0	146	260	19	0	131	219	19
	B22	101	446	1,208	48	14	58	153	48	12	51	132	45
	B19 (I-Dorm2)	0	255	777	17	0	550	1,708	16	.	.	.	0
	B25 (I-Dorm)	165	1,846	6,389	47	357	2,733	7,455	46	571	4,466	10,345	16
	Outflow	149	232	269	49	.	.	.	0	.	.	.	0
E	B8	0	603	2,142	49	0	154	558	49	0	163	559	46
	B13	0	3,934	24,852	49	0	900	5,623	49	0	1,013	6,195	46
	B15	0	496	856	51	0	118	191	47	0	99	161	47
	B17	0	68	145	19	0	86	176	19	0	81	171	19
	B22	0	1,076	3,705	48	0	139	466	48	0	122	401	45
	B19 (I-Dorm2)	0	746	2,905	17	0	1,724	6,369	16	.	.	.	0
	B25 (I-Dorm)	0	2,312	5,698	47	173	3,828	8,937	46	0	5,566	14,154	16
	Outflow	0	260	406	49	.	.	.	0	.	.	.	0

Notes: Data from the fall 2020 and spring 2021 semesters. I-Dorm = COVID-19 isolation dormitory.

As shown in Figure S7, the overall distribution of SARS-CoV-2 viral copy numbers for *N* and *E* for the five dormitories generally aligns with corresponding COVID-19 case data (based on isolation assignment by building) for the time period assessed, with a more pronounced overlap observed for the spring 2021 semester. Case data overlaid with population-adjusted (based on swipe card data) viral copy numbers are shown in Figure S8. For the RT-qPCR results, less-pronounced but somewhat similar trends were observed for the population-adjusted viral copy concentrations for *N*, *E*, and *S* genes, as shown in Figure S9.

Looking at results from the campus wastewater outflow samples, as shown Figure 3, trend lines for SARS-CoV-2 viral copy numbers for *N* and *E* follow a pattern similar to that of COVID-19 case counts based on isolation data for the five dormitories analyzed in this study, as well as for diagnosed cases university-wide – particularly so for the spring 2021 semester. Population-adjusted viral copies from wastewater samples collected from the university’s main COVID-19 isolation dormitory also aligned well with these COVID-19 case trends (Figure S10).

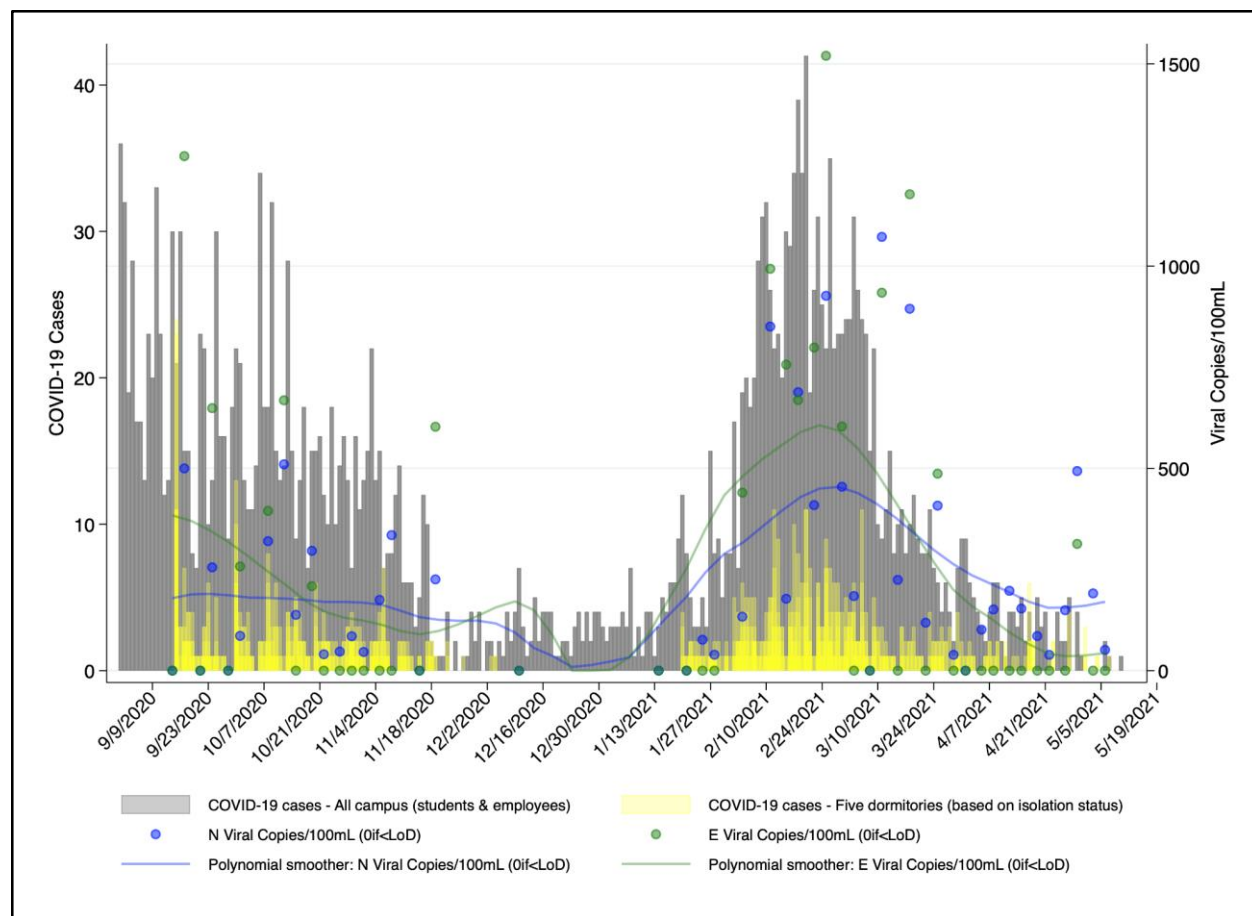


Figure 3: Viral copy concentrations and trends lines for *N* & *E* genes measured by RT-ddPCR for the university wastewater outflow site and COVID-19 cases university-wide (gray) and for students in the five dormitories analyzed in this study (yellow).

Note: *N* and *E* trend lines were fit using polynomial smoothing.

Statistical Modeling & Wastewater-Based Epidemiology

For our initial models, we used an eight-day lead for building-specific COVID-19 cases (the outcome variable, derived from isolation assignment status). As shown in Table 4, using random-effects Poisson regression with cluster-robust SEs, we observed positive (i.e., increases in the likelihood) and statistically significant (at a $p=0.05$ threshold) associations with COVID-19 cases eight days following sample collection for log10 transformed viral copies/100mL for *N* and *E* genes. IRRs from analysis of log10 transformed population-adjusted viral copy concentrations (based on swipe card and weekly housing data) for *N* and *E* were also positive and statistically significant. In particular, one-log increases in swipe-normalized viral copies for *N* and *E* were associated with statistically significant increases of 21% and 27% in the likelihood of observing COVID-19 cases eight days after sample collection (IRR=1.206, $p<0.001$, $n=211$ for *N*; and IRR=1.265, $p<0.001$, $n=211$ for *E*). We also analyzed these associations using binary variables for *N* and *E* (not detected =0, detected =1), and we observed associations with COVID-19 cases that were also clinically and statistically significant. Specifically, detection of *N* genes in wastewater samples were associated with an 85% increase in the likelihood of COVID-19 cases eight days later (IRR=1.845, $p=0.013$, $n=215$), and detection of *E* genes in wastewater samples were associated with a 99% increase in the likelihood of COVID-19 cases eight days later (IRR=1.994, $p=0.007$, $n=215$).

Table 4: Random-effects Poisson regression with cluster-robust standard errors using COVID-19 cases eight days after wastewater sampling as the outcome variable

Model: Predictor Variable	IRR	SE	p-value	n	Clusters (buildings)
<i>N</i> - Log10 Viral Copies/100mL (ND=Zero)	1.199	0.096	0.024	215	5
<i>E</i> - Log10 Viral Copies/100mL (ND=Zero)	1.220	0.101	0.017	215	5
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.206	0.061	<0.001	211	5
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.265	0.082	<0.001	211	5
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.168	0.060	0.002	202	5
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.233	0.087	0.003	202	5
<i>N</i> - Viral Copies as Binary (ND vs Detected)	1.845	0.454	0.013	215	5
<i>E</i> - Viral Copies as Binary (ND vs Detected)	1.994	0.508	0.007	215	5

Notes: Analysis using RT-ddPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. Log10 = log-10 transformed (ND & <LoD = zero). ND = Not detected. Ppl = people.

Because the method of sample collection (composite vs. grab) could be expected to impact our overall findings, we also ran each of the models presented in Table 4 with a binary covariate to control for the potential impacts of sampling method (composite =0, grab =1). As shown in Table S4, after controlling for the type of sampling method, the direction and significance (at a $p=0.05$ threshold) of IRRs remained unchanged across all eight models, and the size of IRRs, SEs, and associated p-values remained very similar, with unadjusted (Table 4) and adjusted (Table S4) IRRs differing by a range of only 0.16% to 3.04% (mean=1.20%, SD=1.02%).

Results for models based on RT-qPCR data from the same subset of wastewater samples are provided in Table S5. Across these models, all IRRs were less than one, indicating negative associations between log10 transformed viral copies for *N*, *E*, and *S* in isolation, or when combined (population-adjusted and unadjusted), and COVID-19 cases eight days following sample collection. Across models with the RT-qPCR data, only the IRR for log10 transformed viral copies for *N* was statistically significant (IRR=0.146, $p<0.001$, $n=212$; and incidentally this was also the lowest IRR observed across models in Table S5). When the association was analyzed using a binary predictor (“Any NES Ct Mean as Binary”), the IRR was close to the null and not significant (IRR=0.996, $p=0.992$, $n=213$).

To evaluate associations between SARS-CoV-2 in wastewater samples and COVID-19 cases seven, eight, and nine days following wastewater sampling, we used the same random-effects Poisson regression models and variables in Table 4, but for the outcome variable we used 3-day bins of cases centered on eight days following wastewater sampling (8 ± 1 day). As shown in Table S6, resulting IRRs were positive for all models using log-10 transformed viral copies as predictors, but only statistically significant for population-adjusted (based on swipe card data) viral copies for the *E* gene (IRR=1.158, $p=0.038$, $n=212$). When using a binary predictor for *N* and *E* (not detected =0, detected =1), IRRs for the 3-day bin of cases were both positive, suggesting a 24% and 55% increase in the likelihood of COVID-19 cases for *N* and *E* detection, respectively; but the association was only statistically significant for *E* (IRR=1.241 $p=0.270$, $n=216$ for *N*; and IRR=1.549, $p=0.012$, $n=216$ for *E*).

Table 5: Random-effects Poisson regression with cluster-robust standard errors where the outcome variable is the number of COVID-19 cases (based on building-specific student assignment to COVID-19 isolation status) four to nine days after wastewater sampling, using RT-ddPCR results as the predictor

Model: Predictor Variable	IRR	SE	P-value	n	Clusters (buildings)
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 4days	0.905	0.437	0.836	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 4days	1.594	0.904	0.411	216	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 5days	2.136	0.246	<0.001	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 5days	2.313	0.517	<0.001	216	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 6days	1.739	0.344	0.005	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 6days	1.901	0.551	0.027	216	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 7days	0.818	0.562	0.770	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 7days	1.083	0.781	0.912	216	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 8days	1.845	0.454	0.013	215	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 8days	1.994	0.508	0.007	215	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 9days	1.112	0.219	0.587	215	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 9days	1.527	0.423	0.127	215	5

Notes: Analysis using RT-ddPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. ND = Not detected.

Analysis of binary variables for *N* and *E* and COVID-19 cases four to nine days (eight days also already presented in Table 4) following wastewater sample collection are provided in Table 5. IRRs were statistically significant (at a $p=0.05$ threshold) using five-, six-, and eight-day leads, with the largest observed IRRs at five-days (IRR=2.136, $p<0.001$, $n=216$ and IRR=2.313, $p<0.001$, $n=216$, for *N* and *E*, respectively) and eight-days (IRR=1.845, $p=0.013$, $n=215$ and IRR=1.994, $p=0.007$, $n=215$, for *N* and *E*, respectively) after wastewater sampling. Running the same swipe card based model shown in Table 4 but using a five day rather than eight day lead, we observed that one-log increases in swipe-normalized viral copies for *N* and *E* were associated with statistically significant increases of 40% and 43% in the likelihood of observing COVID-19 cases five days after sample collection (IRR=1.403, SE=0.157, $p=0.002$, $n=212$ for *N*; and IRR=1.426, SE=0.112, $p<0.001$, $n=212$ for *E*).

Looking at the results around our pre-specified 8-day lead, the size and direction of the IRRs at seven and nine day leads suggested that, from a prediction-based point of view, the eight-day lead yielded larger and more significant effect sizes. As a sensitivity analysis, we ran the seven-, eight-, and nine-day models using data on student assignment to isolation as well as assignment to quarantine status as our COVID-19 case outcome variable. As shown in Table S7, with this broader proxy of COVID-cases as the outcome variable, the IRRs increased slightly when using seven- and eight-day leads, but decreased slightly for *N* when using a nine-day lead (compared to results in Table 5), and the associations were again only significant at an eight-day lead. Finally, we analyzed data from the RT-qPCR results using binary variables for detection of *N*, *E*, and/or *S* Ct values and COVID-19 cases as the outcome (using only isolation status) 8+/-2 days following wastewater sampling (Table S8). Based on these analyses, the IRR was essentially null and not significant at an eight-day lead (IRR=0.996, $p=0.992$, $n=213$), but was positive and statistically significant for the nine-day lead (IRR=1.868, $p=0.023$, $n=213$).

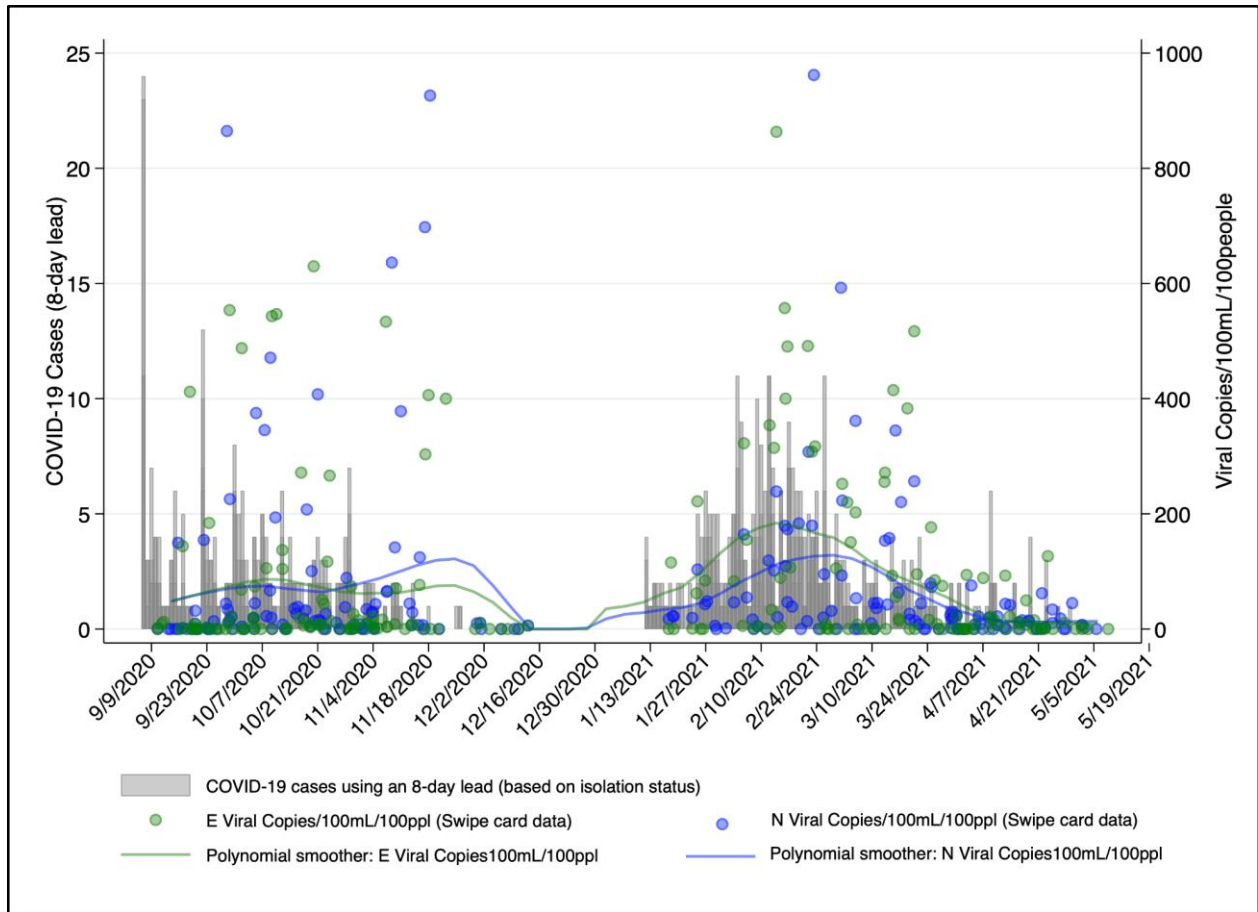


Figure 4: Population-adjusted viral copy concentrations and trend lines for N & E genes measured by RT-ddPCR and COVID-19 cases with an eight-day lead for five dormitories
Notes: Three *N* observations and five *E* observations were removed for improved visualization. Jitter = 3. Ppl = people. *N* and *E* trend lines were fit using polynomial smoothing. Population adjustment based on swipe card data.

Based on these results (Tables 4 & 5), in Figure 4 population-adjusted viral copy concentrations for *N* and *E* (copies/100mL/100people based on swipe card data) are plotted over COVID-19 cases using an eight-day lead (i.e., building-linked SARS-CoV-2 RT-ddPCR results and associated assignment of students to isolation status), with polynomial smoothers to show trend lines for *N* and *E* over the study period. These results are shown disaggregated by dormitory in Figure 5. Figure S11 provides a comparison of population-adjusted viral copy concentration trends between RT-ddPCR and RT-qPCR measurements (using polynomial smoothers) overlaid with COVID-19 case data with an eight-day lead.

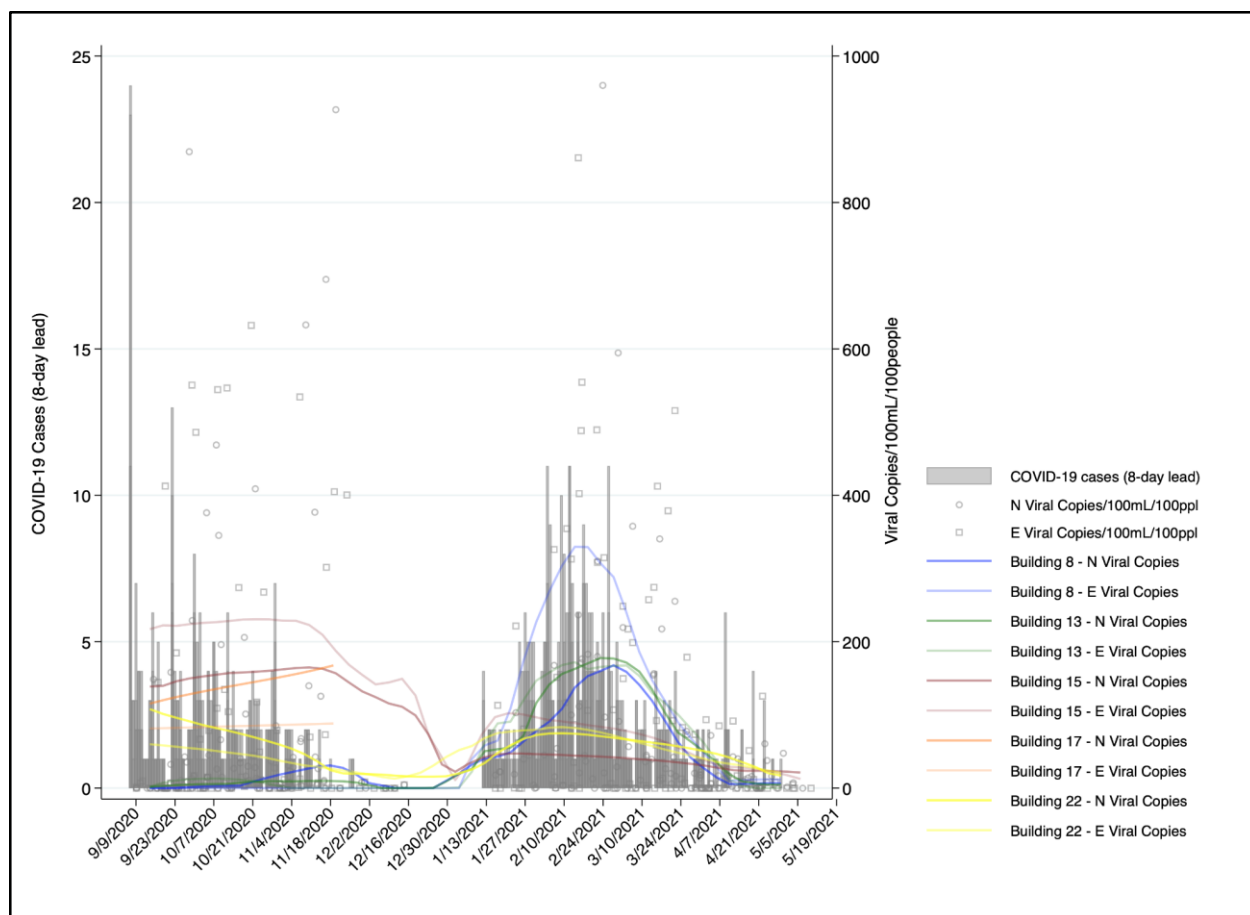


Figure 5: Population-adjusted viral copy concentrations and dormitory specific trend-lines for N & E genes measured by RT-ddPCR and COVID-19 cases with an eight-day lead for five dormitories

Notes: Three *N* observations and five *E* observations were removed for improved visualization. Jitter = 3. Ppl = people. Population adjustment based on swipe card data. *N* and *E* trend lines were fit using polynomial smoothing.

Study Limitations

The layout of many sewers precluded the use of composite samples (e.g., as for B19 and B25) or flow meters. Low flows at times in many sewers (late night and early morning hours in particular) contributed to battery depletion for some composite samplers (an issue compounded by low temperatures during the winter months); in such instances grab samples were used to supplement the volume collected. Due to relatively low occupancy and low wastewater flow in B17, sampling was discontinued after the fall 2020 semester. At times, heterogeneity of the wastewater composition and flow temporarily interfered with sampling equipment (e.g., when toilet paper caught onto sampling tubes/strainers). Logistical constraints (e.g., malfunctioning samplers, construction preventing access to manholes) also prevented consistent collection of samples from all sites on all sampling days.

Working with a clinical lab that was simultaneously processing large numbers of COVID-19 testing swabs for the university allowed for a rapid start to the project as well as the benefit of using the same assay for both clinical and wastewater samples. However, the clinical lab's RT-qPCR techniques were not optimized for wastewater samples, and due to their heavy workload

there were few opportunities to reanalyze samples or optimize methods for analysis of environmental samples. It was not feasible, for example, to analyze multiple dilutions or conduct additional tests to evaluate potential impacts of concentrated inhibitors on RT-qPCR amplification or false negative results (we addressed these issues through a focused re-analysis of a subset of samples using RT-ddPCR).

With regard to the COVID-19 isolation status data, for some students (<15%), their temporary room assignments changed one or more times throughout the duration of their assignment to isolation, which added uncertainty to some of our estimates of daily occupancy for the isolation dormitories, and by extension to estimates of population-adjusted viral copy numbers for the isolation dormitories. However, there were few such issues with data for the start date of isolation status assignments (the primary data point used for our estimation of COVID-19 cases by date and building). Finally, given the nature of the RT-qPCR results, and relatively limited number of observations with Ct values <45 for *N*, *E*, and *S*, we did not have sufficient data to adequately compare equal weightings for *N*, *E*, *S* with a weighting scheme that upweighted results for *N*.

Results in Context & Implications

According to information collated by researchers at UC Merced,²¹ as of January 2022 more than 250 universities worldwide were using, or had used, SARS-CoV-2 WBS. To help contextualize our use of data and methods for this initiative at VT, we conducted a non-systematic review of published peer-reviewed articles describing results from COVID-19 WBS initiatives at universities and colleges in the USA and Canada and extracted summary information on methods from six such publications^{34–39} (see Table S9). Across these universities, the number of wastewater sampling sites ranged from 1-68, sample collection frequencies ranged from 1-7 days/week, and three of the six universities used composite samplers, the others using grab samples. All six used RT-qPCR for SARS-CoV-2 detection, with some analyzing only *N* genes, and others analyzing *N* and *E*. The methods used for concentrating samples as well as the type of recovery controls varied across the studies, as did methods for assessing inhibition (when reported). Normalization was based primarily on the use of human fecal markers or flow/volume (for additional details see Table S9). We did not identify any published university-focused studies that used building-specific occupancy and COVID-19 case data at a daily resolution for WBS or WBE, as was the focus of our study, and relatively few that used statistical modeling for WBE.

Our study identified marked associations between the SARS-CoV-2 signals measured from campus outflow wastewater samples and university-wide COVID-19 case data, as well as similar trends for the five non-isolation dormitories. The clinically and statistically significant associations we observed between population-adjusted SARS-CoV-2 signals in wastewater samples and corresponding COVID-19 case data eight days following sample collection - as well as the higher population-adjusted viral concentrations observed in samples from the university's COVID-19 isolation dormitories - provide additional evidence and a concrete framework for the application of WBS to quantitatively predict COVID-19 trends at sub-sewershed scales, such as a university campus.

Compared with published methods and results from other university-scale COVID-19 WBS initiatives (as of January 2022), our study appeared to be relatively unique with respect to our use

of building-specific occupancy and COVID-19 case data to calculate population-adjusted SARS-CoV-2 results, as well as our use of statistical modeling to assess associations between SARS-CoV-2 signals from wastewater samples and subsequent COVID-19 infections. We expect that through improved modeling of the relationships between SARS-CoV-2 genes measured in sewage and associated COVID-19 infection data at sub-sewershed scales, this framework and approach can help to advance both WBS and WBE efforts at sewershed, sub-sewershed, and building-specific scales.

Supporting Information

Additional figures depicting: isolation and quarantine assignment data by day of week and semester; histograms and QQ-plots of RT-ddPCR data before and after log-10 transformation; Scatterplots of RT-ddPCR and RT-qPCR data for N and E genes; Aggregated occupancy data by week of the year and source; RT-ddPCR and RT-qPCR viral copy concentrations and cases with and without population-adjustment for five buildings and university-wide. Additional tables with data on: wastewater sampling methods by day of week and semester; LoD data for RT-ddPCR results; RT-ddPCR and RT-qPCR results comparison by building/site; statistical modelling results with RT-qPCR data controlling for sampling method, with RT-qPCR data, using 3-day bins of cases (8 ± 1 day), using isolation and quarantine outcomes, and for RT-qPCR data at 6-10 days leads; comparison of university-focused wastewater surveillance methods from selected publications (PDF)

Acknowledgements

This work and research was made possible thanks to the support and contributions of many people across Virginia Tech. In particular, the authors extend their thanks to: Michael Mulhare, Jennifer Averill, and Andrew Marinik (Emergency Management); Bobby Polly (Planning, Infrastructure & Facilities); Jean Luyendijk and Julie Wesel (Hokie Passport Services); Nickie Smith (Housing Services); David Higdon and Leanna House (Statistics); Rachel Silverman and Danielle Short (Population Health Sciences); Kimberly Ascue, Richard Clutter, and Jeremy Larose (Information Systems & IT); and Navid Ghaffarzadegan and other members of the VT COVID-19 Modeling Group.

Funding Sources

This study was primarily supported through internal funding from Virginia Tech. Additional funds were provided through National Science Foundation (NSF) OISE award (#1545756), ECCSS NNCI award (#1542100), CSSI award (#2004751), and an NRT award (#2125798). Funding was also provided by the Center for Science and Engineering of the Exposome at the Virginia Tech Institute for Critical Technology and Applied Science (ICTAS).

Data Availability Statement

The data underlying this study are not publicly available due to the potential to link COVID-19 case data to specific dates and dormitories, and to combine such information with online data sources to potentially re-identify personally identifiable information.

References

1. Metcalf, T. G.; Melnick, J. L.; Estes, M. K. Environmental Virology From Detection of Virus in Sewage and Water by Isolation to Identification by Molecular Biology—A Trip of Over 50 Years. *Annu. Rev. Microbiol.* 1995, 49 (1), 461–487. <https://doi.org/10.1146/annurev.mi.49.100195.002333>.
2. Daughton, C. G. Illicit Drugs in Municipal Sewage Proposed New Nonintrusive Tool to Heighten Public Awareness of Societal Use of Illicit-Abused Drugs and Their Potential for Ecological Consequences. *Pharm. Care Prod. Environ.* 2001, 791, 348–364. <https://doi.org/10.1021/bk-2001-0791.ch020>.
3. Zuccato, E.; Chiabrando, C.; Castiglioni, S.; Calamari, D.; Bagnati, R.; Schiarea, S.; Fanelli, R. Cocaine in Surface Waters: A New Evidence-Based Tool to Monitor Community Drug Abuse. *Environ. Health Glob. Access Sci. Source* 2005, 4, 14–14. <https://doi.org/10.1186/1476-069X-4-14>.
4. Lago, P. M.; Gary, H. E.; Pérez, L. S.; Cáceres, V.; Olivera, J. B.; Puentes, R. P.; Corredor, M. B.; Jiménez, P.; Pallansch, M. A.; Cruz, R. G. Poliovirus Detection in Wastewater and Stools Following an Immunization Campaign in Havana, Cuba. *Int. J. Epidemiol.* 2003, 32 (5), 772–777. <https://doi.org/10.1093/IJE/DYG185>.
5. Lodder, W.; de Roda Husman, A. M. SARS-CoV-2 in Wastewater: Potential Health Risk, but Also Data Source. *Lancet Gastroenterol. Hepatol.* 2020, 5 (6), 533–534. [https://doi.org/10.1016/S2468-1253\(20\)30087-X](https://doi.org/10.1016/S2468-1253(20)30087-X).
6. Zhang, Y.; Cen, M.; Hu, M.; Du, L.; Hu, W.; Kim, J. J.; Dai, N. Prevalence and Persistent Shedding of Fecal SARS-CoV-2 RNA in Patients With COVID-19 Infection: A Systematic Review and Meta-Analysis. *Clin. Transl. Gastroenterol.* 2021, 12 (4), e00343–e00343. <https://doi.org/10.14309/CTG.0000000000000343>.
7. Park, M.; Won, J.; Choi, B. Y.; Lee, C. J. Optimization of Primer Sets and Detection Protocols for SARS-CoV-2 of Coronavirus Disease 2019 (COVID-19) Using PCR and Real-Time PCR. *Exp. Mol. Med.* 2020 526 2020, 52 (6), 963–977. <https://doi.org/10.1038/s12276-020-0452-7>.
8. WHO. Diagnostic Testing for SARS-CoV-2: Interim Guidance, 11 September 2020. 2020.
9. CDC. CDC 2019–Novel Coronavirus (2019-NCoV) Real-Time RT-PCR Diagnostic Panel For Emergency Use Only Instructions for Use. 2020.
10. Won, J.; Lee, S.; Park, M.; Kim, T. Y.; Park, M. G.; Choi, B. Y.; Kim, D.; Chang, H.; Kim, V. N.; Lee, C. J. Development of a Laboratory-Safe and Low-Cost Detection Protocol for SARS-CoV-2 of the Coronavirus Disease 2019 (COVID-19). *Exp. Neurobiol.* 2020, 29 (2), 107–107. <https://doi.org/10.5607/EN20009>.
11. Alsuhaibani, M. A.; Kobayashi, T.; Trannel, A.; Holley, S.; Abosi, O. J.; Jenn, K. E.; Meacham, H.; Sheeler, L.; Etienne, W.; Dains, A.; Kukla, M. E.; Ward, E.; Ford, B.; Edmond, M. B.; Wellington, M.; Diekema, D. J.; Salinas, J. L. Coronavirus Disease 2019 (COVID-19) Admission Screening and Assessment of Infectiousness at an Academic Medical Center in Iowa, 2020. *Infect. Control Hosp. Epidemiol.* 2021, 1–5. <https://doi.org/10.1017/ice.2021.294>.
12. Medema, G.; Heijnen, L.; Elsinga, G.; Italiaander, R.; Brouwer, A. Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the

- Early Stage of the Epidemic in The Netherlands. *Environ. Sci. Technol. Lett.* 2020. <https://doi.org/10.1021/acs.estlett.0c00357>.
13. Jahn, K.; Dreifuss, D.; Topolsky, I.; Kull, A.; Ganesanandamoorthy, P.; Fernandez-Cassi, X.; Bänziger, C.; Stachler, E.; Fuhrmann, L.; Philipp Jablonski, K.; Chen, C.; Aquino, C.; Stadler, T.; Ort, C.; Kohn, T.; Julian, T. R.; Beerenwinkel, N. Detection of SARS-CoV-2 Variants in Switzerland by Genomic Analysis of Wastewater Samples. *medRxiv* 2021. <https://doi.org/10.1101/2021.01.08.21249379>.
 14. Rosa, G. L.; Iaconelli, M.; Mancini, P.; Ferraro, G. B.; Veneri, C.; Bonadonna, L.; Lucentini, L.; Suffredini, E. First Detection of Sars-Cov-2 in Untreated Wastewaters in Italy. *Sci. Total Environ.* 2020, 736 (20), 139652. <https://doi.org/10.1016/j.scitotenv.2020.139652>.
 15. Sharif, S.; Ikram, A.; Khurshid, A.; Salman, M.; Mehmood, N.; Arshad, Y.; Ahmed, J.; Safdar, R. M.; Rehman, L.; Mujtaba, G.; Hussain, J.; Ali, J.; Angez, M.; Alam, M. M.; Akthar, R.; Malik, M. W.; Baig, M. Z. I.; Rana, M. S.; Usman, M.; Ali, M. Q.; Ahad, A.; Badar, N.; Umair, M.; Tamim, S.; Ashraf, A.; Tahir, F.; Ali, N. Detection of SARS-CoV-2 in Wastewater Using the Existing Environmental Surveillance Network: A Potential Supplementary System for Monitoring COVID-19 Transmission. *PLOS ONE* 2021, 16 (6), e0249568. <https://doi.org/10.1371/journal.pone.0249568>.
 16. Ahmed, W.; Tschärke, B.; Bertsch, P. M.; Bibby, K.; Bivins, A.; Choi, P.; Clarke, L.; Dwyer, J.; Edson, J.; Nguyen, T. M. H.; O'Brien, J. W.; Simpson, S. L.; Sherman, P.; Thomas, K. V.; Verhagen, R.; Zaugg, J.; Mueller, J. F. SARS-CoV-2 RNA Monitoring in Wastewater as a Potential Early Warning System for COVID-19 Transmission in the Community: A Temporal Case Study. *Sci. Total Environ.* 2021, 761, 144216–144216. <https://doi.org/10.1016/j.scitotenv.2020.144216>.
 17. Virginia Tech. Virginia Tech Fall 2020 COVID-19 Operational Plan; 2020.
 18. Virginia Tech. Virginia Tech Spring 2021 COVID-19 Operational Plan; 2021.
 19. Betancourt, W. Q.; Schmitz, B. W.; Innes, G. K.; Prasek, S. M.; Pogreba Brown, K. M.; Stark, E. R.; Foster, A. R.; Sprissler, R. S.; Harris, D. T.; Sherchan, S. P.; Gerba, C. P.; Pepper, I. L. COVID-19 Containment on a College Campus via Wastewater-Based Epidemiology, Targeted Clinical Testing and an Intervention. *Sci. Total Environ.* 2021, 779, 146408. <https://doi.org/10.1016/j.scitotenv.2021.146408>.
 20. Harris-Lovett, S.; Nelson, K. L.; Beamer, P.; Bischel, H. N.; Bivins, A.; Bruder, A.; Butler, C.; Camenisch, T. D.; Long, S. K. D.; Karthikeyan, S.; Larsen, D. A.; Meierdiercks, K.; Mouser, P. J.; Pagsuyoin, S.; Prasek, S. M.; Radniecki, T. S.; Ram, J. L.; Roper, D. K.; Safford, H.; Sherchan, S. P.; Shuster, W.; Stalder, T.; Wheeler, R. T.; Korfmacher, K. S. Wastewater Surveillance for SARS-CoV-2 on College Campuses: Initial Efforts, Lessons Learned, and Research Needs. *Int. J. Environ. Res. Public Health* 2021 Vol 18 Page 4455 2021, 18 (9), 4455–4455. <https://doi.org/10.3390/IJERPH18094455>.
 21. UC Merced. COVIDPoops19 Dashboard <https://www.arcgis.com/apps/dashboards/c778145ea5bb4daeb58d31afee389082> (accessed 2022 -01 -25).
 22. Sims, N.; Kasprzyk-Hordern, B. Future Perspectives of Wastewater-Based Epidemiology: Monitoring Infectious Disease Spread and Resistance to the Community Level. *Environ. Int.* 2020, 139.
 23. Kumblathan, T.; Liu, Y.; Uppal, G. K.; Hruday, S. E.; Li, X.-F. Wastewater-Based Epidemiology for Community Monitoring of SARS-CoV-2: Progress and Challenges. *ACS Environ. Au* 2021. <https://doi.org/10.1021/ACSENVIRONAU.1C00015>.

24. Hrudey, S. E.; Conant, B. The Devil Is in the Details: Emerging Insights on the Relevance of Wastewater Surveillance for SARS-CoV-2 to Public Health. *J. Water Health* 2021, 20 (1), 246–270. <https://doi.org/10.2166/wh.2021.186>.
25. Ciesielski, M.; Blackwood, D.; Clerkin, T.; Gonzalez, R.; Thompson, H.; Larson, A.; Noble, R. Assessing Sensitivity and Reproducibility of RT-DdPCR and RT-QPCR for the Quantification of SARS-CoV-2 in Wastewater. *J. Virol. Methods* 2021, 297, 114230. <https://doi.org/10.1016/j.jviromet.2021.114230>.
26. Hamouda, M.; Mustafa, F.; Maraqa, M.; Rizvi, T.; Aly Hassan, A. Wastewater Surveillance for SARS-CoV-2: Lessons Learnt from Recent Studies to Define Future Applications. *Sci. Total Environ.* 2021, 759. <https://doi.org/10.1016/j.scitotenv.2020.143493>.
27. Cohen, A.; Grubb, C. Virginia Tech SARS-CoV-2 Wastewater Surveillance & COVID-19 Epidemiology Initiative: Statistical Analysis Plan. Open Science Framework (<https://osf.io/q69ey/>), 2021.
28. Ahmed, W.; Bertsch, P. M.; Bivins, A.; Bibby, K.; Farkas, K.; Gathercole, A.; Haramoto, E.; Gyawali, P.; Korajkic, A.; McMinn, B. R.; Mueller, J. F.; Simpson, S. L.; Smith, W. J. M.; Symonds, E. M.; Thomas, K. V.; Verhagen, R.; Kitajima, M. Comparison of Virus Concentration Methods for the RT-QPCR-Based Recovery of Murine Hepatitis Virus, a Surrogate for SARS-CoV-2 from Untreated Wastewater. *Sci. Total Environ.* 2020, 739, 139960. <https://doi.org/10.1016/j.scitotenv.2020.139960>.
29. Ceci, A.; Muñoz-Ballester, C.; Tegge, A. N.; Brown, K. L.; Umans, R. A.; Michel, F. M.; Patel, D.; Tewari, B.; Martin, J.; Alcoreza, O.; Maynard, T.; Martinez-Martinez, D.; Bordwine, P.; Bissell, N.; Friedlander, M. J.; Sontheimer, H.; Finkielstein, C. V. Development and Implementation of a Scalable and Versatile Test for COVID-19 Diagnostics in Rural Communities. *Nat. Commun.* 2021, 12 (1), 4400–4400. <https://doi.org/10.1038/s41467-021-24552-4>.
30. Gonzalez, R.; Larson, A.; Thompson, H.; Carter, E.; Cassi, X. F. Redesigning SARS-CoV-2 Clinical RT-QPCR Assays for Wastewater RT-DdPCR; medRxiv, 2021; p 2021.03.02.21252754.
31. Bibby, K.; Bivins, A.; Wu, Z.; North, D. Making Waves: Plausible Lead Time for Wastewater Based Epidemiology as an Early Warning System for COVID-19. *Water Res.* 2021, 202, 117438. <https://doi.org/10.1016/j.watres.2021.117438>.
32. Olesen, S. W.; Imakaev, M.; Duvallet, C. Making Waves: Defining the Lead Time of Wastewater-Based Epidemiology for COVID-19. *Water Res.* 2021, 202, 117433. <https://doi.org/10.1016/j.watres.2021.117433>.
33. Galani, A.; Aalizadeh, R.; Kostakis, M.; Markou, A.; Alygizakis, N.; Lytras, T.; Adamopoulos, P. G.; Peccia, J.; Thompson, D. C.; Kontou, A.; Karagiannidis, A.; Lianidou, E. S.; Avgeris, M.; Paraskevis, D.; Tsiodras, S.; Scorilas, A.; Vasiliou, V.; Dimopoulos, M.-A.; Thomaidis, N. S. SARS-CoV-2 Wastewater Surveillance Data Can Predict Hospitalizations and ICU Admissions. *Sci. Total Environ.* 2022, 804, 150151. <https://doi.org/10.1016/j.scitotenv.2021.150151>.
34. Corchis-Scott, R.; Geng, Q.; Seth, R.; Ray, R.; Beg, M.; Biswas, N.; Charron, L.; Drouillard, K. D.; D'Souza, R.; Heath, D. D.; Houser, C.; Lawal, F.; McGinlay, J.; Menard, S. L.; Porter, L. A.; Rawlings, D.; Scholl, M. L.; Siu, K. W. M.; Tong, Y.; Weisener, C. G.; Wilhelm, S. W.; McKay, R. M. L. Averting an Outbreak of SARS-CoV-2 in a University Residence Hall through Wastewater Surveillance. *Microbiol. Spectr.* 2021, 9 (2), e00792-21. <https://doi.org/10.1128/Spectrum.00792-21>.

35. Juel, M. A. I.; Stark, N.; Nicolosi, B.; Lontai, J.; Lambirth, K.; Schlueter, J.; Gibas, C.; Munir, M. Performance Evaluation of Virus Concentration Methods for Implementing SARS-CoV-2 Wastewater Based Epidemiology Emphasizing Quick Data Turnaround. *Sci. Total Environ.* 2021, 801, 149656. <https://doi.org/10.1016/j.scitotenv.2021.149656>.
36. Karthikeyan, S.; Nguyen, A.; McDonald, D.; Zong, Y.; Ronquillo, N.; Ren, J.; Zou, J.; Farmer, S.; Humphrey, G.; Henderson, D.; Javidi, T.; Messer, K.; Anderson, C.; Schooley, R.; Martin, N. K.; Knight, R. Rapid, Large-Scale Wastewater Surveillance and Automated Reporting System Enable Early Detection of Nearly 85% of COVID-19 Cases on a University Campus. *mSystems* 2021, 6 (4), e00793-21. <https://doi.org/10.1128/mSystems.00793-21>.
37. Reeves, K.; Liebig, J.; Feula, A.; Saldi, T.; Lasda, E.; Johnson, W.; Lilienfeld, J.; Maggi, J.; Pulley, K.; Wilkerson, P. J.; Real, B.; Zak, G.; Davis, J.; Fink, M.; Gonzales, P.; Hager, C.; Ozeroff, C.; Tat, K.; Alkire, M.; Butler, C.; Coe, E.; Darby, J.; Freeman, N.; Heuer, H.; Jones, J. R.; Karr, M.; Key, S.; Maxwell, K.; Nelson, L.; Saldana, E.; Shea, R.; Salveson, L.; Tomlinson, K.; Vargas-Barriga, J.; Vigil, B.; Brisson, G.; Parker, R.; Leinwand, L. A.; Bjorkman, K.; Mansfeldt, C. High-Resolution within-Sewer SARS-CoV-2 Surveillance Facilitates Informed Intervention. *Water Res.* 2021, 204, 117613. <https://doi.org/10.1016/j.watres.2021.117613>
38. Schmitz, B. W.; Innes, G. K.; Prasek, S. M.; Betancourt, W. Q.; Stark, E. R.; Foster, A. R.; Abraham, A. G.; Gerba, C. P.; Pepper, I. L. Enumerating Asymptomatic COVID-19 Cases and Estimating SARS-CoV-2 Fecal Shedding Rates via Wastewater-Based Epidemiology. *Sci. Total Environ.* 2021, 801, 149794. <https://doi.org/10.1016/j.scitotenv.2021.149794>.
39. Scott, L. C.; Aubee, A.; Babahaji, L.; Vigil, K.; Tims, S.; Aw, T. G. Targeted Wastewater Surveillance of SARS-CoV-2 on a University Campus for COVID-19 Outbreak Detection and Mitigation. *Environ. Res.* 2021, 200, 111374. <https://doi.org/10.1016/j.envres.2021.111374>.

Supporting information for: Sub-sewershed SARS-CoV-2 Wastewater Surveillance & COVID-19 Epidemiology Using Building-specific Occupancy & Case Data

SI FIGURES

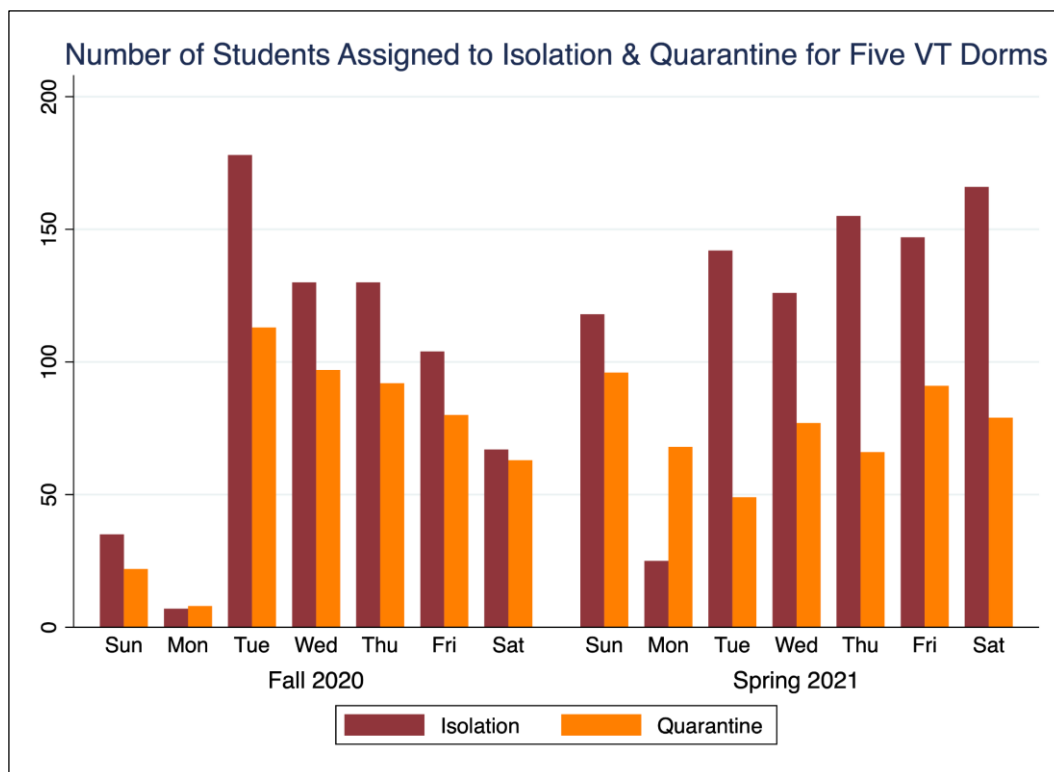
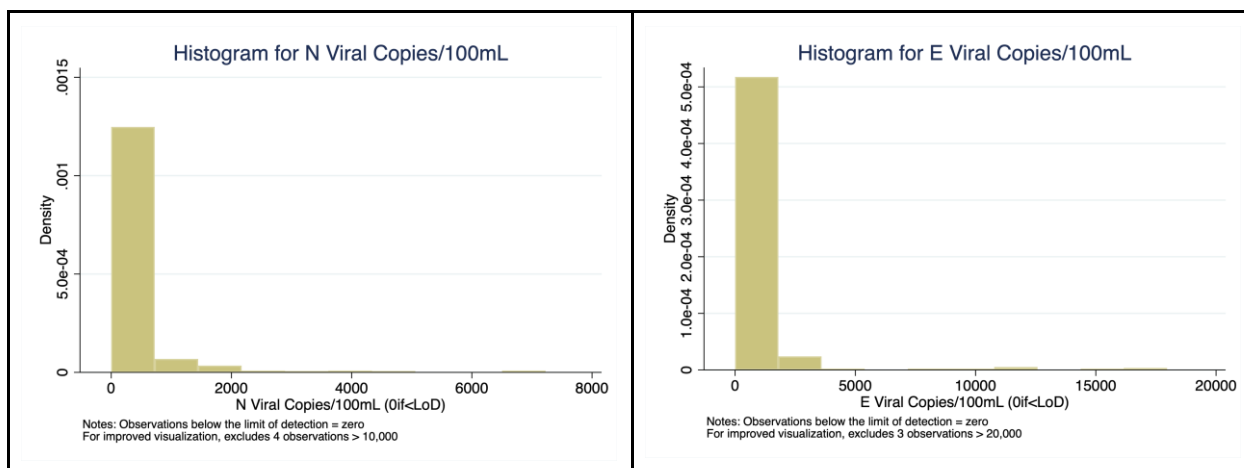


Figure S1: Student assignments to isolation and quarantine status for five VT dormitories during the fall 2020 and spring 2021 semesters.



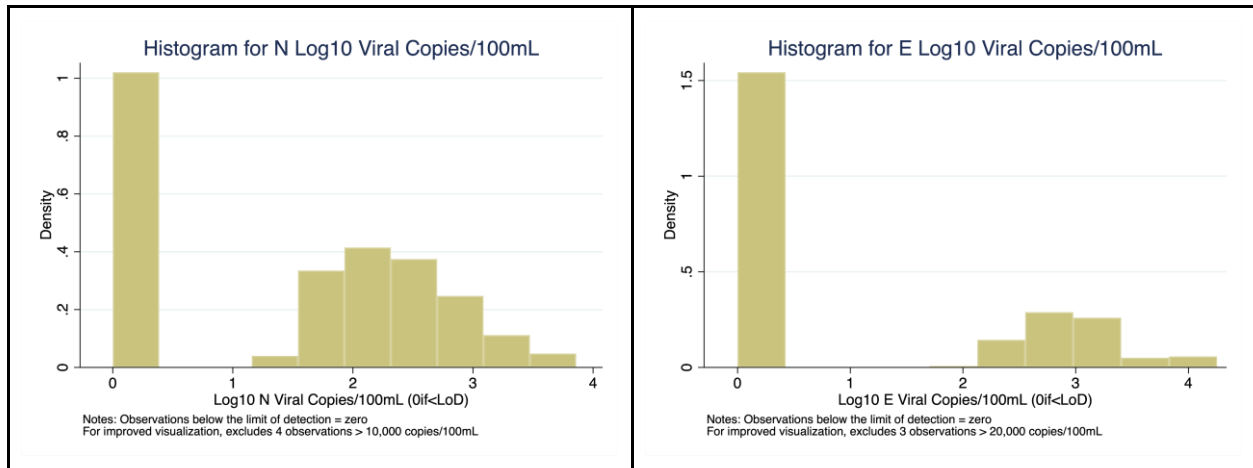
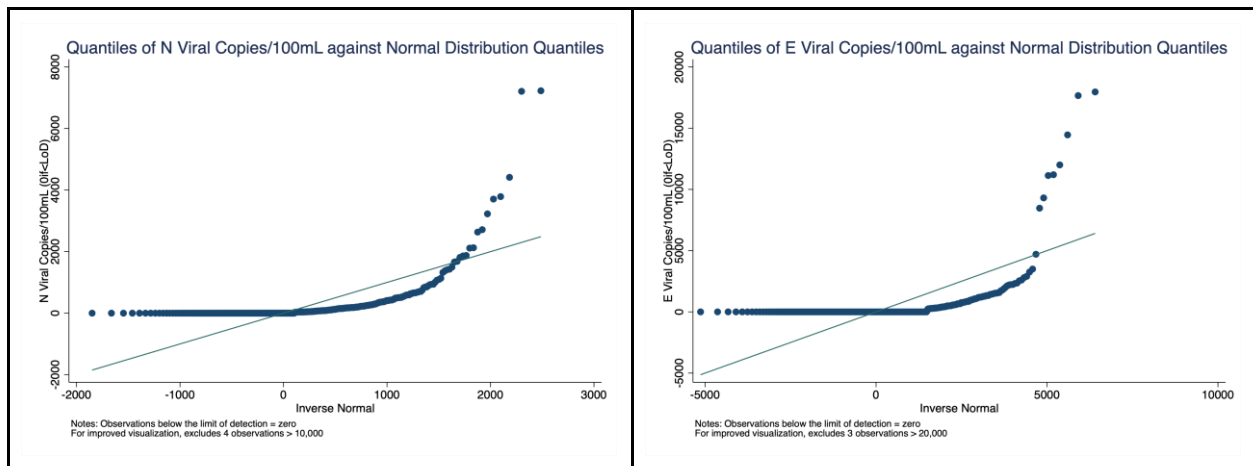


Figure S2: Histograms for RT-ddPCR data for *N* and *E* genes before and after log-10 transformation



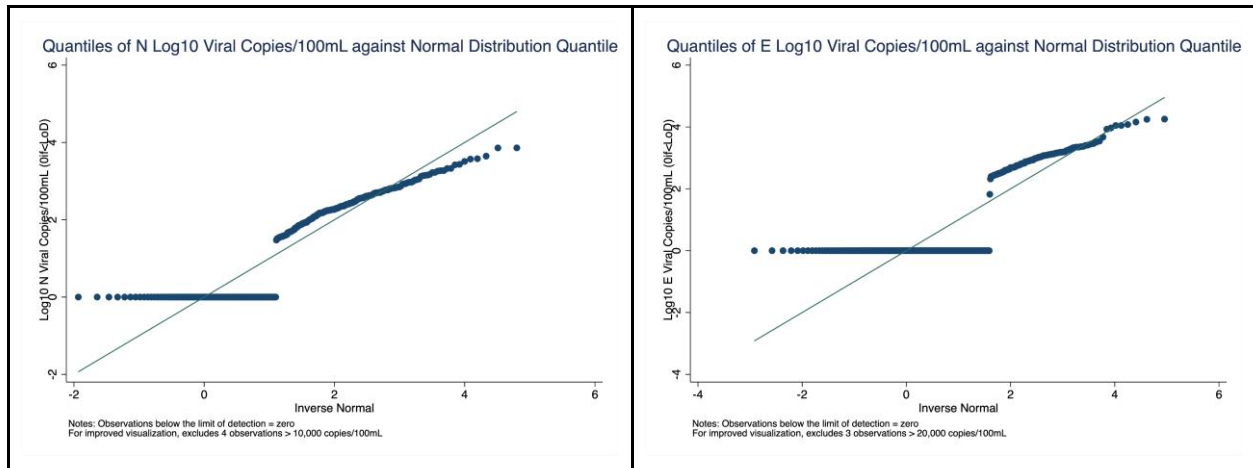


Figure S3: QQ-plots for RT-ddPCR data for *N* and *E* genes before and after log-10 transformation

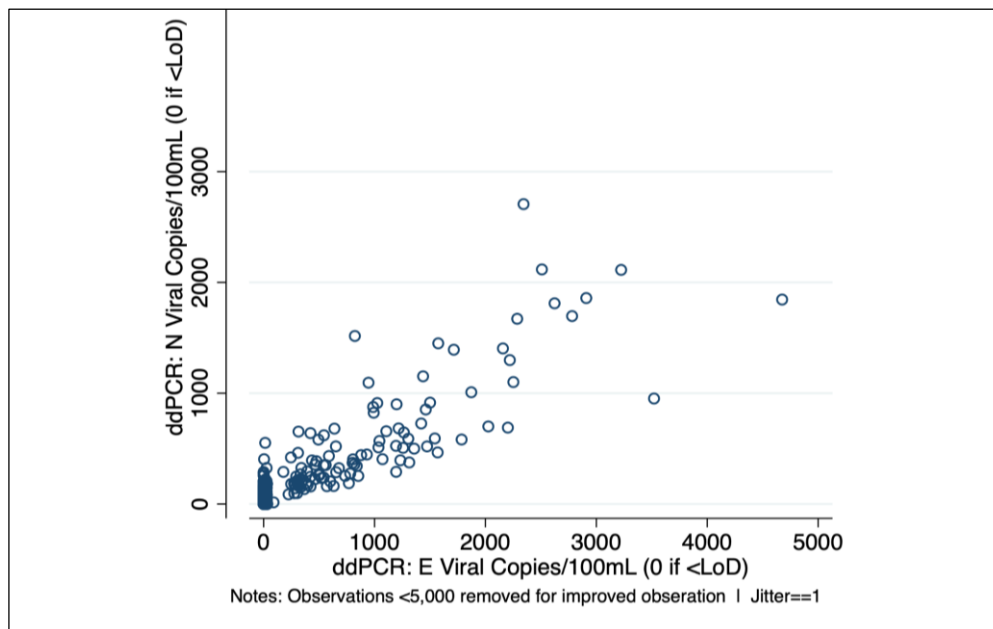


Figure S4: Scatterplot of RT-ddPCR viral copy data for *N* and *E* genes by sample

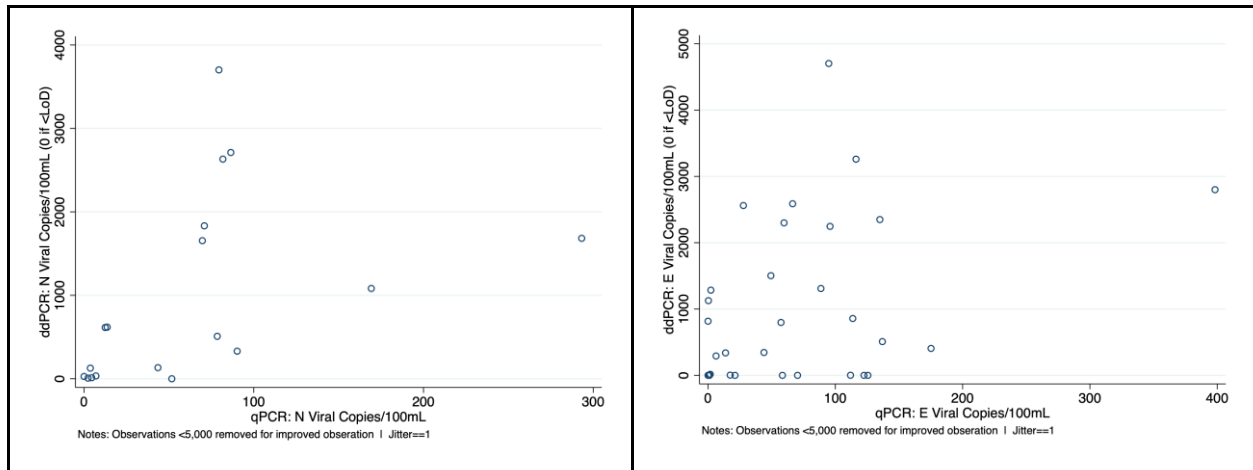


Figure S5: Scatterplots of RT-qPCR and RT-ddPCR viral copy data for *N* and *E* genes by sample

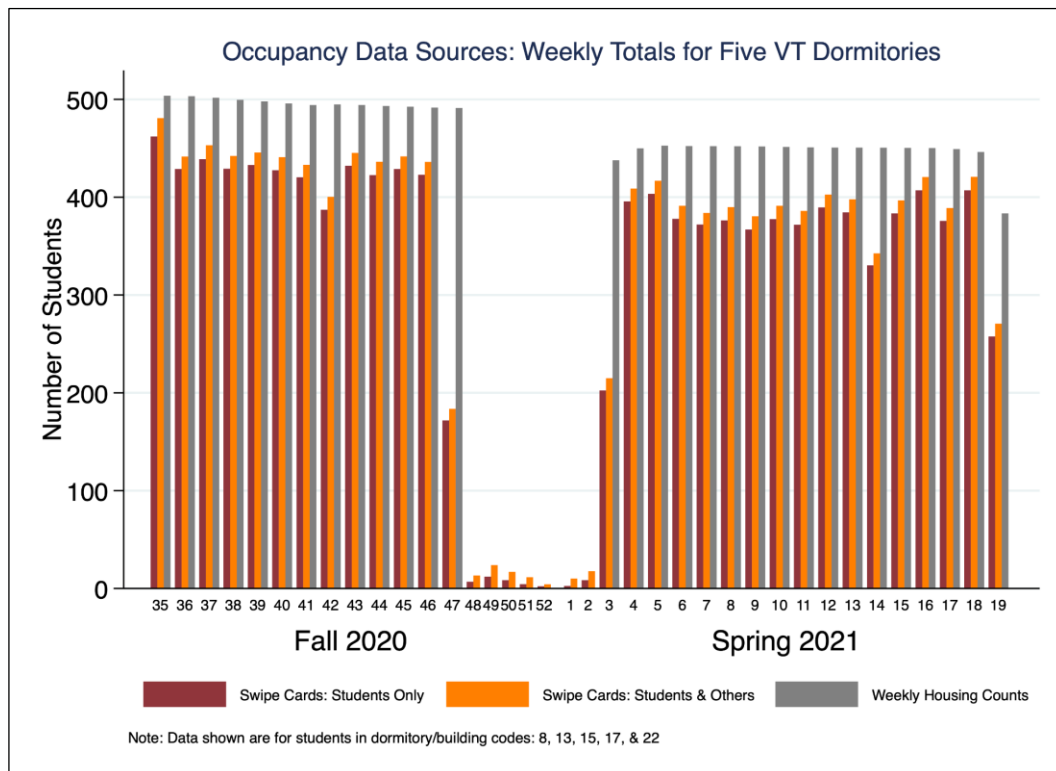


Figure S6: Occupancy data for five dormitories shown by calendar week and semester

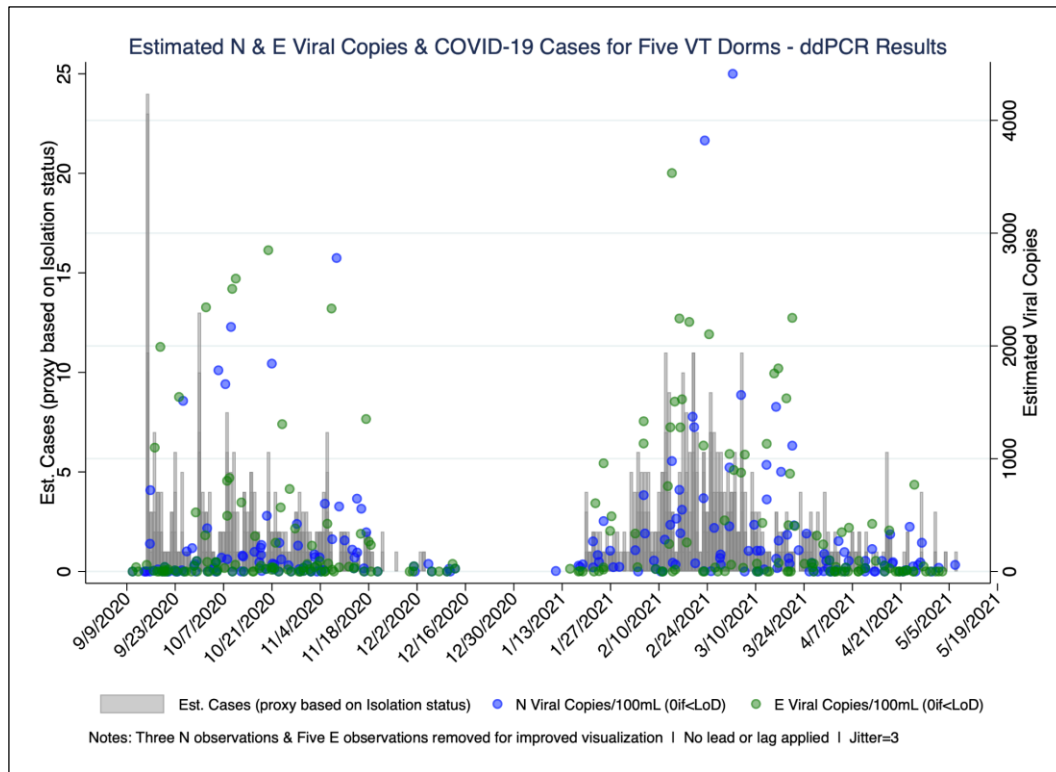


Figure S7: RT-ddPCR viral copy concentrations for *N* & *E* genes and COVID-19 cases for five dormitories

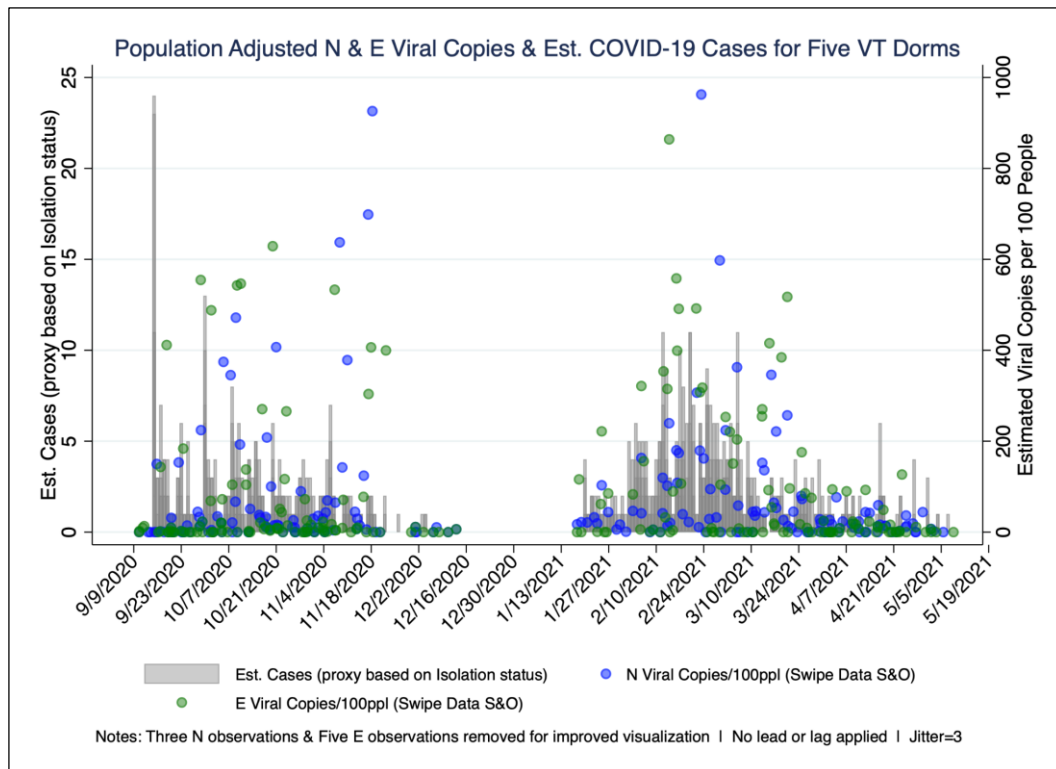


Figure S8: Population-adjusted RT-ddPCR viral copy concentrations for *N* & *E* genes and COVID-19 cases for five dormitories

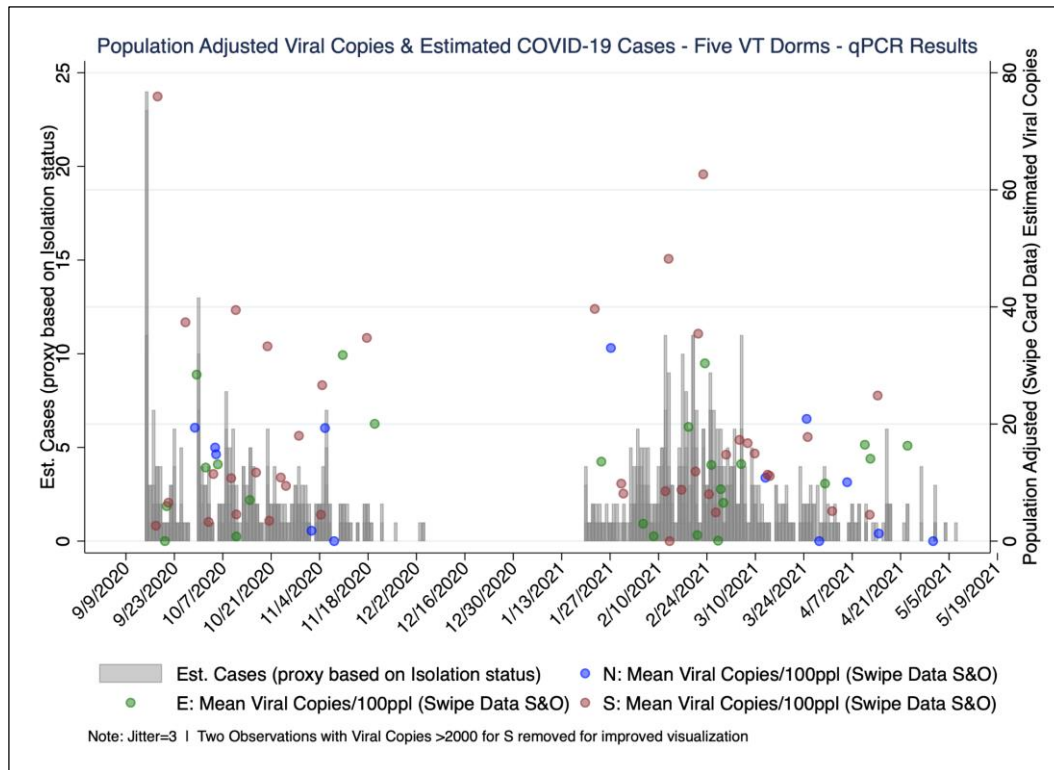


Figure S9: Population-adjusted RT-qPCR viral copy concentrations for *N*, *E*, & *S* genes and COVID-19 cases for five dormitories

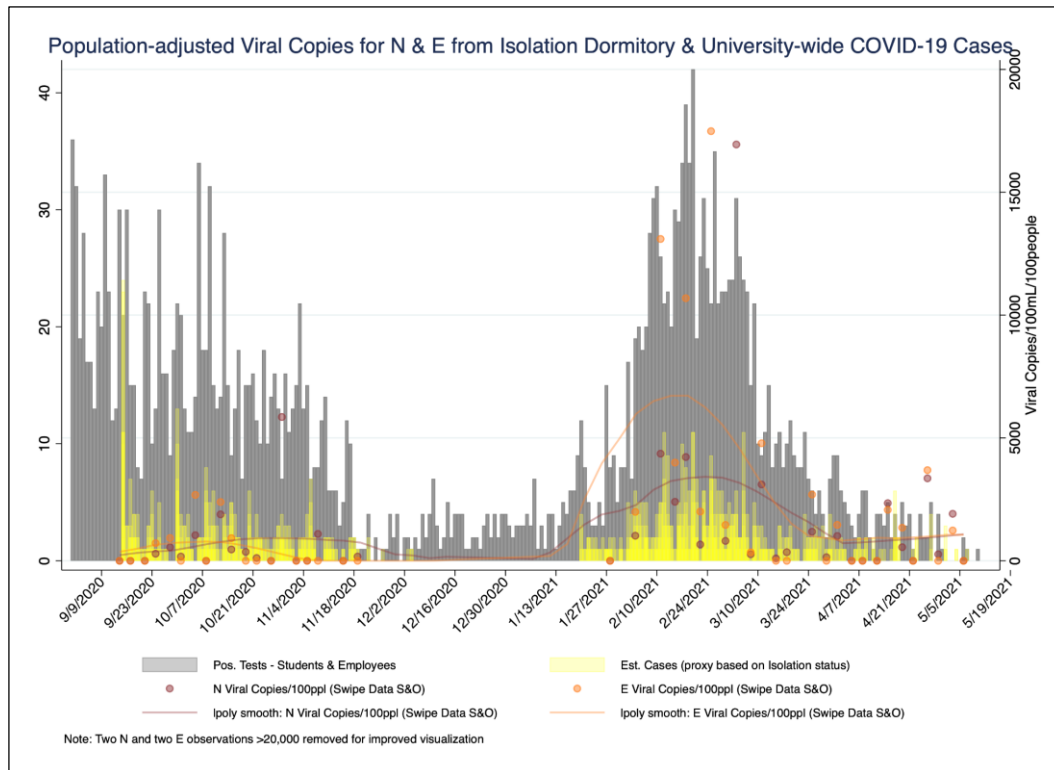


Figure S10: Viral copy concentrations and trends lines for *N* & *E* genes measured by RT-ddPCR for the university isolation dormitory (B25) and COVID-19 cases university-wide (gray) and for students in the five dormitories analyzed in this study (yellow).

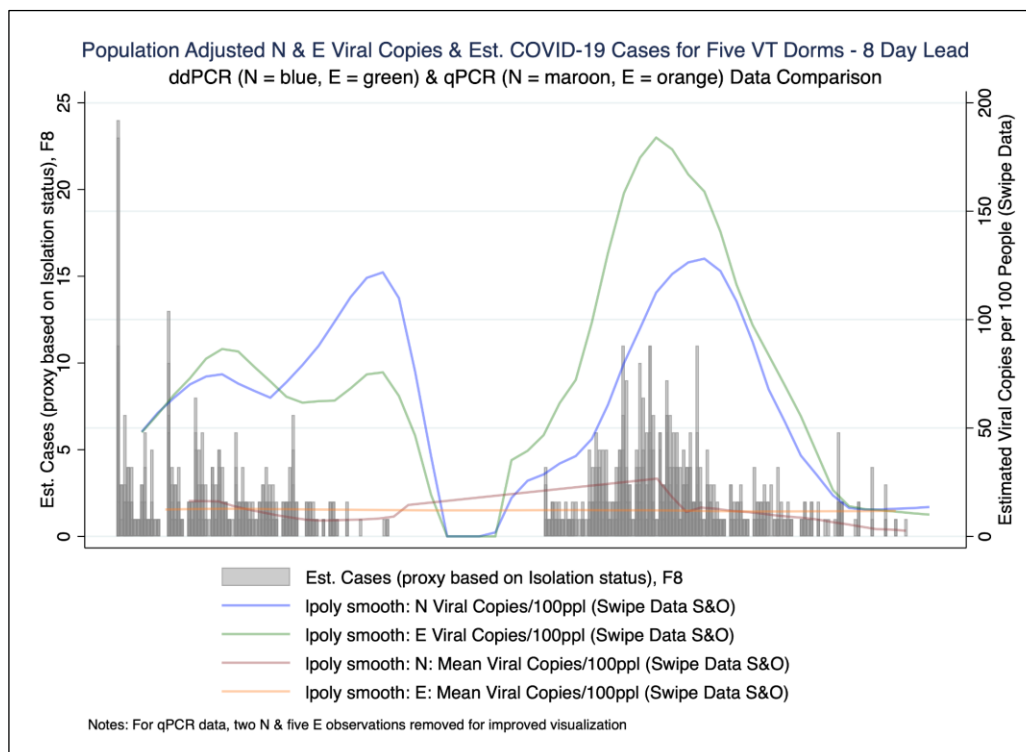


Figure 11: Trend lines for population-adjusted viral copy data for *N* and *E* by PCR method

SI TABLES

Table S1: Wastewater sample types and day of the week collected by building code and semester for selected sites

Building Code	Fall 2020 Semester					Spring 2021 Semester				
	Sample Type		Day of the Week Samples Collected			Sample Type		Day of the Week Samples Collected		
	Composite	Grab	Sun	Mon	Thu	Composite	Grab	Mon	Thu	
B8	13	9	3	7	12	22	6	13	15	
B13	14	8	3	7	12	24	4	13	15	
B15	16	6	3	7	12	19	12	14	17	
B17	8	12	2	8	10	0	0			
B19	0	15	1	7	7	0	3	1	2	
B22	12	10	2	8	12	7	22	14	15	
B25	0	20	2	8	10	0	30	14	16	
B99	15	6	3	6	12	20	12	14	18	
Total	78	86	19	58	87	92	89	83	98	

Table S2: RT-ddPCR Results: *N* & *E* Viral Copies/100mL for all Observations, LoD, & if >LoD

	<i>N</i> - Viral copies/100mL			<i>E</i> - Viral copies/100mL		
	Raw	LoD	if >LoD	Raw	LoD	if >LoD
Min	0	6	0	0	42	0
Median	71	35	71	93	259	0
Max	83,3333	108	83,3333	174,127	799	174,127
Mean	3,052	33	3,049	1,357	244	1,321
SD	45,986	9	45,986	9,983	64	9,988
n	329	329	329	329	329	329
Obs. >0 or >LOD	275	N/A	201	280	N/A	115

Table S3: SARs-CoV-2 Detects for Buildings Used for Primary Analyses: RT-ddPCR & RT-qPCR Results Comparison

Building Code		B8		B13		B15		B17		Overflow I-Dorm B19		B22		Primary I-Dorm B25		Campus Outflow B99		Total	
RT-PCR		n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
ddPCR: N	ND	26	53%	23	47%	19	37%	10	53%	11	65%	15	31%	15	32%	9	18%	128	39%
	Detected	23	47%	26	53%	32	63%	9	47%	6	35%	33	69%	32	68%	40	82%	201	61%
	Total	49	100%	49	100%	51	100%	19	100%	17	100%	48	100%	47	100%	49	100%	329	100%
qPCR: N	ND	46	98%	41	87%	47	90%	19	100%	16	100%	47	96%	42	88%	48	94%	306	93%
	Detected	1	2%	6	13%	5	10%	0	0%	0	0%	2	4%	6	12%	3	6%	23	7%
	Total	47	100%	47	100%	52	100%	19	100%	16	100%	49	100%	48	100%	51	100%	329	100%
ddPCR: E	ND	36	73%	36	73%	31	61%	15	79%	15	88%	26	54%	24	51%	31	63%	214	65%
	Detected	13	27%	13	27%	20	39%	4	21%	2	12%	22	46%	23	49%	18	37%	115	35%
	Total	49	100%	49	100%	51	100%	19	100%	17	100%	48	100%	47	100%	49	100%	329	100%
qPCR: E	ND	44	94%	44	94%	43	83%	19	100%	14	88%	40	82%	39	81%	46	90%	289	88%
	Detected	3	6%	3	6%	9	17%	0	0%	2	12%	9	18%	9	19%	5	10%	40	12%
	Total	47	100%	47	100%	52	100%	19	100%	16	100%	49	100%	48	100%	51	100%	329	100%

Note: Although the number of samples analyzed using RT-qPCR and RT-ddPCR varied slightly for some buildings/sites, the totals coincidentally equaled 329 for both sets of analyzed samples.

Table S4: Random-effects Poisson regression with cluster-robust standard errors using COVID-19 cases eight days after wastewater sampling as the outcome variable, and controlling for grab vs. composite sampling method

Model: Predictor Variable	IRR	SE	p-value	n	Clusters (buildings)
<i>N</i> - Log10 Viral Copies/100mL (ND=Zero)	1.189	0.093	0.026	215	5
[Covariate] Sampling method (composite=0, grab=1)	0.773	0.122	0.103		
<i>E</i> - Log10 Viral Copies/100mL (ND=Zero)	1.218	0.102	0.018	215	5
[Covariate] Sampling method (composite=0, grab=1)	0.755	0.139	0.127		
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.195	0.067	0.001	211	5
[Covariate] Sampling method (composite=0, grab=1)	0.763	0.124	0.095		
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.263	0.081	<0.001	211	5
[Covariate] Sampling method (composite=0, grab=1)	0.750	0.133	0.104		
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.143	0.060	0.011	202	5
[Covariate] Sampling method (composite=0, grab=1)	0.668	0.099	0.007		
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.226	0.086	0.004	202	5
[Covariate] Sampling method (composite=0, grab=1)	0.662	0.114	0.017		
<i>N</i> - Viral Copies as Binary (ND vs Detected)	1.789	0.433	0.016	215	5
[Covariate] Sampling method (composite=0, grab=1)	0.801	0.125	0.157		
<i>E</i> - Viral Copies as Binary (ND vs Detected)	1.959	0.506	0.009	215	5
[Covariate] Sampling method (composite=0, grab=1)	0.782	0.155	0.215		

Notes: Analysis using RT-ddPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. Log10 = log-10 transformed (ND & <LoD = zero). ND = Not detected. Ppl = people.

Table S5: Random-effects Poisson regression with cluster-robust standard errors where the outcome variable is the number of COVID-19 cases (based on building-specific student assignment to COVID-19 isolation status) eight days after wastewater sampling, using RT-qPCR results as the predictor

Model: Predictor Variable	IRR	SE	P-value	n	Clusters (buildings)
<i>N</i> - Log10 Viral Copies/100mL (ND=Zero)	0.146	0.020	<0.001	212	5
<i>E</i> - Log10 Viral Copies/100mL (ND=Zero)	0.750	0.253	0.393	213	5
<i>S</i> - Log10 Viral Copies/100mL (ND=Zero)	0.883	0.180	0.540	212	5
Any <i>NES</i> Ct Mean (inverted scale, ND=zero)	0.970	0.056	0.596	213	5
Any <i>NES</i> Log10 Viral Copies/100mL (ND=Zero)	0.842	0.197	0.462	212	5
Any <i>NES</i> Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	0.701	0.209	0.233	211	5
Any <i>NES</i> Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	0.686	0.202	0.200	212	5
Any <i>NES</i> Ct Mean as Binary (ND vs Detected)	0.996	0.406	0.992	213	5

Notes: Analysis using RT-qPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. Log10 = log-10 transformed (ND = zero). ND = Not detected. Ppl = people.

Table S6: Random-effects Poisson regression with cluster-robust standard errors using COVID-19 cases 8±1 day (i.e., 3-day bins) after wastewater sampling as the outcome variable

Model: Predictor Variable	IRR	SE	p-value	n	Clusters (buildings)
<i>N</i> - Log10 Viral Copies/100mL (ND=Zero)	1.072	0.087	0.389	216	5
<i>E</i> - Log10 Viral Copies/100mL (ND=Zero)	1.129	0.071	0.054	216	5
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.084	0.099	0.373	212	5
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Swipe (ND=Zero)	1.158	0.082	0.038	212	5
<i>N</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.047	0.103	0.638	203	5
<i>E</i> - Log10 Viral Copies/100mL/100ppl - Housing (ND=Zero)	1.139	0.080	0.065	203	5
<i>N</i> - Viral Copies as Binary (ND vs Detected)	1.241	0.243	0.270	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected)	1.549	0.270	0.012	216	5

Notes: Analysis using RT-ddPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. Log10 = log-10 transformed (ND & <LoD = zero). ND = Not detected. Ppl = people.

Table S7: Random-effects Poisson regression with cluster-robust standard errors where the outcome variable is the number of COVID-19 cases (based on building-specific student assignment to COVID-19 isolation & quarantine status) 7, 8, and 9 days after wastewater sampling, using RT-ddPCR results as the predictor

Model: Predictor Variable	IRR	SE	P-value	n	Clusters (buildings)
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 7days	1.158	0.734	0.818	216	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 7days	1.421	1.031	0.628	216	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 8days	1.860	0.478	0.016	215	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 8days	2.029	0.497	0.004	215	5
<i>N</i> - Viral Copies as Binary (ND vs Detected) – 9days	0.901	0.108	0.386	215	5
<i>E</i> - Viral Copies as Binary (ND vs Detected) – 9days	1.221	0.210	0.244	215	5

Notes: Analysis using RT-ddPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation and quarantine status. ND = Not detected.

Table S8: Random-effects Poisson regression with cluster-robust standard errors where the outcome variable is the number of COVID-19 cases (based on building-specific student assignment to COVID-19 isolation status) 6, 7, 8, 9, and 10 days after wastewater sampling, using RT-qPCR results as the predictor

Model: Predictor Variable	IRR	SE	P-value	n	Clusters (buildings)
Any <i>NES</i> Ct Mean as Binary (ND vs Detected) – 6days	1.189	0.295	0.486	214	5
Any <i>NES</i> Ct Mean as Binary (ND vs Detected) – 7days	0.474	0.304	0.245	214	5
Any <i>NES</i> Ct Mean as Binary (ND vs Detected) – 8days	0.996	0.406	0.992	213	5
Any <i>NES</i> Ct Mean as Binary (ND vs Detected) – 9days	1.868	0.515	0.023	213	5
Any <i>NES</i> Ct Mean as Binary (ND vs Detected) – 10days	2.187	1.040	0.100	213	5

Notes: Analysis using RT-qPCR data from samples collected in the fall 2020 and spring 2021 semesters. COVID-19 case data are based on building-specific student assignment to COVID-19 isolation status. ND = Not detected.

Table S9: Comparison of methods described in six published research articles describing COVID-19 wastewater surveillance initiatives at North American universities

First Author	Publication Date	University	Sample Collection	Frequency of Sampling ^b	Number of Sites Sampled	Concentration Method/s	Matrix Recovery Controls ^a	RT-qPCR SARS-CoV-2 Assay/s	Normalization Approach	Human Fecal Indicator Used	Inhibition Assessment
Corchis-Scott ¹	October, 2021	University of Windsor	Grab	3	1	Ultrafiltration via CP Select	NR	<i>N1</i> , D3L mutations on the <i>N</i> gene	Human Fecal Marker Normalization	Pepper mild mottle virus (PMMV)	The template was diluted 1:5 in all reactions
Juel ²	December, 2021	UNC Charlotte	24 hr Composite	7	37 ^c	Ultrafiltration via CP Select & Electronegative membrane filtration	Bovine Coronavirus	<i>N1</i>	Normalized to Wastewater Volume	NA	An internal positive control assay was run on a subset of samples
Karthikeyan ³	August, 2021	UC San Diego	24 hr Composite	7	68	Automated affinity capture magnetic hydrogel particle based concentration method	NR	<i>N1</i> , <i>N2</i> , <i>E</i>	Human Fecal Marker Normalization	Pepper mild mottle virus (PMMV)	CT values were compared for the no dilution to the 1:100,000 dilution
Reeves ⁴	October, 2021	University of Colorado Boulder	24 hr Composite	7	20	Ultrafiltration	Bovine Coronavirus	<i>N1</i> , <i>N2</i> , <i>E</i>	The RNaseP CT value was subtracted from the SARS-CoV-2 <i>E</i> CT value ^e	Genogroup II F+ RNA bacteriophage	NR
Schmitz ⁵	December, 2021	University of Arizona	Grab	2	13	Ultrafiltration	HCoV 229 <i>E</i>	<i>N1</i>	Normalized to Wastewater Volume	NA	NR
Scott ⁶	September, 2021	Tulane University	Grab	1	10 ^d	Polyethylene glycol precipitation	Pseudomonas phage Phi6	<i>N1</i> , <i>N2</i>	Human Fecal Marker Normalization	Pepper mild mottle virus (PMMV)	Known copy numbers of SARS-CoV-2 synthetic RNA and the CT values were used as reference points for uninhibited samples

Notes:

NR= Not reported; NA= Not applicable

^a before RNA extraction

^b Number of samples per week

^c comprised of dorms, Greek village housing, neighborhood sites of on campus nonresidential buildings

^d 6 dorms, isolation dorm, 2 buildings, the library, and student union

^e The bovine coronavirus and F+ bacteriophage signals were used to track sample variability but not to transform the concentration of SARS-CoV-2 RNA.

REFERENCES

- (1) Corchis-Scott, R.; Geng, Q.; Seth, R.; Ray, R.; Beg, M.; Biswas, N.; Charron, L.; Drouillard, K. D.; D'Souza, R.; Heath, D. D.; Houser, C.; Lawal, F.; McGinlay, J.; Menard, S. L.; Porter, L. A.; Rawlings, D.; Scholl, M. L.; Siu, K. W. M.; Tong, Y.; Weisener, C. G.; Wilhelm, S. W.; McKay, R. M. L. Averting an Outbreak of SARS-CoV-2 in a University Residence Hall through Wastewater Surveillance. *Microbiol. Spectr.* 2021, 9 (2), e00792-21. <https://doi.org/10.1128/Spectrum.00792-21>.
- (2) Juel, M. A. I.; Stark, N.; Nicolosi, B.; Lontai, J.; Lambirth, K.; Schlueter, J.; Gibas, C.; Munir, M. Performance Evaluation of Virus Concentration Methods for Implementing SARS-CoV-2 Wastewater Based Epidemiology Emphasizing Quick Data Turnaround. *Sci. Total Environ.* 2021, 801, 149656. <https://doi.org/10.1016/j.scitotenv.2021.149656>.
- (3) Karthikeyan, S.; Nguyen, A.; McDonald, D.; Zong, Y.; Ronquillo, N.; Ren, J.; Zou, J.; Farmer, S.; Humphrey, G.; Henderson, D.; Javidi, T.; Messer, K.; Anderson, C.; Schooley, R.; Martin, N. K.; Knight, R. Rapid, Large-Scale Wastewater Surveillance and Automated Reporting System Enable Early Detection of Nearly 85% of COVID-19 Cases on a University Campus. *mSystems* 2021, 6 (4), e00793-21. <https://doi.org/10.1128/mSystems.00793-21>.
- (4) Reeves, K.; Liebig, J.; Feula, A.; Saldi, T.; Lasda, E.; Johnson, W.; Lilienfeld, J.; Maggi, J.; Pulley, K.; Wilkerson, P. J.; Real, B.; Zak, G.; Davis, J.; Fink, M.; Gonzales, P.; Hager, C.; Ozeroff, C.; Tat, K.; Alkire, M.; Butler, C.; Coe, E.; Darby, J.; Freeman, N.; Heuer, H.; Jones, J. R.; Karr, M.; Key, S.; Maxwell, K.; Nelson, L.; Saldana, E.; Shea, R.; Salveson, L.; Tomlinson, K.; Vargas-Barriga, J.; Vigil, B.; Brisson, G.; Parker, R.; Leinwand, L. A.; Bjorkman, K.; Mansfeldt, C. High-Resolution within-Sewer SARS-CoV-2 Surveillance Facilitates Informed Intervention. *Water Res.* 2021, 204 (1), 117613. <https://doi.org/10.1016/j.watres.2021.117613>.
- (5) Schmitz, B. W.; Innes, G. K.; Prasek, S. M.; Betancourt, W. Q.; Stark, E. R.; Foster, A. R.; Abraham, A. G.; Gerba, C. P.; Pepper, I. L. Enumerating Asymptomatic COVID-19 Cases and Estimating SARS-CoV-2 Fecal Shedding Rates via Wastewater-Based Epidemiology. *Sci. Total Environ.* 2021, 801, 149794. <https://doi.org/10.1016/j.scitotenv.2021.149794>.
- (6) Scott, L. C.; Aubee, A.; Babahaji, L.; Vigil, K.; Tims, S.; Aw, T. G. Targeted Wastewater Surveillance of SARS-CoV-2 on a University Campus for COVID-19 Outbreak Detection and Mitigation. *Environ. Res.* 2021, 200, 111374. <https://doi.org/10.1016/j.envres.2021.111374>.