RICE UNIVERSITY

Advancing methods for wastewater disease surveillance of antibiotic resistance and SARS-CoV-2

By

Esther Ge Lou

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Lauren Stadler (Aug 10, 2022 11:05 CDT)

Lauren Stadler
Assistant professor,
Civil and Environmental Engineering,
Rice University

Adam Smith (Aug 10, 2022 07:49 MDT)

Adam Smith
Associate Professor,
Astani Department of Civil and Environmental Engineering, University of Southern California

Pedro Alvarez
George R. Brown Professor,
Civil and Environmental Engineering,
Rice University

Todd Treangen
Assistant Professor,
Computer Science,
Rice University

HOUSTON, TEXAS

August 2022

ABSTRACT

Advancing methods for wastewater disease surveillance of antibiotic resistance and SARS-CoV-2

by

Esther Ge Lou

Wastewater-based epidemiology (WBE), which involves using biological indicators in sewage to provide information on the overall health of a community, is a powerful tool to monitor public health. WBE offers several advantages that make it complementary to conventional clinical surveillance: it is rapid and resource-efficient, enables broad monitoring of large populations, is able to detect symptomatic and asymptomatic infections, and is not biased by health seeking behavior or access to healthcare resources. Recent studies have shown that WBE is a powerful tool for estimating community-level prevalence of COVID-19 by measuring levels of SARS-CoV-2 RNA in wastewater, and for predicting the prevalence of clinical antibiotic resistance by screening wastewater for antibiotic resistance genes. Furthermore, WBE has enabled global collaboration through national (e.g., National Wastewater Surveillance System (NWSS) on COVID-19) and international (e.g., the Enhanced Gonococcal Antimicrobial Surveillance Program) programs to advance the integration of WBE into public health response. Despite the surge of interest in applying WBE, there are currently no standardized methods for

wastewater disease monitoring, including how and when to collect samples, what methods to use for analysis, and how to interpret the data to inform action. Without a more complete understanding of the methodological challenges involved in characterizing target indicators in wastewater samples, our ability to leverage WBE for routine monitoring and international collaboration is limited. This dissertation aims to evaluate the strengths and weaknesses of several current methods used for wastewater monitoring of antimicrobial resistance (AMR) and SARS-CoV-2, and discuss implications of method selection for future WBE work. The research focuses on four objectives, corresponding to the four chapters presented in this dissertation: (1) characterize the impact of wastewater sampling designs (i.e., grab and composite sampling) on the ARG removal rates achieved by a wastewater treatment plant (WWTP), (2) elucidate the fate of different forms of cell-associated and cell-free ARGs in an emerging wastewater treatment process, (3) compare two targeted methods (i.e., RT-ddPCR and targeted amplicon sequencing) for monitoring SARS-CoV-2 mutations in wastewater, and (4) evaluate short- and long-read metagenomics and a targeted method (epicPCR) for tracking ARG host range across a WWTP.

Sampling design is critical to the collection of representative samples for WBE and for estimating removal rates of genes across wastewater treatment processes. We compared grab and composite sampling in terms of their corresponding calculated removal rates for a suite of genes, including several clinically-relevant ARGs (*bla*NDM-1, *bla*OXA-1, MCR-1, MCR-5, MCR-10, and *qnrA*). We find that the diurnal variation of ARG loading in the WWTP influent and effluent created significantly different instantaneous ARG removal rates among all grab samples collected throughout a day, indicating grab sampling

can introduce bias to ARG removal calculations. Overall, using composite samples are more representative for WBE and for assessing removal of ARGs across wastewater treatment processes as compared to grab sampling which may overestimate ARG removal rates.

The form of the ARG, specifically whether it is cell-free or cell-associated, is critical to understanding ARG removal across wastewater treatment processes. We found that the fraction of cell-associated ARGs decreased whereas the fraction of cell-free ARGs increased in the treated effluent as the influent organic loading rate was gradually increased. The results indicate that the ARGs in treated effluent can transit between cell-associated and cell-free DNA in response to changing operational conditions, which should be considered to better evaluate the total ARGs in the wastewater treatment system.

WBE has been widely applied to track SARS-CoV-2 infections in communities and in some cases to identify circulating variants of concern. There are several different methods that have been applied to screen for variants of concern in wastewater. We compared two targeted methods for screening SARS-CoV-2 variants of concern in wastewater samples. The results demonstrated that RT-ddPCR is more sensitive and should be applied for mutation quantification or variant confirmation in wastewater, whereas detection via targeted amplicon sequencing was influenced by the depth of sequencing, viral load, and mutation concentration. These findings caution the use of quantitative measurements of SARS-CoV-2 variants in wastewater samples determined solely based on targeted amplicon sequencing.

We compared targeted and untargeted methods for ARG detection in wastewater. The results demonstrate that despite its significantly lower sequencing depth, long-read

sequencing outperforms short-read sequencing with higher sensitivity for detecting ARGs, especially for ARGs associated with mobile genetic elements (MGEs). In addition, long-read sequencing consistently revealed a wider range of ARG hosts compared to short-read sequencing. Nonetheless, the host range detected by long-read sequencing represented only a subset of the host range detected by a targeted method, epicPCR (Emulsion, Paired Isolation, and Concatenation PCR).

Taken together, the results have implications for future WBE, particularly in terms of method selection: 1) collect composite samples rather than grab samples to acquire a representative view of the monitoring targets in a population; 2) include different forms of DNA (cell-associated and cell-free) to analyze ARGs because effluent ARGs are present in both forms and can transition between these forms in response to environmental conditions; 3) apply RT-ddPCR for quantitative analysis and early variant detection if targets are known; and 4) use long-read sequencing for routine wastewater AMR surveillance and use epicPCR to obtain a high-resolution host range of clinically-relevant ARGs. The findings provided by this research contribute to establishing a scientific consensus on method selection for WBE, thus advancing it as a routine tool for public health surveillance.

Acknowledgments

First, I would like to give my sincere gratitude and appreciation to my advisor, Dr. Lauren B. Stadler, for her enormous support, incredible mentorship, and invaluable guidance during my study at Rice University. Her trust and encouragement have helped me persevere in dealing with all challenges and work hard to achieve new goals. She treated her students with compassion and understanding, which makes my research activities full of joy and motivations. I have learnt a great deal from her as a scientist, a leader, and a friend. There is certainly no better advisor for my Ph.D. study. Second, I would give my special thanks to Dr. Pedro J.J. Alvarez, Dr. Todd J. Treangen and Dr. Adam L. Smith for sharing their knowledge, insights, and suggestions. In addition, I would like to thank Dr. Tracy Volz for her guidance in improving my writing and communications skills, which is very important to me as an international student.

I would also like to thank my lab colleagues, especially Eric W. Rice, Prashant Kalvapalle and Zachary W. LaTurner. They helped me through many examinations and presentations. We communicated constantly to share perspectives and ideas on different research questions. I also benefited greatly from the conversations with other graduate students in Department of Civil and Environmental Engineering, Dr. Ruonan Sun, and Dr. Peter Zuo, Dr. Pingfeng Yu and Xiaochuan Huang.

Last but not the least, I would like to thank my parents and friends. I am extremely grateful to my parents, Xia Wang and Jinwei Lou, who have always supported me, cared for me, and listened to me unconditionally. I am also grateful to my lifelong friends, Naomi L. Senehi, Lan Lin, Xiaosheng Chen, Siyi Zhou, Dr. Hassan Javed, Kristina Cibor, and

Zhenglyu Lyu. Because of their influences on me, I aspire to be a benevolent, self-disciplined, and optimistic person. This dissertation would not have been possible without their love and inspiration.

Funding for this research was provided by the National Institute of Food and Agriculture (grant no. 2016-68007-25044), the National Science Foundation (CBET 1805901 and 2029025), a Johnson & Johnson WiSTEM2D 2D award, and seed funds from Rice University.

Based on this research, the following papers were published/submitted:

Lou, E. G., Harb, M., Smith, A. L. & Stadler, L. B. Livestock manure improved antibiotic resistance gene removal during co-treatment of domestic wastewater in an anaerobic membrane bioreactor. Environ. Sci.: Water Res. Technol. 6, 2832–2842 (2020).

*This paper is part of the themed collection: *Environmental Science: Water Research & Technology Recent HOT Articles*

Lou, E. G., Sapoval, N., McCall, C., Bauhs, L., Carlson-Stadler, R., Kalvapalle, P., Lai, Y., Palmer, K., Penn, R., Rich, W., Wolken, M., Brown, P., Ensor, K. B., Hopkins, L., Treangen, T. J., & Stadler, L. B. Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 mutation monitoring in wastewater. *Science of The Total Environment* 833, 155059 (2022).

Lou, E. G., Ali, P., Lu, K., Kalvapalle, P., & Stadler, L. B Diurnal variation of ARG removal across the WWTP emphasized the importance of using composite samples for wastewater-based epidemiology.

*submitted to *ES&T Water*

Lou, E. G., Fu, Y., Wang, Q., Treangen, T. J., & Stadler, L. B. Using long- and short-read metagenomics and epicPCR to profile antibiotic resistance genes and their bacterial hosts in wastewater

*Submitted to *Water Research*

In addition, during my studies at Rice, I contributed to the following papers not included in this dissertation:

Sapoval, N., Liu, Y., Lou, E. G., Hopkins, L., Ensor, K. B., Schneider, R., Treangen, T. J., & Stadler, L. B. Enhanced Detection of Recently Emerged SARS-CoV-2 Variants of Concern in Wastewater. medRxiv 2021.09.08.21263279 (2021) doi:10.1101/2021.09.08.21263279.
*Submitted to *Nature Communications*

Harb, M., Lou, E. G., Smith, A. L. & Stadler, L. B. Perspectives on the fate of micropollutants in mainstream anaerobic wastewater treatment. *Current Opinion in Biotechnology* 57, 94–100 (2019).

Table of Contents

Nomenclature

AMR	antimicrobial resistance
AnMBR	anaerobic membrane bioreactor
ARB	antibiotic resistant bacteria
ARG	antibiotic resistance gene
caARG	cell-associated ARG
COVID-19	the causative virus of novel pneumonia
cfARG	cell-free ARG
epicPCR	Emulsion, Paired Isolation, and Concatenation PCR
HGT	horizontal gene transfer
MGE	mobile genetic elements
RT-ddPCR	reverse transcription droplet digital PCR
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
WBE	wastewater-based epidemiology
WWTP	wastewater treatment plant

Chapter 1 Introduction and Objectives

1.1. Problem Statement

Ever since the first case of methicillin-resistant Staphylococcus aureus (MRSA) was identified in the 1960s (Ventola, 2015), antimicrobial resistance (AMR), or antibiotic resistance, has presented a major threat to public health. New global data showing 1.27 million deaths a year in 2019 reveal the urgent need to address antimicrobial resistance (Murray et al., 2022). More than 2.8 million antibiotic-resistant infections occur in the United States each year, and more than 35,000 people die as a result (CDC, 2022b). From a public health perspective, this situation is cause for alarm, especially in the context of the estimated cost of $55 billion every year in the United States due to AMR, including $20 billion for health care and about $35 billion for loss of productivity (Dadgostar, 2019).

Disease monitoring is critical for public health surveillance and response. In the cases of antimicrobial resistance (AMR) and viral diseases such as COVID-19, surveillance primarily occurs in a clinical setting. This often involves isolating resistant serotypes and SARS-CoV-2 viruses in patients' specimens and studying the factors that contribute to the resistance and virulence, respectively. While clinical surveillance provides information about emerging resistance or virulence, this approach generates data from a very small proportion of the population who are ill and symptomatic. Clinical data on AMR surveillance are scattered and lack representativeness due to a large fraction of asymptomatic or mildly symptomatic infections (Ashley et al., 2019) especially for low- and middle-income countries (Iskandar et al., 2021). This is problematic because individuals can carry AMR for months without having any symptoms but still induce

infections and transmissions (Carlet, 2012). Wastewater-based epidemiology (WBE), on the other hand, provides a snapshot of the overall health of the community based on what is being excreted in pooled sewage samples (O'Keeffe, 2021; Polo et al., 2020). Leveraging wastewater to monitor disease indicators such as SARS-CoV-2 has been so effective that in September 2020, the U.S. CDC established the National Wastewater Surveillance System (NWSS) to inform public health responses to COVID-19 and future outbreaks (CDC, 2022a). NWSS plans to expand to include monitoring for antibiotic resistance gene targets in the near future.

WBE is suitable for monitoring population-level AMR. Wastewater samples reflect the community's gut microbiome (Cai et al., 2014; Newton et al., 2015), which is regarded as the "epicenter" of antibiotic resistance (Carlet, 2012; McInnes et al., 2020; van Schaik, 2015). The vast majority of human gut bacteria are commensal, but they can develop resistance due to the selection effect exerted by antibiotics. Antibiotics are taken therapeutically to treat infections caused by pathogenic bacteria, but can also result in resistance development in commensal bacteria, a phenomenon called "standby selection" (Tedijanto et al., 2018). In addition, horizontal gene transfer (HGT) is able to convert the commensal bacteria into resistant opportunistic pathogens (Carlet, 2012). For example, two ARGs each encoding the production of carbapenemase and extended spectrum beta-lactamase (ESBL), which are prevalent in a few opportunistic pathogens, can readily be transferred to and spread among indigenous commensal Proteobacteria in the gut (Goren et al., 2010; Göttig et al., 2015). The gut microbiome, including resistant organisms, are shed in feces. In sewered communities, excrement along with other wastewater enters the sewer system, where it is conveyed to wastewater treatment plants (WWTPs). Untreated

wastewater (WWTP influent) represents a pooled sample of every person that contributed to the waste stream and can be sampled and analyzed to reveal a community's resistome (defined as "all ARGs, including those circulating in pathogenic bacteria, antibiotic producers, and benign nonpathogenic bacteria") (Wright, 2007, 2010) and ARG hosts (defined as the microorganisms that host ARGs, mainly bacteria) (L. Ma et al., 2016), identify clinically-relevant resistance markers, and track the evolution of multidrug resistance in the community.

WBE has proven to be a powerful tool to monitor not only public health threats such as AMR but also viral diseases such as the COVID-19 pandemic. Numerous studies have demonstrated SARS-CoV-2 RNA levels in wastewater are a strong predictive indicator of trends in the nasal positivity rate (Ai et al., 2021; Arora et al., 2020; Stadler et al., 2020). WWTP influent (untreated wastewater) or primary sludge can be used as a near real-time pooled specimen for an entire community, providing rapid information on trends of community-level prevalence and enabling quick responses. Several studies on WBE of SARS-CoV-2 RNA provided information that was used to evaluate the effectiveness of lockdown measures (Hillary et al., 2021) or to guide public health decisions and actions to control the local pandemic (Prado et al., 2021).

In addition to leveraging influent wastewater for WBE, the microbial quality of WWTP effluent has drawn increased attention primarily driven by interests in using treated wastewater for reclamation and water reuse (Hong et al., 2018; Leiva et al., 2021). Irrigation using reclaimed wastewater may lead to the dissemination of antibiotic resistance (X.M. Han et al., 2016; Q. Qin et al., 2015b), which has become one of the primary obstacles to reuse of treated wastewater for agricultural irrigation. To advance applications

regarding reuse of treated wastewater, it is crucial to investigate the profile and abundance of ARGs and ARG hosts in treated effluents of conventional WWTPs and other treatment systems. The profile and abundance of ARGs and host bacteria in the WWTP effluent can also differ from those in the WWTP influent (Majeed et al., 2021a) because of HGT events and selection effects exerted by the treatment processes (Qiu et al., 2018; T. Zhang et al., 2011). In fact, the predicting or modeling the fate of ARGs and ARG hosts across wastewater treatment processes is still challenging, in part because of the vast diversity of ARGs and ARG hosts present in wastewater, and because of the unknown contribution of HGT among the co-mingling enteric and environmental bacteria in WWTPs. Substantial progress has been achieved to elucidate the close association between wastewater resistome and mobilome (Che et al., 2019b; J. Guo et al., 2017; Yin et al., 2022) that potentially facilitates the transfer of antibiotic resistance between WWTP- and human-associated bacteria (Che et al., 2021, 2022). Future WBE studies should continue investigating the profile of ARGs and ARG hosts in treated effluents to improve our understanding of HGT of ARGs across the treatment processes, in addition to assessing the efficacy of wastewater treatment in attenuating the risk of AMR dissemination.

In spite of the significance and benefits in studying these public health indicators (i.e., ARGs for AMR and SARS-CoV-2 RNA for COVID-19) across WWTPs, currently there is no consensus among scientists and practitioners on which methods should be used to characterize the public health indicators in wastewater samples (Liguori et al., 2022). Without a more complete evaluation of the methodologies to investigate the target indicators, our ability to advance WBE for routine monitoring and to study the fate of AMR across wastewater treatment processes is limited. Therefore, it is important to have a

thorough evaluation of methods that can impact the results of target indicators and provide guidelines of method selection for characterizing target indicators in wastewater samples. These guidelines can contribute to closing the knowledge gap regarding risk assessment of AMR and other disease indicators for public health protection.

1.2. Objectives and Significance

The overall objective of this research is to evaluate a set of WBE methods including wastewater sampling method (i.e., grab versus composite sampling), concentration methods for cell-associated and cell-free DNA, and molecular analysis techniques (i.e., RT-ddPCR, epicPCR, targeted amplicon sequencing and metagenomics sequencing). Specifically, this research aims to explore the following subjects:

Investigate the impact of sampling design on the variability and instantaneous removal of ARGs across a wastewater treatment plant (WWTP)

Significance. The nature of wastewater samples poses several challenges that can impact the detection of the signals associated with target indicators. First, the flow and composition of wastewater comprising the influent to a WWTP can fluctuate drastically over a day and due to weather conditions, which can affect the variability of measurements and consistency of detection of target indicators over time. Wastewater sampling is the first step in the WBE workflow, and the representativeness of collected samples influences the quality of the resulting information. Considering variations in the flow and composition of influent, composite sampling appears to be a better strategy compared to grab sampling because it provides a pooled sample matrix consisting of multiple discrete subsamples collected within a period of time (i.e., usually 24 hours), which accounts for diurnal

variations. Recent studies have suggested that diurnal variations of SARS-CoV-2 RNA in wastewater can potentially lead to false-negative detections when grab samples are collected during the low-loading time window (Bivins et al., 2021a; Mendoza Grijalva et al., 2022a). Yet the majority of researchers studying AMR in water and wastewater system continue to use grab instead of composite samples for AMR monitoring likely because of convenience (S. Sun et al., 2021). To assess the importance of collecting representative samples on AMR monitoring results and conclusions on ARG removal across wastewater treatment, it is necessary to investigate the diurnal variations of the loading of a suite of clinically relevant ARGs across the treatment process units of a WWTP and the associated ARG removal rates.

Examine the impact of wastewater sample processing on the quantification results of target indicators

Significance. Another factor that can affect the results of WBE is how a sample is processed and target analytes are concentrated prior to quantification. Target analytes in wastewater are often extremely dilute, and thus a concentration step is required prior to quantification. The choice of sample concentration protocol impacts the recovery of the target analyte in a sample. For example, filtration-based concentration protocols use size selection methods and as a result are biased in the form of analyte they capture. In WWTPs, AMR gene targets can be present in both cell-associated and cell-free forms depending on the status of the host cells. However, the method widely used for wastewater sample concentration (i.e., filtration using 0.22 μm filters) is biased towards capturing the cell-associated fraction (Y. Zhang et al., 2018b). For example, studies have shown that the vast

majority of anaerobic enteric bacteria decay in the sewer system during transport to the WWTP (L. Li et al., 2021) and during the wastewater treatment process (Eregno et al., 2018). Cell decay can release cell-free DNA into the surroundings (Ibáñez de Aldecoa et al., 2017), converting cell-associated ARGs to cell-free ARGs. In fact, cell-free ARGs in treated effluents of WWTPs are of particular importance because they are highly persistent in the downstream environment (Nagler et al., 2018; Oliveira et al., 2021; Sivalingam et al., 2020) and can disseminate via natural transformation, as shown in microcosms (Dong et al., 2019; Mao et al., 2014) and *in situ* experiments (Kittredge et al., 2022). The development of cell-free DNA concentration and extraction methods (Calderón-Franco, van Loosdrecht, et al., 2021; D.-N. Wang et al., 2016; Q.-B. Yuan et al., 2019b) make it feasible to analyze cell-free ARGs from wastewater samples, but a limited number of studies have investigated the fate of cell-associated and cell-free DNA in effluent from treatment systems for wastewater reclamation. Anaerobic membrane bioreactors (AnMBRs) are an emerging biotechnology which is ideally suited for wastewater reclamation because they produce nutrient-rich effluents (nitrogen and phosphorus) that can be utilized for irrigation (Martinez-Sosa et al., 2011; Peña et al., 2019). However, its ability to remove ARGs from wastewater has not been studied carefully, precluding its application for water reuse. Therefore, to advance AnMBRs for water reuse applications, it is important to understand their ability to remove ARGs and the persistence of effluent ARGs in both the cell-associated and cell-free fractions.

Evaluate two targeted methods (RT-ddPCR and targeted amplicon sequencing) in quantifying characteristic mutations of SARS-CoV-2 variants in wastewater

Significance. In addition to sampling and concentration methods, the choice of molecular analysis techniques (i.e., PCR-based quantifications, sequencing technologies) used to detect and quantify target indicators in DNA and RNA isolated from the wastewater samples can also impact WBE results and conclusions. To elaborate, we categorized these techniques into two groups: targeted methods such as RT-ddPCR, epicPCR (Emulsion, Paired Isolation and Concatenation PCR) and targeted amplicon sequencing, and non-targeted methods such as short-read and long-read metagenomics sequencing. Wastewater represents a complex matrix of biological and chemical compounds which can inhibit molecular analysis (Schrader et al., 2012). In addition, target indicators such as SARS-CoV-2 RNA in wastewater can be degraded, fragmentized and diluted (Canh et al., 2021; Kantor et al., 2021; Wurtzer et al., 2021). All these factors make it challenging to detect and recover target indicators. Therefore, for detecting SARS-CoV-2 titers in wastewater, an enrichment step based on PCR amplification is typically needed prior to sequencing to improve sensitivity (Crits-Christoph et al., 2021a). For instance, targeted approaches based on PCR amplification, including RT-ddPCR and amplicon-based targeted sequencing, have been widely applied to detect SARS-CoV-2 RNA in wastewater (Ciesielski et al., 2021; Tyson et al., 2020). Targeted amplicon sequencing (i.e., multiplex tiling PCR coupled with amplicon sequencing) is considered the lower-cost and faster approach (Chiara et al., 2021; X. Lin et al., 2021). In addition, it is a powerful tool to enable comprehensive screening of all potential mutations without any prior knowledge, providing opportunities to proactively discover cryptic variant lineages (Karthikeyan et al., 2022; Smyth et al., 2022) and emerging lineages of concern (Karthikeyan et al., 2022; Sapoval et al., 2021). However, it is unclear how quantitative is this approach as compared

to the gold standard RT-qPCR (Alygizakis et al., 2021; Van Poelvoorde et al., 2021) or RT-ddPCR (Nyaruaba et al., 2022) methods. Thus, a direct comparison between targeted amplicon sequencing and RT-ddPCR (or RT-qPCR) for mutation detection and quantification using wastewater samples is needed to assess the validity of using the results generated via targeted amplicon sequencing to indicate quantitative information of mutations in the community.

Evaluate a targeted method (epicPCR) and a widely used non-targeted method (metagenomic sequencing) in tracking ARG hosts in wastewater

Significance. While targeted methods enable sensitive detection of genetic targets in wastewater, they depend on a priori knowledge of the target sequences to design primers and/or probes for PCR amplification (e.g., primers/probes for RT-ddPCR or RT-qPCR, and primer panels for targeted amplicon sequencing via multiplex tiling PCR). In contrast, non-targeted, unbiased methods such as metagenomics do not require prior knowledge of targets and thus can be used to comprehensively screen of wastewater samples for a variety of disease targets including ARGs (Prieto Riquelme et al., 2022), DNA and RNA viruses (Corpuz et al., 2020), bacterial and fungal pathogens (Y. Yang et al., 2022) etc., albeit with less sensitivity and arguably less quantitatively. Metagenomic sequencing is one of the most widely used methods to analyze AMR in wastewater. It has several advantages such as showing the contextual information of ARGs [e.g., the associations of ARGs with mobile genetic elements (MGEs)] and enabling ARG host tracking (Ju et al., 2019; L. Ma et al., 2016; G. Zhang et al., 2020). The majority of wastewater metagenomic sequencing studies on AMR have used short-read sequencing platforms, generating amplicons of 75-

350 bp. The development and application of long-read based sequencing (e.g., Nanopore technology), which can generate amplicons from 500 bp to up to 2.3 Mb (Payne et al., 2019) can provide contextual genetic information for ARGs, and thus has improved confidence and efficiency in identifying ARGs associated with MGEs, multi-drug resistance, and ARG hosts (Che et al., 2019b; Leggett et al., 2020; Y. Yang et al., 2022). So far, no study has directly compared how short- and long-read sequencing approaches for detection of ARG hosts in wastewater. In addition, it is unknown how non-targeted metagenomic sequencing differ from targeted methods specifically for ARG host tracking. Evaluating these methods in terms of ARG host tracking is important because of the consensus that not just the ARGs, but the ARG and its associated hosts, is the critical to assess AMR threats to public health.

1.3. Dissertation Organization

This dissertation is organized into 6 chapters. **Chapter 1** gives an overview of the general background and contextualizes the research objectives and significance. **Chapter 2**, "Instantaneous ARG removal rates across wastewater treatment plants are not representative due to diurnal variations", assesses the impact of sampling design (grab versus composite sampling) on the conclusions regarding the variability and instantaneous removal of ARGs across a WWTP. **Chapter 3**, "The fate of cell-associated and cell-free antibiotic resistance genes in the effluent of an anaerobic membrane bioreactor co-treating domestic wastewater and cattle manure", investigates effluent ARGs and their presence in the cell-associated and cell-free fractions of the treated effluent of an AnMBR treating a high strength mixture of domestic wastewater and cattle manure. **Chapter 4**, "Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 mutation

monitoring in wastewater", compares the detection sensitivity of two different targeted methods for mutation detection, RT-ddPCR and amplicon-based targeted sequencing (ARTIC v3), for the detection and quantification of five characteristic mutations associated with SARS-CoV-2 variants of concern. **Chapter 5**, "Using long- and short-read metagenomics and epicPCR to profile antibiotic resistance genes and their bacterial hosts in wastewater", evaluates the non-targeted methods, namely, short- and long-read metagenomic sequencing, as well as a targeted approach that uses single-cell emulsion PCR (epicPCR), to identify ARG hosts across a WWTP. **Chapter 6** summarizes the findings of this work, discusses its engineering implications, and makes some suggestions for future research.

Chapter 2 Instantaneous ARG removal rates across wastewater treatment plants are not representative due to diurnal variations

Esther G. Lou, Priyanka Ali, Karen Lu, Prashant Kalvapalle and Lauren B. Stadler

Abstract To evaluate the threat of the environmental dissemination of antibiotic resistance associated with wastewater treatment plants (WWTPs), the removal efficiency of antibiotic resistance genes (ARGs) during wastewater treatment needs to be assessed. Sample collection strategy is one factor that is often overlooked in study design and most studies on ARGs in wastewater perform grab sampling. In this study, we hypothesized that wastewater sampling (i.e., grab sampling and composite sampling) influences the observed ARG concentrations and calculated removal rates across a WWTP. We compared the removal rates calculated based on the two different sampling methods for a suite of genes, including several clinically-relevant ARGs (*bla*NDM-1, blaOXA-1, MCR-1, MCR-5, MCR-10, and *qnrA*). We conducted summer and winter 24-hour sampling campaigns where grab samples were collected every two hours from the influent, secondary effluent, and final effluent. We found the instantaneous removal rate of each target gene calculated based on the 12 grab samples fluctuated by 0.5 – 1.6 log in the winter and 0.9-2.7 log in the summer throughout the day, indicating diurnal variation. Overall, for each target gene, the removal rates calculated based on 24-hour composite samples were approximately equal to the median of the 12 instantaneous removal rates. Our study confirms the importance of using composite rather than grab samples to monitor ARGs in wastewater.

2.1. Introduction

The worldwide dissemination of antibiotic resistance has raised serious public health concerns. According to the CDC, more than 2.8 million antibiotic-resistant infections occur in the US each year and more than 35,000 people die as a result (CDC, 2022b). Antibiotic resistant organisms living in the human gut are excreted via feces and enter wastewater treatment plants (WWTPs) through public sewer systems (Rousham et al., 2018). Therefore, WWTPs are regarded as a hotspot for the environmental dissemination of antibiotic resistance genes (ARGs) (Rizzo et al., 2013b). It is important to incorporate ARG removal efficiency assessment into the body of environmental monitoring of antibiotic resistance, because it could shed light on strategies to mitigate the spread of antibiotic resistance from WWTPs (Liguori et al., 2022). In addition, ARGs have great potential to propagate and spread among bacteria during wastewater treatment, which results in persistence and proliferation of ARGs across the treatment processes (Che et al., 2019a; Ju et al., 2019; Majeed et al., 2021a). Furthermore, the profile and load of ARGs in treated effluent needs to be evaluated, particularly in the context of water reuse, to assess their potential human health risks (Hong et al., 2018; Leiva et al., 2021).

The pattern of ARG removals can be complicated by the diurnal variations in ARG loads at a given WWTP due to flow fluctuations that result from toilet flush frequency, which usually peaks in the morning and early evening (Coutu et al., 2013b). Diurnal variations of many different constituents in wastewater such as suspended solids, chemical oxygen demand (COD), ammonia (Butler et al., 1995; Metcalf & Eddy, 1991; Munksgaard & Young, 1980; Young et al., 1978), pharmaceuticals (Joss et al., 2005; Nelson et al., 2011), antibiotics (Göbel et al., 2005; Plósz et al., 2010; Coutu et al., 2013a), xenobiotics (Joss et

al., 2005; Nelson et al., 2011) have been studied. Daily loads of antibiotics were found to correlate with influent flow and ammonium load (Göbel et al., 2005) and to peak daily in the morning hours during the highest hydraulic loading (Plósz et al., 2010). Viruses in wastewater, such as SARS-CoV-2, also exhibit diurnal variations (Bivins et al., 2021b; Gerrity et al., 2021; B. Li et al., 2021; Mendoza Grijalva et al., 2022b). SARS-CoV-2 detection rate, defined as the proportion of samples with quantifiable RNA concentrations, was the highest at peak flow (Mendoza Grijalva et al., 2022b). A 10-fold increase in RNA concentration in the influent was measured in 24-hr, flow proportional composites samples as compared to grab samples collected in the morning (at minimum observed flow rate), due to diurnal variation (Gerrity et al., 2021). However, to the best of our knowledge, only one study has examined the temporal variations of ARGs in WWTPs that occur over the course of a day. Sun et al. reported that the concentrations of several tetracycline and sulfonamide resistance genes fluctuated across a day (S. Sun et al., 2021). ARG removal rates across different treatment units processes in WWTPs may also differ throughout the day due to the impact of flow variation on the hydraulic residence time in each unit process, similar to what has been observed for antibiotics (Marx et al., 2015).

Previous studies have reported inconsistent ARG removals across WWTPs, and discrepancies in removals underscore the current challenges in predicting the fate of ARGs in WWTPs (Berendonk et al., 2015). It is also unknown if wastewater sampling contributes to the inconsistencies in ARG removal among studies. For example, Sabri et al. found that the concentrations of *sul1* in the secondary effluent increased as compared to the influent in one WWTP (Sabri et al., 2020), whereas Gao et al. reported approximately 2-log of reduction of *sul1* between influent and secondary effluent (Gao et al., 2012b). Both studies

were on conventional activated sludge process, but Sabri et al. used composite sampling and Gao et al. used grab sampling. A myriad of factors, such as wastewater composition, WWTP designs and operational parameters can impact the removal of ARGs (Uluseker et al., 2021; J. Wang & Chen, 2022). In order to accurately compare ARG removal across full-scale systems that experience dynamic loading, it is critical to standardize how removal of ARGs is measured and calculated, and thus it is necessary to understand how sampling methodological decisions impact the interpretation of results reported across studies.

There are two dominant sampling methods used for wastewater monitoring: grab sampling and composite sampling (Simpson et al., 2013). A grab sample is typically a single, discrete sample. Composite samples can be divided into two categories: (1) time-weighted composite samples, and (2) flow-proportional composite samples. These types of composite samples consist of a single sample pooled from a series of individual samples taken at a specific time interval with the same volume (time weighted), or with flow-proportional volumes (flow proportional). A composite sample (most often performed over a 24-hour period) is considered a more representative sample than a grab sample because it is less likely to be affected by an irregularity in the system and is an averaged representation of wastewater characteristics during the compositing period (H. Lee et al., 2007). However, grab sampling is much easier and faster to conduct as compared to composite sampling. In a recent study, the authors reported that out of 55 prior studies that sampled from a total of 263 WWTPs for ARG analysis, 42 of those studies analyzed grab samples (S. Sun et al., 2021). Little is known if these two wastewater sampling methods (grab versus composite sampling) can impact ARG wastewater monitoring results and assessments of ARG removal rates across wastewater treatment processes. A better

understanding of the diurnal variability of ARGs across wastewater treatment processes is needed.

In this study, we assess the results from grab and composite sampling by performing two 24-hour sampling campaigns, one in winter and one in summer. Samples were collected every two hours from the influent, secondary effluent, and final effluent of a WWTP. A suite of clinically relevant genes including ARGs that confer resistance to fluoroquinolone (*qnrA*), carbapenem (*bla*NDM-1), ESBL (*bla*OXA-1), and colistin (MCR-1, MCR-5, and MCR-10) as well as a clinical class 1 integron-integrase gene (*IntI1*) were quantified. We examined: 1) the concentration of ARGs and ARG load across a typical weekday; (2) the variability in calculated ARG removal rates based on intra-day hourly grab samples; (3) the difference between ARG removal rates calculated based on the grab samples and 24-hour composite samples; and (4) the impact of secondary treatment (activated sludge treatment and clarification) and chlorine disinfection on the removal of ARGs.

2.2. Materials and Methods

2.2.1 24-hour sampling campaign

We sampled from the City of West University Place WWTP (Houston, TX) that treats an average of approximately 0.86 (summer) and 0.93 (winter) million gallons of municipal wastewater per day. The instantaneous plant flow rates are provided in the Supporting Information (Table S2.1). The treatment process consists of primary screening, secondary treatment via a contact stabilization process, secondary clarification, chlorine

disinfection using gaseous Cl_2, and sodium bisulfite addition for dechlorination. The hydraulic retention time (HRT) of the plant is approximately 12 hours depending on the flow rate. Two 24-hour continuous sampling campaigns were conducted on two separate weekdays, one in December 2019 and one in June 2022. Both sampling campaigns were conducted during dry weather conditions. Samples were collected from three locations: influent channel, secondary effluent (after contact stabilization and clarification), and final effluent (after dechlorination). Samples were collected every 2 hours over 24-hours, and replicate samples (n=2) were collected during the summer campaign. Instantaneous flow rate was recorded at each sampling event via a digital flow meter measuring the plant's final effluent. Field campaign details including the sampling time, ambient temperature, and number of samples collected are listed in Table S2.2.

Immediately after collection, samples were aliquoted and aliquots were preserved for determination of chemical oxygen demand (COD), ammonia-N (NH_3-N), and for DNA extraction. Sample volumes, a description of the pretreatment and storage of samples, and procedures for COD and NH_3-N quantification are described in the Chapter 2 Appendix Section 1.1. The results of COD and NH_3-N concentrations and removal rates are provided in Table S2.3. DNA extraction was performed using FastDNA SPIN Kits for Soil (MP Biomedicals) following the manufacturer's instructions.

2.2.2 Real-time PCR (qPCR) on target genes

We used qPCR to quantify the absolute abundance (concentration) of each target gene in the samples. Target genes for qPCR quantification included several ARGs [*sul1*, *tet*(W), *AmpC*] and one mobile genetic element (MGE), and clinical *IntI1* (Zheng et al.,

2020), all of which are ubiquitous and abundant in the environment (Gao et al., 2012; Laht et al., 2014; J. Wang et al., 2020). We also analyzed a suite of ARGs that are clinically-relevant, including three colistin resistance gene variants (MCR-1, MCR-5, and MCR-10), one fluoroquinolone resistance gene (*qnrA*), one ESBL-production gene (*bla*OXA-1), and one carbapenem resistance gene (*bla*NDM-1). We quantified the 16S rRNA gene to measure the abundance of bacteria and other microorganisms and use for normalization of ARG and mobile genetic elements (MGE). Each grab sample was analyzed as an individual sample without compositing. Since all grab samples were collected at a constant time interval, and the volume for each type of sample (influent, secondary effluent, and final effluent) was constant, we simulated the target gene concentration of the time-weighted composite sample by averaging the concentrations of the target gene in all grab samples. For the flow proportional composite sample, we calibrated the volume of each individual grab sample proportional to the flow, then pooled them together to calculate the concentration of the target gene in a simulated flow-proportional composite sample.

The copy numbers for all genes except for three MCR variants were assayed using a SYBR green-based method (i.e., one target per assay). The qPCR reactions were carried out in technical triplicates each containing 10 µL 2X Forget-Me-Not EvaGreen qPCR mastermix (Biotium), 50 nM ROX (Biotium), 250 nM of forward and reverse primer, PCR grade H_2O, and 1 µL DNA template. A triplex assay was developed in this study to multiplex MCR-1, MCR-5 and MCR-10 using a probe-based method. The qPCR reactions were carried out in technical triplicates each containing 10 µL 2X qPCRBIO Probe Mix (PCRBIOSYSTEMS), 400 nM forward and reverse primer, 200 nM probe, PCR grade H_2O, and 2 µL DNA template. All targets were assayed using the QuantStudio 3 Real-Time PCR

system (ThermoFisher Scientific) using the fast 2-step mode. Primer sequences, standard curve information, and qPCR reaction conditions are provided in Table S2.4.

2.2.3 Calculations of instantaneous ARG removal rates using grab samples, ARG removal rates of 24-hour composite samples, and the removal of 24-hour ARG loads

We first used the copy number of target genes (16S rRNA gene, ARGs and *IntI1*) in each sample obtained by qPCR and the volume of the corresponding sample to calculate the absolute abundance (i.e., concentration in copies/L), as described in Equation 1. The relative abundance of ARGs and *IntI1* was calculated by normalizing the absolute abundance of the target gene to the absolute abundance of the 16S rRNA gene (Equation 2). The instantaneous removal rate of each target gene via grab sampling was calculated in terms of the absolute abundance and relative abundance in the influent and effluent samples (Equation 3). To calculate the removal rates based on 24-hour composite sampling, the simulated absolute abundance of the target genes in the composite influent and effluent samples as described above were used (Equation 4 & 5). The removal of the load of target genes via 24-hour composite sampling was calculated using Equation 6. A detailed explanation on Equation 6 is provided in the Chapter 2 Appendix Section 1.4 (Figure S2.1). Note that for the removal rates, we calculated the removal by secondary treatment (Equation 3a, 4a, 5a and 6a) and by chlorination (Equation 3b, 4b, 5b and 6b).

Absolute abundance of the target gene: $C_{abs.}$

$$= \frac{\text{copy number of the target gene per reaction (copies)} \times \text{total volume of the DNA extract per sample (μL)}}{\text{volume of sample (L)} \times \text{volume of the DNA extract per reaction (μL)}} \quad \text{(Eq. 1)}$$

Relative abundance: $C_{rel.} = \frac{C_{abs.}}{C_{16S\ rRNA,abs.}}$ (Eq. 2)

Instantaneous removal rate by secondary treatment $= \mathrm{Log}_{10}\left(\frac{\mathrm{C}_{\mathrm{abs.,INF}}}{\mathrm{C}_{\mathrm{abs.,SE}}}\right)$ (Eq. 3a)

Instantaneous removal rate by chlorination $= \mathrm{Log}_{10}\left(\frac{\mathrm{C}_{\mathrm{abs.,SE}}}{\mathrm{C}_{\mathrm{abs.,FE}}}\right)$ (Eq. 3b)

Removal rate by secondary treatment in the $24-$ hour time weighted composite sample

$$= \mathrm{Log}_{10}\left(\frac{\sum_{t=0}^{24} C_{t,INF}}{\sum_{t=0}^{24} C_{t,SE}}\right) \text{(Eq. 4a)}$$

Removal rate by chlorination in the $24-$ hour time weighted composite sample

$$= \mathrm{Log}_{10}\left(\frac{\sum_{t=0}^{24} C_{t,SE}}{\sum_{t=0}^{24} C_{t,FE}}\right) \text{(Eq. 4b)}$$

Removal rate by secondary treatment in the $24-$ hour flow proportional composite sample

$$= \mathrm{Log}_{10}\left(\frac{\sum_{t=0}^{24} \mathrm{C}_{\mathrm{t.,INF}} f_t}{\sum_{t=0}^{24} \mathrm{C}_{\mathrm{t.,SE}} f_t}\right) \text{(Eq. 5a)}$$

Removal rate by chlorination in the $24-$ hour flow proportional composite sample

$$= \mathrm{Log}_{10}\left(\frac{\sum_{t=0}^{24} \mathrm{C}_{\mathrm{t.,SE}} f_t}{\sum_{t=0}^{24} \mathrm{C}_{\mathrm{t.,FE}} f_t}\right) \text{(Eq. 5b)}$$

The removal of ARG loads by secondary treatment in the $24-$ hour sampilng period

$$= \mathrm{Log}_{10}\left(\frac{\mathrm{C}_{0,\mathrm{INF}} f_0 + 2\sum_{t=2}^{22} \mathrm{C}_{\mathrm{t,INF}}\, f_t + \mathrm{C}_{24,\mathrm{INF}} f_{24}}{\mathrm{C}_{0,\mathrm{SE}} f_0 + 2\sum_{t=2}^{24} \mathrm{C}_{\mathrm{t,SE}}\, f_t + \mathrm{C}_{24,\mathrm{SE}} f_{24}}\right) \text{(Eq. 6a)}$$

The removal of ARG loads by chlorination in the $24-$ hour sampilng period

$$= \mathrm{Log}_{10}\left(\frac{\mathrm{C}_{0,\mathrm{SE}} f_0 + 2\sum_{t=2}^{22} \mathrm{C}_{\mathrm{t,SE}}\, f_t + \mathrm{C}_{24,\mathrm{SE}} f_{24}}{\mathrm{C}_{0,\mathrm{FE}} f_0 + 2\sum_{t=2}^{22} \mathrm{C}_{\mathrm{t,FE}}\, f_t + \mathrm{C}_{24,\mathrm{FE}} f_{24}}\right) \text{(Eq. 6b)}$$

t (hour) – the time at which the grab sample was collected with a range from 0 to 24 and an increment of 2,

f_t (L/hour) – the instantaneous flow rate at time t.

2.2.4 Statistical analysis

A Welch two sample t-test was applied to compare two datasets. A p-value of less than 0.05 was the threshold to indicate a "significant" difference between the two datasets.

The relative standard deviation (RSD), which is usually used to represent how data points are scattered around the mean, was calculated to indicate the diurnal variation in instantaneous concentrations of each target ARG across a day.

2.3. Results and Discussion

2.3.1 Instantaneous removal rates via grab sampling varied significantly within a day, demonstrating diurnal variations

We first assessed the instantaneous removal rates of each target gene via grab sampling with respect to time, then compared them to the removal rates calculated based on the simulated 24-hour composite samples (Figure 2.1a). All target genes demonstrated removal rates greater than 90% (i.e., one-log removal) via 24-hour composite sampling,

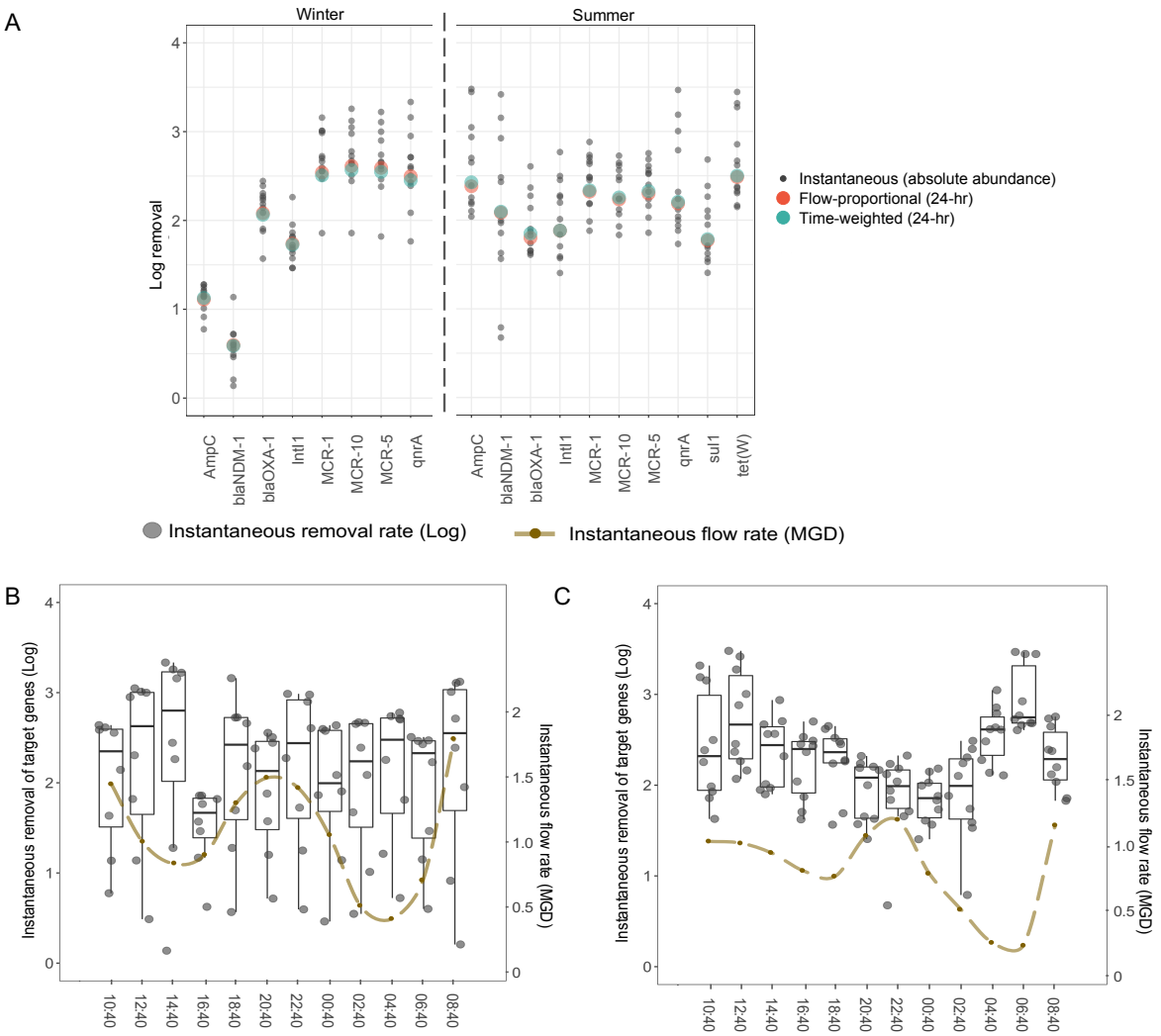


Figure 2.1 Removal rates of target genes in winter and summer sampling. Removal rates of target genes (x-axis) as calculated based on grab samples collected during the winter and summer sampling campaign (a), including instantaneous log removal rates (grey points), log removal rates of the 24-hour time-weighted composite sample (red) and the flow-proportional composite sample (blue). Instantaneous removal rates (left y-axis) of all target genes and the flow rate (right y-axis) over 24 hours during the winter (b) and summer sampling campaigns (c). The scatter points represent the instantaneous removal rates of all target genes at that specific time. Boxes represent the interquartile range, with solid lines as medians. Whiskers represent the standard deviation. The dashed brown line represents the flow rate.

which is consistent with previous studies reporting that WWTPs are effective in reducing ARG concentrations from wastewater (Karkman et al., 2016; K. Qin et al., 2020). However, instantaneous removal rates were highly variable throughout the day (Figure 2.1b & c): the difference between the maximum and minimum observed instantaneous removal rates of each target gene fluctuated by 0.5 – 1.6 log in the winter and 0.9-2.7 log in the summer. This trend was most evident in the summer sampling campaign data (Figure 2.1c), where higher ARG removal rates tended to be observed using grab samples collected during the daytime, whereas lower ARG removal rates were observed for samples collected during the nighttime. The maximum instantaneous removal rates of all target genes were observed most often for the samples collected in the early morning (6:40 AM) or late night (12:40 AM) in summer, and for those samples collected at 2:40 PM in winter. These specific times coincided with times when the plant flow rate was the lowest across the day (Figure 2.1b & c). However, it is hard to disentangle the intrinsically complicated effects of the flow rate on ARG removal. Flow impacts the concentration of genes via dilution (i.e., the fecal content in the wastewater becomes diluted by other liquids that do not contain the genes). Additionally, flow fluctuation can change the residence time of each treatment process, which in turn impacts the removal of the target genes.

The 24-hour time-weighted and flow-proportional samples generated similar removal rates with respect to a given target gene (Figure 2.1a). For each target gene, the percentage difference between the removal rate of the 24-hour time weighted sample and 24-hour flow proportional sample ranged between 1.50-21.4% (mean: 3.40%) in the winter campaign, and from 0.794-26.5% (mean: 11.0%) in the summer campaign. Statistically, the median of the instantaneous removal rates of the 12 grab samples approached the values

of removal rates corresponding to the 24-hour composite samples. To summarize, when the grab sampling approach is used for quantitative ARG measurements across WWTPs, diurnal variation can potentially exclude samples with the lowest ARG removal, leading to biased removal results. Both time and flow composite sampling methods generated comparable ARG removal rates, underscoring that composite samples provide a more balanced and thus representative snapshot of target gene removal trends in a WWTP.

2.3.2 Seasonality of ARG removal was significant as shown by an overall greater removal in relative abundance in the summer

In the summer and winter sampling campaigns, all target genes were successfully removed in absolute abundance as indicated by the average positive removal rates generated by the grab samples collected across the day (Figure 2.1). However, the relative abundance of several target genes (calculated as the absolute abundance of target genes/the absolute abundance of 16S rRNA genes) in the final effluent increased as compared to in the influent, rendering negative removal rates, especially in the winter (Figure S2.2). One possible reason for the increase in relative abundance may be the ecological selection of the corresponding ARG hosts due to their ubiquity and environmental fitness across treatment (Do et al., 2022; Pärnänen et al., 2019). In addition, apart from the ARG hosts, the environmental conditions in the wastewater treatment processes that resulted in increased expression of the ARGs can also contribute to the increase of ARG relative abundance. Those conditions include the selective pressure posed by antibiotics, biocides, heavy metals (Ju et al., 2019; McNamara et al., 2014), and disinfectants or disinfection byproducts that can induce co-selection (D. Li et al., 2016).

In the winter, *AmpC* and *bla*NDM-1 consistently showed negative instantaneous removal rates in relative abundance across a day (n = 12, Figure S2.2). Clinical *IntI1* showed negative instantaneous removal rates in relative abundance as observed in 10 out of 12 grab sampling events. In addition, negative instantaneous removal rate in relative abundance was observed once for *bla*OXA-1 among the 12 events. In the summer, although negative instantaneous removal rates in relative abundance were observed once (*bla*OXA-1, clinical *IntI1*, MCR genes, *sul1*) or twice (*bla*NDM-1) across a day (n=12), the overall removal in relative abundance was greater as compared to in the winter (Figure S2.2).

Specifically, as indicated by instantaneous removal rates calculated from all grab samples, for the target genes *AmpC* and *bla*NDM-1, the removal rates in both absolute and relative abundance were significantly higher in the summer ($p<0.001$, n=12); for clinical *IntI1*, the removal rates in relative abundance were significantly higher in the summer ($p<0.001$, n=12). In terms of removal rates calculated based on 24-hr composite samples, the three target genes, *AmpC*, *bla*NDM-1 and clinical *IntI1*, showed negative removal rates in relative abundance in the winter. This is concerning because *bla*NDM-1 is a clinically relevant ARG that codes for the production of New Delhi metallo-β-lactamase which hydrolyzes carbapenem, one of the most reliable last-resort treatments for bacterial infections (Meletis, 2016). Interestingly, the removal rates in absolute abundance of MCR genes were significantly higher in the winter ($p<0.05$, n=12). The seasonal variations in the removal efficiency for certain ARGs were reported by several studies (An et al., 2018; H. Chen & Zhang, 2013b; Schages et al., 2020; Sui et al., 2017; Zheng et al., 2019). However, considering the negative removal rates in relative abundance and the loads in the treated

effluent (discussed further in Section 2.3.4), certain ARGs such as *bla*NDM-1 and *AmpC* should receive extra attention during the winter, when they may proliferate in the WWTP.

2.3.3 Greater removal of target genes occurred during secondary treatment than chlorine disinfection

We next characterized the diurnal variation of ARGs for each sampling location across the WWTP (i.e., influent, secondary effluent, and final effluent) by calculating the relative standard deviation (RSD) of the target gene concentrations in grab samples collected throughout the day (Table S2.5). The diurnal variations of target gene concentrations in the influent, secondary effluent, and final effluent samples were significant, indicated by the large RSD value of the instantaneous absolute abundance of each target gene calculated for samples across all three sampling locations (ranging from 19.1 – 99.2% for influent, 24.7 – 99.6% for secondary effluent, and 20.3 – 163.1% for final effluent, Table S2.5). These RSD values are comparable to those reported by Sun et al. (S. Sun et al., 2021). Furthermore, we found significant differences in RSD values across the three sample types, where influent RSD < secondary effluent RSD < final effluent RSD ($p<0.0001$), indicating that the wastewater treatment processes magnify the diurnal variability of the concentration of the genes. These findings underscore that composite samples are more representative for assessing the concentrations and loads of ARGs across each wastewater treatment unit process.

To further understand how the treatment process impacted the observed trend in removal rates across the day, we calculated the removal rate across secondary treatment and chlorine disinfection individually for each target gene in absolute abundance (Figure 2.2) and relative abundance (Figure S2.3). In both winter and summer sampling

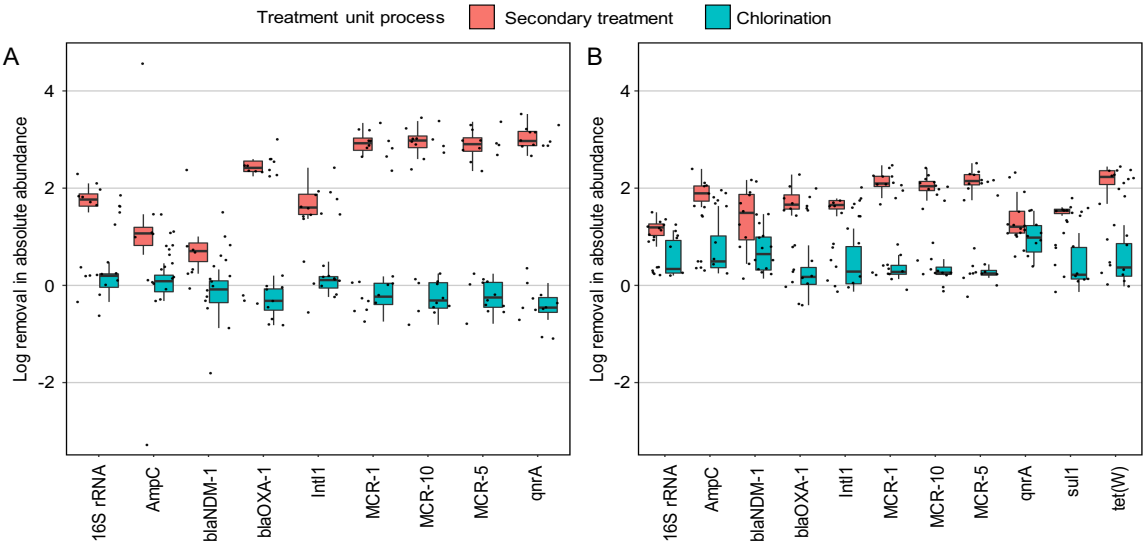


Figure 2.2 Log removal by the WWTP treatment unit process (n=12): secondary treatment and chlorine disinfection. a. Unit process removal of target genes calculated using absolute abundances during the winter sampling campaign. b. Unit process removal of target genes calculated using absolute abundances during the summer sampling campaign. In each panel, instantaneous removal rates as shown as black dots for the target genes. Secondary treatment removal rates are in red, and chlorine disinfection in blue. Boxes represent the interquartile range, with solid lines as medians. Whiskers represent the standard deviation.

campaigns, the instantaneous removal rates of target genes via secondary treatment were significantly higher than via chlorine disinfection ($p<0.0001$). This finding is in contrast to previous studies that reported that disinfection process can provide superior removal of ARGs as compared to secondary treatment alone (X. Cheng et al., 2021; M. Pei et al., 2019; Quach-Cu et al., 2018). However, one study found that biological treatment resulted in greater ARG removal (0.76–1.94 log reduction) as compared to ultraviolet disinfection (less than 0.5-log reduction) (L. Yang et al., 2019). The mechanisms that enable ARGs to persist across each unit process (including secondary treatment using activated sludge process) is still not well understood (Barancheshme & Munir, 2018;

Uluseker et al., 2021). Removal of ARGs during secondary treatment in aerobic systems has been reported in previous studies (Munir et al., 2011; Thakali et al., 2020). Aerobic treatment systems are effective at reducing ARGs for several reasons. First, the high dissolved oxygen concentration causes stress to enteric bacterial hosts of ARGs, which are usually anaerobic or facultative anaerobic microorganisms. For example, a recent study demonstrated the reduction in relative abundance of ARG hosts that are anaerobic (e.g., members of the phyla Bacteroidetes and Firmicutes which are dominant in human gut microbiota) in aerobic sludge as compared to in raw influent wastewater (Dai et al., 2022). In addition, previous studies concluded that ARG removal by activated sludge process is achieved due to the separation of suspended solids in the settlers (H. Chen & Zhang, 2013b; Gao et al., 2012b). Applying coagulation and settlement was found to remove significant abundance of ARGs through removing the biosolids associated with ARGs (Li et al., 2017).

In addition, the diurnal variation of target gene removals was observed for both secondary treatment and chlorine disinfection (Figure 2.2). Many target genes rendered both positive and negative removal rates via secondary treatment or chlorination at different times of the day (Figure 2.2). In addition, in the summer sampling campaign, the standard deviation of the removal rates via secondary treatment was significantly smaller than the standard deviation of the removal rates via chlorination ($p<0.05$), although in the winter sampling campaign, no significant difference was observed. This result showed that in summer, the target gene removals via secondary treatment presented fewer diurnal variations than the target gene removals via chlorination. Overall, secondary treatment contributed significantly to the removal of the vast majority of target genes as compared

to chlorine disinfection (Figure 2.2, $p<0.0001$) which on average showed limited (less than one-log) or even negative removal. There were several notable outliers that did not conform to the general removal patterns observed. For example, in the winter sampling campaign, the removal rates of the relative abundance of *AmpC*, *IntI1*, and *bla*NDM-1 were not significantly higher via secondary treatment as compared to chlorine disinfection ($p>0.05$, Figure S2.3). In the summer sampling campaign, the removal rates of the absolute abundance of *qnrA* and the relative abundance of *qnrA* and *bla*NDM-1 showed no significant difference across unit processes ($p>0.05$, Figure S2.3), either.

2.3.3 Removal of ARGs across the WWTP based on 24-hour loads

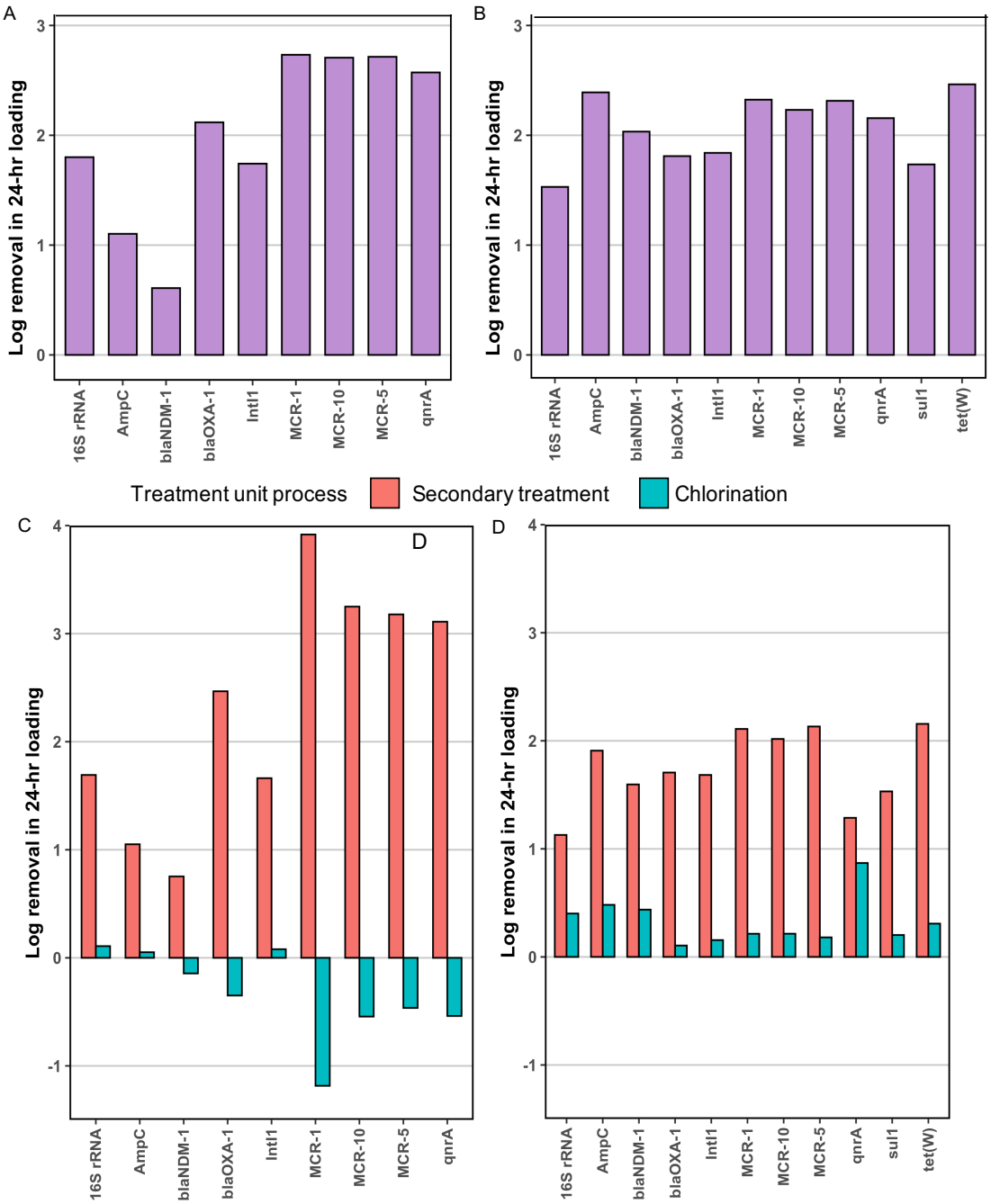


Figure 2.3 Removal of ARGs across the WWTP based on 24-hour loads. The removal rate of each target gene (x-axis) based on 24-hour loads across the WWTP in winter a. and in summer b. The removal rate of the 24-hour loading of each target gene (x-axis) via secondary treatment and chlorination, respectively (c: winter, d: summer). Secondary treatment accounted for the majority of the removal observed across the treatment system

Finally, we calculated the removal of each target gene across the treatment processes during the 24-hour sampling period based on gene loads, which accounted for the plant flow rate (Figure 2.3). The load of ARGs in the treated effluent may be more informative as compared to the concentrations of ARGs to indicate ARG dissemination into effluent-receiving environments, because it represents the actual abundance of ARGs released from a WWTP over a specified period of time (Cacace et al., 2019; Schwermer & Uhl, 2021; P. Wang et al., 2019). We found that all target genes were removed in terms of 24-hour loads (Figure 2.3a & b), but the degree of removal varied for different targets and across different unit processes (Figure 2.3c & d). There were still significant loads of ARGs and *IntI1* present in the final effluent, which is consistent with previous studies (Ju et al., 2019; N. Li et al., 2017; Munir et al., 2011; K. Qin et al., 2020). This finding further highlights the need to elucidate the potential of effluent ARGs to disseminate and transfer in the downstream receiving environments. Notably, the load of certain clinically-relevant ARGs, including *bla*NDM-1, *bla*OXA-1, MCR variants, and qnrA increased as a result of chlorine disinfection during the winter field campaign (Figure 2.3c & d). This finding, consistent with conclusions reported by other studies in which the ARG selection was observed by chlorination (Jin et al., 2020; S.-S. Liu et al., 2018), underscored the need to reduce the loads of ARGs before they reach the chlorine disinfection process.

The relatively weak degree of ARG removal from the disinfection process and the residual abundance of ARGs in the final effluent further emphasize the need for strategies to attenuate the dissemination of effluent ARGs. One modeling study recommended limiting the amount of treated WWTP effluents released in receiving waters to ensure a

minimum dilution rate to control the propagation of ARGs in the receiving waterbody (Corno et al., 2019). Other strategies, such as additional low-cost and effective treatment methods could be applied to remove residues in the treated effluents including ARGs, such as using sewage-sludge biochar and iron-oxide-coated sands (Calderón-Franco, Apoorva, et al., 2021) and electrocoagulation (L. Chen et al., 2020).

2.4. Conclusion

In this study, we quantified the diurnal variation of a suite of ARGs and one MGE across a WWTP. We found that the concentrations of target genes varied across the day, resulting in inconsistent and unrepresentative calculated removal rates based on grab samples taken at different times of day. We also found that the time when most target genes achieved their highest instantaneous removal rate coincided with the time when the lowest flow rate was observed. The results indicate that unintended bias could occur when using grab samples, and caution making conclusions about the removal of ARGs and MGEs based on grab samples. The diurnal variation in target gene concentrations was magnified by the wastewater treatment processes as reflected by the increasing RSD of target gene concentrations in the secondary effluent and final effluent as compared to the influent. Overall, secondary treatment contributed significantly to the removal of target genes whereas chlorine disinfection showed limited or even negative removal especially during the winter sampling campaign. We observed that chlorine disinfection resulted in increased loads of certain clinically-relevant ARGs. To reduce the load of ARGs discharged to the environment via the final effluent, efforts should focus on reducing ARGs before they

reach chlorine disinfection or using different disinfection methods that are more effective at inactivating ARGs.

2.5 Acknowledgments

E.L. and L.B.S. were supported in part by the National Science Foundation (CBET 2029025), and seed funds from Rice University. P.K. and L.B.S. were supported in part by a Johnson & Johnson WiSTEM 2D award. We thank the operators (Paul King, Chad Smith, and Mark Wahlstrom) at the West University City Place WWTP for helping with accessing wastewater samples.

Chapter 3 The fate of cell-associated and cell-free antibiotic resistance genes in the effluent of an anaerobic membrane bioreactor co-treating domestic wastewater and cattle manure

Esther G. Lou, Moustapha Harb, Adam L. Smith and Lauren B. Stadler

Abstract Anaerobic membrane bioreactors (AnMBRs) can manage complex combined waste streams, recover energy, and produce nutrient-rich effluents for irrigation. To advance AnMBRs for water reuse, the removal of antibiotic resistance genes (ARGs) during co-treatment of waste streams requires further attention. In addition, the effluent ARGs in the form of cell-free DNA that are discharged by wastewater treatment systems was recently found to be a neglected but important source of ARGs in the environment. Here, an AnMBR was fed domestic wastewater with increasing amounts of cattle manure. We assessed 1) the removal of target genes including nine ARGs and two mobile genetic elements (MGEs), and 2) the fate of cell-associated ARGs (caARGs) and cell-free ARGs (cfARGs) in the effluent. Manure addition significantly improved the removal of target genes. Further, in the effluent, the concentration of caARGs decreased steadily with increased manure loading whereas cfARGs were predominant when manure loading was greatest. Results imply that to advance AnMBRs to treat waste streams that contain ARGs, AnMBRs likely need to be paired with a downstream disinfection process in agricultural reuse applications to reduce cfARGs and caARGs in the effluent. In addition, our results show that application of AnMBRs for the co-treatment of domestic wastewater and livestock manure could reduce the proliferation risk potential during reuse, as they generate an effluent with relatively low ARG concentrations where cfARGs (which are easier to inactivate than caARGs via disinfection) make up a substantial fraction of the total effluent ARGs and MGEs evaluated in this study.

3.1. Introduction

Seventy percent of the world's freshwater supply is currently used for agricultural irrigation (AQUASTAT, 2014) which is driving the adoption of nontraditional water sources to ensure the long-term supply of water for crop production. Treated wastewater represents a reliable source of water that can offset the rate of freshwater depletion while concurrently reducing nutrient discharges into freshwater systems and redirecting those nutrients to offset fertilizer requirements. Anaerobic membrane bioreactors (AnMBRs) are an emerging biotechnology that are ideally suited to co-manage manure and wastewater because they produce a high-quality, nutrient-rich effluent for reclaimed water irrigation while also recovering energy in the form of biogas.

However, irrigation using reclaimed wastewater may lead to the dissemination of antibiotic resistance (X. M. Han et al., 2016; Q. Qin et al., 2015a), which has become one of the primary obstacles to reuse of treated wastewater for agricultural irrigation. Antibiotic resistance has been deemed one of the most significant human health challenges of the 21st century (Prestinaci et al., 2015). Wastewater treatment plants (WWTPs) represent a hotspot for the dissemination of ARGs in the environment (Rizzo et al., 2013a). Even after treatment, the elevated concentrations of nutrients, salts, disinfectants, disinfection byproducts, and lytic phages in wastewater effluents can contribute to cell competency and transformation, which can further promote the uptake of ARGs by microbial communities in receiving environments (Keen et al., 2017). Therefore, a comprehensive evaluation on antibiotic resistance, namely, antibiotic resistance genes (ARGs) and antibiotic resistant bacteria (ARB) in the treated effluent is crucial to validate the application of any wastewater reuse system.

To ensure the comprehensiveness of the evaluation on the treated effluent in terms of antibiotic resistance, it is necessary to investigate the fate of not only cell-associated ARGs (caARGs), but also cell-free ARGs (cfARGs, i.e., extracellular ARGs that are present due to lysis of dead cells or the secretion of DNA from live cells). As pointed out by a recent publication, a substantial body of previous analyses only focused on caARGs, while cfARGs are largely unexplored albeit highly abundant in the treated effluent discharged by WWTPs (Y. Zhang et al., 2018b). This issue is originated from one limitation of a commonly used methodology to reduce wastewater sample volumes for DNA extraction and analysis: filtration using 0.22 μm filters, because caARGs can be intercepted by the filters thus concentrated while cfARGs can pass through the filters and usually be discarded with the flow-through. In fact, cfARGs in the treated effluents are of particular importance. Several previous studies demonstrated that cfARGs were proven to be highly persistent in the downstream environment, especially when they are adsorbed into soil, sediments, clay minerals and humic substances (Dong et al., 2019; Mao et al., 2014; Nagler et al., 2018). Those cfARGs were found to contain clinically relevant ARGs, such as those encoding resistance to carbapenem, extended-spectrum beta-lactam and fluoroquinolone (Oliveira et al., 2020; Sivalingam et al., 2020). Furthermore, both microcosm (Dong et al., 2019; Mao et al., 2014) and *in situ* (Kittredge et al., 2022) experiments have shown that cfARGs released from dead bacteria can undergo natural transformation and be uptake by the live bacteria. In addition, approximately 65% of the extracellular DNA content of a wastewater metagenome consisted of mobile genetic elements (MGEs), which can directly promote antibiotic resistance development (Calderón-Franco, van Loosdrecht, et al., 2021). Therefore, the acquisition of clinically

significant cfARGs and MGEs by potentially pathogenic bacteria in the environment could transform them into resistant pathogens and may pose a potential threat to human health.

AnMBRs have been successfully applied for both low-strength domestic wastewaters and high-strength organic waste streams (e.g., food waste and manure), illustrating their versatility in managing organic waste streams (Amha et al., 2019; Smith et al., 2013). Furthermore, AnMBR systems in or proximal to agricultural areas could be used in decentralized applications to increase energy recovery while also managing livestock waste by co-treating wastewater with animal manure. Co-management of domestic wastewater with higher strength waste streams, such as animal manure, increases the potential for net energy positive treatment using AnMBR (Smith et al., 2014). In addition, AnMBRs may be uniquely suited to reduce antibiotic resistance proliferation during the co-treatment of wastewater and manure because membrane systems, in general, have been shown to achieve better ARG removal than conventional activated sludge systems (Harb & Hong, 2017a). A lingering question related to ARG removal by AnMBRs, however, is the extent to which cell-free ARGs (cfARGs; i.e. extracellular ARGs that are present due to lysis of dead cells or the secretion of DNA from live cells) are capable of passing through the membranes of these systems (H. Cheng & Hong, 2017a).

To advance the application of AnMBRs for the co-management of wastewater and manure for water reuse applications, we need a better understanding of their ability to remove ARGs and mobile genetic elements (MGEs) during treatment. The goal of this study was to (1) determine if there is an association between the amount of cattle manure added to domestic wastewater and the corresponding removal efficiencies of target genes in

AnMBR, and (2) investigate how manure addition impacted the profile of ARGs and MGEs and their dominant form (i.e., cell-associated vs. cell-free) in AnMBR effluent. Note that to comply with the Environmental Microbiology Minimum Information (EMMI) Guidelines (Borchardt et al., 2021), we spiked all samples with internal DNA standards in both forms of cell-free and cell-associated to account for the recovery of caARGs and cfARGs during sample concentration and DNA extraction. This is the first study to characterize ARG removal during AnMBR co-treatment of real wastewater and manure, and the first study to differentially characterize effluent cell-associated and cell-free ARGs and MGEs from an AnMBR treating real waste streams.

3.2. Materials and Methods

3.2.1 AnMBR set-up and monitoring in stages with increasing manure addition

A bench-scale AnMBR with a liquid volume of 4.5 L (Chemglass Life Science, Vineland, NJ) was operated at ambient temperature (average 17.7 ± 0.39 °C). The pH of reactor mixed liquor was checked regularly, and the observed values ranged from 7.2 to 7.6). The reactor was equipped with three submerged membrane housings that each contained a flat-sheet silicon carbide ultrafiltration membrane (Cembrane, Denmark) with 0.1 μm pore size and 0.015 m^2 effective surface area. Headspace biogas was recirculated via transversally mounted sparging tubes to limit membrane fouling while operating at a sub-critical flux of 5.27 $L/m^2/h$ (LMH), which yielded a hydraulic retention time (HRT) of approximately 19 hrs. The AnMBR was inoculated with anaerobic digester sludge from the Joint Water Pollution Control Plant (Carson, CA). The detailed configuration of the AnMBR and operational parameters are shown in Chapter 3 Appendix Section 1.1.1.

The operation of the AnMBR consisted of 5 stages: Baseline operation, Stage 1, Stage 2, Stage 3, and Stage 4. Domestic wastewater used in this study was collected from the City of West University Place WWTP (Houston, TX), a full-scale activated sludge WWTP that treats an average of 2 million gallons per day. The wastewater samples were collected periodically under dry weather conditions and at the same time of day to avoid diurnal variations. The wastewater was immediately transported to the lab and stored at 4 °C before feeding preparations. In Stages 1 through 4, manure slurry was added to the domestic wastewater in increasing amounts. Manure from beef cattle was collected from McGregor Research Center (McGregor, TX) and then diluted with water to form a manure slurry with a target solids content of 3%. Chapter 3 Appendix Table S3.1 provides detailed influent composition along with the corresponding organic loading rate (OLR) for each stage. Across the different stages of treatment, the OLR increased from 0.6 kg COD/m^3/d in Baseline operation to 2.5 kg COD/m^3/d in Stage 4 operation.

Performance indicators including chemical oxygen demand (COD), soluble COD (sCOD), mixed liquor total and volatile suspended solids (MLSS/MLVSS), pH, volatile fatty acids (VFAs; including acetic acid, formic acid, propionic acid, butyric acid, and valeric acid) and headspace biogas were monitored and described in Chapter 3 Appendix Section 1.1.2. In addition, chemical cleaning using 0.5% sodium hypochlorite was performed at the end of each operational stage to avoid the potential impact of membrane fouling on the experimental results, and to maintain a consistent flux across all operational stages. In addition, effluent tubing was cleaned periodically to remove downstream tube wall biofilms that may have formed.

3.2.2 DNA extraction with internal standards

Cell-associated DNA (caDNA) and cell-free DNA (cfDNA) were separated to quantify cell-associated ARGs (caARGs) and cell-free ARGs (cfARGs), respectively. Internal standards were spiked into all samples prior to DNA extraction. caDNA internal standards were E. coli cells containing a modified engineered plasmid (Chapter 3 Appendix Section 1.2). The plasmid, pReporter_8 (RRID: Addgene_60568; Yang et al., 2014), is a low-copy plasmid that was previously modified by knocking out the gene encoding green fluorescence reporter (GFP) and inserted with the methyl-halide transferase (MHT) gene found in Batis Maritima (H. Y. Cheng et al., 2016). A 112 bp region on MHT gene (primers listed in Chapter 2 Appendix table S4) was selected as the target for qPCR to quantify the initial concentration of standard spiked into the samples (C_o in Equation 1, see below) and the amount recovered in each corresponding sample (C_s in Equation 1, see below). E. coli cells were harvested from overnight culture supplemented with 34 μg/mL chloramphenicol, mixed well and aliquoted into equal volumes for: (1) spiking in the influent and effluent samples, and (2) conducing plasmid extraction on three of the aliquots to get the reference copy number through qPCR (C_o in Equation 1).

cfDNA internal standards were pUC19 plasmids containing a target insertion for qPCR quantification (Chapter 3 Appendix Section 1.2). The insertion was a 183 bp fragment of the ARHGAP11B gene, a human-associated gene that is specific to the brain neocortex (Florio et al., 2016) and was not found to be present in any sample collected in this study (data not shown). Approximately 10^8 copies of plasmids were spiked into each effluent sample prior to filtration (for sample concentration). The initial concentration of spiked cfDNA (C_o in Equation 1) was quantified by performing qPCR on the plasmid stock.

The concentration of recovered internal standard was quantified using qPCR to determine C_s. caDNA and cfDNA standards did not amplify in non-spiked influent, manure, or effluent samples in 40-cycles of qPCR (data not shown).

Influent and effluent samples were passed through membrane filters (mixed cellulose ester, 0.22 μm pore size, Millipore Sigma, MA) and caDNA was defined as DNA extracted from biomass retained on the filters, whereas cfDNA was defined as DNA in the filtrate. Five influent (30 mL each) and five effluent (350 mL each) samples were collected during each operational stage. caDNA was quantified in influent and effluent samples, and cfDNA was quantified in effluent samples only because the influent contained significantly higher cell counts than the effluent indicated by the concentration of rpoB (50 to 10^4-fold higher concentrations in the influent versus effluent). Right before the sample filtration step, both caDNA and cfDNA internal standards were spiked into samples and mixed well.

Membrane filters with caDNA were stored in 50% ethanol at -20 °C until DNA extraction. DNA extraction was performed using FastDNA SPIN Kits for Soil (MP Biomedicals, CA). cfDNA was analyzed in effluent samples by successively collecting the filtrate from the 0.22 μm filtration step and concentrating the DNA in the filtrate using an NAAPs-based method as previously described (D. N. Wang et al., 2016). cfDNA samples were stored at -20 °C until DNA extraction using FastDNA SPIN Kits for Soil (MP Biomedicals, CA). DNA was quantified in caDNA and cfDNA extracts using Qubit 3.0 coupled with dsDNA HS Assay Kit (Invitrogen, CA).

3.2.3 Quantification of ARGs and MGEs

Target genes for qPCR quantification included 9 ARGs (*sul1*, *sul2*, *tet*(O), *tet*(W), *ermB*, *ermF*, *ampC*, *bla*OXA-1 and *bla*NDM-1), 2 MGEs (*IntI1* and *tp614*) and *rpoB* (coding for β-subunit of RNA polymerase) used for normalizing ARGs and MGEs to calculate relative abundance. These ARGs were selected because they are frequently detected in wastewater. In addition, we specifically included, erm because the manure used in this study was collected from beef cattle that were fed with Tylosin, a macrolide-class antibiotic that may have resulted in selective pressure for *ermB* and *ermF* (B. Li, Yang, et al., 2015; Sui et al., 2016). A class 1 integron gene (*IntI1*) and a transposon gene (*tp614*) were also included because they have been found to be associated with the transfer of ARGs and to play an important role in the evolution and proliferation of multiple antibiotic-resistance bacteria (Di Cesare et al., 2016b; Gaze et al., 2011; Stokes & Gillings, 2011). The qPCR reaction was carried out in triplicate with each reaction containing 10.5 µL that included Forget-Me-Not EvaGreen qPCR mastermix (Biotium, CA), 50 nM ROX (Biotium, CA), 500 nM of forward and reverse primers, PCR grade H_2O, and DNA template. Primers and qPCR reaction conditions are provided in Chapter 3 Appendix Table S4 and S5. Ten-fold serial dilutions of cloned plasmids with each target gene were amplified in triplicate for each qPCR assay. The efficiencies of the real-time qPCR assays for the targeted ARGs and MGEs ranged from 89.10% to 102.65%. R^2 values were greater than 0.99 for all qPCR assays. The concentrations of the lowest standard for the target genes ranged from 6 to 920 copies/reaction, which was equivalent to 19 to 3066 copies/mL for 30 mL influent samples and 2 to 263 copies/mL for 350 mL effluent samples. Quality control steps for qPCR are detailed in Chapter 3 Appendix Section 2.3.

The concentration of target genes in each sample (gene copy number/mL) was calculated using the following equation (1):

$$\text{Gene concentration } \left(\frac{\text{copies}}{\text{mL}}\right) = \frac{C_s C_i}{V_s C_o} \quad (1)$$

Where C_s is the copy number of the target gene in the sample's DNA extract (copies) determined by qPCR, V_s is the volume of sample used (mL) to generate the DNA extract, C_i is the copy number of the internal standard (copies) determined by qPCR, and C_o is the copy number of the internal standard (copies) spiked into the sample prior to DNA extraction. Further information on internal standards of cell-associated genes and cell-free genes along with the calibration methods can be found in Chapter 3 Appendix Section 1.2. The recoveries of caDNA and cfDNA are provided in Chapter 3 Appendix Table S7.

3.2.4 Statistical methods

SciPy (https://www.scipy.org) was used for t-test and correlation analysis. The two-tailed unpaired t-test was used to identify significant differences between effluent target gene concentrations between stages (n=5 for cell-associated genes; n=5 for cell-free genes). The log removal values (LRVs) reported in this study for each gene for each operational stage were calculated using the following equation: LRV = log10(influent gene concentration/ effluent gene concentration). The influent and effluent gene concentrations were calculated by taking the average of n=5 samples. The effluent gene concentrations were the sum of cell-associated and cell-free fractions. A t-test was performed to assess whether a given LRV was significantly different between operational stages. Before performing a t-test, the Kolmogorov–Smirnov and Shapiro–Wilk tests were used to ensure

that the dataset followed a normal distribution (α=0.01). P values less than 0.01 were regarded as statistically significant. Pearson's correlation analysis was used for identifying correlations between any pair of two target genes in effluent samples over a 95% confidence interval. Pearson coefficient (r) was used to identify strength of correlations.

3.3. Results and Discussion

3.3.1 Percent COD removal was consistent across all operational stages

Percent COD removal was measured across all operational phases to assess the impact of manure addition on AnMBR performance and effluent water quality (Chapter 3 Appendix Section 2.1).

3.3.2 Wastewater contributed the majority of influent ARGs during manure co-treatment

Although manure contributed the majority of influent COD in Stages 1 through 4 (Figure 3.1a, $p<0.001$), the majority of influent ARGs and MGEs across those same stages remained dominated by the domestic wastewater (Stage 1 shown in Figure 3.1b, $p<0.05$). The most abundant targeted ARGs and MGEs in the wastewater were *intI1*, *sul1* and *sul2*, which was consistent with previous studies (Munir et al., 2010); (Gao et al., 2012a). While approximately 97% of the target ARGs and MGEs in the influent was contributed by the wastewater fraction during Stage 1, erythromycin ribosome methylation genes (*ermB* and

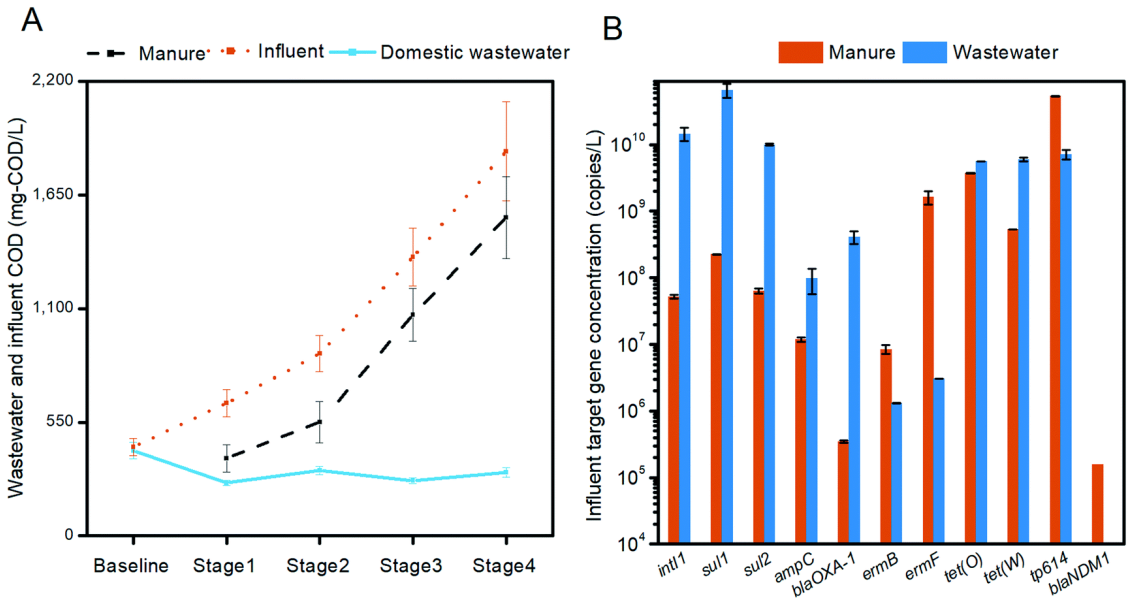


Figure 3.1. AnMBR influent composition: a. COD of wastewater and manure fractions of influent, and total influent cod across all operational stages. Error bars represent the standard deviation of biological replicates within each stage (n = 5); b. concentrations of ARGs and MGEs during a stage 1 loading (influent consisted of approximately 125 g of manure added to 20 L of domestic wastewater). Separate quantification of target genes in wastewater and manure was only performed during stage 1. In all other stages, target genes were quantified in the influent after combining the wastewater and manure.

ermF) that confer resistance to macrolide antibiotics were mainly contributed by the manure fraction (Figure 3.1b). This result is consistent with the fact that the manure was collected from cattle that were fed Tylosin, a macrolide antibiotic. Tylosin is a commonly used in-feed antibiotic in both cattle and swine livestock farms, and previous studies have observed high occurrences of *ermB* and *ermF* in livestock wastes (Sui et al., 2016); (Q. Li et al., 2015). Interestingly, the multi-drug resistance gene, *bla*NDM-1 was below the limit of detection in the influent and effluent samples during Baseline operation when the AnMBR was treating solely domestic wastewater (number of samples = 5) but was detected at concentrations above 10^2 copies/mL when manure was added to the influent (Figure 3.1b). No studies to our knowledge have specifically investigated the presence of the

*bla*NDM-1 gene in livestock manure; however, one study detected two *bla*NDM-1-postitive bacteria strains in the soil around animal farms (B. Wang & Sun, 2015).

3.3.3 The overall removal efficiency of targeted ARGs and MGEs from domestic wastewater was comparable to or greater than reported removal efficiencies of conventional wastewater treatment

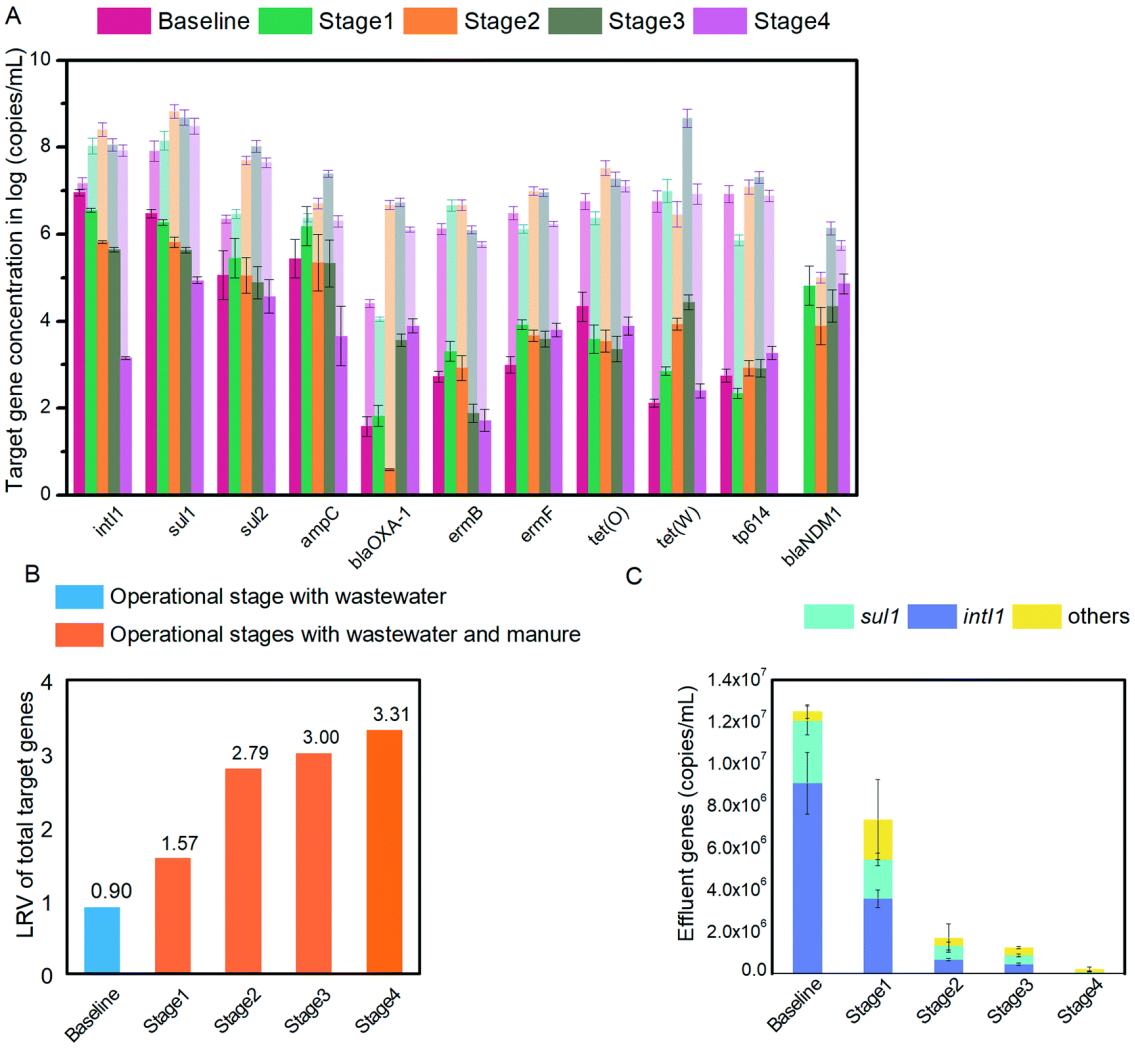


Figure 3.2. ARG and MGE removal across operational stages: a. average influent (shaded bars) and effluent (solid bars) concentrations of target genes (n = 5 for influent and effluent samples); b. log removal values (LRVs) of total target genes across operational stages; c. effluent gene concentrations (sum of cell-associated and cell-free fractions; n = 5). Error bars represent the standard deviations of gene concentrations within each operational stage.

Influent (cell-associated) and effluent (cell-associated and cell-free) ARGs and MGEs were quantified across all operational stages to calculate removal efficiencies. The log removal values (LRVs) of the targeted ARGs and MGEs ranged between 0.20 to 4.13 during Baseline operation when the AnMBR was solely treating domestic wastewater, with significant differences across stages (discussed in Section 3.3.4) and genes (Figure 3.2a). During Baseline operation, 87.4% (0.90 log) of targeted ARGs and MGEs were removed (Figure 3.2b). This removal efficiency is comparable to a study of two full-scale WWTPs, which reported a 89.0% – 99.8% removal of targeted ARGs (Mao et al., 2015a). High LRVs of *tet* genes, *tet*(O) (2.41) and *tet*(W) (4.63) reported here are consistent with previous findings of *tet* gene log reductions in membrane bioreactors treating domestic wastewater (Kappell, Kimbell, Seib, Carey, Choi, Kalayil, Fujimoto, et al., 2018a); (Munir et al., 2010). LRVs of *sul* genes, *sul1* (1.43) and *sul2* (1.28), were comparable to LRVs reported in previous studies on WWTPs, which ranged from 1.2 - 2.7 logs (Munir et al., 2010); (H. Chen & Zhang, 2013a); (Y. Zhang et al., 2018a). In addition, we observed high removal efficiencies of erm genes. The LRVs of *ermB* and *ermF* were 3.39 and 3.48, respectively, and they were both ~1.0 log higher than the LRVs reported from conventional WWTPs (Y. Zhang et al., 2018a); (J. Wang, Mao, et al., 2015). Although erm genes in influent increased across stages due to manure addition, the ultimate concentration in effluent was still less than 100 copies/mL which is lower than typically seen in biological effluent in conventional WWTPs or even in the final disinfected effluents (Y. Zhang et al., 2018a); (J. Lee et al., 2017); (Di Cesare et al., 2016a). All ARGs and MGEs were successfully reduced during baseline operation at over 87% removal, with the exception of *ampC* which actually increased in effluent samples. The enrichment of *ampC* may have

been due to the growth of organisms harboring this gene in the bioreactor. Enrichment of certain target ARGs in terms of their relative abundance (ARG copy number normalized by *rpoB* copies) was also observed for *sul1* and *bla*NDM-1, despite the fact that their absolute abundance decreased during the AnMBR treatment. Other studies have observed enrichment of different ARGs across biological treatment compartments in WWTPs in terms of both absolute abundance (Y. Yang et al., 2014a); (Rafraf et al., 2016a) and in relative abundance (Munir et al., 2010); (McConnell et al., 2018a). The inconsistent patterns of ARG removal/enrichment during wastewater treatment underscore the challenges of predicting the fate of ARGs released to the environment and need to develop a more mechanistic understanding of the factors that control ARG proliferation and attenuation during treatment.

To our knowledge, there is only one study on the fate of ARGs in AnMBRs treating real domestic wastewater (Kappell, Kimbell, Seib, Carey, Choi, Kalayil, & Fujimoto, 2018). In this study, a higher LRV of targeted ARGs (3.3 to 3.6 log) was found in the AnMBR treating primary clarifier effluent, but the mechanisms behind such high ARG removal efficiency were not well understood. This study, in combination with the limited number of previous studies on ARG removal during AnMBR treatment on real and synthetic wastewater (Kappell, Kimbell, Seib, Carey, Choi, Kalayil, & Fujimoto, 2018); (Zarei-baygi et al., 2019), suggests that AnMBRs are effective at removing a large suite of diverse ARGs present in domestic wastewater. This could be an important advantage over conventional wastewater treatment for improving microbial safety during agricultural reuse, especially considering that AnMBRs have also been shown to also surpass pathogen removal rates observed for full-scale aerobic MBRs (Harb & Hong, 2017b).

3.3.4 ARG and MGE removal efficiency increased with increased manure loading

This study is the first to examine the impact of co-treatment of domestic wastewater and manure on ARG removal. Results generally showed that the addition of manure was beneficial to overall targeted ARG and MGE removal rates, which was strongly supported by the consistent trend of decreasing target gene concentration in the effluent through Baseline operation to Stage 4 (Figure 3.2c). Further, as the fraction of manure added to the influent wastewater was incrementally increased, the removal efficiency of overall ARGs and MGEs also increased steadily from Baseline operation to Stage 4 (Figure 3.1b). In addition, the overall ARG and MGE removal rate was largely driven by the removal of *sul1*, *intI1* and *sul2* genes (Chapter 3 Appendix Figure S3.2). The LRVs of *intI1*, *sul1* and *ampC* consistently increased from Baseline operation to Stage 4 due to manure addition ($p<0.01$). In Stage 4, the LRV of all targeted ARGs and MGEs reached 3.31, which was mainly due to the highest LRV of the most abundant influent MGEs and ARGs, namely, *intI1* (4.77) and *sul1* (3.54) (Figure 3.2a). The removal efficiency of targeted ARGs and MGEs observed in Stage 4 was higher than previously reported values from several WWTP studies (J. Wang, Mao, et al., 2015); (Y. Yang et al., 2014a); (Wen et al., 2016). The LRVs of genes *bla*OXA-1, *ermB*, *ermF*, *tet*(O), *tet*(W) and *tp614* were consistently high across all stages (Figure 3.2a; Chapter 3 Appendix Table S3.8). Further, it was interestingly observed that *ermB*, *ermF* and *tp614* were significantly more abundant in the influent manure fraction as compared to the influent wastewater fraction ($p<0.01$, Figure 3.2b). The higher influent concentrations contributed by manure addition may have increased the AnMBR's potential to remove them when manure was added compared to

Baseline operation (when only the wastewater was treated). A previous study found that manure was dominated by cell-associated DNA as opposed to cell-free DNA (Y. Zhang et al., 2013). We observed that *ermB*, *ermF* and *tp614* were significantly more abundant in the manure than the wastewater fraction of the influent ($p<0.01$, Figure 3.2b). Thus, their superior removal when manure was added to the AnMBR may have been because the cell-associated DNA was readily removed via filtration.

The overall targeted ARG and MGE concentration in the effluent decreased consistently with the addition of manure (Figure 3.2c, $p<0.01$), and in Stage 4 effluent, this concentration was approximately 90% lower than that of Baseline operation when the AnMBR was treating domestic wastewater without manure. The removal efficiency of targeted ARGs and MGEs observed in Stage 4 was higher than the previously reported values of several WWTP studies (Wen et al., 2016); (J. Wang, Mao, et al., 2015); (Y. Yang et al., 2014a) and of manure treatment approaches including advanced anaerobic digestion (Wallace et al., 2018); (J. Wang, Ben, et al., 2015) and composting (Zhu et al., 2019). This result, combined with the fact that AnMBRs can recover energy in the form of biogas, underscore that AnMBR is a strong candidate for the co-management of wastewater and manure because they can potentially reduce the proliferation of ARGs in wastewater and animal waste. The results support the application of AnMBRs in decentralized agricultural applications where multiple waste streams must be managed and water and energy reuse could be harnessed.

The improvement in ARG removal with increasing manure loading may have resulted from enhanced microbial activity caused by the increasing OLR. The increase in microbial activity may have impacted ARG removal in several different ways. First,

manure addition could have resulted in a shift in the microbial community to fast-growers that could quickly break down the organics in the substrate. This, in turn, may have selected for microbes that harbor fewer ARGs since they can exert a metabolic burden (i.e. fitness cost) that can result in slower growth rates (Andersson & Levin, 1999). Second, the enhanced biological activity resulting from a greater input of nutrients to the system may have boosted growth generally and resulted in greater turnover of cells and biodegradation of DNA that included ARGs. We also observed a consistent shift in the effluent DNA from

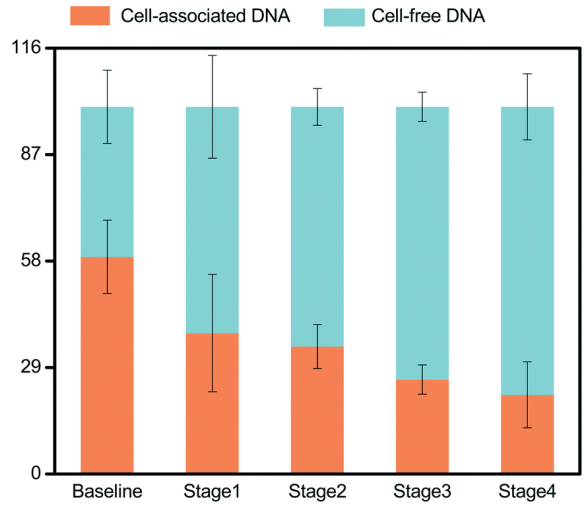


Figure 3.3. Effluent DNA composition: relative abundance (%) of cell-associated and cell-free DNA in the effluent in each operational stage. The ratios were calculated based on averaged concentrations of effluent caDNA and cfDNA within each stage (n = 5). Error bars represent the standard deviations of the corresponding DNA concentrations within each operational stage.

cell-associated to cell-free from Baseline operation to Stage 4 (Figure 3.3). This observation supports the assumption that manure addition resulted in an increase in biological activity in the bioreactor as a previous study (Nagler et al., 2018) in which it was found that the ratio of cfDNA to total DNA content was significantly positively associated with biological activity in anaerobic digesters. Further, the methane COD conversion normalized by the feeding COD (Chapter 3 Appendix Table S3.1) showed consistent increasing trend from Baseline operation to Stage 4, supporting the increasing bioactivity in accordance with manure addition.

3.3.5 The effluent ARG and MGE reservoir shifted from cell-associated to cell-free with increased manure loading

The majority of effluent ARGs and MGEs were cell-associated during all operational stages except for Stage 4 (Chapter 3 Appendix Figure S3.4). In Stage 4, the cell-free fraction of target genes was significantly elevated as compared to the previous stages (t-test, $p<0.001$) and accounted for approximately 89% of the targeted ARGs and MGEs in the effluent. At the same time, the concentration of the cell-associated fraction of targeted ARGs and MGEs in effluent decreased consistently from Baseline operation to Stage 4 ($p<0.01$). These results indicated that cfDNA became the primary reservoir of target genes in the effluent when the manure loading to the system was the highest. The abundance of effluent cell-free ARGs and MGEs confirms that they should not be overlooked in wastewater effluents (or unquantified because of the DNA concentration protocol used), as they can make up a substantial fraction of effluent ARGs (Q.B. Yuan et al., 2019a);(Y. Zhang et al., 2018a). A substantial body of previous studies on ARGs in wastewater did not explicitly capture the cell-free fraction of ARGs, and thus may have

significantly underestimated the risk of ARG propagation from effluents in receiving environments.

In the cell-associated fraction of effluent target genes, the abundance of *intI1* and *sul1* decreased significantly with the addition of manure across all stages ($p < 0.05$) (Figure 3.4a). The removal of intI1 may have been due to the elimination of manure-associated

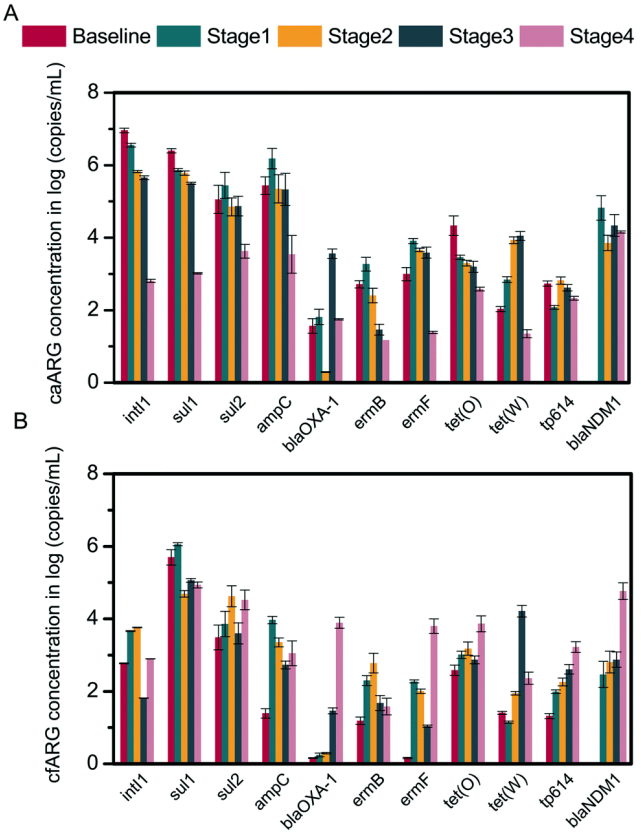


Figure 3.4. Effluent ARG and MGE profiles: concentrations of a. cell-associated and **b.** cell-free fractions of target genes across operational stages (n = 5). Error bars represent the standard deviations of the corresponding gene concentrations within each operational stage.

aerobic hosts of integrons (e.g., Actinomycetales and Bacilli) during anaerobic treatment (W. Sun et al., 2016). Overall, the cell-associated fraction of targeted ARGs and MGEs in the effluent decreased steadily from Baseline operation to Stage 4 (Chapter 3 Appendix Figure S3.4). Indeed, LRVs of the cell-associated fraction of targeted ARGs and MGEs across all five stages ranged from 1.54 - 4.20 logs and were consistent with reported ARG removal efficiencies using membrane-based treatment technologies (Munir et al., 2010);

(Le et al., 2018); (Sui et al., 2018). Given that ultrafiltration membranes (0.01 - 0.1 pore size) retain the vast majority of microbes in AnMBRs, it is likely that some of the caARGs in the effluent are due to microbial regrowth within post-membrane effluent lines. Considering the higher effluent nutrient concentrations during the stages with higher manure addition, it is also possible that the observed reduction of cell-associated target genes was due to effluent selection for microbial groups which are less likely to harbor ARGs and MGEs due to fitness cost. We also performed correlation analysis to identify significant associations between effluent ARGs that were observed in cell-associated and cell-free fractions and the results are discussed in the Chapter 3 Appendix Section 3.3.

3.3.6 Implications for advancing AnMBRs for wastewater reuse

Results of a previous quantitative microbial risk assessment (QMRA) on AnMBR effluents indicate that AnMBR treatment would likely need to be paired with a downstream disinfection process in agricultural reuse applications (Harb & Hong, 2017b). Based on this, the form of effluent ARGs (i.e., cell-associated vs. cell-free) may also influence their inactivation rates during disinfection (here, we define inactivation as the destruction of the ARG such that it is no longer functional). Specifically, a greater proportion of ARGs in the cfDNA fraction may improve ARG inactivation during disinfection. Cell-associated ARGs are more difficult to inactivate because the cell serves as a barrier between the disinfectant and DNA and can thus protect the DNA against damage. A recent study evaluated multiple disinfection methods including free chlorine, monochloramine, chlorine dioxide, ozone, UV, and hydroxyl radicals, and reported that among all these disinfection processes, caARG inactivation always lagged behind cell inactivation (Yoon et al., 2017). This

indicates that the removal of caARGs requires cell inactivation to occur first, whereas cfARGs may be inactivated directly during disinfection. In a few of the very limited amount of studies that distinguished between caARGs and cfARGs during disinfection processes, caARGs were found to be more difficult to remove than cfARGs during chlorination and UV disinfection (Yoon et al., 2017); (He et al., 2019); (McKinney & Pruden, 2012). Further, other studies observed that caARGs became cfARGs during the disinfection process, indicating some disinfection may not be sufficient to completely destroy caARGs (Q.-B. Yuan et al., 2019a); (S. S. Liu et al., 2018).

In this study, the concentration of effluent caARGs decreased steadily with increased manure loading (Chapter 3 Appendix Figure S3.4). In Stage 4, there were 2.34×10^4 copies/mL of caARGs in the effluent, which was much lower than secondary effluent caARG concentrations reported by previous studies of conventional WWTPs (Mao et al., 2015a; McConnell et al., 2018a; Mukherjee et al., 2021). Thus, our results show that application of AnMBRs for the co-treatment of domestic wastewater and livestock manure could reduce the proliferation risk potential during reuse, as they generate an effluent with relatively low ARG concentrations where cfARGs (which are easier to inactivate than caARGs) make up a substantial fraction of the total effluent ARGs and MGEs assessed in this study. However, the abundance of total ARGs in the effluent was still not negligible, particularly, the clinically relevant ARGs encoding resistance to ESBL- and carbapenemase-production (i.e., *bla*OXA-1 and *bla*NDM-1) were still of significant concentrations in both cell-associated and cell-free fractions (Figure 3.4). Therefore, consecutive treatment should be used to remove those ARGs in the AnMBR effluent. Nonetheless, the studies on further treatment of AnMBR effluents for water reuse are

scarce. A recent study has demonstrated that disinfection using UV/H_2O_2 effectively removed extracellular DNA with > 2-log reduction, and the remaining extracellular DNA were not valid to be transformed into competent hosts (Augsburger et al., 2021). In addition, UV/H_2O_2 treatment can effectively remove emerging contaminants and antibiotic resistance elements (ARB and ARGs), avoid formation of disinfection byproducts, and preserve high nutrient content in the AnMBR effluent for water reuse (Augsburger et al., 2021). However, the cost of UV/H_2O_2 application is usually high (Souza et al., 2013) and the energy consumption associated with UV radiation might make AnMBRs less appealing as a decentralized treatment technology. Future studies need to evaluate or develop other tertiary treatment technologies to achieve the goal of advancing AnMBRs' (co-)treatment with the specific tasks to 1) remove caARGs and cfARGs in the AnMBR effluent, 2) preserve nutrients for water reuse, and 3) minimize the required cost and footprint.

3.4. Conclusion

Our results demonstrate that co-management of domestic wastewater and livestock manure using AnMBRs can both improve resource recovery and mitigate the spread of antibiotic resistance in reclaimed water. The removal efficiency of total target genes significantly improved with the increased manure loading in the AnMBR co-treatment process and was greater than many conventional WWTP treatment processes. Increasing manure loading not only decreased total target gene abundance in the effluent but also made cfARGs the dominant form of effluent ARGs. cfARGs require uptake by competent cells to be functional and are easier to inactive during disinfection. Thus, the ability of AnMBR to reduce ARGs during co-treatment of wastewater and manure and generate an effluent with primarily cell-free ARGs (as opposed to cell-associated) may be

advantageous in water reuse applications. Future studies should focus on developing technologies for treating AnMBR effluents to close the application gap.

3.5. Acknowledgement

This work was supported by funding from Water for Agriculture grant no. 2016-68007-25044 from the USDA National Institute of Food and Agriculture. The research was partially supported by Rice University. The authors are grateful to Eric Rice, Laney Baker, Andrew Brower, and Sophie Gao for their help with reactor construction, maintenance, and sample analysis. The authors would also like to thank Paul King and his team of operation from the West University City Place WWTP for helping with logistics related to wastewater sampling.

Chapter 4 Direct Comparison of RT-ddPCR and Targeted Amplicon Sequencing for SARS-CoV-2 Mutation Monitoring in Wastewater

Esther G. Lou, Nicolae Sapoval, Camille McCall, Lauren Bauhs, Russell Carlson-Stadler, Prashant Kalvapalle, Yanlai Lai, Kyle Palmer, Ryker Penn, Whitney Rich, Madeline Wolken, Pamela Brown, Katherine B. Ensor, Loren Hopkins, Todd J. Treangen and Lauren B. Stadler

Abstract. Over the course of the COVID-19 pandemic, variants of SARS-CoV-2 have emerged that are more contagious and more likely to cause breakthrough infections. Targeted amplicon sequencing approach is a gold standard for identification and analysis of variants. However, when applied to environmental samples such as wastewater, it remains unclear how sensitive this method is for detecting variant-associated mutations in environmental samples. Here we directly compare a targeted amplicon sequencing approach (using ARTIC v3; hereafter referred to as sequencing) with RT-ddPCR quantification for the detection of five mutations that are characteristic of variants of concern (VoCs) in wastewater samples. In total, 547 wastewater samples were analyzed using both methods in parallel. When we observed positive mutation detections by RT-ddPCR, 42.6% of the detection events were missed by sequencing, due to negative detection or the limited read coverage at the mutation position. Further, when sequencing reported negative or depth-limited mutation detections, 26.7% of those events were instead positive detections by RT-ddPCR, highlighting the relatively poor sensitivity of sequencing. No or weak associations were observed between quantitative measurements of target mutations determined by RT-ddPCR and sequencing. These findings caution the use of quantitative measurements of SARS-CoV-2 variants in wastewater samples determined solely based on sequencing.

4.1. Introduction

Over the course of the COVID-19 pandemic, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has evolved and numerous lineages have emerged that are more transmissible, cause more severe disease, and/or are better at escaping the immune response system (Garcia-Beltran et al., 2021; Harvey et al., 2021; Q. Li et al., 2020). Tracking the emergence and spread of these variants of concern (VoCs) and variants of interest (VoIs) has become critical to public health response and mitigation strategies for stopping the spread of SARS-CoV-2. Wastewater-based epidemiology (WBE) is one prominent approach that has been adopted by public health departments and water utilities to track infection dynamics in communities by quantifying the amount of SARS-CoV-2 RNA in wastewater samples (Ahmed et al., 2020; Arora et al., 2020). WBE can also be used for monitoring VoCs and VoIs in communities (Bar-Or et al., 2021; Fontenele et al., 2021; Heijnen et al., 2021).

Variant identification in wastewater samples is challenging because the viral genomes are highly fragmented, dilute, and comprised of mixtures of circulating variants. The most common methods used for wastewater variant screening are: (1) quantifying specific characteristic mutations via RT-qPCR or RT-ddPCR (Ciesielski et al., 2021; Heijnen et al., 2021); and (2) enriching and sequencing SARS-CoV-2 genomes in wastewater (Swift et al., 2021; Tyson et al., 2020). Underlying both approaches is the ability to identify and quantify characteristic mutations that define variants. RT-qPCR is regarded as a gold standard method for routine wastewater surveillance (Alygizakis et al., 2021; Rahman et al., 2021; Van Poelvoorde et al., 2021). Compared to RT-qPCR, RT-ddPCR has a superior detection sensitivity (Ahmed et al., 2022; Ciesielski et al., 2021;

Flood et al., 2021) and is less sensitive to inhibitors present in wastewater (Cao et al., 2015; Ciesielski et al., 2021). However, PCR based methods are limited in that the mutations must be known ahead of time for primer and probe design. In addition, in practice, they are limited by the number of targets that can be multiplexed per reaction, and it is difficult to delineate all variants present in the sample by only targeting a few mutations. Next generation sequencing (NGS) enables comprehensive screening of all potential mutations without any prior knowledge, and thus has been frequently applied for characterizing pathogens and viruses (Greninger et al., 2015; J. Yang et al., 2011). The unbiased, non-targeted metagenomics sequencing approaches often require high- or ultra-high coverage in order to obtain enough target sequences (Chiara et al., 2021). On the other hand, targeted sequencing approaches using an enrichment step during library preparation maximize the detection of viruses effectively (X. Deng et al., 2020). For SARS-CoV-2 genome enrichment, multiplex tiling PCR and oligonucleotide capture are the most frequently implemented methods, both demonstrating great performance in terms of genome coverage (Doddapaneni et al., 2021; Tyson et al., 2020) and mutation detection (Bar-Or et al., 2021; Crits-Christoph et al., 2021b). Targeted amplicon sequencing (i.e., multiplex tiling PCR coupled with amplicon sequencing) is considered the lower-cost and faster approach (Chiara et al., 2021; X. Lin et al., 2021). For example, ARTIC Network panels (https://artic.network) are commonly used by laboratories globally to characterize SARS-CoV-2 present in clinical samples (Charre et al., 2020; J. Li et al., 2020; Mboowa et al., 2021). However, targeted amplicon sequencing is frequently limited by coverage and/or quality dropout due to amplification bias and primer knockout by mutations that happen to occur at the priming regions (Davis et al., 2021).

Several studies used amplicon-based sequencing of SARS-CoV-2 from wastewater samples to estimate the prevalence of SARS-CoV-2 variants in the communities (Bar-Or et al., 2021; Layton et al., 2021; Otero et al., 2021). One study used AF (allele frequency) of VoC-associated mutations detected in wastewater sample to estimate the relative abundances of different lineages circulating in a community (Ellmen et al., 2021). However, it is unclear how quantitative mutation AF are that are generated from wastewater genomes via this approach. Thus, a direct comparison between targeted amplicon sequencing and RT-ddPCR (or RT-qPCR) for mutation detection and quantification using wastewater samples is needed. In this study, we quantified five unique mutations using RT-ddPCR and performed targeted amplicon sequencing (ARTIC v3 based) of SARS-CoV-2 in parallel on 547 wastewater samples. We compare the consistency in the approaches in terms of (1) detection vs. no detection; and (2) quantitative information generated by each method. In addition, we evaluated the impact of mutation concentration, single base coverage at the mutation position, the overall SARS-CoV-2 concentrations, and SARS-CoV-2 genome coverage on mutation detection via targeted amplicon sequencing.

4.2. Materials and Methods

4.2.1 Wastewater sample collection, concentration, and RNA extraction

We collected weekly wastewater samples from 39 wastewater treatment plants (WWTPs) in Houston covering a service area of approximately 580 square miles and serving over 2.3 million people. Time-weighted composite samples of raw wastewater were collected every 1 hour for 24 hours from the influent of the WWTPs. The sampling campaign was conducted during two separate periods of time (Phase I and Phase II). Phase

I spanned from February 23, 2021 to April 12, 2021 when Alpha (specifically, B.1.1.7) was the dominant SARS-CoV-2 VoC circulating in Texas (GISAID, https://www.gisaid.org/). Phase II covered from May 24, 2021 to July 12, 2021 when the Delta variant became prevalent, displacing Alpha as reflected by variant confirmed cases in Texas (GISAID, https://www.gisaid.org/). In total, 249 and 298 samples were analyzed during Phase I and Phase II, respectively. SARS-CoV-2 was concentrated in wastewater samples using an electronegative filtration method as previously described (LaTurner et al., 2021a). RNA extraction was performed using a Chemagic™ Prime Viral DNA/RNA 300 Kit H96 (Chemagic, CMG-1433, PerkinElmer) with the PerkinElmer viral RNA/DNA purification protocol and reagents. Finally, 10 μL of sample extract was used for each RT-ddPCR reaction, and 11 μL of sample extract was used for sequencing library preparation. Detailed concentration procedures, RNA extraction procedures, concentration factors (Table S1), and associated quality control measurements are provided in the Supplementary materials following the Environmental Microbiology Minimum Information (EMMI) Guidelines (Borchardt et al., 2021).

4.2.2 RT-ddPCR quantification of SARS-CoV-2 N1, N2 genes, and five characteristic mutations

RT-ddPCR was performed on a QX200 AutoDG Droplet Digital PCR System (Bio-Rad) and a C1000 Thermal Cycler (Bio-Rad) in 96-well optical plates. SARS-CoV-2 N1 and N2 gene targets were quantified in wastewater samples as previously described (LaTurner et al., 2021). Five mutations, namely S:DEL69/70, S:N501Y, S:E484K, S:K417T, S:L452R, were quantified via RT-ddPCR. GT Molecular kits were used for RT-

ddPCR quantification of mutations (kit information provided in Table S2). S:DEL69/70 and S:N501Y, two characteristic mutations associated with the Alpha lineage B.1.1.7, were quantified during Phase I, and S:L452R, S:K417T and S:E484K were quantified during Phase II. The latter three mutations were selected due to their reducing SARS-CoV-2 susceptibility to convalescent and vaccine-elicited sera and mAbs, and their emergence in newly evolved SARS-CoV-2 strains (Jangra et al., 2021; Wilhelm et al., 2021). These SARS-CoV-2 mutations were quantified using one-step RT-ddPCR assays according to the manufacturer's protocol (GT Molecular). A detailed description of the methods, including droplet thresholding and limit of detection (LOD) are described in the Supplementary materials (Section 1.4, Table S3 – S5). For all targets (N1, N2, and the five mutations), positive detection (+) was defined as above the LOD, and a negative detection (-) was defined as below the LOD. RT-ddPCR analysis was used to generate: (1) the concentration of the mutation in copies/L-wastewater, and (2) the fraction of SARS-CoV-2 containing the mutation, which was calculated by normalizing the copies of the mutation by the sum of the copies of the mutation and the wild-type.

4.2.3 Amplicon-based sequencing using ARTIC v3 and data analysis

cDNA was generated using 11 μL RNA extract via reverse transcription using the Superscript IV first-strand synthesis system (ThermoFisher Scientific, 18091050) following the manufacturer's protocol. SARS-CoV-2 genome enrichment via multiplexing PCR was conducted using ARTIC v3 protocol (Tyson et al., 2020). Illumina DNA Prep kit with the manufacturer's manual (DNA Flex) were applied for amplicon tagmentation and flex amplification, followed by library clean-up. Each sample library was then quantitated,

normalized, pooled, and diluted to 6 pM. Finally, sequencing was performed on an Illumina MiSeq instrument using MiSeq Reagent Kit v2 (300-cycles, MS-103-2002) following a 151 + 10 + 10 + 151 cycling recipe. Sequencing read trimming with BBDuk (Bushnell, 2014) was conducted, followed by read mapping with BWA-MEM (H. Li, 2013). Read mapping result was sorted by samtools (H. Li et al., 2009), and primer locations were soft-clipped using iVar (Grubaugh et al., 2019). Finally, variant calls were performed with respect to the Wuhan reference sequence (NC_045512.2) using LoFreq (Wilm et al., 2012).

To compare targeted amplicon sequencing (hereafter referred to as sequencing) with RT-ddPCR, we first calculated the sequencing read coverage for each target mutation in each sample by averaging the single base coverage across the target region used for quantification in RT-ddPCR. For example, for the N1 and N2 RT-ddPCR assays, the CDC 2019-nCoV_N1 probe and 2019-nCoV_N2 probe were applied to RT-ddPCR assays in this study. These probes align to nt 28318 – 28332 and nt 29188 – 29210, respectively (the nt coordinates correspond to the Wuhan reference NC_045512.2). Accordingly, sequencing mapped reads for N1 and N2 were checked at each single base position from nt 28318 to 28332 (N1 region, containing 15 positions) and from nt 29188 to 29210 (N2 region, containing 23 positions), respectively. A positive detection (+) was called for N1 and N2 if (1) the single base coverage at each nt position across the target region (nt 28318 – 28332 for N1 and nt 29188 – 29210 for N2) was at least 1×, and (2) the average single base coverage across the target region was at least 20×; otherwise, a non-detect (ND) was called. For the five mutations, if there was at least 20× read depth at the position that corresponded to the target mutation, a positive detection (+) was called if any reads containing that mutation were observed. A negative detection for sequencing (-) was defined as no reads

containing the target mutation were observed, and there was at least 20× read depth at the mutation location. Finally, for sequencing if less than 20 reads mapped to the mutation position, we defined this as "depth limited (DL)". We used a 20× coverage threshold for sequencing analysis based on previous studies that applied tiled PCR and short-read sequencing (Illumina) for SARS-CoV-2 wastewater analysis (Baaijens et al., 2021; Fontenele et al., 2021).

4.2.4 Statistical Analysis

Welch two sample t-test was applied to compare datasets. Spearman rank correlation analysis was used to study the associations between quantitative results generated by RT-ddPCR and sequencing. For each target mutation, we used Spearman rank correlation to assess the correlations between (1) the mutation concentration as determined by RT-ddPCR (copies/L-wastewater) and the number of reads containing the mutation as determined by sequencing; and (2) the fraction of SARS-CoV-2 containing the target mutation (target mutation concentration/total SARS-CoV-2 concentration as determined by RT-ddPCR) and the AF of the mutation as determined by sequencing. Strength of correlations were identified based on Spearman's correlation coefficient Rho ®.

4.3. Results and Discussion

4.3.1 RT-ddPCR was more sensitive than sequencing for mutation detection

547 wastewater samples (249 samples for Phase I, 298 samples for Phase II) were analyzed using both RT-ddPCR (targeting N1, N2, and five mutations) and sequencing. The wastewater concentrations of SARS-CoV-2 (determined by the average of N1 and N2

concentrations) were significantly higher during Phase I than during Phase II (Figure S1, $p < 0.001$). For sequencing, an average fraction of 0.617 (std: 0.222) of the reads in Phase I, and an average fraction of 0.298 (std: 0.227) of the reads in Phase II mapped to the SARS-CoV-2 Wuhan reference. For the 547 wastewater samples, sequencing generated an average of 155,583 ± 249,909 reads (Phase I: 295,938 ± 325,644; Phase II: 46,784 ± 41,349) that mapped to the reference genome. The average single base coverage across the entire SARS-CoV-2 genome was 973 reads per base (range: 0.03 to 4444). The average breadth of coverage, defined as the percentage of genome bases sequenced per sample, was 66.7% (range: 1.4% to 99.9%; sequencing read statistics detailed in Table S6). Additional information on all samples analyzed, including their detections by RT-ddPCR and sequencing for each target are detailed in Figure S4.2.

To compare RT-ddPCR and sequencing, we categorized detection events into different scenarios. There were four possible scenarios for N1 and N2 detections: positive detections by both RT-ddPCR and sequencing (+/+), positive detection by RT-ddPCR and not detected by sequencing (+/ND), negative detection by RT-ddPCR and positive detection by sequencing (-/+), and negative detection by RT-ddPCR and not detected by sequencing (-/ND). For the five target mutations, there were six possible scenarios: positive detections by both RT-ddPCR and sequencing (+/+), positive detection by RT-ddPCR and negative detection by sequencing (+/-), positive detection by RT-ddPCR and depth limited (single base coverage $< 20\times$ at the mutation position) for sequencing (+/DL), negative detections by both methods (-/-), negative detection by RT-ddPCR and positive detection by sequencing (-/+), and negative detection by RT-ddPCR and depth limited by sequencing (-/DL). We first assessed the relationship between N1 and N2 detections and mutation

detections. Figure 4.1 shows each detection event (center bars) and their corresponding N1 and N2 detections. As expected, almost all samples with RT-ddPCR and/or sequencing positive detections for mutations [(+/+), (+/-), (+/DL); (-/+)] were also positive for N1 and N2 [mainly (+/+) and (+/ND)].

Among all 1094 N1 and N2 detection events in 547 samples, 67.6% had consistent detections for RT-ddPCR and sequencing [(+/+) and (-/ND)]. For RT-ddPCR, we observed very consistent N1 and N2 detections for the vast majority (97.1%) of samples, with N1

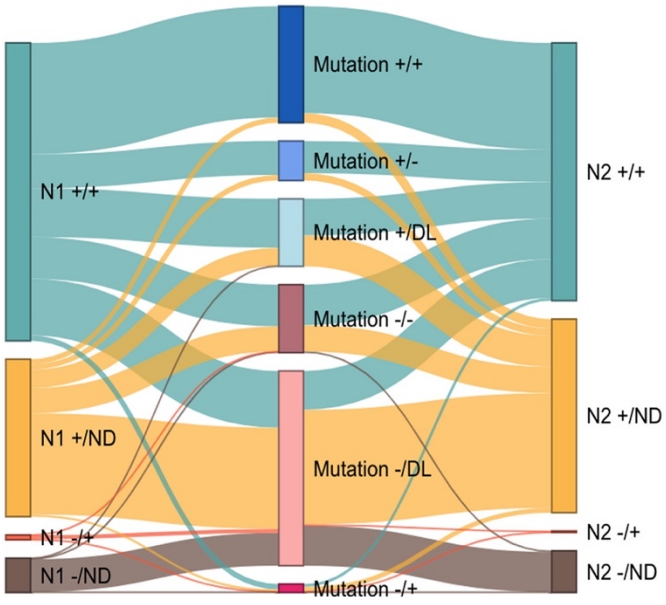


Figure 4.1. Relationship between mutation detection events and N1 and N2 detection events by RT-ddPCR and sequencing. The bars on the left and right group are based on N1 and N2 detection scenarios (format: RT-ddPCR detection/sequencing detection). The bars in the middle group are based on mutation detection scenarios. The height of each node corresponds to the number of detection events in the specific group. The width of the link between each pair of bars represents the number of the shared sample (s) belonging to both detection groups.

and N2 double positive and N1 and N2 double ND events for 531 of the 547 samples. In contrast, N1 and N2 detections via sequencing reported consistent detections for only 403 samples (73.7% of total samples). The inconsistency of sequencing in N1 and N2 detections was likely due to the significantly lower read depth at the N1 region than at the N2 region ($p<0.0001$). Overall, RT-ddPCR detected 499 N1 and N2 double positive events

whereas sequencing only detected 265. These results indicate that sequencing detection is less sensitive than RT-ddPCR detection when focusing on commonly targeted N-gene regions of the SARS-CoV-2 genome.

Unsurprisingly, we also found that sequencing was less sensitive than RT-ddPCR for detecting target mutations. We compared 1,354 possible mutation detections using both RT-ddPCR and sequencing (Table S4.2, Figure S4.2), in terms of their consistency in calling the presence or absence of mutations (Figure 4.2). Among all 1,354 detection events,

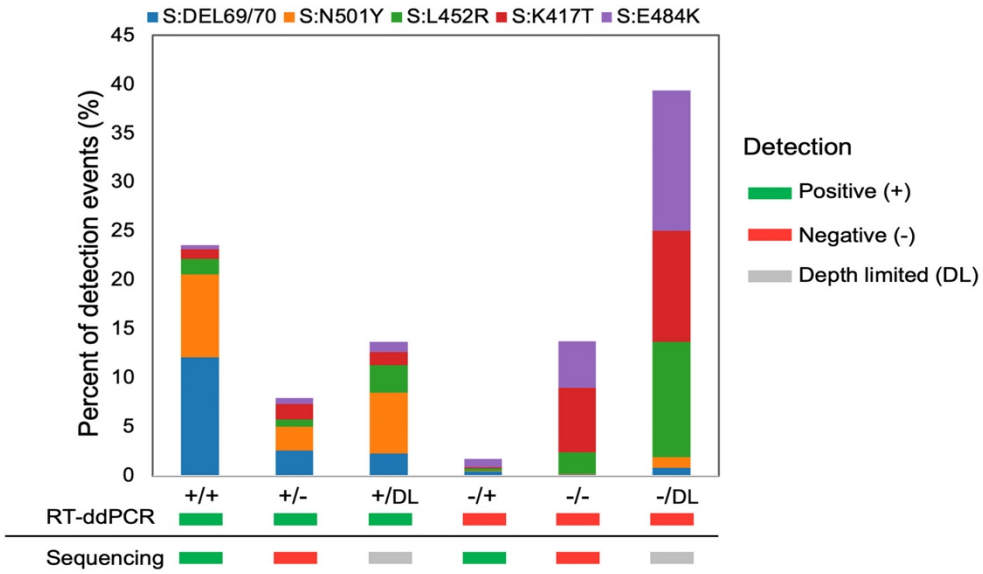


Figure 4.2. Mutation detections based on 1354 detection events via RT-ddPCR and sequencing in parallel for 547 wastewater samples. Percentage of detection events grouped by target mutations (labeled with different colors) is shown on y-axis. The six independent scenarios [(+/+), (+/–), (+/DL), (–/+), (–/–), (–/DL); format (RT-ddPCR detection/sequencing detection)], defined by in-parallel detections via RTddPCR and sequencing are on x-axis. The six scenarios were grouped accordingly based on RT-ddPCR detection and sequencing detection, respectively.

only 39.6% represented consistent detections for both methods [(+/+) and (-/-), Figure 4.2]. The inconsistency was mainly attributed to scenario (-/DL) where mutation detections were confirmed as negative by RT-ddPCR and were limited by the depth at the mutation positions for sequencing. The scenario (-/DL) alone accounted for 39.0% of all detection

events, suggesting sequencing is very likely to miss a negative detection for a mutation. The large number of DL events by sequencing largely occurred during Phase II for the detection of mutations S:E484K, S:K417T and S:L452R (Figure 4.2). During Phase II, the concentrations of SARS-CoV-2 were significantly lower than SARS-CoV-2 concentrations during Phase I when the samples were assayed for S:DEL69/70 and S:N501Y (Figure S4.1, $p<0.0001$). Phase II samples also had lower average single base coverage across the genome by sequencing as compared to Phase I samples (Table S4.6, $p<0.0001$).

When we observed positive detections by RT-ddPCR [(+/+), (+/-), and (+/DL)], 42.6% of the detection events [(+/-) and (+/DL)] were missed by sequencing. When sequencing reported negative or depth limited detections [(+/-), (+/DL), (-/-), and (-/DL)], 26.7% of those events were detected as positive by RT-ddPCR [(+/-), (+/DL)]. These two findings indicate that sequencing was less sensitive than RT-ddPCR for mutation detection. The majority (96.1%) of RT-ddPCR negative detections [(-/+), (-/-), and (-/DL)] were also consistently called as negative or depth-limited by sequencing [(-/-), (-/DL)]. The concentration of the target mutation in a sample is, in part, a function of the number of individuals infected with a SARS-CoV-2 variant containing that mutation. Thus, it would be low in the early stages of a variant outbreak, in which case sequencing may not be as sensitive as RT-ddPCR for early detection.

4.3.2 Impact of mutation concentration, single base coverage at the mutation position, SARS-CoV-2 concentration, and the average single base coverage across the entire genome on mutation detection

To understand when the inconsistencies between sequencing and RT-ddPCR were more likely to occur, we evaluated the impact of the mutation concentration (as determined by RT-ddPCR) and the average single base coverage at the mutation position (as determined by sequencing) on mutation detection events. First, we compared mutation concentrations in the sample as determined by RT-ddPCR for events where both methods reported positive detections (+/+) to events where RT-ddPCR reported positive and sequencing reported negative or depth limited detections [(+/-), (+/DL)] (Figure 4.3a).

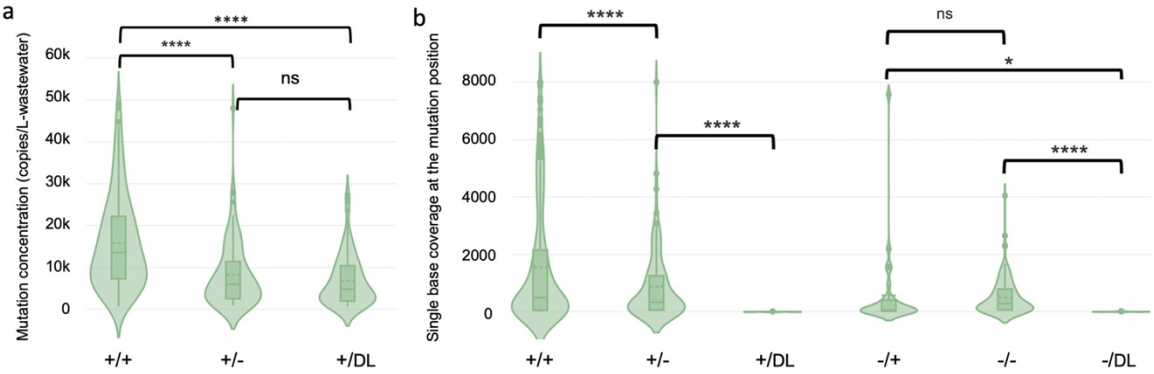
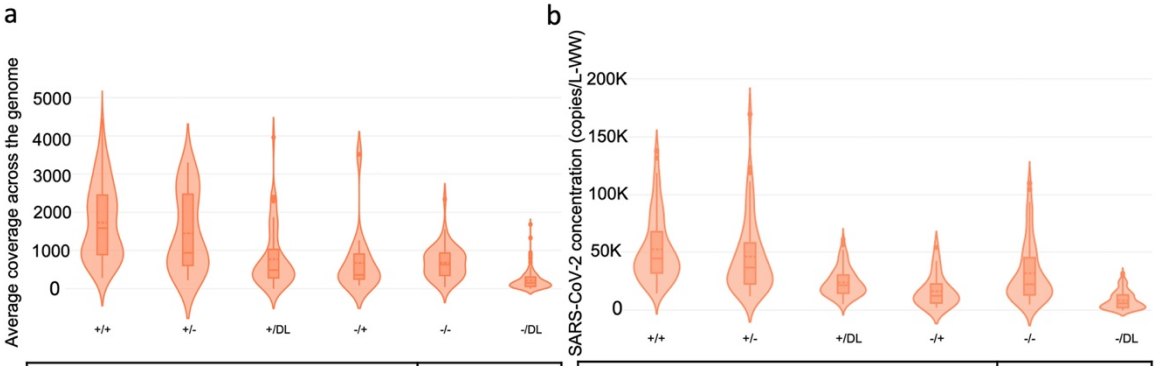


Figure 4.3 Impact of mutation concentration (a) and single base coverage at the mutation position (b) on mutation detection. Violins represent the distribution of detection events in each scenario. Boxes represent the interquartile range, with dashed lines as means and solid lines as medians. Whiskers represent the standard deviation. “ns”, “*”, and “****” indicate “not significant ($p>0.05$)”, “$p < 0.05$” and “$p < 0.0001$”, respectively, based on a t-test.

Then, we compared the single base coverage at the mutation position for these scenarios as determined by sequencing (Figure 4.3b). Results revealed samples with (+/-) and (+/DL) events had significantly lower mutation concentrations and single base coverage than those with (+/+) events (t-test, $p<0.001$; Figure 4.3).

In addition, we evaluated the impact of the sample SARS-CoV-2 concentration (average of N1 and N2 concentrations as determined by RT-ddPCR) and the average single base coverage (read depth) across the entire SARS-CoV-2 genome (as determined by sequencing) on mutation detections (Figure 4.4). We found that the SARS-CoV-2



Results of comparing scenarios in group	Significance level (Welch t-test)
RT-ddPCR positive [(+/+), (+/-), (+/DL)] versus RT-ddPCR negative [(-/+),(-/-),(-/DL)]	p<0.001
Sequencing positive [(+/+), (-/+)] versus sequencing negative [(+/-),(-/-)]	p<0.0001
Sequencing negative [(+/-),(-/-)] versus sequencing depth-limited [(+/DL),(-/DL)]	p<0.005

Results of comparing scenarios in group	Significance level (Welch t-test)
RT-ddPCR positive [(+/+), (+/-), (+/DL)] versus RT-ddPCR negative [(-/+),(-/-),(-/DL)]	p<0.0001
Sequencing positive [(+/+), (-/+)] versus Sequencing negative [(+/-),(-/-)]	p<0.0001
Sequencing negative [(+/-),(-/-)] versus Sequencing depth-limited [(+/DL),(-/DL)]	p<0.001

Figure 4.4 Impact of the average single base coverage (read depth) across the entire SARS-CoV-2 genome and SARS-CoV-2 concentration on mutation detection. Violins represent the distribution of detection events in each scenario. Boxes represent the interquartile range, with dashed lines as means and solid lines as medians. Whiskers represent the standard deviation. (a) The average single base coverage (read depth) across the entire SARS-CoV-2 genome of samples grouped by scenario. (b) SARS-CoV-2 concentration (Copies/L-wastewater) of samples grouped by scenario. The inset table under each panel contains the comparisons of the different groups of scenarios in terms of SARS-CoV-2 concentration (left table) or the average single base coverage (right table) and their significance level.

concentration and the average read depth across the entire genome were significantly higher for RT-ddPCR positive detections [(+/+), (+/-) and (+/DL), n=612] than for samples with RT-ddPCR negative detections [(-/+), (-/-) and (-/DL), n=742; Figure 4.4]. Further, we observed significantly higher SARS-CoV-2 concentrations and the average read depth across the entire genome in samples with sequencing positive detections [(+/+), (-/+),

n=342] than samples with sequencing negative detections [(+/-), (-/-), n=296] or DL [(+/DL), (-/DL), n=716; Figure 4.4].

In addition to SARS-CoV-2 concentration and read depth impacting sequencing detection, multiplexing PCR may also introduce amplification bias due to dimer formation and low Ta dropping-out (Coil et al., 2021; Tyson et al., 2020). This bias can result in an uneven coverage across the genome and impact the number of reads generated at different mutation positions, reducing the reliability of mutation calling. In this study, the two target B.1.1.7 characteristic mutations (S:N501Y and S:DEL69/70) are located in different regions of the SARS-CoV-2 genome, and they were amplified by different ARTIC v3 primers. The substitution S:N501Y (A23063T) was amplified by the primer pair ARTIC.v3_F/R_76, while the deletion S:DEL69/70 (21764ATACATG→A) was amplified by primers ARTIC.v3_F/R_72. We found that the average coverage of these two amplified regions were highly variable across samples (Figure S4.3); and the average coverage across the region amplified by ARTIC.v3_F/R_76 was significantly lower than the region amplified by ARTIC.v3_F/R_72 ($p<0.05$, n=249). This amplification bias helps explain the significantly higher average single base coverage at nt position 21764 (corresponding to S:DEL69/70) than at nt position 23063 (corresponding to S:N501Y) ($p<0.05$, n=249). Correspondingly, sequencing exhibited more sensitive detection for S:DEL69/70 than S:N501Y, detecting 71.2% of S:DEL69/70 RT-ddPCR positive events and 60.2% of S:N501Y RT-ddPCR positive events, respectively (Figure 4.2a).

With the aim of guiding wastewater-based SARS-CoV-2 monitoring in practice, we also attempted to identify whether there was a threshold level of sequencing depth, or the total reads mapped to SARS-CoV-2 reference genome, above which sequencing called

positive detections concomitantly with RT-ddPCR (Figure S4.4). Two groups of samples belonging to scenarios (+/+) and (+/-) were used for this analysis, because there was no significant difference in SARS-CoV-2 concentration or in mutation concentration between them, ensuring that the sample sequencing depth is the only variable that determined mutation detection by sequencing. We were not able to identify a threshold level of reads that could differentiate the sequencing positive and negative detection events. In other words, there was no clear pattern in the number of reads for samples with sequencing positive versus negative detections (Figure S4.4).

3.3 Allele frequency (AF) and the number of reads supporting the mutation from sequencing were not quantitative representations of the mutation concentration as determined by RT-ddPCR

We next asked if data generated from sequencing could be used to quantitatively estimate the proportion of SARS-CoV-2 that contained a target mutation. For each target mutation, we compared (1) the mutation concentration as determined by RT-ddPCR (copies/L-wastewater); and (2) the fraction of SARS-CoV-2 containing the target mutation (target mutation concentration/total SARS-CoV-2 concentration as determined by RT-ddPCR) to (3) the number of reads containing the mutation as determined by sequencing for all sequencing positive and negative detections (*i.e.,* detections with at least 20 × read depth at the mutation position); and (4) the AF of the mutation from sequencing for all sequencing positive and negative detections. We observed weak positive correlations between mutation concentration and raw read counts with the mutation [(1) vs (3); $R=0.30$, $p<0.0001$)], and between fraction of SARS-CoV-2 with the mutation and AF [(2) vs (4);

R=0.32, p<0.0001]. No significant correlation was found between mutation concentration and AF [(1) vs (4); p=0.12)], or between fraction of SARS-CoV-2 with the mutation and raw read counts with the mutation [(2) vs (3); p=0.17].

We further investigated the quantitative relationship (or lack thereof) between mutation levels of the two B.1.1.7 characteristic mutations, S:N501Y and S:DEL69/70, as determined by RT-ddPCR and sequencing during Phase I when B.1.1.7 was the dominate circulating variant in Houston. During this period, the RT-ddPCR-determined concentrations of S:N501Y and S:DEL69/70, as well as the fractions of SARS-CoV-2 containing S:N501Y and S:DEL69/70, were strongly positively correlated (R=0.95, p<0.0001 for S:N501Y and S:DEL69/70 concentrations; R=0.90, p<0.0001 for fractions of SARS-CoV-2 containing S:N501Y and S:DEL69/70; n = 228). These results indicate RT-ddPCR measurements of mutations present in a single, dominant VoC vary consistently with one another. However, when we look at the sequencing data, the correlation between the number of reads containing S:N501Y and S:DEL69/70, and the correlation between the AF values of S:N501Y and S:DEL69/70 were both weaker (R=0.21, p<0.05 for the number of reads containing S:N501Y and S:DEL69/70, n=121; R=0.33, p<0.001 for AF values of S:N501Y and S:DEL69/70, n=121). We also assessed associations between RT-ddPCR and sequencing measurements of N1 and N2 genes. Moderate positive correlations were observed between N1 concentration (via RT-ddPCR) and average single base coverage at N1 region (via sequencing, n=356, R=0.46, p=0.00011), and between N2 concentration and average single base coverage at N2 region (n=311, R=0.46, p=0.00066), respectively. Reasonably, N1 and N2 concentrations determined by RT-ddPCR exhibited a near perfect positive correlation (R=0.98, p<0.0001),

reflecting the robustness of the CDC primers and probes for measuring SARS-CoV-2 levels in wastewater samples. On the other hand, the sequencing average coverage at N1 and N2 regions only displayed a moderate positive correlation with each other ($R=0.46$, $p < 0.0001$). Previous studies that used RT-qPCR for N1 and N2 quantifications in wastewater samples found strong correlations between N1 and N2 signals, reporting correlation coefficients such as 0.952 (Sanjuán & Domingo-Calap, 2021) and above 0.99 (Ahmed et al., 2022). Studies that applied RT-ddPCR for N1 and N2 quantifications in wastewater samples also reported strong correlation coefficients above 0.85 (Feng et al., 2021) and above 0.90 (D'Aoust et al., 2021). These results highlight that AF values and the read counts as determined by sequencing for mutations may not vary consistently with one another, and thus are not appropriate for inferring VoC concentrations or relative abundances in wastewater samples.

3.4 Implications for WBE on SARS-CoV-2

RT-ddPCR or RT-qPCR should be applied for quantitative analyses due to the great sensitivity and consistency of detection. In addition, RT-ddPCR and RT-qPCR generally have much shorter result turnaround time compared to sequencing (Bloom et al., 2021), which is critical for real-time public health response. With knowledge of unique mutations associated with each VOC, it is possible to detect signatures of low levels of VOCs in wastewater samples that may contain a mixture of variants. Numerous studies have emerged that have successfully developed, validated, and applied RT-qPCR or RT-ddPCR assays for the detection of specific VOCs by targeting characteristic mutations. For instance, RT-qPCR assays have been developed to co-monitor B.1.1.7 and B.1.351 by

tracking the trend of a B.1.1.7-specific mutation, D3L, and a B.1.351-specific mutation, the deletion 242-244 (Erster et al., 2021; Yaniv et al., 2021). Recently, allele-specific and multiplex-compatible RT-qPCR assays targeting mutations T19R, D80A, K417N, T478K and E484Q for quantitative detection and discrimination of the Delta, Delta plus, Kappa and Beta variants in wastewater were developed and validated (W. L. Lee et al., 2021).

The lower detection sensitivity of sequencing (Figure 4.2) can be attributed to 1) the low concentration of target mutations in the wastewater sample (Figure 4.3a), 2) the lack of sufficient read depth at the mutation position (Figure 4.3b), 3) low SARS-CoV-2 concentrations (Figure 4.4b), and/or 4) inconsistent single base coverage across the SARS-CoV-2 genome (Figure 4.4a, Figure S4.3). The sensitivity of targeted amplicon sequencing can also be impacted by sample processing and primer choices for genome amplification. The form of SARS-CoV-2 RNA in wastewater has only been characterized in a limited number of studies, and likely exists in both intact and degraded forms (Canh et al., 2021; Wurtzer et al., 2021). The degraded form presents a challenge for amplification. For example, the ARTIC v3 primer scheme used in this study amplifies 400 bp regions of the genome, and thus may fail to amplify short, degraded RNA fragments. Further research is needed to optimize workflows including sample processing and tiled primer design for downstream sequencing and analysis using wastewater samples. For example, improvements in virus recovery and yields during wastewater sample concentration and viral RNA extraction could enhance sequencing sensitivity. In addition, little is known about the impact of concentration method on mutation detection via sequencing. Sensitivity could also be improved by developing multiplexing PCR schemes of higher amplification uniformity and efficiency (Itokawa et al., 2020), or optimizing library preparation protocols

(Coil et al., 2021). PCR inhibition due to other constituents in wastewater is another factor that may impact the sensitivity of sequencing more than RT-ddPCR, as digital PCR is relatively resilient to PCR inhibition (Ahmed et al., 2022; Ciesielski et al., 2021). Another approach for increasing sequencing coverage at specific sites is to target a smaller region of the genome for amplification and sequencing, such as by sequencing only the spike protein region of SARS-CoV-2 instead of the whole genome. For example, the receptor binding domain (RBD) on the spike region of SARS-CoV-2, which is involved in the interactions with human angiotensin-converting enzyme-2 (ACE-2) receptor, can be sequenced instead of the entire genome for mutation or variant analysis (Gregory et al., 2021). The mutations in the RBD are associated with the severity of infection (i.e., ACE-2 binding affinity and virus entry to the host cells)(Andersen et al., 2020; Heald-Sargent & Gallagher, 2012) and potential antibody-escape affecting antigenicity (Harvey et al., 2021). In addition, many of the VoCs are defined by mutations to in the Spike region (Baaijens et al., 2021). Furthermore, the spike region only accounts for approximately 12.8% of the total genome, therefore, may be a more efficient use of sequencing for mutation detection. However, this approach also suffers from amplification and sequencing challenges due to degraded RNA inherent to wastewater samples.

Despite its lower sensitivity and qualitative nature, sequencing still has a clear advantage of being more comprehensive, not limited by a priori knowledge of the target mutations, and enables the discovery of cryptic lineages (Smyth et al., 2022) and emerging lineages of concern (Sapoval et al., 2021). This can be critical for early detection of variants when the availability of primers and probes is limited or delayed due to supply chain challenges. In addition, sequencing data facilitates retrospective analyses, such as

searching for specific mutations or collections of mutations present in samples collected prior to knowledge of the variants in communities (Johnson et al., 2022; La Rosa et al., 2021; Wilton et al., 2021). In practice, WBE systems can benefit from coupling sequencing with quantitative analyses such as RT-ddPCR or RT-qPCR to achieve a comprehensive picture of circulating mutations (using sequencing), and sensitive, quantitative information on variant-associated mutations (using RT-qPCR/RT-ddPCR).

4.4. Conclusion

For WBE work on SARS-CoV-2, sequencing technology has demonstrated irreplaceable advantages in efficient screening and the potentials to detect emerging or cryptic lineages. We performed RT-ddPCR and sequencing analyses in parallel on hundreds of wastewater samples for SARS-CoV-2 monitoring, with a specific focus on mutations associated with VOCs. This is the first study to directly compare mutation detection consistency between these two methods. Results first showed the significantly greater detection sensitivity of RT-ddPCR in detecting five mutations as compared to amplicon-based sequencing. Secondly, quantitative results generated from sequencing, including allele frequency (AF) and single base coverage of specific mutations failed to reflect the concentrations of the corresponding mutations in wastewater, showing poor correlations with RT-ddPCR quantification results. Therefore, caution should be exercised in using sequencing for quantitative assessments of mutation abundance in wastewater samples. RT-ddPCR or RT-qPCR should be applied for quantitative analyses due to the great sensitivity and consistency of detection.

4.5. Acknowledgement

We thank Houston Water for their assistance in the collection of wastewater samples. We acknowledge Rebecca Schneider and Lilian Mojica from Houston Health Department. We also thank Kristina Cibor for her role in project management. This work was supported by the Houston Health Department. E.L. and L.B.S. were supported in part by the National Science Foundation (CBET 2029025), and seed funds from Rice University. T.T. and N.S. were supported in part by C3.ai DTI and P01-AI152999 NIH awards. P.K. and L.B.S. were supported in part by a Johnson & Johnson WiSTEM 2D award. Funding sources for sequencing by Houston Health Department are CDC ELC Enhanced Detection, CDC ELC Enhanced Detection Expansion, and CDC ELC Advanced Molecular Detection.

Chapter 5 Using long- and short-read metagenomics and epicPCR to profile antibiotic resistance genes and their bacterial hosts in wastewater

Esther G. Lou, Yilei Fu, Qi Wang, Todd J. Treangen and Lauren B. Stadler

Abstract. Wastewater surveillance is a powerful tool to evaluate the risks associated with antibiotic resistance. One challenge is selecting which analytical tool to deploy to measure risk indicators, such as antibiotic resistance genes (ARGs) and their respective bacterial hosts. Although metagenomic sequencing is frequently used for resistome analysis, few studies have compared the performance of long-read and short-read sequencing in identifying the host range of ARGs in wastewater. In addition, metagenomic sequencing as an unbiased approach has not been compared to targeted methods such as epicPCR for ARG host detection. In this study, we 1) evaluated long-read and short-read metagenomic sequencing as well as epicPCR for detecting ARG hosts in wastewater, and 2) investigated the fate of ARGs and ARG hosts across the WWTP to evaluate potential risks. Results showed that despite its significantly lower sequencing depth, long-read sequencing outperformed short-read sequencing with higher sensitivity for detecting ARGs, especially for ARGs associated with mobile genetic elements (MGEs). In addition, long-read sequencing consistently revealed a wider range of ARG hosts compared to short-read sequencing. Nonetheless, the host range detected by long-read sequencing only represented a subset of the host range detected by epicPCR. The associations between ARGs and their host phyla were relatively consistent across the locations sampled in the WWTP. However, both long-read sequencing and epicPCR detected new host species that emerged in the treated effluent. Based on these findings, we recommend 1) using long-read sequencing for routine wastewater surveillance, 2) using epicPCR to obtain a high-resolution host range of clinically relevant ARGs, and 3) performing long-term surveillance on specific treatment compartments within a WWTP to understand the emergence of new hosts.

5.1. Introduction

The worldwide propagation and dissemination of antibiotic resistance have raised serious public health concerns. An estimated 1.27 million deaths were attributed to bacterial antibiotic resistant infections in 2019 (Murray et al., 2022). To mitigate this risk, a comprehensive understanding of antibiotic resistance in humans, animals, and the environment is needed (McEwen & Collignon, 2018). Wastewater treatment plants (WWTPs) are regarded as hotspots of antibiotic resistance in the environment, and their roles in impacting the fate of antibiotic resistance are quite complex: (1) wastewater treatment processes generally remove antibiotic resistance genes (ARGs) and ARG hosts from sewage through different mechanisms including but not limited to, anaerobic (Kappell, Kimbell, Seib, Carey, Choi, Kalayil, Fujimoto, et al., 2018b; Lou et al., 2020) and aerobic processes (Du et al., 2015; Munir et al., 2011), coagulation and sedimentation (N. Li et al., 2017), membrane filtration (Cheng & Hong, 2017; Z.H. Li et al., 2019), and disinfection (H. Li et al., 2020; W. Lin et al., 2016); (2) a significant abundance of ARGs and ARG hosts can remain in treated effluent, which may pose a risk to downstream environments (Mao et al., 2015b; McConnell et al., 2018b; Rafraf et al., 2016b). In fact, ARGs and microorganisms discharged by WWTPs were shown to persist in the receiving-river biofilms (Kneis et al., 2022) and sediments (Chu et al., 2017; Quintela-Baluja et al., 2019).

Previous studies have proposed metrics for evaluating the risks associated with antibiotic resistance in the environment to public health (A.N. Zhang et al., 2021; Z. Zhang et al., 2022). For instance, antibiotic resistant, pathogenic ARG hosts, as opposed to just ARGs, are the key factor to assess the antibiotic resistance threats to public health

(Ben et al., 2019; Rice et al., 2020). Furthermore, mobile genetic elements (MGEs), which can directly and indirectly mediate horizontal gene transfer (HGT) of ARGs among microorganisms, are regarded as another significant factor contributing to the risk of ARG dissemination (Che et al., 2019a; L. Ma et al., 2016; Yin et al., 2022). Based on these two aspects, Zhang et al. recently proposed a risk ranking system for ARGs that considers: the anthropogenic carriage of ARGs, the pathogenicity of the ARG hosts, and the mobility of ARGs (A. N. Zhang et al., 2021).

To understand the role of the environmental dissemination of antibiotic resistance on public health, methods are needed to obtain high-resolution and sensitive information on ARG hosts and the associations between ARGs with MGEs. One of the most widely used methods to analyze antibiotic resistance is metagenomic sequencing (Fuhrmeister et al., 2021; Majeed et al., 2021b; Riquelme et al., 2021). Next-generation sequencing (i.e., short-read) coupled with *de novo* assembly recovers ARG-host linkages by screening taxonomical markers on the assembled ARG-carrying contigs (Ju et al., 2019; L. Ma et al., 2016; Yin et al., 2019). However, this method tends to suffer from limited detection sensitivity due to the low percentage of raw reads mapping back to the assembled contigs (Vollmers et al., 2017) and intergenomic assembly errors (Nurk et al., 2017). Importantly, a large portion of reads associated with MGEs fail to assemble because of the extended homologous and mosaic sequences found in those regions (Brown et al., 2021; Juraschek et al., 2021; Maguire et al., 2020). Third-generation sequencing technologies (i.e., long-read) are excellent at tracking ARG hosts in environmental samples (Arango-Argoty et al., 2019; Che et al., 2019a) because long-read sequencing can reveal the genetic context of ARGs without read assembly. For example, two studies that used Oxford Nanopore

Technology (ONT) to sequence wastewater metagenomes reported an average read length of 2-10 kbp, which was longer than the average length of short-read assembled contigs generated from the same samples (Che et al., 2019a; Dai et al., 2022). However, because both long- and short-read metagenomic sequencing are unbiased, non-targeted methods, their ability to detect ARGs, which are generally present at low abundances in the microbial community, is limited (Z. Liu et al., 2019; R. Zhao et al., 2020). On the other hand, targeted methods such as epicPCR (Emulsion, Paired Isolation, and Concatenation PCR) can circumvent this issue. The cornerstone of epicPCR's methodology is single cells are isolated and encapsulated in a polyacrylamide bead within which PCR takes place to fuse a target ARG with the 16S rRNA gene (S. J. Spencer et al., 2016). As a result, PCR amplifies the signal of the ARG and its associated host 16S DNA from the background environmental metagenome, improving detection sensitivity.

In this study, we applied long- and short-read metagenomic sequencing as well as epicPCR to assess ARGs and ARG host range in a WWTP's influent and effluent. Long- and short-read sequencing were compared in terms of their sensitivity to detect ARGs and ARG-host linkages. In addition, the host range of ARGs detected by metagenomic sequencing and epicPCR were compared. Furthermore, the putative pathogenic hosts of ARGs, as well as the genetic contexts of ARGs, were scrutinized to understand and prioritize ARGs based on their potential impact on human health. To the best of our knowledge, this is the first study to explicitly compare ARGs and ARG-host interactions as revealed by long- and short-read sequencing, and epicPCR. Based on these comparisons, we provided method suggestions for wastewater surveillance on antibiotic resistance for public health.

5.2. Materials and Methods

5.2.1 Sample collection, DNA extraction, and pretreatment for epicPCR

Wastewater samples were collected from a conventional WWTP (City of West University Place WWTP, Houston, Texas, USA) that treats an average of 1 million gallons of municipal sewage per day. This WWTP employs a conventional aerobic activated sludge process as secondary treatment, followed by chlorination disinfection (gaseous Cl_2, 2-4 mg/L effective chlorine concentration, 20 mins contact time). Nine grab samples were collected from three sampling locations, WWTP influent, secondary effluent, and final effluent on three consecutive dry days (n=3 for each sampling location). All samples were collected at the same time of the day to avoid diurnal variations. After collection, samples were kept on ice, immediately transported to the lab, and processed within 45 minutes of collection.

DNA was extracted from all samples prior to conducting long- and short-read metagenomic sequencing. A 50 mL influent sample, 250 mL secondary effluent sample, and 500 mL final effluent sample were filtered through a cellulose nitrate membrane filter (pore size 0.22 μm, diameter 47 mm; Millipore Sigma) to concentrate biomass. Next, filters were cut into small pieces using sterilized forceps and transferred to a 2 mL tube containing 0.1 mL glass beads for bead-beating, followed by DNA extraction. A Maxwell RSC Instrument (Cat. Num. AS4500, Promega) using Maxwell RSC PureFood GMO and Authentication kits (Cat. Num. AS1600, Promega) were used to extract DNA. For epicPCR, all influent (n=3) and final effluent samples (n=3) were centrifuged to concentrate cells for cell counting, polymerization, and cell lysis as previously described

(Spencer et al., 2015). Only samples with good cell separation and partitioning in polyacrylamide beads (i.e., one single cell per 35-50 polyacrylamide beads) were used in the downstream experiments to avoid false positive detections. Details of DNA extraction and sample pretreatment for epicPCR are provided in Chapter 5 Appendix section 1.2.

5.2.2 Sequencing epicPCR product using MinION (ONT)

We selected three ARG targets for epicPCR analysis: *sul1*, *ermB*, and *tetO*. They were chosen because of their wide host range as previously reported (Stalder et al., 2019; Wei et al., 2021; A. N. Zhang et al., 2021). For example, *ermB*, the macrolide-lincosamide-streptogramin B (MLSB) resistance gene, is of clinical relevance because it is enriched in human-related environments, harbored by human pathogens, and often carried on MGEs (A. N. Zhang et al., 2021). Primer sequences for the three targets used in this study are listed in Table S1. Details of the epicPCR experiments consisting of fusion PCR and nested PCR are provided in Chapter 5 Appendix section 1.2. After attaining nested PCR products, library preparation and sequencing were performed following the protocol "Native barcoding amplicons (with EXP-NBD104, EXP-NBD114, and SQK-LSK 109)" (ONT). The pooled library was loaded on an R9.4 flow cell (MIN-FLO106, ONT) in a MinION device. The sequencing run was monitored via the software MinKNOW (v.20.10), targeting a >1000X depth per sample.

5.2.3 Metagenomics sequencing (long- and short-read)

DNA extracts of all samples were measured using a Qubit Broad Range dsDNA assay kit and Qubit 2.0 fluorometer. DNA quality was then evaluated using

electrophoresis to ensure a DNA size greater than 3 kbps. For short-read sequencing, DNA extracts were shipped on dry ice to BGI Tech Solutions (Hong Kong) Co., Ltd for DNBseq general DNA library construction and DNBseq platform sequencing. For long-read sequencing, 500 ng of DNA from each of the three sample replicates were combined for library preparation following the protocol "Genomic DNA by ligation (SQK-LSK 109)" (ONT). Each of the three libraries (influent, secondary effluent, and final effluent) was loaded onto a Flow Cell R9.4 (MIN-FLO106, ONT) and sequenced with a MinION device. The sequencing run was controlled via MinKNOW (v.20.10). Long- and short-read sequencing statistics are provided in Table S5.2.

5.2.4 Analysis of epicPCR reads for ARG host range profiling

Raw reads were basecalled via guppy_basecaller (Version 4.4.1+1c81d62). Basecalled reads were trimmed by Porechop (https://github.com/rrwick/Porechop) and filtered using Nanofilt with a minimum quality score of 7 (De Coster et al., 2018). Next, all reads were searched against the corresponding linker primer sequence (RL-*sul1*-519F′, RL-*ermB*-519F′, and RL-*tetO*-519F′) using BLAST. The output reads were filtered using the perfect match criteria (100% identity and 100% length coverage) to exclude partially fused fragments, and only complete ARG-16S rRNA fusion structures were included for the downstream analysis. Then, we used a customized script to split the fusion structures into the ARG and the 16S rRNA gene portions based on the reverse linker position. We then conducted taxonomic classification on the 16S rRNA gene portion using Emu (Curry et al., 2022) and the SILVA ribosomal RNA gene database (release 138, 2019). To avoid false positives, two actions were taken to further filter the reads: 1) The ARG

portion of all split reads was aligned against the SARG database (Yin et al., 2018) using BLAST, 2) for each sample, only hosts that were identified consistently from at least two sample replicates were counted. The results of epicPCR sequencing statistics can be found in Chapter 5 Appendix section 2.1.

5.2.5 Analysis of metagenomic sequencing reads generated by long-read and short-read sequencing

We processed long- and short-read sequencing data in an integrated pipeline as shown in Figure S5.1. Our metagenomic analysis included: (1) identification of ARGs on long reads (via long-read sequencing) or short-read-assembled contigs (via short-read sequencing); (2) filter ARG-carrying long reads and contigs to include only those that were chromosome-associated for the host classification step; (3) identification of MGEs that were located on the same read or contig as the ARGs, and (4) identification of ARG host by taxonomic classification of the chromosomal reads or contigs that were associated with ARGs. Detailed methods describing the pipeline used to detect ARG-carrying long reads via BLAST, ARG-carrying contigs via CARD's Resistance Gene Identifier (RGI, Alcock et al., 2020), and ARG-MGE linkages are provided in Chapter 5 Appendix section 1.3. ARG relative abundance was calculated by normalizing the copy number of ARGs detected on long reads or assembled contigs to the total giga base pairs (Gbp) of the sample (Arango-Argoty et al., 2019).

ARG-carrying reads or contigs were categorized as "chromosome," "plasmid," or "unclassified" via PlasFlow (V1.1) (Krawczyk et al., 2018). The "unclassified" reads and contigs were re-classified via megaBLAST against the NCBI nt database with a

minimum bit score of 50, an E value threshold of e^{-10}, and a 70% sequence similarity cutoff, followed by keyword match ("chromosome") to retrieve chromosome-associated long reads and contigs. All ARG-carrying reads and contigs classified as "chromosome" were subject to taxonomic classification using Centrifuge (V1.0.4; Kim et al., 2016). ARG-host linkages were identified by summarizing the associations between each ARG and the taxonomic classification result of the corresponding ARG-carrying read or contig. Putative pathogens were scanned according to the WHO resistant pathogen list (WHO, 2017). In addition, three publicly available datasets from NCBI SRA were downloaded, each consisting of long-read (via Nanopore) and short-read (via Illumina) data based on sequencing the same wastewater sample (Che et al., 2019a; Fuhrmeister et al., 2021). These datasets were run through identical pipelines for analyzing long- and short-read sequencing data as used in this study to identify ARG-host linkages. Details regarding the three datasets are provided in Table S5.3. Plasmid-associated ARG-carrying reads and contigs were subject to plasmid mobility prediction using MOB-suite (v3.0.3; Robertson & Nash, 2018) and MOBscan (https://castillo.dicom.unican.es/mobscan/). Furthermore, to compare long-read sequencing with epicPCR for ARG host profiling, we processed Centrifuge specifically for those long reads that were found to carry *ermB*, *sul1*, and *tetO*.

5.3. Results and Discussion

5.3.1 Direct comparison of long- and short-read sequencing for resistome analysis

5.3.1.1 Long-read sequencing resulted in the detection of a more diverse and abundant resistome as compared to short-read sequencing

We first compared long- and short-read sequencing in their ability to characterize the diversity of ARGs (defined as the number of unique ARGs) and the relative abundance of ARGs (the copy number of ARGs normalized to sequencing depth) present in the samples. Overall, for raw wastewater (WWTP influent), long-read sequencing detected 347 ARGs with a total ARG relative abundance of 614 reads per billion bases (RPB), whereas short-read sequencing detected 191 ARGs with a total ARG relative abundance of 341 RPB. The total ARG relative abundance generated by both methods was comparable to previous metagenomic analyses that quantified ARGs of wastewater samples collected from western countries (Hendriksen et al., 2019; Kutilova et al., 2021; Riquelme et al., 2021). Therefore, in our study, long-read sequencing detected a significantly more diverse and abundant ARG profile as compared to short-read sequencing, which was surprising since the long-read sequencing depth was much shallower - only 10.12% of that of the short-read counterpart. One explanation for the better performance of long-read sequencing is its higher ARG detection sensitivity, underscored by the significantly higher proportion of ARG-associated reads among all long-read sequencing reads (0.0576%) as compared to the proportion of ARG-associated contigs among all short-read-assembled contigs (0.0119%).

The greater detection sensitivity of long-read sequencing is likely the result of a better preservation of the information of raw reads as compared to short-read sequencing. To elaborate, for short-read sequencing data, only 34.3% of raw reads mapped to the analyzed contigs, indicating a significant read loss during *de novo* assembly, which is a common issue for environmental metagenomes (Deshpande & Fahrenfeld, 2022; L. Ma et al., 2017b; R. Zhao et al., 2020). Of note, the analyzed contigs corresponded to those

passed the length filter (1,500 bp), which accounted for approximately 23.1% of all assembled contigs, indicating the length of the assembled contigs was a limiting factor of ARG detection by short-read sequencing. In this study, short-read assembly was treated as a necessary step to avoid false positives caused by highly similar and relatively short ARG reference sequences. After *de novo* assembly, only 0.0125% of the assembled contigs passed through the filter of the ARG alignment step (based on RGI "perfect" and "strict" matches for ARG calling, Alcock et al., 2020). For long-read processing, 100% of reads were directly subject to ARG alignment because no assembly was needed, and 14.8% of reads passed the filter of the ARG alignment step. Assembly of long reads was not performed due to the limited coverage (data not shown). Therefore, the number of ARG-carrying reads via long-read sequencing was greater than the number of ARG-carrying contigs via short-read sequencing (Table S2). Another reason long-read sequencing may have been more sensitive is because, on average, the size of the ARG-carrying reads (mean size=5,387 bp) was significantly larger than the ARG-carrying contigs (mean size=3,488 bp; $p=1.592e^{-05}$). Longer reads increased the likelihood of detecting multiple ARGs on the same reads. The number and fraction of long reads carrying more than one ARG (413; 23.7%) were greater than the number and fraction of contigs carrying more than one ARG (24; 10.7%). In addition, the contigs that carried more than one ARGs were carrying two or three ARGs, whereas 26.6% of the long reads that carried more than one ARGs were carrying at least three and up to six ARGs. Taken together, long-read sequencing resulted in higher ARG detection sensitivity than short-read sequencing by preventing read loss and through the generation of extended length of ARG-carrying reads.

However, both methods identified a comparable ARG composition with respect to ARG subtypes; each method detected the same suite of 20 ARG subtypes in wastewater (Figure 5.1A). Approximately 90% of the total ARGs detected by each method belonged to subtypes of sulfonamide, macrolide-lincosamide-streptogramin (MLS), tetracycline, multidrug, carbapenem, aminoglycoside, antiseptics, and non-carbapenem-beta-lactams. The consistency between long-read and short-read sequencing in characterizing ARG subtypes has also been reported in other studies of wastewater samples and activated sludge samples (Che et al., 2019), mock bacterial communities (Leggett et al., 2020), and a plant population with a known bacteria spike (Arango-Argoty et al., 2019).

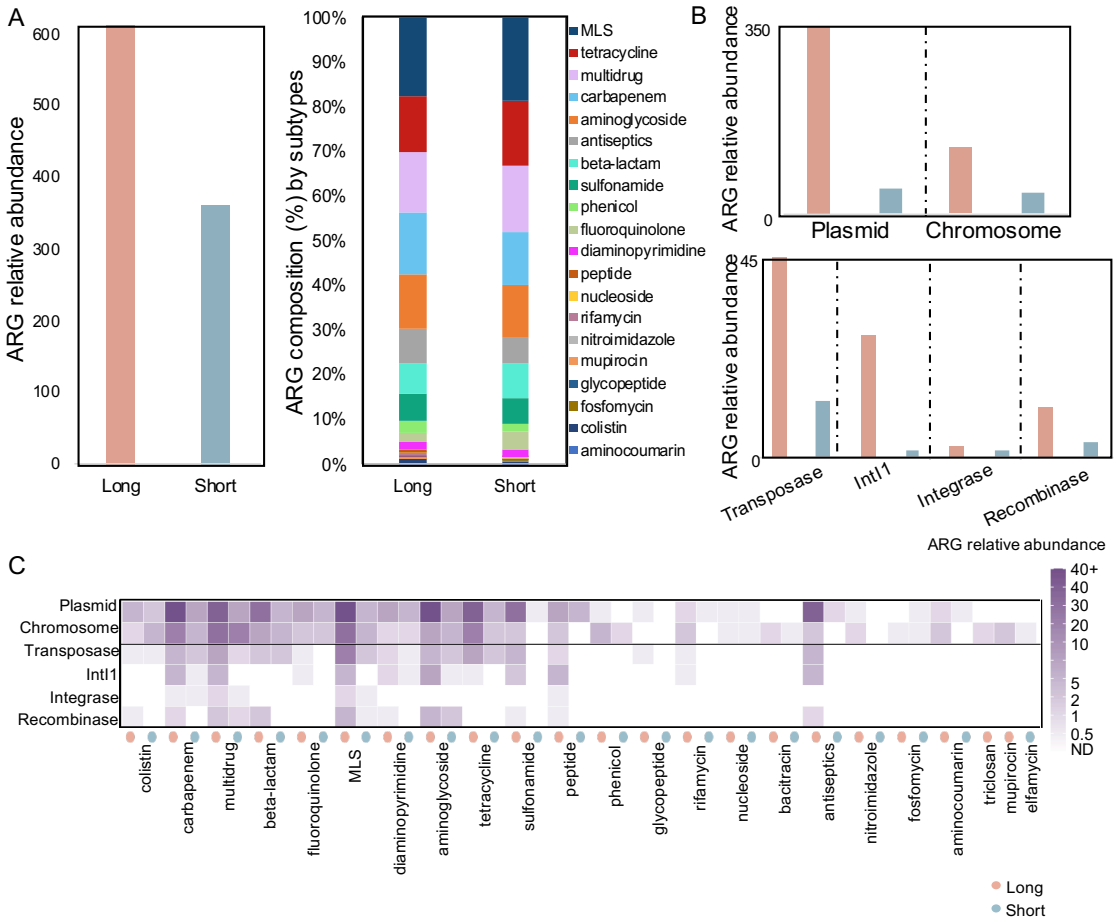


Figure 5.1. Resistome profiles revealed by long- and short-read sequencing on paired wastewater samples (n = 3). A. Total ARG relative abundance revealed by long- and short-read sequencing (left) and ARG composition broken down by drug class subtype (right) according to the relative abundance of ARGs of each subtype. B. Distribution of total ARGs across genetic locations (plasmid or chromosome) and the associations between ARGs and MGEs as determined by long- and short-read sequencing. ARGs associated with more than one MGE were counted separately for each MGE involved. C. Distribution of ARGs (grouped by drug class subtype on the x-axis) across genetic locations and ARG-MGE associations revealed by long- and short-read sequencing.

5.3.1.2 Long-read sequencing detected more ARGs located on chromosomes, plasmids, and on different types of mobile genetic elements (MGEs) as compared to short-read sequencing

Next, we compared the genetic location of ARGs assigned by each sequencing method. Both methods captured ARGs distributed across different genetic locations, namely, plasmid and chromosome. In addition, the associations between ARGs and MGEs (transposases, integrases, recombinases, and integrons) were also recovered. Long-read sequencing exhibited a greater abundance of ARGs that were associated with every single genetic location as compared to short-read sequencing (Figure 5.1B). More specifically, long-read sequencing showed a significantly higher abundance of plasmid-associated ARGs, 6-fold higher than that of short-read sequencing, and a strikingly higher abundance of class 1 integron-integrase genes (*IntI1*)-associated ARGs, 16-fold higher than that of short-read sequencing (Figure 5.1B). The less sensitive detection of MGE-associated ARGs by short-read sequencing was likely the result of the de novo assembly process. To elaborate, the variable copy number and the highly homologous and repetitive sequence compositions of MGEs make it problematic to assemble MGE-associated reads. As shown in a previous study, 82-94% of chromosomal sequences were correctly assembled and binned, but only 38-44% of genomic islands and 1-29% of plasmid sequences were identified in a simulated low-complexity short-read metagenome (Maguire et al., 2020). A similar degree of read loss during short-read assembly was also observed in several other studies of wastewater metagenomes (Deshpande & Fahrenfeld, 2022; J. Liang et al., 2020; Z. Liu et al., 2019; L. Ma et al., 2016). Long-read sequencing, on the other hand, does not require assembly as it generates long reads that can be directly searched against MGE databases. Therefore, long-read sequencing can overcome the data loss issue associated with assembly, making it more feasible to detect ARG-MGE linkages.

So far, only a handful of studies have compared using long-read and short-read sequencing to determine the genetic locations of ARGs in wastewater samples. One study that obtained reads via Nanopore sequencing and contigs assembled from Illumina-sequencing reads found that both resulted plasmid-associated ARGs for all major subtypes of ARGs in wastewater and activated sludge samples (Che et al., 2019). Similarly, in this study, ARGs were found to be primarily located on plasmids rather than chromosomes (Figure 5.1B). In addition, ARGs were mostly co-located with transposases and *IntI1* (Figure 5.1B). We also investigated the distribution of ARGs across different genetic locations with respect to ARG subtypes (Figure 5.1C). For ARGs conferring resistance to carbapenem, multidrug, MLS, diaminopyrimidine, aminoglycoside, tetracycline, nucleoside, and bacitracin, long-read sequencing demonstrated a consistent or slightly wider MGE distribution range compared to short-read sequencing (Figure 5.1C). However, the distribution patterns of sulfonamide resistance genes, peptide resistance genes, rifamycin resistance genes, and antiseptics resistance genes were distinct for each method. Short-read sequencing assigned these ARGs only to plasmids whereas long-read sequencing assigned these ARGs not only to plasmids but also to other MGEs (Figure 5.1C). This inconsistency was likely due to the significantly lower number of ARG-associated contigs detected by short-read sequencing for those specific ARGs (data not shown), which limited its ability to fully capture the potential of those ARGs being associated with MGEs. While this is not the first study to elucidate the genomic locations of ARGs by investigating the genetic context of ARGs, it is the first to explicitly compare long-read and short-read sequencing in profiling the distribution of

ARGs across genomic locations (i.e., chromosomes, plasmids, and other MGEs) in wastewater.

5.3.2 Direct comparison of the ARG host range profiled by long-read and short-read sequencing demonstrated long-read sequencing's superior performance for host tracking

Overall, long-read sequencing identified a greater number of linkages of ARGs and their hosts than short-read sequencing (Figure 5.2A, Figure S5.2). This result highlights that long-read sequencing produced a more diverse host profile than short-read sequencing, even though both methods produced similar total community bacterial composition profiles (Table S5.4) and consistent ARG subtype profiles (Figure 5.1A). However, these two methods showed inconsistency in ARG host identifications (Figure 5.2A). In total, 26 ARG-host family linkages, or 21 ARG subtype-host family linkages, were consistently detected by both methods, which accounts for only a small fraction of the corresponding total linkages detected by each method (Figure 5.2A). Although several studies have focused on the consistency of long-read and short-read sequencing in resistome analysis (Arango-Argoty et al., 2019; Che et al., 2019; Leggett et al., 2020) and sample-wise taxonomic abundance estimation (Brandt et al., 2020; Govender & Eyre, 2022; Sevim et al., 2019; Y. Yang et al., 2022), ours is the first to explicitly compare their ability to characterize ARG host range in wastewater.

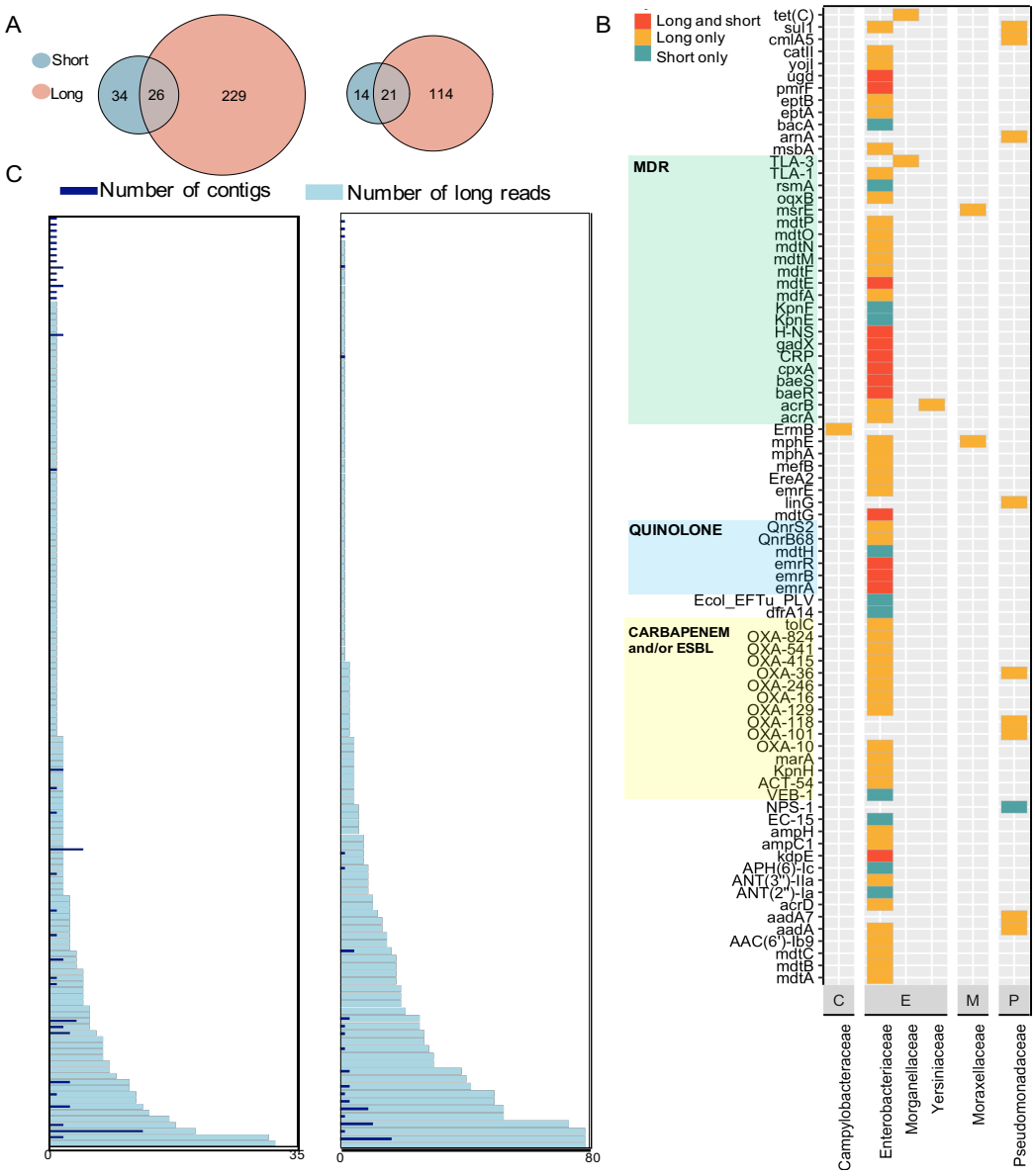


Figure 5.2. Comparison of long- and short-read sequencing for wastewater ARG host identification. A. Venn diagrams illustrating the detections of unique linkages of a specific ARG and its host family (left), and of unique linkages of a specific ARG drug class subtype and its host family (right). B. Profile of ARG hosts that are bacteria on the WHO list for which new antibiotics are urgently needed. Highlighted ARGs are those conferring multidrug resistance (MDR), fluoroquinolone resistance, and those encoding ESBL-production and/or carbapenemase-production. The colors denote a detection of an ARG-host (orange: detected only by long-read sequencing, turquoise: detected only by short-read sequencing, red: detected by both sequencing technologies). Family-level hosts are grouped by Order on x-axis (C: Campylobacterales, E: Enterobacterales, M: Moraxellales, P: Pseudomonadales). C. The number of reads via long-read sequencing (light blue bars) and the number of contigs via short-read sequencing (dark blue bars) supporting each unique linkage of ARG subtype and host (family level). The left panel consists of data generated in this study, and the right panel is using publicly available data of a sample collected from the influent of a WWTP in Boston, MA (sample ID: B_ww_1, Table S5.3).

We further compared the ARG-host linkages profiled by each method, focusing specifically on those putative hosts included on the WHO's resistant pathogen list. Long- and short-read sequencing altogether recovered 117 ARG-host linkages, covering 80 ARGs (corresponding to 17 subtypes) and 31 putative pathogenic species (Figure 5.2B). Both methods identified *Escherichia coli* as the putative pathogenic host that carried the highest abundance of ARGs. In fact, according to global surveillance of clinical cases, among all bacterial pathogens associated with or attributable to antibiotic resistance, *E. coli* ranks first as the cause of direct or indirect deaths (Murray et al., 2022). Not surprisingly, most putative hosts identified were within the family *Enterobacteriaceae* which includes the vast majority of commensal and enteric bacteria that live in the gastrointestinal tract of humans (Abera et al., 2016; K. Lee et al., 2020; Ngbede et al., 2021). Consistent with previous studies, *Enterobacteriaceae* was found to harbor multiple classes of clinically relevant ARGs, especially those encoding ESBL (extended-spectrum beta-lactamase)-production and/or carbapenemase-production (Abera et al., 2016; Castanheira et al., 2021; L. Li et al., 2021; Søraas et al., 2013). Long-read sequencing detected ARG-host connections across six host families and 68 ARGs, whereas short-read sequencing only detected *Enterobacteriaceae* and *Pseudomonadaceae* as the host families for 24 ARGs (Figure 5.2B). Hence, despite having a significantly shallower sequencing depth (Table S5.1), long-read sequencing detected a more comprehensive profile of putative pathogenic hosts of ARGs than short-read sequencing (Table S5.5).

To investigate the inconsistency between the two sequencing methods, we compared the number of reads supporting the ARG-host linkages detected by long-read

sequencing and the number of contigs supporting the ARG-host linkages detected by short-read sequencing (Figure 5.2C). To assess the consistency of our results with other previous datasets, the results from this study and from one publicly available wastewater metagenomic dataset with paired long- and short-read sequencing data [ID: B_ww_1, (Fuhrmeister et al., 2021)] are provided (Figure 5.2C, right; results for the rest of the example publicly available datasets are provided in Figure S5.2). As expected, long-read sequencing demonstrated more ARG subtype-host family linkages as compared to short-read sequencing. Quantitatively, the numbers of long reads were generally greater than the numbers of contigs across the vast majority of the ARG subtype-host family linkages (Figure 5.2C). For those reads and contigs that supported the same linkages, their numbers were moderately correlated ($n=21$, Spearman's $Rho=0.43$, $p<0.05$ for this study; $n=18$, Spearman's $Rho=0.47$, $p<0.05$ for B_ww_1), which indicates some degree of consistency between these two methods in quantifying the linkages of ARG subtypes and host families.

To summarize, long-read sequencing demonstrated its superior performance over short-read sequencing in detecting ARG hosts in two aspects: 1) it captured a wider host range for different ARGs (Figure 5.2B) and ARG subtypes (Figure 5.2C, Figure S5.2); and 2) quantitatively, it detected ARG-host linkages by generating greater numbers of reads supporting the linkages than the number of contigs (Figure 5.2C, Figure S5.2). Of note, the evaluation was based on comparing the detection via long reads with the detection via contigs assembled from short reads, rather than directly comparing the raw reads generated by both sequencing methods. Several previous studies used raw short reads without assembly and identified putative ARG hosts through correlation analysis

that compared the abundance of ARGs and taxonomical markers (Jia et al., 2017; B. Li, Ju, et al., 2015). However, using raw reads to assign ARG hosts for wastewater surveillance has several challenges. First, this approach relies heavily on statistical correlation analysis which requires multiple sample replicates. Obtaining and processing multiple replicate samples significantly increases the amount of work required for sample collection, preparation, and sequencing, as well as the time and cost involved in routine surveillance. Most importantly, the raw read approach is prone to introduce false positives when detecting ARG-host linkages (Deshpande & Fahrenfeld, 2022). Thus, while using assembly is a relatively conservative means to identify ARG-host linkages as compared to using raw reads, it is less likely to generate false positives, which is crucial in wastewater surveillance and risk assessment. Furthermore, the substantial processing time and computational memory requirements of read assembly of short-read sequencing data may be overcome by long-read sequencing. Recent studies have shown that ONT, one of the leading long-read sequencing technologies, can rapidly and reliably detect resistomes and pathogens in one hour in wastewater (Yang et al., 2022) and preterm microbiota (Leggett et al., 2020).

5.3.3 EpicPCR captured more ARG hosts as compared to long-read sequencing

Since long-read sequencing revealed more ARG hosts than short-read sequencing, we next compared the ARG host profile obtained via long-read sequencing to the host range of three ARG targets (*sul1*, *ermB*, and *tetO*) detected by epicPCR. We found that epicPCR detected a greater number of host species for the three ARG targets than the non-targeted, long-read sequencing method (Table 5.1). Note that the host species

generated by long-read metagenomic sequencing were based on classifying chromosomal reads that carry the corresponding ARGs (as discussed in section 2.5; Table 5.1).

Table 5.1. The number of ARG-associated reads and host species detected by epicPCR and long-read sequencing in WWTP influent and effluent samples (n=3). The comprehensive list of hosts detected by epicPCR can be found in Chapter 5 Appendix section 2.2 and Table S5.5.

WWTP influent						
Method	**EpicPCR**	**Long-read metagenomics**				**Both**
	Host species	Total reads	Host species via classifying total reads	Chromosomal reads	Host species via classifying chromosomal reads	Host species
sul1	311	124	32	3	3	2
ermB	57	18	8	3	1	1
tetO	104	30	20	8	4	4
WWTP effluent						
Method	**EpicPCR**	**Long-read metagenomics**				**Both**
	Host species	Total reads	Host species via classifying total reads	Chromosomal reads	Host species via classifying chromosomal reads	Host species
sul1	323	8	7	0	-	-
ermB	14	0	-	0	-	-
tetO	35	0	-	0	-	-

There are two reasons why epicPCR was more sensitive than long-read sequencing in terms of ARG host detection. First, epicPCR is more sensitive as compared to non-targeted metagenomics. Second, epicPCR can include plasmid-associated linkages that may have been overlooked by metagenomics. With epicPCR (Spencer et al., 2016), as long as the target ARG is present in the cell, it can be fused with the taxonomical marker (i.e., 16S rRNA gene) via PCR for host classification regardless of whether the ARG is located on a plasmid or chromosome. However, metagenomics can only classify hosts for ARGs that are associated with chromosomes because it requires the presence of

taxonomic markers co-located on the ARG read, which usually appear on chromosomes rather than on plasmids. Therefore, in theory, epicPCR should generate a more comprehensive host profile than metagenomics, because ARGs are widely distributed on plasmids (Che et al., 2019a; Dai et al., 2022; Rozwandowicz et al., 2018).

The significantly lower number of hosts identified via long-read sequencing as compared to epicPCR was likely attributed to the low fraction of chromosomal reads among the total ARG-associated reads. To further investigate, we selected all long reads that were found to carry *sul1*, *ermB*, and *tetO* disregarding whether they were on chromosomes. As expected, a vast majority of ARG-carrying reads were not classified as chromosomal reads for the three ARG targets (Table 5.1). However, it is not feasible to classify hosts using metagenomics for non-chromosomal ARGs, such as plasmid-borne ARGs, because plasmids do not necessarily resemble their host genomes in either abundance or nucleotide composition. In addition, the possibility of HGT of those plasmid-borne ARGs is unknown. Nevertheless, despite epicPCR's improved detection sensitivity over long-read sequencing, one undeniable value of long-read sequencing is its ability to disclose genetic context of an ARG. For instance, long-read sequencing identified the associations between *ermB* and the gene encoding conjugative transposon proteins in *Clostridioides difficile* (data not shown), highlighting *ermB*'s potential to be horizontally transferred. EpicPCR results in only a short, fused product of the target gene and taxonomic marker gene, retaining no additional contextual information in the sequenced amplicons.

5.3.4 The ARG-host phylum linkages were relatively consistent across WWTP influent and effluent, whereas new ARG-host species linkages appeared in the WWTP effluent

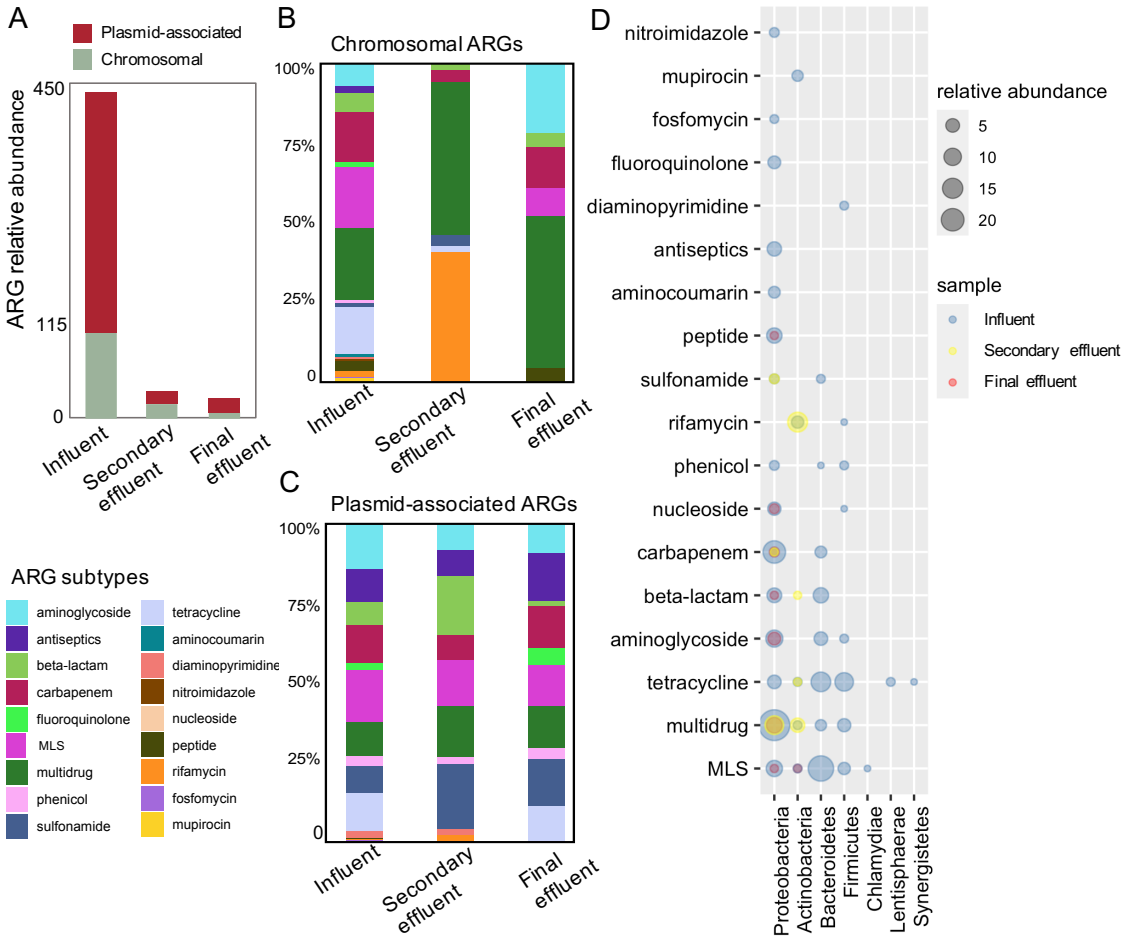


Figure 5.3. Dynamics of resistomes and ARG hosts across the WWTP revealed by long read sequencing. A. The relative abundance of ARGs across the WWTP (influent, secondary effluent and final effluent, n = 3 for each sampling location). ARGs were grouped by their location (red: plasmids, green: chromosome). B. The composition of chromosomal ARGs across samples. ARGs are colored by drug class subtype. C. The composition of plasmid-associated ARGs across samples. ARGs are colored by drug class subtype. D. The ARG host phyla across samples. ARGs are grouped by subtype (y-axis) and hosts are grouped by phyla (x-axis). The size of dots represents the relative abundance of ARGs corresponding to the specific subtype and host phylum. Dot colors indicate sampling location.

ARGs were efficiently removed via conventional activated sludge treatment followed by chlorination disinfection indicated by a 93.6% removal rate based on the relative abundance of total ARGs across the WWTP (Figure 5.3A). The removal rate via activated sludge treatment (91.8%) was comparable to those reported in previous studies

(Mao et al., 2015b; Y. Yang et al., 2014b; Yin et al., 2022). The chlorination process further removed chromosomal ARGs, but slightly increased the relative abundance of plasmid-associated ARGs, leading to a limited removal of total ARG relative abundance (21.0%) from the secondary effluent (Figure 5.3A). Although the role of chlorination remains under debate with respect to its impact on antibiotic resistance (Lin et al., 2016), several studies have shown that chlorination has a limited or even negative effect on the removal of ARGs from secondary effluent (X. Cheng et al., 2021; S.S. Liu et al., 2018; Yuan et al., 2015, 2019).

To further understand the dynamics of resistomes across the treatment processes, the composition of ARGs on chromosomes and plasmids was assessed separately with respect to ARG subtypes (Figure 5.3B). Chromosomal and plasmid-associated ARGs shared 16 ARG subtypes, whereas ARGs encoding resistance to fosfomycin and mupirocin were only detected on chromosomes, and ARGs conferring resistance to colistin and glycopeptide were only found on plasmids. The distribution of ARGs across chromosomes and plasmids was likely influenced by their resistance mechanism. Those ARGs causing antibiotic inactivation or replacement, or protection of the antibiotic's target, tend to be more frequently associated with plasmids than chromosomes, while ARGs associated with efflux pumps often appeared to be located on chromosomes (Figure S5.3). This distribution pattern is consistent with a recent study investigating the distribution of ARGs across chromosomes and plasmids in major groups of Enterobacteriaceae (Y. Wang et al., 2022).

In general, the ARG subtypes present in secondary effluent and final effluent samples were a subset of those in influent samples (Figure 5.3B&C). However,

chromosomal ARGs (Figure 5.3B) demonstrated a less consistent composition profile across the treatment processes as compared to plasmid-associated ARGs (Figure 5.3C). As for the chromosomal ARGs, a spike of rifamycin resistant genes (i.e., *rpoB2*, *RbpA*, and *efpA*) in secondary effluent was likely attributed to the growth of their putative host Actinobacteria (Figure 5.3D), whose relative abundance increased substantially in secondary effluent as compared to in the influent (data not shown). Similarly, the fraction of multidrug resistance genes (MDRs) increased in the secondary effluent (Figure 5.3B), and these MDRs were also carried by Actinobacteria (Figure 5.3D). The growth of Actinobacteria bacteria, which are common aerobes (Barka et al., 2015), was likely the result of the presence of high concentration of dissolved oxygen in the activated sludge treatment process. In contrast, the relative abundance of the obligate anaerobic Bacteroidetes and the facultatively anaerobic Firmicutes decreased in the secondary effluent (data not shown). Consequently, associations between Bacteroidetes or Firmicutes bacteria with ARGs were only observed in the influent samples (Figure 5.3D). Proteobacteria was observed to be the predominant host phylum for ARGs across the wastewater treatment processes (Figure 5.3D), which is consistent with previous studies (Azli et al., 2022; Gu et al., 2022; Z. Liu et al., 2022). Results indicated that the shift of the microbial community in response to the growth or decay of certain phyla drives the resistome across the WWTP (Figure 5.3B&C).

WWTP influent and effluent hosts were similar at the phylum level, as shown by both epicPCR (Figure S5.4) and long-read sequencing (Figure 5.3D, Figure S5.4), which is consistent with a study that used Nanopore sequencing for ARG host detection in WWTP influent and activated sludge (Dai et al., 2022). However, at the species level,

ARG hosts in the WWTP effluent were not entirely a subset of those in the WWTP influent due to the emergence of new hosts in the effluent (Figure S5.4). To gain a deeper understanding of the mechanisms responsible for the removal and selection of ARG hosts by different wastewater treatment unit processes, future research should focus on understanding the relative importance of horizontal gene transfer versus vertical propagation of ARGs via the growth and decay of ARG host, as well as the impact of environmental and operational variables on ARG propagation mechanisms (Barancheshme & Munir, 2018).

5.3.5 ARGs associated with pathogens and mobile genetic elements were present in the final effluent

We narrowed our focus on ARGs that were most likely to pose a public health risk using the following criteria: 1) present in the final effluent, 2) associated with an MGE, and 3) associated with a pathogenic host species. Given the strong performance of long-read sequencing (i.e., high detection sensitivity on resistomes, hosts and MGEs), we performed this analysis using information obtained via long-read sequencing results. Diverse subtypes of ARGs were found to be present in the final effluent (Figures 5.3B&C). Among all ARG-carrying reads in the effluent, 41.3% of them contained MDRs. We next focused on ARGs associated with pathogens. Not surprisingly, the influent contained high abundances of enteric bacteria and as well as a diverse array of ARGs and ARG-carrying pathogens (Figure 5.4). Interestingly, ARG-carrying pathogens that were detected in the secondary effluent, but not found in the influent, were mainly *Mycobacterium* species carrying rifamycin-resistance genes (Figure5.4).

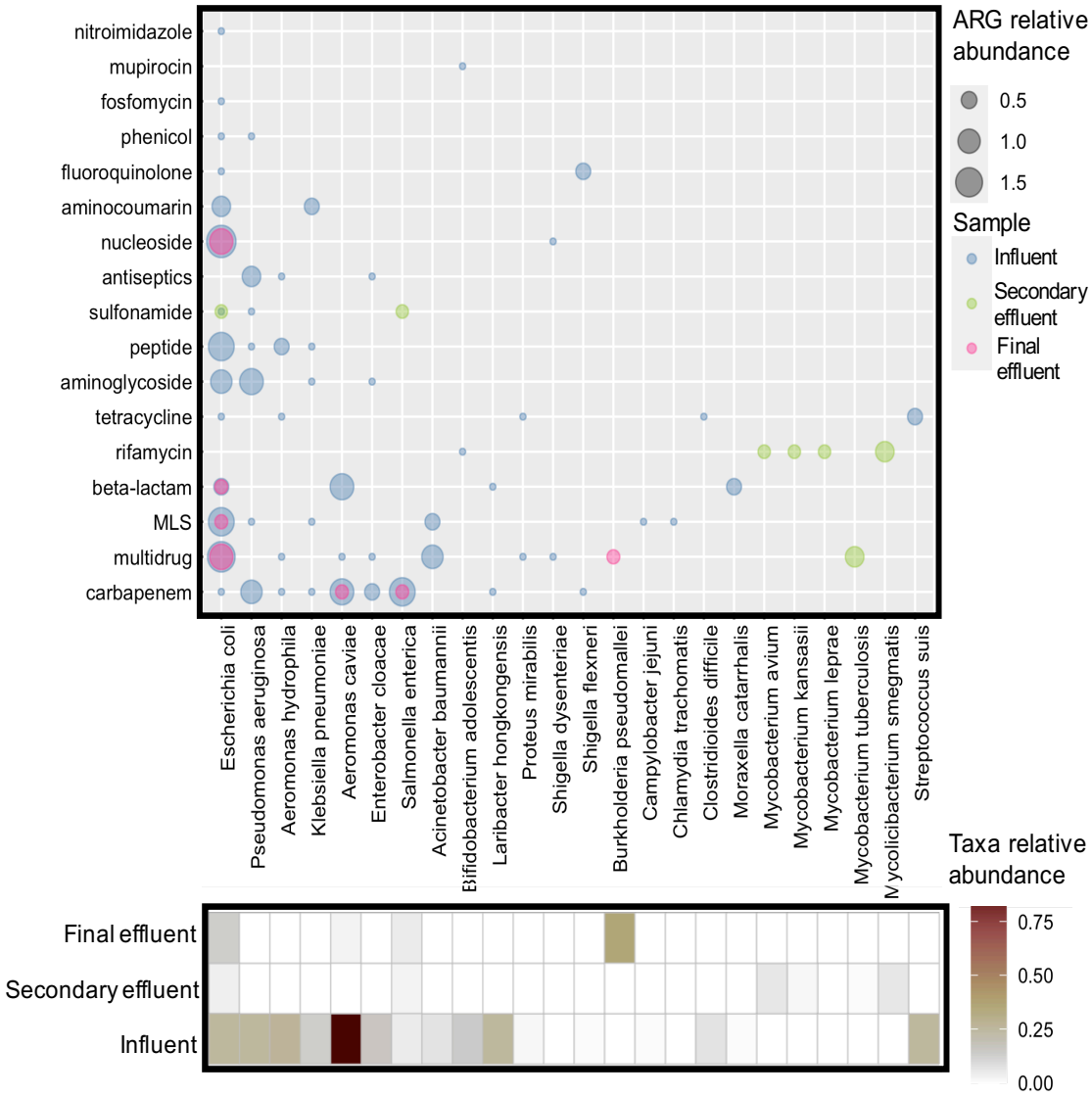


Figure 5.4. ARG-carrying putative pathogens detected in the influent, secondary effluent, and final effluent. ARGs are grouped by drug class subtype on y-axis; pathogenic species are shown on the x-axis. Dot size indicates the relative abundance of ARGs. Dot color indicates the sample location. The heatmap shows the taxa relative abundance of pathogenic species in each sample.

Mycobacterium is ubiquitous in wastewater and activated sludge and is considered a scavenger of insoluble compounds in wastewater (F. Guo et al., 2019; Radomski et al., 2011). One of the detected putative *Mycobacterium* pathogens, *Mycobacterium tuberculosis* (TB), was found to carry *efpA* which encodes an efflux pump system capable of extruding the isoniazid to the exterior of the cell (Rodrigues et al., 2012). This is

particularly concerning because isoniazid is a drug commonly used in TB therapy. However, ARG-carrying *Mycobacterium* species were not found in the final effluent, indicating its effective removal in the disinfection process. *Burkholderia pseudomallei*, which can cause the disease melioidosis (Draper et al., 2010), was an ARG host detected in the final effluent, but not in the influent or secondary effluent. It was associated with *MuxB*, a resistance-nodulation-cell division (RND) antibiotic efflux pump gene that significantly reduces susceptibility to macrolide, beta-lactams, and fluoroquinolones in bacteria.

Overall, six out of seven ARG-carrying putative pathogens present in the final effluent were also detected in the influent (Figure 5.4). It is worth noting that the relative abundance of ARGs carried by *E. coli*, especially those encoding resistance against multidrug, beta-lactam, and nucleoside, were persistent across the entire treatment process (Figure 5.4). In addition, the estimated relative abundance of total *E. coli* decreased significantly from influent to final effluent (Figure 5.4). Together, these results suggest that the chlorination process may have selected for resistant *E. coli*. Previous studies have also observed that multidrug-resistant E. coli was persistent during wastewater treatment (Aslan et al., 2018) and was capable of escaping the oxidation by disinfectants (Mounaouer & Abdennaceur, 2016).

Other than ARG-carrying pathogens, ARGs associated with MGEs can also contribute to public health risks because they can potentially be transferred to pathogenic hosts. To assess whether plasmid-associated ARGs were likely to be mobilizable, we investigated the mobility of the ARG-carrying plasmids by screening for and classifying the mobilization genes (MOB) on the reads. Results showed that all ARG-carrying

plasmids found in secondary effluent (containing 33 ARG-carrying plasmid reads) and final effluent (containing 56 ARG-carrying plasmid reads), as well as most (966 out of 970) ARG-carrying plasmids in influent, were classified as nonmobilizable plasmids due to the lack of a MOB. Only four out of 970 ARG-carrying plasmid reads in influent were found to carry MOB genes. However, although long-read sequencing generated greater length reads as compared to short-read assembled contigs, the average length of ARG-carrying, plasmid-associated reads was 4,885 bps. This suggests incomplete plasmids were assembled and thus may not have contained information needed to call mobility for a plasmid. For example, the length range of 14 representative ARG-bearing conjugative plasmids isolated from WWTPs was reported to be 35,925-290,014 kbps (Che et al., 2022), much longer than the plasmid-associated read length. Therefore, we cannot draw a solid conclusion regarding the mobility of plasmids given that the relatively short plasmid-associated reads captured incomplete plasmid sequences.

A variety of MGEs including *IntI1*s, recombinases, transposases, and integrases were frequently observed in the effluent samples (Table S5.7), which suggests they may be involved in the HGT of ARGs among bacteria (Gillings et al., 2015; Knapp et al., 2008; L. Ma et al., 2017b). Recent studies revealed the striking prevalence of insertion sequences (IS) in resistant pathogens and the relatively consistent linkages between certain IS and specific ARGs across highly diverse bacterial genotypes, indicating the role of IS in mediating the HGT of these ARGs (Che et al., 2021, 2022). Similarly, we also found diverse IS families that were associated with ARGs across samples. Particularly, the IS6 family transposase was found to be frequently associated with two macrolide resistance genes across samples, namely, *msrE* and *mphE* (Table S5.7). In

addition, this specific *mphE*/*msrE*-the IS6 family transposase association was found on a conjugative plasmid equipped with the T4SS and MOBQ machinery, highlighting its potential for HGT via the interaction of IS and a conjugative plasmid.

5.4. Conclusion

The bacterial host and genetic context of an ARG present in our water and wastewater systems is critical to assessing its potential risk to human health. Specifically, ARGs of highest priority for further study are those hosted by pathogenic bacteria and/or with the potential to be horizontally transferred to pathogens (i.e., associated with an MGE). In this study, we evaluated and compared long and short-read metagenomic sequencing-based methods, as well as a targeted method (epicPCR) for identifying ARG hosts and associations with MGEs. We found that long-read sequencing outperformed short-read sequencing by generating a higher relative abundance of ARGs, especially of ARGs associated with MGEs, as well as a more diverse ARG host profile. Moreover, long-read sequencing generally yielded a greater number of reads supporting ARG-host linkages compared to the number of contigs assembled from short reads. EpicPCR outperformed long-read sequencing in terms of the breadth of hosts detected for three ARG targets (*ermB*, *sul1*, and *tetO*), however, it does not provide any additional contextual information (e.g., whether the ARG is associated with an MGE). When we applied these methods to understand ARG host dynamics across the WWTP, we observed consistent trends using long-read sequencing and epicPCR. Overall, the linkages of ARGs and host phyla in the WWTP effluent resembled those in the WWTP influent. However, at the species level, ARG hosts in the WWTP effluent were no longer a subset

of those in the WWTP influent, which reinforces the need for more and longer-term surveillance of emerging effluent ARG hosts, and the importance of understanding the mechanisms of removal and selection of ARGs and ARG hosts across treatment.

These results suggest that for environmental surveillance, long-read sequencing has many advantages as a tool for ARG detection and host tracking due to its high sequencing efficiency and because it does not require assembly. However, if any clinically-relevant ARG targets, such as MCRs (colistin resistance genes), are of particular concern to public health, epicPCR assays could be developed and applied to capture a more comprehensive host profile to complement routine metagenomic screening. Future studies should focus on evaluating standardized methods for wastewater-based surveillance of antibiotic resistance, developing guidelines for better reproducibility, and establishing a risk estimation framework for ARGs in the environment.

5.5. Acknowledgements

This research was supported by funds from the National Institute of Food and Agriculture (grant no. 2016-68007-25044), the National Science Foundation (CBET 1805901 and 2029025, EF-2126387), a Johnson & Johnson WiSTEM2D 2D award, seed funds from Rice University, and a NIH grant from NIAID (P01-AI152999). We thank the operators at the West University City Place WWTP for helping with accessing wastewater samples.

Data Availability

The metagenomic long-read sequencing and short-read sequencing analysis pipeline and the epicPCR analysis pipeline are deposited at https://gitlab.com/treangenlab/wasterwater_arg_metagenomics. All sequencing data can be found at NCBI SRA (project accession number: PRJNA842493).

Chapter 6 Concluding Remarks

6.1 Summary and Conclusions

Antimicrobial resistance (AMR) poses an ongoing threat to public health because it leads to higher medical costs, prolonged hospital days, and increased mortality. Clinical AMR surveillance efforts only capture a fraction of AMR circulating in a community, and one of the major challenges in combating the spread of AMR is the complexity associated with its spread. The spread of AMR is complex because it is driven by ubiquitous bacteria and genetic elements adapted to traverse, persist, and proliferate in a wide array of hosts and environments. One significant point source of AMR in the environment is wastewater. Monitoring wastewater for AMR can provide information in two important ways: (1) influent (sewage, untreated wastewater) represents a pooled sample of all individuals in the community can be used to characterize circulating AMR in the community and inform public health decision-making, and (2) effluent AMR reflects the ability of WWTPs to remove and mitigate the dissemination of AMR in downstream environments. This dissertation examines critical methodological choices associated with wastewater monitoring, in the context of both AMR and SARS-CoV-2, the virus responsible for the COVID-19 pandemic. The results of this work contribute to the establishment of routine WBE for the protection of public health, and advancing our understanding of the risks associated with AMR in the environment.

Elucidated impact of wastewater sampling design and concentration methods on quantitative assessments of AMR in wastewater. Wastewater sampling method (i.e., grab versus composite sampling), can influence the observed concentrations of analytes (in this case, ARGs) in wastewater samples. Furthermore, concentrating target analytes in

wastewater samples prior to nucleic acid extraction and quantification can also impact the form and abundance of the recovered analytes. Currently, the majority of studies on AMR in wastewater have applied grab sampling, likely due to convenience and minimal resources and time as compared to composite sampling. In Chapter 2, we found that the concentrations of each gene fluctuated in grab samples collected throughout the day. This diurnal variation led to significant differences among the intra-day hourly instantaneous removal rates which ranged from 0.5-1.6 logs in winter and 0.9-2.7 logs in summer for each target gene. Additionally, the removal rates calculated based on 24-hour composite samples were approximately equal to the median of the instantaneous removal rates, indicating composite samples provide a more balanced thus representative snapshot of removal trends. Furthermore, the diurnal variation of the concentrations of target genes in the secondary and final effluent samples were more significant than those in the influent samples, indicating the diurnal variation of the target gene concentrations can be amplified by the wastewater treatment train. Our study confirms the importance of using composite rather than grab samples to monitor ARGs for WBE. In addition, the removal rates of target genes across WWTPs should also be assessed via composite samples.

The impact of wastewater sample concentration method on AMR characterization was discussed in Chapter 3. A widely used method to concentrate wastewater samples involves a filtration step using 0.22 μm filters, which can cause cell-free ARGs to be overlooked. This is because cell-free ARGs can pass through the filter pores, and are thus discarded with the flow-through, while cell-associated ARGs are intercepted by the filters. We investigated the profile of ARGs in the effluent of a novel biotechnology for wastewater reclamation and water reuse, an anaerobic membrane bioreactor (AnMBR)

treating high strength mixture of cattle manure and domestic wastewater. We found that the effluent ARG and mobile genetic element (MGE) reservoir shifted from cell-associated to cell-free DNA with the increased organic loading rate. This finding would not have been discovered if the sample concentration and DNA extraction protocols had not captured both cell-associated and cell-free fractions of DNA. Taken together, these two chapters address that methodological choices with regards to wastewater sampling and concentration impacts the results and conclusions drawn from AMR WBE. Thus, method standardization should be established to ensure the validity and reproducibility of WBE results and conclusions generated across laboratories to generate actionable information from AMR monitoring.

Evaluated two widely applied targeted methods, RT-ddPCR and targeted amplicon sequencing, for monitoring SARS-CoV-2 mutations in wastewater. SARS-CoV-2 RNA in wastewater can be degraded and fragmentized and is commonly found in trace amounts. Therefore, for molecular analysis, targeted methods are widely applied because they involve an enrichment step (usually based on PCR) to improve detection sensitivity. Targeted amplicon sequencing is considered a powerful tool to enable comprehensive screening of all potential mutations without any prior knowledge. However, it is unclear how quantitative is this approach as compared to the gold standard RT-qPCR or RT-ddPCR methods in the context of wastewater monitoring. In Chapter 4, we directly compare a targeted amplicon sequencing approach (based on ARTIC v3) with RT-ddPCR quantification for the detection of five characteristic mutations of variants of concern (VoCs) using 547 wastewater samples. When we observed positive mutation detections by

RT-ddPCR, 42.6% of the detection events were missed by targeted amplicon sequencing, due to negative detection or the limited read coverage at the mutation position. Further, when targeted amplicon sequencing reported negative or depth-limited mutation detections, 26.7% of those events were instead positive detections by RT-ddPCR, highlighting the relatively poor sensitivity of targeted amplicon sequencing. No or weak associations were observed between quantitative measurements of target mutations determined by these two methods. These findings caution the use of quantitative measurements of SARS-CoV-2 variants in wastewater samples determined solely based on targeted amplicon sequencing.

Revealed the merits and limitations of metagenomic sequencing for tracking ARG hosts in WWTPs and contrasted them against a targeted method, epicPCR. There is growing interest in deploying unbiased, non-targeted metagenomics sequencing to monitor AMR in WWTPs. In particular, the development and advancement of third generation sequencing that can generate significantly longer read lengths, enables the contextual genetic information of ARGs to be deciphered and can reveal ARG-host information. However, long-read metagenomic sequencing has not been explicitly compared short-read metagenomics for the purposes of ARG-host tracking. In addition, it is unknown how sensitive metagenomics sequencing is as compared to targeted methods (e.g., epicPCR) for ARG host tracking. Identifying ARG hosts is crucial for understanding host-ARG interactions and the potential risks posed by pathogenic resistant bacteria. Chapter 5 shows that despite its significantly lower sequencing depth, long-read sequencing via the Nanopore sequencing platform outperformed short-read sequencing with higher sensitivity for detecting ARGs, especially for ARGs associated with mobile genetic elements (MGEs).

In addition, long-read sequencing consistently revealed a wider range of ARG hosts compared to short-read sequencing. Both long-read sequencing and epicPCR detected new host species that emerged in the treated effluent. However, the host range detected by long-read sequencing only represented a subset of the host range detected by epicPCR. Based on these findings, we recommend 1) using long-read sequencing for routine wastewater surveillance, 2) using epicPCR to obtain a high-resolution host range of clinically relevant ARGs, and 3) performing long-term surveillance on specific treatment compartments within a WWTP to understand the emergence of new hosts.

6.2 Significance and Implications

This dissertation contributes to achieving scientific consensus on method selection for characterizing disease indicators in wastewater samples. Specifically, guidelines of method selection regarding upstream and downstream analyses are provided.

Guidelines for the upstream workflow of AMR characterization in wastewater (sampling and concentration). First, unintended bias could occur with grab sampling, emphasizing the issue of using grab samples to indicate ARG and MGE concentration and loading for WBE, or to calculate ARG and MGE removal rates across the WWTPs. Composite sampling should be conducted at all sample collection locations (i.e.., influent, secondary effluent, and final effluent) across any studied WWTP to ensure the representativeness of those collected samples. Ensuring the representativeness of the samples collected across the WWTP is important, because the samples collected from each of the sampling locations reflect the efficacy of each wastewater treatment process unit in attenuating AMR. It may also help answer the question that which AMR elements are able to escape wastewater

treatment processes and persist in the treated effluent. Furthermore, making composite sampling as the standard sampling method for practicing WBE facilitates data sharing and comparison among research groups. Second, the widely applied, filtration-based sample concentration method is designed to recover the cell-associated fraction of ARGs rather than the cell-free. Our research studied ARGs and MGE in the treated effluent of an AnMBR which was operated as a candidate for wastewater reclamation and water reuse. Our results pointed out that several clinically relevant ARGs can be more abundant in cell-free DNA rather than in cell-associated DNA under certain operating conditions (i.e., high organic loading rate). This finding is significant because it shows that to advance future water reuse applications (e.g., irrigation) using the effluents of wastewater treatment systems, the sample concentration and DNA extraction protocols should be adapted to recover both cell-associated and cell-free ARGs in the effluents in order to achieve a comprehensive evaluation on the potential risks associated with AMR dissemination.

Guidelines for downstream workflow of characterizing disease indicators in wastewater (molecular analysis techniques) – combining the advantages of different analytical methods. Third, we provided guidelines for characterizing SARS-CoV-2 in wastewater to enhance our ability to fight the COVID-19 pandemic using WBE. RT-ddPCR or RT-qPCR should be applied for quantitative analyses due to the great sensitivity and consistency of detection. In addition, RT-ddPCR and RT-qPCR generally have much shorter result turnaround time compared to sequencing, which is critical for real-time public health response. With knowledge of unique mutations associated with each VOC, it is possible to detect signatures of low levels of VOCs in wastewater samples that may contain a mixture

of variants. For example, allele-specific and multiplex-compatible RT-qPCR assays targeting mutations for quantitative detection and discrimination of the Delta, Delta plus, Kappa and Beta variants of SARS-CoV-2 in wastewater were developed and validated. Despite its lower sensitivity and qualitative nature, sequencing still has a clear advantage of being more comprehensive, not limited by a priori knowledge of the target mutations, and enables the discovery of cryptic lineages and emerging lineages of concern. This can be critical for early detection of variants when the availability of primers and probes is limited or delayed due to supply chain challenges. In practice, WBE systems can benefit from coupling sequencing with quantitative analyses such as RT-ddPCR or RT-qPCR to achieve a comprehensive picture of circulating mutations (using sequencing), and sensitive, quantitative information on variant-associated mutations (using RT-qPCR/RT-ddPCR). Finally, for WBE of AMR, long-read metagenomic sequencing has many advantages in ARG detection and host tracking. This is because long-read metagenomic sequencing demonstrates high sequencing efficiency (i.e., rapid result generation in real-time and no need for *de novo* assembly) and a great ability to capture a comprehensive ARG host profile as compared to conventional NGS (short-read). These results highlight that future WBE should consider using long-read metagenomic sequencing for comprehensive screening of AMR in the community. However, if any clinically relevant ARG targets, such as MCRs (colistin resistance genes), are of particular interest according to public health, epicPCR assays could be developed and applied to capture a more comprehensive host profile to complement routine metagenomic screening.

6.3 Suggestions for Future Research

6.3.1. Validating methods for WBE of AMR to achieve standardized routine monitoring

This dissertation includes an evaluation of methods including sampling, sample concentration, and targeted and non-targeted molecular analysis techniques for WBE of AMR. All methods have their merits and limitations, and these methodological choices impact the results and conclusions monitoring efforts. Ensuring the validity of AMR wastewater monitoring data is necessary to assess the risks associated with AMR in communities and track the emergence and evolution of new resistance mechanisms. Here we propose three areas for future research on WBE of AMR:

Advancing long-read metagenomics for ARG host tracking. Metagenomics can be applied as an unbiased method for broad screening of AMR in wastewater, thus recommended for future WBE (Liguori et al., 2022; Prieto Riquelme et al., 2022)**.** However, using long-read metagenomic sequencing for ARG host tracking in wastewater has not been thoroughly evaluated, partly because of its limited application for the detection of pathogens and ARGs in wastewater samples. It is important to systematically benchmark long-read metagenomic sequencing for ARG host tracking, because (1) long-read metagenomics poses several irreplaceable advantages in tracking ARG hosts for WBE compared to short-read metagenomics (e.g., fast result turnaround time, high detection sensitivity), as demonstrated by Chapter 5 of this dissertation; and (2) the relatively high error rate associated with long-read sequencing technology is still an unresolved issue, limiting its applicability for estimating human health risks. Further research is needed to advance long-read metagenomics. First, to track ARG hosts in wastewater metagenomes,

most bioinformatic tools available for metagenomic data have been developed, applied, and evaluated using short-read sequencing data. There are a variety of bioinformatician tools used for metagenomic analysis, serving multiple functions including but not limited to read assembly, sequence alignment, target annotation, and taxonomical classification. For example, Kraken2, a fast classification tool which uses k-mers to assign a taxonomic labels in form of NCBI Taxonomy to the sequence, works efficiently and accurately for analyzing sequences generated by metagenomics using short-read NGS (Wood et al., 2019). However, previous studies have shown that Kraken2 failed to achieve accurate taxonomical classifications at species-level or strain-level for metagenomes sequenced by long-read sequencing methods (e.g., Nanopore) (Dilthey et al., 2019; Pearman et al., 2020). Nevertheless, Kraken2 was recently used for analyzing long reads sequenced via Nanopore platform to identify species-level pathogenic ARG hosts in wastewater (Y. Yang et al., 2022). To the best of our knowledge, no studies have systematically benchmarked the bioinformatic toolbox for long-read metagenomic sequencing specifically for ARG host tracking. The challenge in benchmarking bioinformatic tools specifically for metagenomics is due to the unknown priori knowledge on the exact genetic composition of these complex environmental metagenomes. That is why most validations were based on assessing host-associated microbiomes (Abdelrazik et al., 2021; Fosso et al., 2017; Langille et al., 2013). Fuhrmeister et al. used real-time PCR as a benchmark to assess the sensitivity and specificity of long-read metagenomics via Nanopore platform for the detection of pathogens in wastewater samples (Fuhrmeister et al., 2021). However, while this assessment compared the detections of known pathogens, a similar approach has not been applied to detect hosts of ARGs, which are far more complicated to benchmark due

to the challenges in confirming associations between host taxonomic marker genes and ARGs. Che et al. assessed the ability of Nanopore sequencing to characterize the relative abundance profile of ARG subtypes in wastewater metagenomes using Illumina sequencing. In addition, they compared the taxonomic compositions of the multidrug resistant cultures isolated from wastewater at species level between Illumina and Nanopore sequencing platforms (Che et al., 2019a)**.** However, again, these results were not respect to direct ARG host identification in wastewater metagenomes.

We recommend future research that (1) evaluates or develops novel bioinformatic tools to conduct ARG host tracking via long-read metagenomic sequencing in wastewater; (2) performs a quantitative assessment of the false positive and false negative rates of ARG host detection via long-read metagenomic sequencing using appropriate benchmarks. A good starting point is to design appropriate internal controls. For instance, a recent study spiked the sequenced wastewater metagenomes with *in silico* resistant pathogenic genomes with a gradient of genome coverage, and tracked the detections of internal standards after performing different analysis pipelines (Deshpande & Fahrenfeld, 2022)**.** Furthermore, a recent work found quantitative approaches including experimental procedures to incorporate microbial load (in that case, cell count) variation in downstream analyses significantly improved the sensitivity of detections and reduced false-positives for metagenomic data analysis (Lloréns-Rico et al., 2021). These aforementioned studies indicate that applying *in silico* control and leveraging experimental data for computational analyses hold the promise of enabling benchmarking environmental metagenomics.

Validating novel targeted methods for ARG host tracking. Currently, targeted methods such as epicPCR and Hi-C (based on proximity ligation cross-linking) demonstrate great

sensitivity in identifying diverse ARG hosts in environmental samples with high taxonomic resolution. However, their lengthy protocols and the technical difficulties may easily impair the quality of results, hampering their wide applicability (Roman et al., 2021). In addition, these novel targeted methods have not been appropriately benchmarked. For instance, epicPCR for environmental host tracking has been thoroughly validated using internal controls. Hosts identified by epicPCR are those detected across sample replicates, as performed by Chapter 5 of this dissertation and a couple of previous publications (Roman et al., 2021; Wei et al., 2021). In the case of Hi-C for identification of ARG hosts, the quality of a single Hi-C metagenome-assembled genome (MAG) was validated by performing conventional bacterial culture and sequencing of an *E. coli* isolated from the same fecal sample that generated the *E. coli* Hi-C MAG (Kalmar et al., 2022). Therefore, future research should conduct a more comprehensive validation for these targeted methods to advance their applications for tracking ARG hosts in environmental microbial communities.

6.3.2. Making WBE results actionable for public health decision-making

Wastewater monitoring has emerged as a powerful tool that complements existing public health surveillance systems. In additions to improving methods related to the generation of wastewater monitoring data, research is needed to advance the methods associated with their interpretation. We offer the following recommendations for future research related to informing public health decision-making using WBE.

Detect yet-unknown pathogens via environmental surveillance to proactively protect public health. Ideally, WBE should be a tool to not only respond to and track disease

outbreaks, but also preempt future pandemics. To fulfill this ambitious goal, we need to tackle the "black box" of wastewater metagenomic data and examine cryptic disease transmission. It is necessary to ask the questions including but not limited to, how we can: (1) delineate the genetic determinants that are directly associated with a wide spectrum of relevant diseases, (2) validate the detections of those genetic determinants, and (3) predict future pandemics based on the detection of genetic determinants. Balaji et al. have contributed to answering the first question. by developing SeqScreen, a powerful tool which accurately identifies sequences of concern (SoCs) with respect to emerging pathogens from environmental metagenomic datasets (Balaji et al., 2022). This tool is novel as it addresses pathogenicity of genomic sequences based off of not only searching for known taxonomic markers (which is often a poor proxy for pathogenicity), but also screening for functional pathogenic labels associated with proteins encoding disease functions. Future studies may adapt the methodology of this pipeline by curating a representative database containing sequences of disease functions (i.e., virulence factors that are real threats to the host) (Godbold et al., 2022), and updating this database timely to facilitate proactive disease monitoring. Apart from leveraging this database, we should also develop computational tools that can capture the associations between sequences. This is because certain public health threats are not contributed by a single functional gene but rather a cluster of co-occurring genetic determinants. For example, recently it is reported that the synergy between certain insertion sequences (IS) and several conjugative plasmids is responsible for facilitating the HGT of certain clinically relevant ARGs between pathogenic bacteria (Che et al., 2021, 2022). Monitoring that can detect this synergy and associations between ARGs, ISs, and conjugative elements can help elucidate the clinically

relevant ARGs that are of high risk to be transferred among bacteria in a community. Another recommendation in identifying cryptic AMR is to combine computational mining with experimental verifications. For example, a novel functional metagenomics approach successfully identified a new integron-borne aminoglycoside resistance gene present in clinical pathogens via screening wastewater-impacted environmental metagenomes (Böhm et al., 2020). Then the next question is how to make the process of cryptic AMR screening more efficient, given the fact that functional metagenomics requires lengthy workflow (e.g., functional gene isolation and cloning).

6.3.3. Closing the knowledge gap of risk estimation associated with environmental AMR

As shown by the results of Chapter 2, 3 and 5, ARGs and ARG hosts are ubiquitous in treated effluents. Many previous studies have investigated AMR elements in treated effluents and effluent-receiving environments. However, further research is needed to quantitatively estimate the tangible human health risks associated with environmental AMR. The traditional framework to conduct risk estimation, namely, quantitative microbial risk assessment (QMRA), is not able to be directly used to evaluate risks associated with environmental AMR because of the paucity of dose-response data and the unknown level of acquisition of AMR via HGT in the environment (Garner et al., 2021) . Recently, Schoen et al. used QMRA to estimate the annual risks of colonization, skin infection, bloodstream infection (BSI), and disease burden from exposures to antibiotic-resistant and susceptible *Staphylococcus aureus* (Schoen et al., 2021) . This represents a promising framework and example of how to conduct QMRA for water

reuse applications. Future research can work on refining the framework of QMRA by generating experimental data to disclose the values of those variables (e.g., HGT frequency under specific environmental conditions for each specific strain of interest). However, estimating risks using QMRA only works for viable and culturable pathogenic strains. To estimate the risks associated with an environmental resistome, that includes extracellular DNA (as shown by conclusions of Chapter 3 and a considerable number of previous publications), additional information may be needed to assess the actual risks of HGT (transformation) of the extracellular ARG residues and the risks posed by non-culturable resistant pathogens.

Chapter 2 Appendix: Instantaneous ARG removal rates in across wastewater treatment plants are not representative due to diurnal variations

1. Supplementary methods

1.0 Sampling campaign design

Table S2.1. The instantaneous plant flow rates in winter and summer sampling campaigns.

Time	Flow rate (MGD)	
	Winter	Summer
10:40	1.45	1.0405
12:40	1	1.0285
14:40	0.83	0.95
16:40	0.9	0.811
18:40	1.3	0.7645
20:40	1.5	1.085
22:40	1.42	1.2085
0:40	1.05	0.791
2:40	0.5	0.51
4:40	0.4	0.248
6:40	0.7	0.229
8:40	1.8	1.164

Table S2.2. Description of the methodology used to obtain the hourly samples in the two field campaigns (winter and summer) and sampling conditions.

Time of sampling	Sample aggregation	Location of sampling	Number of samples	Frequency	Ambient temperature
12/18/2019 - 12/19/2019 (winter, dry weather, weekday)	Grab sample for 24 hours	WWTP influent (after screening), secondary effluent (after contact stabilization process and clarification), final effluent	n=12 for each sampling location	Collected every 2 hours	2 – 4 ℃
5/31/2022 - 6/1/2022 (summer, dry weather, weekday)			n=24 for each sampling location (two biological replicates)		26 – 30 ℃

		(after chlorine disinfection)			

Table S2.3. (A) Ammonia-N (NH_3-N) and (B) total COD concentrations and removal rates across the sampling locations in winter and summer sampling campaigns

A.

Time	NH_3-N Concentration (mg/L)			Overall removal rate (%)	Season
	Influent	Secondary effluent	Final effluent		
10:40	48.29	1.70	0.71	98.54	Summer
12:40	46.14	2.62	0.75	98.37	Summer
14:40	38.50	3.61	1.52	96.05	Summer
16:40	37.71	4.56	2.60	93.10	Summer
18:40	24.42	4.34	2.85	88.31	Summer
20:40	35.79	5.38	2.45	93.15	Summer
22:40	23.17	3.93	2.23	90.37	Summer
0:40	26.96	4.08	2.30	91.47	Summer
2:40	19.72	3.55	1.65	91.63	Summer
4:40	26.06	3.10	0.88	96.64	Summer
6:40	29.56	1.53	0.61	97.92	Summer
8:40	28.43	2.02	1.66	94.16	Summer
10:40	50.61	8.20	0.06	99.89	Winter
12:40	37.70	9.54	0.01	99.97	Winter
14:40	43.47	10.30	0.30	99.30	Winter
16:40	44.55	8.23	0.06	1.00	Winter
18:40	46.29	5.66	0.34	99.26	Winter
20:40	32.24	7.73	0.29	99.10	Winter
22:40	69.42	9.29	0.15	99.78	Winter
0:40	33.50	8.10	0.18	0.99	Winter
2:40	31.50	4.59	0.17	99.46	Winter
4:40	36.83	3.97	0.07	99.81	Winter
6:40	30.87	5.12	0.07	99.76	Winter
8:40	47.79	6.72	0.11	99.77	Winter

B.

Time	Total COD Concentration (mg/L)			Overall removal rate (%)	Season
	Influent	Secondary effluent	Final effluent		
10:40	412.11	26.11	22.38	94.57	Summer
12:40	553.62	29.25	23.85	95.69	Summer
14:40	530.71	27.94	23.53	95.57	Summer
16:40	547.07	32.85	24.67	95.49	Summer
18:40	508.63	35.47	27.13	94.67	Summer
20:40	443.19	40.54	24.51	94.47	Summer
22:40	525.80	27.13	26.31	95.00	Summer
0:40	419.47	29.06	25.49	93.92	Summer
2:40	344.21	28.27	25.49	92.59	Summer
4:40	283.68	39.07	27.78	90.21	Summer
6:40	277.96	28.92	26.31	90.53	Summer
8:40	383.48	25.65	26.31	93.14	Summer
10:40	494.81	9.82	16.68	96.63	Winter
12:40	524.72	23.83	48.19	90.82	Winter
14:40	372.26	34.04	Missing data	Missing data	Winter
16:40	244.61	19.16	18.87	92.29	Winter
18:40	521.07	22.08	21.49	95.87	Winter
20:40	352.57	29.52	44.84	87.28	Winter
22:40	421.14	6.47	32.00	92.40	Winter
0:40	368.62	18.58	20.91	94.33	Winter
2:40	378.83	6.91	52.57	86.12	Winter
4:40	283.27	61.61	34.19	87.93	Winter
6:40	358.40	6.61	58.70	83.62	Winter
8:40	470.74	16.10	24.27	94.85	Winter

1.1 Sample pre-treatment and storage for COD and NH_3-N measurement, the procedures for measuring COD and NH_3-N

Pre-treatment of the samples for measurement of COD involved (1) 50 mL influent, secondary effluent and final effluent samples were each acidified with concentrated sulfuric acid (95.0-98.0%) to lower the pH of the sample to less than 2.0; (2) the acidified samples were stored in -4 ℃ till COD measurement. For NH_3-N,10 mL of influent,

secondary effluent and final effluent samples were filtered using 0.45μm pore size syringe filters (Biomed Scientific). The samples were then stored in -4 ℃ till measurement. COD measurements were done through colorimetric analysis, according to Standard method 5220D (Standard Methods, 2005) with COD vial kits (CHEMetrics K7355). Ammonia measurement was done through the phenate method according to Standard method 4500-NH3 F (Standard Methods, 2005).

1.2 Sample filtration for DNA extraction

50 mL of influent sample, 350 mL of secondary effluent sample, and 500 mL final effluent sample were subject to on-site sample concentration. Each sample was slowly poured into a 6-head, Multi-Vac 610-MS Manifold (180310-01, Sterlitech) containing a pre-DI-washed 0.45 μM pore size, electronegative microbiological analysis HA filter (HAWG047S6, MilliporeSigma). A vacuum pump pulled the sample through the filter and was turned off after all filters were completely dry. Next, each membrane filter with retained biomass was submerged in 50% ethanol in a 2 mL round-cap microcentrifuge tube and stored in -20 ℃ for later DNA extraction.

All samples were kept in the storage conditions as previously mentioned during the field sampling. After the sampling was done, samples were transferred back to the Rice University lab to get processed within 30 minutes.

1.3 qPCR conditions and primer/probe sequences

The qPCR program consisted of an initial denaturation for 30 s at 95 °C, followed by 40 cycles of denaturation at 95 °C for 5 s, primer annealing and extension at 65 °C (for

sul1) or 60 °C (for the rest of the targets) for 30 s, and an optional melt-curve stage (for only SYBR green, non-probe-based assays) consisting 15 s at 95 °C, 60 s at 60 °C, and 15 s at 95 °C. All assays were having amplification efficiencies between 90 – 110% and regression coefficients (R^2) greater than 0.99 according to the standard curves. The details of primers/probes can be found in Table S2.2.

Table S2.4. Details of primers and probes used in this study.

Target	Forward primer (5' - 3')	Reverse primer (5' - 3')	Amplicon size (bp)	Reference for primer sequences	Probe (5' - 3')	Reference for probe sequence
16S rRNA	CGGTGAATACGTTCYCGG	GGWTACCTTGTTACGACTT	124	(Suzuki et al., 2000)		
AmpC	CCTCTTGCTCCACATTTGCT	ACAACGTTTGCTGTGTGACG	189	(S. Lee et al., 2022)		
*bla*NDM-1	CGCCATCCCTGACGATCAAA	CTGAGCACCGCATTAGCCG	214	(Luo et al., 2014)		
*bla*OXA-1	TATCTACAGCAGCGCCAGTG	CGCATCAAATGCCATAAGTG	199	(Kennedy et al., 2017)		
sul1	CGCACCGGAAACATCGCTGCAC	TGAAGTTCCGCCGCAAGGCTCG	163	(R. Pei et al., 2006)		
tet(W)	GAGAGCCTGCTATATGCCAGC	GGGCGTATCCACAATGTTAAC	167	(Aminov et al., 2001)		
IntI1 (clinical)	CGAACGAGTGGCGGAGGGTG	TACCCGAGAGCTTGGCACCCA	312	(Zheng et al., 2020)		
qnrA	AGGATTGCAGTTTCATTGAAAGC	TGAACTCTATGCCAAAGCAGTTG	138	(Colomer-Lluch et al., 2014)		
MCR-1	ATCCCATCGCGGACAATCTC	AGACCGTGCCATAAGTGTCA	177	(Mentasti et al., 2021)	56-FAM/ATCAACACA/ZEN/GGCTTTAGCACATAGCGAT/3IABkFQ/	This study
MCR-5	GTATCTCCATGGCATACCTTAC	GAAACAGGTGATCGTGACTTAC	149	This study	/5HEX/AGGTTTATG/ZEN/CCGACCAAGCCTG/3IABkFQ/	This study
MCR-10	GCAATAACCCGACGCTGAAC	GTAACGCGCCTTGCATCATC	133	(Mentasti et al., 2021)	TARMA/GCGTACAGCACTTTCCATTACGTGA/BHQ-2	This study

Table S2.5. The relative standard deviation (RSD) of the target gene concentrations in grab samples collected throughout the day in winter (A) and summer (B).

A. Winter

Target	RSD (%) in absolute concentration		
	Influent	Secondary effluent	Final effluent
AmpC	19.08	60.43	20.33
*bla*NDM-1	43.14	58.12	22.62
*bla*OXA-1	23.54	38.08	85.22
IntI1	28.68	45.45	37.10
MCR-1	37.06	99.58	163.07
MCR-10	35.80	84.90	143.79
MCR-5	38.83	73.95	131.25
qnrA	60.26	49.63	156.05

B. Summer

Target	RSD (%) in absolute concentration		
	Influent	Secondary effluent	Final effluent
AmpC	21.61	39.04	78.33
*bla*NDM-1	99.17	41.23	101.08
*bla*OXA-1	20.31	44.23	55.86
IntI1	37.08	30.58	82.29
MCR-1	52.62	45.86	90.19
MCR-10	35.05	35.13	62.66
MCR-5	34.35	33.39	68.96
qnrA	31.81	52.64	72.02
sul1	25.22	24.67	67.58
tet(W)	24.85	55.04	68.37

1.4 Calculation of 24-hr loading of ARGs and *IntI1*

To calculate the loading of target genes going into and coming out of the WWTP, we came up with the estimation method as described in Figure S5.1. In brief, we first calculated the instantaneous ARG loading rate R (Equation S1), then plotted the curve of R with respect to time (T). Thus, the area bound by the line of R and the axes T represents the total ARG loading within that certain period of time. For example, a 4-hr ARG loading from T=0 to T=4 hour should be the sum of the area of two trapezoids T_1 and T_2 (Equation S2, Figure S5.1). The loading removal is the difference between the influent trapezoid area and the effluent trapezoid area (Equation S3). Extended from this simple example, the 24-hr loading for a specific target gene can be calculated using Equation S4/Equation 5. Similarly, the removal of the 24-hr loading of a specific target gene across the WWTP is calculated using Equation S5/Equation 6.

ARG loading rate $\left(\frac{\text{copies}}{\text{hour}}\right)$:

R = Instantaneous flow rate, f $\left(\frac{\mathrm{L}}{\text{hour}}\right)$ × ARG concentration, C $\left(\frac{\text{copies}}{\mathrm{L}}\right)$

(Eq. S1)

The ARG loading from $\mathrm{T} = 0$ to $\mathrm{T} = 4$ hour:

$$\mathrm{T_1} + \mathrm{T_2} = \frac{(\mathrm{R_0} + \mathrm{R_2}) \times (2 - 0)}{2} + \frac{(\mathrm{R_2} + \mathrm{R_4}) \times (4 - 2)}{2} = \mathrm{R_0} + 2\,\mathrm{R_2} + \mathrm{R_4}$$

(Eq. S2)

The ARG loading removal from $\mathrm{T} = 0$ to $\mathrm{T} = 4$ hour:

$$(\mathrm{T_{1,INF}} + \mathrm{T_{2,INF}}) - (\mathrm{T_{1,FE}} + \mathrm{T_{2,FE}}) = \mathrm{R_{0,INF}} - \mathrm{R_{0,FE}} + 2\,(\mathrm{R_{2,INF}} - \mathrm{R_{2,FE}}) + (\mathrm{R_{4,INF}} - \mathrm{R_{4,FE}}$$

(Eq. S3)

The ARG loading within the $24-$ hour sampling: $\mathrm{R_0} + 2\ \sum_{t=2}^{22} R_t + R_{24}$

(Eq. S4)

The ARG loading removal within the $24-$ hour sampling:

$$Log_{10}\left(\frac{\mathrm{R_{0,INF}} + 2\ \sum_{t=2}^{22} R_{t,INF} + \mathrm{R_{24,INF}}}{\mathrm{R_{0,FE}} + 2\ \sum_{t=2}^{22} R_{t,FE} + \mathrm{R_{24,FE}}}\right)$$

(Eq. S5)

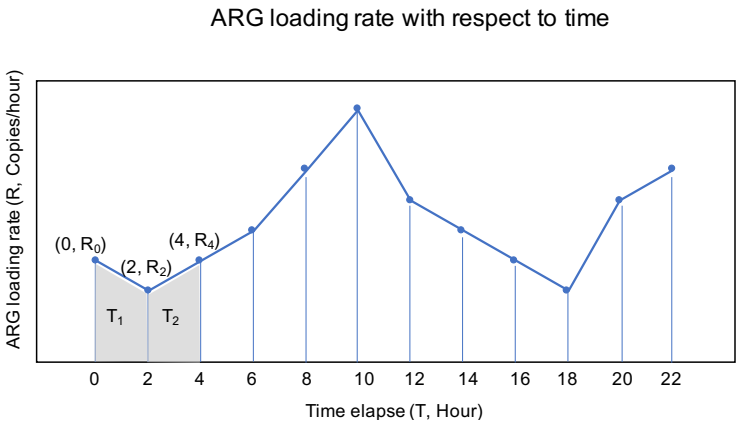


Figure S2.1. Demonstration of the calculation of ARG loading.

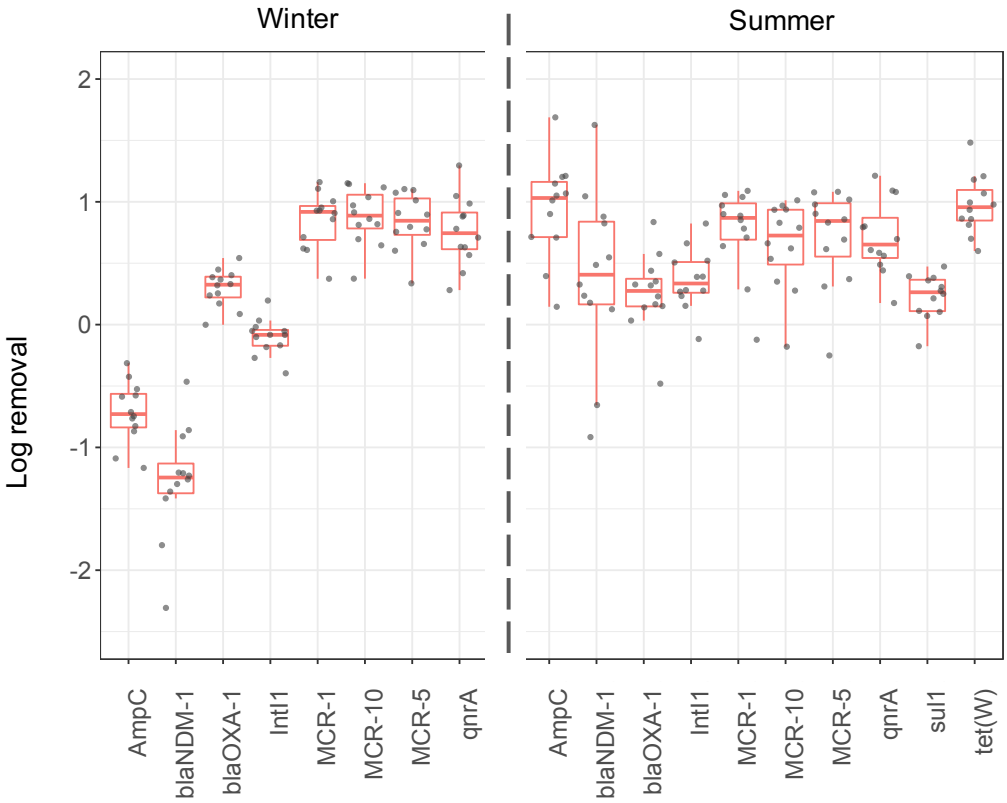


Figure S2.2. Diurnal variation of target genes in terms of relative abundance removal. Instantaneous removal rates (y-axis) of all target genes in terms of relative abundance with respect to time (x-axis) during the winter sampling campaign. The scatter points represent for the instantaneous relative abundance removal rates of all target genes at that specific time when the grab samples were taken. Boxes represent the interquartile range, with solid lines as medians. Whiskers represent the standard deviation. The dashed brown line represents the flow rate. **a.** Relative abundance removal of all target genes with respect to time during the winter field sampling. **b.** Relative abundance removal of all target genes with respect to time during the summer field sampling.

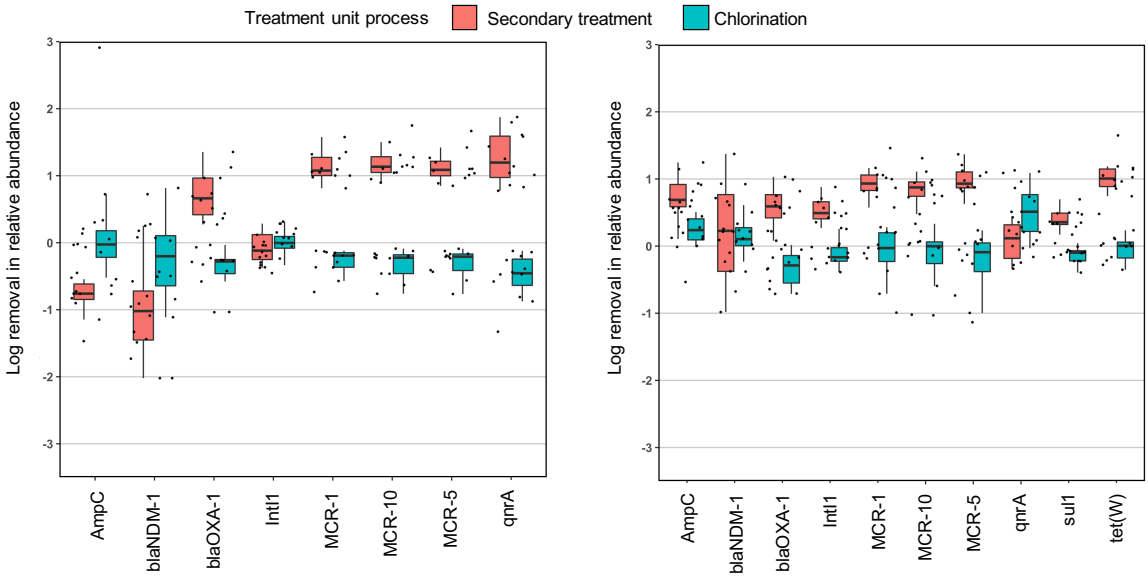


Figure S2.3. Log removal by the WWTP treatment unit process (n=12): secondary treatment and chlorine disinfection. Unit process removal of target genes calculated using relative abundances during the winter (left) and summer (right) sampling campaign. In each panel, instantaneous removal rates as shown as black dots for the target genes. Secondary treatment removal rates are in red, and chlorine disinfection in blue. Boxes represent the interquartile range, with solid lines as medians. Whiskers represent the standard deviation.

Chapter 3 Appendix: Livestock manure improved antibiotic resistance gene removal during co-treatment of domestic wastewater in an anaerobic membrane bioreactor

1. Methods

1.1 AnMBR set-up and monitoring

1.1.1 AnMBR configuration and operational parameters

The AnMBR with a liquid volume of 5 L (Chemglass Life Science, Vineland, NJ) was operated continuously (Fig. S3.1). The hydraulic retention time (HRT) of the AnMBR was maintained at 19 h by controlling the membrane permeate flux. Biomass was only removed from the AnMBR for sampling purposes, resulting in a solids retention time (SRT) of >300 days. Headspace gas was recirculated using a diaphragm pump (KNF Neuberger, Trenton, NJ), and then distributed below each membrane through a horizontally placed sparging tube designed for fouling control. The gas flow rate passing through each sparging tube was controlled via a gas flow meter to maintain the TMPs of all three membranes similar to one another and below 5.5 kPa. The TMP across each membrane was measured using a pressure transducer (Omega Engineering, Stamford, CT). The headspace pressure was monitored using another pressure transducer and the biogas was collected using a Tedlar sampling bag attached to the head plate (Restek, Bellefonte, PA) after a check valve. The influent was stored in a 4 °C refrigerator and pumped into the reactor through a peristaltic pump (Cole-Parmer, Vernon Hills, IL). The effluent was continuously withdrawn with another peristaltic pump with a backwash ratio of 10%. The liquid level was monitored by a sensor switch. The AnMBR was connected to a computer, which operated a control program and LabVIEW (National Instruments, Austin, TX) data

acquisition software. The control program was responsible for operation of all pumps, biogas recirculation, and mixing. The LabVIEW 2017 software (Student Edition) continuously monitored and recorded temperature, TMPs, feed flow rate, and head space pressure.

1.1.2 Operation stages and feeding preparation

After inoculation, the AnMBR fed treating domestic wastewater for 3 weeks until it reached steady-state operation (defined as headspace biogas methane content > 60% and effluent COD < 50 mg/L treating domestic wastewater). Baseline operation data was collected 121 – 135 days after startup. Following the Baseline operation, increasing amounts of manure were added to the influent in Stages 1 through 4. Baseline operation stage, Stage 1, Stage 2, Stage 3 and Stage 4 were operated for 16 days, 20 days, 15 days, 17 days and 18 days, respectively. The influent for Stages 1 - 4 was prepared by defrosting frozen manure slurry in the fridge, homogenizing the manure slurry with a Waring Blender, weighing the slurry, and mixing it with freshly collected domestic wastewater. After adding the manure to the wastewater, the influent was then passed through a 1 mm sieve to remove large solids and prevent influent channel clogging.

Biogas was collected in Tedlar sampling bags with valve and septum fittings (Restek, PA) through a built-in port to the reactor headspace. The volume of biogas produced was measured using a 100 mL BD Slip Tip syringe connected to the gas bag valve after the gas bag had been inflated for approximately a day. A one-way check valve was place between the headspace and the sampling bag to prevent leaking during sampling. For quantitative analysis of the biogas composition, biogas was sampled using a gas-tight glass syringe with lock (Hamilton) and assessed by TRACE™ 1300 Gas Chromatograph

(ThermoFisher Scientific) with pulsed discharge detector (GC-PDD). The standard curve for methane quantification was prepared using analytical grade methane (Airgas).

Chemical oxygen demand (COD) was measured in accordance with USEPA Method 410.4 using Genesys 10S UV-Vis Spectrophotometer (Thermo Scientific) and COD vial kits (CHEMetrics Inc.). Volatile fatty acids (acetate, propionate, formate and valerate), sulfate and nitrate were measured by ion chromatography on an ICS 2100 (Thermo Fisher Scientific, Waltham, MA) using methods described previously (S. Chen & Smith, 2018).

1.2 DNA extraction with internal standards

Internal standards of cell-associated DNA (caDNA) and cell-free DNA (cfDNA) were spiked into samples prior to filtration and DNA extraction to correct for losses during sample processing and DNA extraction. For the caDNA internal standard, we spiked in *Escherichia. coli* DH10β containing an engineered plasmid. The plasmid, *pReporter_8* (RRID: Addgene_60568; Yang et al., 2014), is a low-copy plasmid that was previously modified by knocking out the gene encoding green fluorescence reporter (GFP) and replacing it with the methyl-halide transferase (MHT) gene found in *Batis Maritima* (Cheng et al., 2016). Prior to spiking the samples with the caDNA internal standard, *E. coli* DH10β was grown up on a Luria broth plate containing 34 μg/mL chloramphenicol at 37 °C overnight. A single colony was transferred to a tube containing 2 mL Luria broth with 34 μg/mL chloramphenicol followed by incubation at 200 rpm under 37 °C. After 12 h of incubation, 500 μL liquid culture was added to influent and effluent samples, respectively, right before sample filtration. qPCR was performed on the samples spiked with internal standards to determine the copy number of the recovered caDNA internal standard in the final DNA extracts (C_i in equation 1) and the copy number of target genes

in the final DNA extracts (C_s in equation 1). In addition, a 500 µL aliquot of liquid culture from the same culture tube of internal standard was used in an independent DNA extraction to determine the copy number of caDNA internal standard that was spiked into the sample (C_o in equation 1).

Before DNA extraction, membranes were cut to small pieces and transferred to Lysing Matrix E tubes (MP Biomedicals). All Lysing Matrix E tubes (containing either influent sample, effluent sample or the caDNA internal standards) underwent bead-beating with the maximum intensity for 2 minutes (BioSpec Products, Mini-beadbeater 24, 115 V). After bead-beating, DNA extraction was performed using FastDNA SPIN Kit for Soil (MP Biomedicals) and each sample was eluted to obtain a final volume of 100 µL of DNA extract.

Plasmid pUC19 with an inserted sequence for qPCR was used as the internal standard for cell-free ARG calibration. The insertion is a 183 bp fragment on *ARHGAP11B* gene, a human-associated gene that is specific to the brain neocortex (Florio et al., 2015). The DNA fragment was synthesized (gBlocks, Integrated DNA technology Inc.) and cloned into pMini T2.0 vector and then transferred to NEB 10-β Competent *E. coli* using the PCR Cloning Kit (New England BioLabs Inc., MA). Plasmids were extracted using ZR Plasmid Miniprep kit (ZYMO Research, CA). Approximately 1×10^8 copies of synthesized pMini T2.0 plasmids were added to each effluent sample prior to sample filtration and DNA extraction. The quality of all DNA extracts was tested using 1000 UV-Vis Spectrophotometer (ThermoFisher Scientific, MA). Qubit 3.0 fluorometer and Qubit dsDNA BR Assay Kit (Invitrogen, CA) were applied for DNA quantification.

1.3 Quantifications of target genes and internal standards in real time qPCR

For all target genes (*sul*1, *sul*2, *tet*W, *tet*O, *amp*C, *erm*B, *erm*F, *bla*OXA1, *bla*NDM1, *tp*614, *int*I1) as well as caARG and cfARG internal standards, 10.5 uL qPCR reactions were performed on MicroAmp Fast Optical 96-Well Reaction Plate (0.1 mL, Applied Biosystems) using the QuantStudio 3.0 Real-Time PCR Systems (Applied Biosystems, CA). The standard amplification protocol consisted of an initial denaturation step at 95 °C for 2 min, followed by 40 amplification cycles at 95 °C for 5 s, annealing temperature for 12 s, and 72 °C for 16 s and the melting steps (at 95 °C for 15 s, 60 °C for 1 min, at 95 °C for 15 s). qPCR standards were prepared by inserting the targeted genes into pMiniT 2.0 vector and transformed to NEB 10-β Competent E. coli using the NEB PCR Cloning Kit (New England Biolabs, MA). The inserted target genes, before cloning and transformation, were purified and sequenced PCR products of the AnMBR sludge. The PCR assay was conducted using the exact same primer and condition for qPCR in this study. In addition, PCR products were analyzed on 1% agarose gel electrophoresis to verify the correct amplicon size and the negative presence of non-specific products. The expected PCR products were then cut off from the gel, purified by a Qiagen QIAquick Gel Extraction kit, and sequenced by Sanger method (Genewiz, Inc., TX) to confirm the sequences. After cloning and transformation, transformed E. coli were selected for on AMP selection plates. Grown single colony was picked and cultured overnight again in AMP selection LB overnight. Plasmid extraction was then conducted to acquire plasmids from the cell culture using a ZR Plasmid Miniprep kit (ZYMO Research, CA). Extracted plasmids were then diluted ten-fold to generate standard curve for each qPCR assay. For all qPCR assays performed in this study, three technical replicates were conducted for each biological replicate, melt curves were checked for all reactions, and 3 NTCs for each assay were used

for each run. All assays on actual samples were above the detection limit. The primer efficiency and detection limit of each qPCR assay were reported in ESI Table S5.

2. Results

2.1 Performance

The performance of the AnMBR across all operational stages is shown in Figure A2. In addition, ion-chromatography results showed trace concentration of formic acid (2.21 ± 0.18 mg/L) and acetic acid (2.72 ± 0.16 mg/L) in the effluent during Baseline operation. With the addition of manure starting from Stage 1, effluent COD gradually increased (Fig. S1). Interestingly, in the effluent of Baseline operation, propionate was not detected; however, from Stage 1 through 4, propionate started to accumulate in the effluent significantly due to the addition of manure ($p<0.01$). Previous studies have shown propionate is a key indicator denoting process imbalances in anaerobic digesters treating complex organic waste (Zitomer et al., 2016; Demirel & Yenigün, 2002), which is consistent with the input of manure starting from Stage 1. VFA concentrations were listed in ESI† Table S3.2. Solids concentrations were listed in ESI† Table S3.1.

2.2 The absolute and relative concentrations of target genes in influent and effluent across stages

Relative concentrations are shown in Table S3.10. Absolute concentrations are shown in Table S3.8, 9.

2.3 Correlation analysis of effluent cell-associated and cell-free ARGs

Correlation analysis of effluent caARGs revealed significant associations between different gene types. Effluent cell-associated *int*I1 concentrations were significantly

positively correlated with *sul*1 concentrations across all stages of treatment (Pearsons, r = 0.97, p < 0.01). This suggests *sul*1 may be associated with a Class I integron cassette and co-located on the same plasmids, which was consistent with previous studies on ARG fate in different environments (Zarei-Baygi et al., 2019)(Duan et al., 2018)(Xu et al., 2016). The correlation between *int*I1 and *sul*1 is not surprising because they are both associated with Class I integrons (Mazel, 2006) (Y. Deng et al., 2015). We also observed that the cell-associated concentrations of *amp*C were positively correlated with *rpo*B concentrations (Pearsons r = 0.91, p < 0.01) indicating *amp*C genes are likely cell-associated, which is consistent with the fact that *amp*C genes are frequently detected on chromosomes (Bergstrom et al., 1982)(Mata et al., 2012).

We observed that *int*I1 only correlated with one other ARG (*sul*1) in the cell-associated DNA fraction, but strongly positively correlated with multiple ARGs in the cell-free fraction: *sul*2 (Pearons, r = 0.58, p < 0.01), *amp*C (Pearsons, r = 0.63, p < 0.01) and *erm*B (Pearsons, r = 0.89, p < 0.01). These results are in contrast to some previous studies on fate of ARGs in wastewater environments that reported insignificant associations between *int*I1 and *sul*2, whereas they observed significant positive associations between *int*I1 and *sul*1 (Jang et al., 2018)(Lu et al., 2019)(Xu et al., 2016)(L. Ma et al., 2017a). However, in soil and manure environments, significant positive associations between *int*I1 and *sul*2 have been frequently observed (P. Liu et al., 2017) Guo et al., 2018; (W. Sun et al., 2018)Duan et al., 2018; (X. Zhao et al., 2019); (J. Ma et al., 2019). *Int*I1 has drawn attention in many research studies because it is a proxy for anthropogenic pollution including antibiotic resistance dissemination (Gillings et al., 2014). However, it is challenging to compare our data to these previous studies directly because neither did they explicitly distinguish the

cfARGs from caARGs, nor even capture the cfARG fraction due to the methods they used (Y. Zhang et al., 2016)(Xu et al., 2017)(M. Sun et al., 2015)(Q. Bin Yuan et al., 2018). The concentrations of cell-free *bla*OXA1, *tp*614 and *bla*NDM1 in the effluent increased consistently across all stages (t-test, $p < 0.05$; Fig. 4B). The enrichment of *tp*614 has also been reported in several wastewater treatment processes (Yan et al., 2018), which underscores the challenge of removing it. The difficulty in removing *tp*614 is noteworthy because its concentration has been found to positively correlate with persistent ARGs, particularly tetracycline and extended spectrum beta-lactamase (ESBL) ARGs (Jong et al., 2018). We also observed a significant and positive correlation between *tp*614 and *bla*OXA1 (Pearsons, $r = 0.98$, $p < 0.01$), and *tp*614 and *erm*F (Pearsons, $r = 0.97$, $p < 0.01$). The detailed correlation analysis data using Pearson's correlation analysis can be found in Tables S11 and S12.

Table S1 Solids contents in the influent, effluent and the mixed liquor.

Table S2 Volatile fatty acids (VFAs) concentrations in effluent across all operational stages.

Table S3 Internal standards for caARG and cfARG.

Table S4 Primer sequences and qPCR conditions for all target genes.

Table S5 Primer efficiencies and detection limit of each assay.

Table S6 caDNA and cfDNA recoveries of all stages.

Table S7 LRVs of genes across all stages.

Table S8 Concentrations of target genes in the influent across all stages (copies/mL of influent).

Table S9 Concentrations of targeted genes in the effluent across all stages (copies/mL of effluent) in the (a) Cell-associated fraction and (b) cell-free fraction.

Table S10 Relative abundance of ARGs and MGEs normalized by copies of *rpo*B in the influent and the effluent samples across all stages (copies/ copies of *rpo*B): a. influent; b. effluent.

Table S11 Correlation coefficients for target genes in the effluent cell-associated fraction across all stages, a. r values; b. corresponding p values.

Table S12 Correlation coefficients for target genes in the effluent cell-free fraction across all stages, a. r values; b. corresponding p-values.

Table S1 Solids contents in the influent, effluent and the mixed liquor (n=7).

	Baseline	Stage1	Stage2	Stage3	Stage4
Influent TSS (mg/L)	138.0 ± 2.4	5491.9 ± 104.3	8347.3 ± 164.5	13130.1 ± 261.6	20482.8 ± 405.0
Influent VSS (mg/L)	116.0 ± 2.3	4357.0 ± 85.1	6618.9 ± 133.8	10407.5 ± 208.1	16231.9 ± 321.9
Mixed liquor TSS (mg/L)	7552.4 ± 138.9	8240.3 ± 164.0	8691.4 ± 173.8	10591.7 ± 2111.1	12796.9 ± 256.1
Mixed liquor VSS (mg/L)	5513.2 ± 106.9	5933.0 ± 116.0	6431.7 ± 127.7	8155.6 ± 162.2	9725.6 ± 193.4

Table S2 Volatile fatty acids (VFAs) concentrations in effluent across all operational stages (n=4). The fraction of VFAs in effluent COD is calculated from the theoretical COD of VFAs normalized by the total effluent COD.

	Concentration of VFAs (mg/L)					
Operational Stage	Formate	Acetate	Propionate	Butyrate	Valerate	The fraction of VFAs in effluent COD (%)
Baseline	2.21 ± 0.18	2.72 ± 0.16	-	-	-	6.9
Stage1	5.32 ± 0.11	3.02 ± 0.12	10.00 ± 0.16	-	-	30.0

Stage2	9.06 ± 0.20	2.06 ± 0.05	32.66 ± 0.64	-	-	61.8
Stage3	6.29 ± 0.39	-	44.67 ± 0.86	-	-	49.3
Stage4	8.29 ± 0.17	-	57.12 ± 1.21	-	-	37.25

Table S3 Internal standards for caARG and cfARG.

Internal Standards	Forward primer (5' to 3')	Reverse primer (5' to 3')	Annealing temperature (℃)
caARG, *MHT*	CCCAGATCCCACGGAATCACTT	ATTGCAAAACCATTCCGACCCC	61
cfARG, *ARHGAP11B*	GCCGAGCGGAGTTCAAATTTGA	CGGACACCCTTCACCTTAAT	60

Table S4 The primer sequences and qPCR conditions for all target genes.

Target genes	Forward primer (5' to 3')	Reverse primer (5' to 3')	Annealing temperature (℃)	Reference
*tet*W	GAGAGCCTGCTATATGCCAGC	GGGCGTATCCACAATGTTAAC	60	(Aminov et al., 2001)
*tet*O	ACGGARAGTTTATTGTATACC	TGGCGTATCTATAATGTTGAC	50.3	(Aminov et al., 2001)
*amp*C	CCTCTTGCTCCACATTTGCT	ACAACGTTTGCTGTGTGACG	57.5	(Yang et al., 2012)
*sul*1	CGCACCGGAAACATCGCTGCAC	TGAAGTTCCGCCGCAAGGCTCG	69.5	(Pei et al., 2006)
*sul*2	TCCGGTGGAGGCCGGTATCTGG	CGGGAATGCCATCTGCCTTGAG	65.5	(Pei et al., 2006)
*erm*B	GATACCGTTTACGAAATTGG	GAATCGAGACTTGAGTGTGC	53.5	(Chen et al., 2007)
*erm*F	CGACACAGCTTTGGTTGAAC	GGACCTACCTCATAGACAAG	57.5	Chen et al., 2007)
*bla*OXA1	TATCTACAGCAGCGCCAGTG	CGCATCAAATGCCATAAGTG	60	(Yang et al., 2012)
*int*I1	CTGGATTTCGATCACGGCACG	ACATGCGTGTAAATCATCGTCG	60	(Hardwick et al., 2008)
*tp*614	GGAAATCAACGGCATCCAGTT	CATCCATGCGCTTTTGTCTCT	60	(Zhu et al., 2013a)
*bla*NDM1	CGCCATCCCTGACGATCAAA	CTGAGCACCGCATTAGCCG	57	(Luo et al., 2013)
*rpo*B	AACATCGGTTTGATCAAC	CGTTGCATGTTGGTACCCAT	51	(Dahllof et al., 2000)

Table S5 Primer efficiencies and limit of detection of each assay. qPCR standards were prepared by inserting the targeted genes into pMiniT 2.0 vector and transformed to NEB 10-β Competent E. coli using the NEB PCR Cloning Kit (New England Biolabs, MA). The inserted target genes were amplified from the AnMBR sludge, purified, and sequence-confirmed before cloning. The PCR assay was conducted using the same primers and conditions used for qPCR in this study. In addition, PCR products were analyzed on 1% agarose gel electrophoresis to verify the correct amplicon size and the negative presence of non-specific products. The expected PCR products were exacted from the gel, purified by a Qiagen QIAquick Gel Extraction kit, and sequenced by Sanger method (Genewiz, Inc., TX) to confirm the sequences. After cloning and transformation, transformed *E. coli* were selected for on AMP selection plates. Single colonies were picked and cultured overnight in LB containing AMP. Plasmids were extracted from the cell culture using a ZR Plasmid Miniprep kit (ZYMO Research, CA). Extracted plasmids were used to generate standard curves for each qPCR assay.

Gene	Efficiency (%)	R^2	Detection limit for influent samples (copies/mL)	Detection limit for effluent samples (copies/mL)
rpoB	102.65	0.998	231	20
sul1	97.24	0.999	237	20
sul2	90.19	0.998	3066	263
blaOXA1	107.67	0.996	556	48
ermF	107.91	0.993	275	24
tetW	111.86	0.991	102	9
ampC	114.29	0.998	43	4
ermB	103.23	0.998	19	2
tetO	98.39	0.997	374	32
blaNDM-1	94.10	0.997	282	24
intI1	97.79	0.998	209	18
tp614	95.52	0.998	33	3
iDNA standard	88.89	0.9967	175	15
eDNA standard	99.68	0.9995	203	17

Table S6 caDNA and cfDNA recoveries of all stages. The recovery of cfDNA efficiencies through all stages were around 30% and was consistent across stages. However, The cfDNA extraction method used in this study was originally reported to achieve >90% recovery by the group that first developed this method (D. N. Wang et al., 2016). The discrepancy between these two reported recovery values likely result from experimental steps involved during cfDNA extraction. For example, only cfDNA was targeted without the goal to recover caDNA in Wang et al., 2016. Thus, the raw environmental sample (lake water) was directly passed through the NAAP beads. In contrast, in our study, we used the filtrate harvested consecutively from the filtration step to start NAAP absorption-elution of cfDNA, rather than starting off from raw influent or effluent samples while we collected caDNA on the filter. As a result, a fraction of cfDNA may be associated with the filter directly or with the solids/cells being caught by the filters indirectly and not ending up in

the filtrate. In spite of the relatively lower cfDNA recovery in our study as aforementioned, it is still higher than some wide-applied methods for cfDNA such as alcohol precipitation, CTAB-based extraction and commercial kits (<10%) (Z. Liang & Keeley, 2013); (Eichmiller et al., 2016); (F. Li et al., 2018). Last but not the least, the recoveries in this study are very consistent across all stages, reflecting the reproducibility of the applied cfDNA extraction protocol. Since the goal of applying internal standards for tracking recoveries is to calibrate qPCR results of ARGs and MGEs abundances specifically in cell-associated and cell-free fractions, we believe the reproducibility of DNA recovery is as important as, if not more so, than a high recovery value.

Sample		Recovery Efficiency (%)
Baseline – influent	Cell-associated	66.00±25.69
Baseline – effluent	Cell-associated	49.78±41.66
	Cell-free	33.67±7.69
Stage1 - influent	Cell-associated	38.34±10.27
Stage1 - effluent	Cell-associated	78.22±10.23
	Cell-free	31.50±26.21
Stage2 - influent	Cell-associated	56.26±21.9
Stage2 - effluent	Cell-associated	34.44±1.28
	Cell-free	30.74±12.11
Stage3 - influent	Cell-associated	34.05±12.64
Stage3 - effluent	Cell-associated	50.98±20.11
	Cell-free	29.99±7.70

Table S7 LRVs of genes across all stages.

Genes	Baseline	Stage1	Stage2	Stage3	Stage4
intI1	0.21	1.48	2.20	2.43	4.77
sul1	1.44	1.88	3.00	3.05	3.54
sul2	1.28	1.01	2.65	3.13	3.07
ampC	-0.36	0.20	1.36	2.06	2.64
blaOXA1	2.83	2.22	6.08	3.17	2.21
ermB	3.40	3.35	3.74	4.21	4.04
ermF	3.49	2.20	3.33	3.37	2.44
tet(O)	2.41	2.78	3.98	3.91	3.21
tet(W)	4.63	4.14	2.52	4.22	4.52
tp614	4.18	3.52	4.17	4.40	3.61
blaNDM1	-	-2.18	1.11	1.79	0.88

Table S8 Concentrations of target genes in the influent across all stages (copies/mL of influent).

Target gene	Baseline	Stage 1	Stage 2	Stage 3	Stage 4
*rpo*B	1.32E+05 ± (4.87E+04)	7.95E+05 ± (3.44E+05)	2.79E+05 ± (8.41E+04)	3.83E+06 ± (1.59E+06)	3.22E+06 ± (2.26E+06)
*intI*1	1.47E+07 ± (6.60E+06)	1.08E+08 ± (1.31E+07)	2.53E+08 ± (1.53E+08)	1.15E+08 ± (7.74E+07)	8.38E+07 ± (5.39E+07)
*sul*1	8.23E+07 ± (1.16E+05)	1.42E+08 ± (4.22E+06)	6.62E+08 ± (4.94E+08)	4.85E+08 ± (1.88E+08)	3.04E+08 ± (1.40E+08)
*sul*2	2.23E+06 ± (4.01E+04)	2.93E+06 ± (5.31E+05)	5.04E+07 ± (6.01E+06)	1.05E+08 ± (8.15E+07)	4.46E+07 ± (2.42E+07)
*amp*C	1.18E+05 ± (9.05E+02)	2.41E+06 ± (2.83E+05)	5.12E+06 ± (2.23E+06)	2.46E+07 ± (2.02E+07)	2.01E+06 ± (1.70E+06)
*bla*OXA1	2.59E+04 ± (1.19E+03)	1.11E+04 ± (1.86E+3)	4.71E+06 ± (9.30E+04)	5.44E+06 ± (9.81E+05)	1.28E+06 ± (7.86E+05)
*erm*B	1.34E+06 ± (2.30E+04)	4.62E+06 ± (9.78E+04)	4.71E+06 ± (2.13E+06)	1.25E+06 ± (3.68E+06)	5.88E+05 ± (3.59E+05)
*erm*F	3.07E+06 ± (2.25E+3)	1.32E+06 ± (2.09E+5)	9.91E+06 ± (4.20E+06)	9.07E+06 ± (1.51E+07)	1.73E+06 ± (1.10E+06)
*tet*O	5.66E+06 ± (1.59E+02)	2.35E+06 ± (1.61E+06)	3.34E+07 ± (1.66E+07)	1.87E+07 ± (3.79E+08)	1.26E+07 ± (9.36E+06)
*tet*W	5.71E+06 ± (2.50E+03)	9.94E+06 ± (2.84E+05)	2.87E+06 ± (1.96E+06)	4.61E+08 ± (3.55E+05)	8.44E+06 ± (6.05E+06)
*tp*614	8.42E+06 ± (1.18E+06)	7.30E+05 ± (4.91E+05)	1.23E+07 ± (7.01E+06)	2.05E+07 ± (1.74E+07)	7.72E+06 ± (6.72E+06)
*bla*NDM1	ULC*	4.43E+02 ± (8.17E+00)	1.01E+05 ± (3.92E+04)	1.38E+06 ± (1.17E+06)	5.52E+05 ± (2.57E+05)

Table S9 Concentrations of targeted genes in the effluent across all stages (copies/mL of effluent) in the (a) Cell-associated fraction and (b) cell-free fraction.

Table S9. a

Target gene	Baseline	Stage 1	Stage 2	Stage 3	Stage 4
*rpo*B	6.36E+03 ± (2.83E+03)	1.66E+04 ± (1.96E+04)	1.04E+03± (1.21E+03)	8.62E+03 ± (5.41E+03)	2.42E+02 ± (9.52E+01)
*int*I1	9.07E+06 ± (1.43E+07)	3.57E+06 ± (4.11E+06)	6.67E+05 ± (6.75E+05)	4.51E+05 ± (7.28E+05)	6.43E+02 ± (5.05E+07)
*sul*1	2.50E+06 ± (2.70E+06)	7.31E+05 ± (4.06E+05)	6.08E+05 ± (7.64E+05)	3.18E+05 ± (1.82E+05)	1.04E+03 ± (1.42E+08)
*sul*2	1.13E+05 ± (1.68E+05)	2.76E+05 ± (4.63E+05)	7.08E+04 ± (9.50E+04)	7.40E+04 ± (3.33E+04)	4.24E+03 ± (2.33E+07)
*amp*C	2.74E+05 ± (3.52E+05)	1.52E+06 ± (1.77E+06)	2.22E+05 ± (3.07E+05)	2.14E+05 ± (1.15E+05)	3.45E+03 ± (1.56E+06)
*bla*OXA1	3.65E+01 ± (4.34E+01)	6.49E+01 ± (9.21E+01)	1.97E+00 ± (2.91E+00)	3.63E+03 ± (1.11E+01)	5.57E+01 ± (7.40E+05)
*erm*B	5.18E+02 ± (7.19E+02)	1.86E+03 ± (1.98E+03)	2.55E+02 ± (1.68E+02)	2.90E+01 ± (3.30E+03)	1.49E+01 ± (3.38E+05)
*erm*F	9.91E+02 ± (8.81E+02)	8.13E+03 ± (3.07E+03)	4.59E+03 ± (4.87E+03)	3.87E+03 ± (1.24E+03)	2.46E+01 ± (1.03E+06)
*tet*O	2.15E+04 ± (2.36E+04)	2.89E+03 ± (5.51E+03)	1.99E+03 ± (1.17E+03)	1.57E+03 ± (3.81E+03)	3.81E+02 ± (2.63E+06)
*tet*W	1.07E+02 ± (7.93E+01)	7.01E+02 ± (1.31E+03)	8.51E+03 ± (1.53E+04)	1.12E+04 ± (7.25E+01)	2.24E+01 ± (5.59E+06)
*tp*614	5.39E+02 ± (9.35E+02)	1.20E+02 ± (1.15E+02)	6.55E+02 ± (4.23E+02)	4.17E+02 ± (2.54E+02)	2.14E+02 ± (6.15E+06)
*bla*NDM1	ULC*	6.61E+04 ± (9.66E+04)	7.15E+03 ± (8.92E+03)	2.16E+04 ± (1.19E+04)	1.42E+04 ± (2.59E+05)

Table S9. b

Target gene	Baseline	Stage1	Stage2	Stage3	Stage4
*int*I1	5.91E+02 ± (7.07E+00)	4.58E+03 ± (3.96E+03)	5.79E+03 ± (1.28E+01)	6.50E+01 ± (1.65E+03)	7.90E+02 ± (1.16E+06)
*sul*1	4.98E+05 ± (8.50E+02)	1.15E+06 ± (1.97E+06)	4.88E+04 ± (2.04E+04)	1.15E+05 ± (1.25E+05)	8.62E+04 ± (7.54E+04)
*sul*2	3.08E+03 ± (5.86E+05)	7.21E+03 ± (7.80E+03)	4.19E+04 ± (4.11E+03)	3.99E+03 ± (2.53E+04)	3.32E+04 ± (5.55E+05)
*amp*C	2.46E+01 ± (5.96E+03)	9.27E+03 ± (7.97E+03)	2.25E+03 ± (2.09E+03)	5.35E+02 ± (3.48E+02)	1.12E+03 ± (9.24E+05)
*bla*OXA1	1.45E+00 ± (2.03E+01)	1.75E+00 ± (2.03E+00)	ULC*	2.90E+01 ± (3.50E+02)	7.80E+03 ± (9.70E+03)
*erm*B	1.55E+01 ± (1.29E+01)	1.98E+02 ± (1.90E+02)	6.00E+02 ± (1.78E+01)	4.74E+01 ± (3.54E+01)	3.82E+01 ± (1.33E+01)
*erm*F	1.45E+00 ± (8.69E-01)	1.88E+02 ± (2.41E+02)	1.01E+02 ± (8.01E+01)	1.10E+01 ± (5.55E+02)	6.29E+03 ± (2.53E+03)
*tet*O	3.83E+02 ± (4.70E+02)	1.02E+03 ± (1.41E+03)	1.51E+03 ± (5.68E+01)	7.40E+02 ± (8.89E+03)	7.33E+03 ± (2.60E+03)
*tet*W	2.55E+01 ± (2.30E+01)	1.41E+01 ± (1.38E+01)	8.85E+01 ± (6.01E+01)	1.63E+04 ± (1.12E+01)	2.31E+02 ± (1.65E+02)
*tp*614	2.08E+01 ± (2.88E+01)	9.81E+01 ± (2.53E+02)	1.81E+02 ± (1.56E+02)	4.05E+02 ± (6.01E+10)	1.66E+03 ± (2.82E+03)
*bla*NDM1	ULC*	2.92E+02 ± (4.13E+05)	6.43E+02 ± (5.82E+02)	7.48E+02 ± (3.30E+02)	5.85E+04 ± (4.28E+04)

Table. S10 Relative abundance of ARGs and MGEs normalized by copies of *rpo*B in the influent and the effluent samples across all stages (copies/ copies of *rpo*B): a. influent; b. effluent.

Table S10. a

Target gene	Baseline	Stage1	Stage2	Stage3	Stage4
*int*I1	1.11E+01	1.36E+01	9.08E+01	3.00E+00	2.60E+00
*sul*1	6.22E+01	1.79E+01	2.61E+00	4.23E+00	3.63E+00
*sul*2	1.68E+00	3.69E-01	7.62E-02	2.16E-01	1.47E-01
*amp*C	8.94E-02	3.03E-01	1.02E-01	2.35E-01	4.51E-02
*bla*OXA1	1.95E-02	1.40E-03	9.20E-01	2.21E-01	6.34E-01
*erm*B	1.01E+00	5.81E-01	1.00E+00	2.30E-01	4.61E-01
*erm*F	2.32E+00	1.66E-01	2.10E+00	7.24E+00	2.95E+00
*tet*O	4.28E+00	2.96E-01	3.37E+00	2.06E+00	7.28E+00
*tet*W	4.31E+00	1.25E+00	8.62E-02	2.46E+01	6.68E-01
*tp*614	6.37E+00	9.18E-02	4.29E+00	4.45E-02	9.14E-01
*bla*NDM1	0.00E+00	5.58E-05	8.19E-03	6.75E-02	7.15E-02

Table S10. b

Target gene	Baseline	Stage1	Stage2	Stage3	Stage4
*int*I1	1.43E+02	2.15E+01	6.40E+01	5.23E+00	2.65E-01
*sul*1	3.94E+01	4.40E+00	5.83E+01	3.69E+00	4.30E-01
*sul*2	1.78E+00	1.66E+00	6.79E+00	8.59E-01	1.75E+00
*amp*C	4.31E+00	9.17E+00	2.13E+01	2.48E+00	1.42E+00
*bla*OXA1	5.74E-04	3.91E-04	1.89E-04	4.22E-02	2.30E-02
*erm*B	8.15E-03	1.12E-02	2.44E-02	3.37E-04	6.13E-03
*erm*F	1.56E-02	4.89E-02	4.40E-01	4.50E-02	1.01E-02
*tet*O	3.37E-01	1.74E-02	1.91E-01	1.82E-02	1.57E-01
*tet*W	1.69E-03	4.22E-03	8.16E-01	1.30E-01	9.23E-03
*tp*614	8.47E-03	7.22E-04	6.29E-02	4.84E-03	8.82E-02
*bla*NDM1	0.00E+00	3.98E-01	6.86E-01	2.51E-01	5.87E+00

Table S11 Correlation coefficients for target genes in the effluent cell-associated fraction across all stages, a. r values; b. corresponding p values.

Table S11. a

	*rpo*B	*int*I1	*sul*1	*sul*2	*amp*C	*bla*OXA1	*erm*B	*erm*F	*tet*O	*tet*W	*tp*614	*bla*NDM1
*rpo*B	1.00	0.28	0.09	**0.85**	**0.91**	-0.23	-0.09	0.06	-0.04	-0.01	-0.14	**0.93**
*int*I1	-	1.00	**0.97**	0.39	0.21	-0.34	0.35	-0.11	**0.95**	-0.52	0.17	-0.12
*sul*1	-	-	1.00	0.26	0.06	-0.30	0.20	-0.17	**0.98**	-0.36	0.40	-0.31
*sul*2	-	-	-	1.00	**0.97**	-0.18	**0.97**	**0.82**	0.12	-0.26	-0.41	**0.82**
*amp*C	-	-	-	-	1.00	-0.21	**0.98**	**0.85**	-0.07	-0.29	-0.56	**0.91**
*bla*OXA1	-	-	-	-	-	1.00	-0.36	0.06	-0.26	**0.73**	0.05	0.00
*erm*B	-	-	-	-	-	-	1.00	**0.76**	0.07	-0.43	-0.52	**0.84**
*erm*F	-	-	-	-	-	-	-	1.00	-0.35	0.23	-0.25	**0.81**
*tet*O	-	-	-	-	-	-	-	-	1.00	-0.41	0.37	-0.40
*tet*W	-	-	-	-	-	-	-	-	-	1.00	0.51	-0.18
*tp*614	-	-	-	-	-	-	-	-	-	-	1.00	-0.77
*bla*NDM1	-	-	-	-	-	-	-	-	-	-	-	1.00

Table S11. b

	*rpo*B	*int*I1	*sul*1	*sul*2	*amp*C	*blaOXA1*	*erm*B	*erm*F	*tet*O	*tet*W	*tp*614	*bla*NDM1
rpoB	0.00	0.15	0.65	0.00	0.00	0.58	0.64	0.76	0.83	0.01	0.48	0.00
intI1	-	0.00	0.00	0.23	0.38	0.63	0.87	0.95	0.00	0.01	0.97	0.51
sul1	-	-	-	0.52	0.01	0.77	0.83	0.94	0.00	0.00	0.82	0.95
sul2	-	-	-	-	0.00	0.74	0.00	0.01	0.61	0.91	0.34	0.00
ampC	-	-	-	-	-	-	0.00	0.00	0.57	0.98	0.96	0.00
blaOXA1	-	-	-	-	-	-	0.55	0.92	0.68	0.00	0.93	0.99
ermB	-	-	-	-	-	-	-	0.00	0.97	0.75	0.25	0.00

ermF	-	-	-	-	-	-	-	-	0.28	0.56	0.74	0.00
tetO	-	-	-	-	-	-	-	-	-	0.53	0.86	0.39
tetW	-	-	-	-	-	-	-	-	-	-	0.00	0.99
tp614	-	-	-	-	-	-	-	-	-	-	-	0.00
blaNDM 1	-	-	-	-	-	-	-	-	-	-	-	0.00

Table S12 Correlation coefficients for target genes in the effluent cell-free fraction across all stages, a. r values; b. corresponding p-values.

Table S12. a

	*rpo*B	*int*I1	*sul*1	*sul*2	*amp*C	*bla*OXA1	*erm*B	*erm*F	*tet*O	*tet*W	*tp*614	*bla*NDM1
*rpo*B	1.00	-0.20	0.09	-0.20	-0.21	-0.21	0.47	-0.24	-0.17	**0.77**	-0.09	0.08
*int*I1	-	1.00	0.29	**0.58**	**0.63**	-0.34	**0.89**	-0.31	-0.21	-0.49	-0.38	-0.33
*sul*1	-	-	1.00	-0.55	0.83	-0.47	-0.15	-0.33	-0.37	-0.32	-0.45	-0.36
*sul*2	-	-	-	1.00	-0.12	0.47	0.67	0.47	**0.57**	-0.42	0.46	0.47
*amp*C	-	-	-	-	1.00	-0.22	0.24	-0.20	-0.17	-0.31	-0.28	-0.22
*bla*OXA1	-	-	-	-	-	1.00	-0.32	**0.99**	**0.99**	-0.23	**0.98**	0.99
*erm*B	-	-	-	-	-		1.00	-0.31	-0.19	-0.30	-0.32	-0.32
*erm*F	-	-	-	-	-		-	1.00	**0.99**	-0.25	**0.97**	**0.99**
*tet*O	-	-	-	-	-		-	-	1.00	-0.26	**0.97**	**0.99**
*tet*W	-	-	-	-	-		-	-	-	1.00	-0.04	-0.23
*tp*614	-	-	-	-	-		-	-	-	-	1.00	**0.98**
*bla*NDM1	-	-	-	-	-		-	-	-	-	-	1.00

Table S12. b

	*rpo*B	*int*I1	*sul*1	*sul*2	*amp*C	*bla*OXA1	*erm*B	*erm*F	*tet*O	*tet*W	*tp*614	*bla*NDM1
*rpo*B	0.00	0.35	0.68	0.37	0.34	0.54	0.02	0.27	0.44	0.00	0.69	0.72
*int*I1	-	0.00	0.77	0.00	0.00	0.58	0.33	0.00	0.00	0.40	0.00	0.49
*sul*1	-	-	0.00	0.73	0.75	0.56	0.44	0.70	0.52	0.54	0.80	0.00

*sul*2	-	-	-	0.00	0.00	0.43	0.36	0.00	0.00	0.49	0.00	0.49
*amp*C	-	-	-	-	0.00	0.72	0.37	0.00	0.01	0.43	0.00	0.48
*bla*OXA1	-	-	-	-	-	0.00	0.55	0.00	0.00	0.16	0.00	0.99
*erm*B	-	-	-	-	-	-	0.00	0.27	0.51	0.05	0.87	0.39
*erm*F	-	-	-	-	-	-	-	0.00	0.00	0.33	0.00	0.00
*tet*O	-	-	-	-	-	-	-	-	0.00	0.62	0.00	0.00
*tet*W	-	-	-	-	-	-	-	-	-	0.00	0.89	0.50
*tp*614	-	-	-	-	-	-	-	-	-	-	0.00	0.00
*bla*NDM1	-	-	-	-	-	-	-	-	-	-	-	0.00

Supplementary Figures

Fig. S1. AnMBR performance throughout all stages. Influent COD, effluent COD, COD removal and methane production are shown. Error bars represent standard deviations of all biological replicates within each stage (n>7 for COD data; n = 3 for methane production of each stage).

Fig. S2 Percent removal contributed by the removal of individual target gene across all operational stages (n=5).

Fig. S3. Concentration of target genes (ARGs and MGEs, copies/mL) in the influent across operational stages (n=5).

Fig. S4 Effluent target gene composition: relative abundance (%) of cell-free target genes and cell-associated target genes (n=5).

Fig. S1

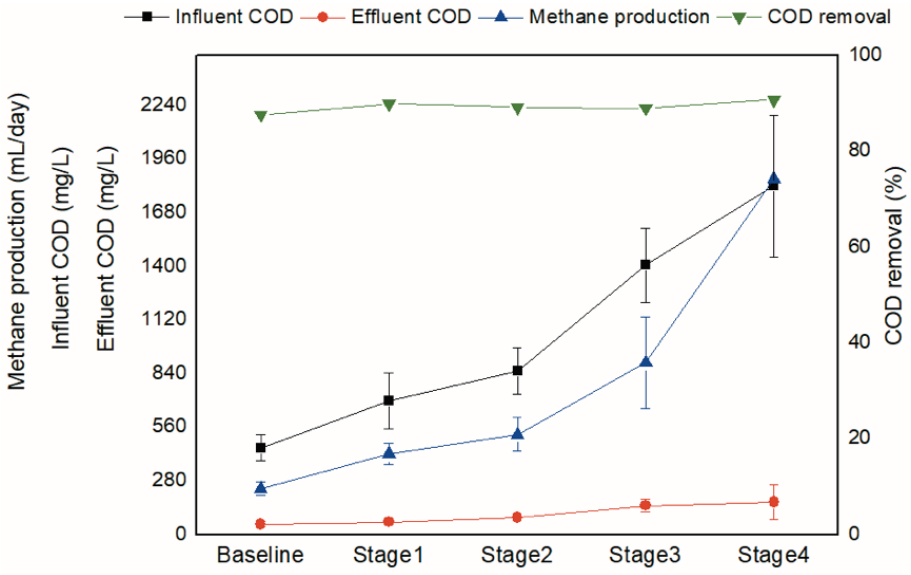


Fig. S2

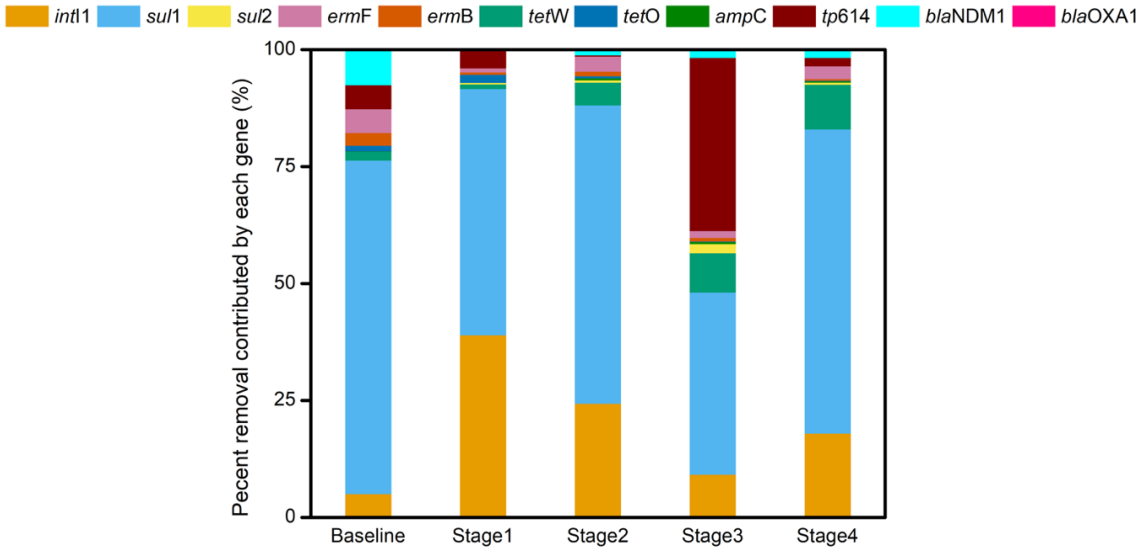


Fig. S3

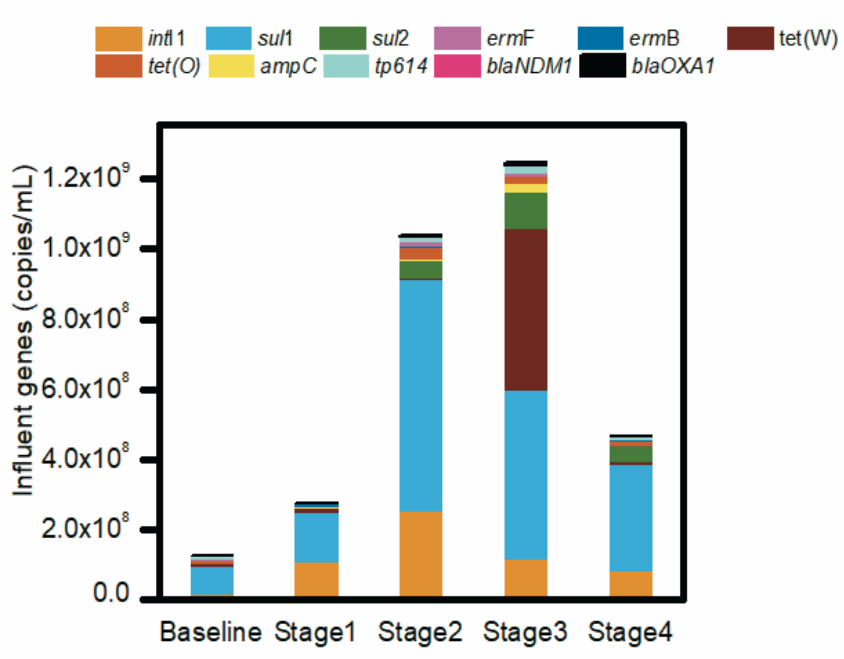
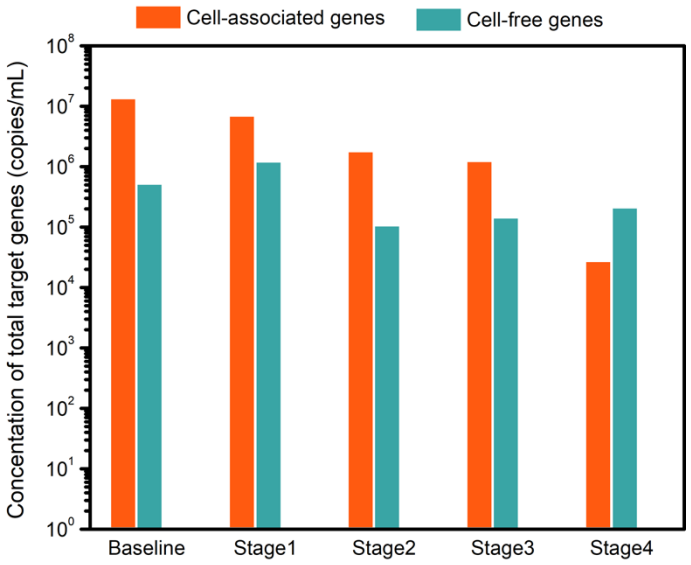


Fig. S4


Cell-associated genes
Cell-free genes
Concentation of total target genes (copies/mL)
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
Baseline
Stage1
Stage2
Stage3
Stage4

Chapter 4 Appendix: Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 mutation monitoring in wastewater

1. Materials and Methods

The protocols for wastewater sample processing, concentration, RNA extraction, RT-ddPCR, and sequencing are described as below. In developing and reporting the procedures, we followed the Environmental Microbiology Minimum Information (EMMI) guidelines (Borchardt et al., 2021),

1.0 Sample collection

During February 23, 2021 to July 12, 2021, we collected wastewater samples from 39 wastewater treatment plants (WWTPs) in Houston covering a service area of approximately 580 square miles and serving over 2.3 million people. Every Monday, time-weighted composite samples of raw wastewater (influent) were collected every 1 hour for 24 hours at each WWTP. After collection from each WWTP site, samples were transported to Houston Water's central laboratory facility on ice. At the facility, samples were aliquoted into bottles, immediately placed back on ice, and transported to Rice University for analysis. Upon arrival, samples were stored at 4 °C prior to sample reduction (concentration) for no longer than 2 hours. Phase I (from February 23, 2021 to April 12, 2021) and Phase II (from May 24, 2021 to July 12, 2021) sampling campaigns included 249 and 298 wastewater samples that were analyzed in this study. Each wastewater sample

was homogenized by agitation and aliquoted into two 50 mL centrifuge tubes before sample concentration.

1.1 Sample concentration (reduction)

1.1.1 Concentration procedure

Prior to concentration, 50 mL wastewater samples were centrifuged for 10 minutes at 4,200 RPM and 4 °C. After centrifugation, the pellet and any remaining solids were discarded, and the supernatant was subject to filtration. Each supernatant was slowly poured into a 6-head, Multi-Vac 610-MS Manifold (180310-01, Sterlitech) containing a pre-DI-washed 0.45 μM pore size, electronegative microbiological analysis HA filter (HAWG047S6, MilliporeSigma). Then, $MgCl_2$*$6H_2O$ was added to the supernatant to achieve a final concentration of 25 mM. The $MgCl_2$ ensures that the viral particles can bind to the negatively charged filter (Ahmed et al., 2020). The samples were gently mixed and allowed to sit for 5 minutes. Subsequently, a vacuum pump pulled the sample through the filter and was turned off after all filters were completely dry. Finally, each filter was folded using sterilized forceps and placed into a bead-beating tube containing 0.1 mm glass beads.

1.1.2 Concentration quality control

To perform quality control on the concentration step, a concentration positive control (CPC) and a concentration negative control (CNC) were implemented. We used a spiked surrogate virus, bovine coronavirus (BCoV), as the positive control (LaTurner et al., 2021b). Briefly, 50 µL BCoV stock solution was spiked into two random wastewater samples immediately before centrifugation. The spiked wastewater samples were processed and concentrated as described above along with other non-spiked, regular

wastewater samples to attain two CPC filters. We used deionized water (DI) as the negative control. Briefly, two 50 mL DI was poured into a manifold head containing a HA filter and processed along with other samples as described above. All control filters were sent to RNA extraction immediately, followed by RT-ddPCR. BCoV was always positive for CPCs and negative for CNCs.

1.2 SARS-CoV-2 extraction

1.2.1 Extraction procedure

We performed RNA extraction using a Chemagic™ Prime Viral DNA/RNA 300 Kit H96 (Chemagic, CMG-1433, PerkinElmer). We followed the manufacturer's recommended protocol, with some modifications in the sample preparation prior to loading. We added the manufacturer's lysis buffer (1000 µL) to each bead-beating tube containing the filter generated from the previous step. The tubes were bead beaten at max speed in a Mini-Beadbeater 24 (3,500 RPM; 112011, BioSpec) for 1 minute, allowed set on ice for 2 minutes, and bead beaten again at max speed for I minute. After bead beating, the tubes were centrifuged to pellet the beads and shredded filters (17,000 g, 4 °C). Subsequently, 300 µL supernatant was loaded into a 96-deep well plate followed by addition of 300 µL lysis buffer and a 14 µL mixture of Proteinase K and Poly(A) RNA reagents as directed by the manufacturer's protocol. Apart from the sample plate, an elution buffer plate, a magnetic bead plate, a wash buffer plate and an eluate collection plate were prepared as directed by the manual. All plates were loaded onto the Chemagic. The extraction program "Chemagic Viral300 360 H96 drying prefilling VD200309.che" was selected for automated RNA extraction. Each sample was eluted to generate a 50 µL RNA extract.

1.2.2 Extraction quality control

In our previous study, we applied the exact same sample collection and concentration procedure (described above), but a different RNA extraction procedure which involved a different automated nucleic acid extraction system, the Maxwell 48 RSC automated platform (AS8500, Promega; LaTurner et al., 2021). Extraction recovery using the Maxwell 48 RSC automated platform (Promega) was evaluated using BCoV spiked samples (LaTurner et al., 2021b). In this current study, before we switched from Promega to Chemagic, we conducted a head-to-head comparison between these two extraction systems using six wastewater samples (data not shown). We found Chemagic produced higher N1 and N2 signals (represented by SARS-CoV-2 concentrations in copies/L-wastewater, obtained from RT-ddPCR on N2 and N2 genes) than Promega. We also included two negative controls in each plate by directly adding 300 µL DI into the sample plate sent for Chemagic extraction and verified that no contamination occurred during RNA extraction by quantifying N1 and N2 genes in these samples via RT-ddPCR.

1.2.3 Extract storage and duration

Immediately after RNA extraction, all sample extracts were sealed and stored at 4 °C for RT-ddPCR analysis within 24 hours. At the same time, 13 µL of each sample extract was aliquoted, placed on ice, and transferred to the Houston Health Department Environmental Microbiology Laboratory for sequencing. Transport time was less than 30 minutes and samples were stored in -80 °C upon arrival until reverse transcription and library preparation.

1.3 Concentration factors calculation (Table S1)

Table S1. The calculation of concentration factors for each step and for the overall sample processing procedure.

Volume of each step during sample concentration and extraction	**Volume**		**Concentration Factor**
Starting volume:	50	ml	
Reagents (lysis buffer) added for bead-beating:	1000	ul	50
Lysate supernatant transferred after bead-beating:	300	ul	
Reagents [lysis buffer, Poly(A) RNA and proteinase K]:	314	ul	
Elution volume:	50	ul	6
The concentration factor for the whole process (concentration and extraction)			**300**

1.4 RT-ddPCR

1.4.1 RT-ddPCR quantification and analysis

All RT-ddPCRs were performed on a QX200 AutoDG Droplet Digital PCR System (Bio-Rad) and a C1000 Thermal Cycler (Bio-Rad) in 96-well optical plates. SARS-CoV-2 N1 and N2 genes were quantified in wastewater samples as previously described (LaTurner et al., 2021). The five target mutations (S:DEL69/70, S:N501Y, S:E484K, S:K417T, S:L452R) were quantified using one-step RT-ddPCR assays detailed in Table S2 according to the manufacturer's protocols (GT Molecular). Briefly, a 22 ul reaction mix containing 10 ul of viral RNA was reverse transcribed using the One-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad). Reaction mix compositions and thermal cycling conditions for each variant assay are detailed in Tables (S3-S5). Each 96-well plate was processed with positive-mutation (positive standards provided by the kit) and no-template controls. A reactions was performed for each replicates for all wastewater samples. Droplets in each

well were read on a QX200 Droplet Reader (Bio-Rad) followed by analysis using QuantaSoft v1.7.4 software. SARS-CoV-2 N1/N2 and Wuhan/UK assays were manually thresholded per channel in Sinplex/Duplex mode while the CA/India and Brazil/South Africa assays were thresholded in cluster mode using the Advanced Classification Method according to GT Molecular's protocol. Manually thresholded N1/N2 and variant data was exported and further analyzed with the concentration factor to obtained gene copies/uL RNA and copies/L-wastewater using custom R scripts. Percent mutations per sample well were calculated for all variant assays by dividing copies/uL RNA per mutation by the total copies/uL RNA in the corresponding well.

1.4.2 Limit of Detection (LOD)

Detection of 3 positive droplets was set for all variant assays and an acceptable total generated droplet count of at least 10,000 was established for all sample wells as recommended by the manufacturer. In addition to the 3 droplets threshold, a per-plate LOD for N1/N2 was calculated by assigning the copy number corresponding to wells having 3 droplets as the initial LOD concentration for that plate. If more than one sample having 3 droplets is detected, then the concentration for all samples having three droplets is averaged. If there are no 3-droplet samples on the plate, a copy number of 0.7 is assumed. A concentration of 0.7 gene copies/uL corresponds to 3 droplets given 10,000 total droplets when taking into consideration a single droplet volume of 0.86 nL, which is the most conservative estimate. The initial LOD concentration is then added to the limit of blank (LOB) to obtain the final LOD for the entire plate. The LOB is the mean concentration of

all negative control samples on the plate plus 1.6 times the standard deviation of the negative controls.

The detailed information of the primer/probe used for mutation detection was not revealed to the authors due to their proprietary nature. Thus, the false positive and false negative rates were not determined for these assays.

Table S2. Information regarding sample collection time, RT-ddPCR mutation targets and kit info.

Sample collection date	**RT-ddPCR target mutations (amino acid change)**	**RT-ddPCR target mutations (nucleotide change)**	**Number of samples assayed**	**GT-molecular Kit for RT-ddPCR (Catalog #)**
Phase I (02/23/2021 - 04/12/2021)	S:DEL69/70	21764ATACATG->A	249	Digital PCR 4-plex (Wuhan + UK B.1.1.7) Differentiation Assay (Not provided)
Phase I (02/23/2021 - 04/12/2021)	S:N501Y	23063A->T	249	
Phase II (05/24/2021 - 07/05/2021)	S:L452R	22917T->G	260	GT-ddPCR 4-plex (Wuhan + B.1.427/B.1.429 + B.1.617) Assay (100144)
Phase II (05/24/2021 - 07/12/2021)	S:K417T	22812A->C	298	SARS-CoV-2 ddPCR Multivariant Discrimination Assay (100134)
Phase II (05/24/2021 - 07/12/2021)	S:E484K	23012G->A	298	
Total number of detection events			1354	

Table S3. Thermal cycling conditions for N1 and N2, S:N501Y and S:DEL69/70.

Cycling Step	Temperature (°C)	Time	Number of Cycles
Reverse Transcription	50	60 min	1

Enzyme activation	95	10 min	1
Denaturation	94	30 sec	40
Annealing/Extension	59	60 sec	
Enzyme Deactivation	98	10 min	1
Droplet Stabilization	4	30 min	1
Hold (optional)	4	24 hrs	1

Table S4. Thermal cycling conditions for S:E484K, S:K417T and S:L452R.

Cycling Step	Temperature °C	Time	Number of Cycles
Reverse Transcription	50	60 min	1
Enzyme activation	95	10 min	1
Denaturation	94	30 sec	45
Annealing/Extension	60	60 sec	
Enzyme Deactivation	98	10 min	1
Droplet Stabilization	4	30 min	1
Hold (optional)	4	24 hrs	1

Table S5. RT-ddPCR reaction compositions for all variant assays

Component	Volume (μL)

One-Step RT-ddPCR Supermix	5.5
Reverse Transcriptase	2.2
300 mM DTT	1.1
GT-4-plex Primer-Probe Solution	1
RNA sample	10
RNase/DNase free water	2.2

1.4.3 cDNA concentration

cDNA concentrations were measured after tile PCR, clean-up, and size selection, right before library preparation. Quantification of cDNA was performed Qubit dsDNA HS assay kit coupled with the Qubit fluorometer (Thermo Fisher Scientific).

2. Results and Discussion

The average concentration of cDNA in samples after tiled PCR and purification was 9.87 ng/μL for Phase I and 5.78 ng/μL for Phase II.

Table S6. Targeted amplicon sequencing statistics. During Phase I and Phase II sample collections, 249 and 298 wastewater samples were sequenced, respectively. For all 547 sequenced samples, the average sequencing depth (the average single base coverage across the entire SARS-CoV-2 reference genome) per sample was 973 (Phase I: 1702, Phase II: 363), the average breadth of genome covered per sample was 66.67%, the average breadth of genome covered across positions with at least 10× depth was 55.8% (Phase I: 83.6%, Phase II: 37.9%).

		Samples with > 90% genome breadth covered*	Samples with > 75% genome breadth covered*	Samples with > 50% genome breadth covered*	Samples with ≥ 1000 × coverage	Samples with ≥ 500 × coverage	Samples with ≥ 100 × coverage
Phase I	Count (n)	165	198	232	171	220	243
	Fraction (n/249; %)	66.3	79.5	93.2	68.7	88.4	97.6
Phase II	Count (n)	16	38	98	11	81	234
	Fraction (n/298; %)	5.4	12.8	32.9	3.7	27.2	78.5

* Only including positions with at least 10× depth across the SARS-CoV-2 genome.

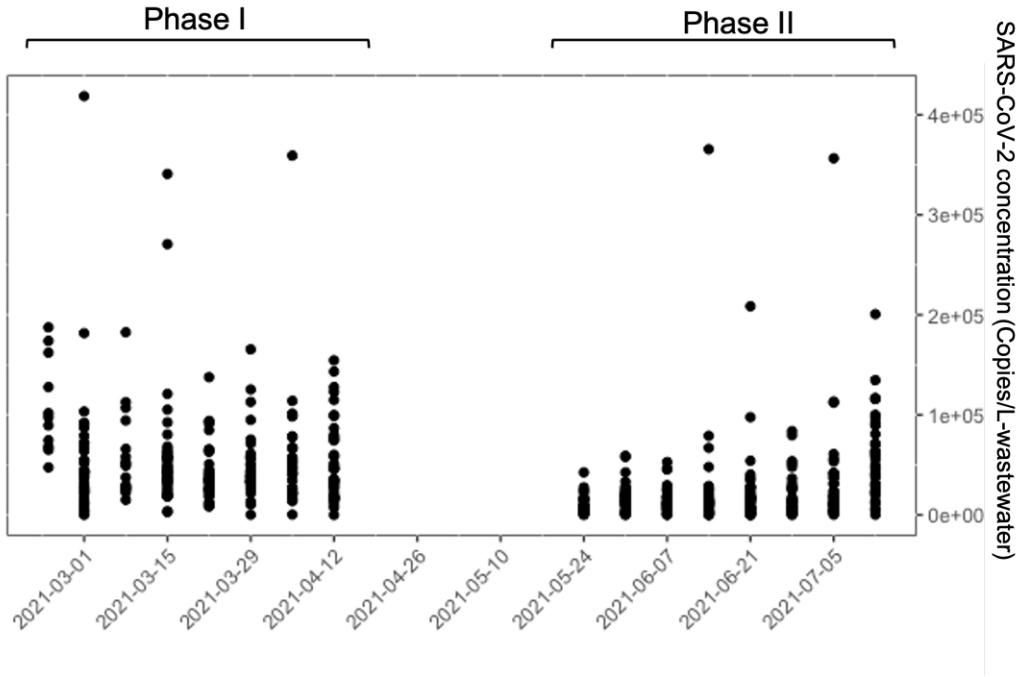
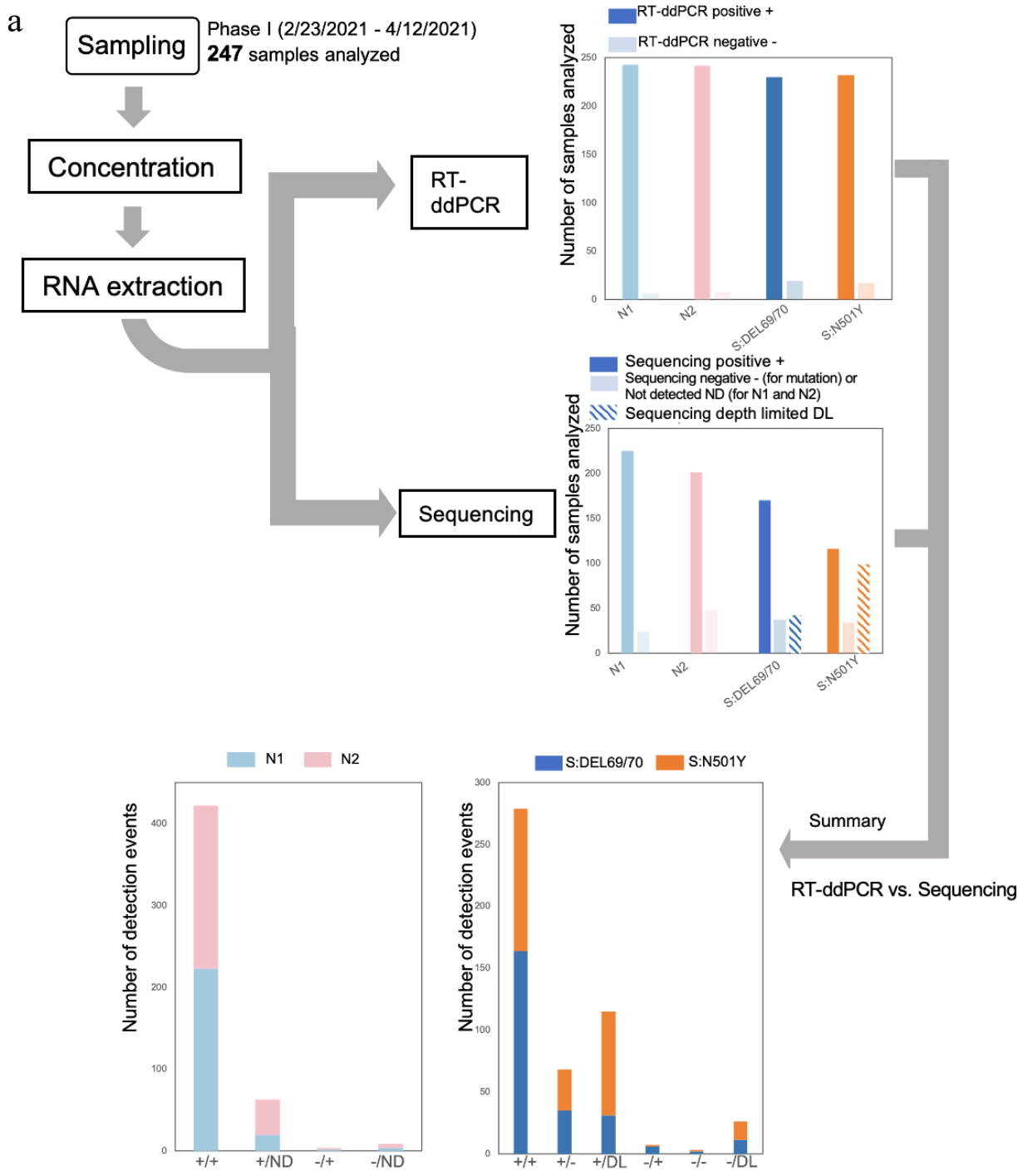


Figure S4.1. SARS-CoV-2 concentrations across Phase I and Phase II (concentrations represent the average of the measured N1 and N2 concentrations, in copies/L-wastewater).


a
Sampling
Phase I (2/23/2021 - 4/12/2021)
247 samples analyzed
Concentration
RNA extraction
RT-ddPCR
Sequencing
RT-ddPCR positive +
RT-ddPCR negative -
Number of samples analyzed
N1
N2
S:DEL69/70
S:N501Y
Sequencing positive +
Sequencing negative - (for mutation) or
Not detected ND (for N1 and N2)
Sequencing depth limited DL
Summary
RT-ddPCR vs. Sequencing
Number of detection events
+/+
+/ND
-/+
-/ND
+/-
+/DL
-/-
-/DL

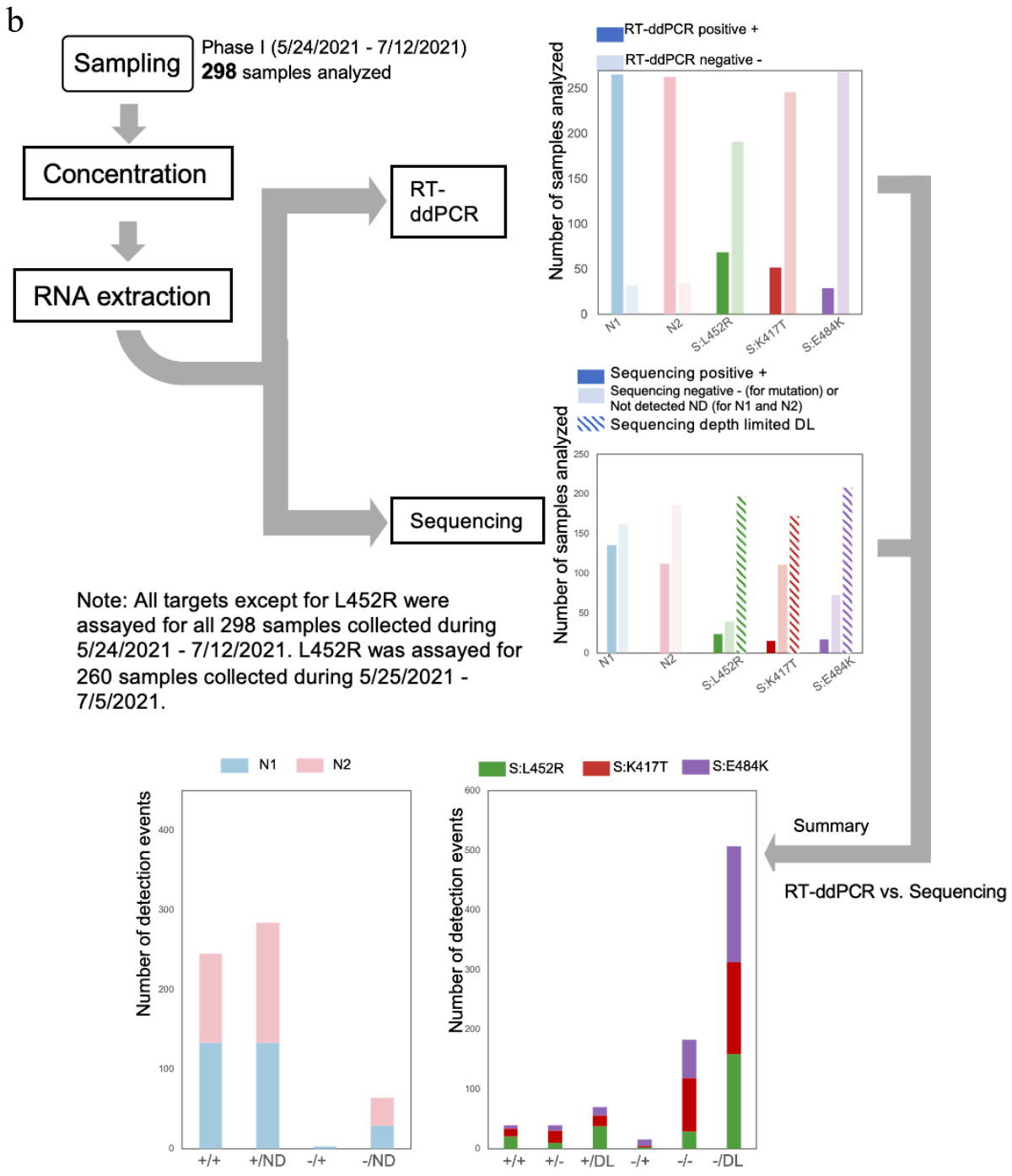


Figure S4.2. Detailed workflow and number of detections by RT-ddPCR and targeted amplicon sequencing for Phase I (a) and Phase II (b) samples.

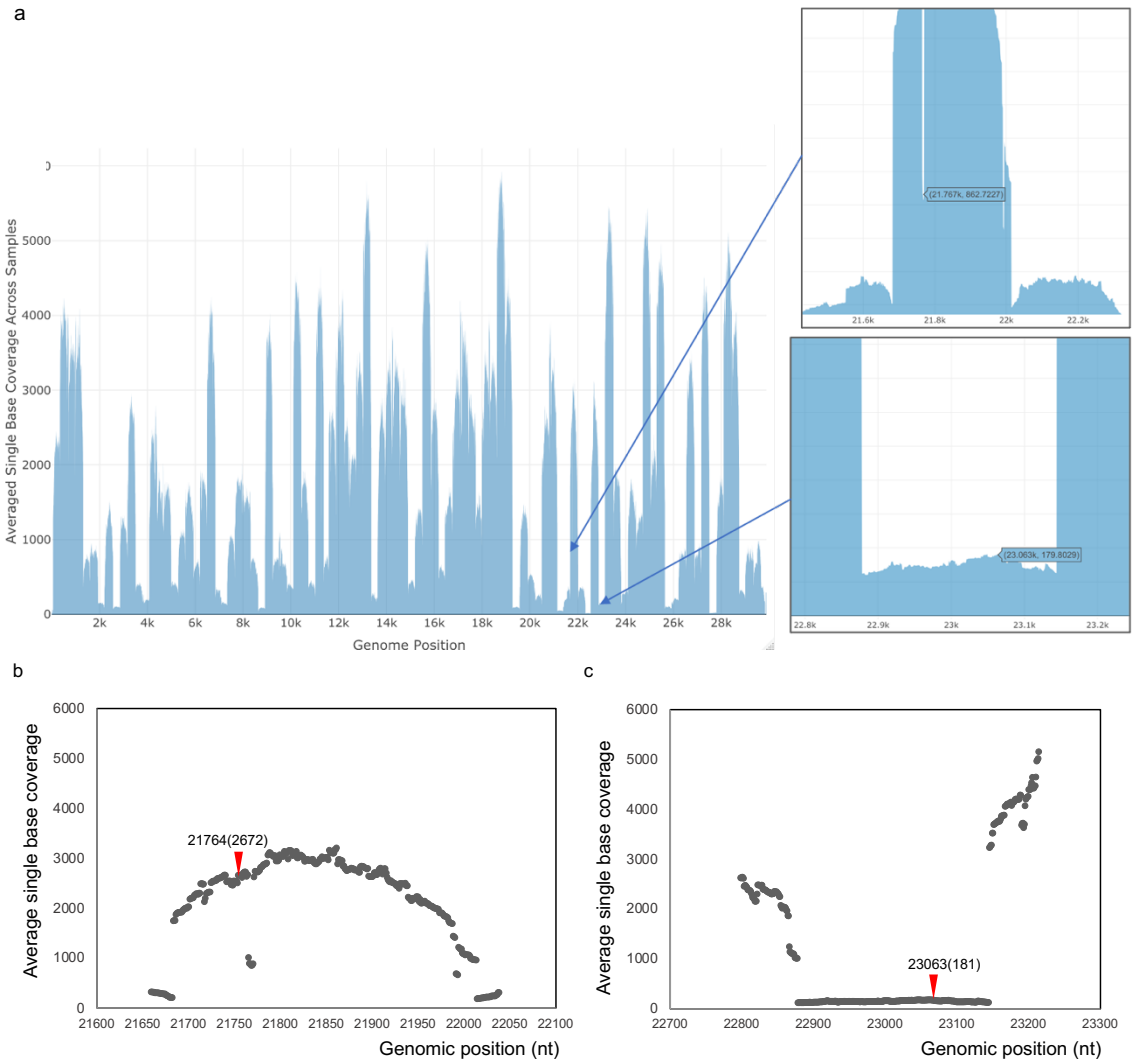


Figure S4.3. The single base coverage across the SARS-CoV-2 genome, based on all 547 samples analyzed (a). The top zoomed-in figure shows the single base coverage around the S:DEL69/70 mutation site and the bottom one shows the single base coverage around the S:N501Y mutation site. The single base coverage profiles cross the specific region amplified by ARTIC.v3_F/R_72 (corresponding to S:del69/70; b) and by ARTIC.v3_F/R_76 (corresponding to S:N501Y; c), based on all Phase I samples analyzed (n = 247), respectively. The red triangles denote the exact mutation position of S:DEL69/70 (nt 21764) and S:N501Y (nt 23063). The numbers in parentheses correspond to the average single base coverage value.

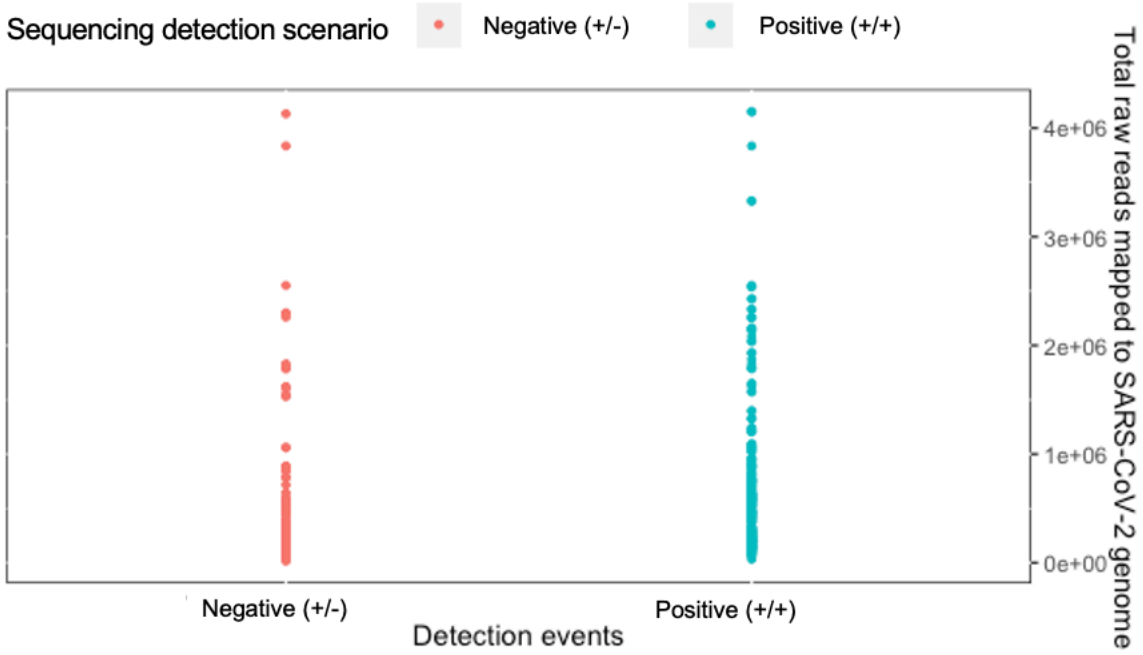


Figure S4.4. The total raw reads mapped to SARS-CoV-2 genome of samples associated with targeted amplicon sequencing negative (+/-) and targeted amplicon sequencing positive (+/+) mutation detection scenarios.

Chapter 5 Using long- and short-read metagenomics and epicPCR to profile antibiotic resistance genes and their bacterial hosts in wastewater

Tables

Figures

1. Methods

1.1 DNA extraction for long-read and short-read sequencing

After biomass concentration, filters were cut into small pieces using sterilized forceps and transferred to a 2 mL tube containing 0.1 mL glass beads for bead-beating. Immediately prior to bead-beating, 1 mL CTAB buffer was added to each sample tube. Sample tubes were vortexed for 30 seconds then incubated at 95°C for 5 minutes. Then, sample tubes were removed from heat and allowed to be cooled at room temperature for no more than 2 minutes. Next, sample tubes were bead beaten at max speed in a Mini-Beadbeater 24 (3,500 RPM; 112011, BioSpec) for 1 minute. After bead-beating, samples were briefly centrifuged and added with 40 µL proteinase K and 20 µL RNase A, then incubated at 70 °C for 10 minutes for lysis treatment. During the incubation, Maxwell RSC cartridges were setup following the manufacture's manual. A volume of 300 µL lysate from each sample tube was transferred to the cartridge. Each cartridge was also added with 300 µL lysis buffer. Finally, all cartridges with added sample lysate and reagents were loaded on to the instrument for automated extraction using the "Maxwell® RSC instrument with the PureFood GMO Protocol". After DNA extraction, DNA was eluted into 100 µL EB and stored in -80 °C before library preparations.

1.2 EpicPCR

1.5 mL WWTP influent sample (n=3) was centrifuged at 10,000 *g* for 1 minute at 4 °C. 600 mL of final effluent sample (n=3) were centrifuged at 5,000 *g* for 10 minutes at 4 °C. Cell pellets were collected in a 2 mL microcentrifuge tube and washed in DNA grade ultrapure water for three times. Next, cells were agitated again using a vortex mixer

at max speed (3000 RPM) for 45 seconds. Then, cells were diluted and stained with DAPI (4',6-diamidino-2-phenylindole) to perform cell count estimation on a Neubauer hemocytometer using a fluorescent microscope (IX 71, Olympus). A final cell count of 1 - 1.4 $\times 10^7$ in 30 µL of cell-water suspension was used for polyacrylamide bead formation and cell lysis treatment as previously described (S. J. Spencer et al., 2016).

Next, fusion PCR and nested PCR were performed for each target per sample (n=3 each sample type). Fusion PCR was conducted within 24 hours after bead formation and cell lysis to prevent DNA degradation. For fusion PCR, PCR mastermix containing fusion templates (45 µL polyacrylamide bead solution), fusion PCR primers (the forward ARG primer F-*sul1*, F-*ermB* or F-*tetO*, the reverse 16S rRNA primer 1492R, and the linker primer RL-*sul1*-519F′, RL-*ermB*-519F′ or RL-*tetO*-519F′), Phusion HF DNA Polymerase (New England Biolabs), and emulsion stabilizers were homogenized in ABIL emulsion oil as previously described (Spencer et al., 2015). Fusion PCR conditions were optimized as follows: initial denaturation at 94 °C for 30 s; 35 cycles of denaturation at 94 °C for 5 s, primer annealing at 55 °C for 30 s, and extension at 72 °C for 30 s; a final extension step at 72 °C for 5 min. The primer sequences and amplicon sizes can be found in Table S1. Immediately after the fusion PCR reaction, 1 mM EDTA was added to the pooled sample, followed by diethyl ether/ethyl acetate wash (Spencer et al., 2015). Next, Monarch PCR & DNA Cleanup Kit (New England Biolabs) was used for DNA extraction from the washed beads. The purified DNA was eluted in 37 µL of EB and was subject to nested PCR. Next, nested PCR was performed using the purified fusion PCR products.

Nested PCR mastermix consisting of Phusion HF DNA polymerase, HF buffer, F'-*sul1*, F-*ermB* or F-*tetO*, and reverse 16S rRNA primer 1391R was divided into

quadruplicate aliquots and combined with the purified fusion PCR products. The nested PCR program consisted of an initial denaturation for 30 s at 98 °C, followed by 38 cycles of denaturation at 98 °C for 5 s, primer annealing at 60 °C (for *sul1*) or 58 °C (for *ermB*) or 55 °C (for *tetO*) for 30 s, extension at 72 °C for 45 s, and a final extension step at 72 °C for 10 min. The final PCR products were loaded onto a 1.5% TAE agarose gel to confirm the expected product via electrophoresis (125 V, 50 minutes). DNA products of approximately 1 kbps in size were extracted using a Monarch gel extraction kit (New England Biolabs). The nested PCR products were purified again using AMPure XP beads and subject to library preparation.

1.3 Integrated pipeline for analyzing long-read and short-read sequencing data

We processed long-read and short-read sequencing reads in an integrated pipeline as shown in Figure S5.1.

1.3.1 Detections of ARG-carrying reads and ARG-carrying contigs, and the associations between ARGs and MGEs

Long-read sequencing reads were screened for ARGs against the CARD database (V3.2.2) using BLAST (https://blast.ncbi.nlm.nih.gov) with a threshold of 70% identity and 70% length coverage. Short-read sequencing reads were assembled and processed using the resistance gene identifier (RGI, version 5.2.1; Alcock et al., 2020). Only contigs carrying the ARGs for which RGI produced “perfect” or “strict” match were selected for further analysis. To explicitly compare long-read and short-read sequencing on resistome characterization, ARG copy numbers generated by both methods were normalized against

sequencing depth as previously described (Arango-Argoty et al., 2019; L. Ma et al., 2016) to attain the relative ARG abundance in reads per billion bases sequenced (RPB). Classification of ARGs was conducted based on the oncology index file curated by CARD database. ARGs corresponding to at least two drug classes were classified as "multidrug" subtype. Beta-lactam resistant ARGs which confer resistance to carbapenem were selected and categorized as the "carbapenem" subtype.

ARG-carrying reads (long-read sequencing) and contigs (short-read sequencing) were subject to BLASTX under the minimum E-value 1×10^{-5} with a length and identity threshold of 70% using a MGE database curated by NanoARG (Arango-Argoty et al., 2019). Long-read and short-read sequencing read statistics are provided in Table S5.2.

To evaluate the results of the comparison between long- and short-read sequencing in terms of ARG-host identification, we downloaded three publicly available datasets from two previous studies (Che et al., 2019b; Fuhrmeister et al., 2021). Both studies conducted long-read and short-read sequencing technologies to sequence the same wastewater samples. Details on the datasets are provided in Table S5.3.

1.3.2 Sample-wise taxonomical abundance estimation for long-read and short-read sequencing data

The sample-wise taxonomical abundance estimation for long-read data was performed via Centrifuge v1.0.4. The program was run directly on ONT and Illumina reads and Centrifuge generated a report that contained the sample abundances.

2. Results and discussion

2.1 EpicPCR sequencing statistics

During read QC, we noticed that even though the size of PCR products was verified via electrophoresis, 50.94% of sequenced DNA still had a read length of shorter than 1 kbps, which may have been due to DNA fragmentation during gel purification and library preparation. We performed an alignment step during which reads were scrutinized for perfect match (i.e., 100% identity and 100% coverage) against the corresponding reverse linker primer sequence. This alignment step is the key to exclude false positives as it filtered out a substantial body of relatively short reads, which were likely partially fused PCR products. As a result, after this alignment step, the vast majority (71.8%) of remaining reads had a length falling within the range of 1007-1089 bps (i.e., the expected length range of nested PCR products given the primer design). Furthermore, the ARG portion of the remaining reads aligned to the corresponding ARG references in SARG database with relatively high sequence similarity and coverage. The average identity of alignment was 93.8 $\pm$ 3.1% for *ermB*, 93.9 $\pm$ 3.0 % for *sul1*, and 94.2 $\pm$ 3.0% for *tetO*. The average length coverage of alignment was 96.4 $\pm$ 9.8% for *ermB*, 99.4 $\pm$ 6.9 % for *sul1*, and 77.5 $\pm$ 5.3% for *tetO*. With respect to the 16S rRNA gene portion of reads, most remaining reads (91.8 $\pm$ 9.4%) passed the 16S rRNA gene alignment criteria of Emu (i.e., the 16S rRNA annotation tool used in this study, Curry et al., 2022) and generated species-level classifications. These results underscore the successful acquisition of ARG-16S rRNA gene fusion structures.

2.2 Host range detected by epicPCR

For *sul1* hosts, epicPCR classified 61 Proteobacteria species and nine Bacteroidetes species (Table S5.5). NCBI Reference Sequence Database (RefSeq) reported consistent *sul1*-host phylum associations, more than 99% (25,195) of the reference sequences associated with *sul1* in bacteria were assigned to Proteobacteria. In previous studies characterizing hosts of *sul1* in wastewater, Bacteroidetes was identified as the dominant host phyla along with Proteobacteria (Wei et al., 2021; G. Zhang et al., 2020). To focus on the host range at the family level, the top three host families of *sul1* classified by epicPCR are *Rhodocyclaceae*, *Aeromonadaceae*, and *Comamonadaceae*, which was consistent with the *sul1* host range profiled by another targeted method using proximity ligation (Stalder et al., 2019). For *ermB* hosts, epicPCR identified 17 Proteobacteria species, 12 Bacteroidetes species, two Firmicute species, and one Fusobacteria species (Table S5.5). In RefSeq, *ermB* was predominantly associated with Proteobacteria (2,415 records), followed by Firmicutes (33 records). Recently, *ermB* has been found more frequently in Bacteroidetes species and was characterized as mobilizable based on its association with certain conjugative transposons (Gupta et al., 2003; Okitsu et al., 2005). Lastly, for *tetO* hosts, epicPCR detected 59 Firmicutes species, 6 Proteobacteria species, and 2 Bacteroidetes species. Consistently, according to NCBI RefSeq, *tetO* was found to be associated with Firmicutes (121 records), Proteobacteria (111 records), and Bacteroidetes (16 records). Almost all hosts detected by long-read sequencing were subsets of the host range detected by epicPCR as discussed in the manuscript. In addition, long-read sequencing and epicPCR demonstrated a consistent host range profile that *sul1* was mainly associated with Proteobacteria, *tetO* was mainly

associated with Firmicutes, and *ermB* was mainly associated with Bacteroidetes and Firmicutes (Table S5.5).

2.3 The profiles of ARG hosts across the WWTP influent and effluent revealed by long-read sequencing and epicPCR

WWTP influent and effluent hosts were the most consistent at the phylum level as shown by epicPCR and long-read sequencing as discussed in the manuscript. In the WWTP effluent, as demonstrated by long-read sequencing, there were 12 ARG host species that were not detected in WWTP influent (hereafter referred to as "new hosts") as well as eight ARG host species which persisted across the whole treatment process (hereafter referred to as "persistent hosts"). In addition, none of the new hosts were found in the secondary effluent samples (data not shown). The persistent hosts included *E. coli* carrying *mdt* genes (i.e., *mdtE*, *mdtN,* and *mdtO*; subtype: efflux pumps) and *Aeromonas caviae* carrying OXA-504 (subtype: multidrug/carbapenem). The new hosts included *Pseudomonas sp. BJP69* carrying *MexD* (subtype: efflux pump), *Pseudomonas oleovorans* carrying *mexF* (subtype: efflux pump), *Salmonella enterica* carrying OXA-256 (subtype: multidrug/carbapenem), *Pandoraea thiooxydans* carrying *ceoB* (subtype: efflux pump), *Enterobacter kobei* carrying *ramA* (subtype: efflux pump), and *Burkholderia pseudomallei* carrying *MuxB* (subtype: efflux pump). Those ARG hosts (*E. coli*, *A. caviae*, *S. enterica*, *E. kobei*, and *B. pseudomallei*) are putative pathogenic species. This finding emphasizes a critical need to include them as the risk indicators, because they were harboring resistance genes of clinical relevance while at the same time poorly responsive to wastewater treatment.

Table S5.1. EpicPCR primers used for fusion PCR and nested PCR.

Fusion PCR				
ARG	Primer	Sequence (5' - 3')	Reference	
sul1	F-sul1	AAATGCTGCGAGTYGGMKCA	(Wei et al., 2018)	
	RL-sul1-519F′	GWATTACCGCGGCKGCTGAACMACCAKCCTRCAGTCCG	(Wei et al., 2021)	
ermB	F-ermB	GAACACTAGGGTTGTTCTTGCA	(Stedtfeld et al., 2018)	
	RL-ermB-519F′	GWATTACCGCGGCKGCTGCTGGAACATCTGTGGTATGGC	The reverse primer portion of ermB was from (Stedtfeld et al., 2018)	
tetO	F-tetO	ACGGARAGTTTATTGTATACC	(Aminov et al., 2001)	
	RL-tetO-519F′	GWATTACCGCGGCKGCTGTGGCGTATCTATAATGTTGAC	The reverse primer portion of tetO was from (Aminov et al., 2001)	
	16S rRNA - 1492R	GGTTACCTTGTTACGACTT	(S. J. Spencer et al., 2016)	
Nested PCR				
ARG	Primer	Sequence (5' - 3')	Reference	Final product size (bp)
sul1	F'-sul1	GACGCCCTGTCCSRTCWGAT	(Wei et al., 2021)	1037
ermB	F-ermB	GAACACTAGGGTTGTTCTTGCA	(Stedtfeld et al., 2018)	1007
tetO	F-tetO	ACGGARAGTTTATTGTATACC	(Aminov et al., 2001)	1043
	U519F-block10	TTTTTTTTTTCAGCMGCCGCGGT AATWC/3SpC3/	(S. J. Spencer et al., 2016)	
	U519R-block10	TTTTTTTTTTGWATTACCGCGGC KGCTG/3SpC3/		
	16S rRNA - 1391R	GACGGGCGGTGTGTRCA	(Lane et al., 1985)	

Table S5.2. Long- and short-read sequencing read statistics.

Method	**Sample**	**Total bases**	**Number of reads/contigs**	**Length N50 (bp)**
Long-read sequencing	Influent (n=3)	4,318,719,998	1,178,106	4,748
Short-read sequencing		42,663,854,700	1,796,758	1,210
Long-read sequencing	Secondary effluent (n=3)	2,885,452,624	464,436	7,159
short-read sequencing	Final effluent (n=3)	1,944,820,196	1,513,866	1,493

Table S5.3. Accession numbers and brief descriptions of the three publicly available wastewater datasets used for comparing long- and short-read metagenomic sequencing

ID	Sequencing platform	Instrument	Total bases	Sample	Sampling region	SRR Accession	Reference
ST _IN	Illumina	Illumina HiSeq 4000 (PE 150)	18G	Municipal wastewater	Hong Kong	SRR8208343	(Che et al., 2019b)
	ONT	ONT MinION	2.4G			SRR7497167	
B_WW_1	Illumina	Illumina HiSeq 2500 (PE150)	5.9G		The greater Boston area, in Massachusetts, USA	SRR12917052	(Fuhrmeister et al., 2021)
	ONT	ONT MinION	1.3G			SRR12917048	
B_WW_2	Illumina	Illumina HiSeq 2500 (PE150)	5.6G			SRR12917051	
	ONT	ONT MinION	0.82 G			SRR12917047	

Table S5.4. Microbial community composition of the WWTP influent sample generated by long- and short-read sequencing

Family	Relative abundance via long-read sequencing	Relative abundance via short-read sequencing
Enterobacteriaceae	4.66	6.24
Bacillaceae	0.58	3.1
Pseudomonadaceae	3.22	3.08
Streptomycetaceae	1.03	2.22
Flavobacteriaceae	0.67	2.16
Burkholderiaceae	1.67	1.9
Lactobacillaceae	0.26	1.88
Streptococcaceae	1.22	1.73
Mycobacteriaceae	0.73	1.56
Microbacteriaceae	0.5	1.45
Xanthomonadaceae	1.05	1.35
Rhizobiaceae	0.63	1.23
Corynebacteriaceae	0.16	1.23
Mycoplasmataceae	0.11	1.22
Sphingomonadaceae	0.49	1.07
Campylobacteraceae	0.49	1.04
Vibrionaceae	0.39	1.03
Paenibacillaceae	0.35	0.98
Comamonadaceae	5.06	0.97
Pasteurellaceae	0.45	0.97
Staphylococcaceae	0.16	0.95
Moraxellaceae	1.82	0.89

Yersiniaceae	0.32	0.89
Clostridiaceae	0.48	0.87
Roseobacteraceae	0.24	0.83
Species	**Relative abundance via long-read sequencing**	**Relative abundance via short-read sequencing**
Salmonella enterica	0.05	2.09
Escherichia coli	0.26	1.16
Bacillus cereus group	0.1	0.57
Helicobacter pylori	0.02	0.47
Bacillus subtilis group	0.04	0.44
pseudomallei group	0.11	0.42
Listeria monocytogenes	0.01	0.38
Pseudomonas syringae group	0.08	0.35
spotted fever group	0.01	0.29
Enterobacter cloacae complex	0.42	0.28
Pseudomonas aeruginosa group	1.08	0.27
Burkholderia cepacia complex	0.26	0.26
Buchnera aphidicola	0.03	0.26
Staphylococcus aureus	0.01	0.26
Burkholderia pseudomallei	0.04	0.25
Yersinia pseudotuberculosis complex	0.01	0.22
Pseudomonas fluorescens group	0.14	0.2
Acinetobacter calcoaceticus/baumannii complex	0.12	0.2
Pseudomonas putida group	0.17	0.19
Klebsiella pneumoniae	0.13	0.19
Pseudomonas aeruginosa	0.25	0.18
Bacillus amyloliquefaciens group	0.01	0.18
Mycobacterium avium complex (MAC)	0.03	0.17
Bacillus thuringiensis	0.01	0.17
Vibrio harveyi group	0.06	0.16

Table S5.5. ARG hosts detected by long-read sequencing, short-read sequencing, and epicPCR
https://rice.box.com/s/d09f8sw7oqj6crn29mk8sbh0w51rn4kt

Table S5.6. Hosts of *sul1*, *ermB* and *tetO* detected by long-read sequencing and epicPCR
https://rice.box.com/s/897x7c05qz8tawmkne0mm45kw9yr3rjh

Table S5.7. Associations between ARGs and MGEs detected by long-read sequencing
https://rice.box.com/s/mv46u9ynqi2cyhc5q978y1tw3m106rwk

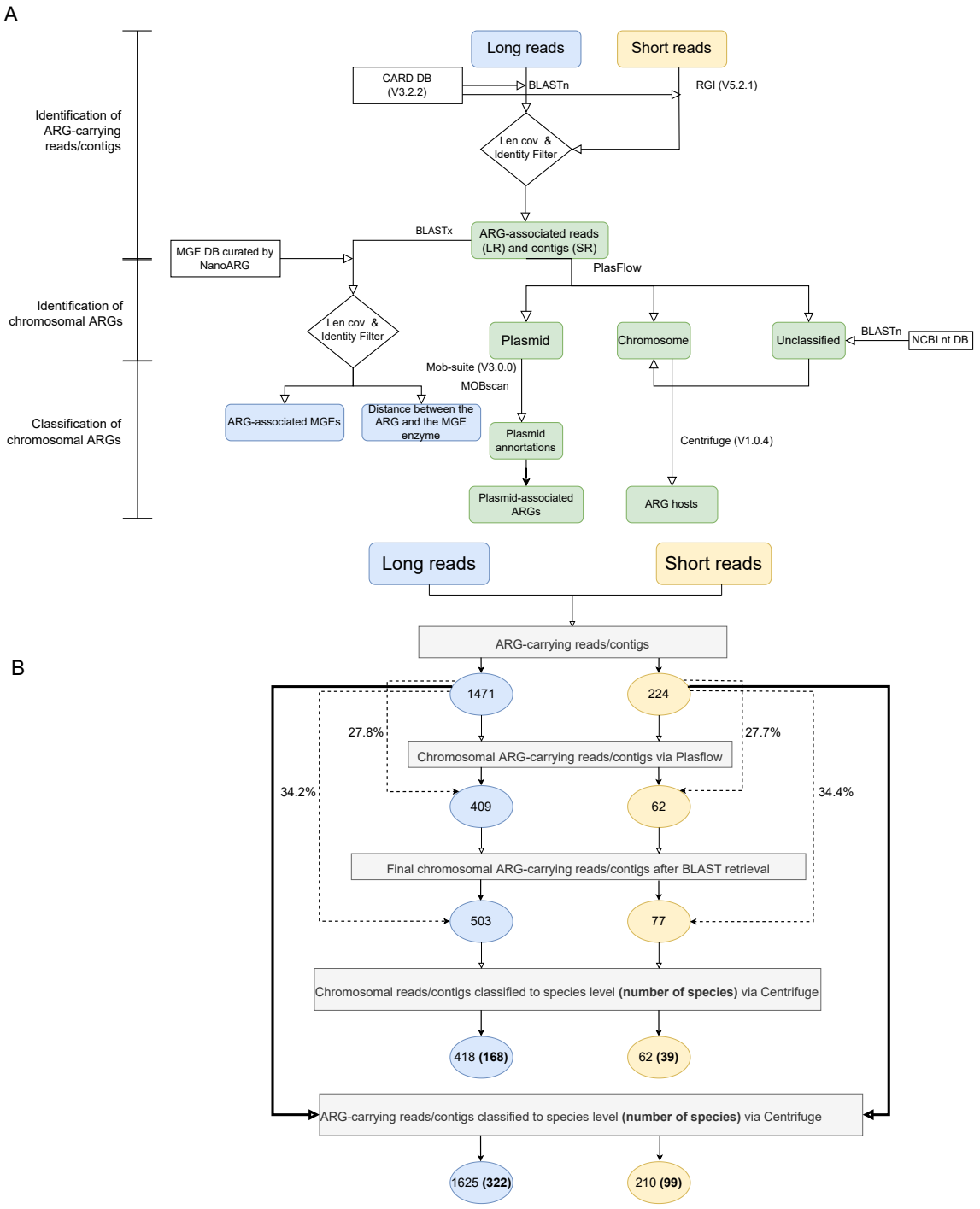


Figure S5.1. Overview of the computational pipeline for analyzing long- and short-read sequencing data. A. The detailed pipeline. B. The output of each step, in terms of the number of reads (long-read sequencing) and contigs (short-read sequencing), and the number of host species classified based on chromosomal ARGs and all ARGs.

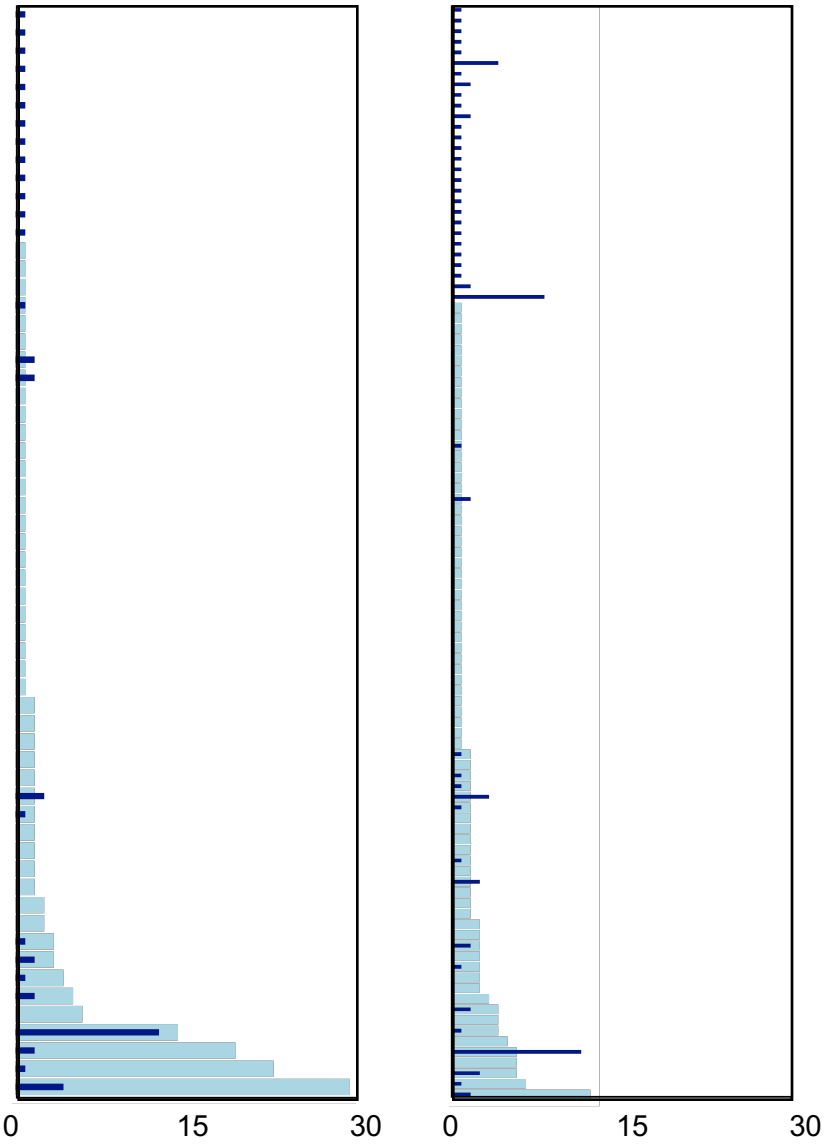


Figure S5.2. Comparison of long- and short-read sequencing in identifying ARG subtypes-host family linkages for other publicly available datasets. Left: sample ID: B_WW_2 (Fuhrmeister et al., 2021), right: ST_IN (Che et al., 2019b).

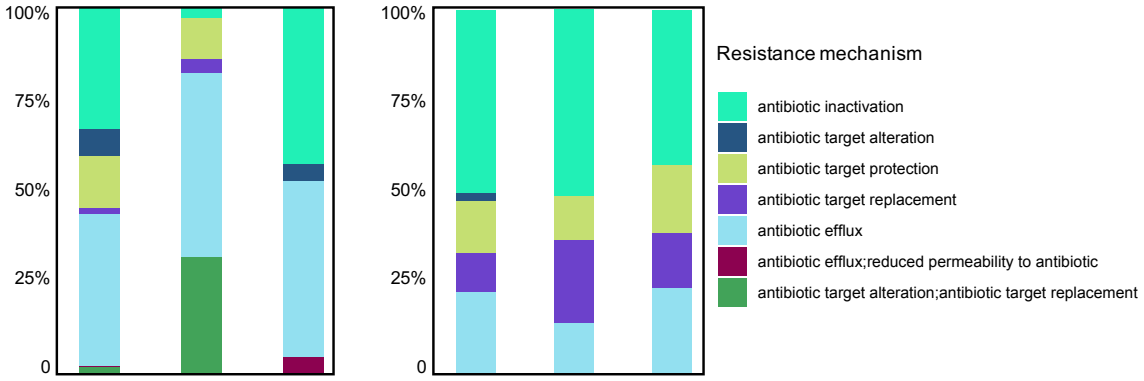


Figure S5.3. Composition of chromosomal ARGs and plasmid-associated ARGs in terms of resistance mechanisms across samples.

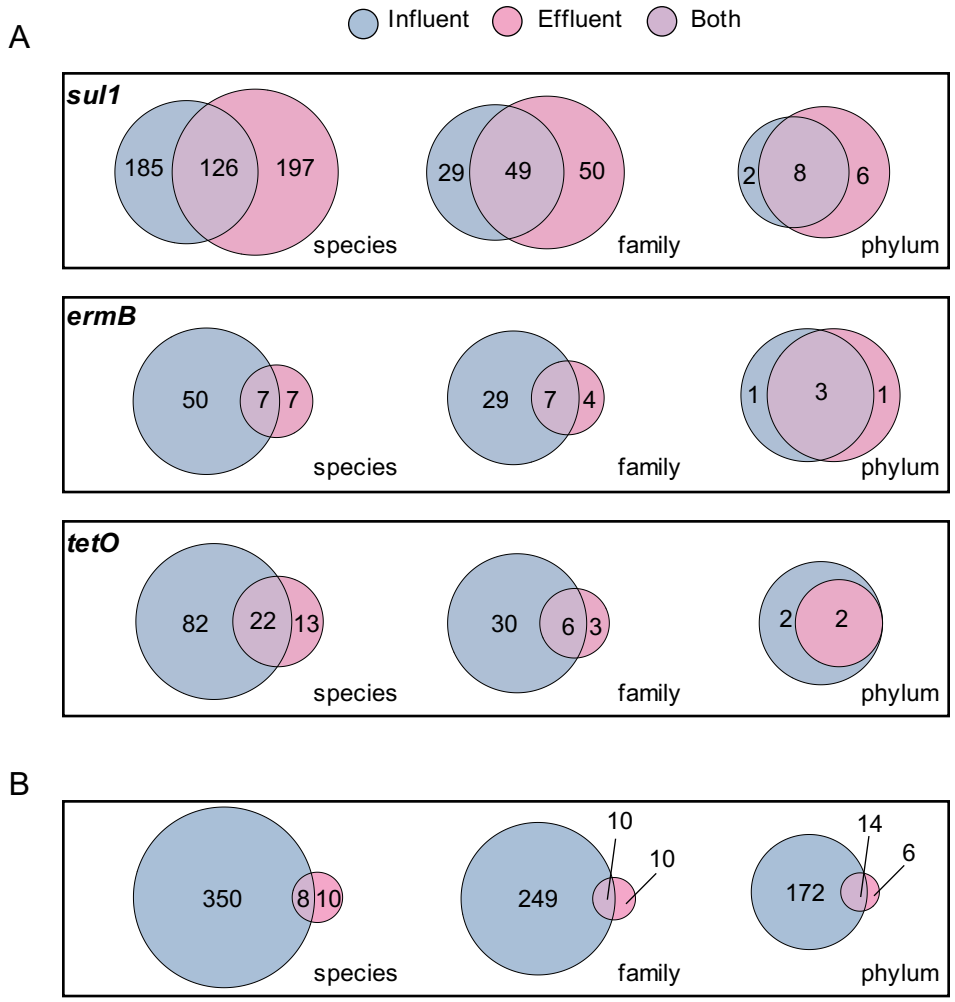


Figure S5.4. WWTP influent and effluent hosts revealed by epicPCR and long read sequencing. a. The count of WWTP influent (blue) and effluent (violet) hosts revealed by epicPCR. The numbers of hosts for the three ARGs (*sul*1, *ermB* and *tetO*) are shown at species, family, and phylum level, respectively. **b.** The count of influent (blue) and effluent (violet) hosts revealed by long-read sequencing. Unique combinations of each ARG and its host counted at the species, family, and phylum level, respectively.

References

Abdelrazik, E., Oweda, M., & El-Hadidi, M. (2021). Benchmarking of Antimicrobial Resistance Gene Detection Tools in Assembled Bacterial Whole Genomes. *2021 3rd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, 273–278. https://doi.org/10.1109/NILES53778.2021.9600515

Abera, B., Kibret, M., & Mulu, W. (2016). Extended-Spectrum beta (β)-Lactamases and Antibiogram in Enterobacteriaceae from Clinical and Drinking Water Sources from Bahir Dar City, Ethiopia. *PLOS ONE*, *11*(11), e0166519. https://doi.org/10.1371/journal.pone.0166519

Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J. W., Choi, P. M., Kitajima, M., Simpson, S. L., Li, J., Tscharke, B., Verhagen, R., Smith, W. J. M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K. V., & Mueller, J. F. (2020). First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of The Total Environment*, *728*, 138764. https://doi.org/10.1016/j.scitotenv.2020.138764

Ahmed, W., Smith, W. J. M., Metcalfe, S., Jackson, G., Choi, P. M., Morrison, M., Field, D., Gyawali, P., Bivins, A., Bibby, K., & Simpson, S. L. (2022). Comparison of RT-qPCR and RT-dPCR Platforms for the Trace Detection of SARS-CoV-2 RNA in Wastewater. *ACS ES&T Water*. https://doi.org/10.1021/acsestwater.1c00387

Ai, Y., Davis, A., Jones, D., Lemeshow, S., Tu, H., He, F., Ru, P., Pan, X., Bohrerova, Z., & Lee, J. (2021). Wastewater SARS-CoV-2 monitoring as a community-level COVID-19 trend tracker and variants in Ohio, United States. *Science of The Total Environment*, *801*, 149757. https://doi.org/10.1016/j.scitotenv.2021.149757

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., … McArthur, A. G. (2020). CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *48*(D1), D517–D525. https://doi.org/10.1093/nar/gkz935

Alygizakis, N., Markou, A. N., Rousis, N. I., Galani, A., Avgeris, M., Adamopoulos, P. G., Scorilas, A., Lianidou, E. S., Paraskevis, D., Tsiodras, S., Tsakris, A., Dimopoulos, M.-A., & Thomaidis, N. S. (2021). Analytical methodologies for the detection of SARS-CoV-2 in wastewater: Protocols and future perspectives. *TrAC Trends in Analytical Chemistry*, *134*, 116125. https://doi.org/10.1016/j.trac.2020.116125

Amha, Y. M., Corbett, M., & Smith, A. L. (2019). Two-Phase Improves Performance of Anaerobic Membrane Bioreactor Treatment of Food Waste at High Organic Loading Rates [Research-article]. *Environmental Science & Technology*, *53*(16), 9572–9583. https://doi.org/10.1021/acs.est.9b02639

Aminov, R. I., Garrigues-Jeanjean, N., & Mackie, R. I. (2001). Molecular Ecology of Tetracycline Resistance: Development and Validation of Primers for Detection of Tetracycline Resistance Genes Encoding Ribosomal Protection Proteins. *Applied*

and Environmental Microbiology, *67*(1), 22–32. https://doi.org/10.1128/AEM.67.1.22-32.2001

An, X.-L., Su, J.-Q., Li, B., Ouyang, W.-Y., Zhao, Y., Chen, Q.-L., Cui, L., Chen, H., Gillings, M. R., Zhang, T., & Zhu, Y.-G. (2018). Tracking antibiotic resistome during wastewater treatment using high throughput quantitative PCR. *Environment International*, *117*, 146–153. https://doi.org/10.1016/j.envint.2018.05.011

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine*, *26*(4), 450–452. https://doi.org/10.1038/s41591-020-0820-9

Andersson, D. I., & Levin, B. R. (1999). The biological cost of antibiotic resistance. *Current Opinion in Microbiology*, *2*(5), 489–493. https://doi.org/10.1016/S1369-5274(99)00005-3

AQUASTAT. (2014). *FAO's Global Water Information System.*

Arango-Argoty, G. A., Dai, D., Pruden, A., Vikesland, P., Heath, L. S., & Zhang, L. (2019). NanoARG: A web service for detecting and contextualizing antimicrobial resistance genes from nanopore-derived metagenomes. *Microbiome*, *7*(1), 88. https://doi.org/10.1186/s40168-019-0703-9

Arora, S., Nag, A., Sethi, J., Rajvanshi, J., Saxena, S., Shrivastava, S. K., & Gupta, A. B. (2020). Sewage surveillance for the presence of SARS-CoV-2 genome as a useful wastewater based epidemiology (WBE) tracking tool in India. *Water Science and Technology*, *82*(12), 2823–2836. https://doi.org/10.2166/wst.2020.540

Ashley, E. A., Shetty, N., Patel, J., van Doorn, R., Limmathurotsakul, D., Feasey, N. A., Okeke, I. N., & Peacock, S. J. (2019). Harnessing alternative sources of antimicrobial resistance data to support surveillance in low-resource settings. *Journal of Antimicrobial Chemotherapy*, *74*(3), 541–546. https://doi.org/10.1093/jac/dky487

Aslan, A., Cole, Z., Bhattacharya, A., & Oyibo, O. (2018). Presence of Antibiotic-Resistant Escherichia coli in Wastewater Treatment Plant Effluents Utilized as Water Reuse for Irrigation. *Water*, *10*(6), 805. https://doi.org/10.3390/w10060805

Augsburger, N., Zaouri, N., Cheng, H., & Hong, P.-Y. (2021). The use of UV/H2O2 to facilitate removal of emerging contaminants in anaerobic membrane bioreactor effluents. *Environmental Research*, *198*, 110479. https://doi.org/10.1016/j.envres.2020.110479

Azli, B., Razak, M. N., Omar, A. R., Mohd Zain, N. A., Abdul Razak, F., & Nurulfiza, I. (2022). Metagenomics Insights Into the Microbial Diversity and Microbiome Network Analysis on the Heterogeneity of Influent to Effluent Water. *Frontiers in Microbiology*, *13*. https://www.frontiersin.org/articles/10.3389/fmicb.2022.779196

Baaijens, J. A., Zulli, A., Ott, I. M., Petrone, M. E., Alpert, T., Fauver, J. R., Kalinich, C. C., Vogels, C. B. F., Breban, M. I., Duvallet, C., McElroy, K., Ghaeli, N., Imakaev, M., Mckenzie-Bennett, M., Robison, K., Plocik, A., Schilling, R., Pierson, M., Littlefield, R., … Baym, M. (2021). *Variant abundance estimation for SARS-CoV-2 in wastewater using RNA-Seq quantification* (p. 2021.08.31.21262938). https://doi.org/10.1101/2021.08.31.21262938

Balaji, A., Kille, B., Kappell, A. D., Godbold, G. D., Diep, M., Elworth, R. A. L., Qian, Z., Albin, D., Nasko, D. J., Shah, N., Pop, M., Segarra, S., Ternus, K. L., & Treangen, T. J. (2022). SeqScreen: Accurate and sensitive functional screening of pathogenic sequences via ensemble learning. *Genome Biology*, *23*(1), 133. https://doi.org/10.1186/s13059-022-02695-x

Barancheshme, F., & Munir, M. (2018). Strategies to Combat Antibiotic Resistance in the Wastewater Treatment Plants. *Frontiers in Microbiology*, *8*. https://doi.org/10.3389/fmicb.2017.02603

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y., & van Wezel, G. P. (2015). Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews : MMBR*, *80*(1), 1–43. https://doi.org/10.1128/MMBR.00019-15

Bar-Or, I., Weil, M., Indenbaum, V., Bucris, E., Bar-Ilan, D., Elul, M., Levi, N., Aguvaev, I., Cohen, Z., Shirazi, R., Erster, O., Sela-Brown, A., Sofer, D., Mor, O., Mendelson, E., & Zuckerman, N. S. (2021). Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel. *Science of The Total Environment*, *789*, 148002. https://doi.org/10.1016/j.scitotenv.2021.148002

Ben, Y., Fu, C., Hu, M., Liu, L., Wong, M. H., & Zheng, C. (2019). Human health risk assessment of antibiotic resistance associated with antibiotic residues in the environment: A review. *Environmental Research*, *169*, 483–493. https://doi.org/10.1016/j.envres.2018.11.040

Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M.-N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., & Martinez, J. L. (2015). Tackling antibiotic resistance: The environmental framework. *Nature Reviews Microbiology*, *13*(5), 310–317. https://doi.org/10.1038/nrmicro3439

Bergstrom, S., Olsson, O., & Normark, S. (1982). Common evolutionary origin of chromosomal beta-lactamase genes in enterobacteria. *Journal of Bacteriology*, *150*(2), 528–534.

Bivins, A., North, D., Wu, Z., Shaffer, M., Ahmed, W., & Bibby, K. (2021a). Within- and between-Day Variability of SARS-CoV-2 RNA in Municipal Wastewater during Periods of Varying COVID-19 Prevalence and Positivity. *ACS ES&T Water*, *1*(9), 2097–2108. https://doi.org/10.1021/acsestwater.1c00178

Bivins, A., North, D., Wu, Z., Shaffer, M., Ahmed, W., & Bibby, K. (2021b). Within- and between-Day Variability of SARS-CoV-2 RNA in Municipal Wastewater during Periods of Varying COVID-19 Prevalence and Positivity. *ACS ES&T Water*, *1*(9), 2097–2108. https://doi.org/10.1021/acsestwater.1c00178

Bloom, J. S., Sathe, L., Munugala, C., Jones, E. M., Gasperini, M., Lubock, N. B., Yarza, F., Thompson, E. M., Kovary, K. M., Park, J., Marquette, D., Kay, S., Lucas, M., Love, T., Sina Booeshaghi, A., Brandenberg, O. F., Guo, L., Boocock, J., Hochman, M., … Arboleda, V. A. (2021). Massively scaled-up testing for SARS-CoV-2 RNA via next-generation sequencing of pooled and barcoded nasal and saliva samples. *Nature Biomedical Engineering*, *5*(7), 657–665. https://doi.org/10.1038/s41551-021-00754-5

Böhm, M.-E., Razavi, M., Marathe, N. P., Flach, C.-F., & Larsson, D. G. J. (2020). Discovery of a novel integron-borne aminoglycoside resistance gene present in

clinical pathogens by screening environmental bacterial communities. *Microbiome*, *8*(1), 41. https://doi.org/10.1186/s40168-020-00814-z

Borchardt, M. A., Boehm, A. B., Salit, M., Spencer, S. K., Wigginton, K. R., & Noble, R. T. (2021). The Environmental Microbiology Minimum Information (EMMI) Guidelines: QPCR and dPCR Quality and Reporting for Environmental Microbiology. *Environmental Science & Technology*, *55*(15), 10210–10223. https://doi.org/10.1021/acs.est.1c01767

Brandt, C., Bongcam-Rudloff, E., & Müller, B. (2020). Abundance Tracking by Long-Read Nanopore Sequencing of Complex Microbial Communities in Samples from 20 Different Biogas/Wastewater Plants. *Applied Sciences*, *10*(21), 7518. https://doi.org/10.3390/app10217518

Brown, C. L., Keenum, I. M., Dai, D., Zhang, L., Vikesland, P. J., & Pruden, A. (2021). Critical evaluation of short, long, and hybrid assembly for contextual analysis of antibiotic resistance genes in complex environmental metagenomes. *Scientific Reports*, *11*(1), 3753. https://doi.org/10.1038/s41598-021-83081-8

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner* (LBNL-7065E). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States). https://www.osti.gov/biblio/1241166

Butler, D., Friedler, E., & Gatt, K. (1995). Characterising the quantity and quality of domestic wastewater inflows. *Water Science and Technology*, *31*(7). https://doi.org/10.1016/0273-1223(95)00318-H

Cacace, D., Fatta-Kassinos, D., Manaia, C. M., Cytryn, E., Kreuzinger, N., Rizzo, L., Karaolia, P., Schwartz, T., Alexander, J., Merlin, C., Garelick, H., Schmitt, H., de Vries, D., Schwermer, C. U., Meric, S., Ozkal, C. B., Pons, M.-N., Kneis, D., & Berendonk, T. U. (2019). Antibiotic resistance genes in treated wastewater and in the receiving water bodies: A pan-European survey of urban settings. *Water Research*, *162*, 320–330. https://doi.org/10.1016/j.watres.2019.06.039

Cai, L., Ju, F., & Zhang, T. (2014). Tracking human sewage microbiome in a municipal wastewater treatment plant. *Applied Microbiology and Biotechnology*, *98*(7), 3317–3326. https://doi.org/10.1007/s00253-013-5402-z

Calderón-Franco, D., Apoorva, S., Medema, G., van Loosdrecht, M. C. M., & Weissbrodt, D. G. (2021). Upgrading residues from wastewater and drinking water treatment plants as low-cost adsorbents to remove extracellular DNA and microorganisms carrying antibiotic resistance genes from treated effluents. *Science of The Total Environment*, *778*, 146364. https://doi.org/10.1016/j.scitotenv.2021.146364

Calderón-Franco, D., van Loosdrecht, M. C. M., Abeel, T., & Weissbrodt, D. G. (2021). Free-floating extracellular DNA: Systematic profiling of mobile genetic elements and antibiotic resistance from wastewater. *Water Research*, *189*, 116592. https://doi.org/10.1016/j.watres.2020.116592

Canh, V. D., Torii, S., Yasui, M., Kyuwa, S., & Katayama, H. (2021). Capsid integrity RT-qPCR for the selective detection of intact SARS-CoV-2 in wastewater. *Science of The Total Environment*, *791*, 148342. https://doi.org/10.1016/j.scitotenv.2021.148342

Cao, Y., Raith, M. R., & Griffith, J. F. (2015). Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality

assessment. *Water Research*, *70*, 337–349. https://doi.org/10.1016/j.watres.2014.12.008

Carlet, J. (2012). The gut is the epicentre of antibiotic resistance. *Antimicrobial Resistance and Infection Control*, *1*(1), 39. https://doi.org/10.1186/2047-2994-1-39

Castanheira, M., Simner, P. J., & Bradford, P. A. (2021). Extended-spectrum β-lactamases: An update on their characteristics, epidemiology and detection. *JAC-Antimicrobial Resistance*, *3*(3), dlab092. https://doi.org/10.1093/jacamr/dlab092

CDC. (2022a, May 16). *National Wastewater Surveillance System*. Centers for Disease Control and Prevention. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html

CDC. (2022b, June 10). *The biggest antibiotic-resistant threats in the U.S.* Centers for Disease Control and Prevention. https://www.cdc.gov/drugresistance/biggest-threats.html

Charre, C., Ginevra, C., Sabatier, M., Regue, H., Destras, G., Brun, S., Burfin, G., Scholtes, C., Morfin, F., Valette, M., Lina, B., Bal, A., & Josset, L. (2020). Evaluation of NGS-based approaches for SARS-CoV-2 whole genome characterisation. *Virus Evolution*, *6*(2), veaa075. https://doi.org/10.1093/ve/veaa075

Che, Y., Xia, Y., Liu, L., Li, A.-D., Yang, Y., & Zhang, T. (2019a). Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome*, *7*(1), 44. https://doi.org/10.1186/s40168-019-0663-0

Che, Y., Xia, Y., Liu, L., Li, A.-D., Yang, Y., & Zhang, T. (2019b). Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome*, *7*(1), 44. https://doi.org/10.1186/s40168-019-0663-0

Che, Y., Xu, X., Yang, Y., Břinda, K., Hanage, W., Yang, C., & Zhang, T. (2022). High-resolution genomic surveillance elucidates a multilayered hierarchical transfer of resistance between WWTP- and human/animal-associated bacteria. *Microbiome*, *10*(1), 16. https://doi.org/10.1186/s40168-021-01192-w

Che, Y., Yang, Y., Xu, X., Břinda, K., Polz, M. F., Hanage, W. P., & Zhang, T. (2021). Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proceedings of the National Academy of Sciences*, *118*(6), e2008731118. https://doi.org/10.1073/pnas.2008731118

Chen, H., & Zhang, M. (2013a). Effects of advanced treatment systems on the removal of antibiotic resistance genes in wastewater treatment plants from Hangzhou, China. *Environmental Science and Technology*, *47*(15), 8157–8163. https://doi.org/10.1021/es401091y

Chen, H., & Zhang, M. (2013b). Effects of Advanced Treatment Systems on the Removal of Antibiotic Resistance Genes in Wastewater Treatment Plants from Hangzhou, China. *Environmental Science & Technology*, *47*(15), 8157–8163. https://doi.org/10.1021/es401091y

Chen, L., Xu, Y., Dong, X., & Shen, C. (2020). Removal of Intracellular and Extracellular Antibiotic Resistance Genes in Municipal Wastewater Effluent by Electrocoagulation. *Environmental Engineering Science*, *37*(12), 783–789. https://doi.org/10.1089/ees.2020.0189

Chen, S., & Smith, A. L. (2018). Methane-driven microbial fuel cells recover energy and mitigate dissolved methane emissions from anaerobic effluents. *Environmental Science: Water Research & Technology*, *4*(1), 67–79. https://doi.org/10.1039/C7EW00293A

Cheng, H., & Hong, P. Y. (2017a). Removal of antibiotic-resistant bacteria and antibiotic resistance genes affected by varying degrees of fouling on anaerobic microfiltration membranes. *Environmental Science and Technology*, *51*(21), 12200–12209. https://doi.org/10.1021/acs.est.7b03798

Cheng, H., & Hong, P.-Y. (2017b). Removal of Antibiotic-Resistant Bacteria and Antibiotic Resistance Genes Affected by Varying Degrees of Fouling on Anaerobic Microfiltration Membranes. *Environmental Science & Technology*, *51*(21), 12200–12209. https://doi.org/10.1021/acs.est.7b03798

Cheng, H.-Y., Masiello, C. A., Bennett, G. N., & Silberg, J. J. (2016). Volatile Gas Production by Methyl Halide Transferase: An In Situ Reporter Of Microbial Gene Expression In Soil. *Environmental Science & Technology*, *50*(16), 8750–8759. https://doi.org/10.1021/acs.est.6b01415

Cheng, X., Xu, J., Smith, G., & Zhang, Y. (2021). Metagenomic insights into dissemination of antibiotic resistance across bacterial genera in wastewater treatment. *Chemosphere*, *271*, 129563. https://doi.org/10.1016/j.chemosphere.2021.129563

Chiara, M., D'Erchia, A. M., Gissi, C., Manzari, C., Parisi, A., Resta, N., Zambelli, F., Picardi, E., Pavesi, G., Horner, D. S., & Pesole, G. (2021). Next generation sequencing of SARS-CoV-2 genomes: Challenges, applications and opportunities. *Briefings in Bioinformatics*, *22*(2), 616–630. https://doi.org/10.1093/bib/bbaa297

Chu, B. T. T., Petrovich, M. L., Chaudhary, A., Wright, D., Murphy, B., Wells, G., & Poretsky, R. (2017). Metagenomics Reveals the Impact of Wastewater Treatment Plants on the Dispersal of Microorganisms and Genes in Aquatic Sediments. *Applied and Environmental Microbiology*, *84*(5), e02168-17, /aem/84/5/e02168-17.atom. https://doi.org/10.1128/AEM.02168-17

Ciesielski, M., Blackwood, D., Clerkin, T., Gonzalez, R., Thompson, H., Larson, A., & Noble, R. (2021). Assessing sensitivity and reproducibility of RT-ddPCR and RT-qPCR for the quantification of SARS-CoV-2 in wastewater. *Journal of Virological Methods*, *297*, 114230. https://doi.org/10.1016/j.jviromet.2021.114230

Coil, D. A., Albertson, T., Banerjee, S., Brennan, G., Campbell, A. J., Cohen, S. H., Dandekar, S., Díaz-Muñoz, S. L., Eisen, J. A., Goldstein, T., Jose, I. R., Juarez, M., Robinson, B. A., Rothenburg, S., Sandrock, C., Stoian, A. M. M., Tompkins, D. G., Tremeau-Bravard, A., & Haczku, A. (2021). SARS-CoV-2 detection and genomic sequencing from hospital surface samples collected at UC Davis. *PLOS ONE*, *16*(6), e0253578. https://doi.org/10.1371/journal.pone.0253578

Colomer-Lluch, M., Jofre, J., & Muniesa, M. (2014). Quinolone resistance genes (qnrA and qnrS) in bacteriophage particles from wastewater samples and the effect of inducing agents on packaged antibiotic resistance genes. *Journal of Antimicrobial Chemotherapy*, *69*(5), 1265–1274. https://doi.org/10.1093/jac/dkt528

Corno, G., Yang, Y., Eckert, E. M., Fontaneto, D., Fiorentino, A., Galafassi, S., Zhang, T., & Di Cesare, A. (2019). Effluents of wastewater treatment plants promote the

rapid stabilization of the antibiotic resistome in receiving freshwater bodies. *Water Research*, *158*, 72–81. https://doi.org/10.1016/j.watres.2019.04.031

Corpuz, M. V. A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., Belgiorno, V., Korshin, G., & Naddeo, V. (2020). Viruses in wastewater: Occurrence, abundance and detection methods. *Science of The Total Environment*, *745*, 140910. https://doi.org/10.1016/j.scitotenv.2020.140910

Coutu, S., Wyrsch, V., Wynn, H. K., Rossi, L., & Barry, D. A. (2013a). Temporal dynamics of antibiotics in wastewater treatment plant influent. *Science of The Total Environment*, *458–460*, 20–26. https://doi.org/10.1016/j.scitotenv.2013.04.017

Coutu, S., Wyrsch, V., Wynn, H. K., Rossi, L., & Barry, D. A. (2013b). Temporal dynamics of antibiotics in wastewater treatment plant influent. *Science of The Total Environment*, *458–460*, 20–26. https://doi.org/10.1016/j.scitotenv.2013.04.017

Crits-Christoph, A., Kantor, R. S., Olm, M. R., Whitney, O. N., Al-Shayeb, B., Lou, Y. C., Flamholz, A., Kennedy, L. C., Greenwald, H., Hinkle, A., Hetzel, J., Spitzer, S., Koble, J., Tan, A., Hyde, F., Schroth, G., Kuersten, S., Banfield, J. F., & Nelson, K. L. (2021a). Genome Sequencing of Sewage Detects Regionally Prevalent SARS-CoV-2 Variants. *MBio*, *12*(1). https://doi.org/10.1128/mBio.02703-20

Crits-Christoph, A., Kantor, R. S., Olm, M. R., Whitney, O. N., Al-Shayeb, B., Lou, Y. C., Flamholz, A., Kennedy, L. C., Greenwald, H., Hinkle, A., Hetzel, J., Spitzer, S., Koble, J., Tan, A., Hyde, F., Schroth, G., Kuersten, S., Banfield, J. F., & Nelson, K. L. (2021b). Genome Sequencing of Sewage Detects Regionally Prevalent SARS-CoV-2 Variants. *MBio*, *12*(1). https://doi.org/10.1128/mBio.02703-20

Curry, K. D., Wang, Q., Nute, M. G., Tyshaieva, A., Reeves, E., Soriano, S., Wu, Q., Graeber, E., Finzer, P., Mendling, W., Savidge, T., Villapol, S., Dilthey, A., & Treangen, T. J. (2022). Emu: Species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nature Methods*, *19*(7), 845–853. https://doi.org/10.1038/s41592-022-01520-4

Dadgostar, P. (2019). Antimicrobial Resistance: Implications and Costs. *Infection and Drug Resistance*, *12*, 3903–3910. https://doi.org/10.2147/IDR.S234610

Dai, D., Brown, C., Bürgmann, H., Larsson, D. G. J., Nambi, I., Zhang, T., Flach, C.-F., Pruden, A., & Vikesland, P. J. (2022). Long-read metagenomic sequencing reveals shifts in associations of antibiotic resistance genes with mobile genetic elements from sewage to activated sludge. *Microbiome*, *10*, 20. https://doi.org/10.1186/s40168-021-01216-5

D'Aoust, P. M., Mercier, E., Montpetit, D., Jia, J.-J., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Langlois, M.-A., Servos, M. R., MacKenzie, M., Figeys, D., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021). Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence. *Water Research*, *188*, 116560. https://doi.org/10.1016/j.watres.2020.116560

Davis, J. J., Long, S. W., Christensen, P. A., Olsen, R. J., Olson, R., Shukla, M., Subedi, S., Stevens, R., & Musser, J. M. (2021). Analysis of the ARTIC Version 3 and

Version 4 SARS-CoV-2 Primers and Their Impact on the Detection of the G142D Amino Acid Substitution in the Spike Protein. *Microbiology Spectrum*. https://doi.org/10.1128/Spectrum.01803-21

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Deng, X., Achari, A., Federman, S., Yu, G., Somasekar, S., Bártolo, I., Yagi, S., Mbala-Kingebeni, P., Kapetshi, J., Ahuka-Mundeke, S., Muyembe-Tamfum, J.-J., Ahmed, A. A., Ganesh, V., Tamhankar, M., Patterson, J. L., Ndembi, N., Mbanya, D., Kaptue, L., McArthur, C., … Chiu, C. Y. (2020). Metagenomic sequencing with spiked primer enrichment for viral diagnostics and genomic surveillance. *Nature Microbiology*, *5*(3), 443–454. https://doi.org/10.1038/s41564-019-0637-9

Deng, Y., Bao, X., Ji, L., Chen, L., Liu, J., Miao, J., Chen, D., Bian, H., Li, Y., & Yu, G. (2015). Resistance integrons: Class 1, 2 and 3 integrons. *Annals of Clinical Microbiology and Antimicrobials*, *14*(1), 45. https://doi.org/10.1186/s12941-015-0100-6

Deshpande, A. S., & Fahrenfeld, N. L. (2022). Abundance, diversity, and host assignment of total, intracellular, and extracellular antibiotic resistance genes in riverbed sediments. *Water Research*, *217*, 118363. https://doi.org/10.1016/j.watres.2022.118363

Di Cesare, A., Eckert, E. M., D'Urso, S., Bertoni, R., Gillan, D. C., Wattiez, R., & Corno, G. (2016a). Co-occurrence of integrase 1, antibiotic and heavy metal resistance genes in municipal wastewater treatment plants. *Water Research*. https://doi.org/10.1016/j.watres.2016.02.049

Di Cesare, A., Eckert, E. M., D'Urso, S., Bertoni, R., Gillan, D. C., Wattiez, R., & Corno, G. (2016b). Co-occurrence of integrase 1, antibiotic and heavy metal resistance genes in municipal wastewater treatment plants. *Water Research*, *94*, 208–214. https://doi.org/10.1016/j.watres.2016.02.049

Dilthey, A. T., Jain, C., Koren, S., & Phillippy, A. M. (2019). Strain-level metagenomic assignment and compositional estimation for long reads with MetaMaps. *Nature Communications*, *10*(1), 3066. https://doi.org/10.1038/s41467-019-10934-2

Do, T. T., Nolan, S., Hayes, N., O'Flaherty, V., Burgess, C., Brennan, F., & Walsh, F. (2022). Metagenomic and HT-qPCR analysis reveal the microbiome and resistome in pig slurry under storage, composting, and anaerobic digestion. *Environmental Pollution*, *305*, 119271. https://doi.org/10.1016/j.envpol.2022.119271

Doddapaneni, H., Cregeen, S. J., Sucgang, R., Meng, Q., Qin, X., Avadhanula, V., Chao, H., Menon, V., Nicholson, E., Henke, D., Piedra, F.-A., Rajan, A., Momin, Z., Kottapalli, K., Hoffman, K. L., Sedlazeck, F. J., Metcalf, G., Piedra, P. A., Muzny, D. M., … Gibbs, R. A. (2021). Oligonucleotide capture sequencing of the SARS-CoV-2 genome and subgenomic fragments from COVID-19 individuals. *PLOS ONE*, *16*(8), e0244468. https://doi.org/10.1371/journal.pone.0244468

Dong, P., Wang, H., Fang, T., Wang, Y., & Ye, Q. (2019). Assessment of extracellular antibiotic resistance genes (eARGs) in typical environmental samples and the transforming ability of eARG. *Environment International*, *125*, 90–96. https://doi.org/10.1016/j.envint.2019.01.050

Draper, A. D. K., Mayo, M., Harrington, G., Karp, D., Yinfoo, D., Ward, L., Haslem, A., Currie, B. J., & Kaestli, M. (2010). Association of the Melioidosis Agent Burkholderia pseudomallei with Water Parameters in Rural Water Supplies in Northern Australia. *Applied and Environmental Microbiology*, *76*(15), 5305–5307. https://doi.org/10.1128/AEM.00287-10

Du, J., Geng, J., Ren, H., Ding, L., Xu, K., & Zhang, Y. (2015). Variation of antibiotic resistance genes in municipal wastewater treatment plant with A2O-MBR system. *Environmental Science and Pollution Research*, *22*(5), 3715–3726. https://doi.org/10.1007/s11356-014-3552-x

Duan, M., Gu, J., Wang, X., Li, Y., Zhang, S., Yin, Y., & Zhang, R. (2018). Effects of genetically modified cotton stalks on antibiotic resistance genes, intI1, and intI2 during pig manure composting. *Ecotoxicology and Environmental Safety*, *147*, 637–642. https://doi.org/10.1016/j.ecoenv.2017.09.023

Eichmiller, J. J., Miller, L. M., & Sorensen, P. W. (2016). Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Molecular Ecology Resources*, *16*(1), 56–68. https://doi.org/10.1111/1755-0998.12421

Ellmen, I., Lynch, M. D. J., Nash, D., Cheng, J., Nissimov, J. I., & Charles, T. C. (2021). *Alcov: Estimating Variant of Concern Abundance from SARS-CoV-2 Wastewater Sequencing Data* (p. 2021.06.03.21258306). https://doi.org/10.1101/2021.06.03.21258306

Eregno, F. E., Tryland, I., Myrmel, M., Wennberg, A., Oliinyk, A., Khatri, M., & Heistad, A. (2018). Decay rate of virus and faecal indicator bacteria (FIB) in seawater and the concentration of FIBs in different wastewater systems. *Microbial Risk Analysis*, *8*, 14–21. https://doi.org/10.1016/j.mran.2018.01.001

Erster, O., Mendelson, E., Levy, V., Kabat, A., Mannasse, B., Asraf, H., Azar, R., Ali, Y., Shirazi, R., Bucris, E., Bar-Ilan, D., Mor, O., Mandelboim, M., Sofer, D., Fleishon, S., & Zukerman, N. S. (2021). *Rapid And high throughput RT-qPCR assay for identification and differentiation between SARS-CoV-2 variants B.1.1.7 and B.1.351* (p. 2021.05.19.21257439). https://doi.org/10.1101/2021.05.19.21257439

Feng, S., Roguet, A., McClary-Gutierrez, J. S., Newton, R. J., Kloczko, N., Meiman, J. G., & McLellan, S. L. (2021). Evaluation of Sampling, Analysis, and Normalization Methods for SARS-CoV-2 Concentrations in Wastewater to Assess COVID-19 Burdens in Wisconsin Communities. *ACS ES&T Water*, *1*(8), 1955–1965. https://doi.org/10.1021/acsestwater.1c00160

Flood, M. T., D'Souza, N., Rose, J. B., & Aw, T. G. (2021). Methods Evaluation for Rapid Concentration and Quantification of SARS-CoV-2 in Raw Wastewater Using Droplet Digital and Quantitative RT-PCR. *Food and Environmental Virology*, *13*(3), 303–315. https://doi.org/10.1007/s12560-021-09488-8

Florio, M., Namba, T., Pääbo, S., Hiller, M., & Huttner, W. B. (2016). A single splice site mutation in human-specific ARHGAP11B causes basal progenitor amplification. *Science Advances*, *2*(12), e1601941. https://doi.org/10.1126/sciadv.1601941

Fontenele, R. S., Kraberger, S., Hadfield, J., Driver, E. M., Bowes, D., Holland, L. A., Faleye, T. O. C., Adhikari, S., Kumar, R., Inchausti, R., Holmes, W. K., Deitrick, S., Brown, P., Duty, D., Smith, T., Bhatnagar, A., Yeager, R. A., Holm, R. H.,

Reitzenstein, N. H. von, … Varsani, A. (2021). High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *MedRxiv*, 2021.01.22.21250320. https://doi.org/10.1101/2021.01.22.21250320

Fosso, B., Santamaria, M., D'Antonio, M., Lovero, D., Corrado, G., Vizza, E., Passaro, N., Garbuglia, A. R., Capobianchi, M. R., Crescenzi, M., Valiente, G., & Pesole, G. (2017). MetaShot: An accurate workflow for taxon classification of host-associated microbiome from shotgun metagenomic data. *Bioinformatics*, *33*(11), 1730–1732. https://doi.org/10.1093/bioinformatics/btx036

Fuhrmeister, E. R., Voth-Gaeddert, L. E., Metilda, A., Tai, A., Batorsky, R. E., Veeraraghavan, B., Ward, H. D., Kang, G., & Pickering, A. J. (2021). *Surveillance of potential pathogens and antibiotic resistance in wastewater and surface water from Boston, USA and Vellore, India using long-read metagenomic sequencing* (p. 2021.04.22.21255864). https://doi.org/10.1101/2021.04.22.21255864

Gao, P., Munir, M., & Xagoraraki, I. (2012a). Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional municipal wastewater treatment plant. *Science of the Total Environment*, *421–422*, 173–183. https://doi.org/10.1016/j.scitotenv.2012.01.061

Gao, P., Munir, M., & Xagoraraki, I. (2012b). Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional municipal wastewater treatment plant. *Science of The Total Environment*, *421–422*, 173–183. https://doi.org/10.1016/j.scitotenv.2012.01.061

Garcia-Beltran, W. F., Lam, E. C., St. Denis, K., Nitido, A. D., Garcia, Z. H., Hauser, B. M., Feldman, J., Pavlovic, M. N., Gregory, D. J., Poznansky, M. C., Sigal, A., Schmidt, A. G., Iafrate, A. J., Naranbhai, V., & Balazs, A. B. (2021). Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell*. https://doi.org/10.1016/j.cell.2021.03.013

Garner, E., Organiscak, M., Dieter, L., Shingleton, C., Haddix, M., Joshi, S., Pruden, A., Ashbolt, N. J., Medema, G., & Hamilton, K. A. (2021). Towards risk assessment for antibiotic resistant pathogens in recycled water: A systematic review and summary of research needs. *Environmental Microbiology*, *23*(12), 7355–7372. https://doi.org/10.1111/1462-2920.15804

Gaze, W. H., Zhang, L., Abdouslam, N. A., Hawkey, P. M., Calvo-Bado, L., Royle, J., Brown, H., Davis, S., Kay, P., Boxall, A. B. A., & Wellington, E. M. H. (2011). Impacts of anthropogenic activity on the ecology of class 1 integrons and integron-associated genes in the environment. *ISME Journal*. https://doi.org/10.1038/ismej.2011.15

Gerrity, D., Papp, K., Stoker, M., Sims, A., & Frehner, W. (2021). Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and incidence/prevalence considerations. *Water Research X*, *10*, 100086. https://doi.org/10.1016/j.wroa.2020.100086

Gillings, M. R., Gaze, W. H., Pruden, A., Smalla, K., Tiedje, J. M., & Ku, J. (2014). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *The ISME Journal*, *9*(6), 1269–1279. https://doi.org/10.1038/ismej.2014.226

Gillings, M. R., Gaze, W. H., Pruden, A., Smalla, K., Tiedje, J. M., & Zhu, Y.-G. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *The ISME Journal*, *9*(6), 1269–1279. https://doi.org/10.1038/ismej.2014.226

GISAID. (n.d.). Retrieved January 5, 2022, from https://www.gisaid.org/

Göbel, A., Thomsen, A., McArdell, C. S., Joss, A., & Giger, W. (2005). Occurrence and Sorption Behavior of Sulfonamides, Macrolides, and Trimethoprim in Activated Sludge Treatment. *Environmental Science & Technology*, *39*(11), 3981–3989. https://doi.org/10.1021/es048550a

Godbold, G. D., Kappell, A. D., LeSassier, D. S., Treangen, T. J., & Ternus, K. L. (2022). Categorizing Sequences of Concern by Function To Better Assess Mechanisms of Microbial Pathogenesis. *Infection and Immunity*, *90*(5), e00334-21. https://doi.org/10.1128/iai.00334-21

Goren, M. G., Carmeli, Y., Schwaber, M. J., Chmelnitsky, I., Schechner, V., & Navon-Venezia, S. (2010). Transfer of Carbapenem-Resistant Plasmid from Klebsiella pneumoniae ST258 to Escherichia coli in Patient. *Emerging Infectious Diseases*, *16*(6), 1014–1017. https://doi.org/10.3201/eid1606.091671

Göttig, S., Gruber, T. M., Stecher, B., Wichelhaus, T. A., & Kempf, V. A. J. (2015). In Vivo Horizontal Gene Transfer of the Carbapenemase OXA-48 During a Nosocomial Outbreak. *Clinical Infectious Diseases*, *60*(12), 1808–1815. https://doi.org/10.1093/cid/civ191

Govender, K. N., & Eyre, D. W. (2022). *Benchmarking taxonomic classifiers with Illumina and Nanopore sequence data for clinical metagenomic diagnostic applications* (p. 2022.01.11.475979). bioRxiv. https://doi.org/10.1101/2022.01.11.475979

Gregory, D. A., Wieberg, C. G., Wenzel, J., Lin, C.-H., & Johnson, M. C. (2021). Monitoring SARS-CoV-2 Populations in Wastewater by Amplicon Sequencing and Using the Novel Program SAM Refiner. *Viruses*, *13*(8), 1647. https://doi.org/10.3390/v13081647

Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke, D., Bouquet, J., Somasekar, S., Linnen, J. M., Dodd, R., Mulembakani, P., Schneider, B. S., Muyembe-Tamfum, J.-J., Stramer, S. L., & Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Medicine*, *7*(1), 99. https://doi.org/10.1186/s13073-015-0220-9

Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., Tan, A. L., Paul, L. M., Brackney, D. E., Grewal, S., Gurfield, N., Van Rompay, K. K. A., Isern, S., Michael, S. F., Coffey, L. L., Loman, N. J., & Andersen, K. G. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biology*, *20*(1), 8. https://doi.org/10.1186/s13059-018-1618-7

Gu, Q., Sun, M., Lin, T., Zhang, Y., Wei, X., Wu, S., Zhang, S., Pang, R., Wang, J., Ding, Y., Liu, Z., Chen, L., Chen, W., Lin, X., Zhang, J., Chen, M., Xue, L., & Wu, Q. (2022). Characteristics of Antibiotic Resistance Genes and Antibiotic-Resistant Bacteria in Full-Scale Drinking Water Treatment System Using Metagenomics and Culturing. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/article/10.3389/fmicb.2021.798442

Guo, F., Zhang, T., Li, B., Wang, Z., Ju, F., & Liang, Y. (2019). Mycobacterial species and their contribution to cholesterol degradation in wastewater treatment plants. *Scientific Reports*, *9*(1), 836. https://doi.org/10.1038/s41598-018-37332-w

Guo, J., Li, J., Chen, H., Bond, P. L., & Yuan, Z. (2017). Metagenomic analysis reveals wastewater treatment plants as hotspots of antibiotic resistance genes and mobile genetic elements. *Water Research*, *123*, 468–478. https://doi.org/10.1016/j.watres.2017.07.002

Gupta, A., Vlamakis, H., Shoemaker, N., & Salyers, A. A. (2003). A New Bacteroides Conjugative Transposon That Carries an ermB Gene. *Applied and Environmental Microbiology*, *69*(11), 6455–6463. https://doi.org/10.1128/AEM.69.11.6455-6463.2003

Han, X. M., Hu, H. W., Shi, X. Z., Wang, J. T., Han, L. L., Chen, D., & He, J. Z. (2016). Impacts of reclaimed water irrigation on soil antibiotic resistome in urban parks of Victoria, Australia. *Environmental Pollution*, *211*, 48–57. https://doi.org/10.1016/j.envpol.2015.12.033

Han, X.-M., Hu, H.-W., Shi, X.-Z., Wang, J.-T., Han, L.-L., Chen, D., & He, J.-Z. (2016). Impacts of reclaimed water irrigation on soil antibiotic resistome in urban parks of Victoria, Australia. *Environmental Pollution*, *211*, 48–57. https://doi.org/10.1016/j.envpol.2015.12.033

Harb, M., & Hong, P. Y. (2017a). Anaerobic membrane bioreactor effluent reuse: A review of microbial safety concerns. *Fermentation*, *3*(3). https://doi.org/10.3390/fermentation3030039

Harb, M., & Hong, P.-Y. (2017b). Molecular-based detection of potentially pathogenic bacteria in membrane bioreactor (MBR) systems treating municipal wastewater: A case study. *Environmental Science and Pollution Research*, *24*(6), 5370–5380. https://doi.org/10.1007/s11356-016-8211-y

Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S. J., & Robertson, D. L. (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*, *19*(7), 409–424. https://doi.org/10.1038/s41579-021-00573-0

He, H., Zhou, P., Shimabuku, K. K., Fang, X., Li, S., Lee, Y., & Dodd, M. C. (2019). Degradation and Deactivation of Bacterial Antibiotic Resistance Genes during Exposure to Free Chlorine, Monochloramine, Chlorine Dioxide, Ozone, Ultraviolet Light, and Hydroxyl Radical. *Environmental Science and Technology*, *53*(4), 2013–2026. https://doi.org/10.1021/acs.est.8b04393

Heald-Sargent, T., & Gallagher, T. (2012). Ready, Set, Fuse! The Coronavirus Spike Protein and Acquisition of Fusion Competence. *Viruses*, *4*(4), 557–580. https://doi.org/10.3390/v4040557

Heijnen, L., Elsinga, G., de Graaf, M., Molenkamp, R., Koopmans, M. P. G., & Medema, G. (2021). Droplet digital RT-PCR to detect SARS-CoV-2 signature mutations of variants of concern in wastewater. *Science of The Total Environment*, *799*, 149456. https://doi.org/10.1016/j.scitotenv.2021.149456

Hendriksen, R. S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S. K., Kjeldgaard, J., Kaas, R. S., Clausen, P. T. L. C., Vogt, J. K., Leekitcharoenphon, P., van de Schans, M. G.

M., Zuidema, T., de Roda Husman, A. M., Rasmussen, S., Petersen, B., … Aarestrup, F. M. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications*, *10*(1), 1124. https://doi.org/10.1038/s41467-019-08853-3

Hillary, L. S., Farkas, K., Maher, K. H., Lucaci, A., Thorpe, J., Distaso, M. A., Gaze, W. H., Paterson, S., Burke, T., Connor, T. R., McDonald, J. E., Malham, S. K., & Jones, D. L. (2021). Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. *Water Research*, *200*, 117214. https://doi.org/10.1016/j.watres.2021.117214

Hong, P.-Y., Julian, T. R., Pype, M.-L., Jiang, S. C., Nelson, K. L., Graham, D., Pruden, A., & Manaia, C. M. (2018). Reusing Treated Wastewater: Consideration of the Safety Aspects Associated with Antibiotic-Resistant Bacteria and Antibiotic Resistance Genes. *Water*, *10*(3), 244. https://doi.org/10.3390/w10030244

Ibáñez de Aldecoa, A. L., Zafra, O., & González-Pastor, J. E. (2017). Mechanisms and Regulation of Extracellular DNA Release and Its Biological Roles in Microbial Communities. *Frontiers in Microbiology*, *8*, 1390. https://doi.org/10.3389/fmicb.2017.01390

Iskandar, K., Molinier, L., Hallit, S., Sartelli, M., Hardcastle, T. C., Haque, M., Lugova, H., Dhingra, S., Sharma, P., Islam, S., Mohammed, I., Naina Mohamed, I., Hanna, P. A., Hajj, S. E., Jamaluddin, N. A. H., Salameh, P., & Roques, C. (2021). Surveillance of antimicrobial resistance in low- and middle-income countries: A scattered picture. *Antimicrobial Resistance & Infection Control*, *10*(1), 63. https://doi.org/10.1186/s13756-021-00931-w

Itokawa, K., Sekizuka, T., Hashino, M., Tanaka, R., & Kuroda, M. (2020). Disentangling primer interactions improves SARS-CoV-2 genome sequencing by multiplex tiling PCR. *PLOS ONE*, *15*(9), e0239403. https://doi.org/10.1371/journal.pone.0239403

Jang, H. M., Kim, Y. B., Choi, S., Lee, Y., Shin, S. G., Unno, T., & Kim, Y. M. (2018). Prevalence of antibiotic resistance genes from effluent of coastal aquaculture, South Korea. *Environmental Pollution*. https://doi.org/10.1016/j.envpol.2017.10.006

Jangra, S., Ye, C., Rathnasinghe, R., Stadlbauer, D., Alshammary, H., Amoako, A. A., Awawda, M. H., Beach, K. F., Bermúdez-González, M. C., Chernet, R. L., Eaker, L. Q., Ferreri, E. D., Floda, D. L., Gleason, C. R., Kleiner, G., Jurczyszak, D., Matthews, J. C., Mendez, W. A., Mulder, L. C. F., … Schotsaert, M. (2021). SARS-CoV-2 spike E484K mutation reduces antibody neutralisation. *The Lancet Microbe*, *0*(0). https://doi.org/10.1016/S2666-5247(21)00068-9

Jia, S., Zhang, X.-X., Miao, Y., Zhao, Y., Ye, L., Li, B., & Zhang, T. (2017). Fate of antibiotic resistance genes and their associations with bacterial community in livestock breeding wastewater and its receiving river water. *Water Research*, *124*, 259–268. https://doi.org/10.1016/j.watres.2017.07.061

Jin, M., Liu, L., Wang, D., Yang, D., Liu, W., Yin, J., Yang, Z., Wang, H., Qiu, Z., Shen, Z., Shi, D., Li, H., Guo, J., & Li, J. (2020). Chlorine disinfection promotes the exchange of antibiotic resistance genes across bacterial genera by natural transformation. *The ISME Journal*, *14*(7), 1847–1856. https://doi.org/10.1038/s41396-020-0656-9

Johnson, R., Sharma, J. R., Ramharack, P., Mangwana, N., Kinnear, C., Viraragavan, A., Glanzmann, B., Louw, J., Abdelatif, N., Reddy, T., Surujlal-Naicker, S., Nkambule, S., Mahlangeni, N., Webster, C., Mdhluli, M., Gray, G., Mathee, A., Preiser, W., Muller, C., & Street, R. (2022). Tracking the circulating SARS-CoV-2 variant of concern in South Africa using wastewater-based epidemiology. *Scientific Reports*, *12*(1), 1182. https://doi.org/10.1038/s41598-022-05110-4

Jong, M. C., Su, J. Q., Bunce, J. T., Harwood, C. R., Snape, J. R., Zhu, Y. G., & Graham, D. W. (2018). Co-optimization of sponge-core bioreactors for removing total nitrogen and antibiotic resistance genes from domestic wastewater. *Science of the Total Environment*, *634*, 1417–1423. https://doi.org/10.1016/j.scitotenv.2018.04.044

Joss, A., Keller, E., Alder, A. C., Göbel, A., McArdell, C. S., Ternes, T., & Siegrist, H. (2005). Removal of pharmaceuticals and fragrances in biological wastewater treatment. *Water Research*, *39*(14), 3139–3152. https://doi.org/10.1016/j.watres.2005.05.031

Ju, F., Beck, K., Yin, X., Maccagnan, A., McArdell, C. S., Singer, H. P., Johnson, D. R., Zhang, T., & Bürgmann, H. (2019). Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes. *The ISME Journal*, *13*(2), 346–360. https://doi.org/10.1038/s41396-018-0277-8

Juraschek, K., Borowiak, M., Tausch, S. H., Malorny, B., Käsbohrer, A., Otani, S., Schwarz, S., Meemken, D., Deneke, C., & Hammerl, J. A. (2021). Outcome of Different Sequencing and Assembly Approaches on the Detection of Plasmids and Localization of Antimicrobial Resistance Genes in Commensal Escherichia coli. *Microorganisms*, *9*(3), 598. https://doi.org/10.3390/microorganisms9030598

Kalmar, L., Gupta, S., Kean, I. R. L., Ba, X., Hadjirin, N., Lay, E. M., Vries, S. P. W. de, Bateman, M., Bartlet, H., Hernandez-Garcia, J., Tucker, A. W., Restif, O., Stevens, M. P., Wood, J. L. N., Maskell, D. J., Grant, A. J., & Holmes, M. A. (2022). HAM-ART: An optimised culture-free Hi-C metagenomics pipeline for tracking antimicrobial resistance genes in complex microbial communities. *PLOS Genetics*, *18*(3), e1009776. https://doi.org/10.1371/journal.pgen.1009776

Kantor, R. S., Nelson, K. L., Greenwald, H. D., & Kennedy, L. C. (2021). Challenges in Measuring the Recovery of SARS-CoV-2 from Wastewater. *Environmental Science & Technology*, *55*(6), 3514–3519. https://doi.org/10.1021/acs.est.0c08210

Kappell, A. D., Kimbell, L. K., Seib, M. D., Carey, D. E., Choi, M. J., Kalayil, T., & Fujimoto, M. (2018). *Environmental Science anaerobic membrane bioreactor treating primary clarifier effluent at 20 ° C*. https://doi.org/10.1039/c8ew00270c

Kappell, A. D., Kimbell, L. K., Seib, M. D., Carey, D. E., Choi, M. J., Kalayil, T., Fujimoto, M., Zitomer, D. H., & McNamara, P. J. (2018a). Removal of antibiotic resistance genes in an anaerobic membrane bioreactor treating primary clarifier effluent at 20 °C. *Environmental Science: Water Research and Technology*, *4*(11), 1783–1793. https://doi.org/10.1039/c8ew00270c

Kappell, A. D., Kimbell, L. K., Seib, M. D., Carey, D. E., Choi, M. J., Kalayil, T., Fujimoto, M., Zitomer, D. H., & McNamara, P. J. (2018b). Removal of antibiotic resistance genes in an anaerobic membrane bioreactor treating primary clarifier

effluent at 20 °C. *Environmental Science: Water Research & Technology*, *4*(11), 1783–1793. https://doi.org/10.1039/C8EW00270C

Karkman, A., Johnson, T. A., Lyra, C., Stedtfeld, R. D., Tamminen, M., Tiedje, J. M., & Virta, M. (2016). High-throughput quantification of antibiotic resistance genes from an urban wastewater treatment plant. *FEMS Microbiology Ecology*, *92*(fiw014). https://doi.org/10.1093/femsec/fiw014

Karthikeyan, S., Levy, J. I., Hoff, P. D., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing uncovers early cryptic SARS-CoV-2 variant transmission. *Nature*, 2021.12.21.21268143. https://doi.org/10.1101/2021.12.21.21268143

Keen, E. C., Bliskovsky, V. V., Malagon, F., Baker, J. D., Prince, J. S., Klaus, J. S., & Adhya, S. L. (2017). Novel "Superspreader" Bacteriophages Promote Horizontal Gene Transfer by Transformation. *MBio*, *8*(1). https://doi.org/10.1128/mBio.02115-16

Kennedy, C.-A., Fanning, S., Karczmarczyk, M., Byrne, B., Monaghan, Á., Bolton, D., & Sweeney, T. (2017). Characterizing the Multidrug Resistance of non-O157 Shiga Toxin-Producing Escherichia coli Isolates from Cattle Farms and Abattoirs. *Microbial Drug Resistance*, *23*(6), 781–790. https://doi.org/10.1089/mdr.2016.0082

Kim, D., Song, L., Breitwieser, F. P., & Salzberg, S. L. (2016). Centrifuge: Rapid and sensitive classification of metagenomic sequences. *Genome Research*, *26*(12), 1721–1729. https://doi.org/10.1101/gr.210641.116

Kittredge, H. A., Dougherty, K. M., & Evans, S. E. (2022). Dead but Not Forgotten: How Extracellular DNA, Moisture, and Space Modulate the Horizontal Transfer of Extracellular Antibiotic Resistance Genes in Soil. *Applied and Environmental Microbiology*, *88*(7), e02280-21. https://doi.org/10.1128/aem.02280-21

Knapp, C. W., Engemann, C. A., Hanson, M. L., Keen, P. L., Hall, K. J., & Graham, D. W. (2008). Indirect Evidence of Transposon-Mediated Selection of Antibiotic Resistance Genes in Aquatic Systems at Low-Level Oxytetracycline Exposures. *Environmental Science & Technology*, *42*(14), 5348–5353. https://doi.org/10.1021/es703199g

Kneis, D., Berendonk, T. U., Forslund, S. K., & Hess, S. (2022). Antibiotic Resistance Genes in River Biofilms: A Metagenomic Approach toward the Identification of Sources and Candidate Hosts. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.2c00370

Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research*, *46*(6), e35. https://doi.org/10.1093/nar/gkx1321

Kutilova, I., Medvecky, M., Leekitcharoenphon, P., Munk, P., Masarikova, M., Davidova-Gerzova, L., Jamborova, I., Bortolaia, V., Pamp, S. J., & Dolejska, M. (2021). Extended-spectrum beta-lactamase-producing Escherichia coli and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: Culture-based and metagenomic approaches. *Environmental Research*, *193*, 110487. https://doi.org/10.1016/j.envres.2020.110487

La Rosa, G., Mancini, P., Bonanno Ferraro, G., Veneri, C., Iaconelli, M., Bonadonna, L., Lucentini, L., & Suffredini, E. (2021). SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Science of The Total Environment*, *750*, 141711. https://doi.org/10.1016/j.scitotenv.2020.141711

Laht, M., Karkman, A., Voolaid, V., Ritz, C., Tenson, T., Virta, M., & Kisand, V. (2014). Abundances of Tetracycline, Sulphonamide and Beta-Lactam Antibiotic Resistance Genes in Conventional Wastewater Treatment Plants (WWTPs) with Different Waste Load. *PLOS ONE*, *9*(8), e103705. https://doi.org/10.1371/journal.pone.0103705

Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences*, *82*(20), 6955–6959. https://doi.org/10.1073/pnas.82.20.6955

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821. https://doi.org/10.1038/nbt.2676

LaTurner, Z. W., Zong, D. M., Kalvapalle, P., Gamas, K. R., Terwilliger, A., Crosby, T., Ali, P., Avadhanula, V., Santos, H. H., Weesner, K., Hopkins, L., Piedra, P. A., Maresso, A. W., & Stadler, L. B. (2021a). Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology. *Water Research*, *197*, 117043. https://doi.org/10.1016/j.watres.2021.117043

LaTurner, Z. W., Zong, D. M., Kalvapalle, P., Gamas, K. R., Terwilliger, A., Crosby, T., Ali, P., Avadhanula, V., Santos, H. H., Weesner, K., Hopkins, L., Piedra, P. A., Maresso, A. W., & Stadler, L. B. (2021b). Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology. *Water Research*, *197*, 117043. https://doi.org/10.1016/j.watres.2021.117043

Layton, B., Kaya, D., Kelly, C., Williamson, K., Bachhuber, S., Banwarth, P., Bethel, J., Carter, K., Dalziel, B., Dasenko, M., Geniza, M., Gibbon, D., Girard, A.-M., Haggerty, R., Higley, K., Hynes, D., Lubchenco, J., McLaughlin, K., Nieto, F. J., … Radniecki, T. (2021). *Wastewater-based epidemiology predicts COVID-19 community prevalence* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-690031/v1

Le, T. H., Ng, C., Tran, N. H., Chen, H., & Gin, K. Y. H. (2018). Removal of antibiotic residues, antibiotic resistant bacteria and antibiotic resistance genes in municipal wastewater by membrane bioreactor systems. *Water Research*, *145*, 498–508. https://doi.org/10.1016/j.watres.2018.08.060

Lee, H., Swamikannu, X., Radulescu, D., Kim, S., & Stenstrom, M. K. (2007). Design of stormwater monitoring programs. *Water Research*, *41*(18), 4186–4196. https://doi.org/10.1016/j.watres.2007.05.016

Lee, J., Jeon, J. H., Shin, J., Jang, H. M., Kim, S., Song, M. S., & Kim, Y. M. (2017). Quantitative and qualitative changes in antibiotic resistance genes after passing

through treatment processes in municipal wastewater treatment plants. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2017.06.250

Lee, K., Kim, D.-W., Lee, D.-H., Kim, Y.-S., Bu, J.-H., Cha, J.-H., Thawng, C. N., Hwang, E.-M., Seong, H. J., Sul, W. J., Wellington, E. M. H., Quince, C., & Cha, C.-J. (2020). Mobile resistome of human gut and pathogen drives anthropogenic bloom of antibiotic resistance. *Microbiome*, *8*(1), 2. https://doi.org/10.1186/s40168-019-0774-7

Lee, S., Fan, P., Liu, T., Yang, A., Boughton, R. K., Pepin, K. M., Miller, R. S., & Jeong, K. C. (2022). Transmission of antibiotic resistance at the wildlife-livestock interface. *Communications Biology*, *5*(1), 1–12. https://doi.org/10.1038/s42003-022-03520-8

Lee, W. L., Gu, X., Armas, F., Chandra, F., Chen, H., Wu, F., Leifels, M., Xiao, A., Chua, F. J. D., Kwok, G. W., Jolly, S., Lim, C. Y., Thompson, J., & Alm, E. J. (2021). *Quantitative SARS-CoV-2 tracking of variants Delta, Delta plus, Kappa and Beta in wastewater by allele-specific RT-qPCR* (p. 2021.08.03.21261298). https://doi.org/10.1101/2021.08.03.21261298

Leggett, R. M., Alcon-Giner, C., Heavens, D., Caim, S., Brook, T. C., Kujawska, M., Martin, S., Peel, N., Acford-Palmer, H., Hoyles, L., Clarke, P., Hall, L. J., & Clark, M. D. (2020). Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. *Nature Microbiology*, *5*(3), 430–442. https://doi.org/10.1038/s41564-019-0626-z

Leiva, A. M., Piña, B., & Vidal, G. (2021). Antibiotic resistance dissemination in wastewater treatment plants: A challenge for the reuse of treated wastewater in agriculture. *Reviews in Environmental Science and Bio/Technology*, *20*(4), 1043–1072. https://doi.org/10.1007/s11157-021-09588-8

Li, B., Di, D. Y. W., Saingam, P., Jeon, M. K., & Yan, T. (2021). Fine-Scale Temporal Dynamics of SARS-CoV-2 RNA Abundance in Wastewater during A COVID-19 Lockdown. *Water Research*, *197*, 117093. https://doi.org/10.1016/j.watres.2021.117093

Li, B., Ju, F., Cai, L., & Zhang, T. (2015). Profile and Fate of Bacterial Pathogens in Sewage Treatment Plants Revealed by High-Throughput Metagenomic Approach. *Environmental Science & Technology*, *49*(17), 10492–10502. https://doi.org/10.1021/acs.est.5b02345

Li, B., Yang, Y., Ma, L., Ju, F., Guo, F., Tiedje, J. M., & Zhang, T. (2015). Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *The ISME Journal*, *9*(11), 2490–2502. https://doi.org/10.1038/ismej.2015.59

Li, D., Zeng, S., He, M., & Gu, A. Z. (2016). Water Disinfection Byproducts Induce Antibiotic Resistance-Role of Environmental Pollutants in Resistance Phenomena. *Environmental Science & Technology*, *50*(6), 3193–3201. https://doi.org/10.1021/acs.est.5b05113

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J., & Zhang, X. (2018). Application of Environmental DNA Metabarcoding for Predicting Anthropogenic Pollution in Rivers. *Environmental Science & Technology*, acs.est.8b03869. https://doi.org/10.1021/acs.est.8b03869

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv:1303.3997 [q-Bio]*. http://arxiv.org/abs/1303.3997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., Zhang, Z., Duan, J., Li, N., Li, B., Song, T., Sardar, M. F., Lv, X., & Zhu, C. (2020). Electrochemical disinfection of secondary effluent from a wastewater treatment plant: Removal efficiency of ARGs and variation of antibiotic resistance in surviving bacteria. *Chemical Engineering Journal*, *392*, 123674. https://doi.org/10.1016/j.cej.2019.123674

Li, J., Wang, H., Mao, L., Yu, H., Yu, X., Sun, Z., Qian, X., Cheng, S., Chen, S., Chen, J., Pan, J., Shi, J., & Wang, X. (2020). Rapid genomic characterization of SARS-CoV-2 viruses from clinical specimens using nanopore sequencing. *Scientific Reports*, *10*(1), 17492. https://doi.org/10.1038/s41598-020-74656-y

Li, L., Nesme, J., Quintela-Baluja, M., Balboa, S., Hashsham, S., Williams, M. R., Yu, Z., Sørensen, S. J., Graham, D. W., Romalde, J. L., Dechesne, A., & Smets, B. F. (2021). Extended-Spectrum β-Lactamase and Carbapenemase Genes are Substantially and Sequentially Reduced during Conveyance and Treatment of Urban Sewage. *Environmental Science & Technology*, *55*(9), 5939–5949. https://doi.org/10.1021/acs.est.0c08548

Li, N., Sheng, G.-P., Lu, Y.-Z., Zeng, R. J., & Yu, H.-Q. (2017). Removal of antibiotic resistance genes from wastewater treatment plant effluent by coagulation. *Water Research*, *111*, 204–212. https://doi.org/10.1016/j.watres.2017.01.010

Li, Q., Wang, Y., Zou, Y. De, Liao, X. Di, Liang, J. B., Xin, W., & Wu, Y. B. (2015). Co-addition of manure increases the dissipation rates of tylosin A and the numbers of resistance genes in laboratory incubation experiments. *Science of the Total Environment*, *527–528*, 126–134. https://doi.org/10.1016/j.scitotenv.2015.04.117

Li, Q., Wu, J., Nie, J., Zhang, L., Hao, H., Liu, S., Zhao, C., Zhang, Q., Liu, H., Nie, L., Qin, H., Wang, M., Lu, Q., Li, X., Sun, Q., Liu, J., Zhang, L., Li, X., Huang, W., & Wang, Y. (2020). The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell*, *182*(5), 1284-1294.e9. https://doi.org/10.1016/j.cell.2020.07.012

Li, Z.-H., Yuan, L., Gao, S.-X., Wang, L., & Sheng, G.-P. (2019). Mitigated membrane fouling and enhanced removal of extracellular antibiotic resistance genes from wastewater effluent via an integrated pre-coagulation and microfiltration process. *Water Research*, *159*, 145–152. https://doi.org/10.1016/j.watres.2019.05.005

Liang, J., Mao, G., Yin, X., Ma, L., Liu, L., Bai, Y., Zhang, T., & Qu, J. (2020). Identification and quantification of bacterial genomes carrying antibiotic resistance genes and virulence factor genes for aquatic microbiological risk assessment. *Water Research*, *168*, 115160. https://doi.org/10.1016/j.watres.2019.115160

Liang, Z., & Keeley, A. (2013). Filtration recovery of extracellular DNA from environmental water samples. *Environmental Science and Technology*. https://doi.org/10.1021/es401342b

Liguori, K., Keenum, I., Davis, B. C., Calarco, J., Milligan, E., Harwood, V. J., & Pruden, A. (2022). Antimicrobial Resistance Monitoring of Water Environments: A Framework for Standardized Methods and Quality Control. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.1c08918

Lin, W., Zhang, M., Zhang, S., & Yu, X. (2016). Can chlorination co-select antibiotic-resistance genes? *Chemosphere*, *156*, 412–419. https://doi.org/10.1016/j.chemosphere.2016.04.139

Lin, X., Glier, M., Kuchinski, K., Mierlo, T. R.-V., McVea, D., Tyson, J. R., Prystajecky, N., & Ziels, R. M. (2021). *Assessing multiplex tiling PCR sequencing approaches for detecting genomic variants of SARS-CoV-2 in municipal wastewater* (p. 2021.05.26.21257861). https://doi.org/10.1101/2021.05.26.21257861

Liu, P., Jia, S., He, X., Zhang, X., & Ye, L. (2017). Different impacts of manure and chemical fertilizers on bacterial community structure and antibiotic resistance genes in arable soils. *Chemosphere*, *188*, 455–464. https://doi.org/10.1016/j.chemosphere.2017.08.162

Liu, S. S., Qu, H. M., Yang, D., Hu, H., Liu, W. L., Qiu, Z. G., Hou, A. M., Guo, J., Li, J. W., Shen, Z. Q., & Jin, M. (2018). Chlorine disinfection increases both intracellular and extracellular antibiotic resistance genes in a full-scale wastewater treatment plant. *Water Research*, *136*, 131–136. https://doi.org/10.1016/j.watres.2018.02.036

Liu, S.-S., Qu, H.-M., Yang, D., Hu, H., Liu, W.-L., Qiu, Z.-G., Hou, A.-M., Guo, J., Li, J.-W., Shen, Z.-Q., & Jin, M. (2018). Chlorine disinfection increases both intracellular and extracellular antibiotic resistance genes in a full-scale wastewater treatment plant. *Water Research*, *136*, 131–136. https://doi.org/10.1016/j.watres.2018.02.036

Liu, Z., Klümper, U., Liu, Y., Yang, Y., Wei, Q., Lin, J.-G., Gu, J.-D., & Li, M. (2019). Metagenomic and metatranscriptomic analyses reveal activity and hosts of antibiotic resistance genes in activated sludge. *Environment International*, *129*, 208–220. https://doi.org/10.1016/j.envint.2019.05.036

Liu, Z., Yao, J., Ma, H., Rukeya, A., Liang, Z., Du, W., & Chen, Y. (2022). Bacterial Hosts and Genetic Characteristics of Antibiotic Resistance Genes in Wastewater Treatment Plants of Xinjiang (China) Revealed by Metagenomics. *Applied Sciences*, *12*(6), 3100. https://doi.org/10.3390/app12063100

Lloréns-Rico, V., Vieira-Silva, S., Gonçalves, P. J., Falony, G., & Raes, J. (2021). Benchmarking microbiome transformations favors experimental quantitative approaches to address compositionality and sampling depth biases. *Nature Communications*, *12*(1), 3562. https://doi.org/10.1038/s41467-021-23821-6

Lou, E. G., Harb, M., Smith, A. L., & Stadler, L. B. (2020). Livestock manure improved antibiotic resistance gene removal during co-treatment of domestic wastewater in an anaerobic membrane bioreactor. *Environmental Science: Water Research & Technology*, *6*(10), 2832–2842. https://doi.org/10.1039/D0EW00387E

Lu, J., Zhang, Y., Wu, J., Wang, J., Zhang, C., & Lin, Y. (2019). Occurrence and spatial distribution of antibiotic resistance genes in the Bohai Sea and Yellow Sea areas, China. *Environmental Pollution*, *252*, 450–460. https://doi.org/10.1016/j.envpol.2019.05.143

Luo, Y., Yang, F., Mathieu, J., Mao, D., Wang, Q., & Alvarez, P. J. J. (2014). Proliferation of Multidrug-Resistant New Delhi Metallo-β-lactamase Genes in Municipal Wastewater Treatment Plants in Northern China. *Environmental Science & Technology Letters*, *1*(1), 26–30. https://doi.org/10.1021/ez400152e

Ma, J., Gu, J., Wang, X., Peng, H., Wang, Q., Zhang, R., Hu, T., & Bao, J. (2019). Effects of nano-zerovalent iron on antibiotic resistance genes during the anaerobic digestion of cattle manure. *Bioresource Technology*, *289*, 121688. https://doi.org/10.1016/j.biortech.2019.121688

Ma, L., Li, A. D., Yin, X. Le, & Zhang, T. (2017a). The Prevalence of Integrons as the Carrier of Antibiotic Resistance Genes in Natural and Man-Made Environments. *Environmental Science and Technology*. https://doi.org/10.1021/acs.est.6b05887

Ma, L., Li, A.-D., Yin, X.-L., & Zhang, T. (2017b). The Prevalence of Integrons as the Carrier of Antibiotic Resistance Genes in Natural and Man-Made Environments. *Environmental Science & Technology*, *51*(10), 5721–5728. https://doi.org/10.1021/acs.est.6b05887

Ma, L., Xia, Y., Li, B., Yang, Y., Li, L.-G., Tiedje, J. M., & Zhang, T. (2016). Metagenomic Assembly Reveals Hosts of Antibiotic Resistance Genes and the Shared Resistome in Pig, Chicken, and Human Feces. *Environmental Science & Technology*, *50*(1), 420–427. https://doi.org/10.1021/acs.est.5b03522

Maguire, F., Jia, B., Gray, K. L., Lau, W. Y. V., Beiko, R. G., & Brinkman, F. S. L. (2020). Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic Islands. *Microbial Genomics, 6*(10), e000436. https://doi.org/10.1099/mgen.0.000436

Majeed, H. J., Riquelme, M. V., Davis, B. C., Gupta, S., Angeles, L., Aga, D. S., Garner, E., Pruden, A., & Vikesland, P. J. (2021a). Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/article/10.3389/fmicb.2021.657954

Majeed, H. J., Riquelme, M. V., Davis, B. C., Gupta, S., Angeles, L., Aga, D. S., Garner, E., Pruden, A., & Vikesland, P. J. (2021b). Evaluation of Metagenomic-Enabled Antibiotic Resistance Surveillance at a Conventional Wastewater Treatment Plant. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/article/10.3389/fmicb.2021.657954

Mao, D., Luo, Y., Mathieu, J., Wang, Q., Feng, L., Mu, Q., Feng, C., & Alvarez, P. J. J. (2014). Persistence of Extracellular DNA in River Sediment Facilitates Antibiotic Resistance Gene Propagation. *Environmental Science & Technology*, *48*(1), 71–78. https://doi.org/10.1021/es404280v

Mao, D., Yu, S., Rysz, M., Luo, Y., Yang, F., Li, F., Hou, J., Mu, Q., & Alvarez, P. J. J. (2015a). Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants. *Water Research*, *85*, 458–466. https://doi.org/10.1016/j.watres.2015.09.010

Mao, D., Yu, S., Rysz, M., Luo, Y., Yang, F., Li, F., Hou, J., Mu, Q., & Alvarez, P. J. J. (2015b). Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants. *Water Research*, *85*, 458–466. https://doi.org/10.1016/j.watres.2015.09.010

Martinez-Sosa, D., Helmreich, B., Netter, T., Paris, S., Bischof, F., & Horn, H. (2011). Anaerobic submerged membrane bioreactor (AnSMBR) for municipal wastewater treatment under mesophilic and psychrophilic temperature conditions. *Bioresource Technology*, *102*(22), 10377–10385. https://doi.org/10.1016/j.biortech.2011.09.012

Marx, C., Günther, N., Schubert, S., Oertel, R., Ahnert, M., Krebs, P., & Kuehn, V. (2015). Mass flow of antibiotics in a wastewater treatment plant focusing on removal variations due to operational parameters. *Science of The Total Environment*, *538*, 779–788. https://doi.org/10.1016/j.scitotenv.2015.08.112

Mata, C., Miró, E., Alvarado, A., Garcillán-Barcia, M. P., Toleman, M., Walsh, T. R., de la cruz, F., & Navarro, F. (2012). Plasmid typing and genetic context of AmpC β-lactamases in enterobacteriaceae lacking inducible chromosomal ampC genes: Findings from a Spanish hospital 1999-2007. *Journal of Antimicrobial Chemotherapy*, *67*(1), 115–122. https://doi.org/10.1093/jac/dkr412

Mazel, D. (2006). Integrons: Agents of bacterial evolution. *Nature Reviews Microbiology*, *4*(8), 608–620. https://doi.org/10.1038/nrmicro1462

Mboowa, G., Mwesigwa, S., Kateete, D., Wayengera, M., Nasinghe, E., Katagirya, E., Katabazi, A. F., Kigozi, E., Kirimunda, S., Kamulegeya, R., Kabahita, J. M., Luutu, M. N., Nabisubi, P., Kanyerezi, S., Bagaya, B. S., & Joloba, M. L. (2021). Whole-genome sequencing of SARS-CoV-2 in Uganda: Implementation of the low-cost ARTIC protocol in resource-limited settings. *F1000Research*, *10*, 598. https://doi.org/10.12688/f1000research.53567.1

McConnell, M. M., Truelstrup Hansen, L., Jamieson, R. C., Neudorf, K. D., Yost, C. K., & Tong, A. (2018a). Removal of antibiotic resistance genes in two tertiary level municipal wastewater treatment plants. *Science of the Total Environment*, *643*, 292–300. https://doi.org/10.1016/j.scitotenv.2018.06.212

McConnell, M. M., Truelstrup Hansen, L., Jamieson, R. C., Neudorf, K. D., Yost, C. K., & Tong, A. (2018b). Removal of antibiotic resistance genes in two tertiary level municipal wastewater treatment plants. *Science of The Total Environment*, *643*, 292–300. https://doi.org/10.1016/j.scitotenv.2018.06.212

McEwen, S. A., & Collignon, P. J. (2018). Antimicrobial Resistance: A One Health Perspective. *Microbiology Spectrum*, *6*(2), 6.2.10. https://doi.org/10.1128/microbiolspec.ARBA-0009-2017

McInnes, R. S., McCallum, G. E., Lamberte, L. E., & van Schaik, W. (2020). Horizontal transfer of antibiotic resistance genes in the human gut microbiome. *Current Opinion in Microbiology*, *53*, 35–43. https://doi.org/10.1016/j.mib.2020.02.002

McKinney, C. W., & Pruden, A. (2012). Ultraviolet disinfection of antibiotic resistant bacteria and their antibiotic resistance genes in water and wastewater. *Environmental Science and Technology*, *46*(24), 13393–13400. https://doi.org/10.1021/es303652q

McNamara, P. J., LaPara, T. M., & Novak, P. J. (2014). The Impacts of Triclosan on Anaerobic Community Structures, Function, and Antimicrobial Resistance. *Environmental Science & Technology*, *48*(13), 7393–7400. https://doi.org/10.1021/es501388v

Meletis, G. (2016). Carbapenem resistance: Overview of the problem and future perspectives. *Therapeutic Advances in Infectious Disease*, *3*(1), 15–21. https://doi.org/10.1177/2049936115621709

Mendoza Grijalva, L., Brown, B., Cauble, A., & Tarpeh, W. A. (2022a). Diurnal Variability of SARS-CoV-2 RNA Concentrations in Hourly Grab Samples of Wastewater Influent during Low COVID-19 Incidence. *ACS ES&T Water*. https://doi.org/10.1021/acsestwater.2c00061

Mendoza Grijalva, L., Brown, B., Cauble, A., & Tarpeh, W. A. (2022b). Diurnal Variability of SARS-CoV-2 RNA Concentrations in Hourly Grab Samples of Wastewater Influent during Low COVID-19 Incidence. *ACS ES&T Water*, acsestwater.2c00061. https://doi.org/10.1021/acsestwater.2c00061

Mentasti, M., David, S., Sands, K., Khan, S., Davies, L., Turner, L., & Wootton, M. (2021). Rapid detection and differentiation of mobile colistin resistance (mcr-1 to mcr-10) genes by real-time PCR and melt-curve analysis. *Journal of Hospital Infection*, *110*, 148–155. https://doi.org/10.1016/j.jhin.2021.01.010

Metcalf & Eddy, I. (1991). *Wastewater engineering: Treatment, disposal, and reuse*. Third edition / revised by George Tchobanoglous, Franklin L. Burton. New York : McGraw-Hill, [1991] ©1991. https://search.library.wisc.edu/catalog/999636476602121

Mounaouer, B., & Abdennaceur, H. (2016). Modeling and kinetic characterization of wastewater disinfection using chlorine and UV irradiation. *Environmental Science and Pollution Research*, *23*(19), 19861–19875. https://doi.org/10.1007/s11356-016-7173-4

Mukherjee, M., Laird, E., Gentry, T. J., Brooks, J. P., & Karthikeyan, R. (2021). Increased Antimicrobial and Multidrug Resistance Downstream of Wastewater Treatment Plants in an Urban Watershed. *Frontiers in Microbiology*, *12*, 657353. https://doi.org/10.3389/fmicb.2021.657353

Munir, M., Wong, K., & Xagoraraki, I. (2010). Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Research*, *45*(2), 681–693. https://doi.org/10.1016/j.watres.2010.08.033

Munir, M., Wong, K., & Xagoraraki, I. (2011). Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Research*, *45*(2), 681–693. https://doi.org/10.1016/j.watres.2010.08.033

Munksgaard, D. G., & Young, J. C. (1980). Flow and Load Variations at Wastewater Treatment Plants. *Journal (Water Pollution Control Federation)*, *52*(8), 2131–2144. JSTOR.

Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Hamadani, B. H. K., Kumaran, E. A. P., McManigal, B., … Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet*, *399*(10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

Nagler, M., Podmirseg, S. M., Griffith, G. W., Insam, H., & Ascher-Jenull, J. (2018). The use of extracellular DNA as a proxy for specific microbial activity. *Applied Microbiology and Biotechnology*, *102*(6), 2885–2898. https://doi.org/10.1007/s00253-018-8786-y

Nelson, E. D., Do, H., Lewis, R. S., & Carr, S. A. (2011). Diurnal Variability of Pharmaceutical, Personal Care Product, Estrogen and Alkylphenol Concentrations in Effluent from a Tertiary Wastewater Treatment Facility. *Environmental Science & Technology*, *45*(4), 1228–1234. https://doi.org/10.1021/es102452f

Newton, R. J., McLellan, S. L., Dila, D. K., Vineis, J. H., Morrison, H. G., Eren, A. M., & Sogin, M. L. (2015). Sewage Reflects the Microbiomes of Human Populations. *MBio*, *6*(2), e02574-14. https://doi.org/10.1128/mBio.02574-14

Ngbede, E. O., Adekanmbi, F., Poudel, A., Kalalah, A., Kelly, P., Yang, Y., Adamu, A. M., Daniel, S. T., Adikwu, A. A., Akwuobu, C. A., Abba, P. O., Mamfe, L. M., Maurice, N. A., Adah, M. I., Lockyear, O., Butaye, P., & Wang, C. (2021). Concurrent Resistance to Carbapenem and Colistin Among Enterobacteriaceae Recovered From Human and Animal Sources in Nigeria Is Associated With Multiple Genetic Mechanisms. *Frontiers in Microbiology*, *12*, 740348. https://doi.org/10.3389/fmicb.2021.740348

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*(5), 824–834. https://doi.org/10.1101/gr.213959.116

Nyaruaba, R., Mwaliko, C., Dobnik, D., Neužil, P., Amoth, P., Mwau, M., Yu, J., Yang, H., & Wei, H. (2022). Digital PCR Applications in the SARS-CoV-2/COVID-19 Era: A Roadmap for Future Outbreaks. *Clinical Microbiology Reviews*, *0*(0), e00168-21. https://doi.org/10.1128/cmr.00168-21

O'Keeffe, J. (2021). Wastewater-based epidemiology: Current uses and future opportunities as a public health surveillance tool. *Environmental Health Review*, *64*(3), 44–52. https://doi.org/10.5864/d2021-015

Okitsu, N., Kaieda, S., Yano, H., Nakano, R., Hosaka, Y., Okamoto, R., Kobayashi, T., & Inoue, M. (2005). Characterization of ermB Gene Transposition by Tn1545 and Tn917 in Macrolide-Resistant Streptococcus pneumoniae Isolates. *Journal of Clinical Microbiology*, *43*(1), 168–173. https://doi.org/10.1128/JCM.43.1.168-173.2005

Oliveira, M., Leonardo, I. C., Nunes, M., Silva, A. F., & Barreto Crespo, M. T. (2021). Environmental and Pathogenic Carbapenem Resistant Bacteria Isolated from a Wastewater Treatment Plant Harbour Distinct Antibiotic Resistance Mechanisms. *Antibiotics*, *10*(9), 1118. https://doi.org/10.3390/antibiotics10091118

Oliveira, M., Nunes, M., Barreto Crespo, M. T., & Silva, A. F. (2020). The environmental contribution to the dissemination of carbapenem and (fluoro)quinolone resistance genes by discharged and reused wastewater effluents: The role of cellular and extracellular DNA. *Water Research*, *182*, 116011. https://doi.org/10.1016/j.watres.2020.116011

Otero, M. C. B., Murao, L. A. E., Limen, M. A. G., Gaite, P. L. A., Bacus, M. G., Acaso, J. T., Corazo, K., Knot, I. E., Sajonia, H., Reyes, F. L. de los, Jaraula, C. M. B., Baja, E. S., & Mundo, D. M. N. D. (2021). *Wastewater-Based Epidemiology and Whole-Genome Sequencing for Community-Level Surveillance of SARS-CoV-2 in Selected Urban Communities of Davao City, Philippines: A Pilot Study* (p. 2021.08.27.21262450). https://doi.org/10.1101/2021.08.27.21262450

Pärnänen, K. M. M., Narciso-da-Rocha, C., Kneis, D., Berendonk, T. U., Cacace, D., Do, T. T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A.,

Martinez, J. L., Michael, S. G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., … Manaia, C. M. (2019). Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances*, *5*(3), eaau9124. https://doi.org/10.1126/sciadv.aau9124

Payne, A., Holmes, N., Rakyan, V., & Loose, M. (2019). BulkVis: A graphical viewer for Oxford nanopore bulk FAST5 files. *Bioinformatics*, *35*(13), 2193–2198. https://doi.org/10.1093/bioinformatics/bty841

Pearman, W. S., Freed, N. E., & Silander, O. K. (2020). Testing the advantages and disadvantages of short- and long- read eukaryotic metagenomics using simulated reads. *BMC Bioinformatics*, *21*, 220. https://doi.org/10.1186/s12859-020-3528-4

Pei, M., Zhang, B., He, Y., Su, J., Gin, K., Lev, O., Shen, G., & Hu, S. (2019). State of the art of tertiary treatment technologies for controlling antibiotic resistance in wastewater treatment plants. *Environment International*, *131*, 105026. https://doi.org/10.1016/j.envint.2019.105026

Pei, R., Kim, S.-C., Carlson, K. H., & Pruden, A. (2006). Effect of River Landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Research*, *40*(12), 2427–2435. https://doi.org/10.1016/j.watres.2006.04.017

Peña, M., do Nascimento, T., Gouveia, J., Escudero, J., Gómez, A., Letona, A., Arrieta, J., & Fdz-Polanco, F. (2019). Anaerobic submerged membrane bioreactor (AnSMBR) treating municipal wastewater at ambient temperature: Operation and potential use for agricultural irrigation. *Bioresource Technology*, *282*, 285–293. https://doi.org/10.1016/j.biortech.2019.03.019

Plósz, B. Gy., Leknes, H., Liltved, H., & Thomas, K. V. (2010). Diurnal variations in the occurrence and the fate of hormones and antibiotics in activated sludge wastewater treatment in Oslo, Norway. *Science of The Total Environment*, *408*(8), 1915–1924. https://doi.org/10.1016/j.scitotenv.2010.01.042

Polo, D., Quintela-Baluja, M., Corbishley, A., Jones, D. L., Singer, A. C., Graham, D. W., & Romalde, J. L. (2020). Making waves: Wastewater-based epidemiology for COVID-19 – approaches and challenges for surveillance and prediction. *Water Research*, *186*, 116404. https://doi.org/10.1016/j.watres.2020.116404

Prado, T., Fumian, T. M., Mannarino, C. F., Resende, P. C., Motta, F. C., Eppinghaus, A. L. F., Chagas do Vale, V. H., Braz, R. M. S., de Andrade, J. da S. R., Maranhão, A. G., & Miagostovich, M. P. (2021). Wastewater-based epidemiology as a useful tool to track SARS-CoV-2 and support public health policies at municipal level in Brazil. *Water Research*, *191*, 116810. https://doi.org/10.1016/j.watres.2021.116810

Prestinaci, F., Pezzotti, P., & Pantosti, A. (2015). Antimicrobial resistance: A global multifaceted phenomenon. *Pathogens and Global Health*, *109*(7), 309–318. https://doi.org/10.1179/2047773215Y.0000000030

Prieto Riquelme, M. V., Garner, E., Gupta, S., Metch, J., Zhu, N., Blair, M. F., Arango-Argoty, G., Maile-Moskowitz, A., Li, A., Flach, C.-F., Aga, D. S., Nambi, I. M., Larsson, D. G. J., Bürgmann, H., Zhang, T., Pruden, A., & Vikesland, P. J. (2022). Demonstrating a Comprehensive Wastewater-Based Surveillance

Approach That Differentiates Globally Sourced Resistomes. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.1c08673

Qin, K., Wei, L., Li, J., Lai, B., Zhu, F., Yu, H., Zhao, Q., & Wang, K. (2020). A review of ARGs in WWTPs: Sources, stressors and elimination. *Chinese Chemical Letters*, *31*(10), 2603–2613. https://doi.org/10.1016/j.cclet.2020.04.057

Qin, Q., Chen, X., & Zhuang, J. (2015a). The fate and impact of pharmaceuticals and personal care products in agricultural soils irrigated with reclaimed water. *Critical Reviews in Environmental Science and Technology*, *45*(13), 1379–1408. https://doi.org/10.1080/10643389.2014.955628

Qin, Q., Chen, X., & Zhuang, J. (2015b). The Fate and Impact of Pharmaceuticals and Personal Care Products in Agricultural Soils Irrigated With Reclaimed Water. *Critical Reviews in Environmental Science and Technology*, *45*(13), 1379–1408. https://doi.org/10.1080/10643389.2014.955628

Qiu, Y., Zhang, J., Li, B., Wen, X., Liang, P., & Huang, X. (2018). A novel microfluidic system enables visualization and analysis of antibiotic resistance gene transfer to activated sludge bacteria in biofilm. *Science of The Total Environment*, *642*, 582–590. https://doi.org/10.1016/j.scitotenv.2018.06.012

Quach-Cu, J., Herrera-Lynch, B., Marciniak, C., Adams, S., Simmerman, A., & Reinke, R. A. (2018). The Effect of Primary, Secondary, and Tertiary Wastewater Treatment Processes on Antibiotic Resistance Gene (ARG) Concentrations in Solid and Dissolved Wastewater Fractions. *Water*, *10*(1), 37. https://doi.org/10.3390/w10010037

Quintela-Baluja, M., Abouelnaga, M., Romalde, J., Su, J.-Q., Yu, Y., Gomez-Lopez, M., Smets, B., Zhu, Y.-G., & Graham, D. W. (2019). Spatial ecology of a wastewater network defines the antibiotic resistance genes in downstream receiving waters. *Water Research*, *162*, 347–357. https://doi.org/10.1016/j.watres.2019.06.075

Radomski, N., Betelli, L., Moilleron, R., Haenn, S., Moulin, L., Cambau, E., Rocher, V., Gonçalves, A., & Lucas, F. S. (2011). Mycobacterium Behavior in Wastewater Treatment Plant, A Bacterial Model Distinct From Escherichia coli and Enterococci. *Environmental Science & Technology*, *45*(12), 5380–5386. https://doi.org/10.1021/es104084c

Rafraf, I. D., Lekunberri, I., Sànchez-Melsió, A., Aouni, M., Borrego, C. M., & Balcázar, J. L. (2016a). Abundance of antibiotic resistance genes in five municipal wastewater treatment plants in the Monastir Governorate, Tunisia. *Environmental Pollution*, *219*, 353–358. https://doi.org/10.1016/j.envpol.2016.10.062

Rafraf, I. D., Lekunberri, I., Sànchez-Melsió, A., Aouni, M., Borrego, C. M., & Balcázar, J. L. (2016b). Abundance of antibiotic resistance genes in five municipal wastewater treatment plants in the Monastir Governorate, Tunisia. *Environmental Pollution*, *219*, 353–358. https://doi.org/10.1016/j.envpol.2016.10.062

Rahman, A., Kang, S., Wang, W., Garg, A., Maile-Moskowitz, A., & Vikesland, P. J. (2021). Nanobiotechnology enabled approaches for wastewater based epidemiology. *TrAC Trends in Analytical Chemistry*, *143*, 116400. https://doi.org/10.1016/j.trac.2021.116400

Rice, E. W., Wang, P., Smith, A. L., & Stadler, L. B. (2020). Determining Hosts of Antibiotic Resistance Genes: A Review of Methodological Advances.

Environmental Science & Technology Letters, *7*(5), 282–291. https://doi.org/10.1021/acs.estlett.0c00202

Riquelme, M. V., Garner, E., Gupta, S., Metch, J., Zhu, N., Blair, M. F., Arango-Argoty, G., Maile-Moskowitz, A., Li, A., Flach, C.-F., Aga, D. S., Nambi, I., Larsson, D. G. J., Bürgmann, H., Zhang, T., Pruden, A., & Vikesland, P. J. (2021). Wastewater Based Epidemiology Enabled Surveillance of Antibiotic Resistance. *MedRxiv*, 2021.06.01.21258164. https://doi.org/10.1101/2021.06.01.21258164

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., Michael, I., & Fatta-Kassinos, D. (2013a). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of The Total Environment*, *447*, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M. C., Michael, I., & Fatta-Kassinos, D. (2013b). Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. *Science of The Total Environment*, *447*, 345–360. https://doi.org/10.1016/j.scitotenv.2013.01.032

Robertson, J., & Nash, J. H. E. (2018). MOB-suite: Software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics*, *4*(8), e000206. https://doi.org/10.1099/mgen.0.000206

Rodrigues, L., Machado, D., Couto, I., Amaral, L., & Viveiros, M. (2012). Contribution of efflux activity to isoniazid resistance in the Mycobacterium tuberculosis complex. *Infection, Genetics and Evolution*, *12*(4), 695–700. https://doi.org/10.1016/j.meegid.2011.08.009

Roman, V. L., Merlin, C., Virta, M. P. J., & Bellanger, X. (2021). EpicPCR 2.0: Technical and Methodological Improvement of a Cutting-Edge Single-Cell Genomic Approach. *Microorganisms*, *9*(8), 1649. https://doi.org/10.3390/microorganisms9081649

Rousham, E. K., Unicomb, L., & Islam, M. A. (2018). Human, animal and environmental contributors to antibiotic resistance in low-resource settings: Integrating behavioural, epidemiological and One Health approaches. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1876), 20180332. https://doi.org/10.1098/rspb.2018.0332

Rozwandowicz, M., Brouwer, M. S. M., Fischer, J., Wagenaar, J. A., Gonzalez-Zorn, B., Guerra, B., Mevius, D. J., & Hordijk, J. (2018). Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *Journal of Antimicrobial Chemotherapy*, *73*(5), 1121–1137. https://doi.org/10.1093/jac/dkx488

Sabri, N. A., van Holst, S., Schmitt, H., van der Zaan, B. M., Gerritsen, H. W., Rijnaarts, H. H. M., & Langenhoff, A. A. M. (2020). Fate of antibiotics and antibiotic resistance genes during conventional and additional treatment technologies in wastewater treatment plants. *Science of The Total Environment*, *741*, 140199. https://doi.org/10.1016/j.scitotenv.2020.140199

Sanjuán, R., & Domingo-Calap, P. (2021). Reliability of Wastewater Analysis for Monitoring COVID-19 Incidence Revealed by a Long-Term Follow-Up Study. *Frontiers in Virology*, *1*. https://www.frontiersin.org/article/10.3389/fviro.2021.776998

Sapoval, N., Lou, E., Hopkins, L., Ensor, K. B., Schneider, R., Treangen, T. J., & Stadler, L. B. (2021). *Enhanced Detection of Recently Emerged SARS-CoV-2 Variants of Concern in Wastewater* (p. 2021.09.08.21263279). https://doi.org/10.1101/2021.09.08.21263279

Schages, L., Wichern, F., Kalscheuer, R., & Bockmühl, D. (2020). Winter is coming – Impact of temperature on the variation of beta-lactamase and mcr genes in a wastewater treatment plant. *Science of The Total Environment*, *712*, 136499. https://doi.org/10.1016/j.scitotenv.2020.136499

Schoen, M. E., Jahne, M. A., Garland, J., Ramirez, L., Lopatkin, A. J., & Hamilton, K. A. (2021). Quantitative Microbial Risk Assessment of Antimicrobial Resistant and Susceptible Staphylococcus aureus in Reclaimed Wastewaters. *Environmental Science & Technology*, *55*(22), 15246–15255. https://doi.org/10.1021/acs.est.1c04038

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors – occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Schwermer, C. U., & Uhl, W. (2021). Calculating expected effects of treatment effectivity and river flow rates on the contribution of WWTP effluent to the ARG load of a receiving river. *Journal of Environmental Management*, *288*, 112445. https://doi.org/10.1016/j.jenvman.2021.112445

Sevim, V., Lee, J., Egan, R., Clum, A., Hundley, H., Lee, J., Everroad, R. C., Detweiler, A. M., Bebout, B. M., Pett-Ridge, J., Göker, M., Murray, A. E., Lindemann, S. R., Klenk, H.-P., O'Malley, R., Zane, M., Cheng, J.-F., Copeland, A., Daum, C., … Woyke, T. (2019). Shotgun metagenome data of a defined mock community using Oxford Nanopore, PacBio and Illumina technologies. *Scientific Data*, *6*(1), 285. https://doi.org/10.1038/s41597-019-0287-z

Simpson, B., Francce, D., & Lewis, B. (2013). *Wastewater Sampling*. https://www.epa.gov/sites/default/files/2015-06/documents/Wastewater-Sampling.pdf

Sivalingam, P., Poté, J., & Prabakar, K. (2020). Extracellular DNA (eDNA): Neglected and Potential Sources of Antibiotic Resistant Genes (ARGs) in the Aquatic Environments. *Pathogens*, *9*(11), 874. https://doi.org/10.3390/pathogens9110874

Smith, A. L., Skerlos, S. J., & Raskin, L. (2013). Psychrophilic anaerobic membrane bioreactor treatment of domestic wastewater. *Water Research*, *47*(4), 1655–1665. https://doi.org/10.1016/j.watres.2012.12.028

Smith, A. L., Stadler, L. B., Cao, L., Love, N. G., Raskin, L., & Skerlos, S. J. (2014). Navigating Wastewater Energy Recovery Strategies: A Life Cycle Comparison of Anaerobic Membrane Bioreactor and Conventional Treatment Systems with Anaerobic Digestion. *Environmental Science & Technology*, *48*(10), 5972–5981. https://doi.org/10.1021/es5006169

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1), 635. https://doi.org/10.1038/s41467-022-28246-3

Søraas, A., Sundsfjord, A., Sandven, I., Brunborg, C., & Jenum, P. A. (2013). Risk Factors for Community-Acquired Urinary Tract Infections Caused by ESBL-Producing Enterobacteriaceae –A Case–Control Study in a Low Prevalence Country. *PLOS ONE*, *8*(7), e69581. https://doi.org/10.1371/journal.pone.0069581

Souza, B. S., Dantas, R. F., Agulló-Barceló, M., Lucena, F., Sans, C., Esplugas, S., & Dezotti, M. (2013). Evaluation of UV/H2O2 for the disinfection and treatment of municipal secondary effluents for water reuse. *Journal of Chemical Technology & Biotechnology*, *88*(9), 1697–1706. https://doi.org/10.1002/jctb.4021

Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., A Weitz, D., Pitkänen, L. K., Vigneault, F., Virta, M. Pj., & Alm, E. J. (2016). Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *The ISME Journal*, *10*(2), 427–436. https://doi.org/10.1038/ismej.2015.124

Spencer, S., Spencer*, S., Tamminen*, M., Preheim, S., Guo, M., Briggs, A., Brito, I., Weitz, D., Pitkänen, L., Vigneault, F., Virta, M., & Alm, E. (2015). EpicPCR (Emulsion, Paired Isolation, and Concatenation PCR). *Protocol Exchange*. https://doi.org/10.1038/protex.2015.094

Stadler, L. B., Ensor, K. B., Clark, J. R., Kalvapalle, P., LaTurner, Z. W., Mojica, L., Terwilliger, A., Zhuo, Y., Ali, P., Avadhanula, V., Bertolusso, R., Crosby, T., Hernandez, H., Hollstein, M., Weesner, K., Zong, D. M., Persse, D., Piedra, P. A., Maresso, A. W., & Hopkins, L. (2020). Wastewater Analysis of SARS-CoV-2 as a Predictive Metric of Positivity Rate for a Major Metropolis. *MedRxiv*, 2020.11.04.20226191. https://doi.org/10.1101/2020.11.04.20226191

Stalder, T., Press, M. O., Sullivan, S., Liachko, I., & Top, E. M. (2019). Linking the resistome and plasmidome to the microbiome. *The ISME Journal*, *13*(10), 2437–2446. https://doi.org/10.1038/s41396-019-0446-4

Stedtfeld, R. D., Guo, X., Stedtfeld, T. M., Sheng, H., Williams, M. R., Hauschild, K., Gunturu, S., Tift, L., Wang, F., Howe, A., Chai, B., Yin, D., Cole, J. R., Tiedje, J. M., & Hashsham, S. A. (2018). Primer set 2.0 for highly parallel qPCR array targeting antibiotic resistance genes and mobile genetic elements. *FEMS Microbiology Ecology*, *94*(9), fiy130. https://doi.org/10.1093/femsec/fiy130

Stokes, H. W., & Gillings, M. R. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. In *FEMS Microbiology Reviews*. https://doi.org/10.1111/j.1574-6976.2011.00273.x

Sui, Q., Jiang, C., Zhang, J., Yu, D., Chen, M., Wang, Y., & Wei, Y. (2018). Does the biological treatment or membrane separation reduce the antibiotic resistance genes from swine wastewater through a sequencing-batch membrane bioreactor treatment process. *Environment International*, *118*, 274–281. https://doi.org/10.1016/j.envint.2018.06.008

Sui, Q., Zhang, J., Chen, M., Tong, J., Wang, R., & Wei, Y. (2016). Distribution of antibiotic resistance genes (ARGs) in anaerobic digestion and land application of swine wastewater. *Environmental Pollution*, *213*, 751–759. https://doi.org/10.1016/j.envpol.2016.03.038

Sui, Q., Zhang, J., Tong, J., Chen, M., & Wei, Y. (2017). Seasonal variation and removal efficiency of antibiotic resistance genes during wastewater treatment of swine

farms. *Environmental Science and Pollution Research*, *24*(10), 9048–9057. https://doi.org/10.1007/s11356-015-5891-7

Sun, M., Ye, M., Wu, J., Feng, Y., Wan, J., Tian, D., Shen, F., Liu, K., Hu, F., Li, H., Jiang, X., Yang, L., & Kengara, F. O. (2015). Positive relationship detected between soil bioaccessible organic pollutants and antibiotic resistance genes at dairy farms in Nanjing, Eastern China. *Environmental Pollution*, *206*, 421–428. https://doi.org/10.1016/j.envpol.2015.07.022

Sun, S., Geng, J., Li, B., Ma, L., Sun, X., Meng, F., Qi, H., & Shen, J. (2021). Temporal variations of antibiotic resistance genes in influents and effluents of a WWTP in cold regions. *Journal of Cleaner Production*, *328*, 129632. https://doi.org/10.1016/j.jclepro.2021.129632

Sun, W., Gu, J., Wang, X., Qian, X., & Tuo, X. (2018). Impacts of biochar on the environmental risk of antibiotic resistance genes and mobile genetic elements during anaerobic digestion of cattle farm wastewater. *Bioresource Technology*, *256*(January), 342–349. https://doi.org/10.1016/j.biortech.2018.02.052

Sun, W., Qian, X., Gu, J., Wang, X. J., & Duan, M. L. (2016). Mechanism and Effect of Temperature on Variations in Antibiotic Resistance Genes during Anaerobic Digestion of Dairy Manure. *Scientific Reports*, *6*(March), 1–9. https://doi.org/10.1038/srep30237

Suzuki, M. T., Taylor, L. T., & DeLong, E. F. (2000). Quantitative Analysis of Small-Subunit rRNA Genes in Mixed Microbial Populations via 5′-Nuclease Assays. *Applied and Environmental Microbiology*, *66*(11), 4605–4614. https://doi.org/10.1128/AEM.66.11.4605-4614.2000

Swift, C. L., Isanovic, M., Correa Velez, K. E., & Norman, R. S. (2021). Community-level SARS-CoV-2 sequence diversity revealed by wastewater sampling. *Science of The Total Environment*, *801*, 149691. https://doi.org/10.1016/j.scitotenv.2021.149691

Tedijanto, C., Olesen, S. W., Grad, Y. H., & Lipsitch, M. (2018). Estimating the proportion of bystander selection for antibiotic resistance among potentially pathogenic bacterial flora. *Proceedings of the National Academy of Sciences*, *115*(51), E11988–E11995. https://doi.org/10.1073/pnas.1810840115

Thakali, O., Brooks, J. P., Shahin, S., Sherchan, S. P., & Haramoto, E. (2020). Removal of Antibiotic Resistance Genes at Two Conventional Wastewater Treatment Plants of Louisiana, USA. *Water*, *12*(6), 1729. https://doi.org/10.3390/w12061729

Tyson, J. R., James, P., Stoddart, D., Sparks, N., Wickenhagen, A., Hall, G., Choi, J. H., Lapointe, H., Kamelian, K., Smith, A. D., Prystajecky, N., Goodfellow, I., Wilson, S. J., Harrigan, R., Snutch, T. P., Loman, N. J., & Quick, J. (2020). Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *BioRxiv*, 2020.09.04.283077. https://doi.org/10.1101/2020.09.04.283077

Uluseker, C., Kaster, K. M., Thorsen, K., Basiry, D., Shobana, S., Jain, M., Kumar, G., Kommedal, R., & Pala-Ozkok, I. (2021). A Review on Occurrence and Spread of Antibiotic Resistance in Wastewaters and in Wastewater Treatment Plants: Mechanisms and Perspectives. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/article/10.3389/fmicb.2021.717809

Van Poelvoorde, L. A. E., Gand, M., Fraiture, M.-A., De Keersmaecker, S. C. J., Verhaegen, B., Van Hoorde, K., Cay, A. B., Balmelle, N., Herman, P., & Roosens, N. (2021). Strategy to Develop and Evaluate a Multiplex RT-ddPCR in Response to SARS-CoV-2 Genomic Evolution. *Current Issues in Molecular Biology*, *43*(3), 1937–1949. https://doi.org/10.3390/cimb43030134

van Schaik, W. (2015). The human gut resistome. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1670), 20140087. https://doi.org/10.1098/rstb.2014.0087

Ventola, C. L. (2015). The Antibiotic Resistance Crisis. *Pharmacy and Therapeutics*, *40*(4), 277–283.

Vollmers, J., Wiegand, S., & Kaster, A.-K. (2017). Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective—Not Only Size Matters! *PLOS ONE*, *12*(1), e0169662. https://doi.org/10.1371/journal.pone.0169662

Wallace, J. S., Garner, E., Pruden, A., & Aga, D. S. (2018). Occurrence and transformation of veterinary antibiotics and antibiotic resistance genes in dairy manure treated by advanced anaerobic digestion and conventional treatment methods. *Environmental Pollution*, *236*, 764–772. https://doi.org/10.1016/j.envpol.2018.02.024

Wang, B., & Sun, D. (2015). Detection of NDM-1 carbapenemase-producing Acinetobacter calcoaceticus and Acinetobacter junii in environmental samples from livestock farms. *Journal of Antimicrobial Chemotherapy*, *70*(2), 611–613. https://doi.org/10.1093/jac/dku405

Wang, D. N., Liu, L., Qiu, Z. G., Shen, Z. Q., Guo, X., Yang, D., Li, J., Liu, W. L., Jin, M., & Li, J. W. (2016). A new adsorption-elution technique for the concentration of aquatic extracellular antibiotic resistance genes from large volumes of water. *Water Research*, *92*, 188–198. https://doi.org/10.1016/j.watres.2016.01.035

Wang, D.-N., Liu, L., Qiu, Z.-G., Shen, Z.-Q., Guo, X., Yang, D., Li, J., Liu, W., Jin, M., & Li, J.-W. (2016). A new adsorption-elution technique for the concentration of aquatic extracellular antibiotic resistance genes from large volumes of water. *Water Research*, *92*, 188–198. https://doi.org/10.1016/j.watres.2016.01.035

Wang, J., Ben, W., Zhang, Y., Yang, M., & Qiang, Z. (2015). Effects of thermophilic composting on oxytetracycline, sulfamethazine, and their corresponding resistance genes in swine manure. *Environmental Sciences: Processes and Impacts*. https://doi.org/10.1039/c5em00132c

Wang, J., & Chen, X. (2022). Removal of antibiotic resistance genes (ARGs) in various wastewater treatment processes: An overview. *Critical Reviews in Environmental Science and Technology*, *52*(4), 571–630. https://doi.org/10.1080/10643389.2020.1835124

Wang, J., Chu, L., Wojnárovits, L., & Takács, E. (2020). Occurrence and fate of antibiotics, antibiotic resistant genes (ARGs) and antibiotic resistant bacteria (ARB) in municipal wastewater treatment plant: An overview. *Science of The Total Environment*, *744*, 140997. https://doi.org/10.1016/j.scitotenv.2020.140997

Wang, J., Mao, D., Mu, Q., & Luo, Y. (2015). Fate and proliferation of typical antibiotic resistance genes in five full-scale pharmaceutical wastewater treatment plants.

Science of the Total Environment, *526*, 366–373. https://doi.org/10.1016/j.scitotenv.2015.05.046

Wang, P., Harb, M., Zarei-Baygi, A., Stadler, L. B., & Smith, A. L. (2019). *Comparative analysis of intracellular and extracellular antibiotic resistance gene abundance in anaerobic membrane bioreactor effluent* [Preprint]. Ecology. https://doi.org/10.1101/702076

Wang, Y., Batra, A., Schulenburg, H., & Dagan, T. (2022). Gene sharing among plasmids and chromosomes reveals barriers for antibiotic resistance gene transfer. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1842), 20200467. https://doi.org/10.1098/rstb.2020.0467

Wei, Z., Feng, K., Li, S., Zhang, Y., Chen, H., Yin, H., Xu, M., & Deng, Y. (2018). Exploring abundance, diversity and variation of a widespread antibiotic resistance gene in wastewater treatment plants. *Environment International*, *117*, 186–195. https://doi.org/10.1016/j.envint.2018.05.009

Wei, Z., Feng, K., Wang, Z., Zhang, Y., Yang, M., Zhu, Y.-G., Virta, M. P. J., & Deng, Y. (2021). High-Throughput Single-Cell Technology Reveals the Contribution of Horizontal Gene Transfer to Typical Antibiotic Resistance Gene Dissemination in Wastewater Treatment Plants. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.1c01250

Wen, Q., Yang, L., Duan, R., & Chen, Z. (2016). Monitoring and evaluation of antibiotic resistance genes in four municipal wastewater treatment plants in Harbin, Northeast China. *Environmental Pollution*, *212*, 34–40. https://doi.org/10.1016/j.envpol.2016.01.043

WHO. (2017). *WHO publishes list of bacteria for which new antibiotics are urgently needed*. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

Wilhelm, A., Toptan, T., Pallas, C., Wolf, T., Goetsch, U., Gottschalk, R., Vehreschild, M. J., Ciesek, S., & Widera, M. (2021). *Antibody-mediated neutralization of authentic SARS-CoV-2 B.1.617 variants harboring L452R and T478K/E484Q* (p. 2021.08.09.21261704). https://doi.org/10.1101/2021.08.09.21261704

Wilm, A., Aw, P. P. K., Bertrand, D., Yeo, G. H. T., Ong, S. H., Wong, C. H., Khor, C. C., Petric, R., Hibberd, M. L., & Nagarajan, N. (2012). LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research*, *40*(22), 11189–11201. https://doi.org/10.1093/nar/gks918

Wilton, T., Bujaki, E., Klapsa, D., Majumdar, M., Zambon, M., Fritzsche, M., Mate, R., & Martin, J. (2021). Rapid Increase of SARS-CoV-2 Variant B.1.1.7 Detected in Sewage Samples from England between October 2020 and January 2021. *MSystems*. https://doi.org/10.1128/mSystems.00353-21

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1), 257. https://doi.org/10.1186/s13059-019-1891-0

Wright, G. D. (2007). The antibiotic resistome: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, *5*(3), 175–186. https://doi.org/10.1038/nrmicro1614

Wright, G. D. (2010). Antibiotic resistance in the environment: A link to the clinic? *Current Opinion in Microbiology*, *13*(5), 589–594. https://doi.org/10.1016/j.mib.2010.08.005

Wurtzer, S., Waldman, P., Ferrier-Rembert, A., Frenois-Veyrat, G., Mouchel, J. M., Boni, M., Maday, Y., Marechal, V., & Moulin, L. (2021). Several forms of SARS-CoV-2 RNA can be detected in wastewaters: Implication for wastewater-based epidemiology and risk assessment. *Water Research*, *198*, 117183. https://doi.org/10.1016/j.watres.2021.117183

Xu, Y., Guo, C., Luo, Y., Lv, J., Zhang, Y., Lin, H., Wang, L., & Xu, J. (2016). Occurrence and distribution of antibiotics, antibiotic resistance genes in the urban rivers in Beijing, China. *Environmental Pollution*, *213*, 833–840. https://doi.org/10.1016/j.envpol.2016.03.054

Xu, Y., Xu, J., Mao, D., & Luo, Y. (2017). Effect of the selective pressure of sub-lethal level of heavy metals on the fate and distribution of ARGs in the catchment scale. *Environmental Pollution*, *220*, 900–908. https://doi.org/10.1016/j.envpol.2016.10.074

Yan, W., Guo, Y., Xiao, Y., Wang, S., Ding, R., Jiang, J., Gang, H., Wang, H., Yang, J., & Zhao, F. (2018). The changes of bacterial communities and antibiotic resistance genes in microbial fuel cells during long-term oxytetracycline processing. *Water Research*, *142*, 105–114. https://doi.org/10.1016/j.watres.2018.05.047

Yang, J., Yang, F., Ren, L., Xiong, Z., Wu, Z., Dong, J., Sun, L., Zhang, T., Hu, Y., Du, J., Wang, J., & Jin, Q. (2011). Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. *Journal of Clinical Microbiology*, *49*(10), 3463–3469. https://doi.org/10.1128/JCM.00273-11

Yang, L., Nielsen, A. A. K., Fernandez-Rodriguez, J., McClune, C. J., Laub, M. T., Lu, T. K., & Voigt, C. A. (2014). Permanent genetic memory with >1 byte capacity. *Nature Methods*, *11*(12), 1261–1266. https://doi.org/10.1038/nmeth.3147

Yang, L., Wen, Q., Chen, Z., Duan, R., & Yang, P. (2019). Impacts of advanced treatment processes on elimination of antibiotic resistance genes in a municipal wastewater treatment plant. *Frontiers of Environmental Science & Engineering*, *13*(3), 32. https://doi.org/10.1007/s11783-019-1116-5

Yang, Y., Che, Y., Liu, L., Wang, C., Yin, X., Deng, Y., Yang, C., & Zhang, T. (2022). Rapid absolute quantification of pathogens and ARGs by nanopore sequencing. *Science of The Total Environment*, *809*, 152190. https://doi.org/10.1016/j.scitotenv.2021.152190

Yang, Y., Li, B., Zou, S., Fang, H. H. P., & Zhang, T. (2014a). Fate of antibiotic resistance genes in sewage treatment plant revealed by metagenomic approach. *Water Research*, *62*, 97–106. https://doi.org/10.1016/j.watres.2014.05.019

Yang, Y., Li, B., Zou, S., Fang, H. H. P., & Zhang, T. (2014b). Fate of antibiotic resistance genes in sewage treatment plant revealed by metagenomic approach. *Water Research*, *62*, 97–106. https://doi.org/10.1016/j.watres.2014.05.019

Yaniv, K., Ozer, E., Shagan, M., Lakkakula, S., Plotkin, N., Bhandarkar, N. S., & Kushmaro, A. (2021). Direct RT-qPCR assay for SARS-CoV-2 variants of concern (Alpha, B.1.1.7 and Beta, B.1.351) detection and quantification in wastewater. *Environmental Research*, *201*, 111653. https://doi.org/10.1016/j.envres.2021.111653

Yin, X., Deng, Y., Ma, L., Wang, Y., Chan, L. Y. L., & Zhang, T. (2019). Exploration of the antibiotic resistome in a wastewater treatment plant by a nine-year longitudinal metagenomic study. *Environment International*, *133*, 105270. https://doi.org/10.1016/j.envint.2019.105270

Yin, X., Jiang, X.-T., Chai, B., Li, L., Yang, Y., Cole, J. R., Tiedje, J. M., & Zhang, T. (2018). ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. *Bioinformatics*, *34*(13), 2263–2270. https://doi.org/10.1093/bioinformatics/bty053

Yin, X., Yang, Y., Deng, Y., Huang, Y., Li, L., Chan, L. Y. L., & Zhang, T. (2022). An assessment of resistome and mobilome in wastewater treatment plants through temporal and spatial metagenomic analysis. *Water Research*, *209*, 117885. https://doi.org/10.1016/j.watres.2021.117885

Yoon, Y., Chung, H. J., Wen Di, D. Y., Dodd, M. C., Hur, H. G., & Lee, Y. (2017). Inactivation efficiency of plasmid-encoded antibiotic resistance genes during water treatment with chlorine, UV, and UV/H2O2. *Water Research*, *123*, 783–793. https://doi.org/10.1016/j.watres.2017.06.056

Young, J. C., Cleasby, J. L., & Baumann, E. R. (1978). Flow and Load Variations in Treatment Plant Design. *Journal of the Environmental Engineering Division*, *104*(2), 289–303. https://doi.org/10.1061/JEEGAV.0000750

Yuan, Q. Bin, Zhai, Y. F., Mao, B. Y., & Hu, N. (2018). Antibiotic resistance genes and intI1 prevalence in a swine wastewater treatment plant and correlation with metal resistance, bacterial community and wastewater parameters. *Ecotoxicology and Environmental Safety*, *161*, 251–259. https://doi.org/10.1016/j.ecoenv.2018.05.049

Yuan, Q.-B., Guo, M.-T., & Yang, J. (2015). Fate of Antibiotic Resistant Bacteria and Genes during Wastewater Chlorination: Implication for Antibiotic Resistance Control. *PLOS ONE*, *10*(3), e0119403. https://doi.org/10.1371/journal.pone.0119403

Yuan, Q.-B., Huang, Y.-M., Wu, W.-B., Zuo, P., Hu, N., Zhou, Y.-Z., & Alvarez, P. J. J. (2019a). Redistribution of intracellular and extracellular free & adsorbed antibiotic resistance genes through a wastewater treatment plant by an enhanced extracellular DNA extraction method with magnetic beads. *Environment International*, *131*(June), 104986. https://doi.org/10.1016/j.envint.2019.104986

Yuan, Q.-B., Huang, Y.-M., Wu, W.-B., Zuo, P., Hu, N., Zhou, Y.-Z., & Alvarez, P. J. J. (2019b). Redistribution of intracellular and extracellular free & adsorbed antibiotic resistance genes through a wastewater treatment plant by an enhanced extracellular DNA extraction method with magnetic beads. *Environment International*, *131*, 104986. https://doi.org/10.1016/j.envint.2019.104986

Zarei-baygi, A., Harb, M., Wang, P., Stadler, L. B., & Smith, A. L. (2019). Evaluating Antibiotic Resistance Gene Correlations with Antibiotic Exposure Conditions in Anaerobic Membrane Bioreactors [Research-article]. *Environmental Science & Technology*, *53*, 3599–3609. https://doi.org/10.1021/acs.est.9b00798

Zarei-Baygi, A., Harb, M., Wang, P., Stadler, L. B., & Smith, A. L. (2019). Evaluating Antibiotic Resistance Gene Correlations with Antibiotic Exposure Conditions in

Anaerobic Membrane Bioreactors [Research-article]. *Environmental Science and Technology*, *53*(7), 3599–3609. https://doi.org/10.1021/acs.est.9b00798

Zhang, A.-N., Gaston, J. M., Dai, C. L., Zhao, S., Poyet, M., Groussin, M., Yin, X., Li, L.-G., van Loosdrecht, M. C. M., Topp, E., Gillings, M. R., Hanage, W. P., Tiedje, J. M., Moniz, K., Alm, E. J., & Zhang, T. (2021). An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nature Communications*, *12*(1), 4765. https://doi.org/10.1038/s41467-021-25096-3

Zhang, G., Guan, Y., Zhao, R., Feng, J., Huang, J., Ma, L., & Li, B. (2020). Metagenomic and network analyses decipher profiles and co-occurrence patterns of antibiotic resistome and bacterial taxa in the reclaimed wastewater distribution system. *Journal of Hazardous Materials*, *400*, 123170. https://doi.org/10.1016/j.jhazmat.2020.123170

Zhang, T., Zhang, X.-X., & Ye, L. (2011). Plasmid Metagenome Reveals High Levels of Antibiotic Resistance Genes and Mobile Genetic Elements in Activated Sludge. *PLOS ONE*, *6*(10), e26041. https://doi.org/10.1371/journal.pone.0026041

Zhang, Y., Li, A., Dai, T., Li, F., Xie, H., Chen, L., & Wen, D. (2018a). Cell-free DNA: A Neglected Source for Antibiotic Resistance Genes Spreading from WWTPs. *Environmental Science and Technology*, *52*(1), 248–257. https://doi.org/10.1021/acs.est.7b04283

Zhang, Y., Li, A., Dai, T., Li, F., Xie, H., Chen, L., & Wen, D. (2018b). Cell-free DNA: A Neglected Source for Antibiotic Resistance Genes Spreading from WWTPs. *Environmental Science & Technology*, *52*(1), 248–257. https://doi.org/10.1021/acs.est.7b04283

Zhang, Y., Li, H., Gu, J., Qian, X., Yin, Y., Li, Y., Zhang, R., & Wang, X. (2016). Effects of adding different surfactants on antibiotic resistance genes and intI1 during chicken manure composting. *Bioresource Technology*, *219*, 545–551. https://doi.org/10.1016/j.biortech.2016.06.117

Zhang, Y., Snow, D. D., Parker, D., Zhou, Z., & Li, X. (2013). Intracellular and extracellular antimicrobial resistance genes in the sludge of livestock waste management structures. *Environmental Science and Technology*, *47*(18), 10206–10213. https://doi.org/10.1021/es401964s

Zhang, Z., Zhang, Q., Wang, T., Xu, N., Lu, T., Hong, W., Penuelas, J., Gillings, M., Wang, M., Gao, W., & Qian, H. (2022). Assessment of global health risk of antibiotic resistance genes. *Nature Communications*, *13*(1), 1553. https://doi.org/10.1038/s41467-022-29283-8

Zhao, R., Yu, K., Zhang, J., Zhang, G., Huang, J., Ma, L., Deng, C., Li, X., & Li, B. (2020). Deciphering the mobility and bacterial hosts of antibiotic resistance genes under antibiotic selection pressure by metagenomic assembly and binning approaches. *Water Research*, *186*, 116318. https://doi.org/10.1016/j.watres.2020.116318

Zhao, X., Wang, J., Zhu, L., & Wang, J. (2019). Field-based evidence for enrichment of antibiotic resistance genes and mobile genetic elements in manure-amended vegetable soils. *Science of the Total Environment*, *654*, 906–913. https://doi.org/10.1016/j.scitotenv.2018.10.446

Zheng, W., Huyan, J., Tian, Z., Zhang, Y., & Wen, X. (2020). Clinical class 1 integron-integrase gene – A promising indicator to monitor the abundance and elimination

of antibiotic resistance genes in an urban wastewater treatment plant. *Environment International*, *135*, 105372. https://doi.org/10.1016/j.envint.2019.105372

Zheng, W., Wen, X., Zhang, B., & Qiu, Y. (2019). Selective effect and elimination of antibiotics in membrane bioreactor of urban wastewater treatment plant. *Science of The Total Environment*, *646*, 1293–1303. https://doi.org/10.1016/j.scitotenv.2018.07.400

Zhu, L., Zhao, Y., Yang, K., Chen, J., Zhou, H., Chen, X., Liu, Q., & Wei, Z. (2019). Host bacterial community of MGEs determines the risk of horizontal gene transfer during composting of different animal manures. *Environmental Pollution*. https://doi.org/10.1016/j.envpol.2019.04.037