

ENVIRONMENTAL MICROBIAL SURVEILLANCE: FROM SOURCE TRACKING IN
WATERSHEDS TO PATHOGEN MONITORING IN SEWERSHEDS

By

Matthew Thomas Flood

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife-Doctor of Philosophy

2022

ABSTRACT

ENVIRONMENTAL MICROBIAL SURVEILLANCE: FROM SOURCE TRACKING IN WATERSHEDS TO PATHOGEN MONITORING IN SEWERSHEDS

By

Matthew Thomas Flood

Understanding of the connections between water and health, through the use of water quality monitoring, surveys and surveillance, can help to address the impacts of anthropomorphic changes on the environment. This study sought to understand these connections through the water quality monitoring within watershed basins as well as pathogen surveillance within sewersheds. Specifically, this dissertation sought to 1) understand the sources of pollution and their connections with land use in the various subsections of watersheds; 2) to find a cost-effective way to surveil the spread of SARS-CoV-2 using wastewater surveillance; and 3) to understand the differences in wastewater surveillance between communities.

Water quality monitoring using microbial source tracking (MST) was performed with a survey of five mixed-use watersheds in Michigan. Through the use of spatial clustering, it was found that temporal contamination was primarily driven by precipitation and its associated variables (e.g., streamflow, turbidity, overland flow), while spatial contamination is driven by land uses (e.g., septic tank density, tile drain proportions, and tillage). Additionally, porcine fecal contamination was more often correlated with nutrients in streams than either bovine or human contamination.

The development of a cost-effective workflow for the detection and quantification of SARS-COV-2 in wastewater was undertaken. Wastewater from communities around Michigan were collected and analyzed along with viral surrogates for SARS-CoV-2 to investigate different workflow options. The *Pseudomonas* phage Phi6 was seeded in different wastewater matrices to

test concentration and recovery by ultrafiltration-based method and polyethylene glycol (PEG) precipitation. The PEG method provided better virus recovery than the ultrafiltration-based methods as measured using RT-ddPCR.

The comparison of two communities (A and B) wastewater results for SARS-CoV-2 analyzed against case data was undertaken. These results were significantly correlated with cases in both communities, but the level of correlation differed based on spatial (e.g., zipcode vs county level cases) and temporal (e.g., date of symptom(s) onset vs. the referral date for cases) resolution. Wastewater surveillance was more representative of higher spatial resolution (zipcode data) of cases in both communities. When examining the temporal resolution of the communities, community B's wastewater results were more closely tied to the onset of symptoms and not the case referral date.

The ability to monitor indicators of pollution in watersheds and surveil etiological agents of disease in sewersheds provide non-intrusive methods for evaluating the potential risks and current burdens to community health. The first part of the work could be considered “downstream” monitoring identifying sources and potential exposures with the goal of reducing waterborne disease. While “upstream” monitoring was used for identifying the disease trends in the community and was focused on public health measures to prevent transmission. This project contributed novel methods, results and analysis providing valuable knowledge ultimately addressing the role of monitoring strategies to protect public health.

This dissertation is dedicated to Emma, Joan, and Chuck.
Their encouragement, support, and advice were indispensable to my success.

TABLE OF CONTENTS

LIST OF TABLES.....	viii
LIST OF FIGURES	x
KEY TO ABBREVIATIONS.....	xii
1.0 Introduction and Literature Review	1
1.1 Water Quality, Monitoring and Health.....	2
1.2 Long-term Monitoring Using Water Quality Indicators	5
1.2.1 Fecal Indicator Bacteria (FIB)	5
1.2.1.1 <i>Escherichia coli</i>	7
1.2.1.2 Enterococci	8
1.2.2 Microbial Source Tracking	9
1.2.2.1 Application of MST markers	12
1.2.3 Surveys for Sources of Fecal Pollution and Their Impact on Water Quality	16
1.2.4 Wastewater Surveillance and Human Health	18
1.2.4.1 COVID-19 (Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2))	18
1.3 Scientific Needs.....	22
1.4 Research Objectives	23
REFERENCES	27
2.0 Connecting microbial, nutrient, physiochemical, and land use variables for the evaluation of water quality within five mixed use watersheds	37
2.1 Abstract.....	38
2.2 Introduction	38
2.3 Materials and methods.....	41
2.3.1 Study area and sample collection.....	41
2.3.2 Flow, physiochemical, and nutrient methods	43
2.3.3 Water sample processing for microbial analysis	44
2.3.4 Microbial molecular analysis methods	45
2.3.5 Landscape data.....	47
2.3.6 Statistical analysis.....	47
2.3.6.1 Spatial clustering.....	48
2.4 Results	49
2.4.1 Water quality summary of five watersheds	49
2.4.2 Spatial and temporal trends in bacterial markers and nutrients	50
2.4.3 Statistical analysis.....	54
2.4.3.1 Cluster analysis	54
2.4.3.2 Correlation Results.....	55
2.5 Discussion.....	57
2.6 Conclusions	61
APPENDIX.....	62
REFERENCES	70

3.0 Methods evaluation for rapid concentration and quantification of SARS-CoV-2 in raw wastewater using droplet digital and quantitative RT-PCR.....	77
3.1 Abstract.....	78
3.2 Introduction	79
3.3 Materials and Methods	82
3.3.1 Wastewater samples and sampling sites	82
3.3.2 Virus stocks.....	82
3.3.3 Virus concentration methods and experiments	83
3.3.4 RNA extraction and quantification by RT-ddPCR and RT-qPCR	85
3.3.4.1 Detection of SARS-CoV-2, Phi6, and coronavirus OC43 using RT-ddPCR	85
3.3.4.2 Detection of SARS-CoV-2 using RT-qPCR.....	87
3.3.5 Data analysis	87
3.4 Results	89
3.4.1 Wastewater characteristics.....	89
3.4.2 Recovery of Phi6 from wastewater samples using ultrafiltration and PEG methods	89
3.4.3 Detection of SARS-CoV-2 in wastewater samples using ultrafiltration and PEG methods	90
3.4.4 Evaluation of rapid PEG approach for the detection of Phi6 and SARS-CoV-2 in wastewater.....	94
3.4.5 Evaluation of PEG precipitation using Phi6 and coronavirus OC43 as potential SARS-CoV-2 surrogates	96
3.5 Discussion.....	96
APPENDIX.....	100
REFERENCES	110
4.0 Understanding the Efficacy of Wastewater Surveillance for SARS-CoV-2 in Two Diverse Communities	119
4.1 Abstract.....	120
4.2 Introduction	120
4.3 Materials and Methods	122
4.3.1 Wastewater sampling and site descriptions	122
4.3.1.1 Wastewater treatment plant descriptions	122
4.3.1.2 Sample Collection Methods.....	123
4.3.2 Viral concentration and processing methods	124
4.3.3 Detection and enumeration of SARS-CoV-2 from wastewater using RT-ddPCR	125
4.3.4 COVID-19 case and vaccination data.....	127
4.3.5 Data analysis	128
4.4 Results	128
4.4.1 Comparison of SARS-CoV-2 concentrations found in wastewater against COVID-19 case data in two communities	128
4.4.1.2 Zipcode vs county level case data varying spatial resolution	134
4.4.2 Impact of vaccination rates on SARS-CoV-2 wastewater signals and case numbers ...	136
4.4.3 Detection of SARS-CoV-2 variants in WWTP A over time	138
4.5 Discussion.....	139
REFERENCES	143
5.0 Synopsis of Monitoring and Surveillance.....	148

REFERENCES	152
------------------	-----

LIST OF TABLES

Table 2.1 Primer and probes for ddPCR MST analysis.....	46
Table 2.A1 Land use percentages for each sampling site's drainage area.....	64
Table 2.A2 Physiochemical summary results by watershed.....	65
Table 2.A3 Ion summary results by watershed.....	66
Table 2.A4 Microbial summary results by watershed.....	67
Table 2.A5 Nutrient summary results by watershed.....	68
Table 3.1 Recovery efficiencies of ultrafiltration and PEG methods for the detection of Phi6 in seeded wastewater samples.	90
Table 3.2 The detection of SARS-CoV-2 genes (N1, N2, E) using ultrafiltration and PEG precipitation (with 16-hr incubation) concentration methods.....	92
Table 3.3 Mean recovery efficiencies of Phi6 in seeded wastewater samples using PEG precipitation method with and without overnight incubation.....	94
Table 3.4 Percent positive and mean concentrations of SARS-CoV-2 gene targets for PEG method with and without overnight incubation as measured using RT-ddPCR.....	95
Table 3.A1 Individual recovery efficiencies of ultrafiltration and PEG methods for the detection of Phi6 in seeded wastewater samples.....	102
Table 3.A2 Individual sample Phi6 percent recoveries for two PEG viral concentration methods.....	103
Table 3.A3 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for centrifugation method 1.....	104
Table 3.A4 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for centrifugation method 2.....	105
Table 3.A5 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for PEG precipitation (with 16-hr hold).....	106
Table 3.A6 Individual sample SARS-CoV-2 gene concentrations for PEG precipitation with 16-hour hold and without holding.....	107

Table 3.A7 Individual coefficients of variations for RT-ddPCR and RT-qPCR for three SARS-CoV-2 gene targets.....	108
Table 3.A8 Individual recovery efficiencies of the SARS-CoV-2 surrogates Phi6 and OC43 at two WWTPs.....	109
Table 4.1 County level demographics, COVID-19 vaccinations, and total COVID-19 cases/1,000 persons to date.....	123
Table 4.2 Primer and probe sequences.....	127
Table 4.3 Physiological measurements for two wastewater treatment plants.....	129
Table 4.4 Summary of wastewater monitoring results for two wastewater treatment plants.....	130

LIST OF FIGURES

Figure 2.1 Maps of studied watersheds showing watershed locations in the state of Michigan, major streams and waterbodies, general land use, and drainage areas for each sampling location. a) Little Pigeon River (LPR) watershed, b) Macatawa (MAC) watershed, c) Sandy Creek (SC) watershed, d) River Raisin (RR) watershed, and e) Kawkawlin (KAW) watershed.....	42
Figure 2.2 Microbial heatmaps for all watersheds and sampling months: a) <i>E. coli</i> Log ₁₀ MPN/100ml, b) <i>B. theta</i> Log ₁₀ GC/100ml, c) CowM2 Log ₁₀ GC/100ml, d) Pig2Bac Log ₁₀ GC/100ml. Cells colored with bright green were above the range depicted on the heatmap. These data points were removed from the depicted ranges to increase visibility of spatial and temporal patterns in the data.....	51
Figure 2.3 Phosphorus and Nitrogen species' heatmaps showing spatial and temporal distributions. a) total dissolved nitrogen (TDN), b) Ammonium (NH ₄ -N), c) total phosphorus (TP), d) total filterable phosphorus (TFP). Individual ranges for each nutrient species are to the right of each heatmaps. Cells colored with dark red were above the range depicted on the heatmap. These data points were removed from the depicted ranges to increase visibility of spatial and temporal patterns in the data.....	52
Figure 2.4 Cluster analysis results for streamflow, markers, land use, tillage, tile drain proportion, and septic tank density. Sites were clustered into up to three clusters representing “low”, “medium”, and “high” relative values for each category. Values for specific variables ranges of values are listed below each cluster.....	55
Figure 2.A1 Water quality variable heatmaps showing spatial and temporal distributions. a) potassium (k) mg/L; b) dissolved oxygen (D) mg/L; c) Conductivity (us/cm).....	69
Figure 3.1 Coefficients of variations for SARS-CoV-2 gene targets; a) sanitary sewer samples, b) WWTP influent samples, c) sanitary sewer samples with all concentration methods, d) WWTP influent samples with all concentration methods. Two-way ANOVA analysis results shown above each graph; ns: Not-significant, * p-value <0.05, ** p-value <0.01, *** p-value <0.001, **** p-value <0.0001.....	93
Figure 3.A1 qPCR standard curves for SARS-CoV-2 gene targets with slope, y intercept and R ² . a) N1 standard curve, b) N2 standard curve, c) E gene standard curve.....	101
Figure 4.1 Wastewater surveillance data (N1 gene target) for WWTP A (N=94) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N1 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.62$ $p < 0.0001$; $n = 86$ paired data points); b) N1 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 85$ paired data points).....	131

Figure 4.2 Wastewater surveillance data (N2 gene target) for WWTP A (N=94) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N2 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 86$ paired data points); b) N2 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.67$ $p < 0.0001$; $n = 85$ paired data points).....	132
Figure 4.3 Wastewater surveillance data (N1 gene target) for WWTP B (N=92) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N1 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.81$ $p < 0.0001$; $n = 61$ paired data points); b) N1 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.48$ $p < 0.0001$; $n = 61$ paired data points).....	133
Figure 4.4 Wastewater surveillance data (N2 gene target) for WWTP B (N=92) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N2 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 61$ paired data points); b) N2 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.38$ $p < 0.0001$; $n = 61$ paired data points).....	134
Figure 4.5 Wastewater surveillance data (N=94) (adjusted by flow and zipcode level population) and county level COVID-19 case data over time. a) WWTP A SARS-CoV-2 gene target results vs. county level case data for COVID-19 (N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.53$ $p < 0.0001$; $n = 93$ paired data points); b) WWTP B SARS-CoV-2 gene target results vs. county level COVID-19 case data (N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.45$ $p < 0.0001$; $n = 58$ paired data points).....	136
Figure 4.6 Vaccination rates and county level cases per 1,000 persons for communities A and B.....	137
Figure 4.7 Percent of population fully vaccinated compared with SARS-CoV-2 gene target loading (GC/Person/Day). a) WWTP A; b) WWTP B.....	138
Figure 4.8 Concentrations of SARS-CoV-2 variant genes for the Alpha, Delta, and Omicron variants over time. Samples positive for the N501Y and DEL 69-70 gene mutations indicate the potential presence of the Alpha variant. Samples positive for the T478K and L452R gene mutations indicate the presence of the Delta variant. Samples positive for the K417N and DEL 69-70 gene mutations indicate the presence of the Omicron variant. Empty squares represent Non-detects (NDs) and X's were samples that were not assayed for that marker.....	139

KEY TO ABBREVIATIONS

B. theta	<i>Bacteroides thetaiotamicron</i> 1-6 alpha mannase
BMP	Best Management Practice
Ca	Calcium
CDC	Centers for Disease Control and Prevention
CDL	Cropland Data Layer
Cl	Chlorine
COVID-19	Coronavirus Disease 2019
CSO	Combined Sewer Overflow
CWA	Clean Water Act
ddPCR	Droplet Digital Polymerase Chain Reaction
DNA	Deoxyribonucleic Acid
DO	Dissolved Oxygen
DON	Dissolved Organic Nitrogen
dPCR	Digital Polymerase Chain Reaction
<i>E. coli</i>	<i>Escherichia coli</i>
FIB	Fecal Indicator Bacteria
FRP	Filterable Reactive Phosphorus
GC	Gene copies
GIS	Global Information Systems
hr	Hour
K	Potassium

KAW	Kawkawlin
LDL	Lower Detection Limit
LLOQ	Lower Limit of Quantification
LPR	Little Pigeon River
LULC	Land Use/ Land Cover
MAC	Macatawa
Mg	Magnesium
MGD	Million Gallons per Day
MHV	Murine Hepatitis Virus
mL	Milliliter
MST	Microbial Source Tracking
MSU	Michigan State University
Na	Sodium
ND	Non-detect
NFQ	Non-fluorescent Quencher
NH ₄	Ammonium
NLCD	National Land Cover Database
NO ₂	Nitrite
NO ₃	Nitrate
NPDES	National Pollutant Discharge Elimination System
NPOC	Non-purgeable Organic Carbon
NTC	Non-template Control
NTU	Nephelometric Turbidity Unit

PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PFU	Plaque Forming Unit
qPCR	Quantitative Polymerase Chain Reaction
QA/QC	Quality Assurance/Quality Control
RED	Rare Event Detection
RdRP	RNA-dependent RNA Polymerase
RNA	Ribonucleic Acid
RR	River Raisin
RT-ddPCR	Reverse Transcription Droplet Digital PCR
RT-qPCR	Reverse Transcription Quantitative PCR
SC	Sandy Creek
SARS-CoV-2	Severe Acute Respiratory Syndrome-Coronavirus-2
SO ₄	Sulfate
SRP	Soluble Reactive Phosphorus
SSO	Sanitary Sewer Overflow
TDN	Total Dissolved Nitrogen
TFP	Total Filterable Phosphorus
TGEV	Transmissible Gastroenteritis Virus
TMDL	Total Maximum Daily Load
TP	Total Phosphorus
TRP	Total Reactive Phosphorus
TSS	Total Suspended Solids

USEPA	United States Environmental Protection Agency
USGS	United States Geological Service
WBE	Wastewater-Based Epidemiology
WHO	World Health Organization
WWTP	Wastewater Treatment Plant
μL	Microliter

1.0 Introduction and Literature Review

1.1 Water Quality, Monitoring and Health

Water quality monitoring is essential to understanding the connections between water and human health. This can be used to address the impact of anthropomorphic changes on the environment and used to address the human condition. According to the World Health Organization (WHO), monitoring can be differentiated and conducted at three different levels. The first of which is *monitoring* as a long-term systemic review of water quality in order to define the status and trends through standardized measurements and observations (WHO, 1998). The second level is that of an intensive *survey* of water quality over a finite duration for a specific purpose (WHO, 1998). The final level is that of water quality *surveillance*. This level involves close examination of specific measurements and observations continuously in order to inform management and operational activities (WHO, 1998). The second and third levels are particularly important in their roles in disease surveillance which not only aim to collect specific measurements and observations but ensure that information and any conclusions drawn from these approaches are disseminated. Specifically, regarding the occurrence of disease(s) in pre-defined populations, this will inform public health actions for the purpose of the reduction of morbidity and mortality (CDC, 2012; Orenstein and Bernier, 1990).

Most long-term microbial water quality monitoring programs focus on surrogates and indicator organisms, such as the bacteria *Escherichia coli* or enterococci (Scott et al., 2002). This is in part due to their ease of use and ubiquitous prevalence in the gastrointestinal tracts and subsequently the fecal matter of mammals and birds. One of the other reasons for the widespread use of microbial surrogates is due to the difficulty in testing and monitoring specific pathogens, yet this is changing because of new technology. As there are many different pathogens of interest, conducting individual tests for each on a routine basis is complicated and can be

expensive especially when disease incidences in the surrounding populations are low. Other new targets used for what is now known as Microbial Source Tracking (MST) identifies species-specific sources of fecal contamination and has developed as a response to this issue for both long-term monitoring and surveys. The idea is that when the source of the fecal pollution/contamination is known then both hazard identification for risk assessment and management of that source can be undertaken (Scott et al., 2002; Heymann and Rodier, 2001).

Both long-term monitoring, surveys and surveillance approaches allow for the examination of the impact on the health of entire populations instead of individuals. For example, the monitoring of waters using appropriate transects for microbial (e.g., indicators, MST markers, and pathogens), chemical (e.g., nutrients such as phosphorus and nitrogen) and physiochemical (e.g., soil runoff) contamination combined with information from global information systems (GIS) on land use and land cover change allows for the evaluation of the whole watershed with regards to human impact on water quality (Nnane et al., 2011; Heaney et al., 2015; Verhoughstraete et al., 2015; Sowah et al., 2017; Pascual-Benito et al., 2020; Ballesté et al., 2020). Additionally, intensive surveys can help to understand the effects and impacts of disturbances which need special diagnostics and generally take a snapshot of the water quality. These results in turn can be assessed in regard to how the water pollution could pose a risk to human health. This same design of intensive surveys evaluating an entire watershed, can also be applied to human wastewater and the constructed environment of sewersheds (Sinclair et al., 2008; Xagorarakis and O'Brien, 2020). The surveillance of wastewater and sewersheds for surrogates and pathogens in order to determine the disease-burden in a population has been referred to as wastewater-based epidemiology (WBE) (Kitajima et al., 2020; Orive et al., 2020). WBE is employed through the indirect surveillance of pathogens which are excreted from

infected individuals providing an estimate of the disease prevalence in the population. This in turn helps to drive more rapid decision making in the form of policy, regulations, and public health orders, for the protection of the public. This has proved especially useful during the current COVID-19 pandemic, where infected individuals excrete the virus independent of the wide variety of symptoms and asymptomatic infections (Kitajima et al., 2020; Orive et al., 2020).

The development of new affordable detection assays using recent advancements in molecular technologies, which provide more accurate and precise measurements of low-level targets in complex matrices, has proved essential in the implementation of monitoring strategies for MST markers and the surveillance of pathogens (Carlson, 2003; Roslev and Bukh, 2011). In particular, the development of digital polymerase chain reaction (dPCR) has been shown to be indispensable in monitoring environmental waters for MST markers and surveilling the complex matrices of human wastewater for pathogens such as SARS-CoV-2, the etiological agent of COVID-19. While the term digital PCR (dPCR) was first used in 1999 and the technique had been independently developed multiple times in 1990 and 1991, the development of quantitative PCR (qPCR) in 1996 overshadowed dPCR until 2007 when new instrumentation allowed for the more widespread use of the technology (Vogelstein and Kinzler, 1999; Morley, 2014). The multitude of variable sources, both point and non-point sources, and the significant impacts of fecal contamination on water and health drove the ongoing development of advanced molecular techniques and assays. These are now crucial in the protection of public health through water quality monitoring and surveillance.

While indicator organisms allow for the standardized determination of the general sense (comparing one water way to another) of contamination in water sources their presence and

abundance is not always correlated with increased risk from waterborne pathogens. The ongoing contamination of surface waters from non-point sources of pollution, namely failed septic tanks and agricultural run-off, in Michigan were the major drivers for taking on an MST study (Dubrovsky et al., 2010; Yang et al., 2016; Verhougstraete et al., 2015; Nshimyimana et al., 2018). This was done in order to understand how land use and human impacts are linked to fecal pollution. While these studies were conducted to understand the connections between the land use and human health, the 2020 worldwide Coronavirus pandemic drove the need to examine and understand the overall health and risk at a community level as initially testing was not meeting the demand for protecting community health. The presence of asymptomatic infected persons complicated the estimation of the disease burdens in communities. Human surveillance conducted by state and federal epidemiologist and the health departments remains difficult. Determining a suitable processing and concentration method for SARS-CoV-2 in wastewater along with a matching robust detection assay which is inhibitor resistant and able to consistently detect low levels of the virus are critical research needs. Furthermore, determining the levels of SARS-CoV-2 present in Michigan wastewaters and how those levels are correlated with known cases of COVID-19 is essential in understanding the value of WBE during the Coronavirus pandemic.

1.2 Long-term Monitoring Using Water Quality Indicators

1.2.1 Fecal Indicator Bacteria (FIB)

Fecal indicator bacteria (FIB) are one of the most broadly used targets for the detection and assessment of water pollution associated with fecal inputs including wastewater. Due to the abundance of diverse waterborne pathogens, FIB are used as surrogates due to their greater

abundance and correspondingly easier detection, while also providing a noticeably lower cost of analysis (Griffin et al., 2001; Horan, 2003). While over the years a set of criteria for an ideal/optimal FIB have been proposed and refined, no single FIB has been able to meet all of them (Bonde, 1966; WHO, 1993; Grabow, 1996; Godfree et al., 1997; Colford et al., 2007). These criteria include 1) the FIB is suitable for use in all waters (e.g., freshwater, marine, streams, lakes, oceans); 2) there is cooccurrence of the FIB and the pathogen(s) of interest; 3) there is a greater abundance of FIB than pathogens; 4) the FIB has greater or equal survivability as pathogens in environmental waters, and also through wastewater and water treatment processes; 5) the FIB do not regrow in the environment (e.g., water and/or sediments); 6) the FIB is easily and reliably detected; 7) the FIB is non-pathogenic; and 8) the method of detection is relatively low cost. While FIB are certainly useful and represent standards used around the world, there are several issues when trying to relate them to the presence and concentration of pathogens. The ability of most bacteria used as FIB to potentially regrow in the environment, the lack of source identification, relatively long incubation times (18-26 hours), and inconclusive relationships between the presence of FIB and pathogens, limit the usefulness of FIB (Schwab, 2007; McLellan et al., 2007). Due to these limitations of individual FIB, approaches using multiple FIB or the combination of FIB with other methods of fecal pollution detection, such as microbial source tracking allows for source identification of pollution sources (McLellan, 2004).

Two of the current most commonly used FIB are *Escherichia coli* (*E. coli*) and enterococci, which were adopted as the leading indicators for fecal pollution in fresh and marine waters in 1986 (USEPA, 1986; USEPA, 2012). These FIB are used to evaluate the recreational water quality of surface waters in the United States. The US EPA has suggested recreational water quality limits based off of three epidemiological studies conducted in 1982 and 1984 in

marine and fresh waters which correlated enterococci and *E. coli* densities with cases of gastroenteritis (Cabelli et al., 1982; Dufour et al., 1984). The US national water quality criteria are 104 enterococci per 100 ml of marine water, a single sample maximum of 61 enterococci per 100 ml in freshwater, and a mean of 235 *E. coli* per 100 ml in freshwater (USEPA, 1986; Wade et al., 2008). While the US EPA set these criteria under the Clean Water Act (1972) each state is responsible for setting their own recreational water quality standard based off of these criteria. While many states chose to directly adopt the criteria set for by the US EPA, Michigan set its recreational water quality standard to a daily maximum as a geometric mean of three individual samples taken from the recreational area which spatially represent that area, of 300 colony forming units (CFUs) of *E. coli* per 100 ml. Additionally, surface waters are also subject to a partial body contact maximum of 1,000 *E. coli* per 100 ml.

1.2.1.1 *Escherichia coli*

Escherichia coli are gram-negative rod-shaped bacteria which are facultative anaerobes and fecal coliforms. These bacteria are distinguished from other fecal coliforms by their ability to grow at 45°C in conjunction with their lack of urease, and their ability to catalyze B-D-glucopyranosiduronic acid through the presence of B-D glucuronidase (Toranzo and McFeters, 1997). *E. coli* are commonly found in the lower intestines of warm-blooded animals including mammals and birds (Winfield and Groisman, 2003). While most strains of *E. coli* are non-pathogenic there are several infectious strains which do cause disease with the primary exposure route being the fecal-oral route (Rice, 2003; Bischoff et al., 2005). The main advantages of using *E. coli* as FIB include its wide adoption and continued use, relative low cost, and its previous use in epidemiological studies where it was correlated with incidences of gastroenteritis recreational

waters in fresh and marine waters (Dufour et al., 1984; Prüss, 1998; Rompré et al., 2002; Wade et al., 2003; Zmirou et al., 2003; Wade et al., 2006; USEPA, 2009). While *E. coli* has several disadvantages (e.g., long incubation time, uneven distribution in the water column), its main disadvantage is that it has been shown to replicate outside of its natural hosts (McLellan et al., 2001; Winfield and Groisman, 2003; Whitman and Nevers, 2004; Vital et al., 2008; Thupaki et al., 2010).

1.2.1.2 Enterococci

Enterococci are gram positive non-spore forming cocci consisting of species from two genera (*Enterococcus* and *Streptococcus*) which are found in the feces of warm-blooded animals. Similar to *E. coli*, enterococci are generally non-pathogenic and are primarily spread through the fecal-oral exposure route. One of the main differences between enterococci and *E. coli* which may determine which is best used as the FIB of choice is enterococci's greater resistance to chlorination and ability to persist longer in the environment, which provides a more protective estimation of water quality compared to *E. coli* (Gleeson and Gray, 1997). Enterococci also share the same disadvantage as *E. coli* in that they are unable to distinguish the sources of fecal contamination. This means while we can use these FIB to quantify pollution, they are limited in providing further information which may help to identify and remediate sources.

While FIB, such as *E. coli* and enterococci, are able to help identify and quantify fecal contamination in water they lack the ability to distinguish specific sources of pollution.

1.2.2 Microbial Source Tracking

Microbial source tracking (MST) markers provide a much-needed approach to distinguish and quantify specific sources of fecal pollution. Microbial source tracking is a field that has matured over the last 20 some years (Scott et al, 2002; Boehm et al., 2013; Steinbacher et al., 2021). MST has been accomplished with two different analysis schemes, library-dependent and library-independent methods. Library-dependent methods rely on a reference library of known gene targets to match sample DNA, while library-independent methods target a single known gene associated with a specific source of pollution and look only for that gene. Library-dependent methods are less widely used than -independent methods due to their limitations including their use of a reference library, complicated analysis of the data, and lack of quantification. Whereas library-independent methods also return less false positive and false-negatives (Griffith et al., 2003).

Using polymerase chain reaction (PCR) assays, MST can be applied with host-specific markers associated with a single species or type of animal to identify sources of pollution (Scott et al., 2002; Santo Domingo et al., 2007). MST molecular approaches have a few distinct advantages over cultivation-based methods including higher target sensitivity and specificity, faster results (4 hours vs 18-24 hours for cultivation), and the potential for a more automated processing of samples (Girones et al., 2010). However, current molecular MST methods do not have the ability to distinguish between viable and non-viable cells/organisms (Girones et al., 2010). This can lead to an over estimation of associated risk due to non-viable organisms or legacy pollution effects being detected. However, the advantages that MST offers make these assays useful diagnostic tools allowing for pollution source identification versus routine monitoring for using indicator organisms.

The development of MST markers starts with the identification of potential target genes followed by, primer and probe development, and finally validation testing to evaluate the specificity of the target sequence for a particular species along with the sensitivity of the assay to detect that target in the environment (Walters and Fields, 2006; Santo Domingo et al., 2007). Validation of each marker is limited by the study design used during the evaluation. The greater the number of fecal samples from different species the assay is tested against along with the number of unique target species fecal samples provide more or less confidence in a new marker. For example, two bovine markers CowM2 and BacCow-UCD were both reported as having >50% host specificity in their initial publications, but while the CowM2 marker was tested against 204 samples for its evaluation the BacCow-UCD marker was only tested against 73 samples (Shanks et al., 2008; Kildare et al., 2007). Additionally, the CowM2 marker was tested against 17 different types of animal feces while the BacCow-UCD was only tested against 7 different species of animals. These differences in initial evaluations of MST markers led to uncertainty in the accuracy of reported sensitivities and specificities. The need for more robust review and testing of new markers across different species, laboratories, and locals were called for. In 2013, Boehm et al. performed a round robin study with 27 laboratories to evaluate and validate 41 different MST markers. The authors used nine different species of animal feces, individual human feces, septage, and wastewater fecal sources (fecal samples were collected from 12 individual animals of each species, nine sewage treatment plants, and six septage collection trucks) to provide single and mixed blinded challenge samples to all participating laboratories to test. Of the 41 different MST markers tested, only 15 (2 human, 2 ruminant, 2 bovine, 1 canine, 2 gull, 2 porcine, 1 horse, 1 deer, and 2 multitarget techniques) were found to be sufficiently sensitive and specific (> 80% sensitivity and specificity) when evaluating by

presence/absence by the majority of the laboratories involved. When the markers were evaluated quantitatively six markers were identified with higher concentrations in their target host feces. These included two human markers (HF183Taqman and BacH), two ruminant (Rum2bac and BacR), one gull (LeeSeaGull), and 1 porcine (Pig2Bac). This study was limited in two ways. First in that the fecal samples and wastewater used for the study were all sourced from a relatively small geographic region (California). Secondly, some of the markers were only tested by a single laboratory during the study. This second limitation was recognized by the authors and further evaluation of these markers was recommended.

Even after markers are evaluated, laboratories need to consider where the marker has been developed and tested to examine broad use geographically (e.g., in the Americas, Europe, tropics, and sub-tropics). Thus, successful microbial source tracking has relied not only on highly sensitive and specific MST markers, but on the selection of the most appropriate marker for the goal of the study and the area in which the study is being conducted. The evaluation of more than a handful of MST markers at a time is rare due to the amount of time, resources, and number of samples required for appropriate comparisons. At the time of this review there have been over 100 different MST markers developed for 15 different targets including human, ruminant, bovine, deer, canine, equine, avian, waterfowl, gull, goose, chicken, porcine, sheep, possum, and universal fecal markers. Human markers are the most prevalent (>35 markers), followed by livestock (>30 markers) (cows, pigs, chickens, ruminant, sheep, horses), then non-chicken avian markers (>20 markers) (general avian, waterfowl, ducks, geese), and finally wildlife markers having the fewest available markers (deer, possum, etc.). This distribution of available MST markers is unsurprising as human fecal contamination is more likely to be associated with increased health risks compared to wildlife feces. While there are over 100 published MST

markers not all are currently and/or widely used. Additionally, some markers have been modified and/or updated since their original publications. For example, the HF183 human marker was originally published as an endpoint PCR reaction MST marker by Scott et al. in 2002 then later updated to a SYBR green qPCR assay by Seurinck et al. in 2005. Then in 2010 Haugland et al. developed a Taqman qPCR version of the HF183 marker which was again modified in 2014 by Green et al. (2014).

1.2.2.1 Application of MST markers

The applications of MST include characterization of pollution sources in watersheds, the quantification of different sources, confirmation of suspected pollution sources, evaluation of large-scale areas for fecal pollution, connecting fecal pollution with pathogens, contamination source identification during outbreaks, and identification of failing sewer infrastructures and illicit connections in urban areas (see the discussion below). The application of MST has increased since its inception and it continues to be a useful tool for water quality monitoring.

A few key papers have been selected and reviewed here which represent these different applications of microbial source tracking. The first two papers that have been selected to represent the characterization of pollution sources in watersheds were published in 2006 and 2007 respectively. Shanks et al. (2006) used five MST markers (two human, two ruminant, and one elk) to analyze fecal contamination from 30 sites totaling nearly 3,000 samples ($n=2,912$) in an Oregon watershed (150,000 ha) over a two year period. This study found that within this watershed the fecal pollution was more closely linked to ruminant sources than human across the whole basin with ruminant markers being detected 75% of the time and rising to a 90% detection rate during spring and fall when precipitation increased. While human markers were found less

frequently than the ruminant markers they were able to be used to identify “hotspots” along waterways where normal sampling had only shown low concentrations due to dilution effects. These results highlight the benefits of characterizing a watershed as a whole, allowing for identification of sources which would otherwise be masked by previous sampling and testing strategies.

Graves et al. (2007) performed a similar study at a smaller scale (n=60 samples) over the course of a year in a smaller Virginia watershed (3,767 ha) which was also dominated by cattle. Graves et al. used a library-dependent MST approach which included fecal samples from humans, cattle, horses, waterfowl, geese (domestic and wild), wood ducks, deer, muskrats, and racoons. While this study used less-specific markers than the library-independent markers used by Shanks et al. (2006), their use of a library dependent method allowed for a wider range of sources to be tested against and identified. Similar to Shanks et al. (2006), Graves et al. (2007) also found that cattle were the most abundant source of fecal pollution in the watershed with 60% of samples found to be positive for cattle feces. With the help of their MST markers. this study also found an unexpectedly high human pollution area but were unable to identify the specific source. This application of MST at the watershed scale provided useful information on the happenings within the watershed but was limited in geographic scope.

Microbial source tracking has been applied at larger geographic scopes as demonstrated by Verhougstraete et al. (2015) and Nshimyimana et al. (2018). These two studies though published three years apart worked off of the same set of samples collected between 2010-2011 from 63 watersheds in Michigan. Samples were collected at single outflow points for each watershed during the fall baseflow (n=63), spring snowmelt (n=63), and a summer rain event (n=63, total n= 189). While Verhougstraete et al. (2015) focused on human sources of pollution

for their analysis, Nshimyimana et al. (2018) focused on animal fecal sources (bovine and porcine) effecting each watershed. These two studies showed the viability of broad-scale approach to the application of MST markers. By taking a step back and looking at single outflow points from each watershed Verhougstraete et al. (2015) was able to correlate human fecal pollution with septic tank numbers across these watersheds. Thus, demonstrating that on-site wastewater treatment not wastewater treatment plants were greater contributors to human fecal pollution across the state of Michigan. Nshimyimana et al. (2018) on the other hand were able to identify relationships between porcine and bovine markers and nutrients (nitrogen and phosphorus) only during baseflow conditions (versus their two other sampling events spring snow melt and summer rain) which suggested that nutrients and the animal markers had different mechanisms of transport during periods of increased streamflow.

The ability to use surrogates to determine the potential risk to human health from pathogens in water is desirable due to the high cost of individual pathogen testing and the potential for multiple pathogens to be present. The use of MST to help identify areas of increased risk and potential pathogen presence is shown in Bradshaw et al. (2016) and Korajkic et al. (2018). Bradshaw et al. (2016) chose to examine the relationship between pathogens and indicators (including MST markers; human, bovine and ruminant) in a mixed-use watershed. While they hypothesized that sediment would be an important source of pathogens and indicators, due to the possibility of resuspension, their study showed that while indicators were found in both sediments and the water column, pathogens were more likely to be found in the water column more often than in the sediments. They also found that there was no consistent relationship between indicators and pathogens, but that using a combination of FIB and MST markers helped to improve the ability to predict pathogen presence/absence. Korajkic et al.

(2018) performed a review of the currently available literature (73 papers) which had attempted to connect microbial indicators (e.g., FIB, MST markers) and pathogens in recreational waters. This review highlighted that while most connections between indicators and pathogens were tenuous at best, under certain conditions relationships could be discerned (after wet weather events or at sites where recent fecal pollution had occurred). Relationships were also more often reported in freshwater environments compared to marine and between bacterial indicators and bacterial or protozoa pathogens.

While direct relationships between pathogens and MST markers are difficult, the use of MST during known outbreaks provides valuable information on the source and quantity of fecal pollution which may be significant factors in the spread of the disease and the health-risk posed. A recent example of this was published by Mattioli et al. in 2021. Mattioli et al. (2021) were able to provide assistance using MST during the 2017 norovirus outbreaks in Pennsylvania when epidemiological investigations were unable to identify a specific source and exposure route for the outbreaks which resulted in 179 illnesses. Using a human MST marker (HF183), Mattioli et al. (2021) were able to demonstrate that a malfunctioning septic system was hydrologically connected to the drinking water well and recreational waters where the outbreaks occurred. This study highlighted the ability of MST to help with outbreak scenarios where traditional epidemiological studies were unable to effectively determine the exposure routes and main source(s) of contamination.

Lastly MST can be applied to human-made water systems as well as environmental areas. This has recently been highlighted best by the work of Gonzalez et al.'s (2020) collection system investigation microbial source tracking (CSI-MST) and Hachad et al.'s (2022) identification of illicit discharges using MST. Gonzalez et al. (2020) CSI-MST used a human MST (HF183)

marker along with extensive sewer collection system information (e.g., sewer line locations, service areas, and storm water systems locations) to survey multiple points within the storm water systems and track down any potential leaks from the sewer systems. The study presented three case studies where MST marker concentrations were used to identify leaks from the sewer systems into the storm water lines which allowed for the local municipality to remediate these failing infrastructure points with minimal disruption to the surrounding areas. This study has acted as proof of concept that with sufficient knowledge of the infrastructure in an urban area contamination from leaking sewer lines can be accurately identified and remediated. While Gonzalez et al (2020) used MST to identify and remediate failing infrastructure, Hachad et al. (2022) used MST to identify and remediate illicit wastewater connections to stormwater systems. MST was used along with other markers of wastewater pollution with a toolbox approach to identify cross connections. By using an index approach with multiple indicators of contamination eight misconnected houses were able to be identified and their connections were corrected. In complex urban settings the use of multiple indicators allowed for the reduction in false positive identification of cross connections.

The application of MST in watersheds, across watersheds, with pathogens, and within sewersheds show how useful these tools are for assessing and correcting water pollution in natural and anthropomorphic settings. Knowledge of the currently available tools available allows for the best results to be obtained for the chosen study.

1.2.3 Surveys for Sources of Fecal Pollution and Their Impact on Water Quality

The 1972 amendment of the Federal Water Pollution Control Act of 1948, also known as the Clean Water Act (CWA), marked a turning point in water quality in the United States. The

ability to regulate and implement pollution control measures resulted in an improvement of water quality across the country. In particular, section 303(d) of the CWA allowed for the development of total maximum daily loads (TMDLs) for waterbodies. These TMDLs set the maximum amount of specific pollutants that were allowed to enter waterbodies. By defining TMDLs for waterbodies, government agencies were able to start regulating the source of pollution impacting water quality. Point sources of pollution, such as wastewater treatment plant outfalls, sanitary sewer outfalls (SSOs), and combined sewer overflows (CSOs), were identified and regulated through the National Pollutant Discharge Elimination System (NPDES) resulting in an overall decrease in the impact of wastewater discharges on water quality. While point sources still contribute to the pollution of water bodies, since 2009 non-point sources of pollution have become the leading contributors to impaired waters (USEPA, 2009). In recent years, point source pollution discharges tend to be cross contamination events between sewage and storm water outfalls or urban runoff into storm water sewers (Kapoor et al., 2014; Staley et al., 2016). For non-point sources of pollution, land use decisions have significant impacts on the sources and transport of microbial pathogens into and through environments (Dreelin et al., 2007). This has shown the importance of land use and land cover (LULC) management. The alteration of natural environments for agriculture in particular can alter the natural percolation and runoff patterns of a watershed. This is problematic in the Great Lakes region where contaminated runoff can lead to nutrient accumulation in waterways and sediments where legacy pollutants can remain (Kinzelman et al., 2004; Smart and Barko 1978; Mortimer 1971; Marvin-DiPasquale and Agee 2003; Weller et al., 2020). In one study, it was found that between 33 and 58% of all nitrogen and phosphorus pollution in the study area was contributed from agricultural land uses (Robertson and Saad, 2011). In recent years, studies have been conducted to determine the

sources and pathways of nutrients and microbes from non-point sources of pollution in watersheds at the regional scale (Verhougstraete et al., 2015; Luszcz et al., 2017; Nshimiyimana et al., 2018).

As mentioned in section 1.2.1 above, water quality standards were determined around recreational exposure to contaminated waters. Throughout the Great Lakes a significant amount of recreational water exposure occurs at beaches. Michigan alone has 1,232 public and 575 private beaches which are required to be monitored regularly to make sure pollution level are below recreational standards. Elevated pollution levels at recreational water access points (e.g., beaches) are a risk to public health and require intervention (either remediation or closures) to mitigate that risk. In the last 10 years there have been 2,268 beach closures in Michigan with a combined total of 14,299 days of beach closures (MI EGLE, 2022).

Monitoring water quality is further complicated with the ability of pollutants to be retained and accumulate in sediments and beach sand. This retention and accumulation of pollutants has been observed in fresh and marine waters where *E. coli* levels were highest in sand and sediments compared to the surrounding waters (Alm et al., 2003; Whitman and Nevers, 2003; Zehms et al., 2008; Cloutier and McLellan 2017). These legacy pollution sources complicate the ability of traditional FIB to determine the current levels of bacterial pollution and risk as well as how to remediate the water quality problems.

1.2.4 Wastewater Surveillance and Human Health

1.2.4.1 COVID-19 (Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2))

In December 2019, a novel coronavirus was identified as the cause of a pneumonia outbreak in the Wuhan Province of China (WHO, 2020a; Zhu et al., 2020; Coronaviridae Study

Group of the International Committee on Taxonomy of Viruses, 2020). Through genetic sequencing this novel coronavirus was determined to be highly similar to severe acute respiratory syndrome coronavirus (SARS-CoV) and was thus named SARS-CoV-2 (Zhu et al., 2020). Coronaviruses are spherical enveloped positive-strand RNA viruses with diameters of approximately 120 nm and genomes ranging in size between 27 and 32 kb. Their lipid envelopes are embedded with spiked glycoproteins.

The disease resulting from infection by SARS-CoV-2 was named coronavirus disease 2019 (COVID-19) and includes symptoms of fever, fatigue, myalgia, dry cough, dyspnea and 2-10% of patients exhibit gastrointestinal symptoms (e.g., diarrhea) (Gu et al., 2020; Chen et al., 2020, Guan et al., 2020; Huang et al., 2020; Wang et al., 2020a; Wang et al., 2020b). While this virus primarily infects the respiratory system, it is also known to infect glandular epithelial cells in the intestinal tract and is subsequently shed in feces of both symptomatic and asymptomatic individuals (Xiao et al., 2020; Gu et al., 2020; Holshue et al., 2020; Song et al., 2020; Park et al., 2020; Wu et al., 2020a). The World Health Organization declared a global pandemic of COVID-19 on March 11, 2020 (WHO, 2020b). In the subsequent two years the pandemic has continued spreading to nearly all countries worldwide resulting in >440,000,000 confirmed cases and nearly 6,000,000 deaths (WHO, 2022). The tracking and surveillance of COVID-19 has become a priority for ensuring public health.

During the start of the pandemic surveillance of SARS-CoV-2 was limited to testing of clinical samples taken from symptomatic individuals. Several genetic markers targeting viral nucleocapsid and envelope genes were developed (CDC, 2020; Corman et al., 2020) However, due to the presence of asymptomatic cases of COVID-19 clinical level surveillance measures were inadequate. The shedding of SARS-CoV-2 in the feces of infected individuals presented an

opportunity to non-intrusively track the spread of COVID-19 at the community level (Kitijima et al., 2020). Early studies testing for the SARS-CoV-2 genes in wastewater showed the ability to consistently detect the genetic signal at concentrations of 10^2 to 10^6 gene copies per liter of wastewater and were able to predict the trends in COVID-19 cases over time (Medema et al., 2020; Ahmed et al., 2020).

Since the beginning of the COVID-19 pandemic there have been several SARS-COV-2 variants of concern that have emerged. The most significant of these variants have been the Alpha variant, the Delta variant, and the Omicron variant (WHO, 2021). These variants have been seen in waves of cases with the Alpha variant first being reported during the fall of 2020 in the United Kingdom, then the Delta variant in India in May 2021, and lastly the Omicron variant in multiple countries during November of 2021 (WHO, 2022). These variants have been monitored clinically as well as by wastewater surveillance. In Michigan the first confirmed case of the Alpha variant was in January 2021, with the Delta variant being detected in June of 2021, and the Omicron variant being detected on December 3rd 2021 (Michigan.gov/Coronavirus, 2022).

Early in the COVID-19 pandemic researchers began to test for SARS-CoV-2 in wastewater. One of the first studies was conducted in the Netherlands where wastewater surveillance was conducted with sewage samples from six cities and an airport (Medema et al., 2020). The authors were able to detect the SARS-CoV-2 genetic signal in wastewater up to six days prior to the first reported cases in the Netherlands. These results showed that wastewater surveillance is able to detect the levels of SARS-CoV-2 in a community and provide an early warning of increasing cases prior to clinical detection methods. Other studies early in the

pandemic were able monitor the spread of the pandemic in their countries with similar results to Medema et al. (Wu et al., 2020b; Ahmed et al., 2020; Lodder et al., 2020).

As the pandemic continued research shifted from proof of concept (e.g., can SARS-CoV-2 be detected in wastewater) to surveillance of the virus as a method of non-intrusive monitoring of community health. These studies ranged from monitoring a single wastewater treatment plant wastewater monitoring (Haramoto et al., 2020) to more widespread surveillance of multiple sites across large geographic regions (e.g., Hata et al., 2021). Hata et al. (2021) were able to monitor their wastewater sites prior to and during the outbreak of SARS-CoV-2 in two Japanese prefectures. They monitored five wastewater treatment plants in two prefectures collecting a total of 45 wastewater samples between March 5th and May 29th 2020. The authors were able to detect SARS-CoV-2 in the wastewater even when cases were less than one in 100,000 persons which was in line with what Medema et al. (2020) saw in the Netherlands. This study also determined that wastewater surveillance was able to detect cases of COVID-19 in communities prior to an increase in accompanying clinical case data. This suggested that in the initial stages of an outbreak or pandemic wastewater surveillance is able to more accurately determine the rate of disease spread in a community. This is particularly important because as cases decline during the pandemic wastewater monitoring may be able to provide an early warning for a resurgence of cases in different communities.

Wastewater surveillance of SARS-CoV-2 has also been used at the building level to help to contain the spread of the disease in close populations. This was performed by Betancourt et al. (2021) for the student dormitories on the campus of the University of Arizona. Betancourt et al. monitored wastewater from 13 dormitories between August 24th and November 20th 2020. When a positive RNA signal for SARS-CoV-2 was detected clinical testing of the individuals living in

the dormitory was performed followed by isolation of infected individuals. The authors were able to observe an 82.0% positive sample predictive value from their monitoring. While this monitoring and intervention plan is unfeasible at larger scales it does provide a methodology to help contain potential outbreaks of COVID-19 in defined populations.

The proliferation of wastewater surveillance of SARS-CoV-2 and its accompanying data has provided the ability to evaluate the efficacy of wastewater-based epidemiology (WBE) on larger scales which have been limited. For example, Morvan et al. (2021) used wastewater surveillance data from multiple studies across 44 sites in England to estimate the prevalence of SARS-CoV-2 to help alleviate some of the shortcoming of isolated clinical case monitoring. Morvan et al. showed that the wastewater results were within 1.1% of prevalence estimates based on case data and preceded clinical testing data by 4-5 days suggesting that wastewater monitoring is a leading indicator of asymptomatic COVID-19 infections. These results show the value of WBE as an additional form of disease surveillance and subsequently an useful tool to preserve public health.

While tragic, the COVID-19 pandemic has served to highlight the usefulness of WBE and the ability for community health to be monitored and observed unobtrusively. Underreporting and asymptomatic cases of disease complicate clinical disease surveillance. The addition of WBE can allow for a more robust surveillance of public health and potentially provide early warnings of disease outbreaks.

1.3 Scientific Needs

There are several areas in water quality monitoring and pathogen surveillance that are in need of further study. Currently, water pollution is more greatly influenced by non-point sources

compared to point sources thus it is essential that source identification be implemented to make progress on remediation of water systems. The identification and testing of locally appropriate genetic markers and understanding the relationships between microbial markers and environmental variables within watersheds are essential. While current strategies using fecal indicators is sufficient to meet many regulatory standards, they are insufficient in addressing water quality associated public health risks (Evans et al., 2019). Pathogen monitoring is now being undertaken to support community strategies for disease control and this is particularly true in response to the COVID-19 pandemic. The development and implementation of methods and monitoring strategies are necessary to promote WBE as an effective means for furtherance of community health (Kitajima et al., 2020). Finally understanding how community structure and diversity effects the efficacy of WBE is necessary to determine the best strategies for future community level pathogen monitoring.

1.4 Research Objectives

Specialized surveys using MST technology for Michigan watersheds were of great interest to examine the major contributors influencing degraded water quality. The use of MST markers and nutrient analysis in various watersheds and their subsections could be further examined over seasonal/temporal scales to provide a better understanding of what was happening within the variety of Michigan watersheds. This was in contrast to previous studies conducted which only looked at a single outflow point from each of the watersheds (Verhougstraete et al., 2015).

In addition, as the COVID pandemic raged on the use of surveillance monitoring approaches around the world exploded. It was hypothesized that SARS-CoV-2 would be able to

be detected from wastewater using molecular methods, specifically ddPCR with greater accuracy and precision than qPCR. Additionally, since enveloped viruses tend to attach to particulate matter, a relatively cheap and easy to use wastewater concentration method like polyethylene glycol (PEG) precipitation would likely be effective at concentrating the SARS-CoV-2 virus to detectable levels with good recovery efficiencies. Unique methodologies for the recovery of a respiratory virus from wastewater needed to be developed and compared with other proposed methods. Also, it was suggested that the detection of SARS-CoV-2 RNA from wastewater would be able to help predict increases in cases across different communities in Michigan.

Four chapters follow this introduction. The first three detail the main studies conducted for this dissertation, and the final chapter provides a summary and a look at future work. The first of these studies focused on microbial source tracking markers and their connections to nutrients in mixed-use watersheds. This chapter has been published in the journal *Water Research* in 2022. The second study focused on the development and evaluation of a method for the concentration and detection of SARS-CoV-2 from wastewater. This chapter has also been published in the *Journal of Food and Environmental Virology* in 2021. Citations for these studies is provided on the cover page of each chapter. The final chapter addressed the surveillance of SARS-CoV-2 in wastewater in two disparate communities. This will be submitted for publication in the near future.

Goal 1

The increased relative input of non-point sources of pollution in watersheds has increased the importance of source identification and the determination of co-occurring contaminants. A survey of five mixed-use watersheds in Michigan was undertaken to investigate the sources of

pollution in the various subsections of each watershed and quantitatively and qualitatively determine which variables correlate.

The objectives of this study were to: 1) determine the spatial and temporal trends in microbial contamination and nutrients in five mixed-use watersheds; 2) to determine if MST markers could be significantly correlated with nutrient levels (e.g., phosphorus and nitrogen); 3) to determine which watershed variables (e.g., nutrient levels, land use, management practices, etc.) within a watershed would be useful in predicting microbial contamination levels.

Goal 2

The onset of the COVID-19 global pandemic left countries and communities scrambling to not only respond to the global health emergency but also to find cost-effective ways to surveil the spread of SARS-CoV-2. The development of a relatively simple workflow for the detection and quantification of SARS-COV-2 in wastewater was undertaken to meet this need. Wastewater from communities around Michigan were collected and analyzed along with viral surrogates for SARS-CoV-2 to investigate different workflow options.

The objectives of this study were to 1) evaluate the efficiencies of polyethylene glycol (PEG) precipitation and ultrafiltration methods to recover *Pseudomonas* phage Phi6, coronavirus OC43, and SARS-CoV-2 from different wastewater matrices; 2) compare two PCR-based methods, reverse transcription quantitative PCR (RT-qPCR) and reverse transcription droplet digital PCR (RT-ddPCR) for the detection of SARS-CoV-2 in different wastewater matrices; 3) develop a rapid, cost-effective, and precise quantification workflow for SARS-CoV-2 in wastewater.

Goal 3

Wastewater-based epidemiology may provide a better method for the surveillance of pathogens in communities of varying sizes. Due to the costly nature of clinical testing and the presence of asymptomatic carries of SARS-CoV-2, monitoring communities as a whole may be able to more accurately determine the incidence of disease and provide a warning when the disease is spreading, prior to an increase in clinical cases.

This study had three main objectives: 1) to evaluate the efficacy of wastewater monitoring of SARS-CoV-2 in two communities with diverse characteristics; 2) to determine if county or zipcode level case data are necessary to successfully correlate with wastewater surveillance results; 2) to compare the spatial resolution of cases (county vs zipcode) and the relationship to SARS-CoV-2 wastewater surveillance data 3) to determine the impact of vaccination rates on SARS-CoV-2 wastewater signals compared to case numbers.

REFERENCES

REFERENCES

- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tschärke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K.V., and Mueller, J.F. (2020). First Confirmed Detection of SARS-CoV2 in Untreated Wastewater in Australia: A Proof of Concept for the Wastewater Surveillance of COVID-19 in the Community. *Sci. Total Environ.* 728, 138764.
- Alm, E.W., Burke, J., and Spain, A. (2003). Fecal Indicator Bacteria are Abundant in Wet Sand at Freshwater Beaches. *Water Research*, 37, 3978-3982.
- Ballesté, E., Demeter, K., Masterson, B., Timoneda, N., Sala-Comorera, L., and Meijer, W.G. (2020). Implementation and Integration of Microbial Source Tracking in a River Watershed Monitoring Plan. *Science of the Total Environment*. 736; 139573.
<https://doi.org/10.1016/j.scitotenv.2020.139573>
- Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Brown, K.M.P., Stark, E.R., Foster, A.R., Sprissler, R.S., Harris, D.T., Sherchan, S.P., Gerba, C.P., and Pepper, I.L. (2021). COVID-19 Containment on a College Campus via Wastewater-Based Epidemiology, Targeted Clinical Testing and an Intervention. *Sci Total Environ* 779, 146408.
<https://doi.org/10.1016/j.scitotenv.2021.146408>
- Bischoff, C., Lüthy, J., Altwegg, M., and Baggi, F. (2005). Rapid Detection of Diarrheagenic *E. coli* by Real-Time PCR. *J Microbiol Meth* 61, 335–341.
<https://doi.org/10.1016/j.mimet.2004.12.007>
- Boehm, A.B., Werfhorst, L.C., Van De, Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Dan, W., and Weisberg, S.B. (2013). Performance of Forty-One Microbial Source Tracking Methods: A Twenty-Seven Lab Evaluation Study. *Water Res.* 47 (18), 6812e6828.
- Bonde, G. (1966). Bacteriological Methods for Estimation of Water Pollution. *Healthy Laboratory Science*, 3, 124-128.
- Bradshaw, J.K., Snyder, B.J., Oladeinde, A., Spidle, D., Berrang, M.E., Meinersmann, R.J., Oakley, B., Sidle, R.C., Sullivan, K., and Molina, M. (2016). Characterizing Relationships among Fecal Indicator Bacteria, Microbial Source Tracking Markers, and Associated Waterborne Pathogen Occurrence in Stream Water and Sediments in a Mixed Land Use Watershed. *Water Res* 101, 498–509. <https://doi.org/10.1016/j.watres.2016.05.014>
- Cabelli, V.J., Dufour, A.P., McCabe, L.J., and Levin, M.A. (1982). Swimming-Associated Gastroenteritis and Water Quality. *American Journal of Epidemiology*, 115, 606-616.
- Carlson, R. (2003). The Pace and Proliferation of Biological Technologies. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science* 1(3), 203-214.

Centers for Disease Control and Prevention (CDC) (2012). Principles of Epidemiology | Lesson 1 - Section 4. Centers for Disease Control and Prevention: Principles of Epidemiology in Public Health Practice, Third Edition An Introduction to Applied Epidemiology and Biostatistics. https://www.cdc.gov/csels/dsepd/ss1978/Lesson1/Section4.html#_ref18

Cloutier, D.D. and McLellan, S.L. (2017). Distribution and Differential Survival of Traditional and Alternative Indicators of Fecal Pollution at Freshwater Beaches. *Appl Environ Microb* 83, e02881-16. <https://doi.org/10.1128/aem.02881-16>

Colford, J.M., Wade, T.J., Schiff, K.C., Wright, C.C., Griffith, J.F., Sandhu, S.K., Burns, S., Sobsey, M., Lovelace, G., and Weisberg, S.B. (2007). Water Quality Indicators and the Risk of Illness at Beaches with Nonpoint Sources of Fecal Contamination. *Epidemiology*, 18, 27-35.

Dreelin, E., McNinch, R., and Rose, J. (2007). Surface Water Summary of *E. coli* and Pathogens in Michigan. Report prepared for the Michigan Department of Environmental Quality Water Bureau

Dubrovsky, N.M., Burow, K.R., Clark, G.M., Gronberg, J.M., Hamilton, P.A., Hitt, K.J., Mueller, D.K., Munn, M.D., Nolan, B.T., Puckett, L.J., Rupert, M.G., Short, T.M., Spahr, N.E., Sprague, L.A., and Wilber, W.G. (2010). The Quality of our Nation's Waters: Nutrients in the Nation's Streams and Groundwater, 1992–2004. USGS Survey Circ. 1350. USGS, Washington, DC.

Dufour, A.P. (1984). Health Effects Criteria for Fresh Recreational Waters. EPA-600/1-84-004, Office of Research and Development, US Environmental Protection Agency, Cincinnati, OH.

Girones, R., Ferrús, M.A., Alonso, J.L., Rodriguez-Manzano, J., Calgua, B., Corrêa, A., Hundesa, A., Carratala, A., and Bofill-Mas, S. (2010). Molecular Detection of Pathogens in Water--the Pros and Cons of Molecular Techniques. *Water Research*, 44, 4325-4339.

Gleeson, C. and Gray, N. (2003). The Coliform Index and Waterborne Disease: Problems of Microbial Drinking Water aAssessment. New York, NY: Taylor and Francis e-Library

Godfree, A.F., Kay, D., and Wyer, M.D. (1997). Faecal Streptococci as Indicators of Faecal Contamination in Water. *J Appl Microbiol* 83, 110–119. <https://doi.org/10.1046/j.1365-2672.83.s1.12.x>

Gonzalez, D., Keeling, D., Thompson, H., Larson, A., Denby, J., Curtis, K., Yetka, K., Rondini, M., Yeargan, E., Egerton, T., Barker, D., and Gonzalez, R. (2020). Collection System Investigation Microbial Source Tracking (CSI-MST): Applying Molecular Markers to Identify Sewer Infrastructure Failures. *J Microbiol Meth* 178, 106068. <https://doi.org/10.1016/j.mimet.2020.106068>

Grabow, W.O.K. (1996) Waterborne Diseases: Update on Water Quality Assessment and Control. *Water SA*, 22, 193-202.

Graves, A.K., Hagedorn, C., Brooks, A., Hagedorn, R.L., and Martin, E. (2007). Microbial Source Tracking in a Rural Watershed Dominated by Cattle. *Water Res* 41, 3729–3739. <https://doi.org/10.1016/j.watres.2007.04.020>

Green, H.C., Haugland, R.A., Varma, M., Millen, H.T., Borchardt, M.A., Field, K.G., Walters, W.A., Knight, R., Sivaganesan, M., Kelty, C.A., and Shanks, O.C. (2014). Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Appl Environ Microb* 80, 3086–94. <https://doi.org/10.1128/aem.04137-13>

Griffin, D.W., Lipp, E.K., Mclaughlin, M.R., and Rose, J.B. (2001). Marine Recreation and Public Health Microbiology: Quest for the Ideal Indicator. *Bio Science*, 51, 817-826.

Griffith, J.F., Weisberg, S.B., and McGee, C.D. (2003). Evaluation of Microbial Source Tracking Methods using Mixed Fecal Sources in Aqueous Test Samples. *Journal of Water and Health*, 1, 141-151.

Hachad, M., Lanoue, M., Duy, S.V., Villemur, R., Sauvé, S., Prévost, M., and Dorner, S. (2022). Locating Illicit Discharges in Storm Sewers in Urban Areas using Multi-parameter Source Tracking: Field Validation of a Toolbox Composite Index to Prioritize High Risk Areas. *Sci Total Environ* 811, 152060. <https://doi.org/10.1016/j.scitotenv.2021.152060>

Haramoto, E., Malla, B., Thakali, O., and Kitajima, M. (2020). First Environmental Surveillance for the Presence of SARS-CoV-2 RNA in Wastewater and River Water in Japan. *Sci Total Environ* 737, 140405. <https://doi.org/10.1016/j.scitotenv.2020.140405>

Hata, A., Hara-Yamamura, H., Meuchi, Y., Imai, S., and Honda, R. (2021). Detection of SARS-CoV-2 in wastewater in Japan during a COVID-19 outbreak. *Sci Total Environ* 758, 143578. <https://doi.org/10.1016/j.scitotenv.2020.143578>

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., and Shanks, O.C. (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by qPCR. *Syst Appl Microbiol* 33, 348–357. <https://doi.org/10.1016/j.syapm.2010.06.001>

Heaney, C.D., Myers, K., Wing, S., Hall, D., Baron, D., and Stewart, J.R. (2015). Source tracking swine fecal waste in surface water proximal to swine concentrated animal feeding operations. *Sci. Total Environ.* 511, 676–683.

Heymann, D.L. and Rodier G.R. (2001). Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases. *Lancet Infect Dis*; 1:345-53; PMID:11871807; [http://dx.doi.org/10.1016/S1473-3099\(01\)00148-7](http://dx.doi.org/10.1016/S1473-3099(01)00148-7)

Horan, N.J. (2003). Handbook of Water and Wastewater Microbiology. Part 1 Basic Microbiol 105–112. <https://doi.org/10.1016/b978-012470100-7/50008-x>

- Kapoor, V., Pitkänen, T., Ryu, H., Elk, M., Wendell, D., and Domingo, J.W.S. (2014). Distribution of Human-Specific Bacteroidales and Fecal Indicator Bacteria in an Urban Watershed Impacted by Sewage Pollution, Determined Using RNA- and DNA-Based Quantitative PCR Assays. *Appl Environ Microb* 81, 91–99. <https://doi.org/10.1128/aem.02446-14>
- Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., and Wuertz, S. (2007). 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: A Bayesian approach. *Water Research*, 16, 3701-3715.
- Kinzelman, J., McLellan, S.L., Daniels, A.D., Cashin, S., Singh, A., Gradus, S., and Bagley, R. (2004). Non-point source pollution: determination of replication versus persistence of *Escherichia coli* in surface water and sediments with correlation of levels to readily measurable environmental parameters. *Journal of Water and Health*, 2, 103-114.
- Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C.P., Hamilton, K.A., Haramoto, E., and Rose, J.B. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of The Total Environment*, 739, 139076. doi:10.1016/j.scitotenv.2020.139076
- Korajkic, A., McMinn, B.R., Ashbolt, N.J., Sivaganesan, M., Harwood, V.J., and Shanks, O.C. (2018). Extended persistence of general and cattle-associated fecal indicators in marine and freshwater environment. *Sci Total Environ* 650, 1292–1302. <https://doi.org/10.1016/j.scitotenv.2018.09.108>
- Lodder, W., and de Roda Husman, A.M. (2020). SARS-CoV-2 in wastewater: potential health risk, but also data source. *Lancet Gastroenterol Hepatol*. 5, 533.
- Luszcz, E.C., Kendall, A.D., and Hyndman, D.W. (2017). A spatially explicit statistical model to quantify nutrient sources, pathways, and delivery at the regional scale. *Biogeochemistry*. 133: 37-57.
- Marvin-DiPasquale, M. and Agee, J.L. (2003). Microbial mercury cycling in sediments of the San Francisco Bay-Delta. *Estuaries*, 26, 1517-1528.
- Mattioli, M.C., Benedict, K.M., Murphy, J., Kahler, A., Kline, K.E., Longenberger, A., Mitchell, P.K., Watkins, S., Berger, P., Shanks, O.C., Barrett, C.E., Barclay, L., Hall, A.J., Hill, V., and Weltman, A. (2021). Identifying septic pollution exposure routes during a waterborne norovirus outbreak - A new application for human-associated microbial source tracking qPCR. *J Microbiol Meth* 180, 106091. <https://doi.org/10.1016/j.mimet.2020.106091>
- McLellan, S.L., Daniels, A.D., and Salmore, A.K. (2001). Clonal populations of thermotolerant enterobacteriaceae in recreational water and their potential interference with fecal *Escherichia coli* counts. *Applied and Environmental Microbiology*, 67, 4934-4938.
- McLellan, S. (2004). Genetic diversity of *E. coli* isolated from urban rivers and beach water. *Applied and Environmental Microbiology*, 70, 4658-4665.

- McLellan, S.L., Hollis, E.J., Depas, M.M., Dyke, M.V., Harris, J., and Scopel, C.O. (2007). Distribution and fate of *Escherichia coli* in Lake Michigan following contamination with urban stormwater and combined sewer overflows. *Journal of Great Lakes Research*, 33, 566-580.
- Michigan Department of Environment, Great Lakes, and Energy (MI EGLE) (2022). Michigan Beachguard System. <https://www.eagle.state.mi.us/beach/Default.aspx>. Accessed March 16, 2022.
- Morley, A.A. (2014). Digital PCR: A brief history. *Biomolecular Detection and Quantification*, 1(1), 1–2. doi:10.1016/j.bdq.2014.06.001
- Mortimer, C.H. (1971). Chemical exchanges between sediments and water in the Great Lakes speculations on probable regulatory mechanisms. *Limnology and Oceanography*, 2, 387-404
- Nnane, D.E., Ebdon, J.E., and Taylor, H.D. (2011). Integrated analysis of water quality parameters for cost-effective faecal pollution management in river catchments. *Water Res.* 45,2235–2246.
- Nshimiyimana J.P., Martin, S.L., Flood, M., Verhougstraete, M.P., Hyndman, D.W., and Rose, J.B. (2018). Regional variations of bovine and porcine fecal pollution as a function of landscape, nutrient, and hydrological factors. *Journal of Environment Quality*. 47:1024-1032.
- Orenstein W.A. and Bernier R.H. (1990). Surveillance: information for action. *Pediatr Clin North Am.* 37:709–34.
- Orive, G., Lertxundi, U., and Barcelo, D. (2020). Early SARS-CoV-2 outbreak detection by sewage-based epidemiology. *Sci Total Environ* 732, 139298. <https://doi.org/10.1016/j.scitotenv.2020.139298>
- Park, S., Lee, C.-W., Park, D.-I., Woo, H.-Y., Cheong, H.S., Shin, H.C., Ahn, K., Kwon, M.-J., and Joo, E.-J. (2020). Detection of SARS-CoV-2 in Fecal Samples from Patients with Asymptomatic and Mild COVID-19 in Korea. *Clin Gastroenterol H.* <https://doi.org/10.1016/j.cgh.2020.06.005>
- Pascual-Benito, M., Nadal-Sala, D., Tobella, M., Ballesté, E., García-Aljaro, C., Sabaté, S., Sabater, F., Martí, E., Gracia, C.A., Blanch, A.R., and Lucena, F. (2020). Modelling the seasonal impacts of a wastewater treatment plant on water quality in a Mediterranean stream using microbial indicators. *J. Environ. Manag.* 261, 110220.
- Prüss, A. (1998). Review of epidemiological studies on health effects from exposure to recreational water. *International Journal of Epidemiology*, 27, 1-9.
- Rice, E.W. (2003). *Escherichia coli*: Pathogenic strains in Encyclopedia of Environmental Microbiology. John Wiley and Sons, NY.

- Robertson, D.M. and Saad, D.A. (2011). Nutrient Inputs to the Laurentian Great Lakes by Source and Watershed Estimated Using SPARROW Watershed Models. *Jawra J Am Water Resour Assoc* 47, 1011–1033. <https://doi.org/10.1111/j.1752-1688.2011.00574.x>
- Rompré, A., Servais, P., Baudart, J., De-Roubin, M.R., and Laurent, P. (2002). Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. *Journal of Microbiological Methods*, 49, 31-54.
- Roslev, P. and Bukh, A.S. (2011). State of the art molecular markers for fecal pollution source tracking in water. *Appl. Microbiol. Biotechnol.* 89; pp. 1341-1355
- Santo Domingo, J.W., Bambic, D.G., Edge, T., and Wuertz, S. (2007). Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research*, 41, 3539-3552.
- Schwab, K.J. (2007). Are existing bacterial indicators adequate for determining recreational water illness in waters impacted by nonpoint pollution? *Epidemiology*, 18, 21-22.
- Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., and Lukasik, J. (2002). Microbial source tracking: Current methodology and future directions. *Applied and Environmental Microbiology*, 68, 5796-5803.
- Seurinck, S., Defoirdt, T., Verstraete, W., and Siciliano, S.D. (2005). Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environ Microbiol* 7, 249–259. <https://doi.org/10.1111/j.1462-2920.2004.00702.x>
- Shanks, O.C., Nietch, C., Simonich, M., Younger, M., Reynolds, D., and Field, K.G., (2006). Basin-Wide Analysis of the Dynamics of Fecal Contamination and Fecal Source Identification in Tillamook Bay, Oregon. *Appl Environ Microb* 72, 5537–5546. <https://doi.org/10.1128/aem.03059-05>
- Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M., and Haugland, R.A. (2008). Quantitative PCR for Detection and Enumeration of Genetic Markers of Bovine Fecal Pollution. *Appl Environ Microb* 74, 745–752. <https://doi.org/10.1128/aem.01843-07>
- Sinclair, R.G., Choi, C.Y., Riley, M.R., and Gerba, C.P. (2008). Pathogen surveillance through monitoring of sewer systems. *Adv. Appl. Microbiol.* 65, 249–269. [https://doi.org/10.1016/S0065-2164\(08\)00609-6](https://doi.org/10.1016/S0065-2164(08)00609-6).
- Smart, R. and Barko, J. (1978). Influence of sediment salinity and nutrients on the physiological ecology of selected salt marsh plants. *Estuarine and Coastal Marine Science*, 7, 487-495.

Sowah, R.A., Habteselassie, M.Y., Radcliffe, D.E., Bauske, E., and Risse, M. (2017). Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States. *Water Res.* 108, 330–338.

Staley, Z.R., Grabuski, J., Sverko, E., and Edge, T.A. (2016). Comparison of Microbial and Chemical Source Tracking Markers to Identify Fecal Contamination Sources in the Humber River (Toronto, Ontario, Canada) and Associated Storm Water Outfalls. *Appl Environ Microb* 82, 6357–6366. <https://doi.org/10.1128/aem.01675-16>

Steinbacher, S.D., Savio, D., Demeter, K., Karl, M., Kandler, W., Kirschner, A.K.T., Reischer, G.H., Ixenmaier, S.K., Mayer, R.E., Mach, R.L., Derx, J., Sommer, R., Linke, R., and Farnleitner, A.H., (2021). Genetic microbial faecal source tracking: rising technology to support future water quality testing and safety management. *Österreichische Wasser- Und Abfallwirtschaft* 73, 468–481. <https://doi.org/10.1007/s00506-021-00811-y>

Thupaki, P., Phanikumar, M.S., Beletsky, D., Schwab, D.J., Nevers, M.B., and Whitman, R.L. (2010). Budget analysis of *Escherichia coli* at a southern Lake Michigan beach. *Environmental Science and Technology*, 44, 1010-1016.

Toranzos, G.A. and McFeters, G.A. (1997). Detection of microorganisms in environmental freshwaters and drinking waters, p249-264. In Hurst, C.J., Knudsen, G., McInerney, M., Stetzenbach, L., and Walter, M. (Editors) *Manual of Environmental Microbiology*. Washing, D.C.: ASM Press.

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., and Lukasik, J. (2002). Microbial Source Tracking: Current Methodology and Future Directions. *Applied and Environmental Microbiology*. 68 (12) 5796-5803; **DOI:** 10.1128/AEM.68.12.5796-5803.2002

United States Environmental Protection Agency (USEPA), O. of W. (1986). Ambient water quality criteria for bacteria - 1986, (EPA440/5-8).

United States Environmental Protection Agency (USEPA) (2009). Review of published studies to characterize relative risks from different sources of fecal. (EPA 822-R-09-001).

United States Environmental Protection Agency (USEPA) (2012). Recreational Water Criteria. (820-F-12-058).

Verhougstraete, M.P., Martin, S.L., Kendall, A.D., Hyndman, D.W., and Rose, J.B. (2015). Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proc. Natl. Acad. Sci.* 112, 10419–10424.

Vital, M., Hammes, F., and Egli, T. (2008). *Escherichia coli* O157 can grow in natural freshwater at low carbon concentrations. *Environmental Microbiology*, 10, 2387-2396.

Vogelstein B. and Kinzler K.W. (1999). Digital PCR. *Proc Natl Acad Sci USA*; 96: 9236–41.

Wade, T.J., Pai, N., Eisenberg, J.N.S., and Colford J.M. (2003). Do US EPA water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and metaanalysis. *Environmental Health Perspectives*, 111, 1102-1109.

Wade, T.J., Calderon, R.L., Sams, E., Beach, M., Brenner, K.P., Williams, A.H., and Dufour, A.P. (2006). Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. *Environmental Health Perspectives*, 114, 24-28.

Wade, T., Calderon, R., Brenner, K., Sams, E., Beach, M., Haugland, R., Wymer, L., and DuFour, A. (2008). High sensitivity of children to swimming-associated gastrointestinal illness: Results using a rapid assay of recreational water quality. *Epidemiology*, 19, 375-383.

Walters, S.P. and Field, K.G. (2006). Persistence and growth of fecal Bacteroidales assessed by bromodeoxyuridine immunocapture. *Applied and Environmental Microbiology*, 72, 4532-4539.

Weller, D., Brassill, N., Rock, C., Ivanek, R., Mudrak, E., Roof, S., Ganda, E., and Wiedmann, M. (2020). Complex Interactions Between Weather, and Microbial and Physicochemical Water Quality Impact the Likelihood of Detecting Foodborne Pathogens in Agricultural Water. *Front Microbiol* 11, 134. <https://doi.org/10.3389/fmicb.2020.00134>

Whitman, R.L. and Nevers, M.B. (2003). Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. *Water*, 69, 5555-5562.

Whitman, R.L. and Nevers, M.B. (2004). Policy analysis *Escherichia coli* sampling reliability at a frequently closed Chicago beach: Monitoring and management implications. *Environmental Science and Technology*, 38, 4241-4246.

Winfield, M.D. and Groisman, E.A. (2003). Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. *Appl Environ Microb* 69, 3687–3694. <https://doi.org/10.1128/aem.69.7.3687-3694.2003>

World Health Organization (WHO) (2003). Guidelines for safe recreational water environments. Volume 1, Coastal and fresh waters. 253 pp.

World Health Organization (WHO) (1998). Water Quality Monitoring A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes. doi:10.4324/9780203476796.ch14

World Health Organization (WHO) (2020a). Pneumonia of unknown cause – China [WWW document]. URL. [https:// www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/](https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/).

World Health Organization (WHO) (2020b). Coronavirus disease (COVID-19) pandemic [WWW document]. URL. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.

World Health Organization (WHO) (2022). WHO Coronavirus (COVID-19) Dashboard.
<https://covid19.who.int/>

Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y., Qu, X., Kuang, L., Fang, X., Mishra, N., Lu, J., Shan, H., Jiang, G., and Huang, X. (2020a). Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *Lancet Gastroenterology Hepatology* 5, 434–435. [https://doi.org/10.1016/s2468-1253\(20\)30083-2](https://doi.org/10.1016/s2468-1253(20)30083-2)

Wu, F., Zhang, J., Xiao, A., Gu, X., Lee, W.L., Armas, F., Kauffman, K., Hanage, W., Matus, M., Ghaeli, N., Endo, N., Duvallet, C., Poyet, M., Moniz, K., Washburne, A.D., Erickson, T.B., Chai, P.R., Thompson, J., and Alm, E.J. (2020b). SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *Msystems* 5, e00614-20.
<https://doi.org/10.1128/msystems.00614-20>

Xagorarakis, I. and O'Brien, E. (2020). Wastewater-based epidemiology for early detection of viral outbreaks. In: O'Bannon, D. (Ed.), *Women in Water Quality*. Springer Nature Switzerland, pp. 75–97. <https://doi.org/10.1007/978-3-030-17819-2>.

Yang, Q., H. Tian, X. Li, W. Ren, B. Zhang, X. Zhang, and Wolf, J. (2016). Spatiotemporal patterns of livestock manure nutrient production in the conterminous United States from 1930 to 2012. *Sci. Total Environ.* 541:1592– 1602. doi:10.1016/j.scitotenv.2015.10.044

Zehms, T.T., Mcdermott, C.M., and Kleinheinz, G.T. (2008). Microbial concentrations in sand and their effect on beach water in Door County, Wisconsin. *Journal of Great Lakes Research*, 34, 524-534

Zmirou, D., Pena, L., Ledrans, M., and Letertre, A. (2003). Risks associated with the microbiological quality of bodies of fresh and marine water used for recreational purposes: Summary estimates based on published epidemiological studies. *Archives of Environmental Health*, 53, 703-711.

2.0 Connecting microbial, nutrient, physiochemical, and land use variables for the evaluation of water quality within five mixed use watersheds

Work presented in this chapter has been published as Flood, M.T.^a, Hernandez-Suarez, J.S.^b, Nejadhashemi, A.P.^b, Martin, S.L.^c, Hyndman, D.^d, and Rose, J.B.^a (2022). Connecting microbial, nutrient, physiochemical, and land use variables for the evaluation of water quality within mixed use watersheds. *Water Res* 219, 118526.

<https://doi.org/10.1016/j.watres.2022.118526>

^a Department of Fisheries and Wildlife, Michigan State University, East Lansing MI 48824, USA

^b Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing MI 48824, USA

^c Department of Earth and Environmental Sciences, Michigan State University, East Lansing MI 48824, USA

^d Department of Geosciences, School of Natural Sciences and Mathematics, University of Texas at Dallas, Richardson TX, 75080, USA

2.1 Abstract

As non-point sources of pollution begin to overtake point sources in watersheds, source identification and complicating variables such as rainfall are growing in importance. Microbial source tracking (MST) allows for identification of fecal contamination sources in watersheds; when combined with data on land use and co-occurring variables (e.g., nutrients, sediment runoff) MST can provide a basis for understanding how to effectively remediate water quality. To determine spatial and temporal trends in microbial contamination and correlations between MST and nutrients, water samples (n=136) were collected between April 2017 and May of 2018 during eight sampling events from 17 sites in 5 mixed-use watersheds. These samples were analyzed for three MST markers (human – *B. theta*; bovine – CowM2; porcine – Pig2Bac) along with *E. coli*, nutrients (nitrogen and phosphorus species), and physiochemical parameters. These water quality variables were then paired with data on land use, streamflow, precipitation and management practices (e.g., tile drainage, septic tank density, tillage practices) to determine if any significant relationships existed between the observed microbial contamination and these variables. The porcine marker was the only marker that was highly correlated (p value <0.05) with nitrogen and phosphorus species in multiple clustering schemes. Significant relationships were also identified between MST markers and variables that demonstrated temporal trends driven by precipitation and spatial trends driven by septic tanks and management practices (tillage and drainage) when spatial clustering was employed.

2.2 Introduction

Non-point sources of pollution have significant impacts on water quality and create health risks through a variety of hazards, including the spread of pathogens, eutrophication,

harmful algal blooms, and increased sedimentation (Bullerjahn et al., 2016; Smith et al., 2015; Vermeulen et al., 2015; Sharpley et al., 2015; Wen et al., 2017; Mateo-Sagasta et al., 2018; Zandaryaa and Mateo-Sagasta, 2018). Overland agricultural runoff is being recognized around the world as having increased impacts on water quality, overtaking known urban and industrial sources as the most prominent contributors to eutrophication of coastal and inland waters particularly in some high-income nations (OECD, 2012; USEPA, 2012; Bonsch et al., 2015; OECD, 2017; Mateo-Sagasta et al., 2018). For example, agricultural runoff has been found to contribute up to 44 and 58% of the phosphorus and nitrogen, respectively entering the Laurentian Great Lakes (Robertson and Saad, 2011). Understanding these impacts on water quality is important in areas such as the state of Michigan, which has three times more agricultural than urban land cover, and has seen an increase in manure application, irrigated land, and the use of organic fertilizers (Michigan Land Use Leadership Council, 2003; USDA, 2019). This increasing trend in organic agricultural practices and their corresponding increase in economic importance for states and local farmers, along with their known impacts on water quality, represent a growing area of uncertainty (USDA, 2019).

Several key agricultural water pollution research needs and knowledge gaps have been identified, including: the need for pollution source identification, identification and testing of locally appropriate markers, and the need to model the pathways of microbial contaminants (Evans et al., 2019). Current strategies using simply *E.coli* to understand fecal and nutrient pollution and monitor large complex watersheds are insufficient to address the most important water quality risks (Evans et al., 2019). Periodic sampling may provide a temporal “snapshot” of water quality, but the ability to sufficiently sample frequently enough is restricted due to the

substantial costs in per sample analysis, in particular when monitoring for multiple water quality variables (e.g., microbial markers, nutrients, streamflow, etc.) (Luszcz et al., 2017).

Current routine fecal indicator bacteria (FIB), such as *Escherichia coli* (*E. coli*), limit the ability to address microbial non-point sources of pollution because they cannot be used to determine the contamination sources. This is because *E. coli* has a ubiquitous presence in the feces of warm-blooded animals and regrows in aquatic environments (Reischer et al., 2013; Mayer et al., 2018; Zhang et al., 2018). Molecular source tracking (MST) markers allow for differentiation of fecal contamination from different hosts, and their presence in environmental waters allows for the identification of pollution sources (Boehm et al., 2013; Harwood et al., 2014; Ahmed et al., 2019).

Previous studies have begun to examine links between land use at various scales with water quality variables, however few studies have attempted to analyze and integrate of MST data with the chemistry, hydrology, geology, and spatial ecology of the system (Strayer et al., 2003; Floyd et al., 2009; Martin et al., 2017). Instream monitoring for microbes and nutrient contaminants and their relationships may be necessary to better understand how non-point sources of pollution within watersheds impact water quality. This study examined the impacts of non-point source pollution, including fecal contamination and nutrient loading within five mixed use Michigan watersheds that were experiencing high nutrient, *E. coli*, and MST levels from multiple sources (e.g., human, animal, chemical fertilizer) (Verhoughstraete et al., 2015, Nshimiyimana et al., 2018). This study had three main objectives: i) to determine the spatial and temporal trends in microbial contamination and nutrients in five mixed-use watersheds, ii) to determine if MST markers could be significantly correlated with nutrient levels (e.g., phosphorus and nitrogen) and iii) to determine which watershed variables (e.g., nutrient levels, land use,

management practices, etc.) within a watershed would be useful in predicting microbial contamination levels.

2.3 Materials and methods

2.3.1 Study area and sample collection

Water samples were collected from five watersheds in Michigan's Lower Peninsula, ranging in area from 14 km² to 2,683 km² (Figure 1). Grab samples (n=136) were collected from 17 sites [Little Pigeon (LPR, n=1), Sandy Creek (SC, n=2), Kawkawlin (KAW, n=3), Macatawa (MAC, n=4), and River Raisin (RR, n=7)] during eight sampling events, between April 2017 and May of 2018, representing the growing season (April – August 2017), fall/winter baseflow (November 2017) and spring snowmelt (March and May 2018). Sandy Creek (SC) is a very small watershed adjacent and just east of the RR. Individual sampling sites were selected based on the ability to subdivide watersheds into distinct land use areas that had adequate streamflow at bridge crossings and lack of interference from lake effects. Due to limited streamflow and small geographic size, the LPR watershed was assigned a single sampling site draining a 13 km² area. The SC watershed had two sampling sites, SC1 and SC2, draining 78 and 13 km², respectively. The MAC watershed was subdivided into four areas with the sampling sites MAC1, MAC2, MAC3, and MAC4 draining 31, 298, 77, and 50 km², respectively. The KAW watershed was subdivided into three areas with sampling sites KAW1, KAW2, and KAW3 draining 213, 567, and 238 km², respectively. The RR watershed was significantly larger than the other four watersheds and was thus subdivided into seven areas, including sites RR1, RR2, RR3, RR4, RR5, RR6, and RR7 draining 2682, 281, 240, 1755, 1205, 210, 348 km², respectively. The sampling sites SC1, MAC2, KAW2, and RR1 were the sampling points at the outlet of the

watersheds downstream of all other sites.

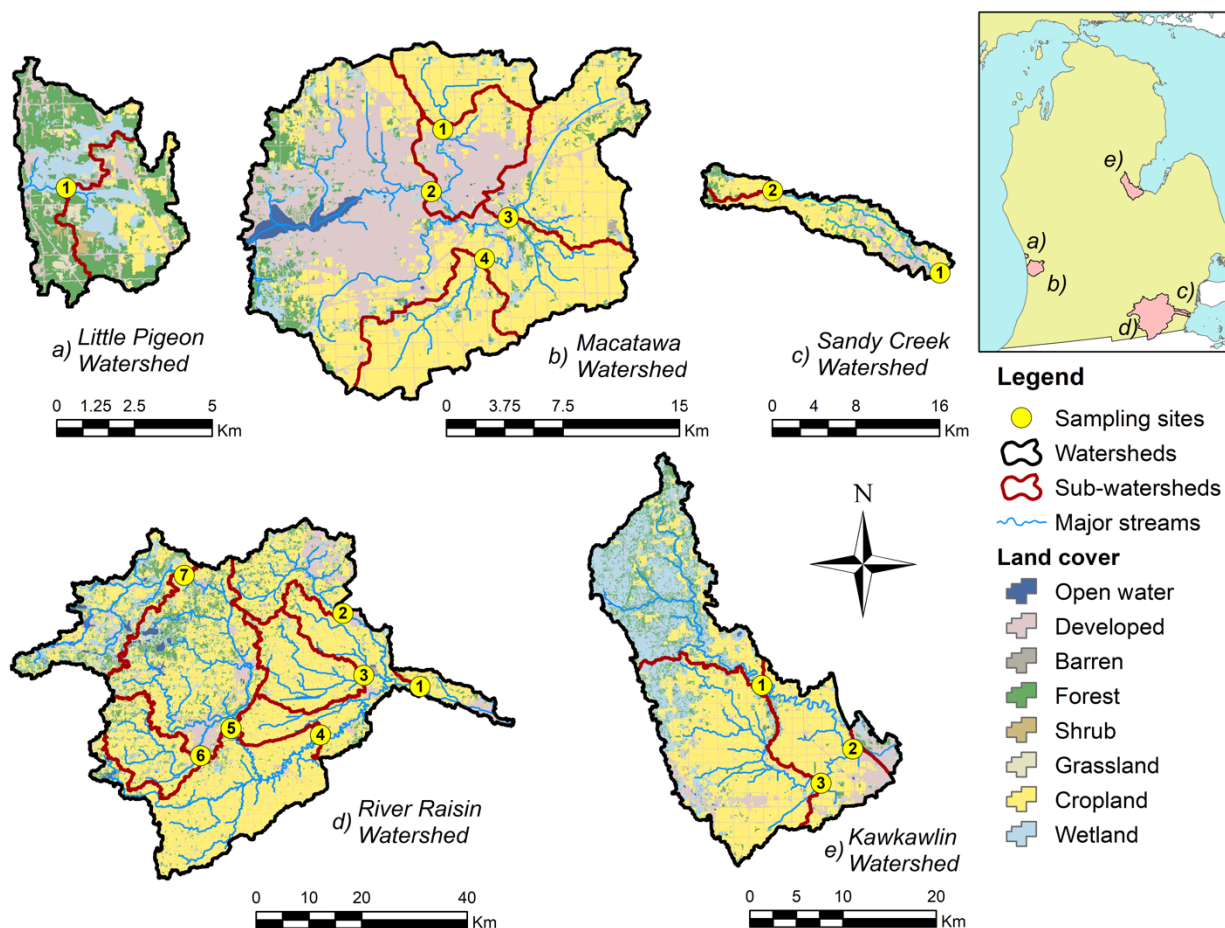


Figure 2.1 Maps of studied watersheds showing watershed locations in the state of Michigan, major streams and waterbodies, general land use, and drainage areas for each sampling location. a) Little Pigeon River (LPR) watershed, b) Macatawa (MAC) watershed, c) Sandy Creek (SC) watershed, d) River Raisin (RR) watershed, and e) Kawkawlin (KAW) watershed.

Watersheds were selected based on elevated microbial and nutrient results from Verhoughstraete et al. (2015) and Nshimiyimana et al. (2018), with the exception of the Kawkawlin Watershed, which was chosen based on known impaired waters of interest in the Saginaw Bay area. A total of 2.6 L was collected from each site during each sampling, with two 1 L volumes collected for microbial analysis and 0.6 L for nutrient and ion analysis. One duplicate sample was collected for nutrient and ion analysis on each day of sampling. Grab

samples were collected on the upstream side of the center of each stream crossing (e.g., bridges). Samples for microbial analysis were then transported on ice to the laboratory at Michigan State University (MSU), where they were stored at 4°C until processing and water samples for nutrient analysis were transported on dry ice and stored at -20°C until analysis.

2.3.2 Flow, physiochemical, and nutrient methods

Streamflow was measured during each sampling event (n=136) at the 17 sites (Figure 1) using either an acoustic Doppler current profiler or a Marsh McBirney Flo-Mate flow meter following US Geological Survey (USGS) protocols (Jarrett, 1991) depending on stream depth with the exception of RR sites 1, 5, and 7 where USGS gage flow data were available. For physiochemical parameters, a YSI 600R sonde (YSI Inc.) was used onsite to measure water temperature (°C), dissolved oxygen (mg L^{-1}), pH, and specific conductance (S/m). Turbidity measurements were performed from grab samples at the MSU laboratory after mixing using a LaMotte 2020we Turbidimeter (LaMotte Inc.). Nutrients [total dissolved nitrogen (TDN), total phosphorus (TP), nitrate (NO_3), nitrite (NO_2), ammonium (NH_4), and soluble reactive phosphorus (SRP)] and ions (Na, K, Mg, Ca, Cl, and SO_4) concentrations (mg/L) were measured in each sample following conventional protocols (Crumpton et al., 1992, Clesceri et al., 1998, Wetzel and Likens 2000, Hamilton et al., 2009) as previously described in Verhougstraete et al. (2015). Nitrogen and phosphorus were partitioned into their different species before analysis, with nitrogen being disaggregated into nitrates ($\text{NO}_2 + \text{NO}_3$), NH_4 , and dissolved organic nitrogen (DON) and phosphorus disaggregated into total reactive phosphorus (TRP), filterable reactive phosphorus (FRP), total filterable phosphorus (TFP), and total phosphorus (TP).

2.3.3 Water sample processing for microbial analysis

A 100ml subsample of each sample was used for *E. coli* and coliform testing using Colilert 18 (IDEXX, ME, USA) according to standard methods, while 900ml was filtered through multiple 47 mm 0.4 μm polycarbonate filters (Whatman, NJ, USA) using a sterile magnetic filter funnel (PALL Corporation, NY, USA) in 100ml aliquots. Individual filters were folded and added to sterile 2.0 ml screw cap tubes (VWR, PA, USA) containing ~ 0.3 g of 212-300 μm acid washed glass beads (Sigma-Aldrich, MO, USA) and stored at for -80°C until DNA extraction. A filtration blank of 100 ml of sterile phosphate buffered water was run with each set of samples. One 100ml filtered subsample was used for DNA extraction. The samples' filters were processed using a modified version of the Environmental Protection Agency Draft Method C (USEPA, 2014) crude DNA extraction method. A total of 590 μl of AE buffer (Qiagen, CA, USA) was added to each tube containing the sample filter and glass beads. The tubes containing each sample filter were then subjected to bead beating for manual cell disruption and DNA extraction at maximum speed for 1 min in a BioSpec Mini-Beadbeater (BioSpec, NH, USA). After bead beating, sample tubes were centrifuged at $12,000 \times g$ for 1 min to pellet any unwanted debris and glass beads. The supernatant (~ 400 μl) was then transferred to a clean 1.5 ml microcentrifuge tube for a 3 min centrifugation at $12,000 \times g$. The supernatant was then transferred to a final 1.5ml microcentrifuge tube and analyzed by Nanodrop (ThermoFischer, MA, USA) to confirm the presence and estimated concentration of total DNA. Whenever possible, sample DNA was analyzed the same day as DNA extraction, with the exception of QA/QC failed runs which were rerun within 24 hours. To avoid any unnecessary degradation of the DNA due to multiple freeze/thaws, each DNA sample was aliquoted into multiple tubes and stored at -80°C .

2.3.4 Microbial molecular analysis methods

The detection of MST markers was performed using droplet digital PCR™ (ddPCR). Three MST markers for this study target human (*B. theta* α -1-6, mannanase), bovine (CowM2), and porcine (Pig2Bac) fecal contamination (Yampara-Iquise et al., 2008; Shanks et al., 2008; Mieszkina et al., 2008) (Table 1). Three replicate ddPCR reactions were performed for each sample with the human and bovine markers analyzed in duplex, while the porcine marker was analyzed alone. Each 22 μ l ddPCR reaction setup contained 1X Supermix for Probes (no dUTP) (Bio-Rad, CA, USA), 900 nmol l⁻¹ of each primer, 250 nmol l⁻¹ of each probe, 0.9 μ l of molecular grade DNase-free water, and 5.5 μ l of template DNA. Microfluidic droplet generation was performed by the Droplet Generator (Bio-Rad, CA, USA) by combining 20 μ l of each reaction mixture with 70 μ l of droplet generation oil resulting in ~20,000 droplets. The resulting 40 μ l oil-reaction mixture emulsions were then transferred to a 96-well PCR plate, heat-sealed with foil and placed into a T100 thermocycler (ramp rate of 2°C s⁻¹) (Bio-Rad, CA, USA) for PCR amplification using the following parameters: 95°C for 10 min, followed by 40 cycles of 94°C for 30 s and 60°C for 1 min then a final cycle of 98°C for 10 min. The plate was then transferred to a QX200 Droplet Reader (Bio-Rad, CA, USA) for the fluorescent detection of positive droplets in each well using the RED (rare event detection) setting.

Table 2.1 Primer and probes for ddPCR MST analysis

Assay and Sequence Type	Primer or Probe Sequence (5' to 3')	Size (bp)	Reference or Source
B. theta α -1-6, mannanase	B.theta 4515901F: CATCGTTCGTCAGCAGTAACA B.theta 4515963R: CCAAGAAAAAGGGACAGTGG B.theta Probe: FAM-CAGCAGGT-NFQ ^{a,b}	63	Yampara-Iquise et al., 2008
CowM2	M2F: CGGCCAAATACTCCTGATCGT M2R: GCTTGTTGCGTTCCTTGAGATAAT M2P: HEX-AGGCACCTATGTCCTTTACCTCATCAACTACAGACA-BHQ1	92	Shanks et al., 2008
Pig2Bac	Pig2Bac41F: GCATGAATTTAGCTTGCTAAATTTGAT Pig2Bac163Rm: ACCTCATACGGTATTAATCCGC Pig2Bac113MGB: FAM-TCCACGGGATAGCC-BHQ1	116	Mieszkin et al., 2009

^aProbe ordered through Roche Universal Probe Library; UPL probe # 62

^bNFQ: Non-fluorescent Quencher

Strict quality control measures were followed for all ddPCR assays. Each assay plate was analyzed with three wells of non-template controls (NTC) (molecular grade DNase-free water), filtration blanks for each batch of samples, and three positive control wells for each assay target. Assay results were only considered for further analysis if >10,000 accepted droplets were achieved in each sample well and within each control well. In addition, any positive NTC wells were considered indicative of possible contamination of the reaction master mixture and all sample results from that plate were rejected and the samples rerun. Samples were only considered true positives if ≥ 3 droplets were positive (above reaction threshold for positive-negative distinction. Samples were considered as positive if at least one technical replicate (1/3) were found to be positive, and close to the calculated detection limit of the assay (354 gene copies (GC) 100ml⁻¹).

2.3.5 Landscape data

Hourly precipitation time series were obtained at the closest NOAA land-based stations to each sampling location (NOAA, 2019). Then, cumulative hourly and multiday precipitation totals (mm) were obtained for the period before each sample collection time (e.g., 6 hr, 12 hr, 18 hr, 24 hr, 2 days, 3 days, 4 days, 6 days, 8 days, 15 days, and 30 days). Land cover proportions for each sampling location's drainage area were obtained by processing the Cropland Data Layer (CDL) 2017 (USDA-NASS, 2017) for details, and the National Land Cover Database (NLCD) for the general land covers. Tillage practice information was obtained from a national survey completed by the USGS spanning 1989 to 2004 (Baker, 2011). The estimated number and density of septic systems per watershed area and tile drainage's proportions were obtained from Luscz et al. (2015, 2017).

2.3.6 Statistical analysis

Multiple methods were utilized to investigate instream variables (e.g., temperature, dissolved oxygen (DO), pH, conductivity, streamflow, non-purgeable organic carbon (NPOC), TDN, Na, K, Mg, Ca, Cl, SO₄, turbidity, NH₄, NO₃, DON, SRP, and TP) and landscape variables (e.g., prior precipitation, land use, tile drainage, septic tank numbers, septic tank density, and tillage) to help explain the levels of bacterial indicators of fecal pollution (i.e., MST markers, and *E. coli*) found in the five watersheds. Bacterial markers were considered as the response variables in all analyses with the geometric means of the log₁₀ transformed technical replicates of each being used for statistical analysis. Non-detect (ND) replicates were included and assigned the assay's detection limit (2.55 log₁₀ GC 100ml⁻¹).

To avoid collinearity, redundant variables were identified and removed from the dataset prior to analysis by determining the pairwise relationships among all explanatory variables (i.e., Temperature, DO, pH, conductivity, streamflow, NPOC, TDN, Na, K, Mg, Ca, Cl, SO₄, turbidity, NH₄, NO₃, DON, SRP, TP, prior precipitation, land use, tile drainage, septic tank numbers, septic tank density, and tillage) across all sampled locations using Spearman's rank correlations ($\rho > 0.7$). The variable that had the lowest average correlation with other predictor variables using the 'findCorrelation' function from the caret package in R (Kuhn, 2020) was retained. Six variables (i.e., TDN, SO₄, Ca, Mg, Cl, and SRP) were found to be collinear with other variables and were thus removed from the dataset.

2.3.6.1 Spatial clustering

The clustering analysis consisted of an agglomerative bottom-up hierarchical approach using standardized Euclidean distances and the Ward's minimum variance method, followed by a single k-means clustering iteration using up to 3 clusters from the hierarchical approach. This analysis was performed to explicitly account for spatial variability using the 'hclust' and 'kmeans' functions in R (R Core Team, 2019). To identify the main factors driving this variability, clusters of sampling locations were determined using six different criteria; *E. coli* and MST marker concentrations, streamflow, land use, tillage, tile drain proportion, and septic tank density. Bacterial (*E. coli* and MST) marker concentrations and instream variable (i.e., temperature, DO, pH, conductivity, streamflow, NPOC, TDN, K, Mg, Cl, turbidity, NH₄, DON, SRP, and TP) raw data were Box-Cox transformed using the 'BoxCoxTrans' function from the caret package in R prior to being used for clustering (Kuhn, 2020).

Correlation analysis was performed using Pearson correlation (r) analysis using GraphPad Prism 8 (GraphPad Software, CA, USA) to ascertain if any significant relationships existed between the three MST markers and nutrient species. These analyses were conducted in two ways. First by using data from all the sites and samples and secondly by using data from the various sampling site clustering configurations (Figure 4).

2.4 Results

2.4.1 Water quality summary of five watersheds

Five watersheds from smallest to largest (LPR 14 km²; SC 82 km²; MAC 292 km²; KAW 582 km²; RR 2,683 km²) were sampled over three seasons (spring, summer, and fall). Selected highlights of important water quality variables collected during this study are presented here with more information provided in Supplemental Materials. The land uses for each sampling site are described in Supplemental Table A1; a detailed land use classification was used for clustering and statistical analysis as described in the materials and methods. The areas drained by each sampling site ranged from 8.16% (LPR1) to 76.37% (MAC4) agricultural land use (Supplemental Table A1).

Average streamflow for all sites ranged from 0.10 m³/s in the LPR to 8.95 m³/s in the RR (Supplemental Table A2). The RR watershed saw the maximum recorded streamflow at 87.50 m³/s in March 2018 while both the KAW and MAC watersheds had streamflows as low as 0 m³/s at multiple sites (KAW1, KAW2, MAC1, MAC4) during June, July, and August 2017 sampling dates. Overall streamflow was lowest during the summer sampling of 2017 and the highest during the spring sampling in 2017 and 2018.

Septic tank densities varied between and within watersheds. The watersheds (LPR, SC, MAC, KAW, and RR) had septic tank density ranges of 14.46, 7.53 to 20.79, 9.10 to 17.14, 6.20 to 19.39, and 6.87 to 23.85 septic tanks/ km², respectively.

The proportion of land within each watershed with tile drains ranged from 0 at LPR1. On average, the proportion of tile drains in LPR, SC, MAC, KAW, and RR were 0, 0.22, 0.30, 0.19, and 0.26, respectively.

Summaries of measured nutrients, MST markers, and physical water quality variables are provided in Supplemental Materials (Tables A2-A5). TP concentrations ranged from 14.77 to 111.24 µg/L, while TDN ranged from 0.66 to 4.90 mg/L across sites. *E. coli*, *B. theta*, CowM2, and Pig2Bac concentrations ranged from 0.30 to 4.30 Log₁₀ MPN/100ml, 2.71 to 2.83 Log₁₀ GC/100ml, 2.60 to 2.77 Log₁₀ GC/100ml, 2.96 to 3.23 log₁₀ GC/100ml, respectively.

2.4.2 Spatial and temporal trends in bacterial markers and nutrients

Individual sample concentrations for the four MST markers, two nitrogen species, and the two phosphorus species were plotted on heat maps to visualize spatial or temporal trends in each dataset (Figures 2-3). The MST markers primarily revealed temporal trends in their datasets, while the nutrient species showed spatial trends and some temporal trends. A few high valued data points were omitted from the scaling on the heatmaps to allow for greater visual resolution of spatial and temporal trends that would have otherwise been camouflaged by scaling the entire range of values.

up to 0.64 at RR3 with the MAC and RR watersheds having the highest proportions of tile drains.

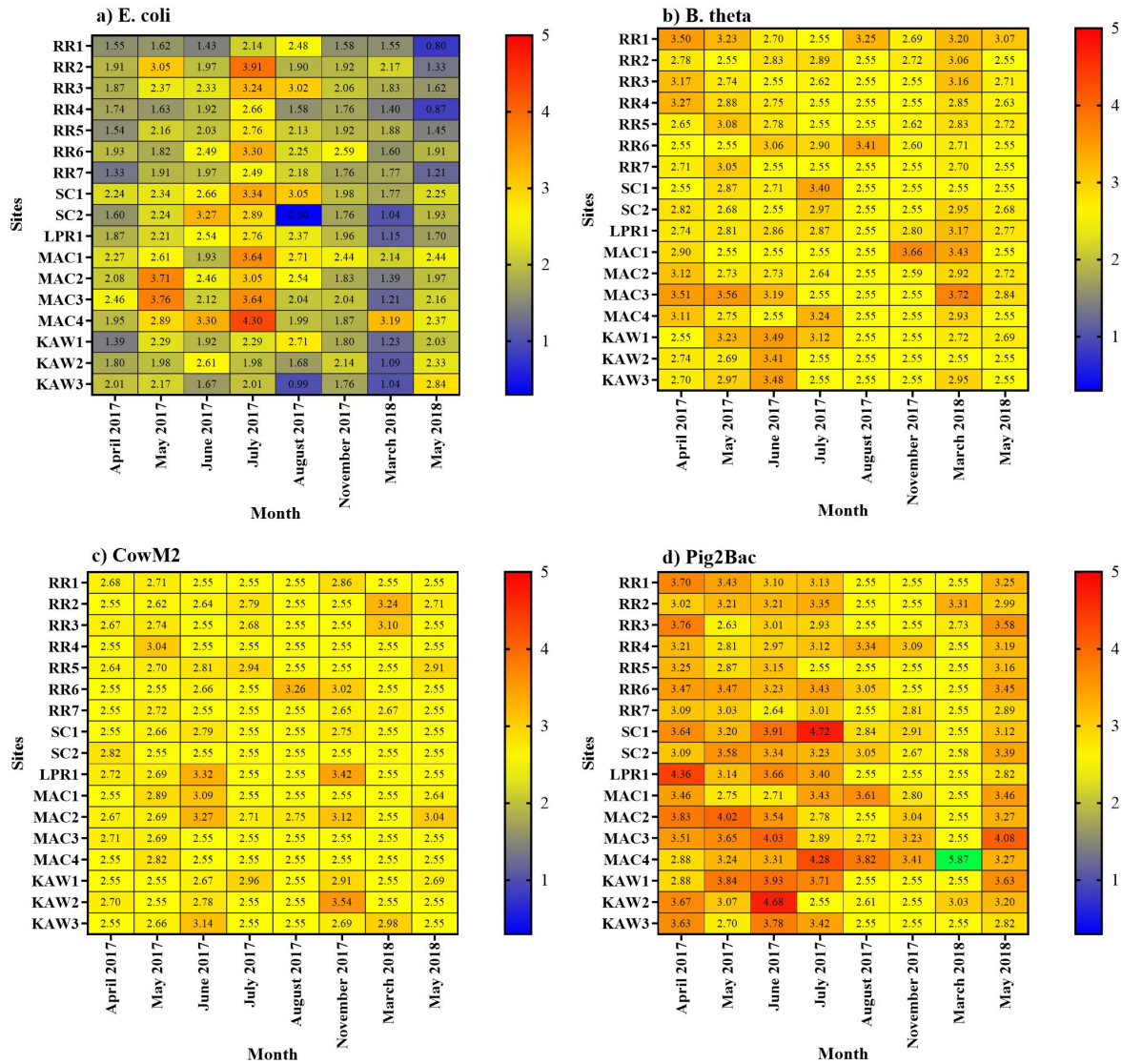


Figure 2.2 Microbial heatmaps for all watersheds and sampling months: a) *E. coli* Log₁₀ MPN/100ml, b) *B. theta* Log₁₀ GC/100ml, c) CowM2 Log₁₀ GC/100ml, d) Pig2Bac Log₁₀ GC/100ml. Cells colored with bright green were above the range depicted on the heatmap. These data points were removed from the depicted ranges to increase visibility of spatial and temporal patterns in the data.

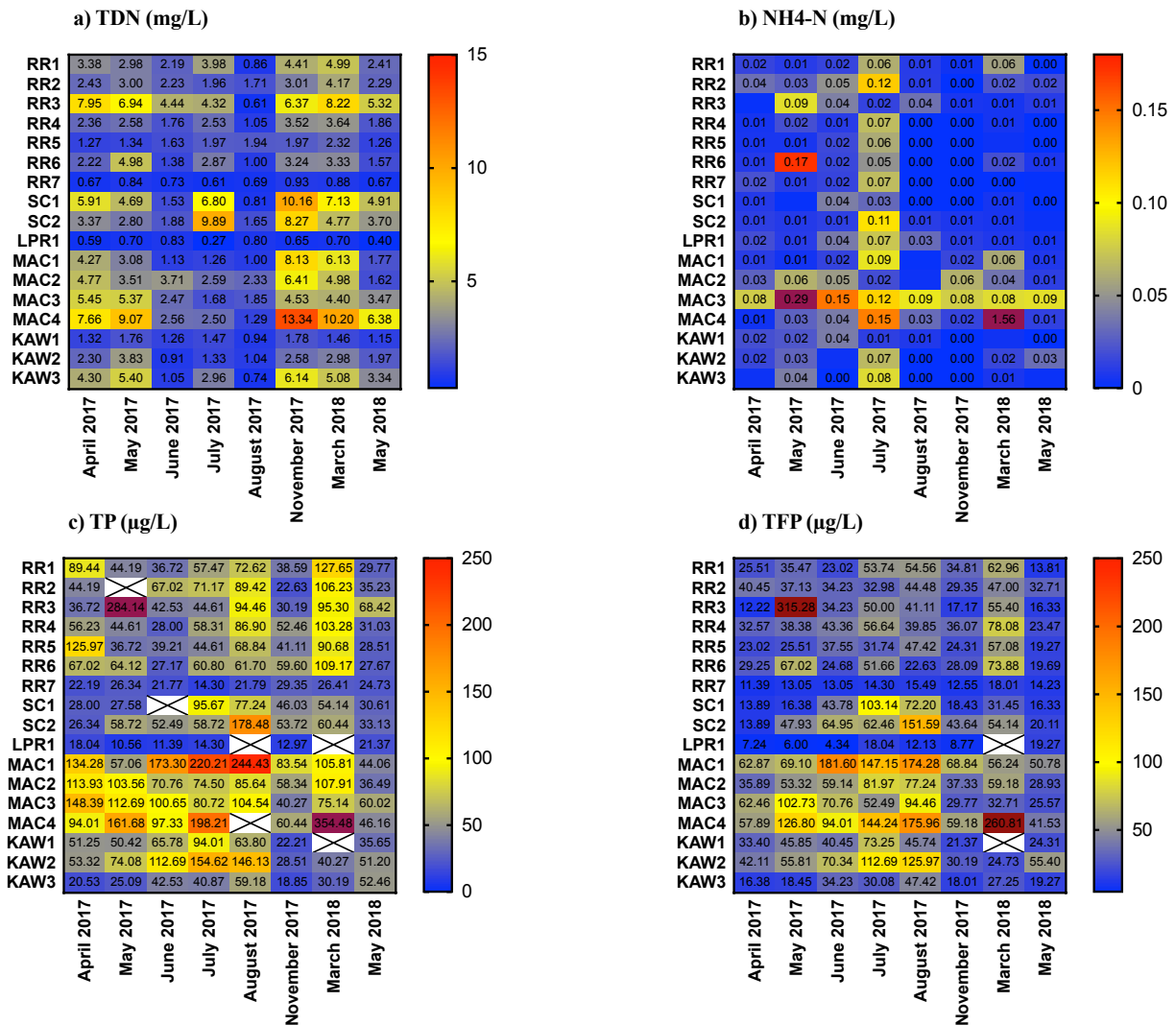


Figure 2.3 Phosphorus and Nitrogen species' heatmaps showing spatial and temporal distributions. a) total dissolved nitrogen (TDN), b) Ammonium (NH4-N), c) total phosphorus (TP), d) total filterable phosphorus (TFP). Individual ranges for each nutrient species are to the right of each heatmaps. Cells colored with dark red were above the range depicted on the heatmap. These data points were removed from the depicted ranges to increase visibility of spatial and temporal patterns in the data.

A general temporal trend was seen for *E. coli* in Figure 2a, with high concentrations during summer (July 2017) and low concentrations in spring (March 2018) across the most of sites. The human marker also showed temporal trends with higher concentrations during the spring and early summer months (i.e., April, May, June, and July 2017, as well as March 2018),

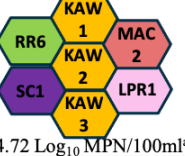
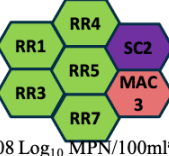
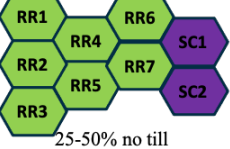
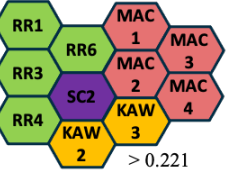
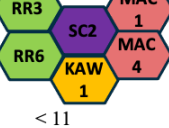
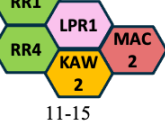
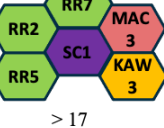
and lower concentrations in the late summer (August 2017), fall (November 2017) and stayed low in the spring May of 2018 (Figure 2b). The bovine marker had higher concentrations in June and November 2017, with its lowest concentrations in July and August 2017 coinciding with the driest (lowest levels of prior precipitation and streamflow) period throughout the study (Figure 2c). The porcine marker also showed higher concentrations between April, July 2017, and May 2018, with lower concentrations during August, November 2017, and March 2018 (Figure 2d).

Spatial and temporal trends were seen for all nitrogen and phosphorus species (Figure 3). The TDN results showed a temporal trend with lower concentrations during June and August 2017 and higher concentrations in November 2017 and March 2018 (Figure 3a). There were also spatial trends with sites such as RR5, RR7, LPR1, and KAW1 that consistently had low TDN concentrations across all sampling events. Ammonium showed a different spatial trend, with MAC3 having considerably higher $\text{NH}_4\text{-N}$ concentrations than all the other sites through time (Figure 3b). There was also a spike in ammonium concentrations across all sites in all watersheds in July 2017. TDN had similar spatial and temporal trends as in the nitrate heatmap, but with lower average concentrations of ammonium vs. nitrate. The phosphorus species showed a spatial trend with the highest concentrations found in the MAC sites for all individual species, but when examining total phosphorus (TP, Figures 3c and 3d), most of the RR sites (excluding RR7) had high phosphorus levels (Figure 3). There were also some temporal trends in the phosphorus species with higher concentrations of TP found in August 2017 and March 2018 (Figure 3c).

2.4.3 Statistical analysis

2.4.3.1 Cluster analysis

The data from individual sites in each watershed were separated by cluster analysis to examine the similarities between watershed sites by six variables, including *E. coli*/ MST marker concentrations, streamflow, land use, tillage, tile drain proportion, and septic tank density. The final cluster analysis resulted in five clustering schemes with land use and tillage producing identical clusters (Figure 4). Clustering sites based on streamflow resulted in three clusters (1, 2, and 3) representing low, medium, and high flows. Concentrations from all four markers *E.coli* and MST (*B. theta*, CowM2, Pig2Bac, and *E. coli*) had 1, 2, and 3 clusters with concentration ranges of 0.99-4.72, 0.30-4.08, 1.33-5.87 Log₁₀ MPN(GC)/100ml, respectively. When cluster analysis was performed using tile drain proportions, only two clusters were identified representing 0-0.172 and >0.221 proportion tile drains (clusters 1 and 2). Landuse/tillage resulted in three clusters with low (25-50 % with no tillage), medium (10 to 25% with no tillage) and high (0-10% with no tillage) tillage practices. Finally, septic tank density split into three clusters representing sites with <11, 11-15, and >17 septic tanks/ km², respectively.

	Cluster 1 (Low values)	Cluster 2 (medium values)	Cluster 3 (High values)
Streamflow	 Low flows	 Medium flows	 High flows
Markers	 0.99 -4.72 Log ₁₀ MPN/100ml ^a	 0.30-4.08 Log ₁₀ MPN/100ml ^a	 1.33-5.87 Log ₁₀ MPN/100ml ^a
Land use/Tillage	 25-50% no till	 10-25% no till	 0-10% no till
Tile drain proportion	 0-0.172	 > 0.221	
Septic Tank Density	 < 11	 11-15	 > 17

^a Concentration ranges include values obtained for all four markers (*B. theta*, CowM2, Pig2bac, and *E. coli*).

Figure 2.4 Cluster analysis results for streamflow, markers, land use, tillage, tile drain proportion, and septic tank density. Sites were clustered into up to three clusters representing “low”, “medium”, and “high” relative values for each category. Values for specific variables ranges of values are listed below each cluster.

2.4.3.2 Correlation Results

The human and bovine markers showed no significant relationships with any of the seven nutrient species when all of the sites were analyzed together. The porcine marker; however, showed statistically significant correlations with all four phosphorus species (i.e., TFP, FRP, TRP, TP) and ammonium with *r* values ranging from 0.22 (*p* = 0.0091) for FRP, to 0.48 (*p* = <0.0001) for ammonium.

Correlation analysis using the three streamflow clusters showed significant correlations between bovine and porcine markers with nutrient species (Figure 4). In streamflow cluster 1 (low flow), the porcine marker was correlated with TFP, FRP, TP, and NH_4 with r values of 0.26 ($p = 0.0192$), 0.26 ($p = 0.0250$), 0.31 ($p = 0.0072$), 0.54 ($p = <0.0001$), respectively. In cluster 2, the bovine marker was correlated with nitrate and TDN with r values of 0.37 ($p = 0.0372$) and 0.37 ($p = 0.0376$), respectively. In cluster 3, however the porcine marker was negatively correlated with TFP ($r = -0.48$, $p = 0.0166$).

Only the porcine marker showed significant relationships with nutrients when clustering by *E.coli*/MST markers or land use/tillage.

Sites were clustered into only two clusters when using land use/tillage. In the first land use/tillage cluster there were no significant correlations between any of the MST markers and nutrients. In the second land use/tillage cluster, the porcine marker was correlated with TFP, FRP, TRP, TP, and NH_4 with r values 0.42 ($p = 0.0083$), 0.42 ($p = 0.0073$), 0.41 ($p = 0.0109$), 0.47 ($p = 0.0030$), and 0.66 ($p = <0.0001$) respectively.

In the first cluster (low concentrations) for the *E.coli*/MST markers there were no significant relationships found between the bacteria and nutrients. In the marker cluster 2, the porcine marker was correlated with NH_4 ($r = 0.30$, p -value 0.0157). In marker cluster 3, the porcine marker was highly correlated with TFP, FRP, TRP, TP and NH_4 with r values of 0.68 ($p = 0.0002$), 0.67 ($p = 0.0003$), 0.60 ($p = 0.0024$), 0.73 ($p = 0.0001$), and 0.82 ($p = <0.0001$), respectively.

The tile drainage clusters showed significant correlations for both the human and porcine markers. In cluster 1, the human marker was correlated with TFP and FRP with r values of 0.40 ($p = 0.0062$) and 0.35 ($p = 0.0187$) respectively. The porcine marker was also correlated with

FRP in cluster 1 with an r value of 0.21 ($p = 0.0485$). In cluster 2, the porcine marker was correlated with TFP, TRP, TP and NH_4 with R values of 0.21 ($p = 0.0467$), 0.26 ($p = 0.0168$), 0.31 ($p = 0.0032$), and 0.56 ($p = <0.0001$).

Septic tank density clustering resulted in the most correlations between markers and nutrient species. The human marker was correlated with TFP, TP, and NH_4 in cluster 3 (highest density) with r values of 0.4 ($p = 0.0037$), 0.43 ($p = 0.0027$), and 0.41 ($p = 0.0036$) respectively. The bovine marker was not significantly correlated in any of the septic tank density clusters. The porcine marker was correlated with TFP, FRP TP, and NH_4 in both cluster 1 and cluster 3. In cluster 1, the porcine marker was correlated with TFP, FRP, TP, and NH_4 had r values of 0.31 ($p = 0.0333$), 0.29 ($p = 0.0475$), 0.39 ($p\text{-value } 0.0068$), and 0.68 ($p = <0.0001$) respectively. In cluster 3, the porcine marker was correlated with TFP, FRP, TP, and NH_4 had r values of 0.34 ($p = 0.0170$), 0.29 ($p = 0.0492$), 0.30 ($p = 0.424$), 0.32 ($p = 0.0247$), respectively.

2.5 Discussion

The data in this study show that individual mixed-use watersheds have unique spatial and temporal trends for both microbial contaminants and nutrients. Our results were in line with previous research where microbial contamination trends are mainly temporal in nature (Lee et al., 2014; Sowah et al., 2017; Nshimiyimana et al., 2018; Badgley et al., 2019; McKee et al., 2020; Hinojosa et al., 2020). These trends are likely associated with the timing of manure applications and microbial transport through the watersheds seen with increased rainfall and overland flows. Nutrients were spatially segregated based on land use and management practices such as tillage and tile drainage.

The MAC watershed, in particular, showed high phosphorus levels at all four sites over the course of the sampling period. The phosphorus pollution in the MAC watershed is linked to agricultural non-point sources of pollution (Steinman et al., 2018). Our results showed correlations between our porcine marker and nutrient species when using the data from the MAC watershed sites in multiple clustering schemes. These results suggest that at least a portion of the nutrients entering the MAC watershed are associated with manure application practices. These high levels of phosphorus are unsurprising as Lake Macatawa, which the watershed drains into, has been hypereutrophic for over 40 years (MWP, 2012). A TMDL of 50 ug/L was set for TP in Lake Macatawa in 1999 by the USEPA, with best management practices (BMPs) aimed at runoff abatement implemented since 2012 to help alleviate nutrient pollution (Walterhouse, 1999; Holden, 2021; Steinman et al., 2018). However, these BMPs have yet to produce the desired results, with our study's finding an average total phosphorus concentration of 58.75 ug/L in the streams draining into the lake. This is consistent with the effects of land use legacy, where changes to the landscape can take decades to propagate through the environment to surface water systems (Martin et al., 2021).

A particularly interesting result of our study is the application of sampling site clustering to elucidate masked correlations between MST markers and nutrients. This is of significance because it shows that water quality monitoring by itself without considering similarities and differences between sampling sites may mask relationships between variables and potential contamination sources. While we identified correlations between the porcine marker and both phosphorus and nitrogen species, the human and bovine markers were less frequently connected to nutrients.

The cow marker was not correlated with nutrients. Cow manure because of its high solids content and lower water activity may release the cow marker at a much different rate compared to nutrients, however this is speculative and no data has been generated to support this. The higher solids content in bovine manure (265 lbs per 1,000 gals) vs. porcine manure (170 lbs per 1,000 gals) changes the availability and uptake of nutrients by crops from the manure (i.e., nitrogen availability in the first year after application from dairy cow manure is 50 to 70% and only 30 to 50% for swine (Lorimor et al., 1980; Zhang, 2017)). Thus, influencing the difference between the animal MST markers and nutrients. The diffuse spreading of cow manure on pastures with grazing livestock or overland spreading of cow manure may have also diluted out the affect. In addition, fertilizer may be a greater source of nutrients compared to manure in these rural areas.

The human marker was correlated with nutrient species when sampling sites were clustered based on septic tank density suggesting that septic tanks are an important source of nutrients in watersheds that have higher septic tank densities (Figure 4). However, the human bacterial marker may be transported through soil in a different time frame compared to nutrients. The input from septic tanks to surface waters was more apparent during low flows, where groundwater contributions to surface waters compared to overland flow are not masked. This is in line with observations from previous studies by Verhougstraete et al. (2015) and Sowah et al. (2017). Additionally, Joseph et al. (2021) showed that human inputs were more highly correlated with stream contamination than bovine sources even in areas with high numbers of cattle.

Another important aspect to consider when examining correlations between MST markers and nutrient levels is the persistence of each marker. The persistence of the MST markers varies and influences the presence (degradation) of the markers. The Pig2Bac marker

T_{90} (time for 90% decay) ranges from 0.90 to 5.11 days, while the human marker the B. theta and CowM2 have reported T_{90} 's of 1.8 and 3.14 days, respectively (Korajkic et al., 2018; He et al., 2016; Brooks et al., 2015; Ballesté et al., 2018). The longer persistence of the porcine marker in water may allow for higher correlations with nutrient species. This correlation is likely influenced by both the source, transport, and fate of the contaminants.

In nearly every cluster, the porcine marker showed correlations with ammonium. This suggests that when porcine manure is applied to the land, a significant portion of the manure and accompanying nutrients make their way into the waterways. This, in turn, along with the lack of bovine markers correlated with nutrients, could mean that the method of application or the physical attributes of the manure sources are significant in the fate of nutrient and microbial contamination to streams.

The correlations between nutrients and MST markers when sampling sites were clustered by septic tank density suggest that septic tank density may help predict where higher levels of fecal pollution will occur. For the human marker, this is in line with a large-scale study conducted by Verhougstraete et al. (2015), which identified a similar correlation between septic tank numbers in watersheds and an increase in the human marker. It is not clear why the pig marker would correlate with nutrients based on septic tank density, but it may be indicative of greater manure application to the available land in rural areas where septic tanks are more widely used.

While this study successfully identified relationships between MST markers and nutrient species, no strong predictive relationships could be determined. This suggests that while MST can help to identify contamination sources within watersheds, there are too many variables in the accompanying water quality data to allow for a strong predictive model to be formed. However,

the abundant correlations between the porcine marker and phosphorus and ammonium show that porcine manure is likely an important source of pollutants from agricultural land that is transported into streams that should be monitored more frequently pre- and post- applications.

To reach desired stream water quality, particularly in problematic agriculturally intensive watersheds, manure and septic tanks both need to be considered for control of microbials and nutrients.

2.6 Conclusions

- Spatial clustering allows for a more accurate analysis of relationships of water quality variables in watersheds.
- Temporal contamination is primarily driven by precipitation and its associated variables (e.g., streamflow, turbidity, overland flow), while spatial contamination is driven by land uses (e.g., septic tank density, tile drain proportions, and tillage).
- Porcine fecal contamination is more often correlated with nutrients in streams than either bovine or human contamination.

APPENDIX

Water temperatures varied over the three seasons, but not between watersheds (11.60 to 16.04°C). DO levels on average were 9 mg/L in four (LPR, SC, MAC, RR) of the watersheds, while KAW sites had an average of 7.97 mg/L. The lowest individual DO measurement was from MAC (3.12 mg/L). Average pH values varied little between watersheds, with averages from 7.82 to 8.17. The lowest measured conductivity occurred in July 2017 in the KAW watershed (212.50 $\mu\text{S}/\text{cm}$), while SC had the highest single sample conductivity (934.60 $\mu\text{S}/\text{cm}$) in August 2017. Average conductivity for all of the watersheds ranged from 310.65 to 600.93 $\mu\text{S}/\text{cm}$. The KAW and LPR watersheds had low average turbidities 2.47 and 2.78 NTU, respectively, while RR, MAC and SC watersheds had high average turbidities (i.e., 10.02, 8.60, and 9.60 NTU, respectively). These varied seasonally along with streamflow.

The MAC watershed had the highest concentration of K, Mg, and Na with concentrations of 6.04 mg/L, 17.82 mg/L, and 28.42 mg/L, respectively (Supplemental Materials Table A3). The LPR watershed had the lowest concentrations of Mg, Na, and non-purgeable organic carbon (NPOC) with concentrations of 1.71 mg/L, 12.24 mg/L, and 11.53 mg/L, respectively. The KAW watershed was found to have the lowest K levels (1.71 mg/L) and the highest NPOC concentration (20.16 mg/L). The RR and SC watersheds showed similar levels of ions as each other. Ca concentrations were found to be highest in RR and SC at 76.01 and 77.60 mg/L, respectively. SO_4 concentrations were found to be highest in the SC and RR watersheds (i.e., 58.72 and 50.24 mg/L respectively) and lowest in the KAW watershed (i.e., 21.02 mg/L).

Table 2.A1 Land use percentages for each sampling site's drainage area

Sampling Site	Agricultural	Developed	Water	Undeveloped^a
KAW1	17.60	4.31	0.57	77.52
KAW2	44.67	10.98	0.45	43.90
KAW3	60.23	11.75	0.32	27.71
LPR1	8.16	13.13	0.22	78.50
MAC1	74.58	8.33	0.19	16.90
MAC2	57.09	21.88	0.28	20.75
MAC3	57.02	17.42	0.36	25.20
MAC4	76.37	8.11	0.10	15.42
RR1	55.07	10.84	1.58	32.51
RR2	45.76	19.65	0.70	33.89
RR3	78.48	6.38	0.09	15.06
RR4	49.95	10.16	2.16	37.73
RR5	36.81	11.84	2.93	48.41
RR6	56.05	5.90	0.38	37.68
RR7	24.28	10.63	4.85	60.25
SC1	52.88	20.58	0.24	26.30
SC2	54.77	6.03	0.15	39.05

^aUndeveloped land use includes the following land use categories: barren, forest, herbaceous, shrubland, and wetland

Table 2.A2 Physiochemical summary results by watershed

Watershed	Water Temperature (°C) Average (Range)	DO (mg/L) Average (Range)	pH Average (Range)	Conductivity (µs/cm) Average (Range)	Streamflow (m³/s) Average (Range)	Turbidity (NTU) Average (Range)
River Raisin (RR)	16.04 (2.14-27.52)	9.69 (6.19-13.78)	8.11 (7.19-8.53)	600.93 (396.40-875.80)	8.95 (0-87.50)	10.02 (1.17-36.40)
Kawkawlin (KAW)	14.88 (0.78-25.20)	7.97 (4.31-13.40)	7.89 (7.40-9.12)	436.66 (212.50-750.70)	1.93 (0-11.44)	2.78 (1.05-7.94)
Macatawa (MAC)	15.61 (2.75-25.31)	9.30 (3.12-14.60)	8.00 (6.64-8.62)	576.34 (403.60-729.10)	0.65 (0-2.67)	8.60 (0.97-46.60)
Sandy Creek (SC)	13.98 (1.08-22.34)	9.37 (3.99-13.32)	7.82 (7.19-8.26)	577.01 (285.40-934.60)	0.33 (0-1.57)	9.60 (1.90-39.10)
Little Pigeon (LPR)	11.60 (3.89-16.80)	9.04 (5.65-12.19)	8.17 (7.54-8.89)	310.65 (220.50-375.80)	0.10 (0.03-0.18)	2.47 (1.62-3.35)

Table 2.A3 Ion summary results by watershed

Watershed	K (mg/L) Average (Range)	Mg (mg/L) Average (Range)	Ca (mg/L) Average (Range)	Cl (mg/L) Average (Range)	Na (mg/L) Average (Range)	NPOC (mg/L) Average (Range)
River Raisin (RR)	2.34 (0.83-4.01)	16.94 (7.57-33.01)	76.01 (43.43-115.92)	37.08 (0.01-66.99)	20.91 (9.31-47.42)	13.54 (4.68-50.17)
Kawkawlin (KAW)	1.71 (0.90-3.48)	13.21 (5.18-24.10)	53.91 (32.64-93.59)	28.44 (0.04-54.77)	17.27 (3.13-35.09)	20.16 (7.42-60.91)
Macatawa (MAC)	6.04 (1.06-10.81)	17.82 (4.96-25.79)	58.36 (28.31-81.18)	43.49 (16.05-68.57)	28.42 (4.74-55.49)	14.05 (5.74-41.54)
Sandy Creek (SC)	3.81 (1.96-8.68)	13.45 (4.74-30.89)	77.60 (37.25-130.10)	36.35 (20.38-57.71)	19.05 (9.03-39.14)	17.38 (3.34-55.04)
Little Pigeon (LPR)	2.07 (1.55-2.77)	9.07 (5.37-11.55)	41.79 (33.71-49.42)	28.24 (27.63-29.85)	12.24 (10.77-13.33)	11.53 (4.56-35.69)

Table 2.A4 Microbial summary results by watershed

		Watersheds				
Markers		RiverRaisin (RR)	Kawkawlin (KAW)	Macatawa (MAC)	Sandy Creek (SC)	Little Pigeon (LPR)
<i>E. coli</i>	% Positive	100 (56/56)	100 (24/24)	100 (32/32)	100 (16/16)	100 (8/8)
	Geomean ^a	1.91	1.84	2.41	1.94	2.01
	(Range)	(0.80-3.91)	(0.99-2.84)	(1.21-4.30)	(0.30-3.34)	(1.15-2.76)
B. theta	% Positive	64.29 (36/56)	54.17 (13/24)	59.38 (19/32)	50 (8/16)	87.5 (7/8)
	Geomean ^b	2.77	2.77	2.83	2.71	2.82
	(Range)	(2.55-3.50)	(2.55-3.49)	(2.55-3.72)	(2.55-3.17)	(2.55-3.17)
CowM2	% Positive	42.86 (24/56)	45.83 (11/24)	40.63 (13/23)	25 (4/16)	50 (4/8)
	Geomean ^b	2.66	2.69	2.67	2.60	2.77
	(Range)	(2.55-3.26)	(2.55-3.54)	(2.55-3.27)	(2.55-2.82)	(2.55-3.42)
Pig2Bac	% Positive	71.43 (40/56)	66.67 (16/24)	87.5 (28/32)	93.75 (15/16)	62.5 (5/8)
	Geomean ^b	2.96	3.07	3.23	3.2	3.07
	(Range)	(2.55-3.76)	(2.55-4.68)	(2.55-5.89)	(2.55-4.72)	(2.55-4.36)

Table 2.A5 Nutrient summary results by watershed

Watershed	TFP (µg/L) Average (Range)	FRP (µg/L) Average (Range)	TRP (µg/L) Average (Range)	TP (µg/L) Average (Range)	NO₃ (mg/L) Average (Range)	NH₄ (mg/L) Average (Range)	TDN (mg/L) Average (Range)	SO₄ (mg/L) Average (Range)
River Raisin (RR)	39.06 (11.39-315.28)	29.16 (1.66-291.47)	39.80 (7.06-303.04)	58.72 (14.30-284.14)	2.34 (0.06-8.06)	0.02 (0-0.17)	2.67 (0.61-8.22)	50.24 (16.26-100.22)
Kawkawlin (KAW)	44.03 (16.38-125.97)	33.20 (5.52-122.00)	39.14 (15.93-139.36)	57.98 (18.85-154.62)	1.80 (0-5.95)	0.02 (0-0.08)	2.38 (0.74-6.14)	21.02 (2.64-50.44)
Macatawa (MAC)	83.60 (25.57-260.81)	66.64 (8.99-224.99)	75.04 (24.42-235.40)	111.24 (36.49-354.48)	3.78 (0.03-11.58)	0.10 (0-1.56)	4.34 (1.00-13.34)	37.41 (11.95-64.15)
Sandy Creek (SC)	48.39 (13.89-151.59)	38.76 (0.89-140.90)	40.51 (15.55-165.98)	58.75 (26.34-178.48)	4.48 (0.01-9.59)	0.01 (0-0.11)	4.90 (0.80-10.16)	58.72 (25.95-153.92)
Little Pigeon (LPR)	10.83 (4.34-19.27)	9.37 (5.13-12.85)	9.76 (5.99-14.00)	14.77 (10.56-21.37)	0.20 (0.14-0.37)	0.03 (0-0.08)	0.66 (0.30-1.05)	37.43 (27.65-45.47)

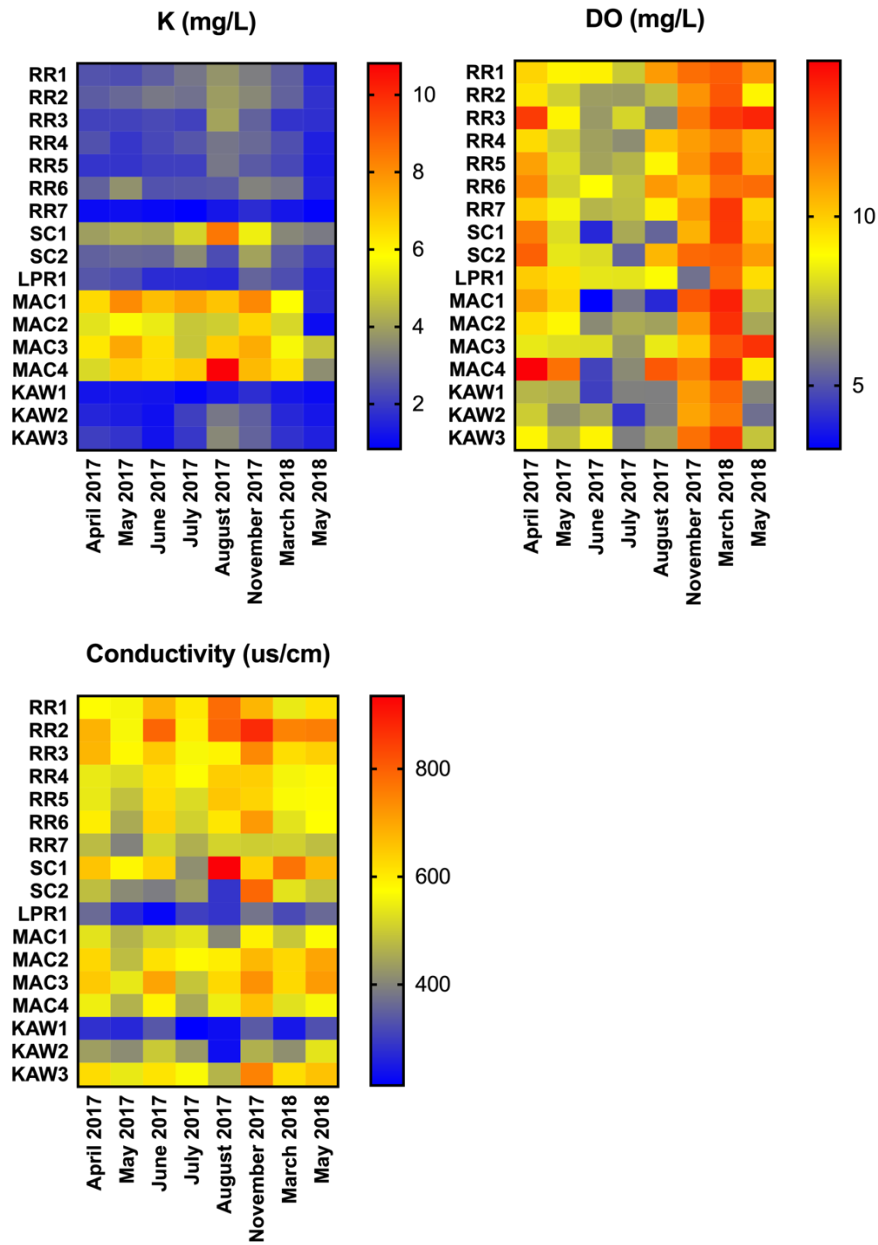


Figure 2.A1 Water quality variable heatmaps showing spatial and temporal distributions. a) potassium (k) mg/L; b) dissolved oxygen (D) mg/L; c) Conductivity (us/cm).

REFERENCES

REFERENCES

- Ahmed, W., Gyawali, P., Feng, S., and McLellan, S.L. (2019). Host specificity and sensitivity of the established and novel sewage-associated marker genes in human and non-human fecal samples. *Appl. Environ. Microbiol.* 85(14): e00641-19; doi: 10.1128/AEM.00641-19
- Anderson, J., Hardy, E., Roach, J., and Witmer, R. (1976). A land use and land cover classification system for use with remote sensor data. USGS Pub. US Gov. Print. Office, Washington, DC.
- Clesceri L.S., Greenberg A.E., Eaton A.D., eds (1998). *Standard Methods for the Examination of Water and Wastewater* (United Book, Baltimore), 20th Ed.
- Badgley, B.D., Steele, M.K., Cappellin, C., Burger, J., Jian, J., Neher, T.P., Orentas, M., and Wagner, R. (2019). Fecal indicator dynamics at the watershed scale: Variable relationships with land use, season, and water chemistry. *Sci Total Environ* 697, 134113. <https://doi.org/10.1016/j.scitotenv.2019.134113>
- Ballesté, E., García-Aljaro, C., and Blanch, A.R. (2018). Assessment of the decay rates of microbial source tracking molecular markers and faecal indicator bacteria from different sources. *J Appl Microbiol* 125, 1938–1949. <https://doi.org/10.1111/jam.14058>
- Baker, N.T. (2011). Tillage practices in the conterminous United States, 1989–2004—Datasets Aggregated by Watershed: U.S. Geological Survey Data Series 573, 13 p.
- Boehm, A.B., Werfhorst, L.C., Van De, Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Dan, W., and Weisberg, S.B. (2013). Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. *Water Res.* 47 (18), 6812e6828.
- Bonsch, M., Popp, A., Biewald, A., Rolinski, S., Schmitz, C., Weindl, I., Stevanovic, M., Högner, K., Heinke, J., Ostberg, S., Dietrich, J.P., Bodirsky, B., Lotze-Campen, H., and Humpenöder, F. (2015). Environmental flow provision: Implications for agricultural water and land-use at the global scale. *Global Environ Change* 30, 113–132. <https://doi.org/10.1016/j.gloenvcha.2014.10.015>
- Brooks, Y., Aslan, A., Tamrakar, S., Murali, B., Mitchell, J., and Rose, J.B. (2015). Analysis of the persistence of enteric markers in sewage polluted water on a solid matrix and in liquid suspension. *Water Res* 76, 201–212. <https://doi.org/10.1016/j.watres.2015.02.039>
- Bullerjahn, G.S., McKay, T.W., Davis, D.B., Baker, G.L., Boyer, L.V., D'Anglada, G.J., Doucette, J.C., Ho, E.G., Irwin, C.L., Kling, R.M., Kudela, R., Kurmayer, A.M., Michalak, J.D., Ortiz, T.G., Otten, H.W., Paerl, B., Qin, B.L., Sohngen, R.P., Stumpf, P.M., Visser, and Wilhelm, S.W. (2016). Global solutions to regional problems: Collecting global expertise to address the

problem of harmful cyanobacterial blooms: A Lake Erie case study. *Harmful Algae* 54:223–238. doi:10.1016/j.hal.2016.01.003

Crumpton W., Isenhardt T., and Mitchell, P. (1992) Nitrate and organic N analyses with second-derivative spectroscopy. *Limnol Oceanogr* 37(4):907–913.

EGLE, (2020). National pollutant discharge elimination system wastewater discharge general permit: concentrated animal feeding operations. State of Michigan Department of Environment, Great Lakes, and Energy (EGLE). https://www.michigan.gov/documents/egle/egle-wrd-cafoggp-2020_674031_7.pdf

Floyd, W.C., Schoenholtz, S.H., Griffith, S.M., Wigington Jr., P.J., and Steiner, J.J. (2009). Nitrate-nitrogen, land use/land cover, and soil drainage associations at multiple spatial scales. *J. Environ. Qual.*, 38 pp. 1473-1482

Hamilton S.K., Bruesewitz D.A., Horst G.P., Weed D.B., and Sarnelle O. (2009) Biogenic calcite–phosphorus precipitation as a negative feedback to lake eutrophication. *Can J Fish Aquat Sci* 66(2):343–350.

Harwood, V.J., Christopher, S., Badgley, B.D., Kim, B., and Asja, K. (2014). Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiol. Rev.* 38 (1), 1e40.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. *Elements of statistical learning* (2nd ed.). New York, New York, USA: Springer US.

He, X., Liu, P., Zheng, G., Chen, H., Shi, W., Cui, Y., Ren, H., and Zhang, X.-X. (2016). Evaluation of five microbial and four mitochondrial DNA markers for tracking human and pig fecal pollution in freshwater. *Sci Rep-uk* 6, 35311. <https://doi.org/10.1038/srep35311>

Holden, S. (2014). Monthly Water Quality Assessment of Lake Macatawa and Its Tributaries, April–September 2012. Michigan Department of Environmental Quality, Water Resources Division; Lansing, MI, USA.

Jamieson, R., Gordon, R., Joy, D., and Lee, H. (2004). Assessing microbial pollution of rural surface waters: A review of current watershed scale modeling approaches. *Agricultural Water Management*. 70:1. 1-17.

Jarrett, R.D. (1991) Wading measurements of vertical velocity profiles. *Geomorphology* 4(3-4):243–247.

Korajkic, A., McMinn, B.R., Ashbolt, N.J., Sivaganesan, M., Harwood, V.J., and Shanks, O.C. (2018). Extended persistence of general and cattle-associated fecal indicators in marine and freshwater environment. *Sci Total Environ* 650, 1292–1302. <https://doi.org/10.1016/j.scitotenv.2018.09.108>

- Lee, D.-Y., Lee, H., Trevors, J.T., Weir, S.C., Thomas, J.L., and Habash, M. (2014). Characterization of sources and loadings of fecal pollutants using microbial source tracking assays in urban and rural areas of the Grand River Watershed, Southwestern Ontario. *Water Res* 53, 123–131. <https://doi.org/10.1016/j.watres.2014.01.003>
- Legendre, P. and Legendre, L., (2012). Chapter 12 Ecological data series. *Dev Environ Model* 24, 711–783. <https://doi.org/10.1016/b978-0-444-53868-0.50012-5>
- Lorimor, J., Powers, W., and Sutton, A. (1980). *Manure Management Systems Series: Manure Characteristics*. Midwest Plan Service: MWPS. 2nd Ed.
https://www.canr.msu.edu/uploads/files/ManureCharacteristicsMWPS-18_1.pdf
- Luszcz, E.C., Kendall, A.D., and Hyndman, D.W. (2017). A spatially explicit statistical model to quantify nutrient sources, pathways, and delivery at the regional scale. *Biogeochemistry*. 133: 37-57.
- Luszcz, E.C., Kendall, A.D., and Hyndman, D.W. (2015). High resolution spatially explicit nutrient source models for the lower peninsula of Michigan. *Journal of Great Lakes Research*. 41. 618-629.
- Mateo-Sagasta, J., Turrall, H., and Burke, J. (2018). Global drivers of water pollution from agriculture. Chapter 2 In *More People, More Food, Worse Water? A Global Review of Water Pollution from Agriculture*. Edited by Mateo-Sagasta J, Marjani S, Turrall H. Rome: FAO and IWMI.
- Martin, S., A. Kendall, Q. Hamlin, L. Wan and Hyndman, D.W. (2021), The Land Use Legacy Effect: Looking back to see a path forward to improve management, *Environmental Research Letters*, DOI: 10.1088/1748-9326/abe14c
- Mayer, R.E., Reischer, G., Ixenmaier, S.K., Derx, J., Blaschke, A.P., Ebdon, J.E., Linke, R., Egle, L., Ahmed, W., and Blanch, A. (2018). Global distribution of human-associated fecal genetic markers in reference samples from six continents. *Environ. Sci. Technol.* 52 (9), 5076e5084.
- Michigan Land Use Leadership Council. (2003). *Michigan’s land, Michigan’s future: final report of the Michigan land use leadership council*.
<https://publicsectorconsultants.com/2003/08/01/michigans-land-michigans-future/>
- McKee, B.A., Molina, M., Cyterski, M., and Couch, A. (2020). Microbial source tracking (MST) in Chattahoochee River National Recreation Area: Seasonal and precipitation trends in MST marker concentrations, and associations with *E. coli* levels, pathogenic marker presence, and land use. *Water Res* 171, 115435. <https://doi.org/10.1016/j.watres.2019.115435>
- Mieszkina, S., Furet, J.-P., Corthier, G.R., and Gourmelon, M.L. (2009). Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific *Bacteroidales* 16S rRNA genetic markers. *Applied and Environmental Microbiology*, 75, 3045-54.

Michigan Department of Transportation (MDOT) (2006). Michigan Land Use Report White Paper.
https://www.michigan.gov/documents/mdot/MDOT_LandUseWhitePaperFinal_397586_7.pdf

Moriasi, D.N., Gitau, M.W., and Daggupati, N. Pai. (2015). Hydrologic and water quality models: performance measures and evaluation criteria. Transactions of the ASABE. 58(6): 1763-1785.

MWP (Macatawa Watershed Project). (2012). *Macatawa Watershed Management Plan*. Macatawa Area Coordinating Council; Holland, MI, USA.
https://www.michigan.gov/documents/deq/wrd-nps-wmp-macatawa-3-6_425599_7.pdf

Nshimiyimana J.P., Martin, S.L., Flood, M., Verhougstraete, M.P., Hyndman, D.W., and Rose, J.B. (2018). Regional variations of bovine and porcine fecal pollution as a function of landscape, nutrient, and hydrological factors. Journal of Environment Quality. 47:1024-1032.

NOAA (2019). Data Tools | Climate Data Online (CDO) | National Climatic Data Center (NCDC). Accessed June 11, 2019. <https://www.ncdc.noaa.gov/cdo-web/datatools/>.

OECD (2012). Water quality and agriculture: meeting the policy challenge. Key Messages and Executive Summary. OECD studies on water. <http://dx.doi.org/10.1787/9789264168060-en>.

OECD (2017). Policy Highlights Diffuse Pollution, Degraded Waters: Emerging Policy Solutions. OECD Environment Directorate.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>.

Robertson, D.M. and Saad, D.A. (2011). Nutrient inputs to the Laurentian Great Lakes by source and watershed estimated using SPARROW watershed models. JAWRA J Am Water Resour Assoc 47(5):1011–1033

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Santo Domingo, J., Seifring, S., Sivaganesan, M., and Haugland, R.A. Quantitative PCR for Detection and Enumeration of Genetic Markers of Bovine Fecal Pollution. Applied and Environmental Microbiology, 74:3. p. 745-752.

Sharpley A.N., Bergström L., Aronsson H., Bechmann M., Bolster C.H., Börling K., Djodjic F., Jarvie H.P., Schoumans O.F., Stamm C., Tonderski K.S., Ulén B., Uusitalo R., Withers P.J.A. (2015) Future agriculture with minimized phosphorus losses to waters: research needs and direction. Ambio. 44:163-179. <http://dx.doi.org/10.1007/s13280-014-0612-x>.

Smith R.A., Schwarz G.E., Alexander R.B. (1997). Regional interpretation of water-quality monitoring data. Water Resour Res. 33(12):2781–2798

Smith, D.R., K.W. King, and M.R. Williams (2015). What is causing the harmful algal blooms in Lake Erie? *J. Soil Water Conserv.* 70:27A–29A. doi:10.2489/jswc.70.2.27A

Strayer, D.L., Beighley, R.E., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., and Naiman R.J. (2003) Effects of land cover on stream ecosystems: roles of empirical models and scaling issues. *Ecosystems*. 6, 407–423

Sowah, R.A., Habteselassie, M.Y., Radcliffe, D.E., Bauske, E., Risse, M., 2017. Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States. *Water Res* 108, 330–338. <https://doi.org/10.1016/j.watres.2016.11.007>

USEPA, (2014). Method C: *Escherichia coli* in Water by TaqMan Quantitative Polymerase Chain Reaction (qPCR) Assay.

USEPA, (2012). Recreational Water Quality Criteria (Washington, DC).

USDA (2019). 2017 Census of Agriculture United States. United States Department of Agriculture. https://www.nass.usda.gov/Publications/AgCensus/2017/index.php#full_report

USDA – NASS (United States Department of Agriculture National Agricultural Statistics Service) (2017). “CropScape—Cropland data layer.” Accessed June 11, 2019. <https://nassgeodata.gmu.edu/CropScape/>

Verhougstraete, M. (2012). Measuring microbial water quality responses to land and climate using fecal indicator bacteria and molecular source tracking in rivers and nearshore surface waters of Michigan. Ph.D. diss., Michigan State University.

Verhougstraete, M.P., S.L. Martin, A.D. Kendall, D.W. Hyndman, and J.B. Rose. 2015. Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proc. Natl. Acad. Sci. USA* 112:10419–10424. doi:10.1073/pnas.1415836112

Vermeulen, L.C., J.D. Kraker, N. Hofstra, C. Kroeze, and G. Medema (2015). Modelling the impact of sanitation, population growth and urbanization on human emissions of *Cryptosporidium* to surface waters—a case study for Bangladesh and India. *Environ. Res. Lett.* 10:094017. doi:10.1088/1748-9326/10/9/094017

Walterhouse, M. (1999). *Total Maximum Daily Load for Phosphorus in Lake Macatawa*. MDEQ Submittal to U.S. Environmental Protection Agency; Lansing, MI, USA. https://www.michigan.gov/documents/deq/wrd-swas-tmdl-macatawa_451047_7.pdf

Wen Y., Schoups G., van de Giesen N. (2017). Organic pollution of rivers: combined threats of urbanization, livestock farming and global climate change. *Sci Rep.* 7:43289. <http://doi.org/10.1038/srep43289>.

Wetzel R, Likens G (2000) *Limnological Analyses* (Springer, New York), 3rd Ed.

Yampara-Iquise, H., Zheng, G., Jones, J.E. and Carson, C.A. (2008) Use of a *Bacteroides* thetaiotaomicron-specific Alpha-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *J Appl Microbiol* 105, 1686–1693.

Zandaryaa, S. and Mateo-Sagasta, J. (2018). Organic matter, pathogens and emerging pollutants. In *More people, more food, worse water?: A global review of water pollution from agriculture*. Edited by Mateo-Sagasta J., Marjani S., and Turrall H. Rome: FAO and IWMI.

Zhang, H. (2017). Fertilizer Nutrients in Animal Manure. Oklahoma State Extension.
<https://extension.okstate.edu/fact-sheets/fertilizer-nutrients-in-animal-manure.html>

Zhang, Y., Wu, R., Zhang, Y., Wang, G., and Li, K. (2018). Impact of Nutrient Addition on Diversity and Fate of Fecal Bacteria. *Sci. Total Environ.* 636, 717e7

3.0 Methods evaluation for rapid concentration and quantification of SARS-CoV-2 in raw wastewater using droplet digital and quantitative RT-PCR

Work presented in this chapter has been published as Flood, M.T.^a, D'Souza, N.^a, Rose, J.B.^a, and Aw, T.G.^b (2021). Methods Evaluation for Rapid Concentration and Quantification of SARS-CoV-2 in Raw Wastewater Using Droplet Digital and Quantitative RT-PCR. Food Environ Virol 13, 303–315. <https://doi.org/10.1007/s12560-021-09488-8>

^a Department of Fisheries and Wildlife, Michigan State University, East Lansing Michigan 48824, USA

^b Department of Environmental Health Sciences, School of Public Health and Tropical Medicine, Tulane University, New Orleans, Louisiana 70112, USA

3.1 Abstract

Wastewater surveillance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an emerging public health tool to understand the spread of Coronavirus Disease 2019 (COVID-19) in communities. The performance of different virus concentration methods and PCR methods needs to be evaluated to ascertain their suitability for use in the detection of SARS-CoV-2 in wastewater. We evaluated ultrafiltration and polyethylene glycol (PEG) precipitation methods to concentrate SARS-CoV-2 from sewage in wastewater treatment plants and upstream in the wastewater network (e.g., manholes, lift stations). Recovery of viruses by different concentration methods was determined using Phi6 bacteriophage as a surrogate for enveloped viruses. Additionally, the presence of SARS-CoV-2 in all wastewater samples was determined using reverse transcription quantitative PCR (RT-qPCR) and reverse transcription droplet digital PCR (RT-ddPCR), targeting three genetic markers (N1, N2 and E). Using spiked samples, the Phi6 recoveries were estimated at 2.6-11.6% using ultrafiltration-based methods and 22.2-51.5% using PEG precipitation. There was no significant difference in recovery efficiencies ($p < 0.05$) between the PEG procedure with and without a 16 hr overnight incubation, demonstrating the feasibility of obtaining same day results. The SARS-CoV-2 genetic markers were more often detected by RT-ddPCR than RT-qPCR with higher sensitivity and precision. While all three SARS-CoV-2 genetic markers were detected using RT-ddPCR, the levels of E gene were almost below the limit of detection using RT-qPCR. Collectively, our study suggested PEG precipitation is an effective low-cost procedure which allows a large number of samples to be processed simultaneously in a routine wastewater monitoring for SARS-CoV-2. RT-ddPCR can be implemented for the absolute quantification of SARS-CoV-2 genetic markers in different wastewater matrices.

3.2 Introduction

Since the emergence and spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (COVID-19), many cities around the world have rapidly expanded their viral surveillance systems, including wastewater monitoring for SARS-CoV-2. This is because SARS-CoV-2 can be shed in the feces of infected individuals from both symptomatic and asymptomatic cases (Park et al., 2020; Wu et al., 2020).

Coronaviruses are positive-strand RNA enveloped viruses with the largest viral genomes of all RNA viruses (27 to 32 kb). They have a spherical virion of about 120 nm in diameter surrounded by a lipid envelope with pronounced spiked glycoproteins (S) embedded. The vast majority of studies on the presence of viruses in human excreta and municipal wastewater have been focused on nonenveloped enteric viruses. There are a number of established methods for the detection of nonenveloped enteric viruses in wastewater, but only fewer evaluated protocols for human enveloped viruses such as SARS-CoV-2 (Haramoto et al., 2018). Analysis of environmental matrices for human viruses often require concentration steps due to the low ambient concentrations of the viruses. Therefore, laboratory methods for the detection of SARS-CoV-2 in wastewater need to examine both sample concentration and RNA quantification methods along with optimizing limits of detection.

Globally, there have been over forty reports on the molecular detection of SARS-CoV-2 in wastewater (e.g., Ahmed et al., 2020a; Ahmed et al., 2020b; Ampeuro et al., 2020; Arora et al., 2020; Balboa et al., 2020; Chavarria-Miró et al., 2020; Curtis et al., 2020; Döhla et al., 2020; Fernández de Mera et al., 2020; Fongaro et al., 2020; Green et al., 2020; Haramoto et al., 2020; Hata et al., 2020; Kocamemi et al., 2020a; Kocamemi et al., 2020b; Kumar et al., 2020; La Rosa et al., 2020a; La Rosa et al., 2020b; Medema et al., 2020; Miyani et al., 2020; Nemudryi et al.,

2020; Bar-Or et al., 2020; Peccia et al., 2020; Prado et al., 2020; Randazzo et al., 2020a; Randazzo et al., 2020b; Rimoldi et al., 2020; Sharif, 2020; Sherchan et al., 2020; Trottier et al., 2020; Vallejo et al., 2020; Wang et al., 2020; Weidhaas et al., 2020; Westhaus et al., 2020; Wu et al., 2020a; Wu et al., 2020b; Wurtzer et al., 2020; Zhang et al., 2020a; Zhang et al., 2020b; Zhou et al., 2020). These studies have had large variability in the numbers of samples from as few as 10 samples collected to over 120 with SARS-CoV-2 RNA being detected at concentrations ranging from 10^2 to 10^6 copies per liter. These SARS-CoV-2 surveillance studies analyzed volumes of raw sewage, treated wastewater and sewage sludge ranging from 2.5 mL to 2000 mL, using various concentration methods such as adsorption-elution based membrane filtration, precipitation (using polyethylene glycol, aluminum hydroxide), ultracentrifugation and ultrafiltration prior to RNA extraction in order to recover the virus. The majority of studies quantified the viral RNA in wastewater using quantitative reverse transcription polymerase chain reaction (RT-qPCR) with external standard curves. Several gene targets specific to the SARS-CoV-2 have been used in wastewater surveillance, including the RNA-dependent polymerase (RdRP), nucleocapsid (N1, N2), envelope protein (E), spike glycoprotein (S), membrane glycoprotein (M) and *ORF1ab* genes (e.g., Lu et al., 2020; Corman et al., 2020).

Currently, cell culture for SARS-CoV-2 requires a Biosafety Level 3 laboratory and specially trained personnel. Therefore, surrogate viruses have been used to mimic SARS-CoV-2 to evaluate virus concentration methods for wastewater. These surrogate viruses include F-specific RNA phages (Balboa et al., 2020; Hata et al., 2020; Medema et al., 2020), mengovirus (Randazzo et al., 2020a), avian coronavirus of infectious bronchitis virus (Kocamemi et al., 2020a), Alphacoronavirus HCoV 229E (La Rosa et al., 2020b), bovine coronavirus BCoV (LaTurner et al., 2021), porcine epidemic diarrhea virus (PEDV) (Randazzo et al., 2020b),

bovine respiratory syncytial virus (BRSV) (Gonzalez et al., 2020), and murine hepatitis virus (Ahmed et al., 2020c). Estimated mean recovery efficiencies for these surrogate viruses ranged from 1% to 73% using different concentration methods originally developed for the detection of enteric viruses in environmental samples (Randazzo et al., 2020a; Medema et al., 2020).

Pseudomonas phage Phi6 has also been used as a model enveloped virus in recovery and persistence studies (Aquino de Carvalho et al., 2017; Ye et al., 2016). Similar to coronaviruses, Phi6 is an enveloped RNA virus, with a segmented genome and glycerophospholipids in its envelope (Vidaver et al., 1973). Since Phi6 is not pathogenic to humans, it is easier to work with than other enveloped animal viruses and no special laboratory biosafety is required.

Rapid, cost-effective, and efficient methods are needed to provide precise data to support public health decision making. This is so that changes in concentrations of SARS-CoV-2 gene markers in wastewater provide meaningful data to inform COVID-19 surveillance. Therefore, the objective of this study was to (i) evaluate the efficiencies of polyethylene glycol (PEG) precipitation and ultrafiltration methods to recover *Pseudomonas* phage Phi6, coronavirus OC43, and SARS-CoV-2 from different wastewater matrices; (ii) compare two PCR-based methods, reverse transcription quantitative PCR (RT-qPCR) and reverse transcription droplet digital PCR (RT-ddPCR) for the detection of SARS-CoV-2 in different wastewater matrices; and (iii) develop a rapid, cost-effective, and precise quantification workflow for SARS-CoV-2 in wastewater.

3.3 Materials and Methods

3.3.1 Wastewater samples and sampling sites

Wastewater samples (500-1000 mL) for this study were collected from 11 sanitary sewer sites and four wastewater treatment plant (WWTP) influent streams (after grit removal) (Supplemental materials Table A1 and A2). A total of twenty sanitary sewer samples were collected as grab samples from the 11 manholes or lift stations. Sanitary sewer samples consisted of wastewater flowing from university dormitories, local communities, and hospital. Influent samples (n=11) from four WWTPs were collected as 24-hr composite samples. Samples used for the comparison of the SARS-CoV-2 surrogates Phi6 and human coronavirus OC43 were collected from two California wastewater treatment plant influents as previously described by Pecson et al. (2021). All samples were kept at 4°C for up to 72 hours. If samples were unable to be processed within 72 hours of collection, then they were frozen at -80°C until analysis.

3.3.2 Virus stocks

Bacteriophage Phi6 and its bacterial host *Pseudomonas syringae* were kindly provided by Dr. Krista Wigginton's lab at University of Michigan. To propagate Phi6, *P. syringae* was grown in King's B medium at 24°C for 6 hours in stationary culture. Phi6 was added to the host and incubated under the same conditions for 16 to 18 hours. After incubation and observed clearing of cell suspension due to lysis, cells and debris were removed from the Phi6 suspension by filtration using 0.22 µm membranes. The Phi6 stocks were stored at 4°C and titered using an overlay method. For the overlay process, 2 ml of host was added to the overlay tube containing King's B agar and 0.5 ml of virus suspension, mixed, and poured onto a plate containing King's

B agar. Plates were incubated at 24°C for 16-24 hours and plaques were counted. Virus titers of approximately 10^9 plaque forming unit (PFU) per ml were routinely obtained.

3.3.3 Virus concentration methods and experiments

Four distinct comparisons were performed in this study. First, three viral concentration methods were tested for their efficiency in recovering Phi6 phages and SARS-CoV-2 in different types of wastewater. Methods 1 (CEN1) and 2 (CEN2) are based on the ultrafiltration principle and used centrifugal filters. Method 3 is a precipitation using polyethylene glycol (PEG). The second comparison was between RT-ddPCR and qPCR using the three viral concentration methods. The third comparison was determining if a rapid PEG precipitation approach (without an overnight incubation) would be able to perform as well or better than PEG precipitation with a 16 hr overnight incubation. Lastly, Phi6 was compared against the human coronavirus OC43 using RT-ddPCR to determine if recovery efficiencies between the two SARS-CoV-2 surrogates were equivalent.

For each experiment, 350 ml of wastewater sample was inoculated with 1 ml of 10^6 plaque forming units (PFU)/ml of Phi6 and homogenized for 10 minutes at 4°C. SARS-CoV-2 was not added to the sample. After homogenization, the sample was subdivided into three 101 ml of aliquots in 250 ml centrifuge bottles for processing with each concentration method. One milliliter of sample was removed from each 250 ml bottle containing the subsample for use in determining the seeded virus level for recovery efficiency of each method. Recovery efficiencies were determined by comparing the concentration of the spiked Phi6 bacteriophage in each subsample prior to processing with the concentration of Phi6 in their final concentrate using RT-

ddPCR. All viral concentration experiments, for each method and each type of wastewater, were conducted in triplicate.

Method 1 (CEN1) was adapted from Ye et al. (2016) but modified to include virus recovery steps from wastewater solids. Briefly, 100 ml of wastewater sample was first centrifuged at 2,500 x g for 5 min at 4°C in order to pellet any solids present in the sample. The supernatant was then collected without disturbing the pellet and filtered through a 0.22 µm polyethersulfone (PES) membrane filter (MilliporeSigma, St. Louis, MO). The sample was then concentrated using a 10 kDa Centricon Plus-70 centrifugal filter unit (MilliporeSigma, St. Louis, MO) according to the manufacturer's protocol. A 1:1 volume of 0.25N glycine buffer was added to the pellet and remaining liquid. The pellet was vortexed every 10 min for 30 min while on ice to dislodge the viruses from suspended solids. After the 30 min incubation the glycine-processed sample was neutralized 1:1 with 2 x PBS. The sample was then centrifuged at 10,000 x g for 30 min at 4°C. The supernatant was processed with the same centrifugal filter and the resulting concentrates were combined.

Method 2 (CEN2) involved the use of the same centrifugal filter but without a pre-filtration step (Medema et al., 2020). In this method, 100 ml of sample was centrifuged at 4,654 x g for 30 min at 4°C without brake. The supernatant was then collected and directly filtered through a 10 kDa Centricon Plus-70 centrifugal filter unit (MilliporeSigma, St. Louis, MO) according to the manufacturer's protocol. The pellet was processed using the same protocol as described in the Method 1 (CEN1).

Method 3 (PEG) was adapted from Borchardt et al. (2017) for the detection of avian influenza virus RNA in groundwater. The samples were mixed with 8% (w/vol) molecular biology grade PEG 8000 (Promega Corporation, Madison WI) and 0.2 M NaCl (w/v). The

samples were mixed slowly on magnetic stirrer at 4°C for 2 hours and then held at 4°C for 16 hours. Following the overnight incubation, samples were centrifuged at 4,700 x g for 45 mins at 4°C. The supernatant was then removed, and the pellet resuspended in the remaining liquid. All sample concentrates were aliquoted and stored at -80°C until further processing.

After the initial comparison of two ultrafiltration methods and PEG precipitation, a rapid PEG precipitation approach (without an overnight incubation) was evaluated with 19 wastewater samples. Each sample was inoculated with Phi6 and homogenized as described above. After mixing the sample with 8% (w/vol) PEG 8000 and 0.2 M NaCl for 2 hours at 4°C, the sample was immediately centrifuged at 4,700 x g for 45 mins at 4°C.

Finally, a comparison between Phi6 and OC43 was performed using wastewater from two California wastewater treatment plants split into 5 subsamples each and processed with the overnight PEG precipitation method.

3.3.4 RNA extraction and quantification by RT-ddPCR and RT-qPCR

Viral ribonucleic acid (RNA) was extracted from wastewater concentrates using the Qiagen QIAmp Viral RNA Minikit according to the manufacturer's protocol with modifications (Qiagen, Germany). In this study, a total of 200 µl of concentrate was used for RNA extraction resulting in a final elution volume of 80 µl. Extracted RNA was stored at -80°C until analysis.

3.3.4.1 Detection of SARS-CoV-2, Phi6, and coronavirus OC43 using RT-ddPCR

One-step RT-ddPCR approach was used to quantify the Phi6 RNA to determine the recovery efficiencies for each concentration method. All the primers and probes used in this study are listed in Table A3. Droplet digital PCR was performed using Bio-Rad's 1-Step RT-

ddPCR Advanced kit with a QX200 ddPCR system (Bio-Rad, CA, USA). Each reaction contained a final concentration of $1 \times$ Supermix (Bio-Rad, CA, USA), 20 U μl^{-1} reverse transcriptase (RT) (Bio-Rad, CA, USA), 15 mM DTT, 900 nmol l^{-1} of each primer, 250 nmol l^{-1} of each probe, 1 μl of molecular grade RNase-free water, and 5.5 μl of template RNA for a final reaction volume of 22 μl . Droplet generation was performed by microfluidic mixing of 20 μl of each reaction mixture with 70 μl of droplet generation oil in a droplet generator (Bio-Rad, CA, USA) resulting in a final volume of 40 μl of reaction mixture-oil emulsions containing up to 20,000 droplets with a minimum droplet count of $> 9,000$. The resulting droplets were then transferred to a 96-well PCR plate which was heat-sealed with foil and placed into a C1000 96-deep well thermocycler (Bio-Rad, CA, USA) for PCR amplification using the following parameters: 25°C for 3 min, 50°C for 1 hr, 95°C for 10 min, followed by 40 cycles of 95°C for 30 s and 60°C for 1 min with ramp rate of 2°C s^{-1} followed by a final cycle of 98°C for 10 min. Following PCR thermocycling, each 96-well plate was transferred to a QX200 Droplet Reader (Bio-Rad, CA, USA) for the concentration determination through the detection of positive droplets containing each gene target by spectrophotometric detection of the fluorescent probe signal.

SARS-CoV-2 RNA and OC43 in wastewater samples were also quantified using the same one-step RT-ddPCR approach except the annealing temperature was set at 55°C. Three SARS-CoV-2 markers were chosen for analysis, the nucleocapsid 1 (N1) and nucleocapsid 2 (N2) gene targets designed by the US Centers for Disease Control and Prevention (CDC) (Lu et al., 2020), the envelope (E) gene from Corman et al. (2020), and OC43 (Table A3). The N1 and N2 gene targets were analyzed in a duplex assay. All analyses were run in triplicate for each

marker. Quality controls were run with every plate including positive and non-template controls, extraction controls, and processing blanks for each batch of samples.

3.3.4.2 Detection of SARS-CoV-2 using RT-qPCR

RT-qPCR approach was also used to quantify SARS-CoV-2 gene markers in wastewater samples. All RT-qPCR reactions were performed using a StepOne Plus™ real-time PCR sequence detector (Applied Biosystems, Foster City, CA). For each assay, a 10-fold diluted standard curve of at least five points, a non-template control, and samples were tested in triplicate. The quantitative synthetic SARS-CoV-2 RNA includes fragments from nucleocapsid and envelope regions (ATCC VR-3276SD) was used to generate standard curves. Amplification reaction mixtures (final total volume of 20 µl) contained 5 µl template RNA, 10 µl of 2 × qScript one-step RT-qPCR ToughMix (QuantaBio), 300 nM, 500 nM and 400 nM of forward primer for N1, N2 and E gene, respectively, 500 nM, 800 nM and 800 nM of reverse primer for N1, N2 and E gene, respectively, and 200 nM of probe. The thermal cycling protocol was as follows: 10 min at 50°C for cDNA synthesis, 3 min at 95°C for initial denaturation, followed by 45 cycles of two steps consisting of 3 s at 95°C and 30 s at 55°C. qPCR amplification efficiencies for the quantification of the N1, N2 and E gene assays were 92.6±4.3%, 95.1±3.4% and 91.6±2.2%, respectively, and the correlation coefficients (R^2) of the standard curves were 0.968±0.002, 0.982± 0.004, and 0.988± 0.0006, respectively.

3.3.5 Data analysis

All SARS-CoV-2, Phi6, and OC43 gene data were converted from gene copies (GC) per reaction to GC per 100 ml before analysis. Non-detects (ND) were assigned their individual

sample's limit of detection. The limit of detection was calculated for each individual sample based on both the molecular assays' theoretical detection limits (i.e., 3 positive droplets for RT-ddPCR; the lowest standard curve concentration for RT-qPCR) and the concentration factor of each processing method examined.

$$\text{Virus GC per 100ml} = \frac{\frac{\text{GC per reaction}}{V_r} \times V_e \times \frac{V_f}{V_c}}{V_i} \times 100$$

Where:

V_i = Initial volume of sample concentration in ml

V_f = Final volume of sample after concentration in ml

V_r = Volume of RNA template used per PCR reaction in μl

V_e = Final volume of RNA eluted from RNA extraction in μl

V_c = Volume of concentrated sample used for RNA extraction in ml

Recovery efficiency was calculated by dividing the total gene copies (GC) / 100 ml concentration of the Phi6 bacteriophage measured in each methods' final concentrate by the concentration (GC/ 100 ml) of Phi6 in each sample before concentration and then multiplying by 100.

Statistics and data visualization were performed using Graphpad Prism 8 (Graphpad Software, CA, USA). Results for the three methods comparison were analyzed with a two-way ANOVA with a Tukey's multiple comparisons test to determine method significance (p value < 0.05). A two-way ANOVA (p < 0.05) and a paired t test (p < 0.05) were performed for the comparison of "normal" (16 hr hold) vs "rapid" (no hold) PEG precipitation methods.

3.4 Results

3.4.1 Wastewater characteristics

Wastewater samples from both sanitary sewer systems and treatment plants were evaluated in this study. All site-specific details including physiochemical data and sampling dates for each sanitary sewer and WWTP site are shown in Table A1 and Table A2, respectively. Wastewater collected from sanitary sewer locations had more variations in each parameter than wastewater collected from WWTP. For example, while sanitary sewer sites showed a wide range of turbidities ranging from 1.87 up to 191 NTU, WWTP influent samples showed less variation (e.g., 20.2 to 111 NTU). Sanitary sewer sites showed little variation in pH and temperature with each ranging from 6.57-8.58 and 13-26.4°C, respectively (Table A1). Influent samples collected from WWTPs had a smaller degree of variation in pH (7.33-7.8) than sanitary sewer sites but had greater variation in temperatures which ranged from 1.40 to 21.67°C (Table A2). Total suspended solids (TSS) and daily flows for each WWTP were also measured. Specifically, samples collected from facility W had the largest range of TSS (48-920 mg L⁻¹) and the highest daily flows ranging from 14.6-27.6 million gallons per day (mgd). Facility E had the smallest range of TSS (164-208 mg L⁻¹) and the lowest daily flow of 2.87 mgd, but facility M had the smallest range of daily flows (3.24-3.86 mgd).

3.4.2 Recovery of Phi6 from wastewater samples using ultrafiltration and PEG methods

Prior to seeding experiments, ambient concentrations of *Pseudomonas* phage Phi6 were determined using RT-ddPCR. All wastewater samples were negative for Phi6.

The mean recovery efficiencies of the two ultrafiltration-based and PEG precipitation methods for the detection of Phi6 using RT-ddPCR in different types of wastewater are

summarized in Table 3.1. For the various wastewater matrices, mean recoveries of ultrafiltration-based Method 1 ranged from 2.6% to 10.6% and Method 2 ranged from 2.7% to 11.6%. The Phi6 virus recovery was statistically higher ($p < 0.0001$) for both sanitary sewers and WWTP influent samples using the PEG method compared to the ultrafiltration methods, with mean recoveries ranging from 22.19% to 51.47% (Table 3.1).

Table 3.1 Recovery efficiencies of ultrafiltration and PEG methods for the detection of Phi6 in seeded wastewater samples.

Wastewater Type	Sampling Site (n=x)	Phi6 phage recovery as measured by RT-ddPCR		
		Mean $\pm$ SD % (range)		
		Method 1/ CEN1	Method 2/ CEN2	Method 3/PEG
Sanitary Sewer	Hospital Lift Station (3)	9.59 $\pm$ 1.14 (8.90-10.91)	4.99 $\pm$ 0.04 (4.95-5.02)	51.47 $\pm$ 26.08 (26.52-78.55)
	Community manhole (6)	10.60 $\pm$ 14.58 (1.98-39.9)	11.64 $\pm$ 6.05 (5.77-22.07)	25.49 $\pm$ 18.46 (3.93-47.49)
Wastewater Treatment Plant Influent	WWTP A (3)	6.05 $\pm$ 4.89 (0.48-9.64)	2.73 $\pm$ 2.04 (1.23-5.05)	36.01 $\pm$ 19.41 (23.03-58.33)
	WWTP E (3)	9.25 $\pm$ 15.72 (0.05-27.41)	9.21 $\pm$ 15.37 (0.10-26.95)	31.98 $\pm$ 7.52 (23.57-38.07)
	WWTP M (3)	2.60 $\pm$ 1.39 (1.03-3.64)	10.37 $\pm$ 12.61 (0.87-24.68)	22.19 $\pm$ 15.72 (4.67-35.04)

The source of wastewater had no significant impact (two-way anova, $n = 18$, p -value = 0.4736) on the recovery efficiency of Phi6, regardless of the virus concentration method yet more variability was seen when testing sanitary sewer samples using PEG (Table A4).

3.4.3 Detection of SARS-CoV-2 in wastewater samples using ultrafiltration and PEG methods

All wastewater samples using the three concentration methods were also analyzed for SARS-CoV-2 using RT-ddPCR and RT-qPCR. The N1 and N2 gene targets showed similar results between the two PCR methods (Table 2). While the E gene target performed satisfactorily

on the RT-ddPCR platform, it showed poor results on the RT-qPCR platform with nearly all samples being identified as non-detects with no detected samples above the lower limit of quantification (LLOQ) (Table 2). The N2 gene target performed the best overall for the RT-qPCR assay. Using RT-ddPCR, the N1, N2, and E gene performed similarly with coefficients of variation for their detection of SARS-CoV-2 of 0.03 and 0.20 for sanitary sewer and WWTP influent samples, respectively (Table 2). Across three concentration methods RT-ddPCR showed fairly consistent patterns of SARS-CoV-2 detection, while the RT-qPCR assays relied heavily on the N2 gene target for SARS-CoV-2 detection (Table 2). Overall RT-ddPCR performed better at detecting SARS-CoV-2 gene targets than RT-qPCR in the wastewater samples tested with the exception of the N2 gene target in sanitary sewer samples which performed better with RT-qPCR (Table 2).

The overall concentrations of SARS-CoV-2 measured by RT-ddPCR for the three gene targets (N1, N2, E) ranged from $< \text{LLOD} - 5.71 \times 10^4 \text{ GC/100ml}$, $< \text{LLOD} - 1.11 \times 10^5 \text{ GC/100ml}$, and $< \text{LLOD} - 3.94 \times 10^4 \text{ GC/100ml}$, respectively (Table A3-A5). The overall concentrations of SARS-CoV-2 measured by RT-qPCR for the three gene targets (N1, N2, E) ranged from $< \text{LLOD} - 1.38 \times 10^5 \text{ GC/100ml}$, $< \text{LLOD} - 2.80 \times 10^5 \text{ GC/100ml}$, and $< \text{LLOQ}$, respectively (Table A3-A5). Slightly higher concentrations of N2 gene target in sanitary sewer and WWTP influent samples were obtained using RT-qPCR as compared to RT-ddPCR (Table 2).

Table 3.2 The detection of SARS-CoV-2 genes (N1, N2, E) using ultrafiltration and PEG precipitation (with 16-hr incubation) concentration methods.

Concentration	% Positive (Mean GC per 100 ml $\pm$ SD)
---------------	--

method	Sample Type	N1 gene		N2 gene		E gene	
		RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR
CEN1	Sanitary Sewer (n=9)	55.6 (2.44± 2.98×10 ³)	11.1 (8.74×10 ³ ± 1.86×10 ⁴)	66.7 (3.27± 3.20×10 ³)	77.8 (2.48± 3.86×10 ⁴)	88.9 (5.06± 6.32×10 ³)	11.1
	WWTP Influent (n=9)	55.6 (1.57± 2.71×10 ³)	ND ^{a,b}	88.9 (1.48±1.63 ×10 ³)	44.4 (1.38± 1.45×10 ³)	55.6 (1.93± 2.68×10 ³)	ND ^c
CEN2	Sanitary Sewer (n=9)	55.6 (3.82± 5.21×10 ³)	11.1 (1.37± 1.56×10 ³)	55.6 (1.48± 3.63×10 ⁴)	100 (4.68± 5.89×10 ³)	66.7 (3.16± 4.95×10 ³)	11.1
	WWTP Influent (n=9)	55.6 (6.55± 8.33×10 ²)	ND	88.9 (2.01± 1.94×10 ³)	77.8 (3.92± 8.62×10 ³)	44.4 (5.01± 3.81×10 ²)	ND
PEG (with 16-hr incubation)	Sanitary Sewer (n=9)	44.4 (1.27± 1.98×10 ⁴)	33.3 (1.85± 4.50×10 ⁴)	33.3 (1.22± 2.08×10 ⁴)	66.7 (2.67± 4.32×10 ⁴)	44.4 (9.27×10 ³ ± 1.31×10 ⁴)	11.1
	WWTP Influent (n=9)	77.8 (6.62± 7.15×10 ³)	22.2 (7.55×10 ³ ±1 .64×10 ⁴)	66.7 (6.76± 7.14×10 ³)	66.7 (4.17± 9.02×10 ⁴)	66.7 (5.87± 8.10×10 ³)	ND
All Methods	Sanitary Sewer (n=27)	51.9 (6.33×10 ³ ± 1.24×10 ⁴)	25.9 (9.53×10 ³ ±2 .80×10 ⁴)	51.9 (1.01± 2.38×10 ⁴)	81.5 (1.87×10 ⁴ ± 3.39×10 ⁴)	66.7 (5.83± 8.90×10 ³)	11.1
	WWTP Influent (n=27)	63 (2.95± 5.03×10 ³)	7.4 (2.93± 9.73×10 ³)	81.5 (3.42± 4.85×10 ³)	63 (1.57± 5.37×10 ⁴)	55.6 (2.77± 5.27×10 ³)	ND

^aND: Non-detect; ^bThe detection limit of N1 gene for qPCR is 4.7 GC per reaction; The lower detection limit of the E gene for qPCR was 47 gene copies per reaction.

Higher precision between gene targets was observed in the sanitary sewer samples versus the WWTP influent samples for both RT-ddPCR and RT-qPCR (Figure 1). RT-ddPCR showed significantly lower coefficients of variations for every combination of concentration method and sample type with the exception of WWTP influents processed by CEN2 and PEG (Figure 1).

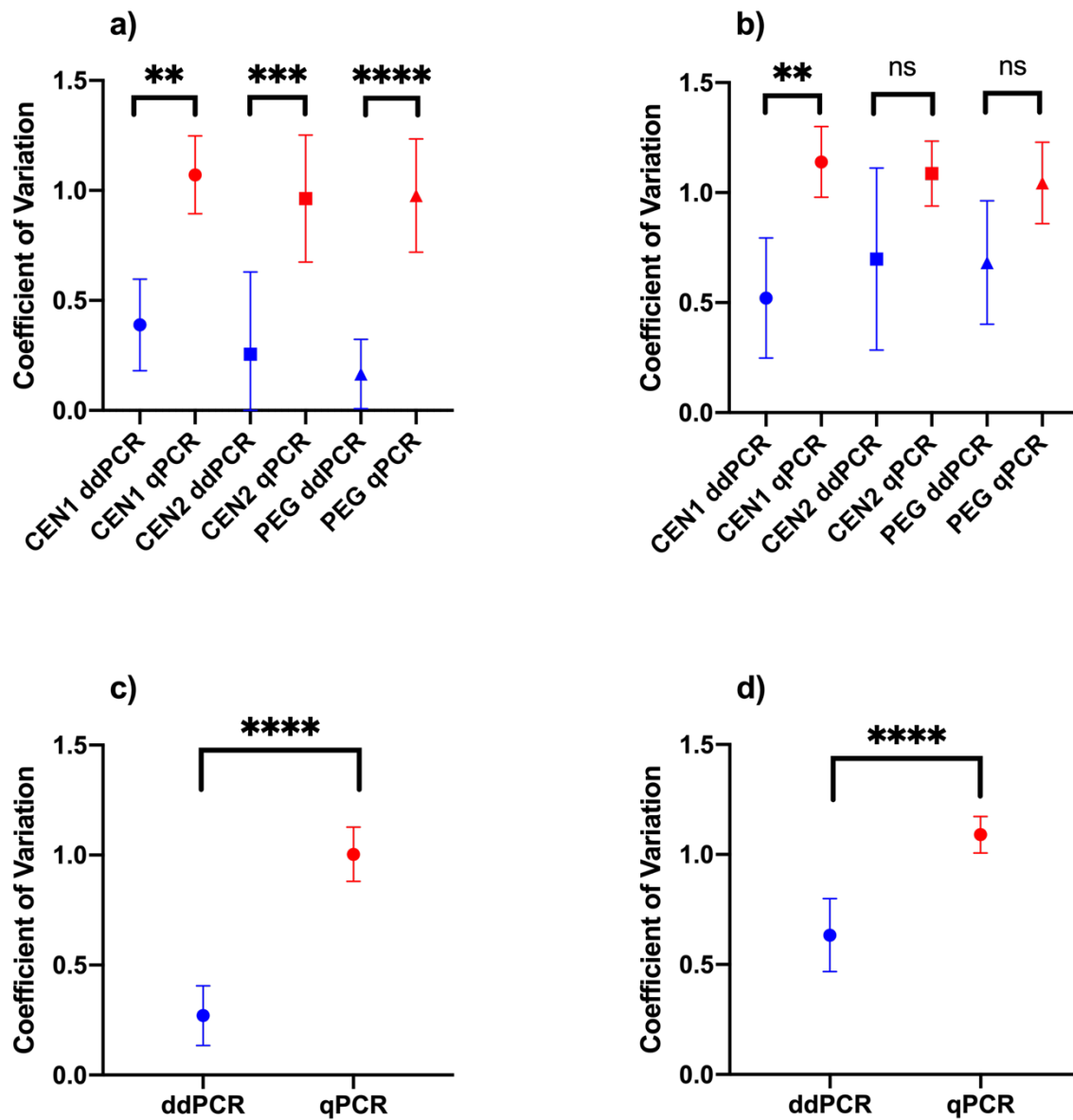


Figure 3.1 Coefficients of variations for SARS-CoV-2 gene targets; a) sanitary sewer samples, b) WWTP influent samples, c) sanitary sewer samples with all concentration methods, d) WWTP influent samples with all concentration methods. Two-way ANOVA analysis results shown above each graph; ns: Not-significant, * p-value <0.05, ** p-value <0.01, *** p-value <0.001, **** p-value <0.0001

3.4.4 Evaluation of rapid PEG approach for the detection of Phi6 and SARS-CoV-2 in wastewater

Based on the comparison between ultrafiltration and PEG methods for wastewater, the PEG precipitation was evaluated further for the detection of Phi6 and SARS-CoV-2 with and without 16 hours of incubation. As shown in Table 3, without an overnight incubation, the PEG method showed an average recovery efficiency of 18.8% for sanitary sewer and of 35% for WWTP influent samples. In general, this approach produced lower recovery for Phi6 as compared with the PEG with overnight incubation, regardless of the type of wastewater samples. However, the difference was not statistically significant (Table 3.3).

Table 3.3 Mean recovery efficiencies of Phi6 in seeded wastewater samples using PEG precipitation method with and without overnight incubation

Wastewater Type	Phi6 mean recovery $\pm$ SD (%)	
	PEG with overnight (16-hr) incubation	Rapid PEG without overnight incubation
Sanitary Sewer (n=15)	32.07 $\pm$ 23.23	18.80 $\pm$ 11.48
WWTP Influent (n=4)	31.31 $\pm$ 11.73	34.99 $\pm$ 26.38
All Samples (n=19)	31.92 $\pm$ 21.27 ^a	21.89 $\pm$ 15.88 ^a

^a No significant difference (n=19, p = 0.1048) in mean recovery efficiencies between methods.

For the detection of SARS-CoV-2 using RT-ddPCR, comparable results were obtained for the N1 and N2 gene targets with and without overnight incubation (Table 4). However, the rapid PEG method produced a lower percentage of WWTP influent samples positive for the E gene. Average SARS-CoV-2 concentrations varied little between sample types, PEG methods (i.e., with and without overnight incubation), and gene targets using RT-ddPCR, ranging from $1.82 \pm 4.55 \times 10^3$ GC/100ml to $8.57 \times 10^3 \pm 1.12 \times 10^4$ GC/100ml (Table 4).

Table 3.4 Percent positive and mean concentrations of SARS-CoV-2 gene targets for PEG method with and without overnight incubation as measured using RT-ddPCR

Wastewater Type	% Positive (Mean GC per 100 ml $\pm$ SD)					
	With overnight (16-hr) incubation			Without overnight incubation		
	N1	N2	E	N1	N2	E
Sanitary Sewer (n=15)	33.3 (4.36 $\pm$ 9.21 $\times 10^3$)	26.7 (4.64 $\pm$ 9.65 $\times 10^3$)	26.7 (3.39 $\pm$ 6.70 $\times 10^3$)	26.7 (4.19 $\pm$ 9.31 $\times 10^3$)	26.7 (3.95 $\pm$ 8.80 $\times 10^3$)	20 (1.97 $\pm$ 4.87 $\times 10^3$)
WWTP Influent (n=4)	75 (8.57 $\times 10^3 \pm$ 1.12 $\times 10^4$)	50 (2.0 $\pm$ 2.15 $\times 10^3$)	75 (8.53 $\pm$ 1.24 $\times 10^4$)	75 (3.97 $\pm$ 3.71 $\times 10^3$)	75 (6.40 $\pm$ 7.99 $\times 10^3$)	25 (7.92 $\times 10^2 \pm$ 3.39 $\times 10^1$)
All (n=19)	42.1 (5.25 $\pm$ 9.49 $\times 10^3$)	31.6 (4.08 $\pm$ 8.62 $\times 10^3$)	36.8 (4.47 $\pm$ 8.06 $\times 10^3$)	36.8 (4.14 $\pm$ 8.35 $\times 10^3$)	36.8 (4.46 $\pm$ 8.48 $\times 10^3$)	21.1 (1.82 $\pm$ 4.55 $\times 10^3$)

3.4.5 Evaluation of PEG precipitation using Phi6 and coronavirus OC43 as potential SARS-CoV-2 surrogates

When recovery efficiencies of Phi6 and OC43 were compared using five replicate influent samples from two WWTPs, no significant difference between the two surrogates was observed (Paired t test P value = 0.6137). These results are in line with recovery efficiencies for OC43 as reported by Pecson et al. (2021). However, the Phi6 recovery efficiencies observed in this study were almost 2 logs higher than the Phi6 recoveries achieved by other laboratories using the PEG method but with solids removal in the previous interlaboratory method assessment study (Pecson et al., 2021).

3.5 Discussion

This study demonstrates that *Pseudomonas* phage Phi6 seeded in different wastewater matrices can be concentrated and recovered by ultrafiltration-based method and PEG precipitation. In general, PEG method provided better virus recovery than the ultrafiltration-based methods as measured using RT-ddPCR. PEG precipitation is usually used as a secondary step for virus concentration in large volumes of water samples (De Keuckelaere et al., 2013; Polaczyk et al., 2008; Cuevas-Ferrando et al., 2021), but has also been used in concentrating enteric viruses directly from sewage as a primary concentration process prior to analysis by cell culture and molecular detection methods (Aw et al., 2010; Hovi et al., 2001; Myrmel et al., 2015; Shieh et al., 1995; Thongprachum et al., 2018). Enveloped viruses such as influenza A viruses and transmissible gastroenteritis virus (TGEV) have been detected in water samples using PEG precipitation combined with an initial filtration step (Blanco et al., 2019; Borchardt et al., 2017; Deboosere et al., 2011; Horm et al., 2012). However, only few studies have evaluated PEG

precipitation as a primary concentration approach for viruses in sewage (e.g., Amdioune et al., 2012; Hovi et al., 2001; Ye et al., 2016).

The exact mechanism of PEG precipitation of viruses from suspension is still not well understood. This could be due to that larger molecular aggregates such as viral particles are preferentially associated with inter-polymer spaces between PEG molecules and thus concentrated until their solubility is exceeded and precipitation occurs (Adams, 1973; Atha et al., 1981). This may also precipitate out viruses attached to other particles in untreated wastewater which is particularly difficult to filter. Thus, this method may be better for those samples with a higher range of TSS and turbidities as evidenced by the wider range in SARS-CoV-2 N1 gene concentrations seen by Pecson et al. (2021) with PEG protocols which removed solids as compared to PEG protocols which retained solids.

In a previous study comparing three methods to concentrate enveloped murine hepatitis virus (MHV) from wastewater samples, PEG and ultracentrifugation recovered 5% of the seeded MHV, whereas the ultrafiltration produced significantly higher recovery, 25% (Ye et al., 2016). A possible explanation to lower virus recovery efficiencies obtained with PEG precipitation could be a high salt (NaCl) concentration (0.5 M) added to the samples. High salt concentrations may inactivate enveloped viruses during PEG precipitation process. For example, Hamelin et al., (1979) showed that infectivity of cytomegalovirus (an enveloped virus) declined from 24.7 to 6.6% as the NaCl concentration was progressively increased from none to 1.0 M NaCl. It has also been reported that infectivity of retro- and lentiviruses decreases significantly in a high salt elution buffer (1M NaCl) (Zimmermann et al., 2011). In this study, 0.2 M NaCl was used in the PEG protocol.

The choice of PEG precipitation over other concentration methods was also based on the affordability of the procedure and the shortages of the filtration materials due to the increasing numbers of laboratories worldwide that monitor SARS-CoV-2 in wastewater. PEG precipitation is a simple and low-cost alternative (e.g., < \$2 USD per sample for PEG method vs. >\$34 USD for ultrafiltration-based method) for the concentration of viruses in wastewater without requiring any preconditioning of the sample. The PEG method used in this study has also been evaluated in a recent interlaboratory methods assessment for SARS-CoV-2 genetic signal in raw sewage using betacoronavirus OC43 as a matrix spike. By comparing 36 standard operating procedures used by 32 participating laboratories, PEG precipitation has shown a high degree of reproducibility across laboratories (Pecson et al., 2021).

Although PEG precipitation provided higher recovery efficiencies for Phi6 and SARS-CoV-2 in wastewater when compared with ultrafiltration, the protocol is slower particularly with an overnight incubation. However, in this study, the results of PEG precipitation with and without an overnight incubation for Phi6 and SARS-CoV-2 were not statistically significant. This is in agreement with other studies that reported a 2-hour precipitation is sufficient for viruses (Deboosere et al., 2011; Polaczyk et al., 2008). Therefore, the PEG protocol could be shortened to increase throughput or accommodate existing analysis workflows for rapid results.

In addition to investigating recovery efficiencies of artificially seeded viruses using different concentration methods, this study compared the detection of SARS-CoV-2 genetic signals in wastewater using RT-qPCR and RT-ddPCR. Overall, RT-ddPCR showed higher sensitivity rate compared to RT-qPCR. While RT-qPCR shows equivalent detection rate of the SARS-CoV-2 N2 gene as RT-ddPCR, RT-ddPCR performed better for the E gene in wastewater. This may be due to RT-ddPCR allowing for greater PCR efficiency when lower concentrations

of the target gene are present and its ability to cope with higher levels of inhibitory substances in wastewater. While a high number of samples in this study were found to be positive for one or more of the SARS-CoV-2 gene targets, a direct comparison of the virus concentrations between sanitary sewer and WWTP influent samples would be inaccurate due to the different sampling methods. For sanitary sewer, grab sampling was used to collect wastewater directly from manholes or lift station whereas composite sampling technique was used for the WWTP. Different wastewater sampling techniques may influence the ability to detect and quantify viral genetic markers using PCR-based methods. For example, a grab sample taken during low flow periods may miss detecting the SARS-CoV-2 genetic markers in wastewater. A similar situation can occur for composite samples particularly for long sampling periods (e.g., 24 hrs) as the viral signals may be diluted. Therefore, determination of the optimal sampling strategy and timing will greatly enhance the ability to accurately detect SARS-CoV-2 in wastewater. Heaton et al. (1992) showed that over 60% of men and women defecated between 5 am and 12 pm each day. These patterns may have changed since the study, but sample collection time is still an important factor to consider when conducting a wastewater surveillance for SARS-CoV-2.

The concentration and detection procedures outlined in this study will facilitate rapid and high-throughput detection of SARS-CoV-2 in wastewater samples. The methods were used successfully in field studies for the detection of SARS-CoV-2 RNA in various wastewater samples.

APPENDIX

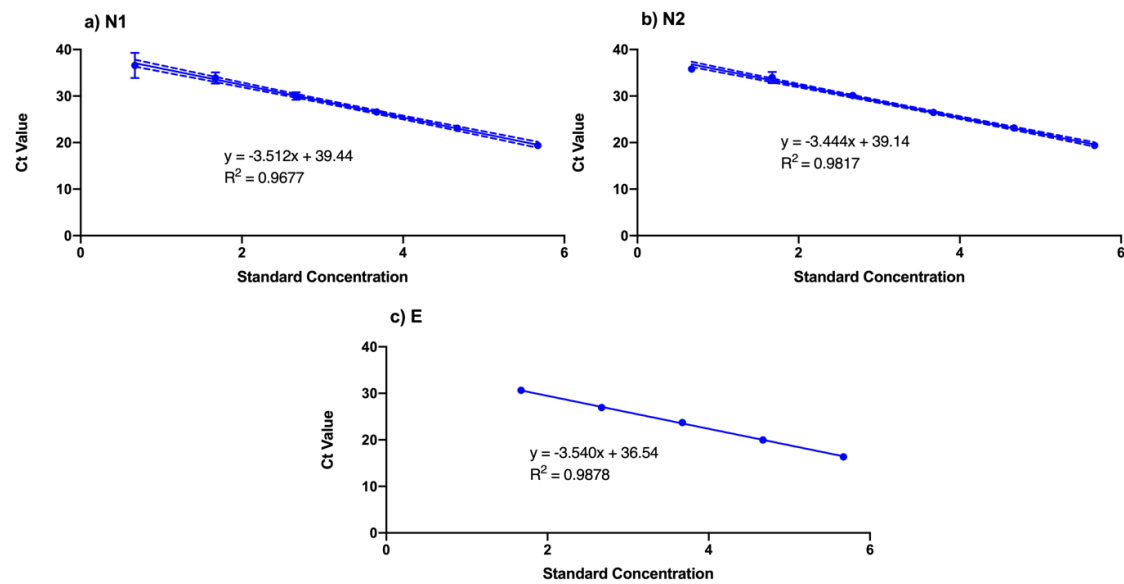


Figure 3.A1 qPCR standard curves for SARS-CoV-2 gene targets with slope, y intercept and R^2 .
a) N1 standard curve, b) N2 standard curve, c) E gene standard curve.

Table 3.A1 Individual recovery efficiencies of ultrafiltration and PEG methods for the detection of Phi6 in seeded wastewater samples.

Wastewater Type	Site ID	Sample Date	Ultrafiltration Method 1 % Recovery	Ultrafiltration Method 2 % Recovery	PEG Precipitation % Recovery
Sanitary Sewer	Hospital Lift Station	3/25/2020	10.91	5.02	78.55
	Hospital Lift Station	3/25/2020	8.97	4.99	49.33
	Hospital Lift Station	3/25/2020	8.90	4.95	26.52
	MSU1	5/11/2020	39.90	22.07	8.40
	MSU2	5/11/2020	4.11	11.13	3.93
	MSU1	8/3/2020	4.28	14.73	21.82
	MSU2	8/3/2020	1.98	6.57	24.61
	MSU3	8/3/2020	3.77	5.77	47.49
	MSU4	8/3/2020	9.56	9.55	46.68
Wastewater Treatment Plant Influent	WWTP A	4/6/2020	8.02	1.90	26.68
	WWTP A	4/20/2020	9.64	1.23	23.03
	WWTP A	6/1/2020	0.48	5.05	58.33
	WWTP E	4/13/2020	27.41	26.95	38.07
	WWTP E	4/20/2020	0.30	0.57	23.57
	WWTP E	6/1/2020	0.05	0.10	34.31
	WWTP M	5/6/2020	3.64	24.68	4.67
	WWTP M	4/29/2020	3.14	5.57	26.87
	WWTP M	6/3/2020	1.03	0.87	35.04

Table 3.A2 Individual sample Phi6 percent recoveries for two PEG viral concentration methods

Wastewater Type	Site ID	Sampling Date	2-hour Spin followed by 16 hr hold % Recovery	2-hour Spin without hold % Recovery
Sanitary Sewer	Hospital Lift Station	3/25/20	18.15	15.56
	MSU3	9/8/20	15.12	5.39
	MSU3	9/14/20	18.91	35.19
	MSU4	9/8/20	12.87	17.65
	MSU4	9/14/20	4.93	35.36
	MSU5	9/8/20	77.79	29.79
	MSU5	9/14/20	17.88	7.52
	MSU6	9/8/20	57.03	32.28
	MSU6	9/14/20	34.52	37.95
	MSU7	9/8/20	13.65	10.52
	MSU7	9/14/20	33.11	9.85
	MSU8	9/8/20	49.79	22.41
	MSU8	9/14/20	12.06	13.22
	LRB2	8/31/20	57.06	15.38
	LRB3	8/31/20	77.68	3.49
Wastewater Treatment Plant Influent (Post-Grit)	WWTP A	4/6/20	21.92	20.92
	WWTP A	4/20/20	22.69	7.17
	WWTP M	8/31/20	33.66	66.85
	WWTP W	8/31/20	46.96	45.01

Table 3.A3 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for centrifugation method 1.

Sample Type	Site ID	Sample Date	N1 GC 100ml ⁻¹		N2 GC 100ml ⁻¹		E GC 100ml ⁻¹	
			RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR	RT-ddPCR	RT_qPCR
Sanitary Sewer	Hospital Lift Station	3/25/2020	7.57E+03	5.67E+04	8.80E+03	1.13E+05	1.22E+04	DNQ ^a
	Hospital Lift Station	3/25/2020	6.94E+03	ND	6.27E+03	2.90E+04	1.55E+04	ND ^b
	Hospital Lift Station	3/25/2020	1.68E+03	1.51E+04	5.62E+03	5.84E+04	2.76E+03	ND
	MSU1	5/11/2020	3.95E+03	ND	4.22E+03	ND	1.23E+04	ND
	MSU2	5/11/2020	7.58E+02	ND	3.41E+03	2.07E+04	1.36E+03	ND
	MSU1	8/3/2020	ND	ND	3.31E+02	DNQ	5.30E+02	ND
	MSU2	8/3/2020	ND	ND	ND	DNQ	2.62E+02	ND
	MSU3	8/3/2020	ND	ND	ND	DNQ	3.70E+02	ND
	MSU4	8/3/2020	ND	ND	ND	ND	ND	ND
Wastewater Plant Influent (Post-grit)	WWTP A	4/6/2020	2.55E+03	ND	2.80E+03	ND	7.95E+03	ND
	WWTP A	4/20/2020	1.43E+03	ND	2.20E+03	ND	2.90E+03	ND
	WWTP A	6/1/2020	ND	ND	2.16E+02	ND	ND	ND
	WWTP E	4/13/2020	8.45E+03	ND	5.05E+03	4.59E+03	4.20E+03	ND
	WWTP E	4/20/2020	ND	ND	1.35E+03	1.46E+03	ND	ND
	WWTP E	6/1/2020	ND	ND	ND	ND	ND	ND
	WWTP M	5/6/2020	4.99E+02	ND	7.82E+02	ND	1.06E+03	ND
	WWTP M	4/29/2020	6.96E+02	ND	2.97E+02	1.46E+03	7.40E+02	ND
	WWTP M	6/3/2020	ND	ND	5.20E+02	5.23E+02	ND	ND

Table 3.A4 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for centrifugation method 2.

Sample Type	Site ID	Sample Date	N1		N2		E	
			GC 100ml ⁻¹		GC 100ml ⁻¹		GC 100ml ⁻¹	
			RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR
Sanitary Sewer	Hospital Lift Station	3/25/2020	6.04E+03	4.91E+03	3.70E+03	4.53E+03	2.48E+03	DNQ ^a
	Hospital Lift Station	3/25/2020	1.58E+04	ND ^b	1.37E+04	8.17E+03	1.56E+04	ND
	Hospital Lift Station	3/25/2020	3.87E+03	ND	3.26E+03	1.82E+04	4.16E+03	ND
	MSU1	5/11/2020	6.87E+03	ND	1.11E+05	7.26E+03	4.27E+03	ND
	MSU2	5/11/2020	ND	ND	ND	1.21E+03	3.17E+02	ND
	MSU1	8/3/2020	ND	ND	ND	7.33E+02	ND	ND
	MSU2	8/3/2020	ND	ND	ND	6.13E+02	ND	ND
	MSU3	8/3/2020	ND	ND	ND	4.95E+02	ND	ND
	MSU4	8/3/2020	3.46E+02	ND	3.41E+02	9.52E+02	3.71E+02	ND
Wastewater Plant Influent (Post-grit)	WWTP A	4/6/2020	5.36E+02	ND	1.69E+03	ND	7.78E+02	ND
	WWTP A	4/20/2020	5.82E+02	ND	5.46E+02	3.04E+03	5.96E+02	ND
	WWTP A	6/1/2020	ND	ND	4.88E+02	3.74E+02	ND	ND
	WWTP E	4/13/2020	ND	ND	9.36E+02	2.68E+04	1.05E+03	ND
	WWTP E	4/20/2020	ND	ND	1.16E+03	8.58E+02	ND	ND
	WWTP E	6/1/2020	ND	ND	ND	1.08E+03	ND	ND
	WWTP M	5/6/2020	4.93E+02	ND	2.96E+03	ND	ND	ND
	WWTP M	4/29/2020	2.82E+03	ND	5.99E+03	8.58E+02	1.06E+03	ND
	WWTP M	6/3/2020	1.92E+02	ND	4.03E+03	1.14E+03	ND	ND

Table 3.A5 Mean concentrations for SARS-CoV-2 Gene Targets for RT-ddPCR and RT-qPCR for PEG precipitation (with 16-hr hold).

Sample Type	Site ID	Sample Date	N1		N2		E	
			GC 100ml ⁻¹		GC 100ml ⁻¹		GC 100ml ⁻¹	
			RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR
Sanitary Sewer	Hospital Lift Station	3/25/2020	2.09E+04	1.38E+05	1.20E+04	1.31E+05	3.49E+04	DNQ ^a
	Hospital Lift Station	3/25/2020	5.71E+04	9.49E+03	6.05E+04	5.64E+04	2.05E+04	ND ^b
	Hospital Lift Station	3/25/2020	2.98E+04	9.83E+03	3.17E+04	2.36E+04	2.22E+04	ND
	MSU1	5/11/2020	2.99E+03	ND	ND	ND	ND	ND
	MSU2	5/11/2020	ND	ND	ND	2.15E+04	ND	ND
	MSU1	8/3/2020	ND	ND	ND	1.58E+03	ND	ND
	MSU2	8/3/2020	ND	ND	ND	ND	ND	ND
	MSU3	8/3/2020	ND	ND	ND	ND	ND	ND
	MSU4	8/3/2020	ND	ND	ND	1.34E+03	6.10E+02	ND
Wastewater Plant Influent (Post-grit)	WWTP A	4/6/2020	8.18E+03	ND	ND	1.94E+04	5.95E+03	ND
	WWTP A	4/20/2020	2.46E+04	ND	5.20E+03	3.63E+04	2.67E+04	ND
	WWTP A	6/1/2020	ND	ND	ND	2.41E+04	ND	ND
	WWTP E	4/13/2020	7.67E+03	5.14E+04	2.27E+03	2.80E+05	6.09E+03	ND
	WWTP E	4/20/2020	4.61E+03	ND	4.93E+03	ND	2.53E+03	ND
	WWTP E	6/1/2020	2.66E+03	1.99E+03	8.83E+03	5.12E+03	ND	ND
	WWTP M	5/6/2020	ND	ND	2.30E+04	5.11E+03	5.90E+03	ND
	WWTP M	4/29/2020	3.64E+03	ND	1.22E+04	ND	ND	ND
	WWTP M	6/3/2020	5.04E+03	ND	ND	ND	1.48E+03	ND

^aDNQ: Detected Non-quantifiable; ^bND: Non-detect.

Table 3.A6 Individual sample SARS-CoV-2 gene concentrations for PEG precipitation with 16-hour hold and without holding.

Sample Type	Site ID	Sample Date	16-hour Hold			No Hold		
			N1 (GC 100ml ⁻¹)	N2 (GC 100ml ⁻¹)	E (GC 100ml ⁻¹)	N1 (GC 100ml ⁻¹)	N2 (GC 100ml ⁻¹)	E (GC 100ml ⁻¹)
Sanitary Sewer	Hospital Lift Station	3/25/20	2.98E+04	3.17E+04	2.22E+04	2.28E+04	2.02E+04	NA ^a
	MSU3	9/8/20	1.34E+03	1.68E+03	1.17E+03	8.53E+02	7.68E+02	5.63E+02
	MSU3	9/14/20	2.39E+04	2.45E+04	1.72E+04	3.08E+04	3.01E+04	1.89E+04
	MSU4	9/8/20	ND ^b	ND	ND	ND	ND	ND
	MSU4	9/14/20	5.60E+02	ND	ND	1.34E+03	ND	ND
	MSU5	9/8/20	ND	ND	ND	ND	ND	ND
	MSU5	9/14/20	ND	ND	ND	ND	ND	ND
	MSU6	9/8/20	ND	ND	ND	ND	ND	ND
	MSU6	9/14/20	1.12E+03	2.63E+03	1.18E+03	ND	9.43E+02	8.97E+02
	MSU7	9/8/20	ND	ND	ND	ND	ND	ND
	MSU7	9/14/20	ND	ND	ND	ND	ND	ND
	MSU8	9/8/20	ND	ND	ND	ND	ND	ND
	MSU8	9/14/20	ND	ND	ND	ND	ND	ND
	LRB2	8/31/20	ND	ND	ND	ND	ND	ND
	LRB3	8/31/20	ND	ND	ND	ND	ND	ND
Wastewater Plant Influent (Post-grit)	WWTP A	4/6/20	8.18E+03	ND	5.95E+03	6.45E+03	6.28E+03	NA
	WWTP A	4/20/20	2.46E+04	5.20E+03	2.67E+04	7.84E+03	1.77E+04	NA
	WWTP M	9/2/20	ND	ND	ND	ND	ND	ND
	WWTP W	8/31/20	6.72E+02	6.72E+02	6.72E+02	7.68E+02	7.68E+02	7.68E+02

^aNA: Not available; ^bND: Non-detect

Table 3.A7 Individual coefficients of variations for RT-ddPCR and RT-qPCR for three SARS-CoV-2 gene targets.

Sample Type	Site ID	Coefficient of Variation					
		CEN1		CEN2		PEG	
		RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR	RT-ddPCR	RT-qPCR
Sanitary Sewer	Hospital Lift Station	0.25	0.73	0.44	0.11	0.51	0.59
	Hospital Lift Station	0.54	1.01	0.08	0.77	0.48	0.87
	Hospital Lift Station	0.61	0.94	0.12	0.84	0.18	0.42
	MSU1	0.70	0.85	1.50	1.08	0.23	1.30
	MSU2	0.75	0.91	0.11	1.02	0.00	0.74
	MSU1	0.35	1.30	0.00	1.20	0.00	1.08
	MSU2	0.11	1.30	0.00	1.28	0.00	1.30
	MSU3	0.19	1.30	0.00	1.38	0.00	1.30
	MSU3	0.00	1.30	0.05	0.99	0.09	1.19
Wastewater Plant Influent (Post-grit)	WWTP A	0.69	1.30	0.61	1.30	0.68	0.74
	WWTP A	0.34	1.30	0.04	0.79	0.63	0.79
	WWTP A	0.43	1.30	0.73	1.21	0.00	0.75
	WWTP E	0.38	1.04	0.24	1.03	0.52	1.18
	WWTP E	1.24	0.80	1.15	0.99	0.32	1.30
	WWTP E	0.00	1.30	0.00	1.10	0.99	0.99
	WWTP M	0.36	1.30	1.21	1.30	1.14	1.05
	WWTP M	0.42	0.84	0.76	1.23	1.01	1.30
	WWTP M	0.83	1.08	1.54	0.83	0.85	1.30

Table 3.A8 Individual recovery efficiencies of the SARS-CoV-2 surrogates Phi6 and OC43 at two WWTPs.

WWTP	Sample Replicate	Recovery Efficiency (%)	
		OC43	Phi6
Hyperion	1	0.72	1.15
	2	14.55	11.85
	3	13.00	8.49
	4	7.33	5.68
	5	4.91	6.93
	All	8.10	6.82
JWPCP	1	6.50	3.71
	2	2.53	3.88
	3	4.83	6.41
	4	5.45	6.48
	5	2.03	3.44
	All	4.27	4.78

REFERENCES

REFERENCES

- Adams, A. (1973). Concentration of Epstein-Barr Virus from Cell Culture Fluids with Polyethylene Glycol. *J Gen Virol* 20, 391–394. <https://doi.org/10.1099/0022-1317-20-3-391>
- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tschärke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K.V., and Mueller, J.F. (2020a). 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. *Sci Total Environ* 728, 138764. <https://doi.org/10.1016/j.scitotenv.2020.138764>
- Ahmed, W., Bertsch, P.M., Angel, N., Bibby, K., Bivins, A., Dierens, L., Edson, J., Ehret, J., Gyawali, P., Hamilton, K., Hosegood, I., Hugenholtz, P., Jiang, G., Kitajima, M., Sichani, H.T., Shi, J., Shimko, K.M., Simpson, S.L., Smith, W.J.M., Symonds, E.M., DSC, K.V.T., Verhagen, R., Zaugg, J., and Mueller, J.F. (2020b). Detection of SARS-CoV-2 RNA in Commercial Passenger Aircraft and Cruise Ship Wastewater: A Surveillance Tool for Assessing the Presence of COVID-19 Infected Travelers. *J Travel Med* 27, taaa116-. <https://doi.org/10.1093/jtm/taaa116>
- Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B.R., Mueller, J.F., Simpson, S.L., Smith, W.J.M., Symonds, E.M., Thomas, K.V., Verhagen, R., and Kitajima, M., (2020c). Comparison of Virus Concentration Methods for the RT-qPCR-based Recovery of Murine Hepatitis Virus, a Surrogate for SARS-CoV-2 from Untreated Wastewater. *Sci Total Environ* 739, 139960. <https://doi.org/10.1016/j.scitotenv.2020.139960>
- Amdioune, H., Maunula, L., Hajjani, K., Faouzi, A., Soukri, A., and Nourlil, J. (2012). Recovery Comparison of Two Virus Concentration Methods from Wastewater Using Cell Culture and Real-Time PCR. *Curr Microbiol* 65, 432–437. <https://doi.org/10.1007/s00284-012-0174-8>
- Ampuero, M., Valenzuela, S., Valiente-Echeverria, F., Soto-Rifo, R., Barriga, G.P., Chnaiderman, J., Rojas, C., Guajardo-Leiva, S., Diez, B., and Gaggero, A. (2020). SARS-CoV-2 Detection in Sewage in Santiago, Chile - Preliminary results. *medRxiv preprint*. <https://doi.org/10.1101/2020.07.02.20145177>
- Arora, S., Nag, A., Sethi, J., Rajvanshi, J., Saxena, S., Shrivastava, S.K., and 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. *medRxiv preprint*. <https://doi.org/10.1101/2020.06.18.20135277>
- Atha, D.H. and Ingham, K.C. (1981). Mechanism of Precipitation of Proteins by Polyethylene Glycols. *Analysis in Terms of Excluded Volume. J Biological Chem* 256, 12108–17.

Aw, T.G. and Gin, K.Y. -H. (2010). Environmental Surveillance and Molecular Characterization of Human Enteric Viruses in Tropical Urban Wastewaters. *J Appl Microbiol* 109, 716–730. <https://doi.org/10.1111/j.1365-2672.2010.04701.x>

Balboa, S., Mauricio-Iglesias, M., Rodríguez, S., Martínez-Lamas, L., Vasallo, F.J., Regueiro, B., and Lema, J.M. (2020). The Fate of SARS-CoV-2 in Wastewater Treatment Plants Points Out the Sludge Line as a Suitable Spot for Incidence Monitoring. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.25.20112706>

Barcelo, D. (2020). An Environmental and Health Perspective for COVID-19 Outbreak: Meteorology and Air Quality Influence, Sewage Epidemiology Indicator, Hospitals Disinfection, Drug Therapies and Recommendations. *J Environ Chem Eng* 8, 104006. <https://doi.org/10.1016/j.jece.2020.104006>

Bar-Or, I., Yaniv, K., Shagan, M., Ozer, E., Erster, O., Mendelson, E., Mannasse, B., Shirazi, R., Kramarsky-Winter, E., Nir, O., Abu-Ali, H., Ronen, Z., Rinott, E., Lewis, Y.E., Friedler, E., Bitkover, E., Paitan, Y., Berchenko, Y., and Kushmaro, A. (2020). Regressing SARS-CoV-2 Sewage Measurements onto COVID-19 Burden in the Population: a Proof-of-concept for Quantitative Environmental Surveillance. *medRxiv preprint*. <https://doi.org/10.1101/2020.04.26.20073569>

Blanco, A., Abid, I., Al-Otaibi, N., Pérez-Rodríguez, F.J., Fuentes, C., Guix, S., Pintó, R.M., and Bosch, A. (2019). Glass Wool Concentration Optimization for the Detection of Enveloped and Non-enveloped Waterborne Viruses. *Food Environ Virol* 11, 184–192. <https://doi.org/10.1007/s12560-019-09378-0>

Borchardt, M.A., Spencer, S.K., Hubbard, L.E., Firnstahl, A.D., Stokdyk, J.P., and Kolpin, D.W. (2017). Avian Influenza Virus RNA in Groundwater Wells Supplying Poultry Farms Affected by the 2015 Influenza Outbreak. *Environ Sci Tech Lett* 4, 268–272. <https://doi.org/10.1021/acs.estlett.7b00128>

Carvalho, N.A. de, Stachler, E.N., Cimabue, N., and Bibby, K. (2017). Evaluation of Phi6 Persistence and Suitability as an Enveloped Virus Surrogate. *Environ Sci Technol* 51, 8692–8700. <https://doi.org/10.1021/acs.est.7b01296>

Chavarria-Miró, G., Anfruns-Estrada, E., Guix, S., Paraira, M., Galofré, B., Sánchez, G., Pintó, R., and Bosch, A. (2020). Sentinel Surveillance of SARS-CoV-2 in Wastewater Anticipates the Occurrence of COVID-19 Cases. *medRxiv preprint*. <https://doi.org/10.1101/2020.06.13.20129627>

Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., Mulders, D.G., Haagmans, B.L., Veer, B. van der, Brink, S. van den, Wijsman, L., Goderski, G., Romette, J.-L., Ellis, J., Zambon, M., Peiris, M., Goossens, H., Reusken, C., Koopmans, M.P., and Drosten, C. (2020). Detection of 2019 Novel Coronavirus (2019-nCoV) by Real-Time RT-PCR. *Eurosurveillance* 25, 2000045. <https://doi.org/10.2807/1560-7917.es.2020.25.3.2000045>

Cuevas-Ferrando, E., Pérez-Cataluña, A., Allende, A., Guix, S., Randazzo, W., and Sánchez, G. (2020). Recovering Coronavirus from Large Volumes of Water. *Science of The Total Environment*, 762, 143101. <https://doi.org/10.1016/j.scitotenv.2020.143101>

Curtis, K., Keeling, D., Yetka, K., Larson, A., and Gonzalez, R. (2020). Wastewater SARS-CoV-2 Concentration and Loading Variability from Grab and 24-Hour Composite Samples. *medRxiv preprint*. <https://doi.org/10.1101/2020.07.10.20150607>

Dare, R.K., Fry, A.M., Chittaganpitch, M., Sawanpanyalert, P., Olsen, S.J., and Erdman, D.D. (2007). Human Coronavirus Infections in Rural Thailand: A Comprehensive Study Using Real-Time Reverse-Transcription Polymerase Chain Reaction Assays. *J Infect Dis* 196, 1321–1328. <https://doi.org/10.1086/521308>

Deboosere, N., Horm, S.V., Pinon, A., Gachet, J., Coldefy, C., Buchy, P., and Vialette, M. (2011). Development and Validation of a Concentration Method for the Detection of Influenza A Viruses from Large Volumes of Surface Water. *Appl Environ Microb* 77, 3802–3808. <https://doi.org/10.1128/aem.02484-10>

Döhla, M., Wilbring, G., Schulte, B., Kümmerer, B.M., Diegmann, C., Sib, E., Richter, E., Haag, A., Engelhart, S., Eis-Hübinger, A.M., Exner, M., Streeck, H., and Schmithausen, R.M. (2020). SARS-CoV-2 in Environmental Samples of Quarantined Households. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.28.20114041>

Fongaro, G., Stoco, P.H., Souza, D.S.M., Grisard, E.C., Magri, M.E., Rogovski, P., Schorner, M.A., Barazzetti, F.H., Christoff, A.P., Oliveira, L.F.V. de, Bazzo, M.L., Wagner, G., Hernandez, M., and Rodriguez-Lazaro, D. (2020). SARS-CoV-2 in Human Sewage in Santa Catalina, Brazil, November 2019. *medRxiv preprint*. <https://doi.org/10.1101/2020.06.26.20140731>

Gendron, L., Verreault, D., Veillette, M., Moineau, S., and Duchaine, C. (2010). Evaluation of Filters for the Sampling and Quantification of RNA Phage Aerosols. *Aerosol Sci Tech* 44, 893–901. <https://doi.org/10.1080/02786826.2010.501351>

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M. H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J., and Gonzalez, D. (2020). COVID-19 Surveillance in Southeastern Virginia using Wastewater-Based Epidemiology. *Water Research*, 186, 116296. <https://doi.org/10.1016/j.watres.2020.116296>

Green, H., Wilder, M., Middleton, F.A., Collins, M., Fenty, A., Gentile, K., Kmush, B., Zeng, T., and Larsen, D.A. (2020). Quantification of SARS-CoV-2 and Cross-Assembly Phage (crAssphage) from Wastewater to Monitor Coronavirus Transmission within Communities. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.21.20109181>

Guerrero-Latorre, L., Ballesteros, I., Villacres, I., Granda-Albuja, M.G., Freire, B., and Rios-Touma, B. (2020). First SARS-CoV-2 Detection in River Water: Implications in Low Sanitation Countries. *medRxiv preprint*. <https://doi.org/10.1101/2020.06.14.20131201>

- Hamelin, C. and Lussier, G. (1979). Concentration of Human Cytomegalovirus from Large Volumes of Tissue Culture Fluids. *J Gen Virol* 42, 193–197. <https://doi.org/10.1099/0022-1317-42-1-193>
- Haramoto, E., Kitajima, M., Hata, A., Torrey, J.R., Masago, Y., Sano, D., and Katayama, H. (2018). A Review on Recent Progress in the Detection Methods and Prevalence of Human Enteric Viruses in Water. *Water Res* 135, 168–186. <https://doi.org/10.1016/j.watres.2018.02.004>
- Haramoto, E., Malla, B., Thakali, O., and Kitajima, M. (2020). First Environmental Surveillance for the Presence of SARS-CoV-2 RNA in Wastewater and River Water in Japan. *Sci Total Environ* 737, 140405. <https://doi.org/10.1016/j.scitotenv.2020.140405>
- Hata, A., Honda, R., Hara-Yamamura, H., and Meuchi, Y. (2020). Detection of SARS-CoV-2 in Wastewater in Japan by Multiple Molecular Assays: Implication for Wastewater-Based Epidemiology (WBE). *medRxiv preprint*. <https://doi.org/10.1101/2020.06.09.20126417>
- Heaton, K.W., Radvan, J., Cripps, H., Mountford, R.A., Braddon, F.E., and Hughes, A.O. (1992). Defecation Frequency and Timing, and Stool Form in the General Population: A Prospective Study. *Gut* 33, 818. <https://doi.org/10.1136/gut.33.6.818>
- Horm, S.V., Gutiérrez, R.A., Sorn, S., and Buchy, P. (2012). Environment: A Potential Source of Animal and Human Infection with Influenza A (H5N1) Virus. *Influenza Other Resp* 6, 442–448. <https://doi.org/10.1111/j.1750-2659.2012.00338.x>
- Hovi, T., Stenvik, M., Partanen, H., and Kangas, A. (2001). Poliovirus Surveillance by Examining Sewage Specimens. Quantitative Recovery of Virus after Introduction into Sewerage at Remote Upstream Location. *Epidemiology Amp Infect* 127, 101–106. <https://doi.org/10.1017/s0950268801005787>
- Jorgensen, A.U., Gamst, J., Hansen, L.V., Knudsen, I.I.H., and Jensen, S.K.S. (2020). Eurofins Covid-19 Sentinel Wastewater Test Provide Early Warning of a Potential COVID-19 Outbreak. *medRxiv preprint*. <https://doi.org/10.1101/2020.07.10.20150573>
- Keuckelaere, A.D., Baert, L., Duarte, A., Stals, A., and Uyttendaele, M. (2013). Evaluation of Viral Concentration Methods from Irrigation and Processing Water. *J Virol Methods* 187, 294–303. <https://doi.org/10.1016/j.jviromet.2012.11.028>
- Kocamemi, B.A., Kurt, H., Hacıoglu, S., Yaralı, C., Saatci, A.M., and Pakdemirli, B. (2020). First Data-Set on SARS-CoV-2 Detection for Istanbul Wastewaters in Turkey. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.03.20089417>
- Kocamemi, B.A., Kurt, H., Sait, A., Sarac, F., Saatci, A.M., and Pakdemirli, B. (2020). SARS-CoV-2 Detection in Istanbul Wastewater Treatment Plant Sludges. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.12.20099358>

Kumar, M., Patel, A.K., Shah, A.V., Raval, J., Rajpara, N., Joshi, M., and Joshi, C.G. (2020). The First Proof of the Capability of Wastewater Surveillance for COVID-19 in India through the Detection of the Genetic Material of SARS-CoV-2. medRxiv preprint. <https://doi.org/10.1101/2020.06.16.20133215>

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., and Stadler, L.B. (2021). Evaluating Recovery, Cost, and Throughput of Different Concentration Methods for SARS-CoV-2 Wastewater-Based Epidemiology. *Water Res* 197, 117043. <https://doi.org/10.1016/j.watres.2021.117043>

Lu, X., Wang, L., Sakthivel, S.K., Whitaker, B., Murray, J., Kamili, S., Lynch, B., Malapati, L., Burke, S.A., Harcourt, J., Tamin, A., Thornburg, N.J., Villanueva, J.M., and Lindstrom, S. (2020). US CDC Real-Time Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome Coronavirus 2 - Volume 26, Number 8—August 2020 - *Emerging Infectious Diseases journal - CDC*. *Emerg Infect Dis* 26, 1654–1665. <https://doi.org/10.3201/eid2608.201246>

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., and Brouwer, A. (2020). Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environ Sci Tech Let* 7, 511–516. <https://doi.org/10.1021/acs.estlett.0c00357>

Mera, I.G.F. de, Río, F.J.R. del, Fuente, J. de la, Sancho, M.P., Hervas, D., Moreno, I., Dominguez, M., Domínguez, L., and Gortázar, C. (2020). COVID-19 in a Rural Community: Outbreak Dynamics, Contact Tracing and Environmental RNA. medRxiv preprint. <https://doi.org/10.20944/preprints202005.0450.v1>

Miyani, B., Fonoll, X., Norton, J., Mehrotra, A., and Xagorarakis, I. (2020). SARS-CoV-2 in Detroit Wastewater. *J Environ Eng* 146, 06020004. [https://doi.org/10.1061/\(asce\)ee.1943-7870.0001830](https://doi.org/10.1061/(asce)ee.1943-7870.0001830)

Myrmel, M., Lange, H., and Rimstad, E. (2015). A 1-Year Quantitative Survey of Noro-, Adeno-, Human Boca-, and Hepatitis E Viruses in Raw and Secondarily Treated Sewage from Two Plants in Norway. *Food Environ Virol* 7, 213–223. <https://doi.org/10.1007/s12560-015-9200-x>

Nemudryi, A., Nemudraia, A., Surya, K., Wiegand, T., Buyukyoruk, M., Wilkinson, R., and Wiedenheft, B. (2020). Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Medrxiv Prepr Serv Heal Sci*. <https://doi.org/10.1101/2020.04.15.20066746>

Park, S., Lee, C.-W., Park, D.-I., Woo, H.-Y., Cheong, H.S., Shin, H.C., Ahn, K., Kwon, M.-J., and Joo, E.-J. (2020). Detection of SARS-CoV-2 in Fecal Samples from Patients with Asymptomatic and Mild COVID-19 in Korea. *Clin Gastroenterol H*. <https://doi.org/10.1016/j.cgh.2020.06.005>

Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., Warren, J.L., Weinberger, D.M., and Omer, S.B. (2020). SARS-CoV-2 RNA Concentrations in Primary Municipal Sewage Sludge as a Leading Indicator of COVID-19 Outbreak Dynamics. medRxiv preprint. <https://doi.org/10.1101/2020.05.19.20105999>

Pecson, B.M., Darby, E., Haas, C.N., Amha, Y.M., Bartolo, M., Danielson, R., Dearborn, Y., Giovanni, G.D., Ferguson, C., Fevig, S., Gaddis, E., Gray, D., Lukasik, G., Mull, B., Olivas, L., Olivieri, A., Qu, Y., and Consortium, SARS-CoV-2 Interlaboratory (2021). Reproducibility and Sensitivity of 36 Methods to Quantify the SARS-CoV-2 Genetic Signal in Raw Wastewater: Findings From an Interlaboratory Methods Evaluation in the U.S. *Environ Sci Water Res Technology* 7, 504–520. <https://doi.org/10.1039/d0ew00946f>

Polaczyk, A.L., Narayanan, J., Cromeans, T.L., Hahn, D., Roberts, J.M., Amburgey, J.E., and Hill, V.R. (2008). Ultrafiltration-Based Techniques for Rapid and Simultaneous Concentration of Multiple Microbe Classes from 100-L Tap Water Samples. *J Microbiol Meth* 73, 92–99. <https://doi.org/10.1016/j.mimet.2008.02.014>

Prado, T., Fumian, T.M., Mannarino, C.F., Maranhão, A.G., Siqueira, M.M., and Miagostovich, M.P. (2020). Preliminary Results of SARS-CoV-2 Detection in Sewerage System in Niterói Municipality, Rio de Janeiro, Brazil. *Memórias Instituto Oswaldo Cruz* 115, e200196. <https://doi.org/10.1590/0074-02760200196>

Randazzo, W., Cuevas-Ferrando, E., Sanjuán, R., Domingo-Calap, P., and Sánchez, G. (2020a). Metropolitan Wastewater Analysis for COVID-19 Epidemiological Surveillance. *Int J Hyg Envir Heal* 230, 113621. <https://doi.org/10.1016/j.ijheh.2020.113621>

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., and Sánchez, G. (2020b). SARS-CoV-2 RNA in Wastewater Anticipated COVID-19 Occurrence in a Low Prevalence Area. *Water Res* 181, 115942. <https://doi.org/10.1016/j.watres.2020.115942>

Rimoldi, S.G., Stefani, F., Gigantiello, A., Polesello, S., Comandatore, F., Mileto, D., Maresca, M., Longobardi, C., Mancon, A., Romeri, F., Pagani, C., Moja, L., Gismondo, M.R., and Salerno, F. (2020). Presence and Vitality of SARS-CoV-2 Virus in Wastewaters and Rivers. medRxiv preprint. <https://doi.org/10.1101/2020.05.01.20086009>

Rosa, G.L., Bonadonna, L., Lucentini, L., Kenmoe, S., and Suffredini, E. (2020a). Coronavirus in Water Environments: Occurrence, Persistence and Concentration Methods - A Scoping Review. *Water Res* 179, 115899. <https://doi.org/10.1016/j.watres.2020.115899>

Rosa, G.L., Iaconelli, M., Mancini, P., Ferraro, G.B., Veneri, C., Bonadonna, L., Lucentini, L., and Suffredini, E. (2020b). First Detection of SARS-CoV-2 in Untreated Wastewaters in Italy. *Sci Total Environ* 736, 139652. <https://doi.org/10.1016/j.scitotenv.2020.139652>

Rosa, G.L., Mancini, P., Ferraro, G.B., Veneri, C., Iaconelli, M., Bonadonna, L., Lucentini, L., and Suffredini, E. (2020). SARS-CoV-2 has been Circulating in Northern Italy since December

2019: Evidence from Environmental Monitoring. medRxiv preprint.
<https://doi.org/10.1101/2020.06.25.20140061>

Sharif, S., Ikram, A., Khurshid, A., Salman, M., Mehmood, N., Arshad, Y., Ahmad, J., Angez, M., Alam, M.M., Rehman, L., Mujtaba, G., Hussain, J., Ali, J., Akthar, Ri., Malik, M.W., Baig, Z.I., Rana, M.S., Usman, M., Ali, M.Q., Ahad, A., Badar, N., Umair, M., Tamim, S., Ashraf, A., Tahir, F., and Ali, N. (2020). Detection of SARS-Coronavirus-2 in Wastewater, using the Existing Environmental Surveillance Network: An Epidemiological Gateway to an Early Warning for COVID-19 in Communities. medRxiv preprint.
<https://doi.org/10.1101/2020.06.03.20121426>

Sherchan, S.P., Shahin, S., Ward, L.M., Tandukar, S., Aw, T.G., Schmitz, B., Ahmed, W., and Kitajima, M. (2020). First Detection of SARS-CoV-2 RNA in Wastewater in North America: A Study in Louisiana, USA. *Sci Total Environ* 743, 140621.
<https://doi.org/10.1016/j.scitotenv.2020.140621>

Shieh, Y.-S.C., Wait, D., Tai, L., and Sobsey, M.D. (1995). Methods to Remove Inhibitors in Sewage and Other Fecal Wastes for Enterovirus Detection by the Polymerase Chain Reaction. *J Virol Methods* 54, 51–66. [https://doi.org/10.1016/0166-0934\(95\)00025-p](https://doi.org/10.1016/0166-0934(95)00025-p)

Thongprachum, A., Fujimoto, T., Takanashi, S., Saito, H., Okitsu, S., Shimizu, H., Khamrin, P., Maneekarn, N., Hayakawa, S., and Ushijima, H. (2018). Detection of Nineteen Enteric Viruses in Raw Sewage in Japan. *Infect Genetics Evol* 63, 17–23.
<https://doi.org/10.1016/j.meegid.2018.05.006>

Trottier, J., Darques, R., Mouheb, N.A., Partiot, E., Bakhache, W., Deffieu, M.S., and Gaudin, R. (2020). Post-Lockdown Detection of SARS-CoV-2 RNA in the Wastewater of Montpellier, France. *One Heal* 10, 100157. <https://doi.org/10.1016/j.onehlt.2020.100157>

Vallejo, J.A., Rumbo-Feal, S., Conde-Pérez, K., López-Oriona, Á., Tarrío-Saavedra, J., Reif, R., Ladra, S., Rodiño-Janeiro, B.K., Nasser, M., Cid, Á., Veiga, M.C., Acevedo, A., Lamora, C., Bou, G., Cao, R., and Poza, M. (2020). Predicting the Number of People Infected with SARS-COV-2 in a Population using Statistical Models Based on Wastewater Viral Load. medRxiv preprint. <https://doi.org/10.1101/2020.07.02.20144865>

Vidaver, A.K., Koski, R.K., and Etten, J.L.V. (1973). Bacteriophage $\phi 6$: a Lipid-Containing Virus of *Pseudomonas Phaseolicola* 1. *J Virol* 11, 799–805. <https://doi.org/10.1128/jvi.11.5.799-805.1973>

Wang, J., Feng, H., Zhang, S., Ni, Z., Ni, L., Chen, Y., Zhuo, L., Zhong, Z., and Qu, T. (2020). SARS-CoV-2 RNA Detection of Hospital Isolation Wards Hygiene Monitoring During the Coronavirus Disease 2019 Outbreak in a Chinese hospital. *Int J Infect Dis* 94, 103–106.
<https://doi.org/10.1016/j.ijid.2020.04.024>

Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J.V., and LaCross, N. (2021).

Correlation of SARS-CoV-2 RNA in Wastewater with COVID-19 Disease Burden in Sewersheds. *Sci Total Environ* 775, 145790. <https://doi.org/10.1016/j.scitotenv.2021.145790>

Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M.A., Bushman, M., Chai, P.R., Duvallet, C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P., Huang, K.H., Lee, W.L., Matus, M., McElroy, K.A., Nagler, J., Rhode, S.F., Santillana, M., Tucker, J.A., Wuertz, S., Zhao, S., Thompson, J., and Alm, E.J. (2020). SARS-CoV-2 Titers in Wastewater Foreshadow Dynamics and Clinical Presentation of New COVID-19 Cases. *Medrxiv Prepr Serv Heal Sci*. <https://doi.org/10.1101/2020.06.15.20117747>

Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y., Qu, X., Kuang, L., Fang, X., Mishra, N., Lu, J., Shan, H., Jiang, G., and Huang, X. (2020). Prolonged Presence of SARS-CoV-2 Viral RNA in Faecal Samples. *Lancet Gastroenterology Hepatology* 5, 434–435. [https://doi.org/10.1016/s2468-1253\(20\)30083-2](https://doi.org/10.1016/s2468-1253(20)30083-2)

Wurtzer, S., Marechal, V., Mouchel, J., Maday, Y., Teyssou, R., Richard, E., Almayrac, J., and Moulin, L. (2020). Evaluation of Lockdown Impact on SARS-CoV-2 Dynamics Through Viral Genome Quantification in Paris Wastewaters. *medRxiv preprint*. <https://doi.org/10.1101/2020.04.12.20062679>

Ye, Y., Ellenberg, R.M., Graham, K.E., and Wigginton, K.R. (2016). Survivability, Partitioning, and Recovery of Enveloped Viruses in Untreated Municipal Wastewater. *Environ Sci Technol* 50, 5077–85. <https://doi.org/10.1021/acs.est.6b00876>

Zhang, D., Ling, H., Huang, X., Li, J., Li, W., Yi, C., Zhang, T., Jiang, Y., He, Y., Deng, S., Zhang, X., Liu, Y., Li, G., and Qu, J. (2020). Potential Spreading Risks and Disinfection Challenges of Medical Wastewater by the Presence of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Viral RNA in Septic Tanks of Fangcang Hospital. *medRxiv preprint*. <https://doi.org/10.1101/2020.04.28.20083832>

Zhang, D., Yang, Y., Huang, X., Jiang, J., Li, M., Zhang, X., Ling, H., Li, J., Liu, Y., Li, G., Li, W., Yi, C., Zhang, T., Jiang, Y., Xiong, Y., Hu, Z., Wang, X., Deng, S., Zhao, P., and Qu, J. (2020). SARS-CoV-2 Spillover into Hospital Outdoor Environments. *medRxiv preprint*. <https://doi.org/10.1101/2020.05.12.20097105>

Zhou, J.-B., Kong, W.-H., Wang, S., Long, Y.-B., Dong, L.-H., He, Z.-Y., and Liu, M.-Q. (2020). Potential Transmission Risk of SARS-CoV-2 Through Medical Wastewater in COVID-19 Outbreak Cities. *medRxiv preprint*. <https://doi.org/10.21203/rs.3.rs-3743/v1>

Zimmermann, K., Scheibe, O., Kocourek, A., Muelich, J., Jurkiewicz, E., and Pfeifer, A. (2011). Highly Efficient Concentration of Lenti- and Retroviral Vector Preparations by Membrane Adsorbers and Ultrafiltration. *Bmc Biotechnol* 11, 55. <https://doi.org/10.1186/1472-6750-11-55>

4.0 Understanding the Efficacy of Wastewater Surveillance for SARS-CoV-2 in Two Diverse Communities

4.1 Abstract

During the COVID-19 pandemic wastewater-based epidemiology (WBE) has been shown to be a useful tool for monitoring the spread of disease in communities and the emergence of new viral variants of concern. As the pandemic enters its third year and clinical testing has declined, WBE offers a consistent non-intrusive way to monitor community health in the long term. However, understanding the best method for the application of WBE in different communities is necessary. This study sought to understand how accurately wastewater monitoring represented the actual burden of disease between communities. Two communities varying in size and demographics in Michigan were monitored for SARS-CoV-2 in wastewater between March of 2020 and February of 2022. Additionally, one community was monitored for SARS-CoV-2 variants of concern from December 2020 to February 2022. Wastewater results were compared with zipcode and county level COVID-19 case data to determine which scope of clinical surveillance was most correlated with wastewater loading. Pearson r correlations were highest in the smaller of the two communities ($r = 0.45-0.81$) with the highest correlations with zipcode level case data. When comparing the date of cases being reported against the date of the onset of symptoms, the smaller community was more highly correlated with the onset date (onset: $r = 0.68-0.81$ vs. referral: $r = 0.38-0.48$), while the larger community showed little variation ($r = 0.62-0.68$). This study has demonstrated that wastewater surveillance in different communities are linked to different geographic and temporal scales.

4.2 Introduction

As the COVID-19 global pandemic enters its third year, the surveillance of SARS-CoV-2, the etiological agent of COVID-19, has begun to shift to less intrusive methods. Wastewater-

based epidemiology (WBE) in particular, has shown its usefulness as large numbers of viral particles are shed in the feces of infected individuals including symptomatic, asymptomatic, and pre-symptomatic persons (Parasa et al., 2020; Wang et al., 2020; Zheng et al., 2020; Lee et al., 2020). The ability of molecular detection techniques to identify and quantify the viral RNA of SARS-CoV-2 in raw wastewater along with the ability to detect spikes in cases prior to the identification of clinical cases is invaluable as the pandemic continues (Peccia et al., 2020; Medema et al., 2020). A number of studies have utilized WBE to track the progress of COVID-19 in communities, and groups have used the sewer to focus on a single building, local area, or wastewater at the treatment facility to represent a city or county geographic scale (Fahrnfeld et al., 2022; Rasero et al., 2022; Lastra et al., 2022).

While previous studies have shown that SARS-CoV-2 levels in wastewater correlate with COVID-19 cases, there has been very little comparative analysis of the wastewater signal across the various communities (Gonzalez et al., 2020; Peccia et al., 2020; Gerrity et al., 2021; Graham et al., 2021). Understanding the how differences in community size and wastewater treatment impact SARS-CoV-2 wastewater results is necessary to properly apply WBE on a wider scale. It is important to understand how wastewater SARS levels reflect the disease and address the impact of new variants and use of vaccinations as clinical testing declines (Martin et al., 2020; Smith et al., 2021).

The goal of this study was to determine how well wastewater surveillance for SARS-CoV-2 addresses the cases of disease in different communities. For this purpose, two communities in Michigan were selected for comparison. These communities vary in population size, demographics, and total numbers of cases of COVID-19 over the course of the pandemic. This study had three main objectives: 1) to evaluate the efficacy of wastewater monitoring of

SARS-CoV-2 in two communities with diverse characteristics; 2) to determine if county or zipcode level case data are necessary to successfully correlate with wastewater surveillance results; 3) to determine the impact of vaccination rates on SARS-CoV-2 wastewater signals compared to case numbers; and 4) examine the occurrence and appearance of new variants in sewage during the waves of COVID-19 in one community.

4.3 Materials and Methods

4.3.1 Wastewater sampling and site descriptions

4.3.1.1 Wastewater treatment plant descriptions

Two communities and their corresponding wastewater treatment plants were selected for sampling and comparison. Wastewater treatment plant B treats wastewater from a city and two surrounding townships within a single county. The WWTP B serves a population of 25,000 persons with an average flow of 2.3 million gallons per day (MGD). Wastewater treatment plant A serves 31 communities, with 25 within its primary county and six others in surrounding counties. The WWTP A serves a population of 110,267 persons with an average flow of 27 MGD. While both the WWTP B and WWTP A use conventional activated sludge followed by disinfection the WWTP A is an approved blending facility which handles wet weather induced inflow. This potentially increases a dilution factor for wastewater during wet weather events.

Table 4.1 County level demographics, COVID-19 vaccinations, and total COVID-19 cases/1,000 persons to date

	County A	County B
Total Population by County (Total Population by Zipcode^a)	405,813 (110,267)	66,699 (25,000)
Population Density (People per sq. mile)	637.13	36.87
Household Size	2.41	2.39
Percent Living in Poverty	19.8	16.4
Percent of Population >65 years	17.97	19.62
Ratio of Male to Female Population	48.2 : 51.8	50.3 : 49.7
Ratio of White to Non-white persons	75.3 : 24.7	93.2 : 6.8
Per Capita Income (2020)	\$46,152	\$44,445
County Level GDP (Thousands of Current Dollars)	16,121,115	2,787,951
Percent Fully Vaccinated as of (3/1/22)	50.2	63.6
Total Number of COVID-19 Cases/1,000 persons as of 3/1/22	247	235
Total Number of COVID-19 Deaths as of 3/1/22	1,692	126
Total Number of COVID-19 Deaths/ 1,000 persons as of 3/1/22	4.2	1.9

^aZipcodes served by WWTP; Sources: US CDC, 2022; BEA, 2022; US Census Bureau, 2022

4.3.1.2 Sample Collection Methods

Wastewater samples for this study were collected over a 24 hr period at the inflows after the primary grit removal of each WWTP. The WWTP B collected composite samples based on their expected daily flow with approximately 65 ml being collected for every 58,000 gallons of wastewater entering the plant for a total of ~2500 ml for a 24 hr period. WWTP A collected composite samples based on a time paced approach collecting 100 ml every 30 mins over a 24 hr period. A total of 1 L of wastewater was then transported to the processing laboratory on ice. A total of 186 samples were collected from WWTP A (N=92) and WWTP B (N=94) between April 2020 and February 2022 at a frequency of once per week. Between April 2020 and January 2021, the samples from WWTP B were shipped overnight on ice to Michigan State University.

Between February 2021 and December 2021, the samples from WWTP B were driven on to Northern Michigan University for processing. A gap in sampling occurred for both WWTPs between January/February and May/July 2021 due to the ending of one project funding and the start of another. All samples from the ARTP were shipped on ice overnight to Michigan State University (April 2020- December 2021). Physiological measurements including temperature, pH, biological oxygen demand (BOD), and total suspended solids (TSS) were taken at the time of sampling by each WWTP's onsite laboratory (Table 2). Turbidity was measured upon arrival at the processing laboratory. Samples collected between April 2020 and October 2020 kept frozen at -80°C until analysis. All samples collected after October 25th, 2020 were kept at 4°C, never frozen and were processed within 72 hours of collection. This change between storage temperatures was due to evidence that the SARS-CoV-2 RNA signal declined in the raw wastewater samples after they had been frozen.

4.3.2 Viral concentration and processing methods

Wastewater samples were processed, and viral particles were concentrated using the polyethylene glycol (PEG) workflow published by Flood et al. (2021). Briefly, samples were inverted to mix 25 times then 100 ml of sample was transferred to a 250 ml polypropylene centrifuge bottle. A total of 8 g of 8% (w/vol) molecular grade PEG 8000 (Promega Corporation, Madison Wisconsin) and 1.17 g NaCl (0.2 M w/v) were added to each sample. The samples were then slowly mixed on magnetic stir plates for 2 hours at 4°C. Samples were either held at 4°C for 16 hrs or immediately transferred to the centrifuge. Samples were centrifuged at 4,700 x g at 4°C for 45 mins. Following centrifugation, the majority of the supernatant was removed, and the remaining pellet was resuspending in the remaining supernatant (2-10 ml). Sample concentrates

were aliquoted and either immediately underwent RNA extraction or were stored at -80°C until further processing.

Viral ribonucleic acid (RNA) was extracted using the QIAmp Viral RNA Minikit (Qiagen, Germany) according to the manufacturers protocol. A total of 200 µl of concentrate was used for each RNA extraction with a final elution volume of 80 µl.

4.3.3 Detection and enumeration of SARS-CoV-2 from wastewater using RT-ddPCR

All genetic targets were analyzed using one-step reverse transcriptase droplet digital PCR. Two general SARS-CoV-2 nucleocapsid 1 (N1) and nucleocapsid 2 (N2) gene targets were analyzed for all samples. The *Pseudomonas* bacteriophage Phi6 was spiked into all samples as either a recovery efficiency control or an inhibition control. The primer and probe sequences for the N1, N2, and Phi6 gene targets are shown in Table 1. Samples from WWTP A were analyzed for genetic markers for SARS-CoV-2 variants of concern starting in December of 2020 using GT Molecular's variant assay kits for digital PCR (GT Molecular, Fort Collins, Colorado, USA). These variants included the Alpha variant (gene targets N501Y and DEL69-70), the Delta variant (gene targets T478K and L452R), and the Omicron variant (gene targets N501Y, DEL69-70, and K417N). The variant assays used the same thermocycling setup as the Phi6 assay. All analyses were run with three technical replicates and a full contingent of quality controls (positive, negative, extraction negative, and non-template controls) on each assay plate.

Droplet digital PCR was performed with a Bio-Rad QX200 ddPCR (Bio-Rad, CA, USA). All assays in this study used the 1-step RT-ddPCR Advanced kit for probes (Bio-Rad, CA, USA) for all ddPCR reaction mixtures. The N1, N2, and Phi6 gene target reaction mixtures all contained a final concentration of 1x Supermix (Bio-Rad, CA, USA), 20 U µl⁻¹ of reverse

transcriptase (RT) (Bio-Rad, CA, USA), 15 mM DTT, 900 nmol of each primer, 250 nmol of each probe. The N1 and N2 gene targets were run in duplex. A total of 5 μ l of sample RNA template was analyzed in technical triplicates for each assay (including each variant assay). The variant assays were run per the manufacturer's protocols.

Droplet generation by microfluidic mixing was performed in a Bio-Rad Automatic Droplet Generator (ADG) (Bio-Rad, CA, USA). Each 20 μ l reaction mixture was combined with 70 μ l of droplet generation oil which resulted in a final volume 40 μ l of reaction mixture-oil emulsions. These emulsions contained up to 20,000 individual oil droplets. After droplet generation the 96-well PCR plates were heat-sealed with foil and placed in a C1000 96-deep well thermocycler (Bio-Rad, CA, USA) for PCR product amplification. The N1 and N2 assay followed the following thermocycling parameters: 25°C for 3 min, 50°C for 1 hr, 95°C for 10 min, followed by 40 cycles of 95°C for 30 s and 55°C for 1 min with ramp rate of 2°C s⁻¹ followed by a final cycle of 98°C for 10 min. The Phi6 and variant assays followed the same thermocycling parameters except their annealing temperature was set to 60°C. After thermocycling was completed the sealed 96-well plates were transferred to the QX200 droplet reader (Bio-Rad, CA, USA) for analysis of the samples' droplets fluorescent probe signals by spectrophotometric detection.

Table 4.2 Primer and probe sequences

Target	Primer/Probe name	Primer/Probe Sequence	Reference
SARS CoV-2	2019-nCoV_N1-F	5'-GACCCCAAAATCAGCGAAAT-3'	CDC, 2020
	2019-nCoV_N1-R	5'-TCTGGTTACTGCCAGTTGAATCTG-3'	
	2019-nCoV_N1-P	5'-FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1-3'	
	2019-nCoV_N2-F	5'-TTACAAACATTGGCCGCAAA-3'	CDC, 2020
	2019-nCoV_N2-R	5'-GCGCGACATTCCGAAGAA-3'	
	2019-nCoV_N2-P	5'-HEX-ACAATTTGCCCCAGCGCTTCAG-BHQ1-3'	
Phi6	Φ6Tfor Φ6Trev Φ6Tprobe	5'-TGGCGGCGGTCAAGAGC-3' 5'-GGATGATTCTCCAGAAGCTGCTG-3' 5'- FAM-CGGTCGTCGTCAGGTCTGACACTCGC-BHQ1-3'	Gendron et al., 2010

4.3.4 COVID-19 case and vaccination data

Data for COVID-19 cases were procured for both zipcode and county levels. Zipcode level case data were provided through an agreement with the Michigan Department of Health and Human Services. Zipcodes serviced by each wastewater treatment plant were provided by plant operators. In the event of missing data for the onset of symptoms, an estimate of onset date was used based on an average of all data with known information. This was calculated by averaging the number of days between onset of symptoms and referral dates for paired data points over the course of the study. The average number of days between onset and referral date was 6.03 days (N= 40,348) for the combined datasets (community A + B). The average number of days between onset and referral date for each community alone were 6.04 days for A and 5.19 days for B with both ranging from 0 to 100 days.

County level case data were obtained from the US Centers for Disease Control and Prevention's (US CDC) COVID Data Tracker website (<https://data.cdc.gov/Public-Health-Surveillance/United-States-COVID-19-Community-Levels-by-County/3nnm-4jni>). COVID-19 vaccination data were obtained from the US CDC's COVID Data Tracker website (<https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh>).

4.3.5 Data analysis

All ddPCR results were converted from gene copies (GC) per reaction (5 µl of sample template) to GC/100 ml prior to analysis as described in Flood et al. 2021. Following conversion to GC/100 ml wastewater results were then normalized for each community based on daily wastewater flows and zipcode level population. Non-detects (ND) replicates were included in statistical analysis results were assigned their lower limits of detection for statistical analysis.

Data visualization and statistical analysis were performed using Graphpad Prism 9 (Graphpad Software, CA, USA). Correlation analyses were performed using pearson correlation (r) analysis. Correlation analyses were compared for results between both community's wastewater results and case data, between wastewater results with zipcode specific and county level cases data, and vaccination rates and case data. To account for lag time between the wastewater signal and cases, both the date of symptom onset and the date of case referral were analyzed against the wastewater signal.

4.4 Results

4.4.1 Comparison of SARS-CoV-2 concentrations found in wastewater against COVID-19 case data in two communities

The data gathered during this study showed that the two wastewater treatment plants (A and B) had distinctly different characteristics (Table 4.3). WWTP A had approximately 10 times the average daily flow (27.39 million gallons per day, MGD) compared to WWTP B which had an average flow of 2.79 MGD. Sample temperatures ranged from 7.4 to 22.6°C for WWTP A and samples from WWTP B ranged from 8.90 to 26.67°C. While WWTP A had a slightly lower average BOD5 levels than WWTP B (161.90 and 206.17, respectively) higher turbidities were

observed at WWTP A (WWTP A: 80.83 vs. WWTP B: 59.32). Wastewater N1 and N2 gene targets average concentrations for SARS-CoV-2 were similar between the two WWTPs (WWTP A (N=94) N1 3.94, N2 3.86; WWTP B (N=92) N1 3.96, N2 3.94 Log₁₀GC/ 100ml) (Table 4.4). However, as expected the loading as calculated by daily average flow at each of the WWTPs and adjusted for population showed that the larger WWTP A had more than twice as much virus (84.06 N1 gene copies per person per day) compared to WWTP B (38.93 GC/Person/Day) and nearly double for the N2 gene as well (69.06 vs. 38.23 GC/Person/Day).

Table 4.3 Physiological measurements for two wastewater treatment plants

WWTP	Estimated Population Served by Zipcode	Flow Rate (MGD)	Temperature (°C)	pH	BOD5 (mg/L)	TSS (mg/L)	Turbidity (NTU)
A	110,267	27.39 (21.10-55.68)	14.81 (7.4-22.6)	7.61 (7.28-7.97)	161.90 (64.0-370.0)	194.00 (90.0-526.0)	80.83 (26.9-158)
B	25,000	2.79 (2.06-4.33)	14.11 (8.90-26.67)	7.29 (7.0-7.7)	206.17 (79.0-341.0)	197.52 (99.0-364.0)	59.32 (17.4-152.0)

Note: A gap in sampling occurred for both WWTPs between January/February and May/July 2021 due to the ending of one project funding and the start of another

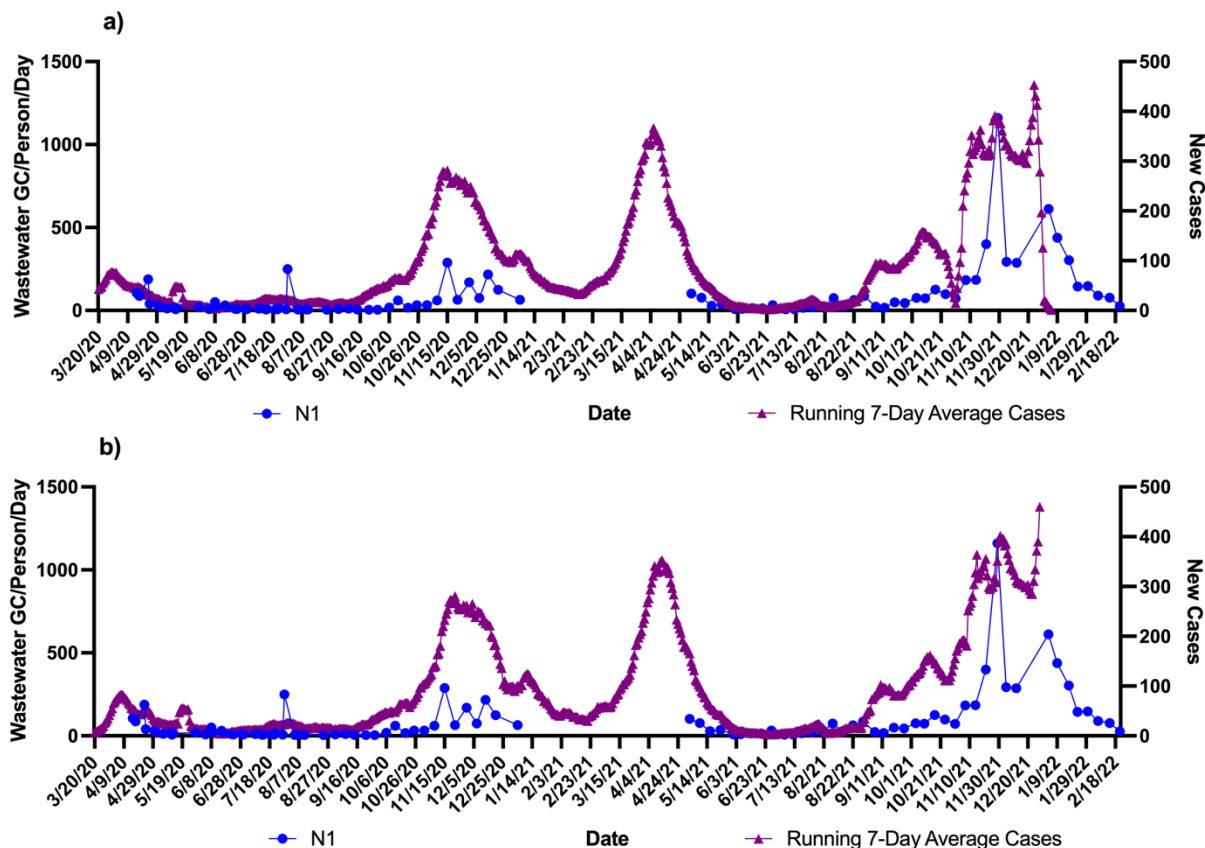
Table 4.4 Summary of wastewater monitoring results for two wastewater treatment plants

		N1 Log₁₀GC/ 100 ml	N2 Log₁₀GC/ 100 ml	N1 GC/Person/ Day	N2 GC/Person/ Day
WWTP A (N=94)	Percent Positive	70.21% (66/94)	68.09% (64/94)	70.21% (66/94)	68.09% (64/94)
	Mean ^a	3.94	3.86	84.06	69.06
	(Range)	(2.70-5.07 ^b)	(2.57-5.00 ^b)	(3.83-1160.17)	(3.83-983.20)
WWTP B (N= 92)	Percent Positive	72.82% (67/92)	77.17% (71/92)	72.82% (67/92)	77.17% (71/92)
	Mean ^a	3.96	3.94	38.93	38.23
	(Range)	(2.78-4.99 ^c)	(2.78-4.95 ^c)	(2.82-341.88)	(2.38-319.96)

^aArithmetic means; ^bDate of peak concentration for WWTP A was 11/29/21; ^cDate of peak concentration for WWTP B was 1/20/21; Note: A gap in sampling occurred for both WWTPs between January/February and May/July 2021 due to the ending of one project funding and the start of another.

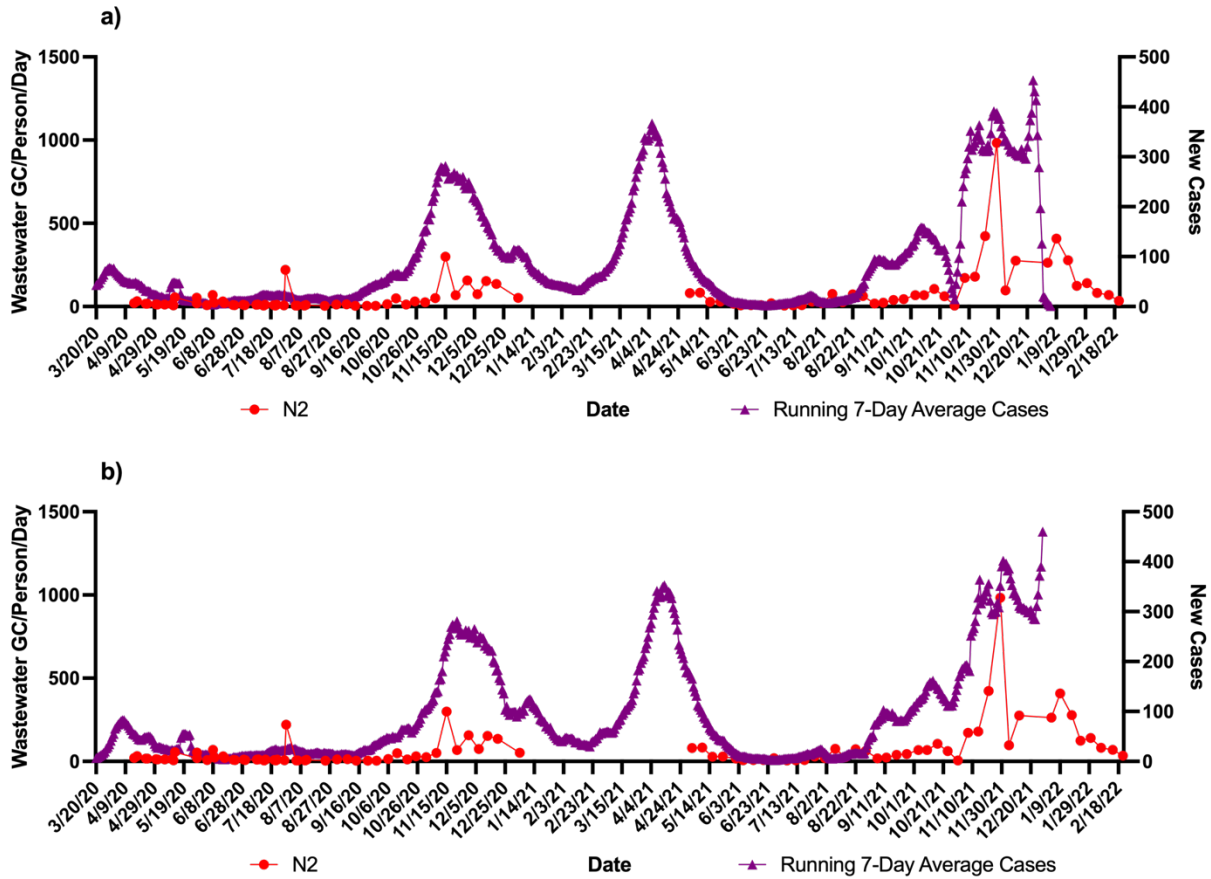
Figures 4.1-4.4 show the results of wastewater surveillance of SARS-CoV-2 graphed with the running 7-day average zipcode level case data comparing the onset of symptoms date for each community versus the date of referral for WWTP A N1; N2 and WWTP B N1; N2. A gap in wastewater data between January/February and May/July 2021 was due to the ending of one project funding and the start of another. Wastewater loading from both communities followed the same trends in case data consistent with the waves of COVID-19 cases in Michigan during the pandemic. The N1 gene results for community A had slightly higher correlation with the referral date (Figure 4.1b: $r = 0.68$ $p < 0.0001$) compared to the onset date (Figure 4.1a: $r = 0.62$ $p < 0.0001$). The N2 gene results correlations for community A were almost identical between the onset and referral dates (Figure 4.2a: onset $r = 0.68$ $p < 0.0001$; Figure 4.2b: referral $r = 0.67$ $p < 0.0001$). However, a larger difference in correlations was observed with community B (Figures 4.3 and 4.4). The N1 gene results were more highly correlated with the onset date (Figure 4.3a: $r = 0.81$ $p < 0.0001$) compared to the referral date (Figure 4.3b: $r = 0.48$ $p < 0.0001$).

This same pattern was seen with the N2 gene results as well with the onset date (Figure 4.4a) showing a correlation of $r = 0.68$ while the referral date (Figure 4.4b) was only $r = 0.38$.



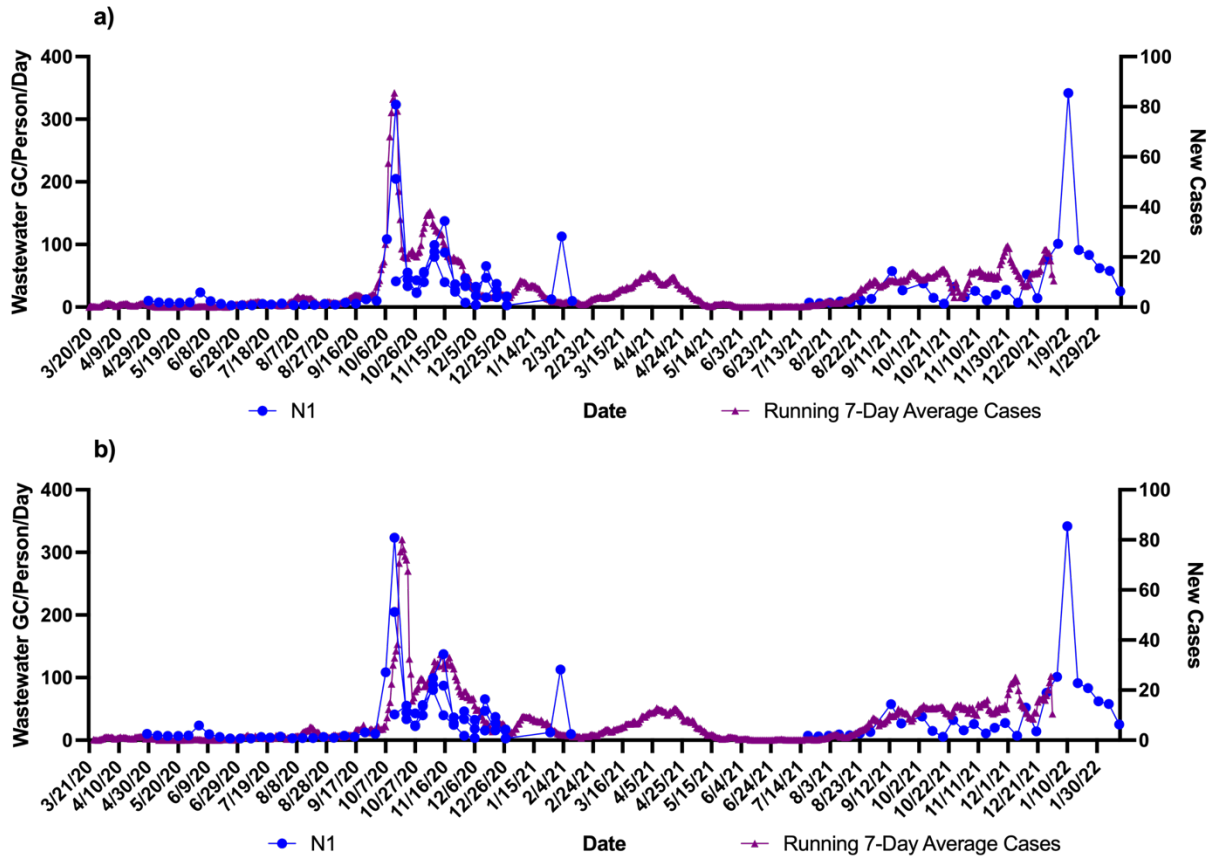
^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

Figure 4.1 Wastewater surveillance data (N1 gene target) for WWTP A (N=94) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N1 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.62$ $p < 0.0001$; $n = 86$ paired data points); b) N1 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 85$ paired data points).



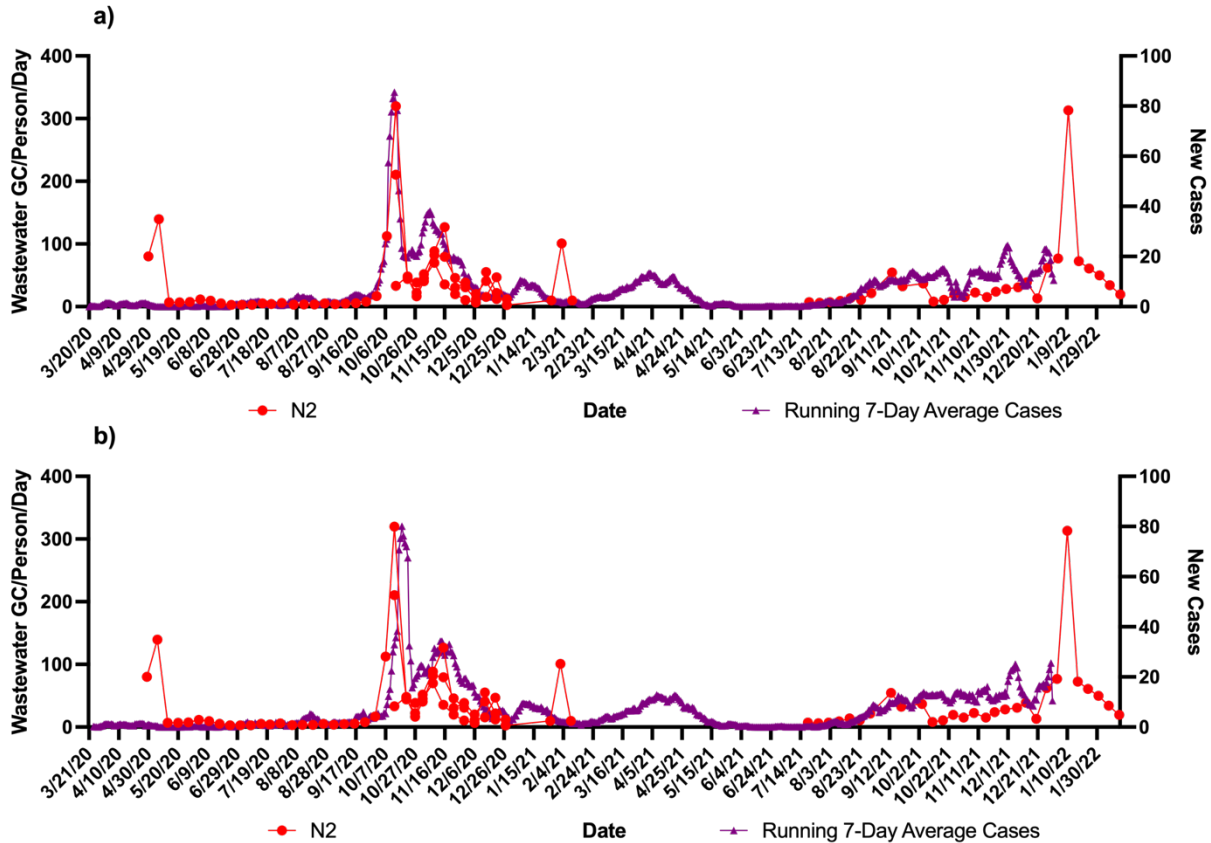
^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

Figure 4.2 Wastewater surveillance data (N2 gene target) for WWTP A (N=94) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N2 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 86$ paired data points); b) N2 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.67$ $p < 0.0001$; $n = 85$ paired data points).



^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

Figure 4.3 Wastewater surveillance data (N1 gene target) for WWTP B (N=92) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N1 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.81$ $p < 0.0001$; $n = 61$ paired data points); b) N1 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.48$ $p < 0.0001$; $n = 61$ paired data points).



^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

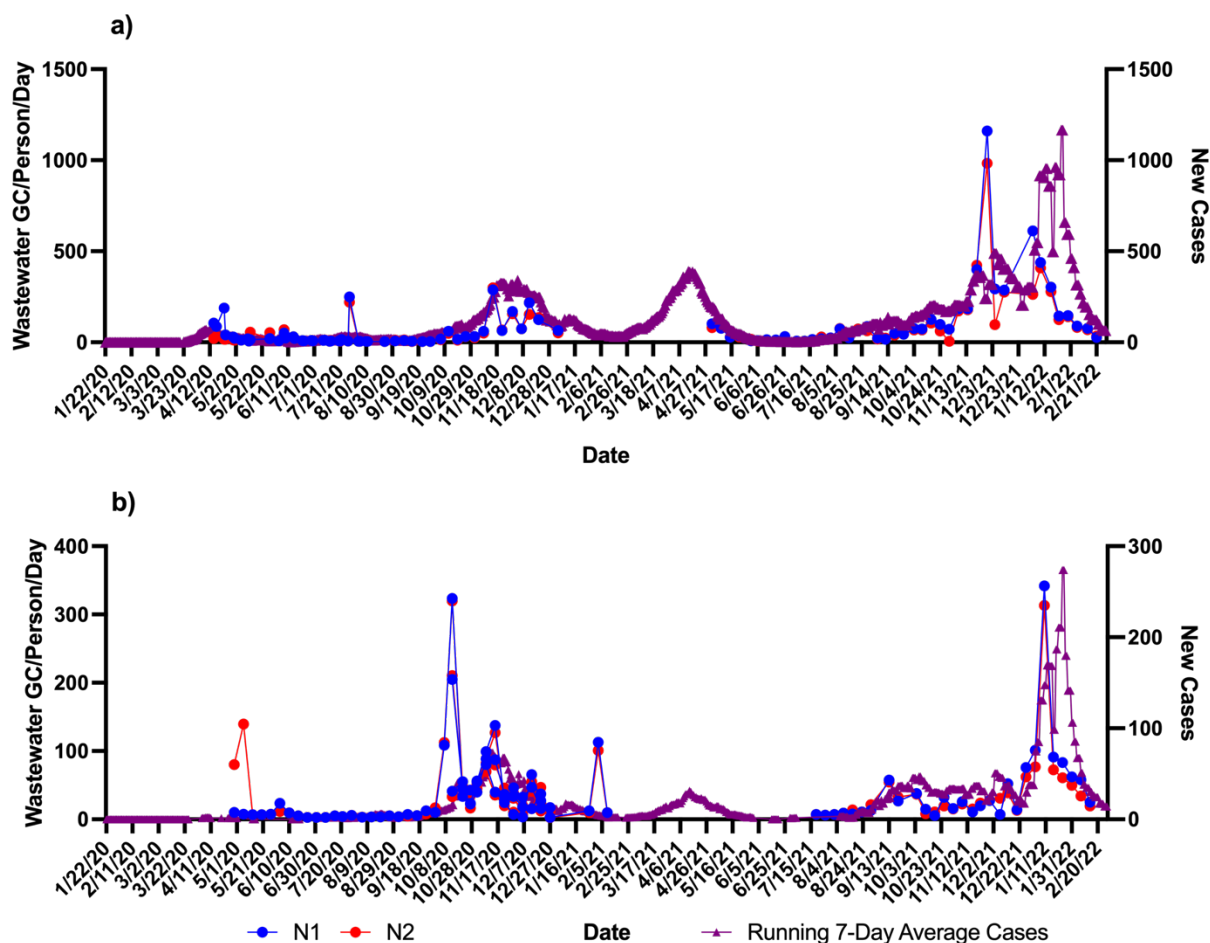
Figure 4.4 Wastewater surveillance data (N2 gene target) for WWTP B (N=92) (GC/Person/Day) and COVID-19 zipcode case data over time. a) N2 vs. case onset of symptoms running 7-day average case data for COVID-19 ($r = 0.68$ $p < 0.0001$; $n = 61$ paired data points); b) N2 vs. referral date for running 7-day average case data for COVID-19 ($r = 0.38$ $p < 0.0001$; $n = 61$ paired data points).

4.4.1.2 Zipcode vs county level case data varying spatial resolution

When comparing the two communities with county level case data the two communities showed similar pearson correlation values of approximately 0.5 (WWTP A: N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.53$ $p < 0.0001$; $n = 93$ paired data points; WWTP B N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.45$ $p < 0.0001$; $n = 58$ paired data points) (Figure 4.5). It is important to note that the discrepancies in the total paired data points and the paired data points in the county level data

was due the presence of censored data for multiple dates in the US CDC database. The zipcode level case data represented 25% of county level data for community A and 37.5% for community B.

Correlations between wastewater loading and case data were compared between communities. The N1 and N2 results for WWTP A with zipcode level referral date case data had pearson r correlation values of 0.68 ($p < 0.0001$) and 0.67 ($p < 0.0001$) while the county level case data had r values of 0.52 ($p < 0.0001$) and 0.53 ($p < 0.0001$) (Figures 4.1b, 4.2b and 4.5a). WWTP B showed greater differences in correlations between zipcode referral dates and county level case data. N1 and N2 results had r values of 0.48 ($p < 0.0001$) and 0.38 ($p < 0.0001$) for zipcode level referral date data compared to only 0.52 ($p < 0.0001$) and 0.45 ($p < 0.0001$) for county level case data (Figures 4.3b, 4.4b, and 4.5b).



^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

Figure 4.5 Wastewater surveillance data (N=94) (adjusted by flow and zipcode level population) and county level COVID-19 case data over time. a) WWTP A SARS-CoV-2 gene target results vs. county level case data for COVID-19 (N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.53$ $p < 0.0001$; $n = 93$ paired data points); b) WWTP B SARS-CoV-2 gene target results vs. county level COVID-19 case data (N1 $r = 0.52$ $p < 0.0001$, N2 $r = 0.45$ $p < 0.0001$; $n = 58$ paired data points).

4.4.2 Impact of vaccination rates on SARS-CoV-2 wastewater signals and case numbers

In this study, the percent vaccination rate at the county level were graphed per day. Vaccination data in this case was for that population fully vaccinated (two doses) for the two counties served by WWTP A and B. The first reported data point for vaccination rate was in

December of 2020. Both counties had rapid increases in vaccination over the following six months (Figure 4.6). However, after June of 2021 vaccination rates drastically declined and has had not significantly increased since then with both counties almost plateauing near 60% of the total population fully vaccinated. While Community B had lower cases/ 1000 persons than Community A vaccinations began there (Community B) almost two months before Community A (Figure 4.6). Even after the introduction of the full vaccine in Community A cases/ 1000 persons continued to rise. After the vaccination rate plateaued in October and November of 2021 viral loading and overall cases were much higher in community A compared to community B (Figures 4.6 and 4.7).

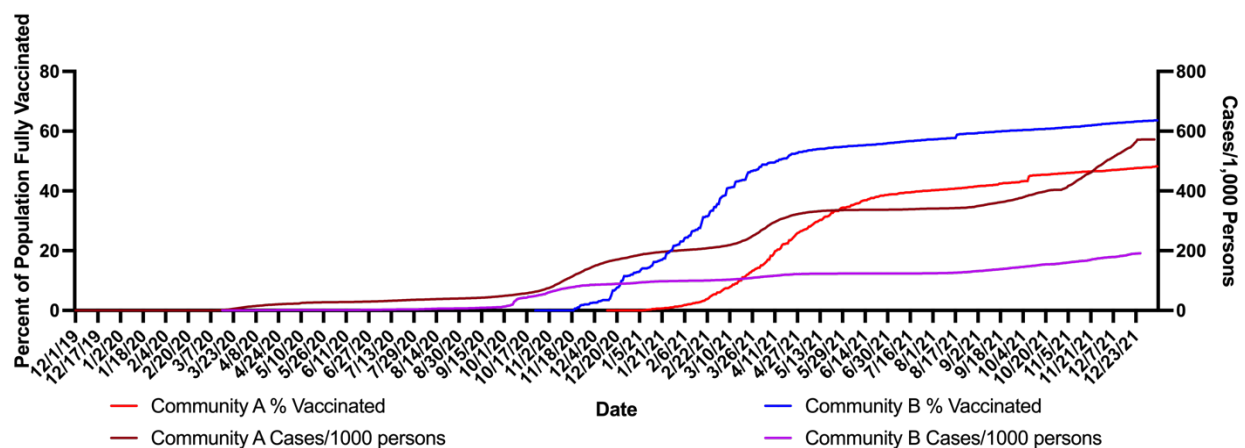
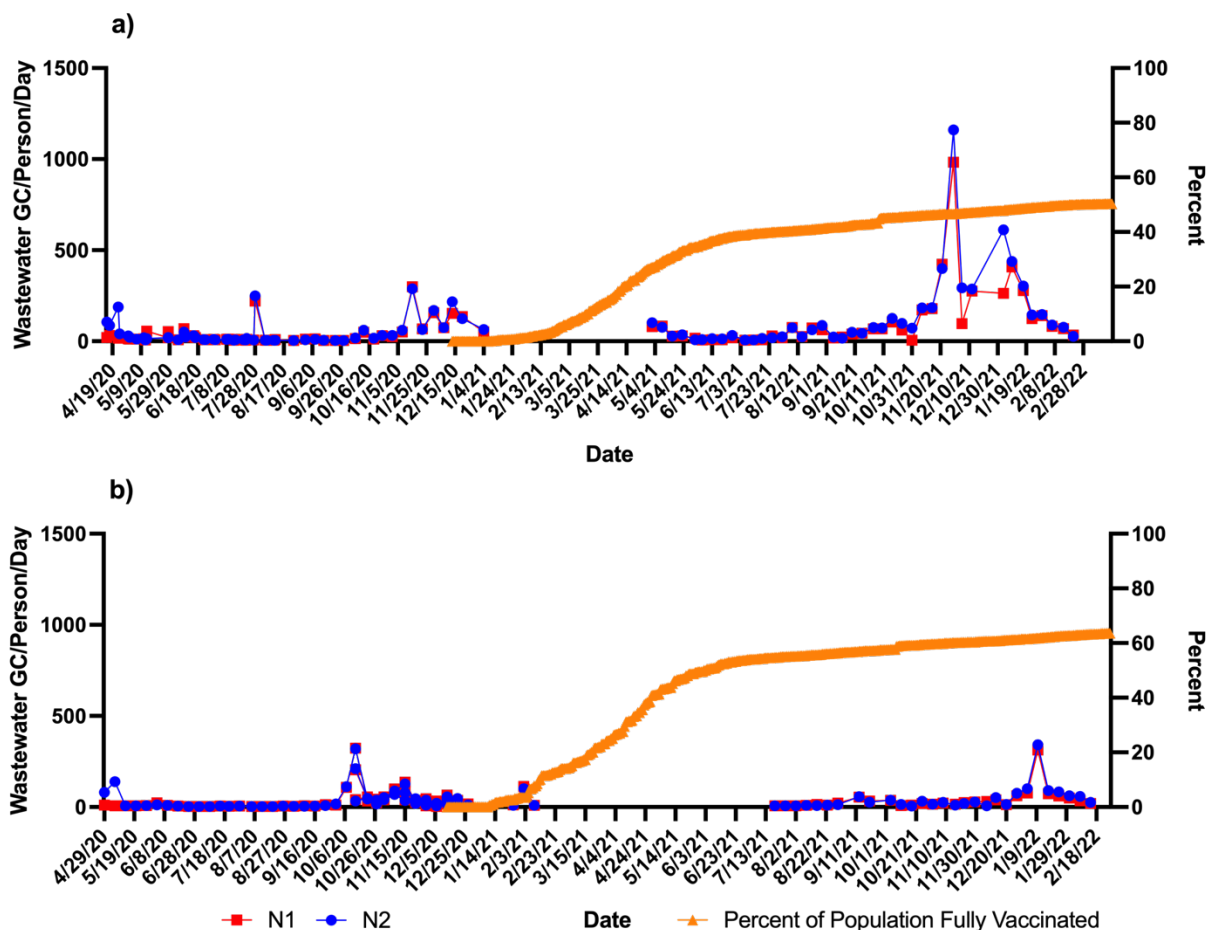


Figure 4.6 Vaccination rates and county level cases per 1,000 persons for communities A and B.



^aZipcode level population data was used for wastewater results normalization; ^bA gap in sampling occurred between January/February and May/July 2021 due to the ending of one project funding and the start of another.

Figure 4.7 Percent of population fully vaccinated compared with SARS-CoV-2 gene target loading (GC/Person/Day). a) WWTP A; b) WWTP B.

4.4.3 Detection of SARS-CoV-2 variants in WWTP A over time

Monitoring for SARS-CoV-2 variants of concern for WWTP A began in December of 2020 with testing for the Alpha variant. In June of 2021, samples from WWTP began to be monitored for mutations associated with the Delta variant and subsequently in January of 2022 samples were monitored for mutations associated with the Omicron variant (Figure 4.8). The N501Y and DEL 69-70 mutations, which indicate the potential presence of the Alpha variant,

were first detected in May of 2021. Levels of N501Y and DEL 69-70 declined as the Delta variant began to spread in Michigan in June of 2021. The Delta variant was first detected in clinical samples in Michigan and confirmed by genetic sequencing on January 16, 2021. However, it was not until July 12, 2021 that the Delta variant was detected in samples from WWTP A. The first detection of the T478K and L452R genes (Delta variant mutations) in WWTP A was on July 12, 2021. Delta variant mutations remained dominant in wastewater samples until January 9, 2022. The K417N and DEL 69-70 mutations (indicative of the Omicron variant) were first detected in wastewater from WWTP A on January 3, 2022 and was first clinically detected in Michigan on December 1, 2021.

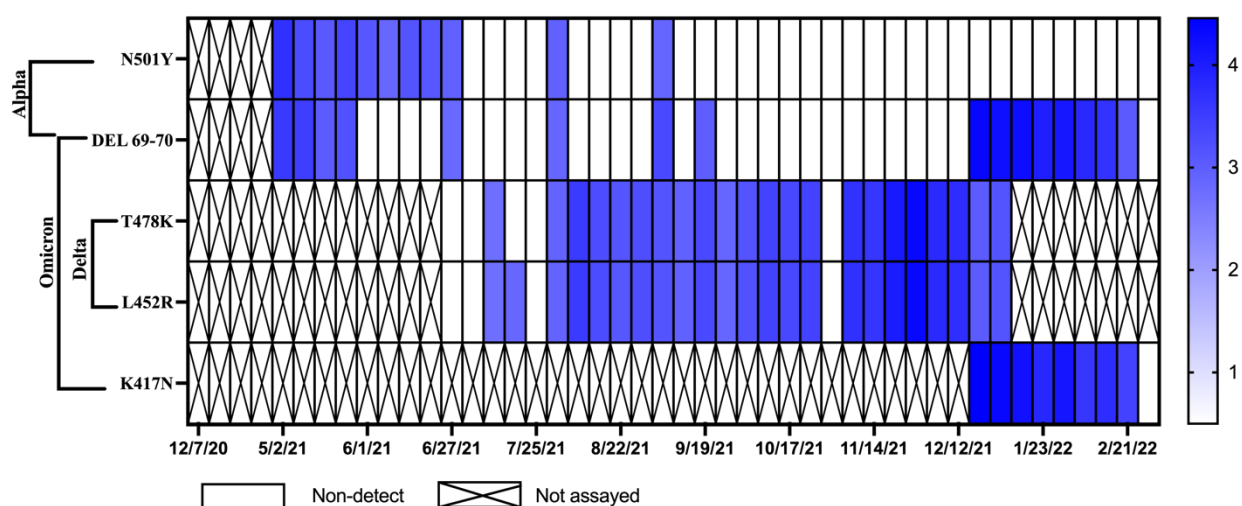


Figure 4.8 Concentrations of SARS-CoV-2 variant genes for the Alpha, Delta, and Omicron variants over time. Samples positive for the N501Y and DEL 69-70 gene mutations indicate the potential presence of the Alpha variant. Samples positive for the T478K and L452R gene mutations indicate the presence of the Delta variant. Samples positive for the K417N and DEL 69-70 gene mutations indicate the presence of the Omicron variant. Empty squares represent Non-detects (NDs) and X's were samples that were not assayed for that marker.

4.5 Discussion

This study demonstrates that the relationship between wastewater surveillance for SARS-CoV-2 with COVID-19 cases differs between communities. While wastewater results were

significantly correlated with the cases in both communities, the level of correlation differed based on spatial (e.g., zipcode vs county level cases) and temporal (e.g., date of symptom(s) onset vs. the referral date for cases) resolution. Both communities (A and B) had higher correlations with zipcode level cases (A: $r = 0.62-0.68$, B: $r = 0.68-0.81$) than county level cases (A: $r = 0.52-0.53$, B: $r = 0.45-0.52$) with the smaller community (B) having the highest levels of correlation overall. However, when the communities' wastewater results were compared against date of symptom(s) onset vs. the referral date for cases community A showed little difference in correlation (onset $r = 0.62-0.68$ vs. referral: $r = 0.67-0.68$). Community B showed a decrease in correlation with cases using the case referral date (onset $r = 0.68-0.81$ vs. report: $r = 0.38-0.48$). These results suggest that for both communities, wastewater surveillance is more representation of higher spatial resolution of cases. When examining the temporal resolution of the communities, the wastewater surveillance results for community A were almost equally as good at representing cases of COVID-19 using either onset or case referral date. However, for community B the wastewater results were more closely tied to the onset of symptoms. These results support the idea that the early warning from wastewater monitoring vary between communities as proposed by Greenwald et al. (2021). These differences between the communities may be due to the amount of septage accepted by each facility (both WWTPs in this study accept septage), the amount of dilution occurring due to stormwater (whether through infiltration or combined treatment), or the percent of the total county population each facility services (WWTP A services 25% of the population while WWTP B services 37.5% of its county's population). Additionally, differences in case "dates" in available datasets may have caused some variation in the relationships observed between wastewater signals and clinical cases. While county level case data from the US CDC consisted of only a single date for each

case, zipcode level datasets provided three dates including “referral date”, “onset date”, and “diagnosis date” with only the “referral dates” being consistently reported.

This study was somewhat limited in its ability to evaluate the impact of vaccination rates on wastewater SARS-CoV-2 surveillance results and community COVID-19 cases. While earlier vaccination rate increases, and lower population levels may have helped curve the increase in cases in community B compared to community A, this current data set is insufficient to statistically evaluate this at this time. Measuring case severity may be a better marker such as hospitalizations or mortality rates, to evaluate the impact of vaccination in communities. The inability to distinguish whether or not the cases, hospitalization, and mortality rates are for vaccinated or unvaccinated individuals is also a limiting factor in accurate examination of the results (Rainey et al., 2022). However, vaccination rates may be able to help in the examination of variants of concern. In community A the gene targets for the Alpha variant were present consistently from May 2, 2021 until June 27, 2021 over which time the vaccination rate for the community increased from 29.3 to 38.4%. After June 27th, the Alpha variant genes were mostly absent from the wastewater samples and were then replaced by the Delta variant mutations. These results are similar to those seen by Yaniv et al. (2021), where an increase in vaccination rates was correlated with the decrease in the prevalence of the Alpha variant, but not the more infectious Delta variant.

Understanding the temporal and geographic resolution of the disease metrics being estimated by wastewater surveillance is paramount for the proper application of WBE. This study found that zipcode data should be used whenever possible compared to county data particularly if the population being served by the wastewater treatment facility is only a proportion of the county. The use of septic tanks could greatly influence this, even if septage is

brought to the WWTP for processing. While there is currently only one study that has investigated SARS-CoV-2 in septic tanks (Zhang et al., 2020) it was focused on treatment and disinfection of hospital wastewater. There is currently no information on the stability of the signal in septage as this would greatly influence what might be expected in individual household wastewater in septic tanks.

Various methods have been used to statistically relate cases of COVID-19 to SARS-CoV-2 concentrations in sewage. Feng et al. (2021) and Ai et al. (2021) have found that use of a fecal indicator does not necessarily improve the correlations. However, Mazumder et al. (2022) and Feng et al. (2021) have found that normalization using loading of the virus per day by population improved the comparisons. In this study this lag was examined by using the onset of symptoms compared to the date of referral. In the larger community (community A) this did not matter, perhaps because of both the greater variability in the case data and SARS-CoV-2 signal in the larger sewer system. Larger complex communities are more difficult to monitor, and detection limits need to be further investigated.

Community A had a much greater mortality than B (based on county level data) over the course of this study. This may have been influenced by access to health care in the greater minority community and as represented by the lower vaccination rates (Alcendor, 2020). Hospitalization data were not available for download by county temporally from the US CDC COVID-19 database. This information will be requested for future analysis. It is clear that increases in COVID-19 as represented by increases in SARS-CoV-2 loading per population in sewage should be considered a warning signal for these disadvantaged communities and should elicit mobilization of health care resources.

REFERENCES

REFERENCES

- Ai, Y., Davis, A., Jones, D., Lemeshow, S., Tu, H., He, F., Ru, P., Pan, X., Bohrerova, Z., and Lee, J. (2021). Wastewater SARS-CoV-2 Monitoring as a Community-level COVID-19 Trend Tracker and Variants in Ohio, United States. *Sci Total Environ* 801, 149757. <https://doi.org/10.1016/j.scitotenv.2021.149757>
- Alcendor, D.J. (2020). Racial Disparities-Associated COVID-19 Mortality among Minority Populations in the US. *J Clin Medicine* 9, 2442. <https://doi.org/10.3390/jcm9082442>
- Bureau of Economic Affairs (BEA) (2022). Personal Income by County, Metro, and Other Areas. <https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas>
- Castiglioni, S., Schiarea, S., Pellegrinelli, L., Primache, V., Galli, C., Bubba, L., Mancinelli, F., Marinelli, M., Cereda, D., Ammoni, E., Pariani, E., Zuccato, E., and Binda, S. (2022). SARS-CoV-2 RNA in Urban Wastewater Samples to Monitor the COVID-19 Pandemic in Lombardy, Italy (March–June 2020). *Sci Total Environ* 806, 150816. <https://doi.org/10.1016/j.scitotenv.2021.150816>
- Fahrenfeld, N.L., Medina, W.R.M., D’Elia, S., Modica, M., Ruiz, A., and McLane, M. (2022). Comparison of Residential Dormitory COVID-19 Monitoring via Weekly Saliva Testing and Sewage Monitoring. *Sci Total Environ* 814, 151947–151947. <https://doi.org/10.1016/j.scitotenv.2021.151947>
- Feng, S., Roguet, A., McClary-Gutierrez, J.S., Newton, R.J., Kloczko, N., Meiman, J.G., and 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, 1955–1965. <https://doi.org/10.1021/acsestwater.1c00160>
- Flood, M.T., D’Souza, N., Rose, J.B., and 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>
- Gendron, L., Verreault, D., Veillette, M., Moineau, S., and Duchaine, C. (2010). Evaluation of Filters for the Sampling and Quantification of RNA Phage Aerosols. *Aerosol Sci Tech* 44, 893–901. <https://doi.org/10.1080/02786826.2010.501351>
- Gerrity, D., Papp, K., Stoker, M., Sims, A., and Frehner, W. (2021). Early-pandemic Wastewater Surveillance of SARS-CoV-2 in Southern Nevada: Methodology, Occurrence, and Incidence/Prevalence Considerations. *Water Res X* 10, 100086. <https://doi.org/10.1016/j.wroa.2020.100086>

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M.H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J., and Gonzalez, D. (2020). COVID-19 Surveillance in Southeastern Virginia using Wastewater-based Epidemiology. *Water Res* 186, 116296.

<https://doi.org/10.1016/j.watres.2020.116296>

Graham, K.E., Loeb, S.K., Wolfe, M.K., Catoe, D., Sinnott-Armstrong, N., Kim, S., Yamahara, K.M., Sassoubre, L.M., Grijalva, L.M.M., Roldan-Hernandez, L., Langenfeld, K., Wigginton, K.R., and Boehm, A.B. (2021). SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with COVID-19 Cases in a Large Urban Sewershed. *Environ Sci Technol* 55, 488–498.

<https://doi.org/10.1021/acs.est.0c06191>

Greenwald, H.D., Kennedy, L.C., Hinkle, A., Whitney, O.N., Fan, V.B., Crits-Christoph, A., Harris-Lovett, S., Flamholz, A.I., Al-Shayeb, B., Liao, L.D., Beyers, M., Brown, D., Chakrabarti, A.R., Dow, J., Frost, D., Koekemoer, M., Lynch, C., Sarkar, P., White, E., Kantor, R., and Nelson, K.L. (2021). Tools for Interpretation of Wastewater SARS-CoV-2 Temporal and Spatial Trends Demonstrated with Data Collected in the San Francisco Bay Area. *Water Res X* 12, 100111. <https://doi.org/10.1016/j.wroa.2021.100111>

Lastra, A., Botello, J., Pinilla, A., Urrutia, J.I., Canora, J., Sánchez, J., Fernández, P., Candel, F.J., Zapatero, A., Ortega, M., and Flores, J. (2022). SARS-CoV-2 Detection in Wastewater as an Early Warning Indicator for COVID-19 Pandemic. Madrid Region Case Study. *Environ Res* 203, 111852–111852. <https://doi.org/10.1016/j.envres.2021.111852>

Lee, S., Kim, T., Lee, E., Lee, C., Kim, H., Rhee, H., Park, S.Y., Son, H.-J., Yu, S., Park, J.W., Choo, E.J., Park, S., Loeb, M., and Kim, T.H. (2020). Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *Jama Intern Med* 180, 1447–1452. <https://doi.org/10.1001/jamainternmed.2020.3862>

Martin, J., Klapsa, D., Wilton, T., Zambon, M., Bentley, E., Bujaki, E., Fritzsche, M., Mate, R., and Majumdar, M. (2020). Tracking SARS-CoV-2 in Sewage: Evidence of Changes in Virus Variant Predominance during COVID-19 Pandemic. *Viruses* 12, 1144. <https://doi.org/10.3390/v12101144>

Mazumder, P., Dash, S., Honda, R., Sonne, C., and Kumar, M. (2022). Sewage surveillance for SARS-CoV-2: Molecular Detection, Quantification and Normalization Factors. *Curr Opin Environ Sci Heal* 100363. <https://doi.org/10.1016/j.coesh.2022.100363>

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., and Brouwer, A. (2020). Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environ Sci Tech Let* 7, 511–516. <https://doi.org/10.1021/acs.estlett.0c00357>

Parasa, S., Desai, M., Chandrasekar, V.T., Patel, H.K., Kennedy, K.F., Roesch, T., Spadaccini, M., Colombo, M., Gabbiadini, R., Artifon, E.L.A., Repici, A., and Sharma, P. (2020). Prevalence

of Gastrointestinal Symptoms and Fecal Viral Shedding in Patients with Coronavirus Disease 2019. *Jama Netw Open* 3, e2011335. <https://doi.org/10.1001/jamanetworkopen.2020.11335>

Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., Warren, J.L., Weinberger, D.M., Arnold, W., and Omer, S.B. (2020). Measurement of SARS-CoV-2 RNA in Wastewater Tracks Community Infection Dynamics. *Nat Biotechnol* 38, 1164–1167. <https://doi.org/10.1038/s41587-020-0684-z>

Rainey, A.L., Loeb, J.C., Robinson, S.E., Lednicky, J.A., McPherson, J., Colson, S., Allen, M., Coker, E.S., Sabo-Attwood, T., Maurelli, A.T., and Bisesi, J.H. (2022). Wastewater Surveillance for SARS-CoV-2 in a Small Coastal Community: Effects of Tourism on Viral Presence and Variant Identification Among Low Prevalence Populations. *Environ Res* 208, 112496–112496. <https://doi.org/10.1016/j.envres.2021.112496>

Rasero, F.J.R., Ruano, L.A.M., Real, P.R.D., Gómez, L.C., and Lorusso, N. (2022). Associations Between SARS-CoV-2 RNA Concentrations in Wastewater and COVID-19 Rates in Days After Sampling in Small Urban Areas of Seville: A Time Series Study. *Sci Total Environ* 806, 150573. <https://doi.org/10.1016/j.scitotenv.2021.150573>

Smith, T., Cassell, G., and Bhatnagar, A. (2021). Wastewater Surveillance Can Have a Second Act in COVID-19 Vaccine Distribution. *Jama Heal Forum* 2, e201616. <https://doi.org/10.1001/jamahealthforum.2020.1616>

US Census Bureau. (2022). County Population by Characteristics: 2010-2019. United States Census Bureau Website. <https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html>

US Centers for Disease Control and Prevention (CDC) (2020). 2019-Novel Coronavirus (2019-nCoV) Real-time rRT-PCR Panel.

US Centers for Disease Control and Prevention (CDC) (2022). COVID Data Tracker. US Centers for Disease Control and Prevention. <https://covid.cdc.gov/covid-data-tracker/#datatracker-home>

Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., and Tan, W. (2020). Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *Jama* 323, 1843–1844. <https://doi.org/10.1001/jama.2020.3786>

Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J.V., and LaCross, N. (2021). Correlation of SARS-CoV-2 RNA in Wastewater with COVID-19 Disease Burden in Sewersheds. *Sci Total Environ* 775, 145790. <https://doi.org/10.1016/j.scitotenv.2021.145790>

Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., Bonneau, R., Brown, M.A., Bushman, M., Chai, P.R., Duvallet, C., Erickson, T.B., Foppe, K., Ghaeli, N., Gu, X., Hanage, W.P., Huang, K.H., Lee, W.L., Matus, M., McElroy, K.A., Nagler, J., Rhode, S.F., Santillana,

M., Tucker, J.A., Wuertz, S., Zhao, S., Thompson, J., and Alm, E.J. (2022). SARS-CoV-2 RNA Concentrations in Wastewater Foreshadow Dynamics and Clinical Presentation of New COVID-19 Cases. *Sci Total Environ* 805, 150121. <https://doi.org/10.1016/j.scitotenv.2021.150121>

Yaniv, K., Ozer, E., Lewis, Y., and Kushmaro, A. (2021). RT-qPCR Ssays for SARS-CoV-2 Variants of Concern in Wastewater Reveals Compromised Vaccination-induced Immunity. *Water Res* 207, 117808–117808. <https://doi.org/10.1016/j.watres.2021.117808>

Zhang, D., Ling, H., Huang, X., Li, J., Li, W., Yi, C., Zhang, T., Jiang, Y., He, Y., Deng, S., Zhang, X., Wang, X., Liu, Y., Li, G., and Qu, J. (2020). Potential Spreading Risks and Disinfection Challenges of Medical Wastewater by the Presence of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Viral RNA in Septic Tanks of Fangcang Hospital. *Sci Total Environ* 741, 140445. <https://doi.org/10.1016/j.scitotenv.2020.140445>

Zheng, S., Fan, J., Yu, F., Feng, B., Lou, B., Zou, Q., Xie, G., Lin, S., Wang, R., Yang, X., Chen, W., Wang, Q., Zhang, D., Liu, Y., Gong, R., Ma, Z., Lu, S., Xiao, Y., Gu, Y., Zhang, J., Yao, H., Xu, K., Lu, X., Wei, G., Zhou, J., Fang, Q., Cai, H., Qiu, Y., Sheng, J., Chen, Y., and Liang, T. (2020). Viral Load Dynamics and Disease Severity in Patients Infected with SARS-CoV-2 in Zhejiang Province, China, January-March 2020: Retrospective Cohort Study. *Bmj* 369, m1443. <https://doi.org/10.1136/bmj.m1443>

5.0 Synopsis of Monitoring and Surveillance

Overall, this project sought to further the understanding of the connections between water and health and to help address the impacts of anthropomorphic changes on the environment through water quality monitoring. This was approached in two ways; the first was an examination downstream of the fecal contamination using microbial source tracking (MST) to evaluate exposure and risk and the second was upstream surveillance at the sewage treatment facility evaluating pathogen excretion and the health of communities as a whole.

The use of MST to evaluate the impact of agricultural practices on water quality in five mixed use watersheds demonstrated the temporal fecal contamination was primarily driven by streamflow/precipitation while spatial contamination was driven by land use. These conclusions were possible through the use of spatial clustering of individual sampling sites allowing for more robust and accurate evaluation of the relationships between variables. Through these analyses porcine pollution was identified as the MST marker most often associated with nutrient contamination. This is of interest as manure for fertilizer use has been increasing in recent years. Additionally, with the implementation of the Food Safety Modernization Act (FSMA), which was signed into law in 2011 and began to take effect in 2015, understanding the sources of fecal contamination in agricultural waters has become even more crucial to protecting food safety and human health. The provisions within the FSMA stipulating safe levels of fecal microbial contamination allowed in agricultural waters and requirements around testing for these pollutants means that source identification and remediation will be even more necessary in the future in particular for smaller producers who have limited water resources. Protecting water quality for food production and recreational use requires the collaboration and coordination of all stakeholders and policymakers. Effective communication between parties, understanding the needs of stakeholders, knowing the abilities and limitations of detection and remediation

methods, and the levels of risk associated with varying levels of contamination are all necessary to protect and preserve water resources. The MST methods allow for a better knowledge on the source of the contamination thus improving the communication and decision making.

The applications of MST are continuing to expand and demonstrate their value. The use of MST for the detection of leaking sewer lines by Gonzalez et al. (2020) showed the value of MST outside of environmental monitoring and led to remediation through infrastructure repairs. On the other hand, the use of MST for differentiation of fecal sources remains a critical function for directing remediation efforts within watersheds. This is demonstrated well by Nguyen et al. (2018) who were able to determine that high levels of FIB found in a Florida watershed were not coming from human sources as was previously assumed, but from animal sources including birds and deer. These source identifications in turn allowed for the more accurate implementation of a TMDL for these impaired waters.

The COVID-19 pandemic presented an abrupt need for virus concentration methods for wastewater to help monitor the etiological agent SARS-CoV-2 for the surveillance of community health. The development of a reliable easy to use workflow for the concentration and detection of SARS-CoV-2 in wastewater was needed. Through the use of a surrogate virus (Phi6 bacteriophage), and field studies polyethylene glycol (PEG) precipitation and RNA detection using ddPCR were demonstrated to be a viable method for the recovery and detection of SARS-CoV-2 from wastewater samples. This study showed that when developing a new workflow and/or method for widespread use across multiple laboratories, accessibility in terms of ease of use and cost along with sufficient sensitivity and specificity were all necessary.

Following the development of the SARS-CoV-2 PEG precipitation and ddPCR workflow samples from two unique communities in Michigan were collected, analyzed and compared to

determine the ability of wastewater surveillance to correlate with cases of COVID-19. This study has shown that wastewater loading of SARS-CoV-2 more accurately correlate with higher resolution (zipcode vs. county level cases) case data. Additionally, as the pandemic progressed the waves of variant strains of SARS-CoV-2 were able to be detected and monitored in one of the communities. This study allowed us to learn that the resolution of case data analyzed along with differences in population demographics can change the efficiency and accuracy of wastewater monitoring across communities.

The ability to monitor indicators of pollution in watersheds and surveil etiological agents of disease in sewage provide non-intrusive methods for evaluating the potential risks and current burdens to community health. While this project was able to accomplish both of these tasks and do so in a way that provided valuable knowledge and methods there still remains many ways to expand on this work in the future. This includes but is not limited to the expansion of species specific MST markers, further understanding the connections between MST markers, pathogens, and nutrients in watersheds. There is always a need for the development of additional methods and workflows particularly now for the surveillance of other pathogens in wastewater. More wastewater-based epidemiology will be undertaken in the future and understanding the impacts of community demographics on the spread and surveillance of disease can be elucidated via sewage testing. While the work in this dissertation focused mainly on research, expanding the lines of communication and the knowledge shared between the scientific community, regulators, and policy makers will be pivotal for the success of any long-term monitoring plan.

REFERENCES

REFERENCES

Gonzalez, D., Keeling, D., Thompson, H., Larson, A., Denby, J., Curtis, K., Yetka, K., Rondini, M., Yeargan, E., Egerton, T., Barker, D., and Gonzalez, R. (2020). Collection System Investigation Microbial Source Tracking (CSI-MST): Applying Molecular Markers to Identify Sewer Infrastructure Failures. *J Microbiol Meth* 178, 106068.

<https://doi.org/10.1016/j.mimet.2020.106068>

Nguyen, K.H., Senay, C., Young, S., Nayak, B., Lobos, A., Conrad, J., and Harwood, V.J. (2018). Determination of wild animal sources of fecal indicator bacteria by microbial source tracking (MST) influences regulatory decisions. *Water Res* 144, 424–434.

<https://doi.org/10.1016/j.watres.2018.07.034>