IDENTIFICATION OF ENDEMIC AND EMERGING VIRAL INFECTIONS IN DETROIT-MI AND TRUJILLO-PERU USING WASTEWATER-BASED-EPIDEMIOLOGY METHODS

By

Brijen Babulal Miyani

A DISSERTATION

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

Environmental Engineering – Doctor of Philosophy

2024

ABSTRACT

The recent COVID-19 pandemic has highlighted the importance of wastewater-based-epidemiology (WBE) methods to effectively monitor and predict infectious viral disease outbreaks. Traditional disease detection systems rely on identification of outbreaks by diagnostic analysis of clinical samples. Since it is impossible to test every individual in a community for all potential viral infections those systems are lacking in their ability to predict outbreaks. Environmental based surveillance that includes WBE can help reduce the load of medical systems, guide clinical testing, and provide early warnings. Untreated wastewater can serve as a community-based excrement sample that can be tested to identify the diversity of human viruses prevalent in the community. This dissertation presents innovative screening tools based on molecular methods, high throughput sequencing, and bioinformatics analysis that can be applied in the analysis of wastewater samples to identify viral diversity in the corresponding catchment community. Confirmation of this method was done using traditional PCR methods and all results were correlated with clinical samples. Further, population biomarker methods were developed to normalize the signals. The proposed methods were applied to identify endemic and emerging viral infections in a major metropolitan city in US (Detroit-MI) and Latin America (Trujillo-Peru) using wastewater-based-epidemiology methods.

The first chapter of the dissertation focuses on an application of bioinformatics-based screening tool to reveal high abundance of rare human herpesvirus 8 in Detroit wastewater. The second chapter focuses on early warning of COVID-19 second wave in Detroit MI. The third chapter focuses on surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States and assessing per capita SARS-CoV-2 estimations and COVID-19 incidence. The fourth chapter uses the bioinformatics-based screening tool identify a wide variety of human viruses in Trujillo-Peru wastewater and confirms Covid-19, monkeypox, and diarrheal disease outbreaks.

ACKNOWLEDGEMENTS

I would like to express sincere thanks to my advisor Dr Irene Xagoraraki for her constant guidance, support, and motivation throughout my PhD program at Michigan State University. Like a mother, she has been always there to guide and shape me into a better researcher. She accepted me for who I am and gave me novel opportunities to work on, for which I am deeply grateful.

A sincere thanks to my committee members: Dr Thomas Voice, Dr Syed Hashsham, and Dr Rene Hinojosa. Their expert advice and encouragement in my graduate life are invaluable to me. I am also grateful to Dr Morgan Jenahvive for being my mentor when I was a teaching assistant for EGR 100. Sincere thanks to Dr Anna Mehrotra, Xavier Fonoll, and Dr John Norton for collaborating for SARS-CoV-2 work. Additional thanks to Dr William Cunningham, Sabrina Vieyra, Dr Ruben Briceno De La Cruz for collaborating for Peru work. Grateful to Dr Russell Allen Faust for collaborating for polio work.

A most sincere thanks to my colleagues for welcoming me to the lab, teaching me the basics, and collaborating on several research projects. Colleagues include Camille McCall, Huiyun Wu, Evan O'Brien, Liang Zhao, Sejal Buch, Zachary Gentry, Maddie Spooner, Yangyang Zou, Yabing Li, and Heidy Peidro Guzman.

A special thanks to all my friends who have been invaluable with their support throughout my time at MSU. George Viju, Sri Manoharan, Aditi, Monik, Sanket, Harikrishna, Divya, Abhijeet, Praveen, Aniket, Henry, Dhruv, Akshay, Sachin, Kushal, Taurean, Alex, Bobby, Noah, Aakash, Rahul, Sarthak, Vishwanath, Manu, Augusto, Patricia, Lindsey, Pooja, Anirudh, Tejas, and Ruchita kept me sane at MSU. Thanks to all my friends and faculty from boarding school, Vallabh Ashram and bachelor's university, SSASIT, for building a strong foundation in me. A special thanks to football coach Rajan Mishra for always pushing me and imparting persistence, patience, and resiliency in me. Special thanks to my parents (Babubhai and Varshaben) for bringing me into this world and always supporting me financially and emotionally. My brother, Jayarth, taught me to be practical. I am grateful for my family that taught me to always keep my feet on the ground. The above mentioned and other countless unmentioned people were vitally important during my PhD program at MSU.

TABLE OF CONTENTS

INTRODUCTION……...1

CHAPTER 1: HIGH ABUNDANCE OF HUMAN HERPESVIRUS 8 IN WASTEWATER FROM A LARGE URBAN AREA……4
BIBLIOGRAPHY……..17

CHAPTER 2: EARLY WARNINGS OF COVID-19 SECOND WAVE IN DETROIT MI...….21
BIBLIOGRAPHY……..35

CHAPTER 3: SURVEILLANCE OF SARS-COV-2 IN NINE NEIGHBORHOOD SEWERSHEDS IN DETROIT TRI-COUNTY AREA, UNITED STATES: ASSESSING PER CAPITA SARS-COV-2 ESTIMATIONS AND COVID-19 INCIDENCE……..37
BIBLIOGRAPHY……..65

CHAPTER 4: BIOINFORMATICS-BASED SCREENING TOOL IDENTIFIES A WIDE VARIETY OF HUMAN VIRUSES IN TRUJILLO-PERU WASTEWATER AND CONFIRMS COVID-19, MONKEYPOX, AND DIARRHEAL DISEASE OUTBREAKS…….70
BIBLIOGRAPHY……..89

CONCLUSIONS AND SIGNIFICANCE…….....97

INTRODUCTION

The first chapter assesses the diversity and abundance of Human Herpesviruses (HHVs) in the influent of an urban wastewater treatment plant by using shotgun sequencing, metagenomic analysis, and qPCR. Influent wastewater samples were collected from the three interceptors that serve the City of Detroit and Wayne, Macomb, and Oakland counties between November 2017 to February 2018. The samples were subjected to a series of processes to concentrate viruses which were further sequenced and amplified using qPCR. All nine types of human herpesviruses were detected in wastewater. Human Herpesvirus 8 (HHV-8), known as Kaposi's sarcoma herpesvirus, which is only prevalent in 5-10% of USA population, was found to be the most abundant followed by Human Herpesvirus 3 or Varicella-zoster virus. The high abundance of HHV-8 in the Detroit metropolitan area may be attributed to the HIV-AIDS outbreak that was ongoing in Detroit during the sampling period. The approach described in this paper can be used to establish a baseline of viruses secreted by the community as a whole. Sudden changes in the baseline would identify changes in community health and immunity.

The second chapter focuses on using wastewater-based-epidemiology to provide early warnings of the second COVID-19 wave in Detroit metropolitan area in MI, USA. SARS-CoV-2 RNA from untreated wastewater samples was compared to reported public health records. Untreated wastewater samples were collected from the Great Lakes Water Authority (GLWA) Water Resource Recovery Facility (WRRF), located in southeast Michigan, between Aug 6, 2020 and Dec 14, 2020. The WRRF receives wastewater from its service area via three main interceptors: Detroit River Interceptor (DRI), North Interceptor-East Arm (NIEA), and Oakwood-Northwest-Wayne County Interceptor (ONWI). A total of 144 untreated wastewater samples were collected (45, 48, and 51 for ONWI, NIEA and DRI respectively) at the point of intake into the WRRF. Virus-selective sampling was conducted, and viruses were isolated from wastewater using electropositive NanoCeram column filters. For each sample, an average of 33 L of wastewater was passed through NanoCeram electropositive cartridge filters at an average rate of 11 L/m. Viruses were eluted and concentrated and SARS-CoV-2 RNA concentrations were quantified with RT-qPCR. SARS-CoV-2 RNA was detected in 98% of samples and measured concentrations were in the range of 4.45E+04 to 5.30E+06 genomic copies/L. Early warnings of COVID-19 peaks were observed approximately four weeks prior to reported publicly available clinical data.

The third chapter focuses on surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States and assessing per capita SARS-CoV-2 estimations and COVID-19 incidence. Appropriate population biomarkers were screened for wastewater SARS-CoV-2 normalization and the normalized SARS-CoV-2 values were compared across locations with different demographic characteristics in southeastern Michigan. Wastewater samples were collected between December 2020 and October 2021 from nine neighborhood sewersheds in the Detroit Tri-County area. Using reverse transcriptase droplet digital polymerase chain reaction (RT-ddPCR), concentrations of N1 and N2 genes in the studied sites were quantified, with N1 values ranging from 1.92×10^2 genomic copies/L to 6.87×10^3 gc/L and N2 values ranging from 1.91×10^2 gc/L to 6.45×10^3 gc/L. The strongest correlations were observed with between cumulative COVID-19 cases per capita (referred as COVID-19 incidences thereafter), and SARS-CoV-2 concentrations normalized by total Kjeldahl nitrogen (TKN), creatinine, 5- hydroxyindoleacetic acid (5-HIAA) and xanthine when correlating the per capita SARS-CoV-2 and COVID-19 incidences. When SARS-CoV-2 concentrations in wastewater were normalized and compared with COVID-19 incidences, the differences between neighborhoods of varying demographics were reduced as compared to differences observed when comparing non-normalized SARS-CoV-2 with COVID-19 cases. This indicates when studying the disease burden in communities of different demographics, accurate per capita estimation is of great importance. The study suggests that monitoring selected water quality parameters or biomarkers, along with RNA concentrations in wastewater, will allow adequate data normalization for spatial comparisons, especially in areas where detailed sanitary sewage flows and contributing populations in the catchment areas are not available. This opens the possibility of using WBE to assess community infections in rural areas or the developing world where the contributing population of a sample could be unknown.

The fourth chapter uses the bioinformatics-based screening tool to identify a wide variety of human viruses in Trujillo-Peru wastewater and confirms Covid-19, monkeypox, and diarrheal disease outbreaks. The immense burden of human and zoonotic viral infections in Peru is widely recognized. Peru was one of the worst-hit countries during the COVID-19 pandemic. Moreover, multiple other viral diseases (enteric, respiratory, bloodborne, and vector-borne) are endemic and arising. According to Peru's Ministry of Health, the COVID-19 pandemic reverted the current health facilities of the country leading to reduced action to curb other diseases. Furthermore,

many viral diseases in the area are under-reported and not recognized. One significant aspect of combating infections is early detection. The One Health approach, in addition to clinical testing, incorporates environmental surveillance for early detection of infectious disease outbreaks. The purpose of this study is to use wastewater surveillance methods to identify virus-related diseases circulating in Trujillo-Peru. A screening tool, based on molecular methods, high throughput sequencing and bioinformatics analysis is applied. We collected nine untreated wastewater samples from the Covicorti wastewater utility in Trujillo-Peru. High throughput metagenomic sequencing followed bioinformatic analysis was used to assess viral diversity of the samples. Our results revealed the presence of sequences associated with multiple human and zoonotic viruses including *Orthopoxvirus, Hepatovirus, Rhadinovirus, Parechovirus, Mamastrovirus, Enterovirus, Varicellovirus, Norovirus, Kobuvirus, Bocaparvovirus, Simplexvirus, Spumavirus, Orthohepevirus, Cardiovirus, Molliscipoxvirus, Salivirus, Parapoxvirus, Gammaretrovirus, Alphavirus, Lymphocryptovirus, Erythroparvovirus, Sapovirus, Cosavirus, Deltaretrovirus, Roseolovirus, Flavivirus, Betacoronavirus, Rubivirus, Lentivirus, Betapolyomavirus, Rotavirus, Hepacivirus, Alphacoronavirus, Mastadenovirus, Cytomegalovirus* and *Alphapapillomavirus*. For confirmation purposes, we tested for the presence of selective viruses belonging to the genera detected above. PCR based molecular methods confirmed the presence of SARS-CoV-2, MPXV, noroviruses, and RoA in our samples. Furthermore, publicly available clinical data for selected viruses confirm our findings. Wastewater, or other environmental media surveillance, combined with bioinformatics has a vast potential to serve as a systematic screening tool for the identification of a myriad of human or zoonotic viruses that may cause disease. The results of this practical method can guide further clinical surveillance efforts and allocation of resources. Incorporation of this bioinformatic-based screening tool by public health officials in Peru and other Latin American countries will help manage endemic and emerging diseases that could save human lives and resources.

CHAPTER 1: HIGH ABUNDANCE OF HUMAN HERPESVIRUS 8 IN WASTEWATER FROM A LARGE URBAN AREA

This chapter was published in Journal of Applied Microbiology, Volume 130, Issue 5, 1 May 2021, Pages 1402–1411, B. Miyani, C. McCall and I. Xagoraraki, High abundance of human herpesvirus 8 in wastewater from a large urban area

ABSTRACT

This study assesses the diversity and abundance of Human Herpesviruses (HHVs) in the influent of an urban wastewater treatment plant by using shotgun sequencing, metagenomic analysis, and qPCR. Influent wastewater samples were collected from the three interceptors that serve the City of Detroit and Wayne, Macomb and Oakland counties between November 2017 to February 2018. The samples were subjected to a series of processes to concentrate viruses which were further sequenced and amplified using qPCR. All nine types of human herpesviruses were detected in wastewater. Human Herpesvirus 8 (HHV-8), known as Kaposi's sarcoma herpesvirus, which is only prevalent in 5-10% of USA population, was found to be the most abundant followed by Human Herpesvirus 3 or Varicella-zoster virus. The high abundance of HHV-8 in the Detroit metropolitan area may be attributed to the HIV-AIDS outbreak that was ongoing in Detroit during the sampling period. The approach described in this paper can be used to establish a baseline of viruses secreted by the community as a whole. Sudden changes in the baseline would identify changes in community health and immunity.

1. INTRODUCTION

Untreated wastewater can be considered as a community-based excrement sample, which if monitored in a timely manner, can identify spikes in excreted viruses that can be related to outbreaks. Wastewater-based epidemiology can be used for endemic viral disease identification, early detection of potential viral outbreaks and monitoring of community health (O'Brien and Xagoraraki, 2019; Xagoraraki and O'Brien, 2020). Contrary to classic epidemiology and traditional disease detection and management systems, which rely on clinical symptoms to detect and describe an outbreak, wastewater-based epidemiology methods can monitor community health and forecast viral outbreaks in advance (Xagoraraki and O'Brien, 2020). Viruses of particular concern are herpesviruses since they are abundant in human populations and have been associated with multiple diseases. Furthermore, increased shedding of certain human herpesviruses (HHVs) has been associated with immunocompromised populations. Monitoring

herpesvirus diversity and abundance may serve as an indicator of overall population health and immunity.

Herpesviruses are DNA viruses belonging to the *Herpesviridae* family. *Herpesviridae*, based on biological and molecular properties, are further subdivided into Alphaherpesvirinae (α-Herpesviruses), Betaherpesvirinae (β-Herpesviruses), and Gammaherpesvirinae (γ-Herpesviruses) subfamilies (Pellett and Roizman, 2013; Roizman *et al.*, 1981). Nine herpesviruses are known to infect humans and include Human herpesvirus 1 HHV-1 (or Herpes simplex virus HSV-1), HHV-2 (or HSV-2), HHV-3 (or Varicella-zoster virus VZV), HHV-4 (or Epstein-Barr virus EBV), HHV-5 (or human cytomegalovirus HCMV/CMV), HHV-6 (HHV-6A and HHV-6B), HHV-7, and HHV-8 (or Kaposi's sarcoma herpesvirus KSHV). Currently the two variants of HHV-6 are recognized as separate viruses (Adams *et al.*, 2013). α-Herpesviruses include HHV-1, HHV-2, and HHV-3. β-Herpesviruses include HHV-5, HHV-6, and HHV-7. γ-Herpesviruses include carcinogenic HHV-4 and HHV-8 (Contreras and Slots, 2000; Fratini *et al.*, 2014).

Age adjusted seroprevalence of HHV-1 and HHV-2 is 48.1% and 12.1%, respectively (McQuillan *et al.*, 2018). Currently rather than primary infection, recurrent form of HHV-3 infection is quite prevalent in USA with 1 million cases of it occurring every year (CDC, 2019). Age adjusted seroprevalence of HHV-4, HHV-5, HHV-6 and HHV-7 is at least 50% (Ablashi *et al.*, 1994; Bate *et al.*, 2010; Braun *et al.*, 1997; Campadelli-Fiume *et al.*, 1999; Dowd *et al.*, 2013; CDC, 2020; R. *et al.*, 2012) whereas HHV-8 is prevalent only in 5% of the general population (Gao *et al.*, 1996; Martin, 2007).

Shedding of HHVs occur in symptomatic individuals and otherwise healthy asymptomatic individuals (Roizman *et al.*, 2013). However, some herpesvirus species such as HHV-8 and HHV-4 are shed at higher amounts in immunocompromised populations, solid organ transplant patients, cancerous patients and critically ill non immunocompromised patients compared to healthy controls (Gianella *et al.*, 2016; Gianella, Morris, *et al.*, 2013; Gianella, Smith, *et al.*, 2013; Libert et al., 2015; Lisco *et al.*, 2012; Lucht *et al.*, 1993; Lucht, Biberfeld and Linde, 1995; Miller *et al.*, 2006; Palmieri et al., 2018; Sarmento et al., 2018). Furthermore, reactivation of HHVs is common in immunocompromised individuals like HIV patients (Kim *et al.*, 2006). HHV-8 prevalence in the US is the lowest compared to other HHV species (Gao *et al.*, 1996; Martin, 2007). Therefore, it is conceivable that monitoring the excretion of HHV-8

could be used as an indicator of population immunity. Measurements could be conducted to establish base excretion levels of a certain population, compare between populations, and possibly monitor sudden increases that may indicate changes in community health and immunity.

HHVs have been detected in wastewater, sludge and biosolids using sequencing and metagenomic analysis (Aw *et al*., 2014; Bibby and Peccia, 2013; Bibby *et al.* 2011; Brisebois *et al.*, 2018; Di Bonito *et al.* 2017; O'Brien *et al* 2017). Published papers have focused on Herpesvirus genus or the family *Herpesviridae.* To the best of our knowledge, no previous study validated next generation sequencing results using qPCR and no previous study has surveyed all nine HHV types in environmental samples. The objective of this paper is to identify the diversity of HHV species in wastewater by using metagenomic sequencing and qPCR and investigate the possibility of tracking community health by monitoring HHVs in wastewater.

2. MATERIALS AND METHODS

2.1 Field sampling

Samples were collected from the Detroit wastewater treatment utility that serves Wayne, Oakland and Macomb counties of Michigan. Three main sewers (interceptors): North Interceptor—East Arm (NI-EA), Detroit River Interceptor (DRI), and Oakwood-Northwest-Wayne County Interceptor (O-NWI) collect wastewater from the City of Detroit and Wayne, Oakland and Macomb counties. All three interceptors were sampled at their discharge points in the wastewater utility, prior to any treatment. Samples were collected in triplicate at each interceptor, between November 2017 to February 2018 (11/17/2017, 12/1/2017, 12/14/2017, 1/19/2018, 2/2/2018 and 2/16/2018). A total of nine samples (triplicates for each interceptor) were collected per sampling date for a total of 54 samples. Approximately 12L of influent wastewater was passed through NanoCeram electropositive cartridge filters at a rate of 11-12 L min^{-1} using a previously described method (Kuo *et al.*, 2010; O'Brien *et al.*, 2017; U.S. EPA, 2001). Samples were transported to Michigan State University on ice for further processing.

2.2 Sample processing

Viruses were eluted from the cartridge filters with 1.5% w/v beef extract (0.05 mol l^{-1} glycine, pH 9.5). Elution was carried out using a previously described method (Kuo *et al.,* 2010; O'Brien *et al.,* 2017; U.S. EPA, 2001) within 48 h. All eluted samples were stored in 2 ml cryogenic tubes in -80 °C. Viral DNA and RNA was extracted using QIAamp Viral RNA Mini

Kit (Qiagen). All the extracts were preserved in -20 °C. Quantitation of viral DNA was done using a Qubit Fluorometer.

2.3 Next generation sequencing and metagenomics analysis

After nucleic acid extraction, sample replicates were combined. This resulted in eighteen samples. To detect HHV-related contigs, samples were sequenced with Illumina NextSeq Sequencing at the Research Technology Support Facility (RTSF) at Michigan State University. Libraries were prepared through Rubicon ThruPLEX DNA-Seq kit, which were further quantified and checked for quality control. All samples were loaded onto one lane of an Illumina NextSeq 500 and sequencing was performed from both ends for an output of 150bp.

Trimming of reads and removal of adaptors was done using Trimmomatic (Bolger *et al.*, 2014). To reduce the chance of false positives, the trimmed reads were merged and assembled into contig files using IDBA-UD *de novo* assembler (Peng *et al.*, 2012). The contigs were blasted against NCBI's RefSeq viral database using BLAST+/2.2.31 with a minimum E-value 10^{-3}. A PYTHON code was written to extract contigs related to all 9 types of HHVs from all BLAST output files using the NCBI accession number. The contig dataset was filtered to remove low quality contigs. A minimum base length of 200bp was used as threshold based on previous papers (Aw *et al.*, 2014; Rosario *et al.*, 2009). A BitScore (Dora *et al.*, 2015), of 100 was used as a second filter to select reliable HHVs contigs.

To detect HIV-related contigs in wastewater samples, nucleic acid was merged as stated above, reverse transcribed, and subjected to random amplification as previously described (Wang *et al.*, 2003). After DNA quantification with Qubit Fluorometer, the eighteen cDNA samples were sequenced with Illumina HiSeq Sequencing. Samples were processed on a HiSeq 4000 sequencing platform to generate pair-end 150bp reads. Trimming, removal of adaptors, assembling using IDBA-UD *de novo* and blasting against merged NCBI viral database followed. A PYTHON code, as described above, was used to extract contigs of HIV from the BLAST output using the NCBI accession number.

2.4 qPCR analysis

All samples were analyzed using qPCR for Human Herpesvirus 8 and 6 on Eppendorf realplex 2 instrument. The assay used for HHV-8 was based on previously described method (Lallemand *et al.*, 2000). The total reaction mixture was 20 µl containing 10 µl of 1X Roche Lightcycler qPCR Master mix, 1.8 µl of 10 µmol l^{-1} of each forward and reverse primer (final

concentration 900nm), 1 ul of 5 µmol l^{-1} probe (final concentration 250nm), 0.4 µl of PCR grade water and 5 µl of DNA extract. For each sample, the reaction was carried out in triplicates under the following conditions: denaturation at 95 °C for 15 min and 45 amplification cycles of 95 °C, 15 s and 65 °C for 1 min.

The assay used for HHV-6 was based on previously described method (Gautheret-Dejean *et al.*, 2002). The total reaction mixture was 20 µl containing 10 µl of 1X Roche Lightcycler qPCR Master mix, 0.8 µl of 10 µmol l^{-1} of each forward and reverse primer (final concentration 400nm), 0.4 µl of 10 µmol l^{-1} probe (final concentration 200nm), 3 µl of PCR grade water and 5 µl of DNA extract. For each sample, the reaction was carried out in triplicates under the following conditions: denaturation at 95 °C for 15 min and 45 amplification cycles of 95 °C, 15 s and 60 °C for 1 min. All samples were analyzed in triplicate. The standard curves obtained for HHV-8 and HHV-6 are illustrated by figure 1.1.

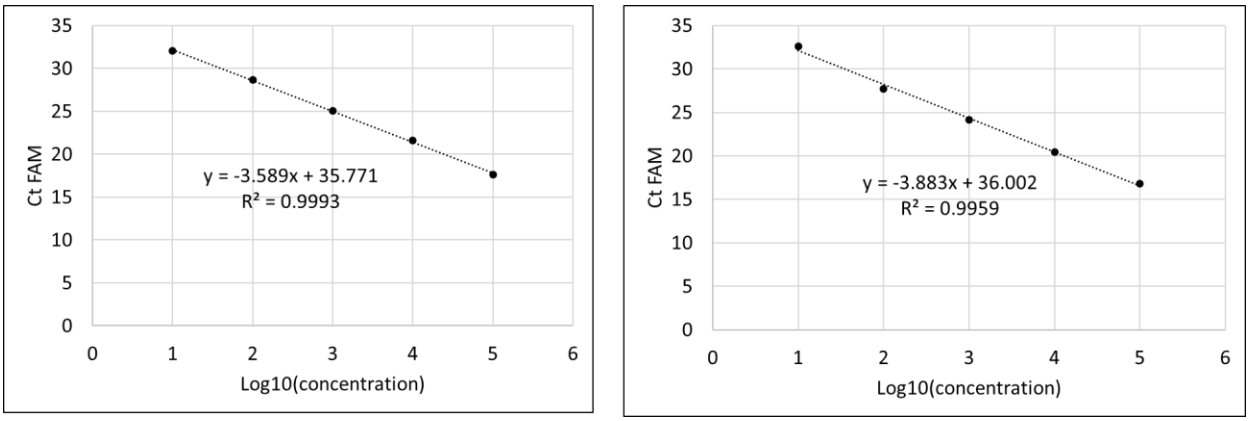

Figure 1.1. Standard curves of qPCR HHV-8 (Left) and HHV-6 (Right)

2.5 Statistical analysis

Oneway ANOVA (comparing means) and a Tukey post hoc test were done in SPSS to check if the datasets were statistically significant compared to each other.

3. RESULTS

Trimming resulted in an average of 22 million read pairs per sample. An average of 6% of those reads was eliminated by Trimmomatic. The remaining 94% of reads were assembled into contigs by IDBA-UD *de novo* assembler. BLAST results were viewed with MEGAN. Table 1.1 illustrates the percentage of affiliated viral sequences and number of contigs in all samples. The percentage of affiliated sequences ranged from 44 to 75%. Viral sequences were divided in

dsDNA, ssDNA, ssRNA, unclassified archaeal, unclassified bacterial and unclassified viruses. Diversity of viruses present in the samples is also illustrated in Table 1.1.

Table 1.1. Classification of Viruses using Illumina Sequencing

Sample Date)	Number of contigs	% of Affiliated Viral Sequences	Assigned viruses	dsDNA viruses	ssDNA viruses	ssRNA viruses	Unclassified archaeal viruses	Unclassified bacterial viruses	Unclassified viruses
O-NWI 11/17/17	256178	68.5	9.7	97.4	0.15	0.09	0	0.27	2.12
NI-EA 11/17/17	306405	66.8	8.8	97.5	0.17	0	0	0.27	2.06
DRI 11/17/17	284815	64.3	8.3	97.2	0.21	0	0	0.24	2.39
O-NWI 12/1/17	231217	61.9	9.5	96.6	0.23	0.14	0	0.28	2.72
NI-EA 12/1/17	189833	57.7	10.1	95.6	0.34	0.18	0.13	0.31	3.41
DRI 12/1/17	294031	55.52	8.8	95.5	0.43	0.17	0.11	0.26	3.49
O-NWI 12/14/17	260841	74.3	9.2	98	0.08	0	0	0.27	1.65
NI-EA 12/14/17	257241	74.4	9.05	98	0.11	0.12	0	0.24	1.55
DRI 12/14/17	268942	75.9	9.07	98.1	0	0	0	0.29	1.55
O-NWI 1/19/18	234981	69.9	10.2	97.3	0.17	0.1	0	0.25	2.2

Table 1.1. (cont'd)

NI-EA 1/19/18	281807	73	11.7	97.7	0.11	0	0	0.28	1.86
DRI 1/19/18	237258	67.85	10.5	97	0.22	0.09	0	0.27	2.4
O-NWI 2/2/18	234535	73.4	11.7	97.6	0.14	0	0	0.26	2.03
NI-EA 2/2/18	252170	73.2	11.4	97.8	0.11	0	0	0.24	1.85
DRI 2/2/18	221373	68	11.1	96.7	0.21	0.11	0	0.27	2.67
O-NWI 2/16/18	190586	47.55	8.4	94.3	0.55	0.26	0.18	0.27	4.38
NI-EA 2/16/18	322559	23.83	2.8	94	0.39	0	0	0	5.64
DRI 2/16/18	228384	44.1	7.6	94.4	0.54	0.23	0.2	0.29	4.37

Contigs were matched to all HHVs (figure 1.2). The highest number of matches was observed for HHV-8 and the lowest for HHV-1, 2 and 6. The second higher abundance was observed for HHV-3. The number of HHV-8 related contigs was significantly greater than the number of observed HHV-3 related contigs ($P = 0.0$).

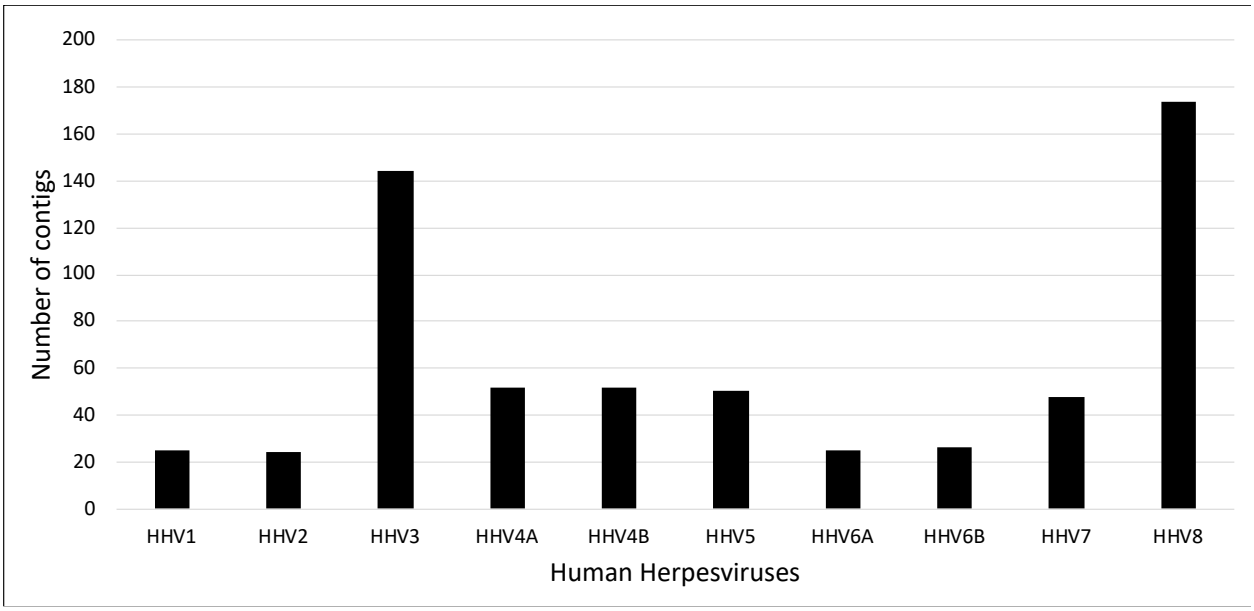

Figure 1.2. Average number of contigs in all samples (O-NWI, NI-EA and DRI for all six dates) of different HHVs

HHV species with lowest abundance (HHV-6) and highest abundance (HHV-8) based on metagenomics analysis, were further analyzed with qPCR. HHV-6 was detected in 9 out of 54 samples as shown in figure 1.3. Sampling dates 1/19/2018 and 2/16/2018 were the only dates when HHV-6 was detectable in more than one interceptor, as shown in figure 1.3. HHV-8 was found in 51 out of 54 samples as shown in figure 1.4. The highest number of genomic copies of HHV-8 per L, 112, was found on 12/1/2017 at NI-EA interceptor.

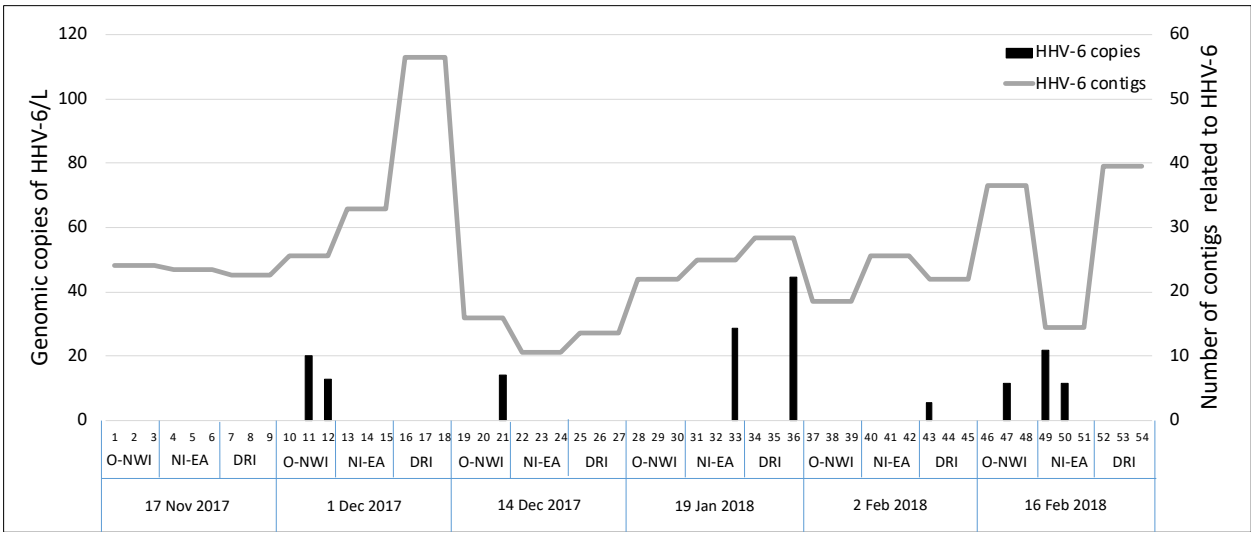

Figure 1.3. Genomic copies of HHV-6 L^{-1} (measured by qPCR) and number of contigs related to HHV-6 (detected with Illumina sequencing) in all wastewater interceptors for all sampling dates. O-NWI = Oakwood-Northwest Wayne County Interceptor, NI-EA = North Interceptor-East Arm, DRI = Detroit River Interceptor

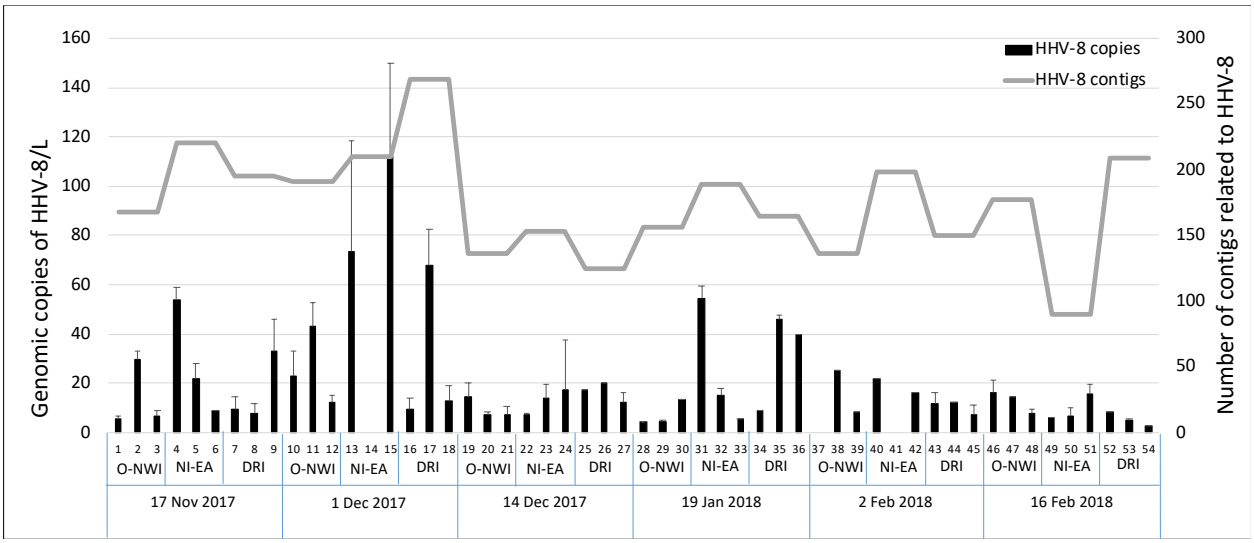

Figure 1.4. Genomic copies of HHV-8 L^{-1} (measured by qPCR) and number of contigs related to HHV-8 (detected with Illumina sequencing) in all wastewater interceptors for all sampling dates. O-NWI = Oakwood-Northwest Wayne County Interceptor, NI-EA = North Interceptor-East Arm, DRI = Detroit River Interceptor

NI-EA interceptor accounted for the maximum percentage of HHV-8 copies L^{-1} for the majority of dates as depicted in figures 1.5 and 1.7. The highest number of genomic copies L^{-1} of HHV-8 was detected on 12/1/2017 (figure 1.4). One-way ANOVA (comparing means) and a

Tukey post hoc test in SPSS resulted in genomic copies L^{-1} measured in the NI-EA interceptor being statistically significantly higher over other interceptors (P = 0.0). Out of all interceptors only samples collected at the NI-EA interceptor on 12/1/17 showed statistically significantly results (P < 0.05) as compared to the other five sampling dates.

To correlate with HHV-8 data, the number of HIV contigs were identified. The PYTHON code resulted in contigs of HIV of low parameters (BitScore 22 to 56, and E-value 1E-3 to 1E-10). Figure 1.6 illustrates the HIV contigs for different interceptors. Contigs related to NI-EA were higher than contigs related to O-NWI and DRI (figure 1.6), but there was no statistical significance found between any of the interceptors.

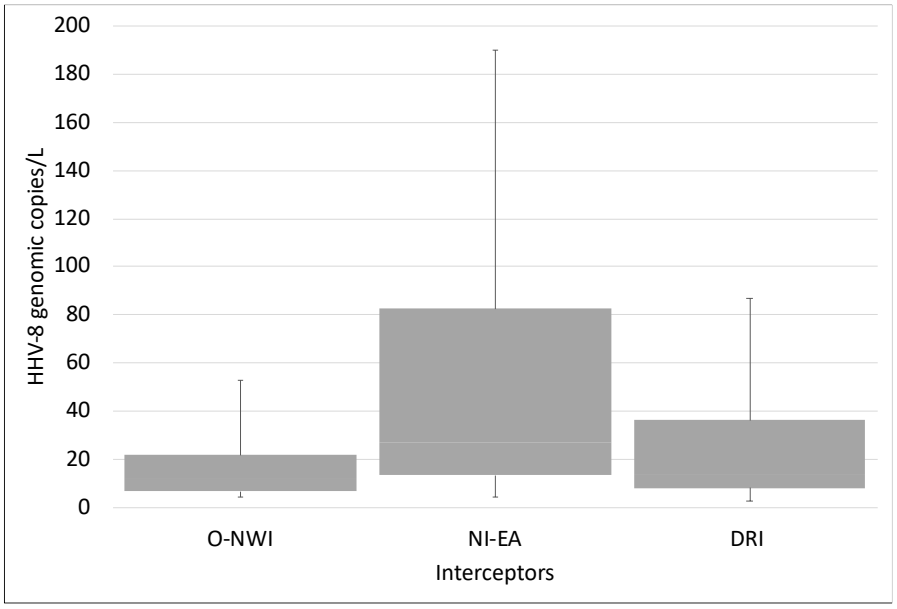

Figure 1.5. Distribution of HHV-8 Concentrations between interceptors for all sampling dates. O-NWI = Oakwood-Northwest Wayne County Interceptor, NI-EA = North Interceptor-East Arm, DRI = Detroit River Interceptor

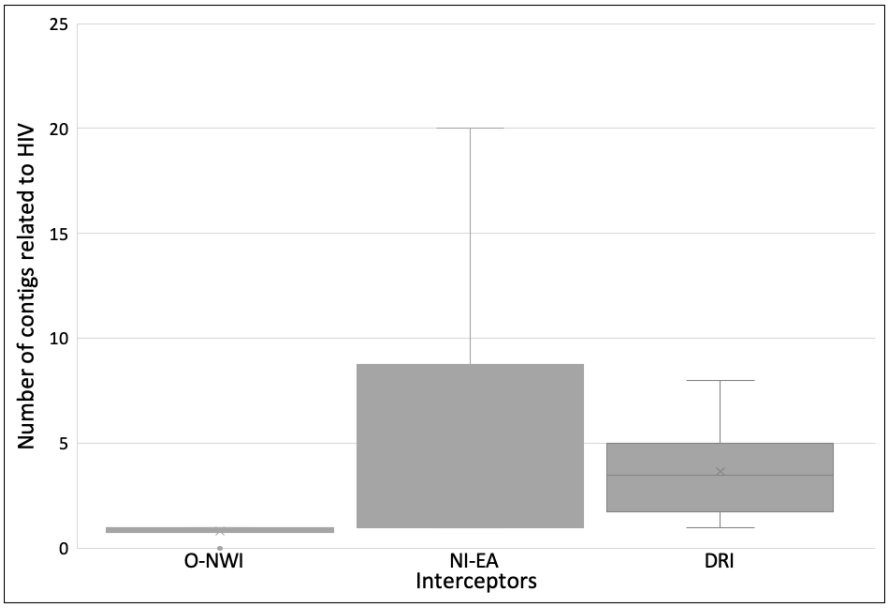

Figure 1.6. Average number of HIV-related contigs in all the samples (O-NWI, NI-EA and DRI for all six dates). O-NWI = Oakwood-Northwest Wayne County Interceptor, NI-EA = North Interceptor-East Arm, DRI = Detroit River Interceptor

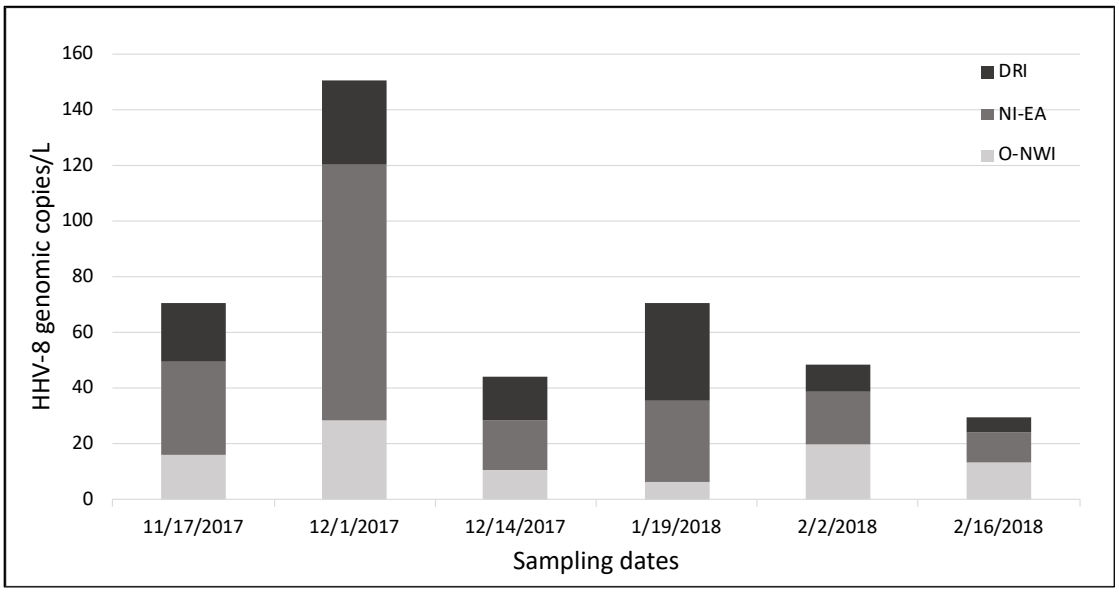

Figure 1.7. Average number of genomic copies/L of HHV-8 for all sampling dates. O-NWI = Oakwood-Northwest Wayne County Interceptor, NI-EA = North Interceptor-East Arm, DRI = Detroit River Interceptor

4. DISCUSSION

Untreated wastewater samples were collected from the influent of the Detroit wastewater treatment utility that services the city of Detroit and Oakland, Wayne and Macomb counties. All

nine types of HHVs were detected in the samples using molecular methods. According to Michigan Department of Health and Human Services (MDHHS, 2018), there have been HHV-3 related cases observed in these counties during the same timeframe (2017 and 2018) as our sampling period. The highest number of HHV-3 related (chicken pox, shingles and unspecified) cases were observed in Oakland County followed by almost equivalent number of cases for Macomb and Wayne County. MDHHS only reports HHV-3 related cases (Varicella/chickenpox, Shingles and VZV infection unspecified). HHV-3, according to metagenomic analysis, was found as the second most abundant virus from all samples (figure 1.3). Other HHVs related cases are not currently monitored by MDHHS.

HHV-8 related contigs were observed in all samples and their abundance was the highest between all nine HHV species. The second highest abundance was observed for HHV-3. One-way ANOVA indicated that HHV-8 related contigs were significantly higher than HHV-3 related contigs with a p-value of 0.0. Furthermore, HHV-8 was detected in 51 out of 54 samples with qPCR. This was an unexpected outcome since HHV-8, compared to other HHVs, is relatively less abundant in the USA with a seroprevalence of 5-10% (Martin, 2007). On the contrary, in this study we observed that HHV8 was the most prevalent observed species by both Illumina sequencing and qPCR. One of the reasons of high prevalence could be due to higher probability of reactivation of HHV-8 in immunocompromised individuals like HIV/AIDS patients or patients undergoing any kind of surgical transplantation, compared to healthy individuals (Hudnall *et al.*, 1998; Mendez *et al.*, 1999). HIV-positive patients shed HHV-8 at a higher rate compared to healthy individuals (Gianella *et al.*, 2016; Gianella, Morris, *et al.*, 2013; Gianella, Smith, *et al.*, 2013; Lisco *et al.*, 2012; Lucht *et al.*, 1993; Lucht, Biberfeld and Linde, 1995; Miller *et al.*, 2006).

This is supported by the ongoing HIV outbreak in the Detroit Metropolitan area at the time of sampling. MDHHS reported an HIV outbreak in Detroit that started in 2016 (MDHHS, 2020). The prevalence rates in Detroit city, 713.3 per 100,000 residents, was 4.3 times greater than average prevalence rate in Michigan (MDHHS, 2020). HIV prevalence rates in Wayne county except Detroit City, Oakland county and Macomb county at the end of 2018 were 191.2, 164.5 and 141.9 respectively per 100,000 residents (MDHHS, 2020). Bioinformatics analysis confirmed presence of HIV-related contigs in our samples.

Regular monitoring of HHVs such as HHV-8 in wastewater influent for a particular city might create a baseline of HHVs shed as a whole by the community. Spikes or sudden increases might possibly forecast other ongoing diseases in the community since immunocompromised individuals shed HHVs at a higher rate as compared to healthy individuals.

In summary, all HHVs, except HHV-8, are quite prevalent in USA and have been associated with multiple potential diseases. These viruses are being regularly shed from both asymptomatic and symptomatic individuals and ultimately end up in wastewater. Influent wastewater samples were collected from the three interceptors (NI-EA, DRI and O-NWI) serving the City of Detroit and Wayne, Macomb and Oakland counties between November 2017 to February 2018. The samples were subjected to a series of processes to concentrate viruses which were further sequenced and amplified using qPCR. All nine types of human herpesviruses were detected in wastewater using Illumina sequencing, bioinformatics and qPCR, with HHV-8 being the most abundant. According to literature, HHVs are shed at higher rate in HIV immunocompromised patients, solid organ transplant recipients, cancerous patients and critically ill non-immunocompromised patients compared to healthy controls. The high abundance of HHV-8 in the Detroit metropolitan area is attributed to the HIV-AIDS outbreak that was ongoing in Detroit during the sampling period. The approach described in this paper can be used to monitor HHVs to establish a baseline secreted by the community. Sudden changes in the baseline would identify changes in community health and immunity.

BIBLIOGRAPHY

Ablashi, D. V. *et al.* (1994) Human herpesvirus-7 (HHV-7). *In Vivo (Brooklyn).*

Adams, M. J., King, A. M. Q. and Carstens, E. B. (2013) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). *Arch. Virol.* doi: 10.1007/s00705-013-1688-5.

Aw, T. G., Howe, A. and Rose, J. B. (2014) Metagenomic approaches for direct and cell culture evaluation of the virological quality of wastewater. *J. Virol. Methods*. doi: 10.1016/j.jviromet.2014.09.017.

Bate, S. L., Dollard, S. C. and Cannon, M. J. (2010) Cytomegalovirus Seroprevalence in the United States: The National Health and Nutrition Examination Surveys, 1988–2004. *Clin. Infect. Dis.* doi: 10.1086/652438.

Bibby, K. and Peccia, J. (2013) Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ. Sci. Technol.* doi: 10.1021/es305181x.

Bibby, K., Viau, E. and Peccia, J. (2011) Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett. Appl. Microbiol.* doi: 10.1111/j.1472-765X.2011.03014.x.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. doi: 10.1093/bioinformatics/btu170.

Braun, D. K., Dominguez, G. and Pellett, P. E. (1997) Human herpesvirus 6. *Clin. Microbiol. Rev.* doi: 10.1146/annurev.me.44.020193.002113.

Brisebois, E. *et al.* (2018) Human viral pathogens are pervasive in wastewater treatment center aerosols. *J. Environ. Sci. (China)*. doi: 10.1016/j.jes.2017.07.015.

Campadelli-Fiume, G., Mirandola, P. and Menotti, L. (1999) Human herpesvirus 6: An emerging pathogen. *Emerg. Infect. Dis.* doi: 10.3201/eid0503.990306.

CDC (2019) *Shingles (Herpes Zoster)*. Available at: https://www.cdc.gov/shingles/index.html (Accessed: 7 June 2020).

CDC (2020) *Cytomegalovirus (CMV) and Congenital CMV Infection*. Available at: https://www.cdc.gov/cmv/index.html (Accessed: 7 June 2020).

Contreras, A. and Slots, J. (2000) Herpesviruses in human periodontal disease. *J. Periodontal Res.*, pp. 3–16. doi: 10.1034/j.1600-0765.2000.035001003.x.

Di Bonito, P. *et al.* (2017) Detection of oncogenic viruses in water environments by a Luminex-based multiplex platform for high throughput screening of infectious agents. *Water Res.* doi: 10.1016/j.watres.2017.06.088.

Dora, P. *et al.* (2015) Important Databases and Tools To Identify Promising Drug Targets By Subtractive Genomics Approach – a Review. *Int. J. Res. Eng. Technol.*, 04(06), pp. 453–455. doi: 10.15623/ijret.2015.0406077.

Dowd, J. B. *et al.* (2013) Seroprevalence of Epstein-Barr Virus Infection in U.S. Children Ages 6-19, 2003-2010. *PLoS One*. doi: 10.1371/journal.pone.0064921.

Fratini, M., Di Bonito, P. and La Rosa, G. (2014) Oncogenic Papillomavirus and Polyomavirus in Water Environments: Is There a Potential for Waterborne Transmission? *Food Environ. Virol.* doi: 10.1007/s12560-013-9134-0.

Gao, S. J. *et al.* (1996) KSHV antibodies among Americans, Italians and Ugandans with and without Kaposi's sarcoma. *Nat. Med.* doi: 10.1038/nm0896-925.

Gautheret-Dejean, A. *et al.* (2002) Development of a real-time polymerase chain reaction assay for the diagnosis of human herpesvirus-6 infection and application to bone marrow transplant patients. *J. Virol. Methods*, 100(1–2), pp. 27–35. doi: 10.1016/S0166-0934(01)00390-1.

Gianella, S. *et al.* (2016) Replication of Human Herpesviruses Is Associated with Higher HIV DNA Levels during Antiretroviral Therapy Started at Early Phases of HIV Infection. *J. Virol.* doi: 10.1128/jvi.02638-15.

Gianella, S., Morris, S. R., *et al.* (2013) Herpes viruses and HIV-1 drug resistance mutations influence the virologic and immunologic milieu of the male genital tract. *AIDS*. doi: 10.1097/QAD.0b013e3283573305.

Gianella, S., Smith, D. M., *et al.* (2013) Shedding of HIV and human herpesviruses in the semen of effectively treated HIV-1-infected men who have sex with men. *Clin. Infect. Dis.* doi: 10.1093/cid/cit252.

Hudnall, S. D. *et al.* (1998) Serologic and Molecular Evidence of Human Herpesvirus 8 Activation in Renal Transplant Recipients. *J. Infect. Dis.* doi: 10.1086/314482.

Kim, H. N. *et al.* (2006) Oral Herpes Simplex Virus Type 2 Reactivation in HIV-Positive and -Negative Men. *J. Infect. Dis.* doi: 10.1086/505879.

Kuo, D. H. W. *et al.* (2010) Assessment of human adenovirus removal in a full-scale membrane bioreactor treating municipal wastewater. *Water Res.* doi: 10.1016/j.watres.2009.10.039.

Lallemand, F. *et al.* (2000) Quantitative analysis of human herpesvirus 8 viral load using a real-time PCR assay. *J. Clin. Microbiol.*, 38(4), pp. 1404–1408. doi: 10.1128/JCM.38.4.1404-1408.2000.

Libert, Nicolas, Christine Bigaillon, Cyrus Chargari, Mourad Bensalah, Violaine Muller, Stephane Merat, and Stephane de Rudnicki. "Epstein-Barr virus reactivation in critically ill immunocompetent patients." *Biomedical journal* 38, no. 1 (2015).

Lisco, A. *et al.* (2012) Semen of HIV-1-infected individuals: Local shedding of herpesviruses and reprogrammed cytokine network. *J. Infect. Dis.* doi: 10.1093/infdis/jir700.

Lucht, E. *et al.* (1993) Human immunodeficiency virus type 1 and cytomegalovirus in saliva. *J. Med. Virol.* doi: 10.1002/jmv.1890390213.

Lucht, E., Biberfeld, P. and Linde, A. (1995) Epstein-Barr Virus (EBV) DNA in saliva and EBV serology of HIV-1-infected persons with and without hairy leukoplakia. *J. Infect.* doi: 10.1016/S0163-4453(95)80025-5.

Martin, J. N. (2007) The epidemiology of KSHV and its association with malignant disease. in *Hum. Herpesviruses Biol. Ther. Immunoprophyl.* doi: 10.1017/CBO9780511545313.055.

McQuillan, G. *et al.* (2018) Prevalence of Herpes Simplex Virus Type 1 and Type 2 in Persons Aged 14-49: United States, 2015-2016. *NCHS Data Brief*.

MDHHS Michigan Department of Health and Human Services (2018) MMWR MDSS Morbidity and Mortality Weekly Report, Michigan Disease Surveillance System Weekly Disease Report for The Week Ending December 29th, 2018 (MMWR Week 52)

MDHHS Michigan Department of Health and Human Services (2020) *HIV Statistics and Data Reports*. Available at: https://www.michigan.gov/mdhhs/0,5885,7-339-71550_2955_2982_72251-350111--,00.html (Accessed: 7 June 2020).

Mendez, J. C. *et al.* (1999) Relationship of HHV8 replication and Kaposi's sarcoma after solid organ transplantation. *Transplantation*. doi: 10.1097/00007890-199904270-00022.

Miller, C. S. *et al.* (2006) High prevalence of multiple human herpesviruses in saliva from human immunodeficiency virus-infected persons in the era of highly active antiretroviral therapy. *J. Clin. Microbiol.* doi: 10.1128/JCM.00256-06.

O'Brien, E. and Xagoraraki, I. (2019) A water-focused one-health approach for early detection and prevention of viral outbreaks. *One Heal.* doi: 10.1016/j.onehlt.2019.100094.

O'Brien, E. *et al.* (2017) Diversity of DNA viruses in effluents of membrane bioreactors in Traverse City, MI (USA) and La Grande Motte (France). *Water Res.* doi: 10.1016/j.watres.2017.01.014.

Palmieri, Michelle, Mariana Ornaghi, Victor Adriano de Oliveira Martins, Luciana Correa, Thais Bianca Brandao, Ana Carolina do Prado Ribeiro, Laura Masami Sumita, Tania Regina Tozetto-Mendoza, Claudio Sergio Pannuti, and Paulo Henrique Braz-Silva. "Oral shedding of human herpesviruses in patients undergoing radiotherapy/chemotherapy for head and neck squamous cell carcinoma is not affected by xerostomia." *Journal of oral microbiology* 10, no. 1 (2018): 1476643

Pellett, P. E. and Roizman, B. (2013) Herpesviridae. in *Fields Virol. Sixth Ed.* doi: 10.1201/9781351071642-20.

Peng, Y. *et al.* (2012) IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. doi: 10.1093/bioinformatics/bts174.

R., A.-Z. *et al.* (2012) *Cytomegalovirus complicating inflammatory bowel disease: A 10-year experience in a community-based, university-affiliated hospital*. *Gastroenterol. Hepatol.*

Roizman, B. *et al.* (1981) Herpesviridae: Definition, provisional nomenclature, and taxonomy. *Intervirology*. doi: 10.1159/000149269.

Roizman, B., Knipe, D. M. and Whitley, R. J. (2013) Herpes simplex viruses. in *Fields Virol. Sixth Ed.*

Rosario, K. *et al.* (2009) Metagenomic analysis of viruses in reclaimed water. *Environ. Microbiol.* doi: 10.1111/j.1462-2920.2009.01964.x.

Sarmento, Dmitry José de Santana, Rubens Caliento, Amabile Oliveira de Souza, Tânia Regina Tozetto-Mendoza, Michelle Palmieri, Victor Adriano de Oliveira Martins, Paulo

Henrique Braz-Silva, and Marina Gallottini. "Salivary shedding of herpesviruses in renal transplant recipients." *Journal of Investigative and Clinical Dentistry* 9, no. 4 (2018): e12356

U.S. EPA (2001) *Manual of Methods for Virology (Chapter 14)*.

Wang, D. *et al.* (2003) Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol.* doi: 10.1371/journal.pbio.0000002.

Xagoraraki, I. and O'Brien, E. (2020) Wastewater-Based Epidemiology for Early Detection of Viral Outbreaks. in. doi: 10.1007/978-3-030-17819-2_5.

CHAPTER 2: EARLY WARNINGS OF COVID-19 SECOND WAVE IN DETROIT MI

This chapter was published in Journal of Environmental Engineering Volume 147, Issue 8, 3 June 2021, Brijen Miyani, Liang Zhao, Maddie Spooner, Sejal Buch, Zachary Gentry, Anna Mehrotra, John Norton, and Irene Xagoraraki, Early Warnings of COVID-19 Second Wave in Detroit

ABSTRACT

This study focuses on using wastewater-based-epidemiology to provide early warnings of the second COVID-19 wave in Detroit metropolitan area in MI, USA. SARS-CoV-2 RNA from untreated wastewater samples was compared to reported public health records. Untreated wastewater samples were collected from the Great Lakes Water Authority (GLWA) Water Resource Recovery Facility (WRRF), located in southeast Michigan, between Aug 6, 2020 and Dec 14, 2020. The WRRF receives wastewater from its service area via three main interceptors: Detroit River Interceptor (DRI), North Interceptor-East Arm (NIEA), and Oakwood-Northwest-Wayne County Interceptor (ONWI). A total of 144 untreated wastewater samples were collected (45, 48, and 51 for ONWI, NIEA and DRI respectively) at the point of intake into the WRRF. Virus-selective sampling was conducted, and viruses were isolated from wastewater using electropositive NanoCeram column filters. For each sample, an average of 33 L of wastewater was passed through NanoCeram electropositive cartridge filters at an average rate of 11 L/m. Viruses were eluted and concentrated and SARS-CoV-2 RNA concentrations were quantified with RT-qPCR. SARS-CoV-2 RNA was detected in 98% of samples and measured concentrations were in the range of 4.45E+04 to 5.30E+06 genomic copies/L. Early warnings of COVID-19 peaks were observed approximately four weeks prior to reported publicly available clinical data.

1. INTRODUCTION

The possibility and methodological approaches of using wastewater-based epidemiology for early detection of viral outbreaks have been previously discussed (Xagoraraki, 2020; Xagoraraki and O'Brien 2020; O'Brien and Xagoraraki 2019). The methodology has been applied and validated for the Detroit area population for multiple enteric, respiratory and bloodborne viruses (McCall *et al*. 2020; Miyani *et al*. 2020b, McCall *et all,* in press), and recently for SARS-CoV-2 (Miyani *et al.* 2020a). Wastewater surveillance has been used to

indicate the presence of Influenza A (H1N1) (Heijnen *et al*. 2011) and Hepatitis E virus, Adenovirus, Astrovirus and Rotavirus (Hellmér *et al*. 2014).

Wastewater-based epidemiology can be used as an additional resource for health officials to monitor and predict COVID-19 prevalence in the community. The final product of wastewater-based-epidemiology that is of interest to health officials is the lag time i.e., the time difference between SARS-CoV-2 RNA detection in wastewater and the time a person in the community gets tested positive for COVID-19. As reported in the scientific community, the lag time seems to be varying between 3 to 28 day (Saguti *et al*. 2021, Nemudryi *et al*. 2020, Peccia *et al*. 2020, Weidhaas *et al*. 2020, Ahmed *et al*. 2020, Randazzo *et al*. 2020, Medema *et al*. 2020, Rimoldi *et al*. 2020, Harvey *et al*. 2020, Ahmed *et al*. 2021).

A critical component of molecular testing in environmental samples is concentrating and cleaning the sample to remove chemical or microbial contamination typically not found in clinical samples. For this work we used virus-selective sampling of large volumes of wastewater followed by elution and concentration (Miyani 2020a). The method is based on electropositive column filters (Nanoceram filters Argonide, Sanford, FL) which have an advantage over conventional 1MDS filters (USEPA 2001) in terms of low cost and higher viral recovery efficiency from wastewater (Soto-Beltran *et al*. 2013). In this paper we also used wastewater interceptor flow rates and service area population to normalize observed concentrations of SARS-CoV-2 RNA in wastewater. Our SARS-CoV-2 RNA measurements and normalized concentrations provided early warnings of about four weeks prior to medically reported data of the second COVID-19 wave in Detroit, Michigan.

2. MATERIAL AND METHODS

2.1 Field Site

Untreated wastewater samples were collected from the Great Lakes Water Authority (GLWA) Water Resource Recovery Facility (WRRF) located in southeast Michigan between Aug 6, 2020 and Dec 14, 2020. The GLWA WRRF is the largest single-site wastewater treatment plant in the US. It has a primary and secondary treatment capacity of 6,435 million liters per day (ML/d; 1,700 million gallons per day [MGD]) and 3,520 ML/d (930 MGD), respectively. GLWA's WRRF has a semi-combined sewer-shed system, which collects and treats stormwater along with residential, industrial, and commercial waste, depending on service area. The WRRF serves the three largest counties, by population, in Michigan: Wayne, Oakland, and

Macomb. The WRRF receives wastewater from its service municipalities via three main interceptors: Detroit River Interceptor (DRI), North Interceptor-East Arm (NIEA), and Oakwood-Northwest-Wayne County Interceptor (ONWI) – figure 2.1. Untreated wastewater samples were collected from all three interceptors at the point of discharge into the WRRF. Estimated population served by each interceptor, daily flows, and other characteristics of the three interceptors, between Aug 6 and Dec 14, 2020, are shown in table 2.1.

Table 2.1. Interceptor serviced population, flows, and wastewater quality parameters between Aug 6 and Dec 14, 2020

Characteristic	ONWI	NIEA	DRI
Population	840,600	1,482,000	492,000
Flow (ML/d)	677 ± 272	625 ± 191	758 ± 210
Flow (MGD)	179 ± 72	165 ± 51	200 ± 55
Estimated Fraction of Flow that is Sanitary (%)	33 ± 6.6	53 ± 12.3	27 ± 5.4
BOD (mg/L)	113 ± 39	179 ± 69	70 ± 31
TSS (mg/L)	100 ± 48	191 ± 79	100 ± 47
TP (mg/L)	3 ± 0.8	3 ± 0.8	1 ± 0.4

Notes:

1. ML/d = 10^6 liters per day; MGD = million gallons per day; cBOD = carbonaceous biochemical oxygen demand; TSS = total suspended solids; TP = total phosphorus.
2. Values for flow, BOD, TSS and TP are shown as average ± one standard deviation.

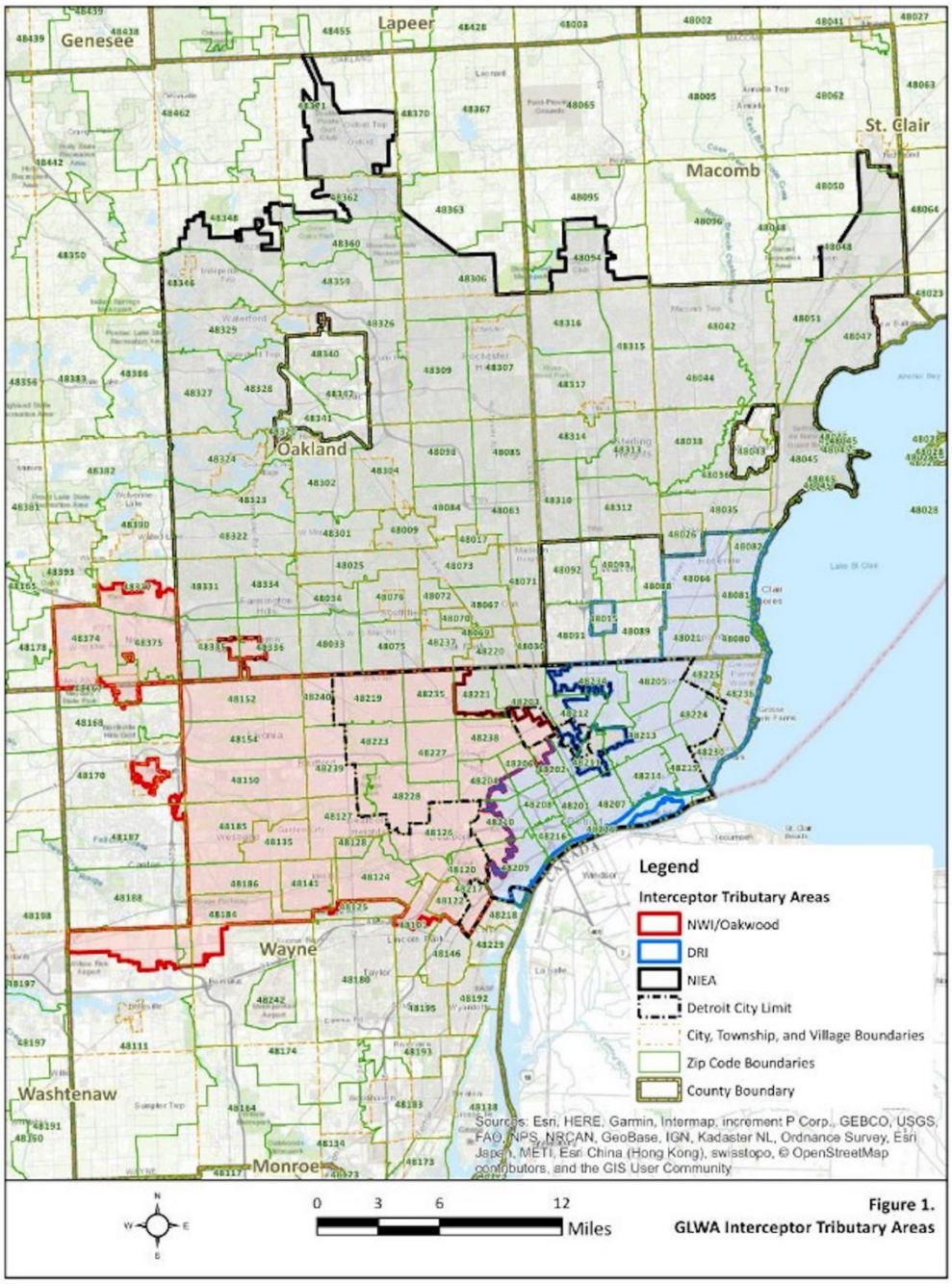

Figure 2.1. Areas served by different GLWA WRRF interceptors

Note: Seventy two percent of the population of Wayne, Macomb and Oakland Counties is serviced by the WWRF

2.2 Virus-Selective Sample Collection and Elution

A total of 144 untreated wastewater samples were collected at the influent of the GLWA

WRRF (45, 48, and 51 for ONWI, NIEA, and DRI respectively). The sampling took place on 18 different dates between Aug 6 and Dec 14, 2020 (table 2.2) and 94% of the samples were collected in triplicate. Viruses were isolated from wastewater using electropositive NanoCeram® column filters (Argonide, Sanford, FL). Depending on the suspended solids of wastewater, approximately 10-60 L of influent wastewater was passed through the column filters at a rate not more than 11.3 L/min using a previously described method (Miyani *et al*. 2020a). After sampling, all NanoCeram® cartridges were stored at 4^{o}C until elution, which was performed within 48 hours using 1.5% beef extract (0.05 M glycine, pH 9.5) and a previously described method (Miyani *et al*. 2020a). Following elution, each sample was aliquoted into multiple 2ml corning tubes. 140 µL of sample was used from one of the corning tubes for RNA extraction

Table 2.2. SARS-CoV-2 RNA concentrations (genomic copies/liter) quantified in untreated wastewater

Sampling Date	ONWI Interceptor		NIEA Interceptor		DRI Interceptor	
	Average	STDEV	Average	STDEV	Average	STDEV
8/6/20	9.90E+04	2.66E+04	1.64E+05	5.46E+04	1.06E+05	6.13E+04
8/11/20	5.66E+05	4.68E+05	5.75E+05	3.52E+05	8.71E+04	
8/18/20	1.18E+06		NA		9.23E+04	3.71E+04
8/26/20	8.37E+04	3.05E+04	9.88E+04	6.77E+03	1.39E+05	
9/1/20	NA		NA		1.04E+05	3.06E+04
9/8/20	2.88E+05	3.44E+05	8.34E+05	5.27E+05	3.94E+05	5.13E+03
9/15/20	6.50E+05	7.00E+05	8.59E+05	6.34E+05	9.42E+04	3.23E+04
9/21/20	NA		NA		NA	
9/29/20	2.35E+06	2.66E+06	1.28E+06	9.01E+05	NA	
10/7/20	3.53E+06	1.76E+06	2.04E+06	2.10E+05	2.80E+06	1.29E+06
10/14/20	1.04E+06	6.92E+05	1.21E+06	1.28E+06	6.08E+05	8.25E+05
10/21/20	NA		1.24E+06	1.43E+06	1.13E+06	1.34E+06
10/28/20	NA		1.33E+06	1.57E+06	1.80E+05	3.26E+04
11/4/20	2.69E+05	1.12E+04	2.25E+05	8.38E+04	1.33E+05	7.45E+04
11/11/20	1.43E+05	4.49E+03	1.86E+05	4.28E+03	2.50E+05	1.33E+05
11/18/20	4.25E+05	2.80E+05	2.53E+05	8.96E+04	1.54E+05	
11/23/20	NA		NA		NA	
11/30/20			1.81E+05	9.27E+03	1.33E+05	
12/7/20	2.78E+05	2.42E+05	2.43E+05	3.23E+04	1.07E+05	2.45E+04
12/14/20	6.53E+04		7.99E+04	3.95E+04	1.01E+05	

Note: NA=non-available (sample not collected)

2.3 RNA extraction

Viral RNA was extracted using Viral RNA QIAGEN kits following manufacturer's protocol. A volume of 140 µL of elution buffer was used to elute RNA from 140 µL of sample.

2.4 RT-qPCR

The extracted viral RNA was reverse transcribed using iScript RT-qPCR Supermix (Bio-Rad). A two-step quantitative reverse transcription polymerase chain reaction (RT-qPCR) was

used to quantify SARS-CoV-2 in the samples. qPCR was performed on a Mastercycler ep realplex2 (Eppendorf) in 96-well optical plates. The primers and probe targeting the N1 gene of SARS-CoV-2 were used based on a previously described assay (Miyani *et al*. 2020a).

American Type Culture Collection synthetic SARS-CoV-2 RNA: ORF, E, N VR-3276SD was used as a positive control. Sterile nuclease-free water was used as a negative control. For each sample, five microliters of cDNA were transferred to a 15 µL reaction mix containing a final concentration of 150 nM of each primer, 125 nM of probe, 1× Lightcycler 480 probes master, and sterile nuclease free water. All reactions were performed in four replicates with the following amplification conditions: denaturation at 95ºC for 10 min, followed by 45 cycles of 95ºC for 10 s and 55ºC for 30 s. All qPCR runs included a positive and a negative control.

The protocol mentioned above was used to prepare standard curves with 10-fold serial dilutions of positive SARS-CoV-2 RNA control ranging from 10^1 to 10^5 genome copies/ul. A standard curve was prepared once every two weeks. The average standard curve was used to calculate SARS-CoV-2 concentrations since biweekly standard curves were very similar. The standard curve used to quantify SARS-CoV-2 RNA in field samples has a slope and R-squared value of -3.4279 and 0.9977 respectively.

Sampling volume and all dilutions starting from field sampling to RT-qPCR were accounted when calculating copies of SARS-CoV-2 RNA per liter. These included: volume of wastewater passed through the NanoCeram filter, 1 L of beef extract used for elution, 30 mL of Na_2HPO_4, volume of RNA used in reverse transcription (10 µL), dilution factor, and volume of cDNA used in qPCR (5 µL).

To estimate the recovery percentage of reverse transcription and qPCR, three older samples (free of SARS-CoV-2) were spiked with SARS-CoV-2 RNA. The samples used for the recovery study were obtained from the three interceptors on 2/16/2018 using the same method mentioned in the study and stored in -80C freezer (Miyani et al. 2020b). Viral RNA was extracted, reverse transcribed and amplified for SARS-CoV-2 using the aforementioned method. Five ul of 5.8E+04 gc/ul SARS-CoV-2 RNA was spiked in the extracted viral RNA samples which were reverse transcribed and amplified for SARS-CoV-2 detection. No amplification of SARS-CoV-2 was observed in the samples before spiking. Additionally, 5 ul of 5.8E+04 gc/ul SARS-CoV-2 RNA was spiked in PCR-grade water, reverse transcribed and amplified for SARS-CoV-2. Recovery percentage of SARS-CoV-2 RNA, for reverse transcription, dilution

and qPCR was observed to be 54%, 46% and 30% for O-NWI, NI-EA and DRI respectively.

The limit of detection of RT-qPCR was determined using 10^{-1}, $10^{-0.5}$, 10^{0}, $10^{0.5}$, 10^{1}, 10^{2}, 10^{3}, 10^{4}, 10^{5} dilutions of positive SARS-CoV-2 RNA gc/ul control based the method described in FDA approved CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel document (CDC Coronavirus 2020). No amplification was observed in 10^{-1}, $10^{-0.5}$ and 10^{0} dilutions of SARS-CoV-2 RNA. The limit of detection was determined to be $10^{0.5}$ SARS-CoV-2 gc/ul.

The qPCR efficiency was found to be 95.76% based on the standard curve and the following equation $E = 10^{(-1/slope)} - 1$.

2.5 Publicly-Available Clinical Data

Publicly-available online data reporting confirmed COVID-19 infections in the City of Detroit, and Wayne, Oakland, and Macomb counties were used for this study (https://www.michigan.gov/coronavirus/). The database was accessed on Feb 6, 2020. The state of MI reports the data as follows: "county" is based on the county of residence; "cases" are aggregated by the date of onset of COVID-19 symptoms, if known, otherwise by laboratory specimen date, if known, otherwise by case referral date; "confirmed cases" only include individuals who have had a positive diagnostic laboratory test for COVID-19. The data are shown in figure 2.2.

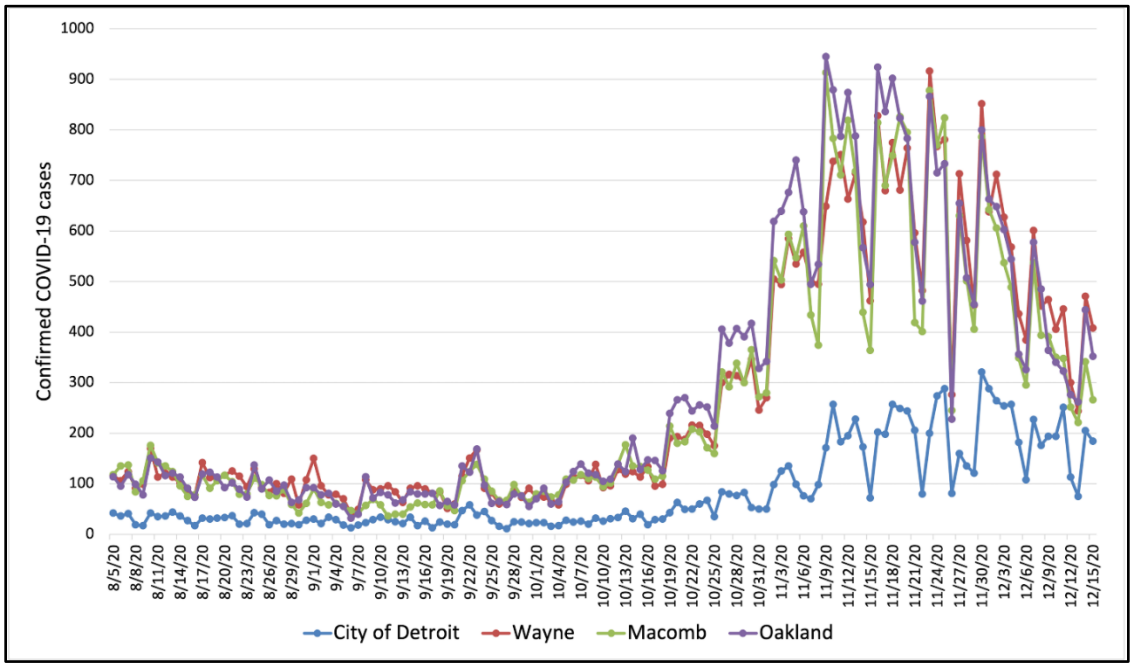

Figure 2.2. Observed COVID-19 cases between August 5 and December 15, 2020

2.6 Interceptor Contributing Populations and Flow Rates

The population served by each interceptor was estimated from 2020 populations provided by the Southeast Michigan Council of Governments by traffic analysis zone (TAZ). Geographic information systems (GIS) analysis was used to intersect the TAZ boundaries with ZIP Code boundaries and proportionally allocate population from each TAZ to the intersected areas. ZIP Code boundaries were also intersected with the interceptor service areas to allow a calculation of population by interceptor area.

Daily flow rates for the three interceptors were estimated from the daily influent flow to the WRRF (calculated from GLWA-reported primary influent flow minus WRRF recycle flows) and a calibrated hydrologic and hydraulic model developed for the GLWA collection system. The collection system model was developed with US EPA's Stormwater Management Model (SWMM) 5 as part of the GLWA Wastewater Master Plan. The SWMM model represents sanitary wastewater and infiltration/inflow, hydraulics in all physical assets of the collection system and at the WRRF entrances, and stages and flows in the Rouge and Detroit rivers.

2.7 Correlations Between COVID-19 Cases and Wastewater SARS-CoV-2 Concentrations

One-Way ANOVA and *Post-Hoc Tukey* test (comparing means) by SPSS was used to evaluate the lag time between SARs-CoV-2 concentrations in wastewater and confirmed clinical cases. Time series analysis was not possible because of data gaps. All the confirmed COVID-19 clinical cases and SARS-CoV-2 RNA qPCR measurements (gc/liter) were pooled together in the following groups: Aug 13-Aug 17, Aug 17-Aug 31, Aug 31-Sept 14, Sept 14-Sept 28, Sept 28-Oct 12, Oct 12-Oct 26, Oct 26-Nov 9, Nov 9-Nov 23, Nov 23-Dec 7 and Dec 7- Dec 21. The reason for choosing biweekly groups was to compensate for the dates on which sampling was not conducted. All the groups were analyzed by One-Way ANOVA and *Post-Hoc Tukey* test to determine statistically significant differences.

3. RESULTS AND DISCUSSION

A total of 144 samples were collected and processed between Aug 6, and Dec 14, 2020. SARS-CoV-2 was detected in 98% of samples. Figure 2.3 illustrates the average and standard deviation of genomic copies per liter of SARS-CoV-2 in all three interceptors. The data are also shown in table 2.2. SARS-CoV-2 concentrations, for all three interceptors, were in the range of 10^4 to 10^6 genomic copies per liter. The lowest and highest concentration, for ONWI, was

observed on Sept 8, 2020 (4.45E+04 copies/L) and Sept 29, 2020 (5.36E+06 copies/L), respectively. For NIEA, the lowest and highest was observed on Dec 14, 2020 (5.2E+04 copies/L) and Oct 14, 2020 (2.65E+06 copies/L), respectively. Comparatively, in the case of DRI, the lowest and highest was observed on Sept 15, 2020 (6.33E+04 copies/L) and October 7, 2020 (3.71E+06 copies/L), respectively. Figure 2.2 shows confirmed COVID-19 cases for the City of Detroit and Macomb, Oakland and Wayne Counties. As seen in figure 2.2, the number of confirmed COVID-19 cases for the second wave in the region peaked around Nov 9, 2020, a month later than the observed peak in SARS-CoV-2 RNA concentrations in wastewater (figure 2.3).

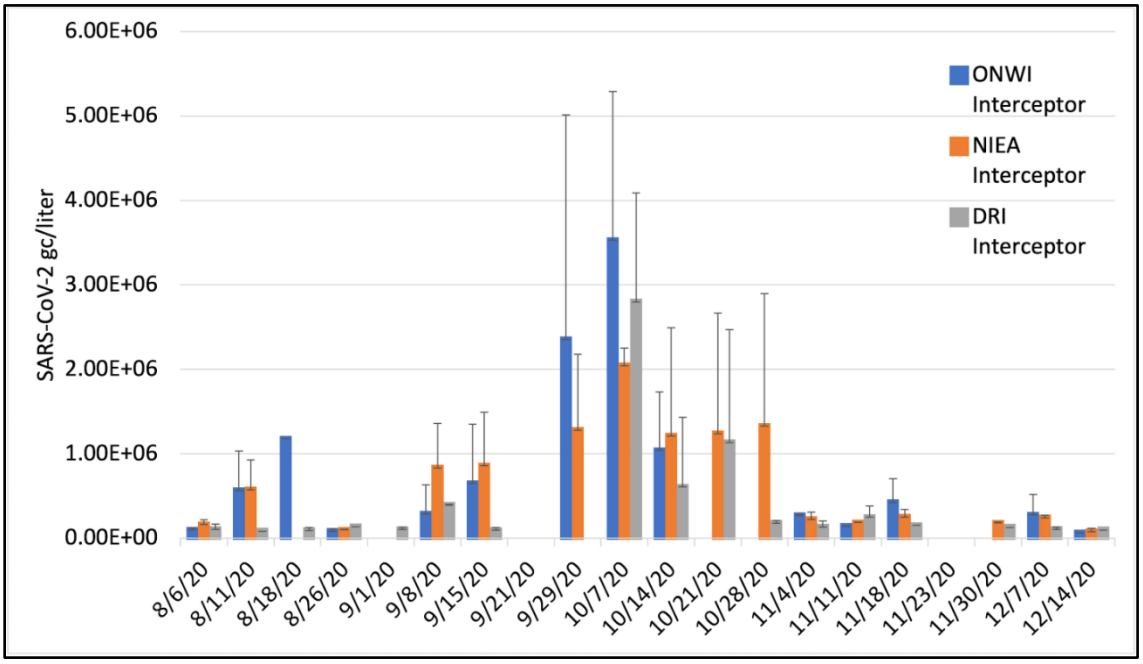

Figure 2.3. SARS-CoV-2 RNA concentrations quantified in the ONWI, NIEA and DRI interceptors

Figures 2.4 and 2.5 show boxplots of biweekly grouped measured SARS-CoV-2 RNA concentrations in wastewater and confirmed COVID-19 cases. For SARS-CoV-2 RNA concentrations (figure 2.4), the Sept 28-Oct 12 group was found to be significantly different compared to all other groups (all p-value = 0). Similarly, for biweekly confirmed COVID-19 cases (figure 2.5), the Oct 26-Nov 9 group was found significantly different compared to all other groups (all p-value = 0). These observations indicate a lag time of approximately one month (Sept 28 to Oct 26).

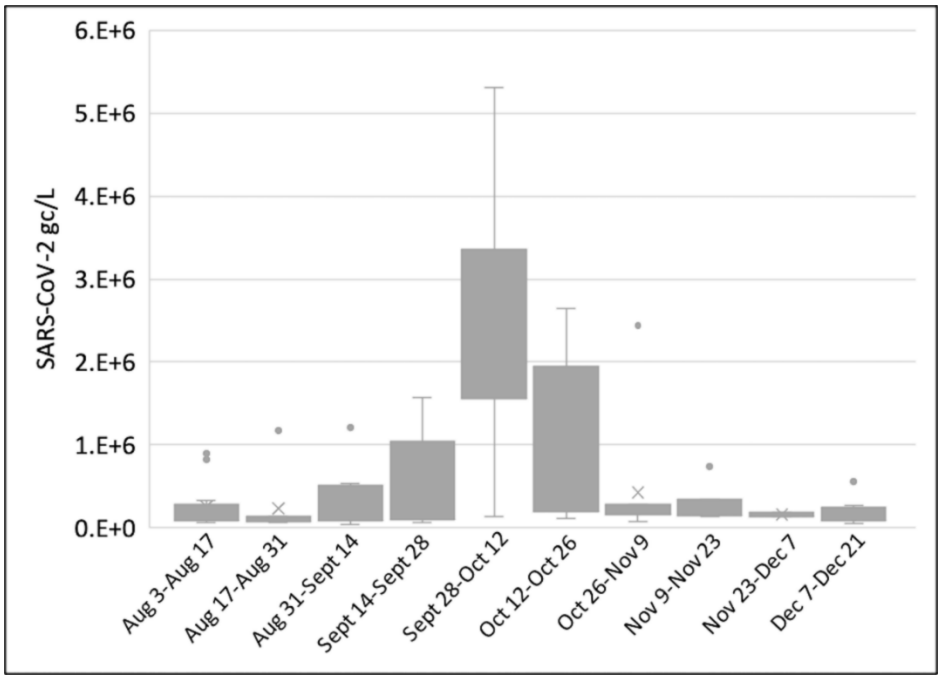

Figure 2.4. Biweekly SARS-CoV-2 concentrations at three interceptors

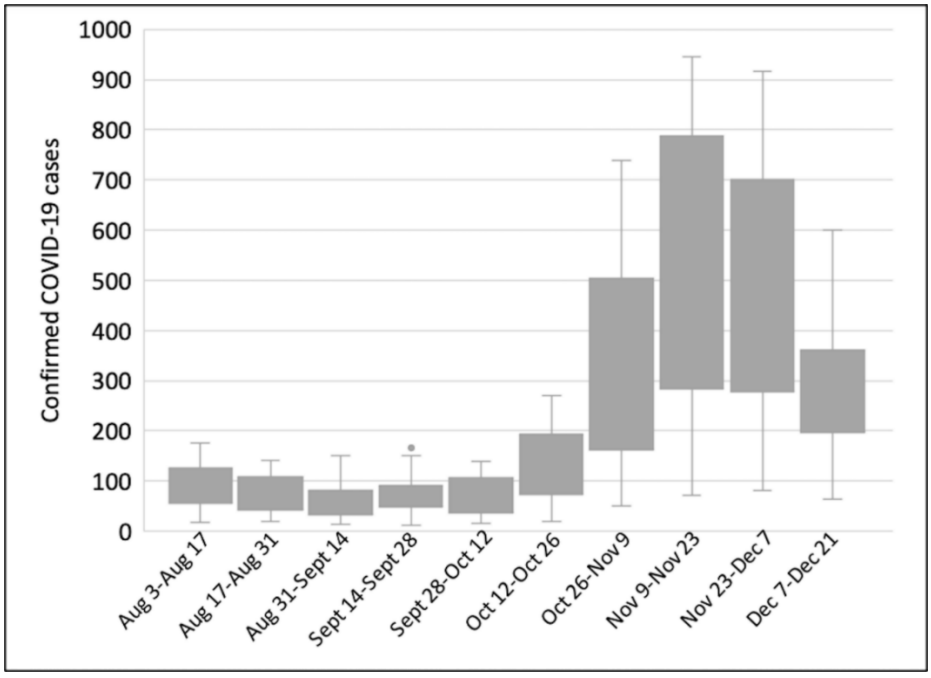

Figure 2.5. Biweekly confirmed COVID-19 cases in City of Detroit and Wayne, Macomb and Oakland Counties

Additionally, qPCR data were normalized (and expressed as genomic copies per day per person) by multiplying with the wastewater flow rate on the specific day and dividing by the population served by each interceptor (figure 2.6). Normalized SARS-CoV-2 RNA

concentrations in the WRRF influent peaked on the same dates as non-normalized concentrations (figure 2.6) both prior to reported COVID-19 cases. However, the magnitude of the normalized and non-normalized peak concentrations is different, eliciting the importance of understanding other parameters relevant to viral fate and transport within the sewer-sheds and the change in population. One-Way ANOVA and *Post-Hoc Tukey* results were obtained for normalized data and, similarly to non-normalized data, the Sept 28-Oct 12 group was significantly different as to compared to other sampling dates for biweekly combined SARS-CoV-2 RNA concentrations.

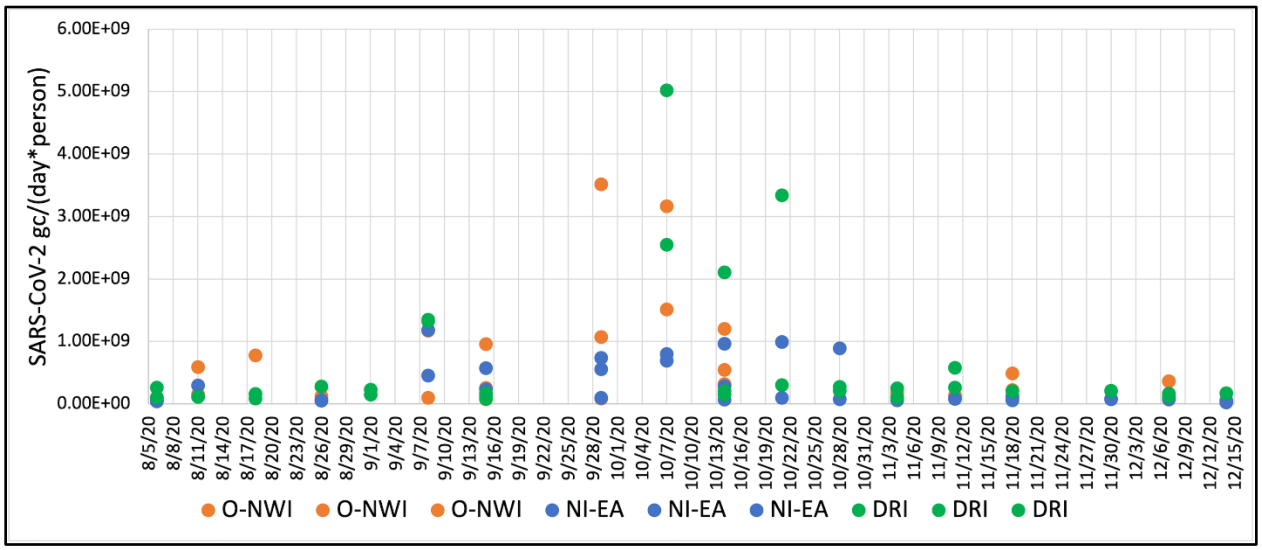

Figure 2.6. Normalized SARS-CoV-2 RNA concentrations in genomic copies per day, per person, gc/(day*person)

In summary, our SARS-CoV-2 RNA measurements suggested early warnings of about four weeks prior to the rise of clinically reported COVID-19 cases in Detroit, Michigan. This agrees with observations by Saguti et al. 2021 and Ahmed et al. 2021. Saguti *et al.* 2021 collected influent wastewater samples, daily, from the city of Gothenburg and surrounding municipalities in Sweden. The daily samples were combined into weekly samples and passed through Nano-Ceram filter which were then concentrated and eluted by milk powder and ultracentrifugation respectively. Saguti *et al.* 2021 found a time lag of three to four weeks between increases in SARS-CoV-2 RNA in wastewater and the number of newly hospitalized patients with COVID-19. Additionally, Ahmed *et al.* 2021 observed SARS-CoV-2 RNA in wastewater from Southern Brisbane wastewater treatment plant, Australia in late February 2020, three weeks before the first clinical case was reported.

Other studies report shorter lag times. Randazzo *et al.* 2020 investigated the occurrence of SARS-CoV-2 RNA in six wastewater treatments plants in Spain, and indicated that the positive influent wastewater samples were detected 12 to 16 days prior to the reported clinical COVID-19 cases in municipalities. Medema *et al.* 2020 collected sewages samples from six cities and an airport in the Netherlands. It was found that SARS-CoV-2 RNA measurements were detected in Amersfoort 6 days before the first cases were reported. Peccia *et al.* 2020 found that the SARS-CoV-2 RNA concentrations in sludge were 0 to 2 days ahead of COVID-19 positive test results. The time difference between SARS-CoV-2 RNA concentrations in sludge to local hospital admissions was 1 to 4 days, whereas 6 to 8 days for reporting of SARS-CoV-2 positive test results. Nemudryi *et al.* 2020 collected samples from the municipal wastewater treatment plant in Bozeman, Montana and was able to identify that SARS-CoV-2 RNA levels in wastewater were leading clinical PCR test results by 2 to 4 days.

The type of samples and their source, viral retention in the wastewater collection network, including dilution, sorption to solids, and decay, and the demographics of the population may greatly influence the observed lag times. At this point generalizations of expected lag times between different locations and populations cannot be made. Accurate determination of lag times and prediction involves the incorporation of many other measurements, data, and processes, including but not limited to, estimation of the dilution, retention and fate of viral particles in the wastewater collection network and the estimation of characteristics of contributing population (Xagoraraki 2020). Establishing early warning of upcoming fluctuations in the greater Metropolitan area of Detroit facilitates planning of resources and further clinical testing. Our team is currently conducting targeted sewer-shed sampling in communities with high cumulative COVID-19 cases per 100.000 people and communities with varying demographics.

4. CONCLUSIONS

Untreated wastewater samples were collected and processed between Aug 6, and Dec 14, 2020, from the Detroit metropolitan area in southeast Michigan. SARS-CoV-2 RNA was detected in 98% of samples with concentrations ranging from 10^4 to 10^6 copies/L. Early warnings of the second wave of COVID19 were observed ahead of clinical data reporting. Statistical analysis indicated a lag time of four weeks between observed SARS-CoV-2 RNA concentrations in wastewater and COVID-19 reported cases in the community. Along with

clinical diagnostic testing, wastewater-based epidemiology may be a helpful resource for health officials in predicting the incidence of SARS-CoV-2 in community. Accurate prediction models can be created by including processes that affect the fate of viruses in the collection network, demographic information, and shedding rate and duration data.

BIBLIOGRAPHY

Ahmed, Warish, Ben Tscharke, Paul M. Bertsch, Kyle Bibby, Aaron Bivins, Phil Choi, Leah Clarke et al. (2021). SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: A temporal case study. *Science of The Total Environment* 761: 144216.

CDC (Centers for Disease Control and Prevention). (2020). Coronavirus (COVID-19). Accessed February 8, 2021. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html

Harvey, A. P., Fuhrmeister, E. R., Cantrell, M. E., Pitol, A. K., Swarthout, J. M., Powers, J. E., Pickering, A. J. (2020). Longitudinal monitoring of SARS-CoV-2 RNA on high-touch surfaces in a community setting. *Environmental Science & Technology Letters*.

Heijnen, L., & Medema, G. (2011). Surveillance of influenza A and the pandemic influenza A (H1N1) 2009 in sewage and surface water in the Netherlands. *Journal of water and health*, *9*(3), 434-442.

Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., ... & Norder, H. (2014). Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. *Applied and environmental microbiology*, *80*(21), 6771-6781.

Karim, M. R., Rhodes, E. R., Brinkman, N., Wymer, L., & Fout, G. S. (2009). New electropositive filter for concentrating enteroviruses and noroviruses from large volumes of water. *Applied and Environmental Microbiology*, *75*(8), 2393-2399.

McCall C., Wu H., Miyani B., Xagoraraki I. (2020) Identification of Multiple Potential Viral Diseases in a Large Urban Center using Wastewater Surveillance. *Water Research*, 184: 116160

McCall, C., Wu, H., O'Brien, E., & Xagoraraki, I. (in press). Assessment of Enteric Viruses During a Hepatitis Outbreak in Detroit MI using Wastewater Surveillance and Metagenomic Analysis. *Journal of Applied Microbiology*.

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., & 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. *Environmental Science & Technology Letters*, *7*(7), 511-516.

Miyani B., Fonoll X., Norton J., Mehrotra A., Xagoraraki I. (2020a) SARS-CoV-2 in Detroit Wastewater. *Journal of Environmental Engineering*, 146(11): 06020004

Miyani B., McCall C., Xagoraraki I. (2020b) High Abundance of Human Herpesvirus 8 in Wastewater from a Large Urban Area, *Journal of Applied Microbiology*, 1455: 1-10

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., ... & Wiedenheft, B. (2020). Temporal detection and phylogenetic assessment of SARS-CoV-2 in municipal wastewater. *Cell Reports Medicine*, *1*(6), 100098.

O'Brien E., Xagoraraki I. (2019) A Water-Focused One-Health Approach for Early Detection and Prevention of Viral Outbreaks. *One Health*, 7: 100094 1-9

Peccia, J., Zulli, A., Brackney, D. E., Grubaugh, N. D., Kaplan, E. H., Casanovas-Massana, A., Omer, S. B. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature Biotechnology*, *38*(10), 1164-1167.

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., & Sánchez, G. (2020). SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water research*, *181*, 115942.

Rimoldi, Sara Giordana, Fabrizio Stefani, Anna Gigantiello, Stefano Polesello, Francesco Comandatore, Davide Mileto, Mafalda Maresca et al. (2020) Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers. *Science of the Total Environment* 744: 140911.

Saguti, F., Magnil, E., Enache, L., Churqui, M. P., Johansson, A., Lumley, D., ... & Norder, H. (2021). Surveillance of wastewater revealed peaks of SARS-CoV-2 preceding those of hospitalized patients with COVID-19. *Water research*, *189*, 116620.

Soto-Beltran, M., Ikner, L. A., & Bright, K. R. (2013). Effectiveness of poliovirus concentration and recovery from treated wastewater by two electropositive filter methods. *Food and environmental virology*, *5*(2), 91-96.

USEPA., (2001). Concentration and processing of waterborne viruses by positive charge 1MDS cartridge filters and organic flocculation. Chap. 14 in *USEPA manual of methods of virology*. Washington, DC: USEPA

Weidhaas, J., Aanderud, Z., Roper, D., VanDerslice, J., Gaddis, E., Ostermiller, J., ... & LaCross, N. (2020). Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds.

Xagoraraki I., (2020) Can we predict viral outbreaks using wastewater surveillance? *Journal of Environmental Engineering*, 146(11): 01820003 - INVITED EDITORIAL

Xagoraraki I., O'Brien E. (2020) Wastewater-Based Epidemiology for Early Detection of Viral Outbreaks. In: Women in Water Quality (O'Bannon D., Editor): *Women in Engineering and Science*, Springer Nature, Switzerland

CHAPTER 3: SURVEILLANCE OF SARS-COV-2 IN NINE NEIGHBORHOOD SEWERSHEDS IN DETROIT TRI-COUNTY AREA, UNITED STATES: ASSESSING PER CAPITA SARS-COV-2 ESTIMATIONS AND COVID-19 INCIDENCE

This chapter was published in Science of The Total Environment Volume 851, Part 2, 10 December 2022, 158350, Yabing Li, Brijen Miyani, Liang Zhao, Maddie Spooner, Zach Gentry, Yangyang Zou, Geoff Rhodes, Hui Li, Andrew Kaye, John Norton, Irene Xagoraraki, Surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States: Assessing per capita SARS-CoV-2 estimations and COVID-19 incidence

ABSTRACT

Wastewater-based epidemiology (WBE) has been suggested as a useful tool to predict the emergence and investigate the extent of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In this study, we screened appropriate population biomarkers for wastewater SARS-CoV-2 normalization and compared the normalized SARS-CoV-2 values across locations with different demographic characteristics in southeastern Michigan. Wastewater samples were collected between December 2020 and October 2021 from nine neighborhood sewersheds in the Detroit Tri-County area. Using reverse transcriptase droplet digital polymerase chain reaction (RT-ddPCR), concentrations of N1 and N2 genes in the studied sites were quantified, with N1 values ranging from 1.92×10^2 genomic copies/L to 6.87×10^3 gc/L and N2 values ranging from 1.91×10^2 gc/L to 6.45×10^3 gc/L. The strongest correlations were observed with between cumulative COVID-19 cases per capita (referred as COVID-19 incidences thereafter), and SARS-CoV-2 concentrations normalized by total Kjeldahl nitrogen (TKN), creatinine, 5-hydroxyindoleacetic acid (5-HIAA) and xanthine when correlating the per capita SARS-CoV-2 and COVID-19 incidences. When SARS-CoV-2 concentrations in wastewater were normalized and compared with COVID-19 incidences, the differences between neighborhoods of varying demographics were reduced as compared to differences observed when comparing non-normalized SARS-CoV-2 with COVID-19 cases. This indicates when studying the disease burden in communities of different demographics, accurate per capita estimation is of great importance. The study suggests that monitoring selected water quality parameters or biomarkers, along with RNA concentrations in wastewater, will allow adequate data normalization for spatial comparisons, especially in areas where detailed sanitary sewage flows and contributing populations in the catchment areas are not available. This opens the possibility of using WBE to

assess community infections in rural areas or the developing world where the contributing population of a sample could be unknown.

1. INTRODUCTION

It has been more than two years since the start of the coronavirus disease 2019 (COVID-19) pandemic which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). During this period, scientists from multiple disciples including epidemiology, evolutionary biology and environmental engineering have come a long way in understanding the distribution, transmission, genomic characteristics and risks of SARS-CoV-2 (Ahmed et al., 2022a; Ahmed et al., 2022c; Koelle et al., 2022; Laborde et al., 2020; Li et al., 2022; Lu et al., 2020; Sills et al., 2020; Subbaraman, 2021). Although our understanding about SARS-CoV-2 has improved, challenges ahead remain, with new emerging variants of concern, high transmissibility and COVID-19 risks, and developing control strategies in rural areas and the developing world.

Wastewater-based epidemiology (WBE) is suggested to be an important and useful tool to investigate the prevalence and spatial distribution of viruses such as hepatitis virus, poliovirus noroviruses in the environment. The corresponding wastewater surveillance strategy can serve as an "early warning" system to provide scientific evidence for public health decision-making (Ahmed et al., 2022b; Ahmed et al., 2021b; Bibby et al., 2021; Hovi et al., 2012; Lodder et al., 2012; McCall et al., 2021; Miyani et al., 2020; Xagoraraki and O'Brien, 2020; Zhao et al., 2022). Wastewater surveillance systems have been launched to monitor SARS-CoV-2 RNA concentrations and transmission in the various environments, when coping with the COVID-19 pandemic (Ahmed et al., 2021a; Gerrity et al., 2021; Medema et al., 2020; Wu et al., 2021; Xie et al., 2022). Studies associated with SARS-CoV-2 vary in scales, ranging from small campuses (Betancourt et al. 2021; Gibas et al 2021) to the large national scale programs (Bivins and Bibby, 2021; Gonzalez et al., 2020; Li et al., 2021b; Miyani et al., 2020; Miyani et al., 2021; Scott et al., 2021; Wu et al., 2021; Wu et al. 2021; Zhao et al., 2022). By conducting WBE coupled with targeted clinical testing in 13 dorms in the University of Arizona, both SARS-CoV-2 RNA in the wastewater and COVID-19 cases in the community were investigated, and the results provide evidence for the application of WBE in defined communities (Betancourt et al., 2021). A similar building-level SARS-CoV-2 surveillance study was implemented and found that identification of positive COVID-19 cases can be indicated by WBE (Gibas et al., 2021).All these instances

validate the application of WBE as a promising and cost-effective tool to assist in public health responses to disease outbreaks.

However, a critical factor associated with successful application of WBE is the accurate population estimation (Sims and Kasprzyk-Hordern, 2020). It is highly recommended by the United States Centers for Disease Control and Prevention (CDC) that normalizing the viral concentrations by the number of people served by the sewer system enables the viral levels comparisons across different sampling locations (CDC, 2022). The variability of population size within a catchment area is dynamic due to many factors such as commuting and tourism. This contributes to various uncertainties in real-time WBE. Additionally, if WBE were to be applied in rural areas or the developing world, where information on the characteristics of catchment areas and populations may not be available, estimating the contributing population will be of great importance.

Although both viral (e.g., pepper mild mottle virus, PMMoV) and human biomarkers have been suggested to normalize the SARS-CoV-2 RNA in the wastewater, however, it is still debated which marker is the most ideal to perform the normalization (Xie et al., 2022). In a most recent study, using coprostanol as a potential biomarker, the strongest correlation between normalized SARS-CoV-2 RNA concentrations and COVID-19 incidences was observed, which demonstrated the promising application of population biomarkers in wastewater surveillance studies (Reynolds et al., 2022). Additionally, 5-HIAA was selected to assess population size in an exploratory study, in which surveillance of SARS-CoV-2 RNA in wastewater from two municipalities in Latvia were discussed (Gudra et al., 2022). To evaluate the application of population biomarkers in normalizing SARS measurements, one can review the efforts and conclusions drawn of the study of drug abuse estimation using WBE, where biomarkers have been widely studied all over the world (Choi et al., 2018).

Two approaches were commonly used to estimate population sizes and compare with census data: measurement of water quality parameters in sewage (biochemical oxygen demand [BOD], nitrogen, etc.) (Tscharke et al., 2019; Zheng et al., 2019) and measurement of population biomarkers (Chen et al., 2014; Gracia-Lor et al., 2017). The advantage of using water quality parameters is that they can be easily measured and tracked, which makes them easily applicable especially in rural areas and the developing world. However, water quality parameters, such as BOD, total nitrogen (TN) or total phosphorous (TP) levels may be influenced by both human and

non-human contributions (e.g., industrial discharges, agricultural activities) (Been et al., 2014; Daughton, 2012; Rico et al., 2017). However, population biomarkers in sewage, which are either endogenous compounds, such as creatinine, 5-HIAA, or exogenous substances like caffeine and its metabolites, are contributed by human metabolism only and, therefore, may be surrogates for estimating sewershed populations (Chen et al., 2014; Chiaia et al., 2008; Choi et al., 2018). Compared with the census approach, they may reflect the real-time fluctuation of population.

For example, a breakdown compound of muscle tissues, creatinine was proposed as a potential population biomarker and is widely used clinically (Chen et al., 2014; Chiaia et al., 2008). However, degradability of creatinine in sewer systems was reported, and this affects its potential as a population biomarker (Thai et al., 2014). Neurotransmitter metabolite 5-HIAA was suggested as an eligible population biomarker that is cultural independent and may be reliable for comparisons of population sizes among different countries (Chen et al., 2014). Caffeine is one of the most ubiquitous micro-contaminants found in the untreated wastewater due to its wide usage in many globally popular products (e.g., tea, cola drinks or coffee) as a stimulating agent. It has been suggested as a human biomarker for assessing the real-time fluctuations of population (Daughton, 2012; Froehner et al., 2010). Metabolism of caffeine is extensive and at least 17 urinary metabolites of caffeine are identified (Gracia-Lor et al., 2017). Occurrences of caffeine metabolites in wastewater are still scarce, which limit their application in population estimation. Meanwhile, concentrations of the metabolites depend on human habits, it was found that the average loads of caffeine and its metabolites were slightly lower during the weekends, which may result from the relatively lower consumption of coffee (Senta et al., 2015).

This study focuses on the evaluation of the relationship between the COVID-19 incidences and the measured and estimated per capita SARS-CoV-2 RNA concentrations. It is hypothesized that population normalization with water quality markers and population biomarkers could be a reliable approach to calculate the per capita SARS-CoV-2 RNA concentrations. To test the hypothesis, we collected 486 samples from nine sewersheds of different demographics in the Detroit Tri-County area, MI from December 2020 through October 2021, and quantified SARS-CoV-2 RNA concentrations at a neighborhood level. The population of each community was assessed using water quality parameters (BOD, TKN, total suspended solids [TSS], volatile suspended solids [VSS]) and human biomarkers (creatinine, 5-HIAA, caffeine and its metabolites), which was further validated with the census-reported population in

the catchment areas. Correlations between COVID-19 incidences and per capita SARS-CoV-2 RNA concentrations were performed to identify relatively reliable population biomarkers that can be applied in sewage surveillance studies.

Moreover, using normalized SARS-CoV-2 concentration and COVID-19 incidences, spatial differences between nine sewersheds of different demographics were assessed. Previously, SARS-CoV-2 RNA concentrations in wastewater collected from the Great Lakes Water Authority (GLWA) Water Resource Recovery Facility (WRRF) influent in Detroit, Michigan were quantified and ranged from 10^4~10^5 gc/L (Miyani et al., 2020). Considering sewage surveillance can capture the presence of virus before the onset of symptoms, an "early warning" system was proposed to try and forecast the second COVID-19 wave in the Detroit metropolitan area by combining the sewage surveillance methodology with local public health records (Miyani et al., 2021; Richardson, 2021). However, sewage surveillance research of SARS-CoV-2 incorporating the contributed population of the studied sewersheds is still limited. To our knowledge, a spatial comparison across the Detroit neighborhoods sewersheds has not been performed. Such spatial investigations of SARS-CoV-2 RNA that take into consideration the social-demographic characteristics of various communities within the Detroit Tri-County would provide important information for understanding county-level comparisons of COVID-19 incidences as well as inform the public health decision-making process.

The study suggests that monitoring selected water quality parameters or biomarkers along with RNA concentrations in wastewater will allow adequate data normalization, especially in areas where detailed sanitary sewage flows in the catchment areas are not available. This opens the possibility of using WBE to assess community infections in rural areas or the developing world where the contributing population of a sample could be unknown.

2. MATERIAL AND METHODS

2.1. Sampling sites and sample collection

The Detroit metropolitan area, often referred to as Metro Detroit, is a major metropolitan area in the U.S. and the largest in the state of Michigan. It is known for its developed economics and cultural diversity. The City of Detroit serves as the metropolitan area's core, and the metropolitan area extends into three adjacent counties: Macomb, Oakland, and Wayne. This area is also referred as the Detroit Tri-County area. Social and demographic characteristics of communities in this area are varied. Although the underlying mechanisms of disparities are

unknown, significant racial/ethnic differences in COVID-19 cases in the U.S. were revealed and social-demographic factors including economic status, racial/ethnic status, household composition should be included when assessing the COVID-19 burdens (Karmakar et al., 2021; Kim and Bostwick, 2020).

Sewer wastewater samples from three locations in Macomb County (EP, MT, SH), three locations in the City of Detroit, located in Wayne County (D1, D2, D3) and three locations in Oakland County (SF, WB, OP) were collected to conduct the neighborhood sewershed surveillance of SARS-CoV-2 in the Detroit Tri-County area (figure 3.1). Sampling locations were selected in the Detroit Tri-County area to ensure data from neighborhoods with varying demographics. Census tract level population and demographic information obtained from the Southeast Michigan Council of Governments (SEMCOG) website (https://semcog.org/) were evaluated along with sample location catchment areas to define sewershed level demographics (table 3.1).

Site	County	City/Township	ZIP Code
EP	Macomb	Eastpointe	48021
MT	Macomb	Macomb Township	48044,48042,48049,48095
SH	Macomb	Sterling Heights	48310,48312,48313
D1	Wayne	Detroit	48205
D2	Wayne	Detroit	48210
D3	Wayne	Detroit	48235
SF	Oakland	Southfield	48076
WB	Oakland	West Bloomfield	48322
OP	Oakland	Oak Park	48237

Figure 3.1. Locations of nine sewershed sites selected from the Detroit Tri-County area in Michigan in the United States. Decisions were made with local health departments and sewersheds selected represent different demographic characteristics from County Macomb (EP, MT and SH), Wayne (D1, D2 and D3) and Oakland (SF, WB and OP) in the Detroit Tri-County area MI

Sampling locations in the City of Detroit (D1, D2 and D3) have relatively higher population densities and higher poverty rates and relatively lower household income than those in Macomb County (EP, MT and SH) and Oakland County (SF, WB and OP). Sample locations

MT and SH in Macomb County have the largest catchment areas and populations. The nine sample locations include a demographic variety in the Detroit Tri-County area, and represent areas with differing racial makeups, education levels, poverty, and income levels (table 3.1).

Table 3.1. Demographic characteristics of the catchment areas for the nine sampling sewersheds in the Detroit Tri-County area MI

Sample Tributary Catchment											
Site	Area (Acres)	Population	Density[1]	Asian (%)	Black (%)	Hispanic (%)	White (%)	Poverty (%)	65+[2] (%)	Degree[3] (%)	Total Household Income ($)
EP	278	2,400	8.6	0	37	5	54	5	16	17	56,450
MT	29,264	99,970	3.4	3	5	3	88	4	14	34	102,850
SH	6,246	37,560	6.0	10	5	2	80	12	17	27	65,700
D1	135	1,690	12.5	0	95	0	4	19	11	5	39,300
D2	372	5,190	14.0	0	8	76	14	32	6	5	35,900
D3	127	1,300	10.2	0	95	0	2	44	17	14	22,100
SF	717	3,080	4.3	0	61	3	31	10	18	56	92,100
WB	1,218	5,800	4.8	15	19	0	63	5	27	55	100,520
OP	286	2,270	7.9	2	85	3	6	15	16	20	51,680

[1]: Unit of population density, people per acre.

[2]: Percent of population older than 65.

[3]: Percent of population with bachelor's degree or higher.

Sewer wastewater samples were collected during 18 sampling events from December 2020 to October 2021. For each site, triplicate samples were collected each sampling event for a total of 54 samples collected for each site. Dates for each sampling event are shown in the supplementary table 3.2. Sewer flow values for each sampling event were measured by existing flow meters at two sites in Macomb County (MT and SH) and two sites in Oakland County (SF and WB). A hydraulic model [GLWA's Reginal Wastewater Collection System (RWCS) Model] was updated with rainfall data for the sampling period and applied to estimate flow values for sewersheds D1, D2 and D3 in the City of Detroit (Wayne County) for each sampling event.

There was no available flow meter data or hydraulic model that could be used to estimate flows for sites EP in Macomb County and OP in Oakland County for the study period.

Table 3.2. Specific dates of sampling events occurred for the nine neighborhood sewersheds (County Macomb: EP, MT and SH; County Wayne: D1, D2 and D3; and County Oakland: SF, WB and OP) in the Detroit Tri-County Area MI

Week	EP	MT	SH	D1	D2	D3	SF	WB	OP
12/7/20	12/9/20	12/9/20	12/9/20	12/10/20	12/10/20	12/10/20	12/8/20	12/8/20	12/9/20
12/14/20	12/17/20	12/17/20	12/17/20	12/15/20	12/15/20	12/15/20	12/17/20	12/18/20	12/18/20
1/18/21	1/20/21	1/19/21	1/19/21	1/20/21	1/20/21	1/20/21	1/21/21	1/21/21	1/21/21
1/25/21	1/27/21	1/26/21	1/26/21	1/27/21	1/27/21	1/27/21	1/28/21	1/28/21	1/28/21
2/1/21	2/3/21	2/2/21	2/2/21	2/3/21	2/3/21	2/3/21	2/4/21	2/4/21	2/4/21
2/8/21	2/10/21	2/9/21	2/9/21	2/10/21	2/10/21	2/10/21	2/11/21	2/11/21	2/11/21
3/1/21	3/2/21	3/2/21	3/2/21	3/3/21	3/3/21	3/3/21	3/4/21	3/4/21	3/4/21
3/8/21	3/9/21	3/9/21	3/9/21	3/10/21	3/10/21	3/10/21	3/11/21	3/11/21	3/11/21
3/15/21	3/16/21	3/16/21	3/16/21	3/17/21	3/17/21	3/17/21	3/18/21	3/18/21	3/18/21
3/22/21	3/23/21	3/23/21	3/23/21	3/24/21	3/24/21	3/24/21	3/25/21	3/25/21	3/25/21
8/30/21	8/31/21	8/31/21	8/31/21	9/1/21	9/1/21	9/1/21	9/2/21	9/2/21	9/2/21
9/6/21	9/7/21	9/7/21	9/7/21	9/8/21	9/8/21	9/8/21	9/9/21	9/9/21	9/9/21
9/13/21	9/14/21	9/14/21	9/14/21	9/15/21	9/15/21	9/15/21	9/16/21	9/16/21	9/16/21

Table 3.2. (cont'd)

9/20/21	9/21/21	9/21/21	9/21/21	9/22/21	9/22/21	9/22/21	9/23/21	9/23/21	9/23/21
9/27/21	9/28/21	9/28/21	9/28/21	9/29/21	9/29/21	9/29/21	9/30/21	9/30/21	9/30/21
10/4/21	10/5/21	10/5/21	10/5/21	10/6/21	10/6/21	10/6/21	10/7/21	10/7/21	10/7/21
10/11/21	10/12/21	10/12/21	10/12/21	10/13/21	10/13/21	10/13/21	10/14/21	10/14/21	10/14/21
10/18/21	10/19/21	10/19/21	10/19/21	10/20/21	10/20/21	10/20/21	10/21/21	10/21/21	10/21/21

Viruses were collected and isolated from wastewater using electropositive NanoCeram column filters (Argonide, Sanford, FL, USA) based on the EPA Virus Adsorption-Elution (VIRADEL) method (Miyani et al., 2020; Miyani et al., 2021; Xagoraraki et al., 2014; Zhao et al., 2022). Specifically, depending on the quantity of suspended solids in the wastewater, approximately 20 to 50 L of raw wastewater were passed through NanoCeram electropositive cartridge filters at a rate not more than 11 L/min. Flow meter readings were recorded at the beginning and termination of each sampling event to measure the total volume raw wastewater that passed through the filter. After sampling, the NanoCeram column filters were placed in sealed plastic bags on ice and transported to the laboratory for elution and downstream molecular analysis within 24 hours. In addition, for each sampling event, triplicate grab samples of raw sewage were collected using 1L autoclavable polythene plastic bottles. To prevent degradation of biomarkers, the pH in the bottles was adjusted to 2. The bottle was transported to the laboratory within 24 hours for the biomarkers analysis. Solid phase extraction for biomarkers analysis was performed within two weeks of sample collection. Another set of samples were collected for each sampling event and transported to Paragon Laboratories, Inc. in Livonia, MI for the analyses of water quality markers (BOD, TKN, TSS and VSS). Standards methods were applied to perform the analyses as follows: BOD: SM5210B; TKN: SM 4500-N_{org} B; TSS: SM 2540D and VSS: SM 2540 E.

2.2. Quantification of SARS-CoV-2 in the wastewater samples

2.2.1 Virus elution and RNA extraction

Virus elution from the cartridge filters was conducted within 48 hours of each sampling event. Viruses were eluted based on a previously described method (Miyani et al., 2020; Miyani et al., 2021) using 1.5% beef extract (0.05 M glycine). Following elution, each sample was aliquoted into multiple 2 mL Corning tubes. Subsequently, 140 µL of sample was used from one of the corning tubes for RNA extraction. Viral RNA was extracted using QIAGEN QIAamp Viral RNA QIAGEN kits (QIAGEN, Hilden, Germany), following the manufacturer's protocol with the volume of final eluting reagent (buffer AVE) modified from 60 µL to 140 µL as in the previous study (Miyani et al., 2021; Zhao et al., 2022). Bacteriophage Phi6 was spiked to estimate the losses of virus in elution and concentration. And the recoveries were decided from 10.37% to 58.96%, with a mean recovery of 24.91% (±22.89%). RNA extracts were stored at -80 °C and RT-ddPCR were performed within 24 hours after the extraction.

2.2.2 Reverse transcriptase droplet digital PCR (RT-ddPCR)

A QX200 AutoDG Droplet Digital PCR system (Bio-Rad, Hercules, CA, USA) was applied to perform RT-ddPCR and the One-step RT-ddPCR Advanced Kit was used for Probes (Bio-Rad, Hercules, CA, USA). Primers and probe targeting N1 and N2 of SARS-CoV-2 used were summarized in table 3.3. The N1 N2 gene Duplex Assay Reaction Mixture was prepared with 5.5 µL of One-Step RT-Supermix (20x) (final volume ratio: 0.25), 2.2 µL of Reverse Transcriptase (RT) (final volume ratio: 0.1), 1.1 µL 300 mM DTT (final volume ratio: 0.05), 3.3 µL of N1 primer probe mix (final volume ratio: 0.15), 3.3 µL N2 primer probe mix (final volume ratio: 0.15), and 1.1 µL of PCR-grade water (final volume ratio: 0.05) in a final volume of 16.5 µL per reaction. Amounts of the mixture was prepared according to the sample number. After mixing thoroughly, the reagents were pipetted into each well of a 96-well plate. Then, 5.5 µL of RNA product was added to each well reaching a total reaction volume of 22 µL.

Table 3.3. Primers and probe targeting at gene N1 and N2 of SARS-CoV-2 used in reverse transcriptase droplet digital PCR (RT-ddPCR)

Primers/probe	N1	N2
Forward primer	GAC CCC AAA ATC AGC GAA AT	TTA CAA ACA TTG GCC GCA AA
Reverse primer	TCT GGT TAC TGC CAG TTG AAT CTG	GCG CGA CAT TCC GAA GAA
Probe	[FAM] ACC CCG CAT TAC GTT TGG TGG ACC [BHQ1]	[FAM] ACA ATT TGC CCC CAG CGC TTC AG [BHQ1].

The 96-well plate was sealed on a PX1 PCR Plate Sealer (Bio-Rad, Hercules, CA, USA), subsequently vortexed and centrifuged at 1000 rpm for 30 seconds. Oil droplets were generated using an Automated Droplet Generator (Bio-Rad, Hercules, CA, USA). Samples were then run on a C1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA, USA) using the following conditions for the N1 N2 Duplex: 25 ℃ for 3 min, 50 ℃ for 60 min, 95 ℃ for 10 min, following 40 cycles of 95 ℃ for 30 s and 55 ℃ for 1 min with a ramp speed of 2 ℃/s, 98 ℃ for 10 min, and finally 4℃ until the next step. Plates were transferred to a QX200 Droplet Reader (Bio-Rad, Hercules, CA, USA) for a measurement of fluorescence in each droplet.

For each RT-ddPCR run, three positive controls (PTCs) and three negative controls (NTCs), and process negative controls (including virus elution and RNA extraction process controls) were included. Twist Bioscience Twist Synthetic SARS-CoV-2 RNA Control 2 (MN908947.3) with a concentration of 10^2 gc/µL was used for PTCs. Nuclease-free water was used for NTCs. Nano-pure water was used as a substitute for 1.5% beef extract in virus elution, as process negative control. Sterile nuclease-free water was used as a substitute for 140 µL of sample for RNA extraction, as process negative control. All samples were run in triplicate.

Determination of Limit of Blank (LOB) and Limit of Detection (LOD) was based on the methods described in the manufacturer's (Bio-Rad) guidelines for evaluating analytical sensitivity and validation of RT-ddPCR (Bio-Rad, Hercules, CA, USA). The Limit of Blank (LOB) was determined by testing three types of samples using RT-ddPCR, across four consecutive days including the prior-to COVID-19 pandemic samples collected from the same interceptors, nuclease-free water, and negative process control samples from elution and extraction processes. The LOB for N1 gene ddPCR was determined as 0.09 gc/µL, and the LOB

for N2 gene ddPCR was 0.08 gc/µL. Furthermore, an LOD of 0.1 gc/µL with 72.9% confidence for the N1 gene and 0.1 gc/µL with 81.3% confidence for the N2 gene were determined.

2.3. Population biomarkers analysis

Population biomarkers were extracted from wastewater using solid-phase extraction (SPE). Specifically, 19.95 mL of wastewater sample passed through pre-conditioned 200 mg/6cc Waters Oasis HLB cartridges connected to a 12-port vacuum manifold. The SPE cartridges were conditioned by sequentially washing with 5 mL of methanol and 5 mL of deionized (DI) water. The flow rate was controlled at approximately 2 mL/min. After the wastewater sample passed through the cartridge, the sample vial was rinsed with 3 mL of DI water and the rinse water was loaded to the SPE cartridge. Clean vials were placed in the vacuum manifold beneath each SPE cartridge to collect the eluate. The cartridge was eluted with 5 mL of methanol at a rate of 1 mL/min. The samples were analyzed within two weeks using a Shimadzu Prominence high-performance liquid chromatography coupled to a Sciex 4500 triple quadruple mass spectrometer (LC-MS/MS) within two weeks.

The target population biomarkers included creatinine, 5-HIAA, caffeine and its metabolites xanthine, methylxanthine, theobromine, paraxanthine and theophylline. These biomarkers were analyzed by the LC-MS/MS under the positive ionization mode. The ionspray voltage, temperature, curtain gas pressure and entrance potential were 5000 V, 700 °C, 20 psi and 10 V. An Agilent Eclipse Plus C18 column (50 mm × 2.1 mm, particle size 5 µm) was used for separation. The binary mobile phase consisted of phase A water containing 0.3% formic acid and phase B methanol (acetonitrile) with a flow rate of 0.3 mL/min. The injection volume was 5 µL. In the tandem mass spectrometer, multiple reaction monitoring acquisition mode was set up for precursor and product ion transitions, and the highest-intensity transition was selected for quantification. The biomarker concentrations in the extracts were quantified against the external calibration curves.

2.4. Sources of demographic and clinical data

Demographic and socioeconomic data were downloaded from the Southeast Michigan Council of Governments (SEMCOG) website (https://maps.semcog.org/CommunityExplorer/). Daily clinically confirmed cases data for sites in Macomb County (EP, MT and SH) were downloaded from Macomb County Health Department (https://mcmap.maps.arcgis.com/apps/dashboards/439123f4ca934ceb893927776fcda9f8), and

for sites in Oakland County (SF, WB and OP) were downloaded from Oakland County Michigan (https://www.oakgov.com/covid/casesByZip.html), for sites in Wayne County (D1, D2 and D3), daily clinically confirmed cases data were downloaded from Wayne County Michigan (https://www.waynecounty.com/covid19dashboard/). Sum of the weekly/monthly new-confirmed cases for the duration of the 18 sampling events were calculated and referred as cumulative clinical cases thereafter.

2.5. Statistical analysis

Concentrations and loads of SARS-CoV-2 RNA targeting at both N1 and N2 were calculated using formula 3.1. R function Shapiro test was applied to check the normality of the data obtained in this study. Mean values of SARS-CoV-2 for the nine sewersheds were compared among the nine neighborhood sewersheds using one-way analysis of variance (ANOVA), followed by a Tukey's post hoc test to explore the differences of SARS-CoV-2 burdens within the nine sewersheds in the Detroit Tri-County area. Pearson correlation coefficients were applied to assess the relationship between population markers and the population sizes estimated from analysis of census data. Normalized SARS-CoV-2 RNA concentration were obtained through using SARS-CoV-2 RNA concentration divided by concentration of the population marker. Furthermore, correlations between the normalized SARS-CoV-2 RNA concentrations and COVID-19 incidences were also conducted to assess the COVID-19 burden in communities of the Detroit Tri-County area. All the calculations and statistical analysis were performed using Microsoft Excel and R version 4.1.1.

$$C_S = C_{PCR} \times \frac{V_{EL} \times \frac{1000\mu L}{mL}}{V_S} \times \frac{V_{EX\prime}}{V_{EX}} \times \frac{V_{PCR\prime}}{V_{PCR}} \qquad (3.1)$$

C_S: Concentrations of SARS-CoV-2 RNA in samples collected; gc/L.

C_{PCR}: Concentration of SARS-CoV-2 RNA obtained from RT-ddPCR instrument; gc/μL.

V_{EL}: Sample volume after elution, Na_2HPO_4 solution (0.15 M) was used to dissolve virus precipitate and the volume is 30 mL.

V_S: Sample volume collected by cartridge filters, and value was calculated from pump readings and converted to number with unit of L.

V_{EX}: Volume of sample used for RNA extraction; 140 μL.

$V_{EX\prime}$: Volume of RNA products after RNA extraction; 140 μL.

V_{PCR}: Volume of RNA product used for RT-ddPCR; 5.5 μL.

V_{PCR},: Final reaction volume of RT-ddPCR; 22 µL.

3. RESULTS AND DISCUSSION

3.1. Prevalence of SARS-CoV-2 and clinical COVID-19 cases in the communities

SARS-CoV-2 concentrations were quantified using N1 and N2 RT-ddPCR assays for samples collected from Macomb County Sites (EP, MT and SH), City of Detroit (Wayne County) (D1, D2 and D3) and Oakland County sites (SF, WB and OP) in the Detroit Tri-County area. Concentrations of both N1 and N2 in all the studied sites were found to be greater than 1.91×10^2 gc/L and up to 6.87×10^3 gc/L. Average concentration of N1 for the nine sample sites ranged from 348 ± 14.2 gc/L in D1 to 937 ± 110 gc/L in MT. The average concentration of N2 ranged from 372 ± 17.0 gc/L in D1 to 879 ± 161 gc/L in SH. Concentrations of SARS-CoV-2 in MT and SH were relatively higher than those at other sites, which may relate to the larger population in these sewersheds. Tukey's post hoc analysis support this observation and indicated concentrations of SARS-CoV-2 in MT and SH are significantly higher than those in the remaining sites (figure 3.2A and 3.2B). Meanwhile, N1 and N2 measurements were strongly correlated ($r>0.76$) for every site except site SF, where more high outliers of N2 measurements were found compared with N1 measurements.

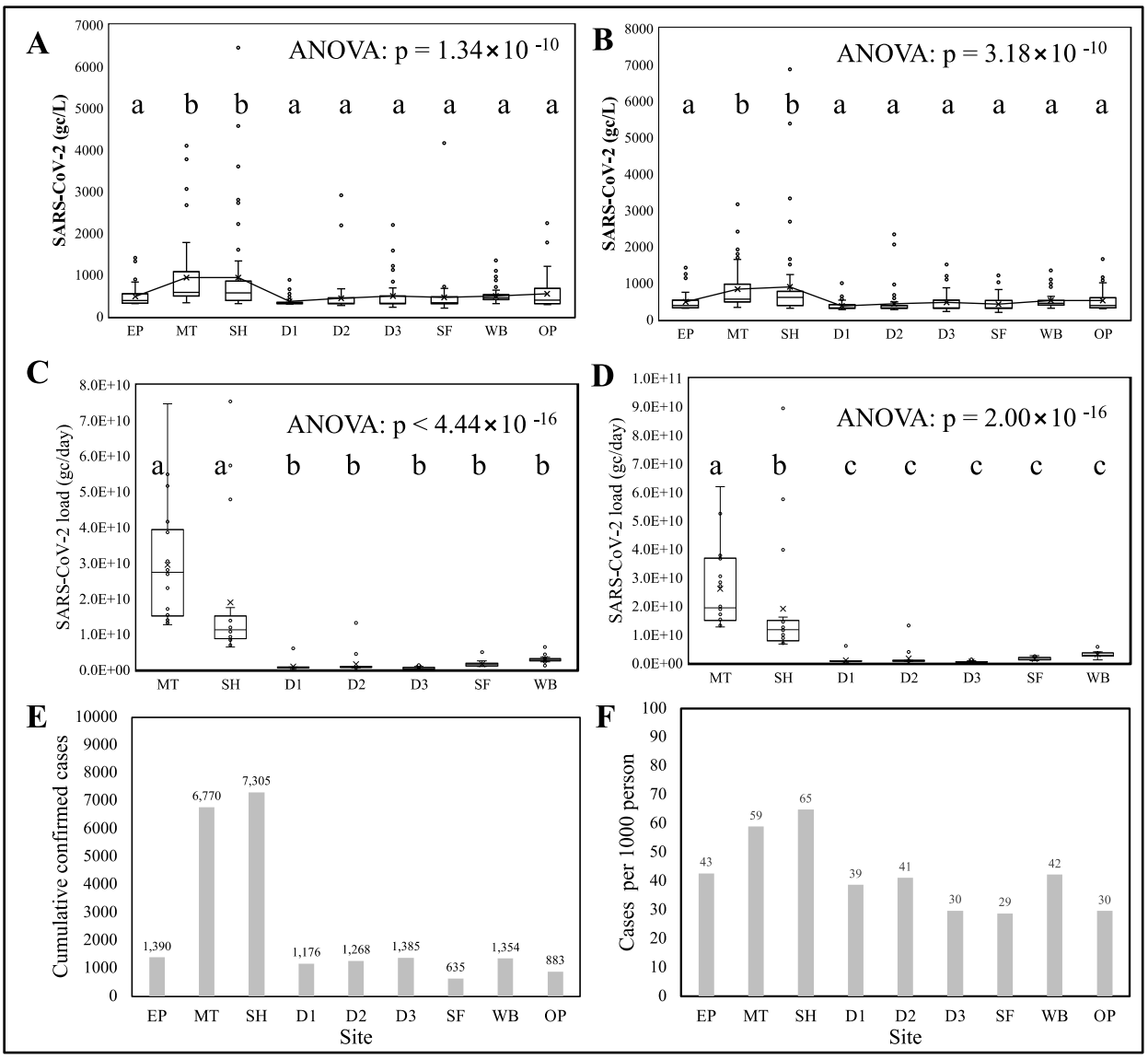

Figure 3.2. A&B: Concentrations of SARS-CoV-2 (A: N1, B: N2) detected in samples collected from sites of County Macomb (EP, MT and SH), Wayne (D1, D2 and D3) and Oakland (SF, WB and OP) in Detroit Tri-County area, MI. One-way analysis of variance (ANOVA), followed by a Tukey's post hoc test indicates SARS-CoV-2 RNA concentrations in sewershed MT and SH are significantly higher than those in the remaining sewersheds. **C&D:** Loads of SARS-CoV-2 (C: N1, D: N2) calculated for sites with flow data available (Macomb: MT and SH, Wayne: D1, D2 and D3, and Oakland: SF and WB). EP and OP are absent because their flow rates are unavailable. MI. One-way analysis of variance (ANOVA), followed by a Tukey's post hoc test indicates loads of SARS-CoV-2 in sewershed MT and SH are significantly higher than those in the remaining sewersheds. **E:** Cumulative clinical confirmed cases for the studied period

Figure 3.2. (cont'd)

corresponding to occurrence of the 18 sampling events. **F:** The COVID-19 cases rate expressed as cases per 1000-person in nine neighborhood sewersheds (Macomb County: EP, MT and SH; City of Detroit (Wayne County): D1, D2 and D3; and Oakland County: SF, WB and OP) in the Detroit Tri-County area (F)

Flow rates of seven sewersheds [MT and SH in Macomb County, D1, D2 and D3 in the City of Detroit (Wayne County), and SF and WB in Oakland County] were available and used to investigate SARS-CoV-2 loads in the study period. Flow data for the four sites located at existing flow meters (MT, SH, SF, and WB) were available in 5-minute intervals. The hydraulic model used to estimate flows for the three City of Detroit sites (D1, D2, and D3) had a 15-minute reporting time step. Due to meter flow variability, either 30-min or 1 hour moving average flow values were applied for the analyses. Averages of the flow rates for the seven sewersheds ranged from 7.88×10^5 L/day in D3 to 3.16×10^7 L/day in MT (table 3.4).

Table 3.4. Descriptive statistics of flow rates (30-min or 1-hour moving average, L/day) for the seven neighborhood sewersheds (County Macomb: MT and SH; County Wayne: D1, D2 and D3; and County Oakland: SF and WB) in the Detroit Tri-County Area MI

	MT	SH	D1	D2	D3	SF	WB
Minimum	3.01×10^7	1.54×10^7	1.35×10^6	1.44×10^6	4.89×10^5	2.06×10^6	2.78×10^6
Maximum	3.58×10^7	2.97×10^7	1.89×10^7	4.22×10^7	3.92×10^6	7.72×10^6	6.43×10^6
Average	3.16×10^7	2.01×10^7	2.51×10^6	4.04×10^6	7.88×10^5	3.50×10^6	5.84×10^6
Std	1.47×10^6	3.16×10^6	4.10×10^6	9.52×10^6	7.88×10^5	1.20×10^6	7.90×10^5

The average SARS-CoV-2 loads using N1 gene assay ranged from 3.64×10^8 gc/day in D3 to 2.94×10^{10} gc/day in MT, while for N2 gene assay, they ranged from 3.53×10^8 gc/day in D3 to 2.60×10^{10} gc/day in MT (figure 3.2C and 3.2D). Further Tukey's post hoc analysis indicated SARS-CoV-2 loads in MT and SH were significantly larger than those in the remaining sites.

Cumulative COVID-19 cases in the nine sewersheds for the entire study period are presented in figure 3.2E. COVID-19 cases for MT and SH sewersheds (6,770 and 7,305 cases, respectively) were larger than those in the other seven sewersheds. Confirmed cumulative cases in the SF sewershed (635) were the lowest reported during the study period. Cumulative COVID-19 cases per capita was calculated (figure 3.2F) using the population data for the nine

sewersheds. On a per capita basis, COVID-19 incidences in MT and SH remain to be the highest; however, the difference between these two sewersheds and the remaining seven sewersheds decreased compared to the cumulative cases without normalizing to population.

Both concentrations and loads of SARS-CoV-2 RNA in MT and SH sewersheds were higher than the other sewersheds, which is consistent with their high cumulative clinically confirmed cases. Relationships between SARS-CoV-2 RNA loads (targeting at N1 and N2) and cumulative confirmed cases were explored using spearman correlation analysis (figure 3.3A and 3.3C). Also, linkages between SARS-CoV-2 RNA loads and populations that the sewersheds served were investigated through spearman correlation analysis (figure 3.3B and 3.3D). The results suggest SARS-CoV-2 RNA loads are correlated with both the cumulative clinical COVID-19 cases and population served by the sewersheds and agrees with previous published work (Bertels et al., 2022; Wilder et al., 2021; Wu et al., 2021). A linear relationship between SARS-CoV-2 RNA concentration and population size in the catchment area in New York, US was identified (Wilder et al., 2021). With the wastewater surveillance study of SARS-CoV-2 across 40 states in the U.S. from February to June 2020, a positive correlation between SARS-CoV-2 RNA detection rates and population sizes were found (Bertels et al., 2022; Wu et al., 2021). However, as mentioned above, difference of COVID-19 incidences between sewershed MT and SH, and the remaining seven sewersheds decreased when normalized to population, which signifies the importance of assessing COVID-19 burden to per capita level.

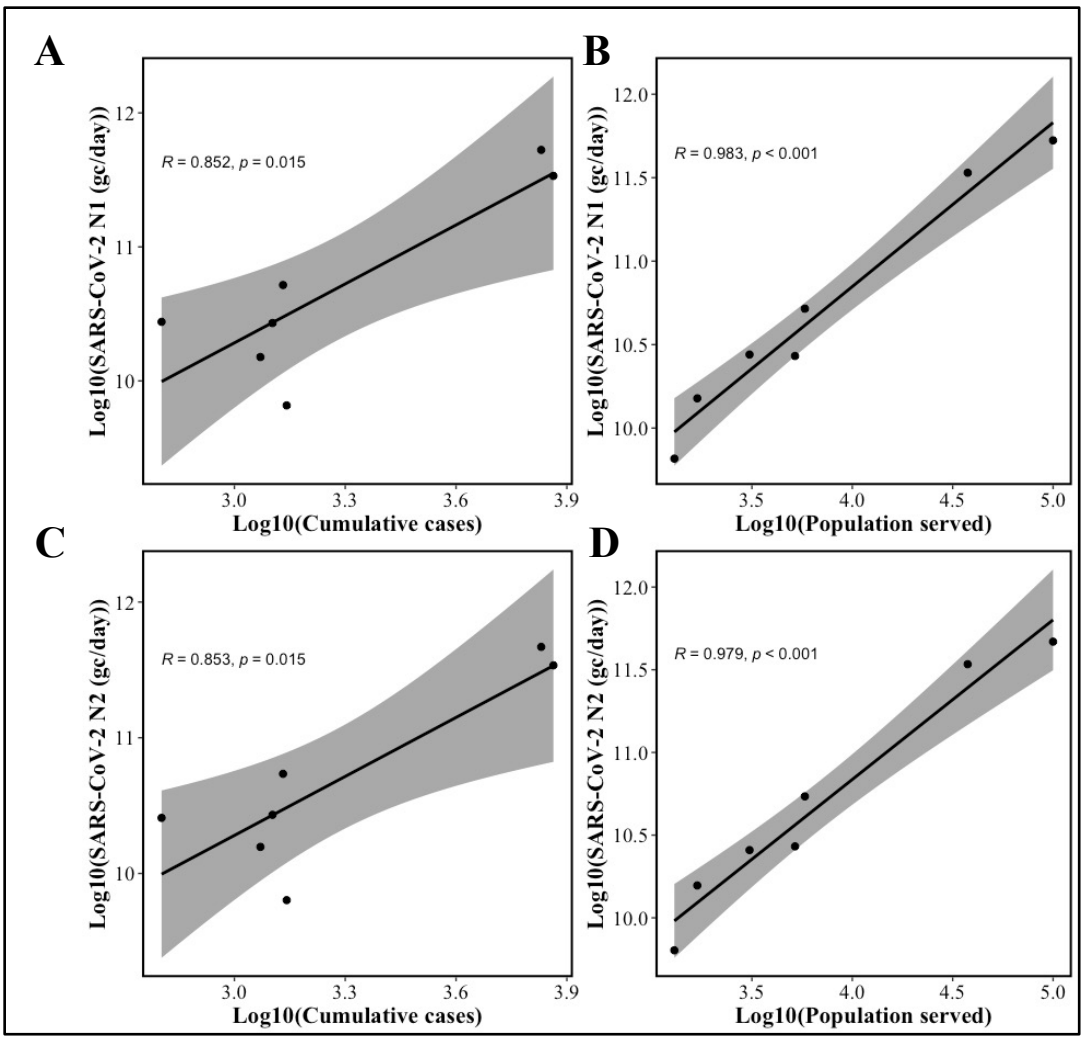

Figure 3.3. A: Spearman correlation analysis between SARS-CoV-2 gene N1 load and cumulative cases. **B:** Spearman correlation analysis between SARS-CoV-2 gene N1 load and population served by each sewershed. **C:** Spearman correlation analysis between SARS-CoV-2 gene N2 load and cumulative cases. **D:** Spearman correlation analysis between SARS-CoV-2 gene N2 load and population served by each sewershed. Correlations were performed for sewersheds with flow data available (Macomb: MT and SH, Wayne: D1, D2 and D3, and Oakland: SF and WB)

3.2. Assessment of the potential population markers in wastewater-based surveillance

In addition to the census approach, contributing population to a sewershed can be assessed using water quality parameters (BOD, etc.) and population biomarkers (creatinine, etc.). In this study, both water quality and population biomarkers were studied to evaluate potential

population markers in per-capita SARS-CoV-2 assessments and cross-site comparisons. Water quality constituent and population biomarker concentrations from the nine sites are shown in table 3.5.

Table 3.5. Concentrations of water quality markers and population biomarkers in the wastewater samples collected from nine sewersheds in the Detroit Tri-County area MI. Values are described as average ± standard deviation

Markers	Sites								
	EP	MT	SH	D1	D2	D3	SF	WB	OP
BOD (mg/L)	228±61.0	208±47.0	134±28.0	120±43.0	68.0±19.0	79.0±25.0	84.0±37.0	215±44.0	125±59.0
TKN (mg/L)	40.0±9.00	58.0±6.00	45.0±8.00	25.0±7.00	25.0±7.00	30.0±10.0	31.0±10.0	34.0±4.00	40.0±16.0
TSS (mg/L)	97.0±50.0	126±78.0	99.0±21.0	39.0±20.0	24.0±12.0	37.0±17.0	56.0±30.0	107±27.0	67.0±49.0
VSS (mg/L)	85.0±41.0	106±56.0	91.0±17.0	35.0±17.0	18.0±4.00	33.0±15.0	43.0±19.0	98.0±24.0	58.0±41.0
Creatinine (µg/L)	32.6±15.9	34.1±10.9	22.5±10.9	24.8±10.0	22.2±9.1	25.6±8.4	20.5±12.3	24.1±9.4	37.4±20.3
5-HIAA (µg/L)	8.10±4.80	11.1±5.80	4.90±5.10	5.10±3.90	5.30±3.90	6.50±4.20	6.60±3.90	6.60±3.70	9.50±6.30
Caffeine (µg/L)	137±60.4	182±77.0	113±50.9	55.4±24.8	46.7±18.4	33.6±11.6	106±64.1	156±78.3	91.3±36.4
Xanthine (µg/L)	29.5±21.2	33.4±20.6	23.9±16.0	20.8±12.9	17.7±11.5	21.3±14.4	22.6±14.7	26.3±20.7	35.8±23.6
Methylxanthine (µg/L)	64.6±56.2	103±74.2	63.0±50.2	32.4±24.8	30.5±24.5	36.3±27.0	49.0±40.4	49.7±42.0	67.6±50.6
Theophylline (µg/L)	35.2±11.3	60.0±17.3	40.9±12.8	16.5±5.30	23.0±9.30	19.3±7.20	30.8±9.40	34.4±13.3	35.2±11.1

Table 3.5. (cont'd)

Theobromine (μg/L)	37.4±17.8	67.2±25.4	41.5±18.6	21.2±7.20	24.3±10.0	22.9±9.00	31.6±13.6	40.6±19.6	38.0±17.6
Paraxanthine (μg/L)	28.1±8.10	46.6±14.1	31.8±10.2	13.5±4.10	18.6±7.00	15.8±5.40	25.0±7.60	27.0±10.2	28.0±8.70

The highest concentrations of BOD were found in EP (228 ± 61.0 mg/L) and WB (215 ± 44.0 mg/L), which have populations of approximately 2,400 and 5,800, respectively, based on the census estimate. The two sites with the highest concentrations of TKN were MT (58.0 ± 6.00 mg/L) and SH (45.0 ± 8.00 mg/L), which also have the largest catchment level populations (99,970 and 35,560, respectively). High levels of TSS and VSS were found in site MT (TSS: 126 ± 78.0 mg/L; VSS: 106 ± 56.0 mg/L) and WB (TSS: 107 ± 27.0 mg/L; VSS: 98.0 ± 24.0 mg/L), however, the population estimated for the MT sewershed (99,970) is much larger than the WB sewershed (5,800).

The highest concentration of creatinine was detected in sewer samples collected from the OP site (37.4 ± 20.3 μg/L), followed by the MT site (34.1 ± 10.9 μg/L). The estimated population for the MT sewershed is the largest (99,970), which is consistent with its high creatinine concentration. However, the population estimated for the catchment area of the OP site is relatively small (2,270). Similar results were found for 5-HIAA, where concentrations were found to be very high at sites MT and OP (MT:11.1 ± 5.80 μg/L; OP:9.50 ± 6.30 μg/L), even though the estimated populations of the two sites vary significantly. Concentrations of caffeine and its metabolites were identified in microgram per liter, and concentrations of the metabolites were found to be slightly lower than the parent compound (Chen et al., 2014; Choi et al., 2018; Gracia-Lor et al., 2017).

Correlations analysis between the loads (g/day) of population markers and the population sizes obtained from census approach were performed (figure 3.4). The results showed that BOD, TKN, creatinine, 5-HIAA and three of caffeine's metabolites are correlated strongly with the population sizes obtained from census ($r > 0.90$) (figure 3.4).

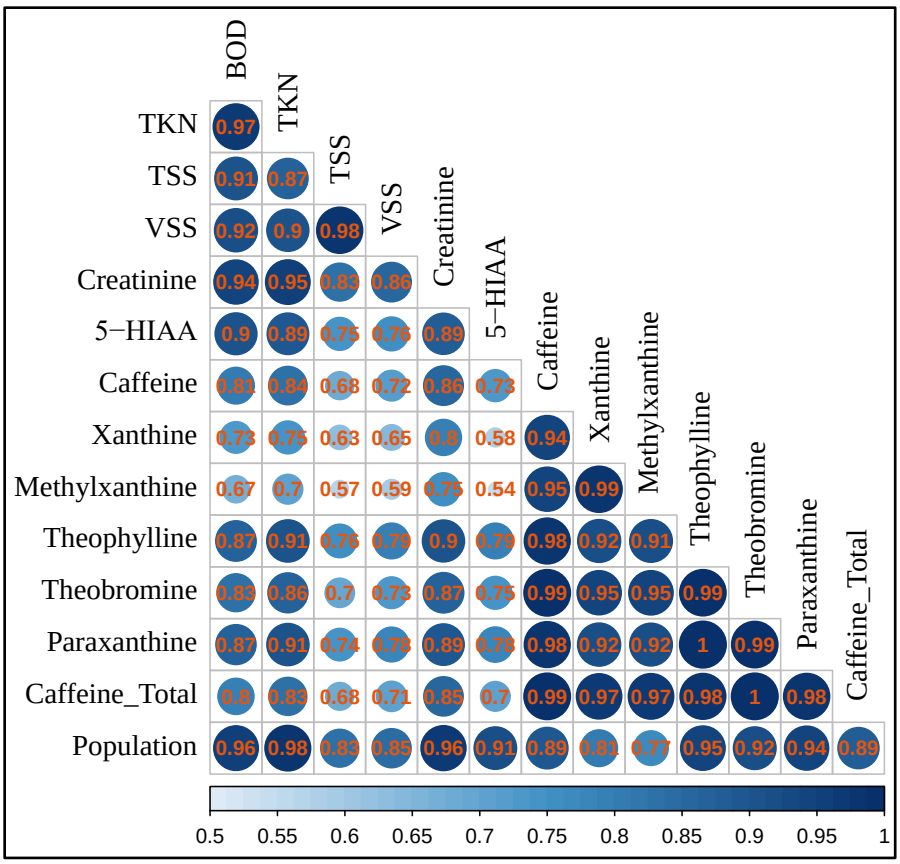

Figure 3.4. Correlation analysis between loads of potential population markers (including water quality parameters and human biomarkers) and population sizes obtained from census. Numbers in the shapes are the correlation coefficients. Unit of all the population markers is kept consistent and is g/day

Concentration of the wastewater RNA may increase with the larger population sizes, as in our study, concentrations of SARS-CoV-2 RNA in MT and SH wastewater are much higher than those in other sites (figure 3.2A and 3.2B). To understand the per capita viral contribution, it is critical to have an accurate estimate of population to be able to normalize measured SARS-CoV-2 concentrations in wastewater.

3.3. Normalized SARS-CoV-2 and the clinical cases rate

To understand the per capita viral contribution and perform spatial comparisons, SARS-CoV-2 RNA concentrations were normalized to population served in order to relate with COVID-19 incidences in the nine sampling locations. Descriptive characteristics of the normalized SARS-CoV-2 were summarized in table 3.6.

Table 3.6. Normalized SARS-CoV-2 in nine neighborhood sites (County Macomb: EP, MT and SH; County Wayne: D1, D2 and D3; and County Oakland: SF, WB and OP) in the Detroit Tri-County area MI. Values were described as average ± standard deviation

Normalizing factors	Sites								
	EP	MT	SH	D1	D2	D3	SF	WB	OP
SARS/BOD (gc/mg BOD)	N1:2.15±0.94 N2:2.17±1.00	N1:4.73±4.34 N2:4.16±3.17	N1:6.94±9.61 N2:7.04±10.70	N1:3.32±2.07 N2:3.48±2.04	N1:7.18±7.44 N2:7.12±6.28	N1:6.85±5.26 N2:6.43±3.82	N1:6.53±8.12 N2:5.79±3.73	N1:2.29±0.70 N2:2.39±0.98	N1:5.07±3.90 N2:4.78±2.99
SARS/TKN (gc/mg N)	N1:11.9±4.34 N2:12.0±4.76	N1:15.9±12.7 N2:14.0±9.30	N1:20.8±28.7 N2:21.1±32.1	N1:15.8±11.1 N2:16.6±11.1	N1:19.5±17.8 N2:19.6±16.2	N1:20.9±22.7 N2:20.0±21.3	N1:17.4±19.1 N2:16.7±15.1	N1:14.5±4.93 N2:15.2±6.64	N1:18.2±2.47 N2:17.5±2.14
SARS/TSS (gc/mg TSS)	N1:5.53±2.49 N2:5.53±2.53	N1:11.6±19.2 N2:10.3±13.0	N1:10.2±17.2 N2:10.1±18.3	N1:11.0±6.01 N2:11.8±6.66	N1:20.6±25.0 N2:20.9±21.5	N1:16.5±15.0 N2:15.2±10.5	N1:11.3±16.9 N2:9.57±7.56	N1:4.84±2.17 N2:5.08±2.94	N1:9.90±6.75 N2:9.40±5.16
SARS/VSS (gc/mg VSS)	N1:6.31±2.91 N2:6.31±2.98	N1:13.2±21.2 N2:11.8±14.2	N1:10.9±17.4 N2:10.8±18.6	N1:12.3±7.17 N2:13.2±7.89	N1:26.2±29.8 N2:26.1±25.3	N1:18.6±17.1 N2:17.2±12.3	N1:13.9±19.8 N2:12.1±9.42	N1:5.25±2.27 N2:5.53±3.18	N1:11.4±7.67 N2:10.8±6.08

Table 3.6. (cont'd)

SARS/Creatinine (gc/µg Creatinine)	N1:19.1±16.1 N2:19.4±18.4	N1:31.5±28.0 N2:27.5±21.8	N1:57.8±96.4 N2:58.0±103	N1:16.2±8.06 N2:16.9±8.20	N1:24.9±24.9 N2:24.7±22.2	N1:20.2±16.1 N2:18.9±9.75	N1:29.5±22.4 N2:29.3±20.7	N1:23.8±13.1 N2:25.2±16.4	N1:18.8±13.0 N2:18.6±13.2
SARS/5-HIAA (gc/µg 5-HIAA)	N1:437±876 N2:452±931	N1:926±2083 N2:803±1808	N1:2584±6172 N2:2661±6620	N1:458±854 N2:461±852	N1:515±996 N2:503±989	N1:449±953 N2:430±928	N1:529±1234 N2:513±1211	N1:424±838 N2:401±729	N1:422±878 N2:419±879
SARS/Caffeine (gc/µg Caffeine)	N1:5.15±3.62 N2:5.25±4.36	N1:8.10±8.23 N2:6.96±6.67	N1:10.1±15.9 N2:9.63±16.6	N1:8.23±3.68 N2:9.18±4.35	N1:11.3±10.2 N2:11.1±7.78	N1:22.0±21.9 N2:20.5±17.4	N1:5.21±2.18 N2:5.64±2.52	N1:4.27±2.64 N2:4.73±3.10	N1:8.30±7.27 N2:7.77±5.51
SARS/Xanthine (gc/µg Xanthine)	N1:27.0±18.7 N2:26.8±20.3	N1:46.0±42.9 N2:39.4±35.3	N1:52.8±88.3 N2:51.0±92.7	N1:22.3±8.24 N2:24.7±10.2	N1:29.7±20.4 N2:29.5±16.1	N1:36.5±30.7 N2:33.1±21.3	N1:24.2±9.11 N2:26.1±10.8	N1:29.3±19.2 N2:32.6±24.0	N1:23.0±18.9 N2:21.4±14.0

Table 3.6. (cont'd)

SARS/Methylxanthine (gc/µg Methylxanthine)	N1:15.2±12.1 N2:15.3±13.5	N1:17.1±16.5 N2:14.9±14.1	N1:24.9±46.8 N2:24.1±49.1	N1:16.3±7.48 N2:18.3±9.78	N1:20.9±17.9 N2:20.7±15.0	N1:23.4±19.9 N2:21.2±14.1	N1:13.2±6.98 N2:14.1±8.16	N1:17.2±12.6 N2:19.2±15.5	N1:13.8±12.5 N2:12.7±9.19
SARS/Theophylline (gc/µg Theophylline)	N1:18.4±11.9 N2:18.2±12.8	N1:20.8±18.1 N2:18.2±15.1	N1:25.9±40.3 N2:24.9±42.3	N1:25.5±9.56 N2:28.4±11.7	N1:22.0±17.2 N2:21.7±13.0	N1:34.6±26.8 N2:33.0±20.7	N1:15.6±4.84 N2:16.9±6.05	N1:19.0±12.5 N2:21.4±15.7	N1:19.7±15.0 N2:18.5±10.8
SARS/Theobromine (gc/µg Theobromine)	N1:19.5±14.0 N2:19.3±15.1	N1:20.6±20.1 N2:17.7±16.1	N1:28.2±46.2 N2:27.1±48.4	N1:19.9±7.88 N2:22.1±9.13	N1:20.8±17.4 N2:20.8±13.3	N1:30.0±25.1 N2:27.7±16.7	N1:15.9±5.32 N2:17.2±6.26	N1:16.7±10.7 N2:18.7±13.5	N1:20.4±17.5 N2:18.9±12.3
SARS/Paraxanthine (gc/µg Paraxanthine)	N1:22.6±14.3 N2:22.3±15.0	N1:27.5±25.2 N2:23.6±20.0	N1:32.5±48.8 N2:31.3±51.3	N1:30.9±11.6 N2:34.3±13.8	N1:26.0±18.2 N2:25.8±13.5	N1:41.9±32.4 N2:40.0±25.1	N1:19.2±5.92 N2:20.8±6.92	N1:24.0±16.0 N2:26.8±19.2	N1:24.7±19.3 N2:23.1±13.1

Clinical case per 1000-person per site were calculated using the cumulative clinical cases divided by the population size served by each sewershed. Clinical cases per 1000-person for the nine sites ranged from 29 cases to 65 cases per 1000-person (figure 3.2F). Even though SH and MT have a relatively higher confirmed case rate, the deviation of these two sites to the remaining seven sites has reduced significantly compared to cumulative clinical cases in the study periods (figure 3.2E).

Correlations between normalized SARS-CoV-2 and the clinical cases were used to evaluate the application of the population markers. Results indicated that SARS-CoV-2 RNA normalized on the basis of TKN, creatinine, 5-HIAA and xanthine were correlated strongly with

the clinical cases per 1000-person (figure 3.5). Furthermore, when TKN and xanthine were used to normalize the SARS-CoV-2 RNA concentrations, no significant differences were found among the nine sites. When creatinine and 5-HIAA were used, SH still stood out from the nine sampling locations (figure 3.6 and 3.7). Findings in this study may promote the per capita viral assessment and benefit the public health, especially in rural areas and the developing world, since factor like TKN can be measured easily onsite.

Concentrations of SARS-CoV-2 RNA in wastewater are affected by several factors including shedding-related factors, sewershed population, in-sewer factors (e.g., load and physiochemical properties of solid particles and organic matters, influx of rainwater/stormwater/groundwater), and sampling strategies (Bertels et al., 2022). Adjusting for these factors can help reduce uncertainties of the wastewater data. Among these factors, population normalization is crucial for accurate wastewater surveillance and confident viral assessment. By evaluating various water quality parameters and human biomarkers as the normalizing factors, SARS-CoV-2 RNA normalized by TKN, creatinine, 5-HIAA and xanthine correlated strongly with the clinical cases per 1000-person.

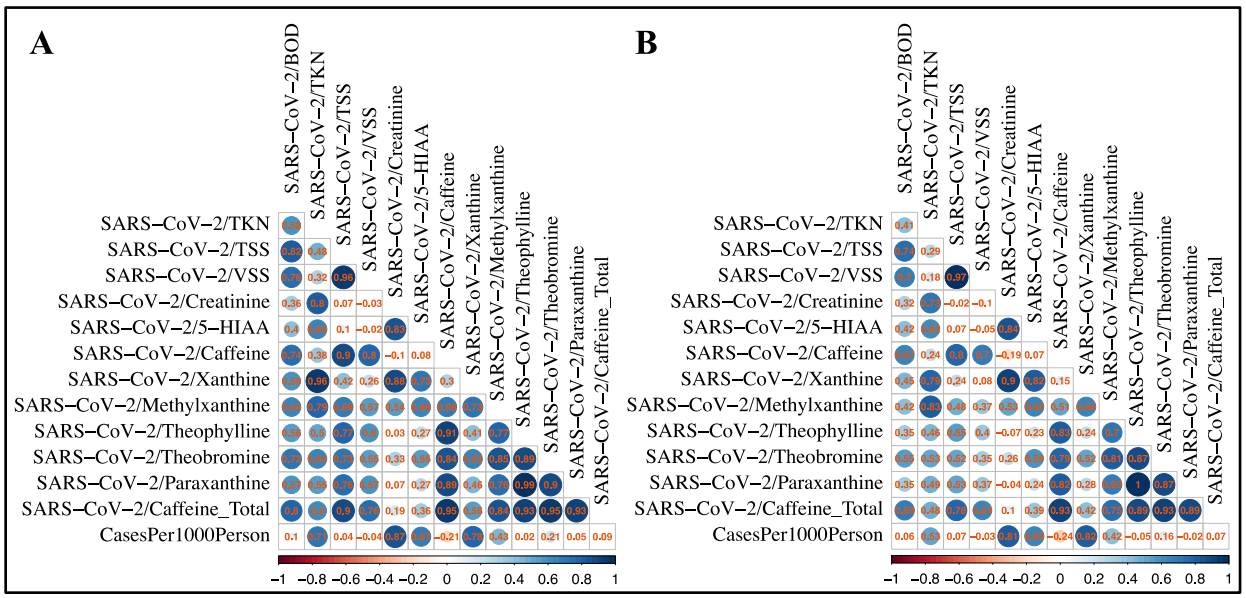

Figure 3.5. Normalized SARS-CoV-2 (A: N1; B: N2) by population markers to assess the COVID burdens to the per capita level. Correlation analyses were performed between normalized SARS-CoV-2 and COVID-19 cases per 1,000-person. Events with water quality markers, creatinine, 5-HIAA, and caffeine and its metabolites available were included to do the correlation. For each sewershed and normalizing factor, sum of the normalized SARS-CoV-2 for

Figure 3.5. (cont'd)

the 8 events were considered to do the correlation. COVID-19 cases rates were calculated for the periods corresponding to those 8 events. Unit of all the normalized SARS-CoV-2 concentrations is kept consistent and is gc/g population marker

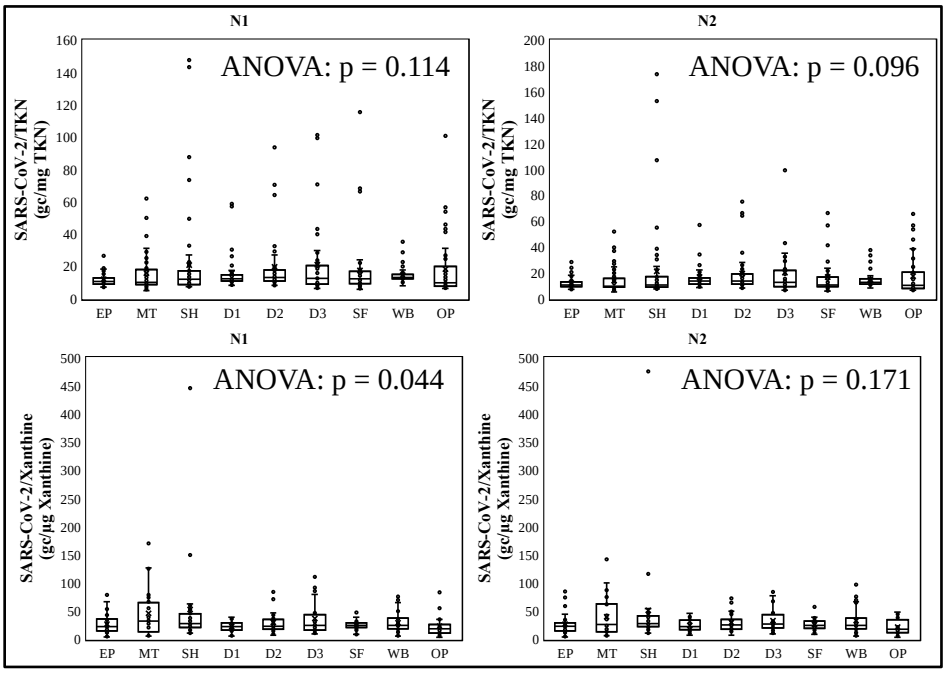

Figure 3.6. Boxplots of normalized SARS-CoV-2 RNA concentrations by TKN and Xanthine in nine neighborhood sites (County Macomb: EP, MT and SH; County Wayne: D1, D2 and D3; and County Oakland: SF, WB and OP) in the Detroit Tri-County Area MI

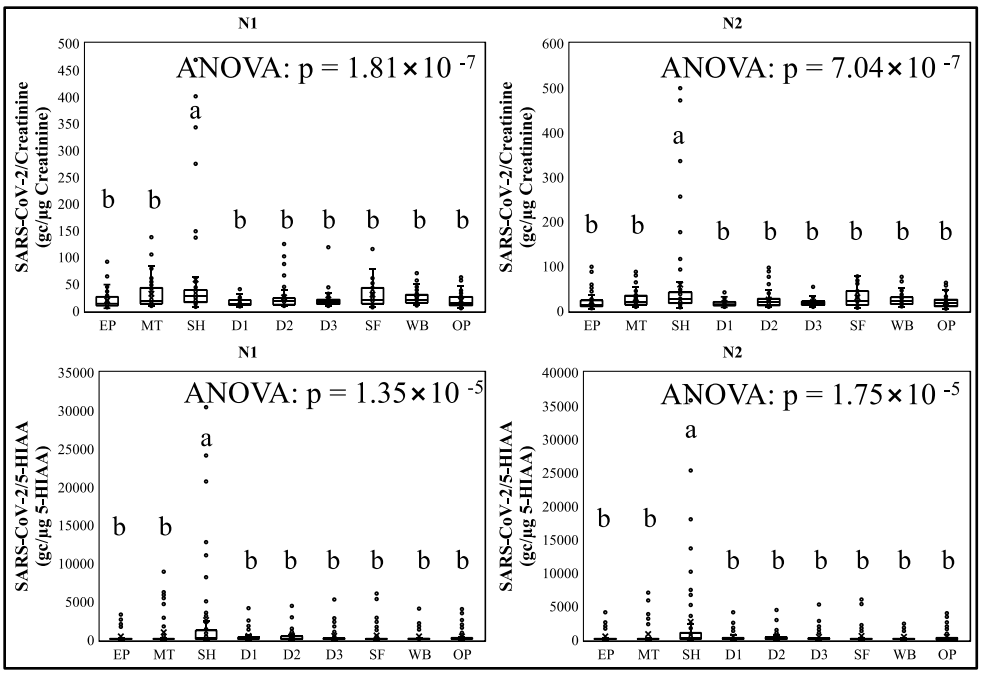

Figure 3.7. Boxplots of normalized SARS-CoV-2 RNA concentrations by TKN and Xanthine in nine neighborhood sites (County Macomb: EP, MT and SH; County Wayne: D1, D2 and D3; and County Oakland: SF, WB and OP) in Detroit MI

4. CONCLUSIONS

Average concentrations of N1 (SARS-CoV-2) in wastewater samples collected from nine sites ranged from 3.48×10^2 to 9.37×10^2 gc/L, and for N2, averages ranged from 3.72×10^2 to 8.79×10^2 gc/L. Both levels and loads of SARS-CoV-2 RNA in wastewater in two neighborhoods (MT and SH) were found to be higher than that in the other seven sites. The differences were statistically significant. Comparisons between normalized SARS-CoV-2 by population markers and COVID-19 incidences, indicated that normalization of SARS-CoV-2 RNA concentrations with TKN, creatinine, 5-HIAA and xanthine correlated positively with the COVID-19 incidences.

BIBLIOGRAPHY

Ahmed, W., Bibby, K., D'Aoust, P.M., Delatolla, R., Gerba, C.P., Haas, C.N., Hamilton, K.A., Hewitt, J., Julian, T.R., Kaya, D., Monis, P., Moulin, L., Naughton, C., Noble, R.T., Shrestha, A., Tiwari, A., Simpson, S.L., Wurtzer, S. and Bivins, A. 2021a. Differentiating between the possibility and probability of SARS-CoV-2 transmission associated with wastewater: empirical evidence is needed to substantiate risk. FEMS Microbes 2. https://doi.org/10.1093/femsmc/xtab007

Ahmed, W., Bivins, A., Metcalfe, S., Smith, W.J.M., Ziels, R., Korajkic, A., McMinn, B., Graber, T.E. and Simpson, S.L. 2022a. RT-qPCR and ATOPlex sequencing for the sensitive detection of SARS-CoV-2 RNA for wastewater surveillance. Water Research 220, 118621. https://doi.org/https://doi.org/10.1016/j.watres.2022.118621

Ahmed, W., Bivins, A., Simpson, S.L., Bertsch, P.M., Ehret, J., Hosegood, I., Metcalfe, S.S., Smith, W.J.M., Thomas, K.V., Tynan, J. and Mueller, J.F. 2022b. Wastewater surveillance demonstrates high predictive value for COVID-19 infection on board repatriation flights to Australia. Environment International 158, 106938. https://doi.org/https://doi.org/10.1016/j.envint.2021.106938

Ahmed, W., Bivins, A., Smith, W.J.M., Metcalfe, S., Stephens, M., Jennison, A.V., Moore, F.A.J., Bourke, J., Schlebusch, S., McMahon, J., Hewitson, G., Nguyen, S., Barcelon, J., Jackson, G., Mueller, J.F., Ehret, J., Hosegood, I., Tian, W., Wang, H., Yang, L., Bertsch, P.M., Tynan, J., Thomas, K.V., Bibby, K., Graber, T.E., Ziels, R. and Simpson, S.L. 2022c. Detection of the Omicron (B.1.1.529) variant of SARS-CoV-2 in aircraft wastewater. Science of The Total Environment 820, 153171. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.153171

Ahmed, W., Tscharke, B., Bertsch, P.M., Bibby, K., Bivins, A., Choi, P., Clarke, L., Dwyer, J., Edson, J., Nguyen, T.M.H., O'Brien, J.W., Simpson, S.L., Sherman, P., Thomas, K.V., Verhagen, R., Zaugg, J. and Mueller, J.F. 2021b. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: A temporal case study. Science of The Total Environment 761, 144216. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.144216

Been, F., Rossi, L., Ort, C., Rudaz, S., Delemont, O. and Esseiva, P. 2014. Population normalization with ammonium in wastewater-based epidemiology: application to illicit drug monitoring. Environ Sci Technol 48(14), 8162-8169. https://doi.org/10.1021/es5008388

Bertels, X., Demeyer, P., Van den Bogaert, S., Boogaerts, T., van Nuijs, A.L.N., Delputte, P. and Lahousse, L. 2022. Factors influencing SARS-CoV-2 RNA concentrations in wastewater up to the sampling stage: A systematic review. Sci Total Environ, 153290. https://doi.org/10.1016/j.scitotenv.2022.153290

Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Pogreba Brown, K.M., Stark, E.R., Foster, A.R., Sprissler, R.S., Harris, D.T., Sherchan, S.P., Gerba, C.P. and Pepper, I.L. 2021. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. Science of The Total Environment 779, 146408. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.146408

Bibby, K., Bivins, A., Wu, Z. and North, D. 2021. Making waves: Plausible lead time for wastewater based epidemiology as an early warning system for COVID-19. Water Research 202, 117438. https://doi.org/https://doi.org/10.1016/j.watres.2021.117438

Bivins, A. and Bibby, K. 2021. Wastewater Surveillance during Mass COVID-19 Vaccination on a College Campus. Environmental Science & Technology Letters 8(9), 792-798. https://doi.org/10.1021/acs.estlett.1c00519

CDC 2022, pp. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/data-reporting-analytics.html.

Chen, C., Kostakis, C., Gerber, J.P., Tscharke, B.J., Irvine, R.J. and White, J.M. 2014. Towards finding a population biomarker for wastewater epidemiology studies. Sci Total Environ 487, 621-628. https://doi.org/10.1016/j.scitotenv.2013.11.075

Chiaia, C.A., Banta-Green, C. and Field, J. 2008. Eliminating Solid Phase Extraction with Large-Volume Injection LC:MS:MS- Analysis of Illicit and Legal Drugs and Human Urine Indicators in US Wastewaters. Environ. Sci. Technol. 42(23), 7. https://doi.org/doi.org/10.1021/es802309v

Choi, P.M., Tscharke, B.J., Donner, E., O'Brien, J.W., Grant, S.C., Kaserzon, S.L., Mackie, R., O'Malley, E., Crosbie, N.D., Thomas, K.V. and Mueller, J.F. 2018. Wastewater-based epidemiology biomarkers: Past, present and future. TrAC Trends in Analytical Chemistry 105, 453-469. https://doi.org/10.1016/j.trac.2018.06.004

Daughton, C.G. 2012. Real-time estimation of small-area populations with human biomarkers in sewage. Sci Total Environ 414, 6-21. https://doi.org/10.1016/j.scitotenv.2011.11.015

Froehner, S., Souza, D.B., Machado, K.S. and Da Rosa, E.C. 2010. Tracking anthropogenic inputs in Barigui River, Brazil using biomarkers. Water, Air, & Soil Pollution 210(1), 8. https://doi.org/10.1007/s11270-009-0220-8

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

Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Roppolo Brazell, L., Hinton, K., Lontai, J., Stark, N., Young, I., Quach, C., Russ, M., Kauer, J., Nicolosi, B., Chen, D., Akella, S., Tang, W., Schlueter, J. and Munir, M. 2021. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. Science of The Total Environment 782, 146749. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.146749

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

Gracia-Lor, E., Castiglioni, S., Bade, R., Been, F., Castrignano, E., Covaci, A., Gonzalez-Marino, I., Hapeshi, E., Kasprzyk-Hordern, B., Kinyua, J., Lai, F.Y., Letzel, T., Lopardo, L., Meyer, M.R., O'Brien, J., Ramin, P., Rousis, N.I., Rydevik, A., Ryu, Y., Santos, M.M., Senta, I., Thomaidis, N.S., Veloutsou, S., Yang, Z., Zuccato, E. and Bijlsma, L.

2017. Measuring biomarkers in wastewater as a new source of epidemiological information: Current state and future perspectives. Environ Int 99, 131-150. https://doi.org/10.1016/j.envint.2016.12.016

Gudra, D., Dejus, S., Bartkevics, V., Roga, A., Kalnina, I., Strods, M., Rayan, A., Kokina, K., Zajakina, A., Dumpis, U., Ikkere, L.E., Arhipova, I., Berzins, G., Erglis, A., Binde, J., Ansonska, E., Berzins, A., Juhna, T. and Fridmanis, D. 2022. Detection of SARS-CoV-2 RNA in wastewater and importance of population size assessment in smaller cities: An exploratory case study from two municipalities in Latvia. Science of The Total Environment 823, 153775. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.153775

Hovi, T., Shulman, L.M., van der Avoort, H., Deshpande, J., Roivainen, M. and EM, D.E.G. 2012. Role of environmental poliovirus surveillance in global polio eradication and beyond. Epidemiol Infect 140(1), 1-13. https://doi.org/10.1017/S095026881000316X

Karmakar, M., Lantz, P.M. and Tipirneni, R. 2021. Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. JAMA Netw Open 4(1), e2036462. https://doi.org/10.1001/jamanetworkopen.2020.36462

Kim, S. and Bostwick, W. 2020. Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. Health Educ Behav 47(4), 4. https://doi.org/https://doi.org/10.1177/1090198120929677

Koelle, K., Martin, M.A., Antia, R., Lopman, B. and Dean, N.E. 2022. The changing epidemiology of SARS-CoV-2. Science 375(6585), 1116-1121. https://doi.org/doi:10.1126/science.abm4915

Laborde, D., Martin, W., Swinnen, J. and Vos, R. 2020. COVID-19 risks to global food security. Science 369(6503), 500-502. https://doi.org/10.1126/science.abc4765

Li, D., Gaynor, S.M., Quick, C., Chen, J.T., Stephenson, B.J.K., Coull, B.A. and Lin, X. 2021a. Identifying US County-level characteristics associated with high COVID-19 burden. BMC Public Health 21(1), 1007. https://doi.org/10.1186/s12889-021-11060-9

Li, J., Ahmed, W., Metcalfe, S., Smith, W.J.M., Tscharke, B., Lynch, P., Sherman, P., Vo, P.H.N., Kaserzon, S.L., Simpson, S.L., McCarthy, D.T., Thomas, K.V., Mueller, J.F. and Thai, P. 2022. Monitoring of SARS-CoV-2 in sewersheds with low COVID-19 cases using a passive sampling technique. Water Research 218, 118481. https://doi.org/https://doi.org/10.1016/j.watres.2022.118481

Li, X., Kulandaivelu, J., Zhang, S., Shi, J., Sivakumar, M., Mueller, J., Luby, S., Ahmed, W., Coin, L. and Jiang, G. 2021b. Data-driven estimation of COVID-19 community prevalence through wastewater-based epidemiology. Science of The Total Environment 789, 147947. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.147947

Lodder, W.J., Buisman, A.M., Rutjes, S.A., Heijne, J.C., Teunis, P.F. and de Roda Husman, A.M. 2012. Feasibility of quantitative environmental surveillance in poliovirus eradication strategies. Appl Environ Microbiol 78(11), 3800-3805. https://doi.org/10.1128/AEM.07972-11

Lu, R.J., Zhao, X., Li, J., Niu, P.H., Yang, B., Wu, H.L., Wang, W.L., Song, H., Huang, B.Y., Zhu, N., Bi, Y.H., Ma, X.J., Zhan, F.X., Wang, L., Hu, T., Zhou, H., Hu, Z.H., Zhou,

W.M., Zhao, L., Chen, J., Meng, Y., Wang, J., Lin, Y., Yuan, J.Y., Xie, Z.H., Ma, J.M., Liu, W.J., Wang, D.Y., Xu, W.B., Holmes, E.C., Gao, G.F., Wu, G.Z., Chen, W.J., Shi, W.F. and Tan, W.J. 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395(10224), 565-574. https://doi.org/10.1016/s0140-6736(20)30251-8

McCall, C., Wu, H., O'Brien, E. and Xagoraraki, I. 2021. Assessment of enteric viruses during a hepatitis outbreak in Detroit MI using wastewater surveillance and metagenomic analysis. J Appl Microbiol 131(3), 1539-1554. https://doi.org/10.1111/jam.15027

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. Environmental Science & Technology Letters 7(7), 511-516. https://doi.org/10.1021/acs.estlett.0c00357

Miyani, B., Fonoll, X., Norton, J., Mehrotra, A. and Xagoraraki, I. 2020. SARS-CoV-2 in Detroit Wastewater. Journal of Environmental Engineering 146(11). https://doi.org/10.1061/(asce)ee.1943-7870.0001830

Miyani, B., Zhao, L., Spooner, M., Buch, S., Gentry, Z., Mehrotra, A., Norton, J. and Xagoraraki, I. 2021. Early Warnings of COVID-19 Second Wave in Detroit. Journal of Environmental Engineering 147(8). https://doi.org/10.1061/(asce)ee.1943-7870.0001907

Reynolds, L.J., Sala-Comorera, L., Khan, M.F., Martin, N.A., Whitty, M., Stephens, J.H., Nolan, T.M., Joyce, E., Fletcher, N.F., Murphy, C.D. and Meijer, W.G. 2022. Coprostanol as a Population Biomarker for SARS-CoV-2 Wastewater Surveillance Studies. Water 14(2), 225.

Richardson, H. 2021. How waste water is helping South Africa fight COVID-19. Nature 593(7860), 616-617. https://doi.org/10.1038/d41586-021-01399-9

Rico, M., Andres-Costa, M.J. and Pico, Y. 2017. Estimating population size in wastewater-based epidemiology. Valencia metropolitan area as a case study. J Hazard Mater 323(Pt A), 156-165. https://doi.org/10.1016/j.jhazmat.2016.05.079

Scott, L.C., Aubee, A., Babahaji, L., Vigil, K., Tims, S. and Aw, T.G. 2021. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. Environ Res 200, 111374. https://doi.org/10.1016/j.envres.2021.111374

Senta, I., Gracia-Lor, E., Borsotti, A., Zuccato, E. and Castiglioni, S. 2015. Wastewater analysis to monitor use of caffeine and nicotine and evaluation of their metabolites as biomarkers for population size assessment. Water Res 74, 23-33. https://doi.org/10.1016/j.watres.2015.02.002

Sills, J., Prather, K.A., Marr, L.C., Schooley, R.T., McDiarmid, M.A., Wilson, M.E. and Milton, D.K. 2020. Airborne transmission of SARS-CoV-2. Science 370(6514), 303-304. https://doi.org/doi:10.1126/science.abf0521

Sims, N. and Kasprzyk-Hordern, B. 2020. Future perspectives of wastewater-based epidemiology: Monitoring infectious disease spread and resistance to the community level. Environ Int 139, 105689. https://doi.org/10.1016/j.envint.2020.105689

Subbaraman, N. 2021. How do vaccinated people spread Delta? What the science says. Nature 596(7872), 327-328. https://doi.org/10.1038/d41586-021-02187-1

Thai, P.K., O'Brien, J., Jiang, G., Gernjak, W., Yuan, Z., Eaglesham, G. and Mueller, J.F. 2014. Degradability of creatinine under sewer conditions affects its potential to be used as biomarker in sewage epidemiology. Water Res 55, 272-279. https://doi.org/10.1016/j.watres.2014.02.035

Tscharke, B.J., O'Brien, J.W., Ort, C., Grant, S., Gerber, C., Bade, R., Thai, P.K., Thomas, K.V. and Mueller, J.F. 2019. Harnessing the Power of the Census: Characterizing Wastewater Treatment Plant Catchment Populations for Wastewater-Based Epidemiology. Environ Sci Technol 53(17), 10303-10311. https://doi.org/10.1021/acs.est.9b03447

Webb Hooper, M., Napoles, A.M. and Perez-Stable, E.J. 2020. COVID-19 and Racial/Ethnic Disparities. JAMA 323(24), 2466-2467. https://doi.org/10.1001/jama.2020.8598

Wilder, M.L., Middleton, F., Larsen, D.A., Du, Q., Fenty, A., Zeng, T., Insaf, T., Kilaru, P., Collins, M., Kmush, B. and Green, H.C. 2021. Co-quantification of crAssphage increases confidence in wastewater-based epidemiology for SARS-CoV-2 in low prevalence areas. Water Res X 11, 100100. https://doi.org/10.1016/j.wroa.2021.100100

Wu, F., Xiao, A., Zhang, J., Moniz, K., Endo, N., Armas, F., 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., McElroy, K.A., Rhode, S.F., Matus, M., Wuertz, S., Thompson, J. and Alm, E.J. 2021. Wastewater surveillance of SARS-CoV-2 across 40 U.S. states from February to June 2020. Water Res 202, 117400. https://doi.org/10.1016/j.watres.2021.117400

Xagoraraki, I. and O'Brien, E. 2020. Wastewater-Based Epidemiology for Early Detection of Viral Outbreaks. Women in Water Quality. https://doi.org/10.1007/978-3-030-17819-2_5

Xagoraraki, I., Yin, Z. and Svambayev, Z. 2014. Fate of Viruses in Water Systems. Journal of Environmental Engineering 140(7). https://doi.org/10.1061/(asce)ee.1943-7870.0000827

Xie, Y., Challis, J.K., Oloye, F.F., Asadi, M., Cantin, J., Brinkmann, M., McPhedran, K.N., Hogan, N., Sadowski, M., Jones, P.D., Landgraff, C., Mangat, C., Servos, M.R. and Giesy, J.P. 2022. RNA in Municipal Wastewater Reveals Magnitudes of COVID-19 Outbreaks across Four Waves Driven by SARS-CoV-2 Variants of Concern. ACS ES&T Water. https://doi.org/10.1021/acsestwater.1c00349

Zhao, L., Zou, Y., Li, Y., Miyani, B., Spooner, M., Gentry, Z., Jacobi, S., David, R.E., Withington, S., McFarlane, S., Faust, R., Sheets, J., Kaye, A., Broz, J., Gosine, A., Mobley, P., Busch, A.W.U., Norton, J. and Xagoraraki, I. 2022. Five-week warning of COVID-19 peaks prior to the Omicron surge in Detroit, Michigan using wastewater surveillance. Science of The Total Environment 844, 157040. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.157040

Zheng, Q.D., Wang, Z., Liu, C.Y., Yan, J.H., Pei, W., Wang, Z. and Wang, D.G. 2019. Applying a population model based on hydrochemical parameters in wastewater-based epidemiology. Sci Total Environ 657, 466-475. https://doi.org/10.1016/j.scitotenv.2018.11.426

CHAPTER 4: BIOINFORMATICS-BASED SCREENING TOOL IDENTIFIES A WIDE VARIETY OF HUMAN VIRUSES IN TRUJILLO-PERU WASTEWATER AND CONFIRMS COVID-19, MONKEYPOX, AND DIARRHEAL DISEASE OUTBREAKS

This chapter was published in One Health Volume 18, June 2024, 100756, Brijen Miyani, Yabing Li, Heidy Peidro Guzman, Ruben Kenny Briceno, Sabrina Vieyra, Rene Hinojosa, Irene Xagoraraki, Bioinformatics-based screening tool identifies a wide variety of human and zoonotic viruses in Trujillo-Peru wastewater

ABSTRACT

Peru was one of the most affected countries during the COVID-19 pandemic. Moreover, multiple other viral diseases (enteric, respiratory, bloodborne, and vector-borne) are endemic and rising. According to Peru's Ministry of Health, various health facilities in the country were reallocated for the COVID-19 pandemic, thereby leading to reduced action to curb other diseases. Many viral diseases in the area are under-reported and not recognized. The One Health approach, in addition to clinical testing, incorporates environmental surveillance for detection of infectious disease outbreaks. The purpose of this work is to use a screening tool that is based on molecular methods, high throughput sequencing and bioinformatics analysis of wastewater samples to identify virus-related diseases circulating in Trujillo-Peru.

To demonstrate the effectiveness of the tool, we collected nine untreated wastewater samples from the Covicorti wastewater utility in Trujillo-Peru on October 22, 2022. High throughput metagenomic sequencing followed by bioinformatic analysis was used to assess the viral diversity of the samples. Our results revealed the presence of sequences associated with multiple human and zoonotic viruses including Orthopoxvirus, Hepatovirus, Rhadinovirus, Parechovirus, Mamastrovirus, Enterovirus, Varicellovirus, Norovirus, Kobuvirus, Bocaparvovirus, Simplexvirus, Spumavirus, Orthohepevirus, Cardiovirus, Molliscipoxvirus, Salivirus, Parapoxvirus, Gammaretrovirus, Alphavirus, Lymphocryptovirus, Erythroparvovirus, Sapovirus, Cosavirus, Deltaretrovirus, Roseolovirus, Flavivirus, Betacoronavirus, Rubivirus, Lentivirus, Betapolyomavirus, Rotavirus, Hepacivirus, Alphacoronavirus, Mastadenovirus, Cytomegalovirus and Alphapapillomavirus. For confirmation purposes, we tested the samples for the presence of selective viruses belonging to the genera detected above. PCR based molecular methods confirmed the presence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), monkeypox virus (MPXV), noroviruses GI and GII (NoVGI and NoVGII), and

rotavirus A (RoA) in our samples. Furthermore, publicly available clinical data for selected viruses confirm our findings.

Wastewater or other environmental media surveillance, combined with bioinformatics methods, has the potential to serve as a systematic screening tool for the identification of human or zoonotic viruses that may cause disease. The results of this method can guide further clinical surveillance efforts and allocation of resources. Incorporation of this bioinformatic-based screening tool by public health officials in Peru and other Latin American countries will help manage endemic and emerging diseases that could save human lives and resources.

1. INTRODUCTION

Peru is the third largest South American country comprising of a diverse range of landscapes and a population of 31,331,228 (Peru National geography kids). Half of Peru is covered by the world's largest rainforest, the Amazon. Considering the country's unique landscapes and ecosystems, diverse range of wildlife, and large population, the country could be an epicenter for a wide variety of human and zoonotic diseases. The United States Agency for International Development (USAID) Emerging Pandemic Threats PREDICT program has included Peru, Bolivia, Brazil, and Colombia in training personnel for surveillance of outbreaks of infectious diseases associated with wildlife pathogens in the community (Uhart et al., 2013).

A quick historical review of emerging infectious diseases of human pathogens points out that 60% or more have originated from animals (Jones et al., 2008). Furthermore, 75% of human pathogens studied are known to be viruses that are associated with wildlife reservoirs (Woolhouse et al., 2007). SARS-CoV-2, apart from ebola hemorrhagic fever, nipah viral encephalitis, hantavirus pulmonary syndrome, H5N1 highly pathogenic avian influenza, and the pandemic 2009 H1N1 influenza virus are recent examples of emerging infectious diseases of wildlife origin (Karesh et al., 2005; Flanagan et al., 2012). The One Health approach (also known as 'One world, One health'), that focuses on the innate relation between human, animal, and environmental health, can provide solutions to effective monitoring and early detection of infectious viral diseases by incorporating environmental surveillance in addition to clinical testing (O'Brien et al., 2019).

Traditional disease detection systems are based on diagnostic analysis of clinical samples. This approach, however, assumes that patients are examined at a clinical setting after symptoms have developed. Testing every individual for active infections or immunity, especially in

underserved communities, is impossible. Environmental based surveillance, such as wastewater testing, provides a means of collecting and analyzing community-composite-samples (O'Brien et al., 2019). Wastewater testing may provide early warnings of potential upcoming viral outbreaks as well as predictions of fluctuations of established outbreaks for specific geographical areas. Wastewater-based epidemiology has the potential to identify multiple endemic respiratory, enteric, bloodborne and vector-borne diseases circulating in the community (McCall et al., 2020; Miyani et al., 2021; Li et al., 2024a; Li et al., 2024b) as well as predicting viral outbreaks such as Covid-19 (Xagoraraki et al., 2020a; Xagoraraki et al., 2020b; Li et al., 2022; Zhao et al., 2022; Miyani et al., 2021). Early detection of emerging viral diseases in the community could help boost the preparedness of public health officials and save lives.

The immense burden of human and zoonotic viral infections in Peru is widely recognized. Respiratory diseases, from both viral and bacterial infections, were the leading cause of death across all age group from the year 2019 to 2020 (before the Covid-19 outbreak) (Peru WHO Mortality Database). Furthermore, Peru was one of the most affected countries during the Covid-19 pandemic. In Peru, from January 2020 to December 2023, as reported by the World Health Organization (WHO), there were 4,520,102 confirmed cases of Covid-19 with 221,564 deaths (Peru WHO Coronavirus Dashboard). During the Covid-19 pandemic, Peru experienced the worst infection rate (99%) (IHME Covid-19) and death rate (6400 deaths per million population) (Peru Covid-19 statistics; Peru NPR WAMU 88.5; Johns Hopkins CORONAVIRUS RESEARCH CENTER). Moreover, multiple other viral diseases (enteric, respiratory, bloodborne, and vector-borne) are endemic and rising. Arboviruses, such as Dengue virus, have been on a rise in South American countries including Peru (Hasan et al., 2022). Other arboviruses like Chikungunya and Zika have also been reported in the region (Centro Ministry Salud 2021). Many viral diseases in the area are under-reported and not recognized. According to Peru's Ministry of Health, various health facilities in the country were reallocated for the COVID-19 pandemic, thereby leading to reduced action to curb other diseases.

The purpose of this study is to analyze wastewater samples collected at Trujillo-Peru, with a screening tool that involves high throughput metagenomic sequencing and bioinformatics analysis, to identify potential virus-related diseases circulating in the community. This method may supplement and guide clinical surveillance efforts. Incorporation of this bioinformatic-based

screening tool by public health officials in Peru and other Latin American countries will help manage endemic and emerging diseases that could save human lives and resources.

2. MATERIAL AND METHODS

2.1 Study area and wastewater sampling

Trujillo, located in northwestern part of Peru, is the capital city of La Libertad province. According to 2017 census (Instituto Republica del Peru 2017), about 970 thousand people live in Trujillo making it the third largest city in the country. SEDALIB S.A (https://sedalib.com.pe/), a waste utility company, has divided the city's wastewater into five drainage collectors: Covicorti, Cortijo, Valdiva, Industrial Park, and the basin of La whitewashed. More than 90% of the city's wastewater is treated by these plants (Avila et al., 2016). Covicorti is the largest wastewater treatment utility in Trujillo and collects wastewater from 3 districts of Trujillo-Peru: Trujillo, Victor Larco Herrera, and El Porvenir. Annually, it treats an average flow of 59,166 m3 per day (685 L/sec) and an organic load of 24,350 kg of chemical oxygen demand per day (I.D.E.C 2006). The utility discharges treated wastewater to the ocean along Playa de Buenos Aires. Wastewater samples were collected in triplicates from the Covicorti wastewater utility on October 22, 2022.

Covicorti wastewater utility connects the grit removal tank and aeration tank by a long open 600m aqueduct. Untreated influent grab wastewater samples were collected in triplicates at 100m (A), 300m (B), and 500m (C) from the grit removal tank, for a total of 9 samples. For each sample, 200 mL of wastewater was collected in labelled corning polypropylene bottle. Samples were labeled as COVIA1, COVIA2, COVIA3, COVIB1, COVIB2, COVIB3, COVIC1, COVIC2, COVIC3. Samples were also collected at the Playa de Buenos Aires beach waters resulting in no signals for SARS-CoV-2 and hence no further analysis was conducted. Samples were shipped overnight on dry ice to the Environmental Virology Laboratory at Michigan State University (MSU) and stored in freezer (-80C) once delivered. Virus concentration were conducted within 72 hours of sampling

2.2 Virus concentration using Polyethylene Glycol (PEG) precipitation and RNA extraction

Polyethylene Glycol (PEG) precipitation was performed on grab samples using a previously described method (Li et al., 2024a; Flood et al., 2021). This involved adding 1.17g of sodium chloride and 8g of polyethene glycol to 100 mL sample and mixing in 4C at 110rpm for

2 hours (Flood et al., 2021). Samples were centrifuged for 45 mins at 4700xg at 4C . The supernatant was discarded without disturbing the pellets. Virus pellets were resuspended with 1-4 mL of liquid sample. Final volumes were noted, and the concentrated samples were stored in the freezer (-80C) for subsequent RNA extraction. Type I water (Barnstead Nano pure water system) was used as a negative control (NTC) in PEG precipitation process.

Viral nucleic acids were extracted according to manufacturer's protocol using QIAGEN QIAamp Viral RNA Mini Kit (QIAGEN cat# 52904, Germantown, Maryland) to obtain sufficient viral genetic materials. PCR-grade water was included in the RNA extraction processes as a negative control (NTC).

2.3 High throughput metagenomic sequencing and downstream bioinformatics analysis

High throughput metagenomic sequencing and downstream bioinformatics was applied as described previously (Li et al., 2024a; Li et al., 2024b) to assess viral diversity in the samples. RNA extract from the triplicates of each site were combined to result in total of 3 samples (COVIA, COVIB, COVIC). The RNA samples were reverse transcribed to cDNA using a random-primer protocol developed to identify viral pathogens (Wang et al., 2002, Wang et al., 2003).

Viral cDNA for the wastewater samples (n = 3) were sent to the MSU Research Technology Support Facility's Genomics Core for library preparation and sequencing. Quality of the raw reads were assessed using FastQC (Andrews et al., 2010). Every sample, for about 89% of both R1 and R2 reads, had quality score of greater than 30. An average of about 118 million reads and 35.4 Gb yield were obtained for each sample. Trimming, assembly and three different types of annotations were conducted using a previously described method (Li et al., 2024a; Li et al., 2024b). Briefly, adapters and low-quality reads were trimmed with Trimmomatic (Bolger et al., 2014). The proportion of viral reads in the samples were determined by aligning the trimmed reads against the National Center for Biotechnology Information (NCBI) BLAST non-redundant database using Kaiju (v. 1.9.0) (Menzel et al., 2016). Contigs were assembled using Megahit (Bibby et al., 2011) and aligned with NCBI RefSeq virus database (retrieved on December 1, 2022) with DIAMOND Blastx (Bağcı et al., 2021) to determine the virus composition. Furthermore, the assembled contigs were aligned again to custom Swiss-Prot human virus database using BLASTx to identify viral diversity at genus level (McCall et al., 2020). MEGAN

software was used for taxonomic annotation of viruses at genus level (Huson et al., 2016).

2.4 Municipal wastewater characterization by quantification of pepper mild mottle virus (PMMoV) and crAssphage virus using digital droplet qPCR

PMMoV virus was quantified using a GT molecular ddPCR kit for the Bio-Rad QX200™ Droplet Digital™ PCR System (cat# 100320) following the manufacturer's protocol. The kit contains single plex assay mastermix and positive control of PMMoV. Briefly 5.5 µL One-Step RT-ddPCR Supermix (20x), 2.2 µL of Reverse Transcriptase, 1.1 µL of 300mM DTT, 1 µL of PMMoV mastermix, and 9 µL RNA were mized. The total volume of reaction was made 22 µL by adding PCR-grade water. PMMoV was amplified using 50 °C for an hour, 95 °C for 10 mins, and 45 cycles (94 °C for 30 sec and 55 °C for 1min at slow ramp speed of 2 °C/second). The samples were subjected to a last extension cycle for 98 °C for 10mins followed by 4 °C for 30 min for droplet stabilization.

crAssphage virus was quantified using a GT molecular ddPCR assay kit (cat# 100285) using the manufacturer's protocol. Briefly, the mastermix contained 11 µL of 2X Supermix for Probes, 1 µL of crAssphage mastermix, and 10 µL RNA. crAssphage virus was amplified using 95 °C for 10 mins, and 45 cycles of (94 °C for 30 sec and 58 °C for 1min at slow ramp speed of 2 °C/second). The samples were subjected to a last extension cycle for 98 °C for 10 mins followed by 4 °C for 30 min for droplet stabilization. Each reaction consisted of samples run in triplicates, and NTCs from elution, extraction and ddPCR step.

2.5 Quantification of SARS-CoV-2 and monkeypox virus using digital droplet qPCR

SARS-CoV-2 was quantified by a duplex (N1 and N2 gene) digital droplet PCR (ddPCR) technology using Bio-Rad's One-Step RT-ddPCR Advanced kit with a QX200 ddPCR system (Bio-Rad, CA, USA) using a previously described method (Li et al., 2022; Zhao et al., 2022). The assay contained 5.5 µL One-Step RT-ddPCR Supermix (20x), 2.2 µL of Reverse Transcriptase, 1.1 µL of 300mM DTT, and 5.5 µL RNA. The final concentration of N1 and N2 gene forward and reverse primer was 900 nm, whereas the N1 and N2 probe was at 250 nm. The total volume of reaction was made 22 µL by adding PCR-grade water. Each reaction consisted of samples run in triplicates, PEG precipitation NTC, extraction NTC and ddPCR controls (SARS-CoV-2 from Twist Bioscience Twist Synthetic SARS-CoV-2 RNA Control 2 MN908947.3 as PTC and PCR-grade water as NTC). SARS-CoV-2 was amplified using 25 °C for 3 mins, 50 °C for an hour, 95 °C for 10 mins, and 40 cycles of (95 °C for 30 sec and 55 °C for 1min at slow

ramp speed of 2 °C/second). The samples were subjected to a last extension cycle for 98 °C for 10mins. Limit of blank (LOB) and Limit of detection (LOD) were calculated according to manufacturer's guidelines (Bio-Rad) as shown previously (Li et al., 2022; Zhao et al., 2022). LOB for SARS-CoV-2 N1 gene and N2 gene ddPCR was determined to be 0.09 and 0.08 copies/µL respectively (Li et al., 2022; Zhao et al., 2022). Limit of detection (LOD) of 0.1 copies/µL with 72.92 % confidence for the N1 gene ddPCR and 0.1 copies/µL with 81.25 % confidence for the N2 gene ddPCR were determined as shown previously (Li et al., 2022; Zhao et al., 2022).

Monkeypox virus (MPXV) was quantified using digital droplet PCR (ddPCR) technology using Bio-Rad's ddPCR™ Supermix for Probes following the manufacture's protocol. The primers and probe for amplifying are recommended by CDC (CDC PRB) and were used from a previously published study (Li et al., 2010). The assay contained 10 µL ddPCR Supermix for probes (20x, No dUTP), 900nm forward and reverse primer, 250nm probe, and 8.7 µL sample. The total volume of reaction was made 22 µL by adding PCR-grade water. Monkeypox viral DNA was purchased from ATCC (ATCC Number: VR-3270SD). The stock concentration was diluted to 10^3 copies/µL and used as a PTC. PCR-grade water was used as an NTC. Each reaction consisted of samples and controls run in triplicates. MPXV was amplified using 95 °C for 10 mins, and 45 cycles of (94 °C for 30 seconds and 58 °C for 1min at slow ramp speed of 2 °C/second). The samples were subjected to a last extension cycle for 98 °C for 10 mins followed by 4 °C for 30 min for droplet stabilization. LOB and LOD for MPXV was determined to be 0 copies/µL and 0.08 copies/µL (95% confidence) respectively.

2.6 Detection of noroviruses and rotavirus using conventional PCR, gel electrophoresis and sanger sequencing

The samples were tested for Norovirus GI (NoVGI), Norovirus GII (NoVGII) and Rotavirus A (RoA) using gene specific primers (Thongprachum et al., 2018). Reverse transcription (RT) was performed using Invitrogen SuperScript™ IV Reverse Transcriptase (cat# 18090010) protocol. A final concentration of 1 µM random primer (cat# 48190011 Invitrogen Thermo Fisher) was added in the reaction. cDNA was generated using a cycle of 23 °C for 10 min, 50 °C for 1 hr, 94 °C for 5 min. Each reaction consisted of samples and NTC from all steps (PEG precipitation, extraction, RT). Norovirus G1 RNA (ATCC number: VR-3234SD) and human rotavirus RNA (ATCC number: VR-2018DQ) were included as RT PTCs.

PCR was performed using Invitrogen Platinum™ SuperFi II DNA Polymerase (cat# 12361010) protocol. The final concentrations of forward and reverse primers for NoVGI, NoVGII, and RoA were 1 μM, 0.8 μM, 1 μM respectively. 10 μL RNA and PCR-grade water were added to make a total volume of reaction of 20 μL. NoVGI was amplified at one cycle of 94 °C for 5 mins, 40 cycles (94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min) and a final extension cycle of 72 °C for 7 mins. NoVGII and RoA were amplified at one cycle of 94 °C for 3 mins, 35 cycles (94 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 1 min) and a final extension cycle of 72 °C for 7 mins. An additional NTC was included for PCR step to check for cross contamination.

The results were confirmed by running the samples on 1.5% agarose gels. All PTC for NoVGI and RoA assays were confirmed with gel electrophoresis. The samples were PCR purified using QIAquick PCR Purification Kit (QIAGEN cat#28104), quantified by Thermo Scientific Qubit Fluorometer and submitted to RTSF at Michigan State University for sanger sequencing. The result sequences were analyzed using FinchTV 1.5.0 version. After performing quality control, the sequences were blasted using NCBI BLASTn to confirm presence of the viral species. The top BLASTn result under default conditions was chosen as the best possible match. A phylogenetic tree was generated using similar hits from BLAST and other common human viruses within the genus. All sequences are aligned using UGENE software with the MUSCLE algorithm. The aligned regions are extracted for generation of a phylogenetic tree using MEGA11 software (Tamura et al., 2021).

2.7 Clinical data collection

Clinical cases of viral diseases for the year 2022 for Peru, La Libertad, and Trujillo were retrieved from online publicly available sources (Hospital Regional Trujillo, Peru Ministerio de Salud, Peru Ministerio de Salud – cumulative Incident cases, Unique digital platform of the Peruvian State).

3. RESULTS

3.1 Municipal wastewater characterization

Water quality parameters of untreated wastewater were provided by the utility as shown in table 4.1.

PMMoV co-occurs with multiple pathogens of interest and has been suggested as a promising index for enteric viruses (Symonds et al., 2018). Similarly, crAssphage has been

detected abundantly in wastewater and has been proposed as a viral fecal indicator (Dutilh et al., 2014; Stachler et al., 2014). A combination of PMMoV and crAssphage has been used by many researchers as a fecal and enteric virus indicator (Malla et al., 2019; Tandukar et al., 2020; Holm et al 2022; Meuchi et al., 2023).

Levels of PMMoV and crAssphage in wastewater samples were tested and the results are shown in figure 4.1. All the samples tested positive for both viruses. Average concentrations of PMMoV and crAssphage ranged around 10^8 and 10^6 genomic copies per 100ml of wastewater respectively as shown in figure 4.1. The results fall in the range found in untreated wastewater from all around the world (Symonds et al., 2018).

Table 4.1. Water Quality parameters at Covicorti wastewater treatment plant

Description/test	**Influent**	**Final Effluent**
Sampling date	6/9/2022	6/9/2022
Biological oxygen demand (mg/L)	320.69	26.03
Chemical oxygen demand (mg O_2/L)	683.75	122.18
Total suspended solids (mg/L)	376	205
Oils and fats (mg/L)	43.7	5.2
Thermotolerant coliforms (NMP/100 ml)	$1.7 * 10^8$	$3.4 * 10^5$
pH	7.98	7.95

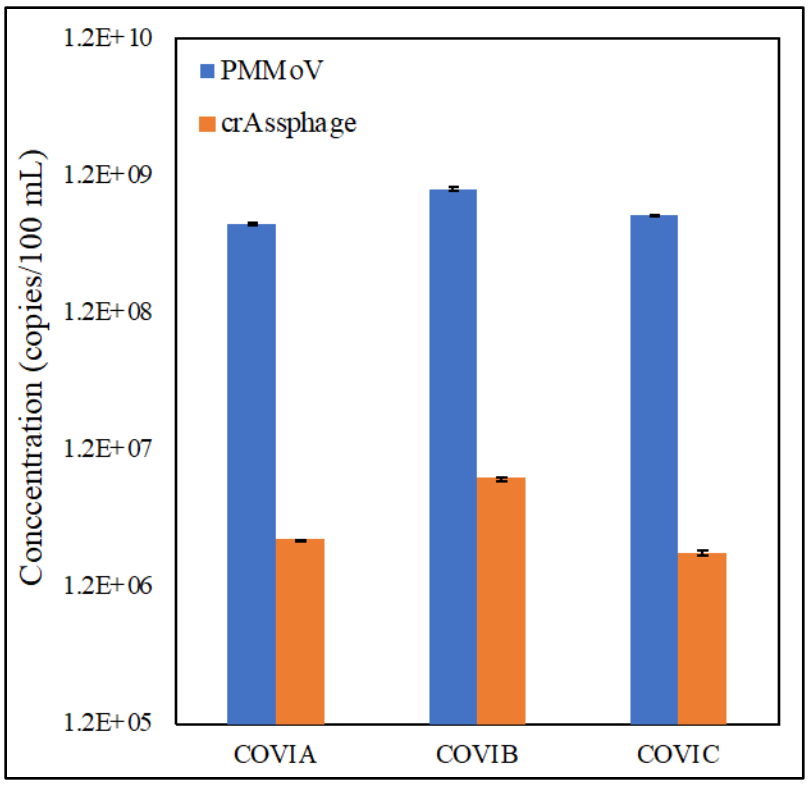

Figure 4.1. Concentrations of enteric virus indicators (PMMoV and crAssphage) for wastewater characterization

3.2 Sequencing and bioinformatics analysis for viral diversity identification

A description of the metagenomic sequencing data after each bioinformatic analysis step is summarized in table 4.2. After trimming, at least 80 million reads were obtained for each sample, among these reads, more than 1.2 million were classified as viruses with kaiju. To achieve substantial gains in taxonomic mapping, long contiguous sequences (contigs) generated by MEGAHIT, and more than 400 thousand contigs were obtained. When comparing these contigs against the NCBI RefSeq virus protein database, more than 17 thousand contigs were assigned to virus with MEGAN.

Table 4.2. Sequencing yield, number of clean reads, viral reads and viral contigs of each wastewater sample

Sample	Yield (Gbp)	Number of clean reads	Number of viral reads	Percent of viral reads (%)	Number of contigs	Number of viral contigs	Percent of viral contigs (%)
COVIA	36.6	118,495,760	6,068,836	5.12	884,618	140,316	15.86
COVIB	33.2	107,865,545	11,395,312	10.56	704,133	114,245	16.22
COVIC	32.0	104,061,454	3,720,569	3.58	669,569	83,766	12.51

Virus composition in the family level was analyzed by comparing the assembled contigs against the NCBI RefSeq virus protein database. The results showed the viral populations that took a large proportion in the wastewater samples were affiliated to bacteriophage, and are members of *Myoviridae, Siphoviridae* and *Microviridae* families.

To improve the detection of human viruses in samples collected from the wastewater utility in Peru, assembled contigs were compared against a customized Swiss-Prot human virus protein database. Human viruses identified in less than 2 samples were excluded. Figure 4.2 shows the diversity of viruses detected at genus level. Among the human viral contigs, *Orthopoxvirus* took a large proportion in all the samples, followed by *Hepatovirus, Rhadinovirus, Parechovirus, Mamastrovirus, Enterovirus, Varicellovirus, Norovirus, Kobuvirus, Bocaparvovirus, Simplexvirus, Spumavirus, Orthohepevirus, Cardiovirus, Molliscipoxvirus, Salivirus, Parapoxvirus, Gammaretrovirus, Alphavirus, Lymphocryptovirus, Erythroparvovirus, Sapovirus, Cosavirus, Deltaretrovirus, Roseolovirus, Flavivirus, Betacoronavirus, Rubivirus, Lentivirus, Betapolyomavirus, Rotavirus, Hepacivirus, Alphacoronavirus, Mastadenovirus, Cytomegalovirus* and *Alphapapillomavirus*

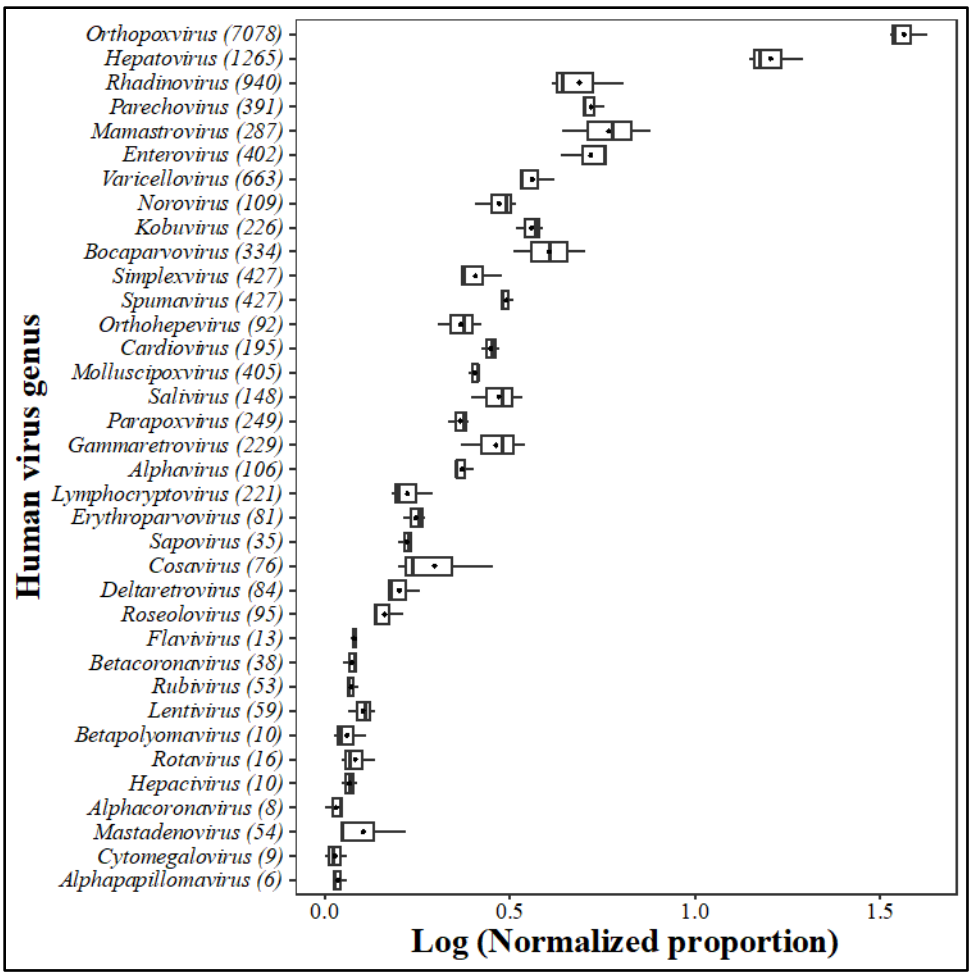

Figure 4.2. Log (Normalized proportion) of human viruses in wastewater collected from Trujillo-Peru. Contigs in three samples (COVIA, COVIB and COVIC) were combined together to do the calculation. Numbers in the brackets after the genus names are the sum of human viral contigs identified in the three samples

3.3 Quantification of SARS-CoV-2 and MPXV

A total of nine samples were taken at Covicorti wastewater utility by grab sampling followed by polyethylene glycol (PEG) precipitation. SARS-CoV-2 and MPXV virus's assays were run in triplicates for all nine samples along with positive and negative controls. Figure 4.3 shows SARS-CoV-2 and MPXV concentrations in copies/100 mL of wastewater. COVIB had the highest concentrations for both SARS-CoV-2 and MPXV with about 2500 and 700 copies/100 mL respectively. MPXV concentrations were almost three times less than SARS-CoV-2 by ddPCR.

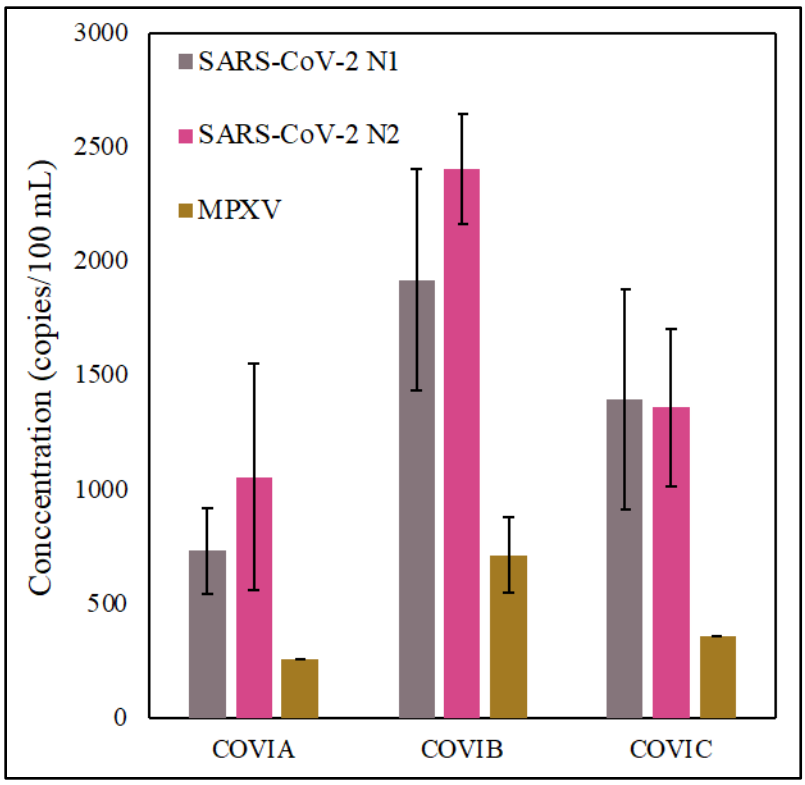

Figure 4.3. Confirmation of selected human viruses (SARS-CoV-2 [N1 and N2 genes] and monkeypox viruses) in wastewater collected from Trujillo-Peru by digital droplet PCR

3.4 Detection of Norovirus GI (NoVGI), Norovirus GII (NoVGII), and Rotavirus A (RoA) by PCR, gel electrophoresis, BLASTn analysis

Presence of NoVGI, NoVGII, and RoA was detected by PCR. For gel electrophoresis, all nine samples tested positive for NoVGI, NoVGII, and RoA. The amplicon size for NoVGI, NoVGII, and RoA was 330, 387, and 395 respectively (Thongprachum et al., 2018). The samples along with positive controls were PCR purified and submitted for sanger sequencing with forward primer and reverse primer separately. The resulting sequence was blasted against viral database on NCBI Blast website. The top 10 BLASTn sequences producing significant alignments results for all samples and positive controls amplified using gene specific primers (NoVGI, NoVGII, RoA) resulted in the respective viruses with at least 94% percent identity and 98% query coverage.

Using the BLASTn result, similar hits and other common human viruses were downloaded and aligned. Phylogenetic trees were generated using the aligned regions with MEGA11 software (Tamura et al., 2021). Phylogenetic trees generated with selected human viral related sequences identified in COVIC1, COVIC2, COVIC3 samples, reference sequences they

are affiliated with, and PTCs purchased from ATCC were shown in figure 4.4. All the nine samples BLASTn results were found to be closely associated with the respective viruses as shown in phylogenetic trees.

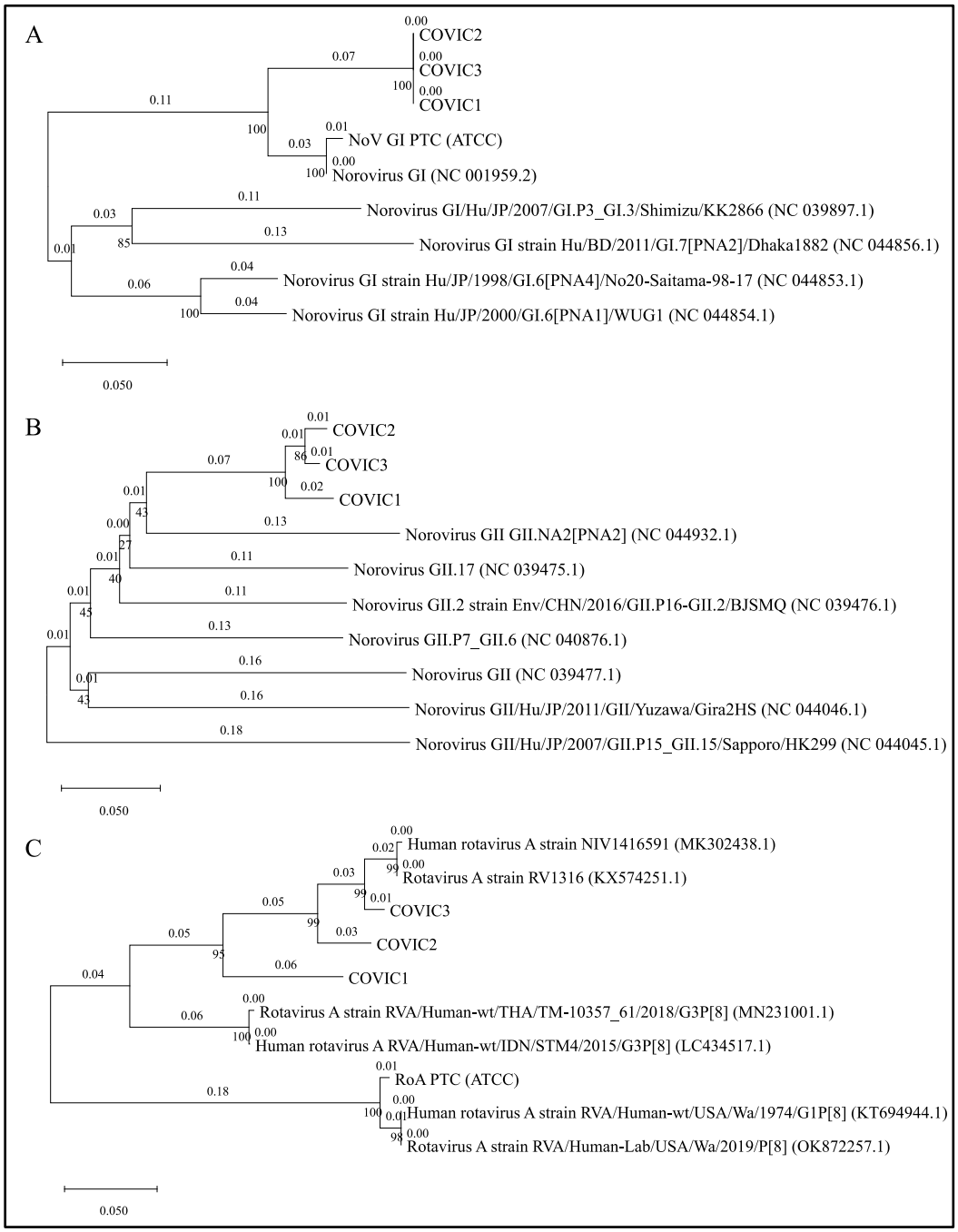

Figure 4.4. Phylogenetic analysis of selected human viruses identified in wastewater in Trujillo-Peru. A: Norovirus GI. B: Norovirus GII. C: Rotavirus A. Reference genomes they are affiliated with, positive controls used in this study and human viral related sequences identified

Figure 4.4. (cont'd)

in the samples are included. The tree was constructed using the neighbor-joining method with the replicates of bootstrap test as 1000. Percentage of replicate trees in which the associated sequences clustered together are shown below the branches. The evolutionary distances were computed with the Kimura 2-parameter method for Norovirus GI and GII related sequences. Tamura 3-parameter method was used for Rotavirus A related sequences. Gaps and missing data were eliminated (complete deletion option). The scale bar represents the estimated number of base substitutions per site. MEGA 11 was used to perform the phylogenetic analysis. Accession numbers of the reference sequences in NCBI are shown in the brackets after the names

3.5 Publicly available health records of clinical confirmed viral disease in the area

Table 4.3 contains all clinical cases available for Peru, La Libertad, and Trujillo (Hospital Regional Trujillo, Peru Ministerio de Salud, Peru Ministerio de Salud – cumulative Incident cases, Unique digital platform of the Peruvian State). During 2022, Trujillo had 2, 5, 21, 151, and 148, cases of chicken pox, monkeypox, dengue, acute diarrheal disease, and acute respiratory infections respectively. There were no reported cases of Yellow fever, Hepatitis B, Zika, Chikungunya in Trujillo for the year 2022. Comparatively, Peru as a whole country had 3, 811, 10, 164 cases of Yellow fever, Hepatitis B, Zika, and Chikungunya respectively. Of all diseases acute diarrheal, acute respiratory, and Dengue cases are the highest in the country of Peru, the state of La Libertad, and the city of Trujillo. It is important to note that most viral diseases in the area are unreported.

Table 4.3. Clinical cases of viral diseases for the year 2022 in Peru

Location/ Viral disease	**Peru**	**La Libertad**	**Trujillo**
Acute respiratory infections	622,259	448	148
Acute diarrheal disease	306,201	288	151
Dengue	38,877	162	21
Monkey pox	3,367	20	5
Hepatitis B	811	8	0
Chicken pox	376	46	2
Chikun-gunya	164	0	0
Zika	10	0	0
Yellow fever	3	0	0

Note: All clinical cases were obtained from BIBLIOGRAPHY (Hospital Regional Trujillo, Peru Ministerio de Salud, Peru Ministerio de Salud – cumulative Incident cases, Unique digital platform of the Peruvian State)

4. DISCUSSION

Since the COVID-19 pandemic multiple efforts have focused on wastewater surveillance. Most studies focused on selected known pathogens, such as SARS-CoV2 (Zhao et al., 2022; Miyani et al., 2020; Randazzo et al., 2020) or MPXV (Girón et al., 2023; de Jonge et al., 2022; Wurtzer et al., 2022). In this study we apply a tool that includes broad screening of viruses at the genus level (McCall et al., 2020; Li et al., 2024a; Li et al., 2024b). The first level of screening includes testing wastewater, or other environmental samples, with high throughput sequencing followed by bioinformatics analysis. If signals identify potential presence of viruses of concern, further PCR testing targeting the specific virus follows. Final confirmation with clinical surveillance is recommended. We have successfully applied this method in samples collected in Detroit MI to identify endemic and emerging diseases in the area (McCall et al., 2020; Miyani et al., 2021b; Li et al., 2024a; Li et al., 2024b; McCall et al., 2021). Application of such a screening tool has the potential to identify endemic and emerging diseases that may otherwise be missed by regular clinical testing.

The screening analysis presented in this study revealed the presence of a wide variety of human and zoonotic, respiratory, bloodborne, enteric and vector-borne viruses circulating in the population (figure 4.2). For example, genera *Betacoronavirus* and *Orthopoxvirus* (that includes species SARS-CoV-2 and MPXV respectively) were detected by high throughput sequencing (figure 4.2). To further investigate and validate these findings SARS-CoV-2 and MPXV were quantified with ddPCR (figure 4.3). Clinical cases of COVID-19 and MPX reported by Peru Ministry of Health (table 4.3) confirm our findings. Until November 2022, Peru reported the greatest number of MPX cases among countries where the disease is not endemic (like Africa) (PAHO WHO). Similarly, NoV GI, NoV GII, and RoA viruses that belong to genera *Norovirus* and *Rotavirus* were selected for further testing and confirmation to target part of the large number of diarrheal cases detected in Trujillo-Peru (table 4.3). Confirmation tests for SARS-CoV2, MPXV, norovirus and rotavirus validate the utility of the method. Importantly, the genomic sequences shown in figure 4.2 reveal the potential presence of a wider range of human and zoonotic viruses.

Many of the viral-related sequences identified in this study (figure 4.2) correspond to viruses that have been detected in wastewater or other environmental samples in Latin American countries. For example, *Orthopoxvirus* was the largest genus detected in all three wastewater samples tested in this study. MPXV, a viral species in genus Orthopoxvirus, has been detected in wastewater in Chile (Ampuero et al., 2023). Vaccinia virus, another viral species of genus Orthopoxvirus caused outbreaks in animals and workers in the dairy industry in Brazil (Peres et al., 2013; Abrahão et al., 2015; Kroon et al., 2011) and Colombia (Styczynski et al., 2019). Hepatovirus genus, consisting of different types of hepatitis viruses, was the second largest genera to be detected in our samples. Hepatitis A and E virus has been detected in wastewaters in Brazil (Prado et al., 2012), Venezuela (Rodríguez-Díaz et al., 2009), Argentina (Yanez et al., 2014; Castro et al., 2023), Colombia (Baez et al., 2017), and Ecuador (Guerrero-Latorre et al., 2018). SARS-CoV-2 (belonging to genus Betacoronavirus which was detected in our samples) has been detected in wastewater in Peru (Pardo-Figueroa et al., 2022; Reyes-Calderón et al., 2022), Brazil (Fongaro et al., 2021), and Ecuador (Guerrero-Latorre et al., 2020).

Gastroenteritis related viruses belonging to genera *Cosavirus, Kobuvirus, Mamastrovirus, Norovirus, Rotavirus, Salivirus,* and *Sapovirus* were detected in our samples. Cosavirus was detected in river waters in Argentina (López et al., 2021) and Ecuador (Guerrero-Latorre et al.,

2018) and wastewater samples in Venezuela (Zamora-Figueroa et al., 2024). Aichivirus (belonging to genus *Kobuvirus*) was tested positive in wastewater samples of Uruguay (Burutarán et al., 2016) and Ecuador (Guerrero-Latorre et al., 2018). Astroviruses were present in wastewater of Brazil (Guimarães et al., 2008), Venezuela (Rodríguez-Díaz et al., 2009), Ecuador (Guerrero-Latorre et al., 2018), and Uruguay (Lizasoain et al., 2015). Wastewater of Latin American countries like Brazil (Fumian et al., 2019), Chile (Díaz et al., 2012), Argentina (Fernández et al., 2011), Nicaragua (Bucardo et al., 2011), Uruguay (Victoria et al., 2014), Ecuador (Guerrero-Latorre et al., 2018), and Venezuela (Rodríguez-Díaz et al., 2009) tested positive for norovirus. Comparatively, rotavirus was detected in wastewater samples of Uruguay (Tort et al., 2015), Brazil (Mehnert et al., 1993), and Argentina (Barril et al., 2010). Klassevirus and sapovirus were detected in river waters (Calgua et al., 2013) and wastewater (Guerrero-Latorre et al., 2018; Fioretti et al., 2016) of Brazil and Ecuador respectively.

Genera such as *Bocaparvovirus* and *Mastadenovirus* have been known to be associated with both gastroenteritis and respiratory illnesses. Human bocaparvovirus has been detected in wastewaters of Uruguay (Salvo et al., 2018; Salvo et al., 2019) and Ecuador (Guerrero-Latorre et al., 2018), whereas various types of human mastadenovirus have been detected in Brazil (Santos et al., 2004), Argentina (Ferreyra et al., 2015), Ecuador (Guerrero-Latorre et al., 2018), and Venezuela (Rodríguez-Díaz et al., 2009). Other examples include Saffold virus that belongs to Cardiovirus genus and has been detected in wastewater of Argentina (López et al., 2021) and Ecuador (Guerrero-Latorre et al., 2018). Human parechovirus, belonging to genus Parechovirus, has been detected in wastewater in Ecuador (Guerrero-Latorre et al., 2018). Human polyomaviruses belonging to genus Betapolyomavirus have been detected in raw sewage in Argentina (Torres et al., 2016), Brazil (Fumian et al., 2010), and Chile (Levican et al., 2019).

This proposed screening method has the potential to identify non-reportable (not required to be tested for at clinical settings) human and zoonotic viruses that may be emerging in the community. For some of those emerging viruses there may not be validated clinical tests available. While this study demonstrates the advantages of wastewater surveillance with molecular methods, high-throughput sequencing, and bioinformatics to identify endemic and emerging diseases in the contributing population, there are several limitations. Particularly in resource-limited nations, obtaining all the requisite resources for such endeavors can pose challenges. Another limitation of this study is the sample size, as only three pooled samples were

tested with high-throughput sequencing and bioinformatics methods. This restricted number of samples may limit the generalizability of the findings that can only serve as a pilot study. The results of this study only indicate the possibility of occurrence of the detected viral genera in the community. To develop a tool that can predict viral disease fluctuations over space a time multiple more samples have to be collected, analyzed, normalized, and compared with clinical data in the catchment area. This will require a long-term full-scale investigation that needs to involve local public health officials and environmental scientists and engineers.

5. CONCLUSIONS

Wastewater, or other environmental media surveillance, combined with bioinformatics has a vast potential to serve as a systematic screening tool for the identification of a myriad of human or zoonotic viruses that may cause disease. One common problem in Peru and other Latin America countries is that health agencies do not have the resources and tools to anticipate and systematically monitor disease outbreaks. The methodology presented here provides a practical method to assist with clinical surveillance and prediction.

The viral species detected with bioinformatics methods (figure 2) are at the genus level and need to be confirmed by additional PCR based methods for species identification. The results highlight the importance of the method as an initial screening tool. The results may be used by local professionals to guide further clinical monitoring and develop health care policies. Incorporation of this bioinformatic-based screening tool by public health officials in Latin America will help identify endemic and emerging diseases that could save human lives and resources.

BIBLIOGRAPHY

"Peru - National geographic kids." [Online]. Available: https://kids.nationalgeographic.com/geography/countries/article/peru. [Accessed: 09-Dec-2023].

Uhart M, Pérez A, Rostal M, Robles EA, Mendoza AP, Nava A, de Paula CD, Miranda F, Iñiguez V, Zambrana C, Durigon E. A 'One Health'approach to predict emerging zoonoses in the Amazon. Saúde Silvestre e Humana: Experiências e Perspectivas. Rio de Janeiro, FIOCRUZ. 2013:65-73.

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. Global trends in emerging infectious diseases. Nature. 2008 Feb;451(7181):990-3.

Woolhouse M, Gaunt E. Ecological origins of novel human pathogens. Critical reviews in microbiology. 2007 Jan 1;33(4):231-42.

Karesh WB, Cook RA. The human-animal link. Foreign Aff.. 2005;84:38.

Flanagan ML, Parrish CR, Cobey S, Glass GE, Bush RM, Leighton TJ. Anticipating the species jump: surveillance for emerging viral threats. Zoonoses and public health. 2012 May;59(3):155-63.

O'Brien E, Xagoraraki I. A water-focused one-health approach for early detection and prevention of viral outbreaks. One Health. 2019 Jun 1;7:100094.

Xagoraraki I, O'Brien E. Wastewater-based epidemiology for early detection of viral outbreaks. Women in water quality: Investigations by prominent female engineers. 2020a:75-97.

Xagoraraki I. Can we predict viral outbreaks using wastewater surveillance? Journal of Environmental Engineering. 2020b Nov 1;146(11):01820003.

Li Y, Miyani B, Zhao L, Spooner M, Gentry Z, Zou Y, Rhodes G, Li H, Kaye A, Norton J, Xagoraraki I. Surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States: Assessing per capita SARS-CoV-2 estimations and COVID-19 incidence. Science of The Total Environment. 2022 Dec 10;851:158350.

Zhao L, Zou Y, Li Y, Miyani B, Spooner M, Gentry Z, Jacobi S, David RE, Withington S, McFarlane S, Faust R. Five-week warning of COVID-19 peaks prior to the Omicron surge in Detroit, Michigan using wastewater surveillance. Science of The Total Environment. 2022 Oct 20;844:157040.

Miyani B, Zhao L, Spooner M, Buch S, Gentry Z, Mehrotra A, Norton J, Xagoraraki I. Early warnings of COVID-19 second wave in Detroit. Journal of Environmental Engineering. 2021 Aug 1;147(8):06021004.

McCall C, Wu H, Miyani B, Xagoraraki I. Identification of multiple potential viral diseases in a large urban center using wastewater surveillance. Water research. 2020 Oct 1;184:116160.

Miyani B, McCall C, Xagoraraki I. High abundance of human herpesvirus 8 in wastewater from a large urban area. Journal of Applied Microbiology. 2021 May 1;130(5):1402-11.

Li Y, Miyani B, Childs KL, Shiu SH, Xagoraraki I. Effect of wastewater collection and

concentration methods on assessment of viral diversity. Science of The Total Environment. 2024a Jan 15;908:168128.

Li Y, Miyani B, Faust RA, David RE, Xagoraraki I. A broad wastewater screening and clinical data surveillance for virus-related diseases in the metropolitan Detroit area in Michigan. Human Genomics. 2024b Dec;18(1):1-8.

"Peru WHO Mortality Database Leading causes of death per year."

"Peru: WHO Coronavirus (COVID-19) Dashboard."

Hasan MM, Hernández-Yépez PJ, de los Angeles Rivera-Cabrera M, Sarkar A, dos Santos Costa AC, Essar MY. Concurrent epidemics of dengue and COVID-19 in Peru: Which way forward?. The Lancet Regional Health–Americas. 2022 Aug 1;12.

Centro Nacional de Epidemiologia Prevencion y control de enfermedades, "Incremento de incidencia de casos y ocurrencia de brotes por Dengue y otras arbovirosis en el país," Minist. Salud, vol. 19, no. 2, p. 6, 2021

"Instituto Nacional de Estadistica e Informatica, INEI, Republica del Peru (2017). La Libertad: Resultados Definitivos, Censo Nacional," 2017

Avila Calderón SL. Caracterización físico-química del efluente final de uso agrícola de la Planta de Tratamiento El Cortijo, Trujillo, Perú.

I. D. E. C. Profesional, "ca lio te de In g ie ría uí m ica ca lio te de Bi In g en ie ría m," 2006

Flood MT, D'Souza N, Rose JB, Aw TG. 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. 2021 Sep;13:303-15.

Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL. Microarray-based detection and genotyping of viral pathogens. Proceedings of the National Academy of Sciences. 2002 Nov 26;99(24):15687-92.

Wang D, Urisman A, Liu YT, Springer M, Ksiazek TG, Erdman DD, Mardis ER, Hickenbotham M, Magrini V, Eldred J, Latreille JP. Viral discovery and sequence recovery using DNA microarrays. PLoS biology. 2003 Nov;1(2):e2.

Andrews S. FastQC: a quality control tool for high throughput sequence data.2010

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1;30(15):2114-20.

Menzel P, Ng KL, Krogh A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature communications. 2016 Apr 13;7(1):11257.

Bibby K, Viau E, Peccia J. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. Letters in applied microbiology. 2011 Apr 1;52(4):386-92.

Bağcı C, Patz S, Huson DH. DIAMOND+ MEGAN: fast and easy taxonomic and functional analysis of short and long microbiome sequences. Current protocols. 2021 Mar;1(3):e59.

Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R.

MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. PLoS computational biology. 2016 Jun 21;12(6):e1004957.

Centers for Disease Control and Prevention., “Centers for Disease Control & Prevention Poxvirus & Rabies Branch (PRB),” 01, no. 01, 2022

Li Y, Zhao H, Wilkins K, Hughes C, Damon IK. Real-time PCR assays for the specific detection of monkeypox virus West African and Congo Basin strain DNA. Journal of virological methods. 2010 Oct 1;169(1):223-7.

Thongprachum A, Fujimoto T, Takanashi S, Saito H, Okitsu S, Shimizu H, Khamrin P, Maneekarn N, Hayakawa S, Ushijima H. Detection of nineteen enteric viruses in raw sewage in Japan. Infection, Genetics and Evolution. 2018 Sep 1;63:17-23.

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Molecular biology and evolution. 2021 Jul 1;38(7):3022-7.

“Hospital Regional Docente De Trujillo - Epidemiological Bulletins.” [Online]. Available: https://www.hrdt.gob.pe/site/index.php/servicios-linea/publicaciones/boletines-epidemiologicos. [Accessed: 01-Dec-2023]

“Peru Ministerio de Salud. National Center for Epidemiology, Disease Prevention and Control.” [Online]. Available: https://www.dge.gob.pe/portalnuevo/publicaciones/boletines-epidemiologicos/. [Accessed: 01-Dec-2023].

Peru Ministerio de Salud. National Center for Epidemiology, Disease Prevention and Control – Cumulative Incident Cases.” [Online]. Available: https://www.dge.gob.pe/portalnuevo/informacion-publica/casos-de-incidencias-acumuladas/. [Accessed: 01-Dec-2023].

“Unique digital platform of the Peruvian State - Ministry of Health.” [Online]. Available: https://www.gob.pe/minsa. [Accessed: 01-Dec-2023].

Symonds EM, Nguyen KH, Harwood VJ, Breitbart M. Pepper mild mottle virus: a plant pathogen with a greater purpose in (waste) water treatment development and public health management. Water research. 2018 Nov 1;144:1-2

Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GG, Boling L, Barr JJ, Speth DR, Seguritan V, Aziz RK, Felts B. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nature communications. 2014 Jul 24;5(1):4498.

Stachler E, Bibby K. Metagenomic evaluation of the highly abundant human gut bacteriophage CrAssphage for source tracking of human fecal pollution. Environmental Science & Technology Letters. 2014 Oct 14;1(10):405-9.

Malla B, Ghaju Shrestha R, Tandukar S, Sherchand JB, Haramoto E. Performance evaluation of human-specific viral markers and application of pepper mild mottle virus and crAssphage to environmental water samples as fecal pollution markers in the Kathmandu Valley, Nepal. Food and Environmental Virology. 2019 Sep 1;11:274-87.

Tandukar S, Sherchan SP, Haramoto E. Applicability of crAssphage, pepper mild mottle virus,

and tobacco mosaic virus as indicators of reduction of enteric viruses during wastewater treatment. Scientific Reports. 2020 Feb 27;10(1):3616.

Holm RH, Nagarkar M, Yeager RA, Talley D, Chaney AC, Rai JP, Mukherjee A, Rai SN, Bhatnagar A, Smith T. Surveillance of RNase P, PMMoV, and CrAssphage in wastewater as indicators of human fecal concentration across urban sewer neighborhoods, Kentucky. FEMS microbes. 2022;3:xtac003.

Meuchi Y, Nakada M, Kuroda K, Hanamoto S, Hata A. Applicability of F-specific bacteriophage subgroups, PMMoV and crAssphage as indicators of source specific fecal contamination and viral inactivation in rivers in Japan. Plos one. 2023 Jul 14;18(7):e0288454.

Miyani B, Fonoll X, Norton J, Mehrotra A, Xagoraraki I. SARS-CoV-2 in Detroit wastewater. Journal of Environmental Engineering. 2020 Nov 1;146(11):06020004.

Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water research. 2020 Aug 15;181:115942.

Girón-Guzmán I, Díaz-Reolid A, Truchado P, Carcereny A, García-Pedemonte D, Hernáez B, Bosch A, Pintó RM, Guix S, Allende A, Alcamí A. Spanish wastewater reveals the current spread of Monkeypox virus. Water research. 2023 Mar 1;231:119621.

de Jonge EF, Peterse CM, Koelewijn JM, van der Drift AM, van der Beek RF, Nagelkerke E, Lodder WJ. The detection of monkeypox virus DNA in wastewater samples in the Netherlands. Science of The Total Environment. 2022 Dec 15;852:158265.

Wurtzer S, Levert M, Dhenain E, Boni M, Tournier JN, Londinsky N, Lefranc A, OBEPINE SIG, Ferraris O, Moulin L. First detection of monkeypox virus genome in sewersheds in France: the potential of wastewater-based epidemiology for monitoring emerging disease. Environmental Science & Technology Letters. 2022 Oct 20;9(11):991-6.

McCall C, Wu H, O'Brien E, Xagoraraki I. Assessment of enteric viruses during a hepatitis outbreak in Detroit MI using wastewater surveillance and metagenomic analysis. Journal of Applied Microbiology. 2021 Sep 1;131(3):1539-54.

"Pan American Health Organization (PAHO) World Health Organization (WHO) - Peru Responds to Monkeypox by Engaging Affected Communities." [Online]. Available: https://www.paho.org/en/stories/peru-responds-monkeypox-engaging-affected-communities. [Accessed: 07-Feb-2024].

Ampuero M, Martínez-Valdebenito C, Ferrés M, Soto-Rifo R, Gaggero A. Monkeypox Virus in Wastewater Samples from Santiago Metropolitan Region, Chile. Emerging Infectious Diseases. 2023 Nov;29(11):2358.

Peres MG, Bacchiega TS, Appolinário CM, Vicente AF, Allendorf SD, Antunes JM, Moreira SA, Legatti E, Fonseca CR, Pituco EM, Okuda LH. Serological study of vaccinia virus reservoirs in areas with and without official reports of outbreaks in cattle and humans in São Paulo, Brazil. Archives of Virology. 2013 Dec;158:2433-41.

Abrahão JS, Campos RK, de Souza Trindade G, Da Fonseca FG, Ferreira PC, Kroon EG.

Outbreak of severe zoonotic vaccinia virus infection, Southeastern Brazil. Emerging infectious diseases. 2015 Apr;21(4):695.

Kroon EG, Mota BE, Abrahão JS, da Fonseca FG, de Souza Trindade G. Zoonotic Brazilian Vaccinia virus: from field to therapy. Antiviral Research. 2011 Nov 1;92(2):150-63.

Styczynski A, Burgado J, Walteros D, Usme-Ciro J, Laiton K, Farias AP, Nakazawa Y, Chapman C, Davidson W, Mauldin M, Morgan C. Seroprevalence and risk factors possibly associated with emerging zoonotic vaccinia virus in a farming community, Colombia. Emerging Infectious Diseases. 2019 Dec;25(12):2169.

Prado T, Fumian TM, Miagostovich MP, Gaspar AM. Monitoring the hepatitis A virus in urban wastewater from Rio de Janeiro, Brazil. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2012 Feb 1;106(2):104-9.

Rodríguez-Díaz J, Querales L, Caraballo L, Vizzi E, Liprandi F, Takiff H, Betancourt WQ. Detection and characterization of waterborne gastroenteritis viruses in urban sewage and sewage-polluted river waters in Caracas, Venezuela. Applied and Environmental Microbiology. 2009 Jan 15;75(2):387-94.

Yanez LA, Lucero NS, Barril PA, Díaz MD, Tenaglia MM, Spinsanti LI, Nates SV, Isa MB, Ré VE. Evidence of hepatitis A virus circulation in central Argentina: seroprevalence and environmental surveillance. Journal of Clinical Virology. 2014 Jan 1;59(1):38-43.

Baez PA, Lopez MC, Duque-Jaramillo A, Pelaez D, Molina F, Navas MC. First evidence of the Hepatitis E virus in environmental waters in Colombia. PLoS One. 2017 May 16;12(5):e0177525.

Castro IL, Espul C, Paula VS, Altabert NR, Gonzalez JE, Lago BV, Villar LM. High prevalence of hepatitis A and E viruses in environmental and clinical samples from West Argentina. Brazilian Journal of Infectious Diseases. 2023 May 19;27:102738.

Guerrero-Latorre L, Romero B, Bonifaz E, Timoneda N, Rusiñol M, Girones R, Rios-Touma B. Quito's virome: Metagenomic analysis of viral diversity in urban streams of Ecuador's capital city. Science of the Total Environment. 2018 Dec 15;645:1334-43.

Pardo-Figueroa B, Mindreau-Ganoza E, Reyes-Calderon A, Yufra SP, Solorzano-Ortiz IM, Donayre-Torres AJ, Antonini C, Renom JM, Quispe AM, Mota CR, Chernicharo CA. Spatiotemporal surveillance of SARS-CoV-2 in the sewage of three major urban areas in Peru: generating valuable data where clinical testing is extremely limited. Acs Es&T Water. 2022 May 20;2(11):2144-57.

Reyes-Calderón A, Mindreau-Ganoza E, Pardo-Figueroa B, Garcia-Luquillas KR, Yufra SP, Romero PE, Antonini C, Renom JM, Mota C, Santa-Maria MC. Evaluation of Low-Cost SARS-CoV-2 Purification Methods for Viral Quantification and NGS Analysis: Implications for Wider Wastewater-Based Epidemiology Adoption. Available at SSRN 4308586. 2022 Dec 23.

Fongaro G, Stoco PH, Souza DS, Grisard EC, Magri ME, Rogovski P, Schörner MA, Barazzetti FH, Christoff AP, de Oliveira LF, Bazzo ML. The presence of SARS-CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. Science of The Total Environment. 2021 Jul 15;778:146198.

Guerrero-Latorre L, Ballesteros I, Villacrés-Granda I, Granda MG, Freire-Paspuel B, Ríos-Touma B. SARS-CoV-2 in river water: Implications in low sanitation countries. Science of the Total environment. 2020 Nov 15;743:140832.

López GR, Martinez LM, Freyre L, Freire MC, Vladimirsky S, Rabossi A, Cisterna DM. Persistent detection of cosavirus and saffold cardiovirus in Riachuelo River, Argentina. Food and Environmental Virology. 2021 Mar;13:64-73.

Zamora-Figueroa A, Rosales RE, Fernández R, Ramírez V, Bastardo M, Farías A, Vizzi E. Detection and diversity of gastrointestinal viruses in wastewater from Caracas, Venezuela, 2021-2022. Virology. 2024 Jan 1;589:109913.

Burutarán L, Lizasoain A, García M, Tort LF, Colina R, Victoria M. Detection and molecular characterization of Aichivirus 1 in wastewater samples from Uruguay. Food and environmental virology. 2016 Mar;8:13-7.

Guimarães FR, Ferreira FF, Vieira CB, Fumian TM, Shubo T, Leite JP, Miagostovich MP. Molecular detection of human astrovirus in an urban sewage treatment plant in Rio de Janeiro, Brazil. Memórias do Instituto Oswaldo Cruz. 2008;103:819-23.

Lizasoain A, Tort LF, García M, Gómez MM, Cristina J, Leite JP, Miagostovich MP, Victoria M, Colina R. Environmental assessment of classical human astrovirus in Uruguay. Food and Environmental Virology. 2015 Jun;7:142-8.

Fumian TM, Fioretti JM, Lun JH, Dos Santos IA, White PA, Miagostovich MP. Detection of norovirus epidemic genotypes in raw sewage using next generation sequencing. Environment International. 2019 Feb 1;123:282-91.

Díaz J, Solari V, Cáceres O, Mena J, Baeza S, Muñoz X, O'Ryan M, Galeno H, Maldonado A, Mamani N. Brote de gastroenteritis aguda en la Región de Antofagasta, Chile: 2010. Revista chilena de infectología. 2012 Feb;29(1):19-25.

Fernández MD, Torres C, Martínez LC, Giordano MO, Masachessi G, Barril PA, Isa MB, Campos RH, Nates SV, Mbayed VA. Genetic and evolutionary characterization of norovirus from sewage and surface waters in Cordoba City, Argentina. Infection, Genetics and Evolution. 2011 Oct 1;11(7):1631-7.

Bucardo F, Lindgren PE, Svensson L, Nordgren J. Low prevalence of rotavirus and high prevalence of norovirus in hospital and community wastewater after introduction of rotavirus vaccine in Nicaragua. PLoS One. 2011 Oct 7;6(10):e25962.

Victoria M, Tort LL, García M, Lizasoain A, Maya L, Leite JP, Miagostovich MP, Cristina J, Colina R. Assessment of gastroenteric viruses from wastewater directly discharged into Uruguay River, Uruguay. Food and Environmental Virology. 2014 Jun;6:116-24.

Tort LF, Victoria M, Lizasoain A, García M, Berois M, Cristina J, Leite JP, Gómez MM, Miagostovich MP, Colina R. Detection of common, emerging and uncommon VP4, and VP7 human group A rotavirus genotypes from urban sewage samples in Uruguay. Food and environmental virology. 2015 Dec;7:342-53.

Mehnert DU, Stewien KE. Detection and distribution of rotavirus in raw sewage and creeks in São Paulo, Brazil. Applied and Environmental Microbiology. 1993 Jan;59(1):140-3.

Barril PA, Giordano MO, Isa MB, Masachessi G, Ferreyra LJ, Castello AA, Glikmann G, Nates SV. Correlation between rotavirus A genotypes detected in hospitalized children and sewage samples in 2006, Córdoba, Argentina. Journal of medical virology. 2010 Jul;82(7):1277-81.

Calgua B, Fumian T, Rusinol M, Rodriguez-Manzano J, Mbayed VA, Bofill-Mas S, Miagostovich M, Girones R. Detection and quantification of classic and emerging viruses by skimmed-milk flocculation and PCR in river water from two geographical areas. Water research. 2013 May 15;47(8):2797-810.

Fioretti JM, Rocha MS, Fumian TM, Ginuino A, da Silva TP, de Assis MR, Rodrigues JD, Carvalho-Costa FA, Miagostovich MP. Occurrence of human sapoviruses in wastewater and stool samples in Rio De Janeiro, Brazil. Journal of Applied Microbiology. 2016 Sep 1;121(3):855-62.

Salvo M, Lizasoain A, Castells M, Bortagaray V, Castro S, Colina R, Tort FL, Victoria M. Human bocavirus: detection, quantification and molecular characterization in sewage and surface waters in Uruguay. Food and environmental virology. 2018 Jun;10:193-200.

Salvo M, Mir D, Tort LF, Lizasoain A, Colina R, Victoria M. Description of a human Bocavirus recombinant strain in the Americas. Memórias do Instituto Oswaldo Cruz. 2019 Oct 21;114.

Santos FM, Vieira MJ, Garrafa P, Monezi TA, Pellizari VH, Hársi CM, Mehnert DU. Discrimination of adenovirus types circulating in urban sewage and surface polluted waters in São Paulo city, Brazil. Water Science and Technology: Water Supply. 2004 Apr 1;4(2):79-85.

Ferreyra LJ, Giordano MO, Martinez LC, Barril PA, Masachessi G, Isa MB, Poma R, Rajal V, Biganzoli P, Nates SV, Pavan JV. Tracking novel adenovirus in environmental and human clinical samples: no evidence of endemic human adenovirus type 58 circulation in Cordoba city, Argentina. Epidemiology & Infection. 2015 May;143(7):1427-31.

Torres C, Barrios ME, Cammarata RV, Cisterna DM, Estrada T, Novas SM, Cahn P, Fernández MD, Mbayed VA. High diversity of human polyomaviruses in environmental and clinical samples in Argentina: detection of JC, BK, Merkel-cell, Malawi, and human 6 and 7 polyomaviruses. Science of the Total Environment. 2016 Jan 15;542:192-202.

Fumian TM, Guimarães FR, Pereira Vaz BJ, da Silva MT, Muylaert FF, Bofill-Mas S, Gironés R, Leite JP, Miagostovich MP. Molecular detection, quantification and characterization of human polyomavirus JC from wastewater in Rio De Janeiro, Brazil. Journal of Water and Health. 2010 Sep 1;8(3):438-45.

Levican J, Levican A, Ampuero M, Gaggero A. JC polyomavirus circulation in one-year surveillance in wastewater in Santiago, Chile. Infection, Genetics and Evolution. 2019 Jul 1;71:151-8.

Institute for Health Metrics and Evaluation (IHME) Covid-19 (https://covid19.healthdata.org/global?view=cumulative-deaths&tab=trend)

Peru Covid-19 statistics worldometer (https://www.worldometers.info/coronavirus/country/peru/)

Peru has the world's highest COVID death rate. Here's why NPR WAMU 88.5 (https://www.npr.org/sections/goatsandsoda/2021/11/27/1057387896/peru-has-the-worlds-highest-covid-death-rate-heres-why)

Johns Hopkins University & Medicine CORONAVIRUS RESEARCH CENTER (https://coronavirus.jhu.edu/data/mortality)

World Health Organization Disease Outbreak News; Multi-country monkeypox outbreak in non-endemic countries. 2022. (https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON390)

Antinori A, Mazzotta V, Vita S, Carletti F, Tacconi D, Lapini LE, D'Abramo A, Cicalini S, Lapa D, Pittalis S, Puro V. Epidemiological, clinical and virological characteristics of four cases of monkeypox support transmission through sexual contact, Italy, May 2022. Eurosurveillance. 2022 Jun 2;27(22):2200421.

de Jonge EF, Peterse CM, Koelewijn JM, van der Drift AM, van der Beek RF, Nagelkerke E, Lodder WJ. The detection of monkeypox virus DNA in wastewater samples in the Netherlands. Science of The Total Environment. 2022 Dec 15;852:158265.

Polson SW, Wilhelm SW, Wommack KE. Unraveling the viral tapestry (from inside the capsid out). The ISME journal. 2011 Feb;5(2):165-8.

CONCLUSIONS AND SIGNIFICANCE

The first chapter assesses the diversity and abundance of Human Herpesviruses (HHVs) in the influent of an urban wastewater treatment plant by using shotgun sequencing, metagenomic analysis, and qPCR. All HHVs, except HHV-8, are quite prevalent in USA and have been associated with multiple potential diseases. These viruses are being regularly shed from both asymptomatic and symptomatic individuals and ultimately end up in wastewater. Influent wastewater samples were collected from the three interceptors (NI-EA, DRI and O-NWI) serving the City of Detroit and Wayne, Macomb and Oakland counties between November 2017 to February 2018. The samples were subjected to a series of processes to concentrate viruses which were further sequenced and amplified using qPCR. All nine types of human herpesviruses were detected in wastewater using Illumina sequencing, bioinformatics and qPCR, with HHV-8 being the most abundant. According to literature, HHVs are shed at higher rate in HIV immunocompromised patients, solid organ transplant recipients, cancerous patients and critically ill non-immunocompromised patients compared to healthy controls. The high abundance of HHV-8 in the Detroit metropolitan area is attributed to the HIV-AIDS outbreak that was ongoing in Detroit during the sampling period. The approach described in this paper can be used to monitor HHVs to establish a baseline secreted by the community. Sudden changes in the baseline would identify changes in community health and immunity.

The second chapter focuses on using wastewater-based-epidemiology to provide early warnings of the second COVID-19 wave in Detroit metropolitan area in MI, USA. Untreated wastewater samples were collected and processed between Aug 6, and Dec 14, 2020, from the Detroit metropolitan area in southeast Michigan. SARS-CoV-2 RNA was detected in 98% of samples with concentrations ranging from 10^4 to 10^6 copies/L. Early warnings of the second wave of COVID19 were observed ahead of clinical data reporting. Statistical analysis indicated a lag time of four weeks between observed SARS-CoV-2 RNA concentrations in wastewater and COVID-19 reported cases in the community. Along with clinical diagnostic testing, wastewater-based epidemiology may be a helpful resource for health officials in predicting the incidence of SARS-CoV-2 in community. Accurate prediction models can be created by including processes that affect the fate of viruses in the collection network, demographic information, and shedding rate and duration data.

The third chapter focuses on surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States and assessing per capita SARS-CoV-2 estimations and COVID-19 incidence. Average concentrations of N1 (SARS-CoV-2) in wastewater samples collected from nine sites ranged from 3.48×10^2 to 9.37×10^2 gc/L, and for N2, averages ranged from 3.72×10^2 to 8.79×10^2 gc/L. Both levels and loads of SARS-CoV-2 RNA in wastewater in two neighborhoods (MT and SH) were found to be higher than that in the other seven sites. The differences were statistically significant. Comparisons between normalized SARS-CoV-2 by population markers and COVID-19 incidences, indicated that normalization of SARS-CoV-2 RNA concentrations with TKN, creatinine, 5-HIAA and xanthine correlated positively with the COVID-19 incidences.

The fourth chapter uses the bioinformatics-based screening tool to identify a wide variety of human viruses in Trujillo-Peru wastewater and confirms Covid-19, monkeypox, and diarrheal disease outbreaks. Wastewater, or other environmental media surveillance, combined with bioinformatics has a vast potential to serve as a systematic screening tool for the identification of a myriad of human or zoonotic viruses that may cause disease. One common problem in Peru and other Latin America countries is that health agencies do not have the resources and tools to anticipate and systematically monitor disease outbreaks. The methodology presented here provides an economical and practical method to assist with surveillance and prediction. The viral species detected with bioinformatics methods are at the genus level and need to be confirmed by additional PCR based methods for species identification. The results highlight the importance of the method as an initial screening tool. The results may be used by local professionals to guide further clinical monitoring and develop health care policies. Incorporation of this bioinformatic-based screening tool by public health officials in Latin America will help identify endemic and emerging diseases that could save human lives and resources.