Advancement of Sampling Methods and Genomic Analyses for SARS-CoV-2 Wastewater and Environmental Monitoring

Md Pervez Kabir

A doctoral thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Environmental Engineering

Ottawa-Carleton Institute for Environmental Engineering
Department of Civil Engineering
Faculty of Engineering
University of Ottawa

uOttawa

Abstract

Wastewater and environmental monitoring (WEM) have emerged as a pivotal tool for monitoring Coronavirus disease 2019 (COVID-19) within population, offering a non-invasive and cost-effective early-warning system to track infection dynamics. WEM for COVID-19 involves collection of wastewater samples from in-premise plumbing within buildings, within community wastewater infrastructure such as sewer systems, influent and primary sludge from wastewater resource recovery facilities (WWRFs). Various sampling methods are employed for wastewater collection, including 24-hour composite, grab, and passive sampling whereas primary sludge is collected using only 24-hour composite sampling method. Following wastewater samples collection, samples are analyzed to quantify the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) genetic material (RNA) using molecular technique such as reverse transcription polymerase chain reaction (RT-PCR) for understanding the disease prevalence at the population level. Although WEM showed promises, SARS-CoV-2 genome continues to evolve with the progression of the pandemic, leading to the emergence of multiple variants of concern (VOCs). To monitor these VOCs, wastewater genomic surveillance (WWGS) has emerged as another public health tool using composite influent wastewater samples. However, genomic analysis using influent wastewater presents significant challenges, including low concentrations of viral RNA, the presence of PCR inhibitors, RNA degradation in wastewater matrices, and high operational costs associated with sequencing. It is critical to address these challenges to improve the reliability and scalability of WWGS for monitoring emerging SARS-CoV-2 variants. The objective of this dissertation is therefore to evaluate different sampling methods, sample processing strategies such as concentration and RNA extraction methods and genomic analyses for comprehensive SARS-CoV-2 WEM. In particular, the first specific objective is to assess the SARS-CoV-2 RNA concentrations in wastewater solids collected using an autosampler, passive samplers and primary sludge

samples. Results show SARS-CoV-2 RNA concentrations in wastewater solids from passive samplers can be effectively compared ($p > 0.05$) to conventional autosampler and primary sludge samples. The second specific objective is to evaluate the feasibility of passive sampling for WWGS of SARS-CoV-2 in high-flow WWRFs. The findings indicate that single nucleotide variants (SNVs) profile and SARS-CoV-2 lineage prevalence is similar ($p > 0.05$) across auto, COSCa-ball, and Torpedo passive samplers, showing concordance with clinical surveillance data. Notably, the genomic recovery of SARS-CoV-2 from passive samplers is shown to be significantly influenced by sequencing read length, where shorter reads (300 bp) results in lower genomic recovery compared to longer reads (600 bp). The third specific objective centers on optimizing a primary sludge concentration and RNA extraction method for SARS-CoV-2 genome sequencing that yields comparable or improved results compared to conventional WWGS. The findings demonstrate that our optimize sludge processing method consistently recovers near-complete ($\geq 90\%$) SARS-CoV-2 genomes from influent wastewater and primary sludge. Genomic analyses reveal that lineages and SNV profiles are comparable between influent wastewater and primary sludge. However, primary sludge exhibits a higher likelihood of rare (low prevalence) and Canadian cryptic SNVs detection compared to influent wastewater, emphasizing its potential to enhance variant monitoring and genomic resolution in WWGS. The fourth and final specific objective of this dissertation is to compare the diagnostic performance of allele-specific (AS)-RT-qPCR and sequencing-based methods to determine their accuracy in WWGS. The research found that the frequency estimation of single allele using AS-RT-qPCR, amplicon sequencing as well as haplotype frequency estimations are similar and contain sufficient information to describe the trajectory of variant prevalence in wastewater across time. Youden's index further confirms that the diagnostic performance nearly identical across the methods. Overall, this PhD dissertation advances sampling method, sample processing strategy and genomic analyses of SARS-CoV-2 by demonstrating its

potential to deliver real-time, scalable, cost-efficient, and comprehensive data, contributing to improved preparedness for future pandemics.

Preface

This dissertation is an original work prepared by Md Pervez Kabir under the supervision of Drs. Robert Delatolla and Tyson E. Graber. This dissertation comprises of four manuscripts: two have been published, and the remaining two are currently under review in peer-reviewed journals for publication. Versions of these manuscripts are presented in chapters 3 to 6 of this dissertation. A brief introduction of each manuscript, along with author contributions are presented below.

Chapter 3: Manuscript 1

A version of the following manuscript in this chapter has been published in Journal of Environmental Chemical Engineering in 2025: Kabir, M. P., Renouf, E., Pisharody, L., Mercier, E., D'Aoust, P. M., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Tomalty, E., Graber, T. E., and Delatolla, R. *Wastewater solids drive comparability of sampling methods for SARS-CoV-2 wastewater and environmental surveillance*.

Md Pervez Kabir: developed research question, experimental design, sample collection, laboratory experiments, data analysis, prepared and revised the manuscript.

Elizabeth Renouf: reviewed statistical analysis and revised the manuscript.

Lakshmi Pisharody: supported data analysis and revised the manuscript.

Élisabeth Mercier: supported laboratory experiments and revised the manuscript.

Patrick M. D'Aoust: supported laboratory experiments and revised the manuscript.

Shen Wan: conducted laboratory experiments and revised the manuscript.

Nada Hegazy: revised the manuscript.

Tram Nguyen: revised the manuscript.

Chandler Wong: revised the manuscript.

Felix Addo: revised the manuscript.

Emma Tomalty: revised the manuscript.

Tyson E. Graber: supervised the study, reviewed experimental design & data analysis, review, and revised the manuscript.

Robert Delatolla: supervised the study, reviewed experimental design & data analysis, funding acquisition, review and revised the manuscript.

Chapter 4: Manuscript 2

A version of the following manuscript in this chapter has been submitted to the Journal of Water Research in 2025: Kabir, M. P., Plaza-Diaz, J., D'Aoust, P.M., Mercier, E., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Renouf, E., Lawal, O. U., Goodridge, L., Graber, T. E., and Delatolla, R. *Passive sampling for genomic surveillance of SARS-CoV-2 in wastewater resource recovery facility: Insights for pandemic preparedness*

Md Pervez Kabir: developed research question, experimental design, sample collection, laboratory experiments, bioinformatics analysis, data analysis, prepared and revised the manuscript.

Julio Plaza-Diaz: supported laboratory experiments, bioinformatics analysis and revised the manuscript.

Élisabeth Mercier: supported laboratory experiments and revised the manuscript.

Shen Wan: conducted laboratory experiments and revised the manuscript.

Nada Hegazy: revised the manuscript.

Chandler Wong: revised the manuscript.

Felix Addo: revised the manuscript.

Elizabeth Renouf: reviewed statistical analysis and revised the manuscript.

Opeyemi U. Lawal: bioinformatics analysis and revised the manuscript.

Lawrence Goodridge: bioinformatics analysis and revised the manuscript.

Tyson E. Graber: supervised the study, reviewed experimental design & data analysis, review, and revised the manuscript.

Robert Delatolla: supervised the study, reviewed experimental design & data analysis, funding acquisition, review and revised the manuscript.

Chapter 5: Manuscript 3

A version of the following manuscript in this chapter has been submitted to the Journal of ACS ES&T Water in 2025: Kabir, M. P., Plaza-Diaz, J., Mercier, E., D'Aoust, P. M., Goodridge, L., Lawal, O.U., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Renouf, E., Graber, T. E., and Delatolla, R. *Optimization of primary sludge processing method for wastewater genomic surveillance of SARS-CoV-2.*

Md Pervez Kabir developed research question, experimental design, sample collection, laboratory experiments, bioinformatics analysis, data analysis, prepared and revised the manuscript.

Julio Plaza-Diaz assisted in performing laboratory experiments, data analysis, bioinformatics analysis, and revised the manuscript.

Élisabeth Mercier supported laboratory experiments and revised the manuscript.

Patrick M. D'Aoust supported laboratory experiments and revised the manuscript.

Lawrence Goodridge supported bioinformatics analysis and revised the manuscript.

Opeyemi U. Lawal supported bioinformatics analysis and revised the manuscript.

Shen Wan revised the manuscript.

Nada Hegazy revised the manuscript.

Tram Nguyen: revised the manuscript.

Chandler Wong: revised the manuscript.

Felix Addo: revised the manuscript.

Elizabeth Renouf: reviewed statistical analysis and revised the manuscript.

Tyson E. Graber: supervised the study, reviewed experimental design & data analysis, review, and revised the manuscript.

Robert Delatolla: supervised the study, reviewed experimental design & data analysis, funding acquisition, review and revised the manuscript.

Chapter 6: Manuscript 4

A version of the following manuscript in this chapter has been published in Eco-Environment and Health in 2025. Kabir, M. P., Mercier, E., Eid, W., Plaza-Diaz, J., D'Aoust, P. M., Landgraff, C., Goodridge, L., Lawal, O. U., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Thakali, O., Pisharody, L., Stephenson, S., E. Graber, T. E., and Delatolla, R. *Diagnostic performance of allele-specific RT-qPCR and genomic sequencing in wastewater-based surveillance of SARS-CoV-2*.

Md Pervez Kabir developed research question, experimental design, sample collection, laboratory experiments, bioinformatics analysis, data analysis, prepared and revised the manuscript.

Élisabeth Mercier supported laboratory experiments and revised the manuscript.

Walaa Eid supported laboratory experiments and revised the manuscript.

Julio Plaza-Diaz assisted in performing laboratory experiments, data analysis, bioinformatics analysis, and revised the manuscript.

Patrick M. D'Aoust supported laboratory experiments and revised the manuscript.

Chrystal Landgraff supported laboratory experiments and revised the manuscript.

Lawrence Goodridge supported bioinformatics analysis and revised the manuscript.

Opeyemi U. Lawal supported bioinformatics analysis and revised the manuscript.

Shen Wan conducted laboratory experiments and revised the manuscript.

Nada Hegazy revised the manuscript.

Tram Nguyen revised the manuscript.

Chandler Wong revised the manuscript.

Ocean Thakali revised the manuscript.

Lakshmi Pisharody revised the manuscript.

Sean Stephenson revised the manuscript.

Tyson E. Graber: supervised the study, reviewed experimental design & data analysis, review, and revised the manuscript.

Robert Delatolla: supervised the study, reviewed experimental design & data analysis, funding acquisition, review and revised the manuscript.

I am aware of the University of Ottawa Academic Regulations (C-7), and I certify that I have obtained written permission from each co-author to include the above materials in my dissertation. The above material describes work completed during my full-time registration as a graduate student at the University of Ottawa.

Acknowledgment

With deep gratitude, I first and foremost would like to thank the Almighty Allah, the most gracious and merciful creator, for granting me the strength to complete this PhD dissertation.

I would like to express my sincere gratitude to my supervisors, Dr. Robert Delatolla and Dr. Tyson E. Graber, for their unwavering supports. Your constant encouragement, guidance and availability has been pivotal throughout this journey. The moral and financial support you provided, especially during the challenging moments of my PhD, is something I will always treasure and remember for a lifetime. I gratefully acknowledge the financial support from Ontario's Ministry of Environment, Conservation, and Parks through the SARS-CoV-2 surveillance initiative, as well as the CIHR Applied Public Health Chair in Environment, Climate Change, and One Health, awarded to Dr. Robert Delatolla. I also extend my gratitude to the University of Ottawa for the admission scholarship, which provided financial assistance throughout my program. To my incredible research team, thank you for your unwavering support throughout this journey.

I am deeply thankful to my family for their endless love and support at every step. To my parents, I am forever grateful; words alone cannot express my appreciation. A special mention goes to my cousin, Abu Zafor, whose constant encouragement from my childhood has been a driving force in my life.

Finally, I would like to express my deepest gratitude to my beloved wife, Tasneem Bint Abdul Majid, for her unwavering support, patience, and encouragement throughout my PhD journey. Her love, sacrifices, and belief in me, especially during the most difficult times, have been my greatest source of strength. This achievement would not have been possible without her by my side, and for that, I am eternally grateful. Thank you for being my constant source of support and pillar of strength always.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

ALOD	Assay's Limit of Detection
ALOQ	Assay's Limit of Quantification
AMR	Antimicrobial Resistance
ANOVA	One-Way Analysis of Variance
ARGs	Antibiotic-Resistant Genes
AS-RT-qPCR	Allele-Specific Reverse Transcription Quantitative Polymerase Chain Reaction
BAM	Binary Alignment Map
CCDC	Chinese Center for Disease Control and Prevention
CDC	Centers for Disease Control and Prevention
cDNA	Complementary Deoxyribonucleic Acid
CNV	Copy Number Variation
COVID-19	Coronavirus disease 2019
Ct	Cycle Threshold
DNA	Deoxyribonucleic Acid
dsDNA	Double stranded DNA
ENMF	Electronegative Membrane Filtration
EPMF	Electropositive Membrane Filtration
FN	False Negative
FP	False Positive
GISAID	Global Initiative on Sharing All Influenza Data
INDELs	Insertions and Deletions
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
NGS	Next-Generation Sequencing
NML	National Microbiology Laboratory

PMMoV	Pepper Mild Mottle Virus
PEG	Polyethylene Glycol
PCR	Polymerase Chain Reaction
PLA	Polylactic Acid
QC	Quality Control
RBD	Receptor-Binding Domain
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
ROPEC	Robert O. Pickard Environmental Center
RT	Reverse Transcriptase
RT-ddPCR	Reverse Transcription Droplet Digital Polymerase Chain Reaction
RT-qPCR	Reverse Transcription Quantitative Polymerase Chain Reaction
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SAM	Sequence Alignment Map
SARS-CoV-1	Severe Acute Respiratory Syndrome Coronavirus 1
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SMF	Skimmed-Milk Flocculation
SNV	Single Nucleotide Variant
TN	True Negative
TP	True Positive
TS	Total Solids
TSS	Total Suspended Solids
UF	Ultrafiltration
USEPA	United States Environmental Protection Agency
VCF	Variant Call Format
VIRADEL	Viral Adsorption Elution

VOC	Variant of Concern
VOI	Variant of Interest
VUM	Variant under Monitoring
VS	Volatile Solids
VSS	Volatile Suspended Solids
WBE	Wastewater-Based Epidemiology
WEM	Wastewater Environmental Monitoring
WGS	Whole Genome Sequence
WWS	Wastewater Surveillance
WWGS	Wastewater Genomic Surveillance
WHO	World Health Organization
WWTP	Wastewater Treatment Plant
WWRF	Wastewater Resource Recovery Facility

1 Chapter 1: Introduction

1.1 Background

An unknown disease with flu symptoms (i.e., high fever, cough, chest discomfort, dyspnea, and bilateral lung infiltration) first appeared in Wuhan, China in December 2019, concerning the Wuhan municipal government (Chen et al., 2020; Wang et al., 2020). The Wuhan Municipal Health Commission immediately informed the Chinese Center for Disease Control and Prevention (CCDC) and the World Health Organization (WHO), and WHO dispatched a medical team to Wuhan (WHO, 2020a; Wu & McGoogan, 2020). The medical team collected swab samples from infected patients and following whole genome sequencing (WGS), they identified betacoronavirus as the causative agent of the disease (Wu et al., 2020; Zhou et al., 2020). The virus is an enveloped, positive-sense, single-stranded RNA virus genome and the viral genome is roughly 29.9kb long (Kim et al., 2020). The virus is highly transmissible and can spread through respiratory droplets and aerosols from infected individuals (Zhu et al., 2020). The Chinese government implemented various mitigation measures throughout the Wuhan municipality to prevent the spread of the virus, but the disease quickly spread throughout China within a few weeks (Hu et al., 2021). WHO declared the new disease as a public health emergency of international concern on January 30, 2020, (Alhama et al., 2021). The International Committee on Taxonomy subsequently designated the new betacoronavirus as SARS-CoV-2 due to its genetic resemblance with SARS-CoV-1 on February 11, 2020, and named the disease as COVID-19 (Gorbalenya et al., 2020). The COVID-19 reached its initial outbreak peak in China in February 2020 (Fisher & Heymann, 2020; Lai et al., 2020) while international detection began in late February 2020, and WHO declared COVID-19 as a global pandemic on March 11, 2020, (WHO, 2020b). As the pandemic progressed, a total of 776 million confirmed cases and 7.05 million deaths were reported globally (WHO, 2024) and throughout the pandemic, clinical surveillance has been used as common standard for

monitoring and spread of the COVID-19 disease. Clinical surveillance involves collecting swab samples from symptomatic infected individuals, followed by RT-PCR testing to confirm the infections (Chau et al., 2020). However, clinical surveillance overlooks asymptomatic carriers as well as individuals without access to clinical testing, hesitancy about clinical testing, trust in the healthcare system as well as stigma and fear of discrimination leading to a gap in understanding the full extent of population-level infections.

The causative agent of COVID-19 disease is a respiratory virus whose genetic material has been detected in various bodily fluids and excretions, including sputum, urine, and feces of the asymptomatic, symptomatic, and presymptomatic patients (Fei Xiao et al., 2020; Jones et al., 2020). These findings led to hypothesizing that monitoring municipal wastewater could serve as a valuable tool for providing aggregated, population-level information on SARS-CoV-2 infection dynamics. The hypothesis was first validated by a group of Dutch scientists in February 2020 who confirmed the presence of SARS-CoV-2 RNA in wastewater and established a correlation between RNA signals in wastewater and reported clinical cases (Medema et al., 2020). The presence of SARS-CoV-2 RNA in wastewater also confirmed in Australia (Ahmed et al., 2020), Asia (Haramoto et al., 2020), North America (Sherchan et al., 2020), South America (Fongaro et al., 2021), and Africa (Johnson et al., 2021). The monitoring of viral RNA in wastewater to track the disease prevalence in populations is commonly referred to as WEM. The WEM was previously limited to illicit drugs, gastrointestinal infections caused by enteric viruses, but the COVID-19 pandemic saw the most widespread adoption of WEM for respiratory illness such as SARS-CoV-2, influenza, RSV (D' Aoust et al, 2021; Mercier et al., 2022; Mercier et al., 2023). Over the course of the COVID-19 pandemic, WEM has been extended to 4648 sites in 72 countries around the world (Naughton et al., 2024). This enormous number of research and surveillance sites include monitoring of SARS-CoV-2 RNA in polluted surface waters (Guerrero-Latorre et al., 2020; Haramoto et al., 2020; Rimoldi et al., 2020),

analysis of wastewaters to access the occurrence, trends and correlation with epidemiological metrics (Chavarria-Miró et al., 2021; D'Aoust et al., 2021; Gerrity et al., 2021).

The emergence of SARS-CoV-2 WEM has prompted the development of several wastewater sampling methods including time or flow-proportional 24-hours composite and grab sampling. Composite samples are often collected using an autosampler over the course of 24 hours while grab sample is taken at a certain time of the day, usually in the morning. However, implementation of an autosampler for composite sampling is labor intensive, expensive, and challenging at geo-specific locations, while grab sample is prone to intraday variation, making it less representative to population-level disease prevalence (Augusto et al., 2022; Harris-Lovett et al., 2021; Kitajima et al., 2020). As an alternative to autosampler and grab samples, passive sampling has been used as a promising tool to capture wide ranges of pollutants since 1987. The Moore swab is the most primitive form of passive sampler, used to trace *Salmonella Paratyphi B* from effluent sewage (Sears et al., 1984). Later, it was adopted to capture microorganisms such as fecal-borne pathogens, poliovirus, and human norovirus in wastewater (Hayes & Gagnon, 2024; Bivins et al., 2022). During the CoV-19 pandemic, passive sampling has been adopted as a practical, affordable, and reliable sampling method for SARS-CoV-2 WEM (Mejías-Molina et al., 2023; Habtewold et al., 2022; Kevill et al., 2022). In passive sampling, sampling mediums (i.e., medical gauzes, cheesecloth, cellulose sponges, or electronegative membrane filters) are packed within styrene-like plastic devices and deployed in sewer networks to sorb wastewater solids over time (Bivins et al., 2022). Wastewater is comprised of a variety of suspended materials and biosolids that are likely to compete with SARS-CoV-2 viral particles for sorption in the medium. Therefore, characteristics of the captured solids on passive samplers may affect SARS-CoV-2 RNA measurements, but this association is not yet well comprehended. Most of the earlier research (Breulmann et al., 2023; Cha et al., 2023; Mejías-Molina et al., 2023; Habtewold et al., 2022) homogenized collected

wastewater solids into deionized water or soluble buffer, measured SARS-CoV-2 RNA, along with other disease targets using the solute, and reported as copies/mL or copies/L. The SARS-CoV-2 RNA partitions favorably to the solids fraction (Espinosa et al., 2022), and therefore, directly analyzing solids from passive samplers could yield higher concentrations compared to the solute. Besides, the SARS-CoV-2 RNA is found in higher concentrations in settled solids (i.e., primary sludge) compared to influent wastewater (Kim et al., 2022; D'Aoust et al., 2021). Quantifying SARS-CoV-2 RNA concentrations in wastewater solids mitigates the influence of variable hydraulic loading and dilution effects, thereby providing a more accurate representation of community-level viral shedding. Furthermore, mass-based normalization facilitates integration with endogenous population biomarkers (e.g., PMMoV, crAssphage, or chemical tracers), enhances temporal and spatial comparability across diverse sampling strategies, and supports cross-matrix analysis between solids and influent (D'Aoust et al., 2021). However, it has yet to be determined whether the SARS-CoV-2 RNA concentrations in solids collected using passive samplers can be effectively compared to composite wastewater and primary sludge samples.

The WEM using composite and passive sampling offers critical insights into the SARS-CoV-2 infection landscape within the populations, but with the progression of the COVID-19 pandemic, SARS-CoV-2 continues to evolve, resulting in the emergence of VOCs (González-Candelas et al., 2021; Safari & Elahi, 2022). These emerging VOCs can't be tracked by quantifying SARS-CoV-2 RNA concentrations in wastewaters. To track the VOCs and their associated lineages, another approach known as WWGS has been adopted to support clinical genomic surveillance of SARS-CoV-2. WWGS generally entails the collection of wastewater samples, followed by viral RNA extraction and sequencing to identify and characterize the distinct strains or variants of SARS-CoV-2 circulating within the populations (Ai et al., 2021; Li et al., 2022; Vo et al., 2022). This approach allows for the early detection of prevalent SARS-

CoV-2 variants on both regional and national scales (Amman et al., 2022; Crits-Christoph et al., 2021; Izquierdo-Lara et al., 2021). WWGS also demonstrated the detection of circulating and cryptic (not detected through clinical surveillance) mutations and lineages in the population level (Jahn et al., 2022; Karthikeyan et al., 2022; Smyth et al., 2022). WWGS has predominantly relied on composite wastewater samples. Although passive sampling has demonstrated success in WEM, it remains underutilized in WWGS. To date, only five studies have focused on SARS-CoV-2 genomic analyses using passive sampling. Among these, (Corchis-Scott et al., 2021; Mangwana et al., 2022) employed AS-RT-qPCR for the detection of lineage-specific mutations, while (Alamin et al., 2024; Overton et al., 2024; Cha et al., 2023) sequenced the SARS-CoV-2 genome for comprehensive variant identification. However, Cha et al., (2023) was unable to recover near-complete genome (≥ 90%) from any of the wastewater samples collected by Moore swabs, whereas Alamin et al. (2024) reported successful recovery of SARS-CoV-2 genome from the majority of wastewater samples collected through Tampon swabs. Despite the potential of passive sampling in WWGS, their applications were primarily limited to tracking variants in congregate settings with low sewer flow rates such as student dorms at university campuses and wastewater from airport. However, the implementation of passive sampling in larger, high-flow wastewater environments such as WWRFs for genomic analysis remains unexplored. In these settings, passive sampling may face challenges, as the high-flow, turbulent conditions could reduce the efficiency of capturing and retaining viral particles on the sampling medium. Besides, there are concerns about the integrity of viral RNA in such conditions, as high-flow wastewater environments may degrade RNA more rapidly, potentially impacting the quality of genomic data retrieved through passive sampling. Moreover, earlier studies (Fontenele et al., 2021; Nemudryi et al., 2020) suggested that recovering SARS-CoV-2 genomes requires at least 2.8×10^5 viral copies per liter in wastewater. However, later findings indicated that genome recovery depends heavily on

sample processing, concentration and RNA extraction methods (Chen et al., 2024; Feng et al., 2023). Thus, it remains uncertain whether passive sampling in WWRFs and current processing methods can be effectively utilized for genomic analysis of SARS-CoV-2. Additionally, it is unclear whether SARS-CoV-2 lineage profiles obtained from passive sampling would be comparable to those from conventional wastewater auto sampling or clinical surveillance strategies. This underscores the need for further research to evaluate the efficacy of passive sampling in high-flow wastewater contexts in order to support the continued and broader application of WWGS for monitoring respiratory pathogens and emerging disease targets.

In addition to knowledge gaps in wastewater sampling methods, the challenge of selecting appropriate sample types for effective genomic analyses remains unresolved. One of the major challenges in WWGS using influent wastewater is the recovery of near-complete SARS-CoV-2 genomes, largely due to the fragmentation nature and limited amount of the RNA target in wastewater (Bar-Or et al., 2021; Fontenele et al., 2021; Patel et al., 2021). Besides, the detection and tracking of low-frequency variants of SARS-CoV-2 in wastewater pose significant challenges. To resolve this issue, clarified primary sludge is theoretically a more appropriate location for sample collection compared to influent wastewater for WWGS, as SARS-CoV-2 RNA is found in higher concentrations in primary sludge as compared to influent wastewater (Balboa et al., 2021; D'Aoust et al., 2021; Kim et al., 2022). However, in light of the higher concentrations of SARS-CoV-2 genomic targets in primary sludge, sequencing efforts are more complex and challenging due to the anticipated presence of enriched PCR inhibitors in sludge matrices (Peccia et al., 2020). Consequently, limited efforts have been documented in scientific literature to sequence the SARS-CoV-2 genome directly from primary sludge. Early in the pandemic, Lin et al., (2021) attempted to sequence the whole genome of SARS-CoV-2 from primary sludge but failed to produce library for recovering near-complete (≥ 90%) genome. They expressed the reasons to be the presence of PCR inhibitors in primary

sludge which inhibits library preparation, while other impurities caused undesired amplification resulting in poor quality sequencing data along with incomplete genomic coverage. They also assumed that the SARS-CoV-2 genome is less intact in primary sludge, which led to significant RNA degradation during direct sludge processing. Although, later Baaijens et al., (2022) revealed the recovery of near complete genome of SARS-CoV-2 from the selected primary sludge samples (< 30% of those samples) where the cycling threshold (Ct) value was less than 31, and the genomic recovery decreased with the Ct value increased. However, a critical comparison between these two studies suggests that differences in sludge concentration and RNA extraction methods may be the most influential factor affecting the successful recovery of near-complete genomes from primary sludge samples. This highlighted the need for optimized processing methods to recover near-complete genomes from primary sludge, and resolve challenges associated with WWGS of SARS-CoV-2 using primary sludge samples.

The application of WWGS for SARS-CoV-2 in low-resource settings is limited due to the high costs associated with sequencing technologies as well as the requirement for specialized technical expertise. To overcome this challenge, researchers have adopted AS-RT-qPCR to monitor SARS-CoV-2 variants and mutations in wastewater, focusing on specific small segments of the viral genome (Graber et al., 2021; Yaniv et al., 2021). The AS-RT-qPCR entails two parallel PCR reactions in which a wild-type or universal gene region and a mutant-type gene region are amplified using distinct primers and probes (Kalendar et al., 2022) to characterize the proportionality of the mutation and/or associated variant (Graber et al., 2021) This method is rapid and gives interpretable results in a short period, but can't delineate all the mutations present in a sample (Ellmen et al., 2021). However, two studies have shown that AS-RT-qPCR and AS-RT-dPCR are more sensitive compared to sequencing-based methods for variant detection and quantification in wastewater (Ahmed, et al., 2022; Lou et al., 2022).

Previous studies (Izquierdo-Lara et al., 2021; Wurtzer et al., 2022) also revealed that single-allele frequency estimations obtained by AS-RT-dPCR were strongly correlated with single-allele or haplotype frequency estimates derived from sequencing. However, the diagnostic performance (i.e., accuracy) of AS-RT-qPCR and sequencing methods with respect to the perceived "ground truth" of clinical surveillance remains unclear. The diagnostic performance of an assay usually expressed as Youden's index, which reports sensitivity and specificity in a single metric. Assigning this index to test the presence or absence of SARS-CoV-2 variants in wastewater samples could help assay development teams in choosing the best allele or set of alleles to estimate variant frequencies.

1.2 Research Objectives

The primary aim of this research is to evaluate different sampling methods, sample processing strategies such as concentration and RNA extraction methods and genomic analyses for comprehensive SARS-CoV-2 RNA monitoring in wastewaters. The specific research objectives are as follows:

- The first objective is to evaluate the SARS-CoV-2 RNA concentrations in wastewater solids collected using autosampler, passive samplers and primary sludge samples.
- The second objective is to evaluate the feasibility of passive sampling for WWGS in high-flow wastewater settings.
- The third objective is to develop a primary sludge concentration and RNA extraction method for whole genome sequencing of SARS-CoV-2.
- The fourth and final objective is to compare the diagnostic performance of AS-RT-qPCR and sequencing-based methods to determine their application in WWGS of SARS-CoV-2.

1.3 Thesis Organization

The dissertation is structured as a manuscript-based thesis as specified by the school of Graduate and Postdoctoral Studies at the University of Ottawa. The organization of this thesis is as follows:

Chapter 1 offers a brief introduction to the study, outlining the overall and specific research objectives. It also provides a detailed description of the thesis structure, setting the stage for the subsequent chapters.

Chapter 2 provides a comprehensive overview of the global efforts in WEM of SARS-CoV-2, detailing the various sampling methods, sample storage conditions, sample processing techniques, quantification methods, sequencing strategies, and the tools and pipelines used for variant detection.

Chapter 3 is a research article entitled "*Wastewater solids drive comparability of sampling methods for SARS-CoV-2 wastewater and environmental surveillance*". This manuscript has been published in the journal of Environmental Chemical Engineering in 2025. This study evaluates the SARS-CoV-2 RNA concentrations in wastewater solids collected by autosampler, passive samplers, and primary sludge. The study also explores the influences of wastewater solids content on SARS-CoV-2 RNA measurement using a linear mixed-effects model.

Chapter 4 is a research article entitled "*Passive sampling for genomic surveillance of SARS-CoV-2 in wastewater resource recovery facility: Insights for pandemic preparedness".* This manuscript has been submitted for publication to the journal of Water Research in 2025. This study evaluates the applicability of passive sampling for WWGS in high-flow wastewater settings. In particular, the study focuses on the genomic recovery of SARS-CoV-2 from passive

samplers and compares the SNVs profile and lineages prevalence with conventional auto sampling as well as available clinical surveillance data.

Chapter 5 is a research article entitled "*Optimization of primary sludge processing method for wastewater genomic surveillance of SARS-CoV-2"*. This manuscript has been submitted for publication to the journal of ACS ES&T Water in 2025. This study focuses on developing of a primary sludge concentration and RNA extraction method for SARS-CoV-2 genome sequencing that yields comparable or better results to conventional WWGS. This study also compares SNVs profiles detected in influent wastewater and primary sludge with Canadian and global clinical sequences to identify rare and cryptic SARS-CoV-2 SNVs.

Chapter 6 is a research article entitled "*Diagnostic performance of allele-specific RT-qPCR and genomic sequencing in wastewater-based surveillance of SARS-CoV-2"*. This manuscript has been published in the journal of Eco-Environment & Health in 2025. This study compares single-allele frequency estimation using AS-RT-qPCR, to single-allele or haplotype frequency estimations derived from amplicon-based sequencing to estimate variant prevalence in municipal wastewater samples. Finally, this study assesses the diagnostic performance of each method using Youden index with respect to the perceived clinical surveillance.

1.4 References

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

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

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

Alamin, M., Oladipo, P., Hartrick, J., Islam, N., Bahmani, A., Turner, C. L., Shuster, W., & Ram, J. L. (2024). Improved passive sampling methods for wastewater to enable more sensitive detection of SARS-CoV-2 and its variants. *Science of the Total Environment*, *950*. https://doi.org/10.1016/j.scitotenv.2024.175044

Alhama, J., Maestre, J. P., Martín, M. Á., & Michán, C. (2021). Monitoring COVID-19 through SARS-CoV-2 quantification in wastewater: progress, challenges and prospects. *Microbial Biotechnology*, *0*, 1–10. https://doi.org/10.1111/1751-7915.13989

Amman, F., Markt, R., Endler, L., Hupfauf, S., Agerer, B., Schedl, A., Richter, L., Zechmeister, M., Bicher, M., Heiler, G., Triska, P., Thornton, M., Penz, T., Senekowitsch, M., Laine, J., Keszei, Z., Klimek, P., Nägele, F., Mayr, M., … Bergthaler, A. (2022). Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nature Biotechnology*, *40*(12), 1814–1822. https://doi.org/10.1038/s41587-022-01387-y

Augusto, M. R., Claro, I. C. M., Siqueira, A. K., Sousa, G. S., Caldereiro, C. R., Duran, A. F. A., de Miranda, T. B., Bomediano Camillo, L. de M., Cabral, A. D., & de Freitas Bueno, R. (2022). Sampling strategies for wastewater surveillance: Evaluating the variability of

SARS-COV-2 RNA concentration in composite and grab samples. *Journal of Environmental Chemical Engineering*, *10*(3). https://doi.org/10.1016/j.jece.2022.107478

Baaijens, J. A., Zulli, A., Ott, I. M., Nika, I., van der Lugt, M. J., Petrone, M. E., Alpert, T., Fauver, J. R., Kalinich, C. C., Vogels, C. B. F., Breban, M. I., Duvallet, C., McElroy, K. A., Ghaeli, N., Imakaev, M., Mckenzie-Bennett, M. F., Robison, K., Plocik, A., Schilling, R., … Baym, M. (2022). Lineage abundance estimation for SARS-CoV-2 in wastewater using transcriptome quantification techniques. *Genome Biology*, *23*(1). https://doi.org/10.1186/s13059-022-02805-9

Balboa, S., Mauricio-Iglesias, M., Rodriguez, S., Martínez-Lamas, L., Vasallo, F. J., Regueiro, B., & Lema, J. M. (2021). The fate of SARS-COV-2 in WWTPS points out the sludge line as a suitable spot for detection of COVID-19. *Science of the Total Environment*, *772*, 145268. https://doi.org/10.1016/j.scitotenv.2021.145268

Bar-Or, I., Weil, M., Indenbaum, V., Bucris, E., Bar-Ilan, D., Elul, M., Levi, N., Aguvaev, I., Cohen, Z., Shirazi, R., Erster, O., Sela-Brown, A., Sofer, D., Mor, O., Mendelson, E., & Zuckerman, N. S. (2021). Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel. *Science of the Total Environment*, *789*, 1–14. https://doi.org/10.1016/j.scitotenv.2021.148002

Bivins, A., Kaya, D., Ahmed, W., Brown, J., Butler, C., Greaves, J., Leal, R., Maas, K., Rao, G., Sherchan, S., Sills, D., Sinclair, R., Wheeler, R. T., & Mansfeldt, C. (2022). Passive sampling to scale wastewater surveillance of infectious disease: Lessons learned from COVID-19. *Science of the Total Environment*, *835*. https://doi.org/10.1016/j.scitotenv.2022.155347

Bowes, D. A., Driver, E. M., & Halden, R. U. (2022). A framework for wastewater sample collection from a sewage cleanout to inform building-scale wastewater-based

epidemiology studies. *Science of the Total Environment*, *836*. https://doi.org/10.1016/j.scitotenv.2022.155576

Breulmann, M., Kallies, R., Bernhard, K., Gasch, A., Müller, R. A., Harms, H., Chatzinotas, A., & van Afferden, M. (2023). A long-term passive sampling approach for wastewater-based monitoring of SARS-CoV-2 in Leipzig, Germany. *Science of the Total Environment*, *887*. https://doi.org/10.1016/j.scitotenv.2023.164143

Cha, G., Graham, K. E., Zhu, K. J., Rao, G., Lindner, B. G., Kocaman, K., Woo, S., D'amico, I., Bingham, L. R., Fischer, J. M., Flores, C. I., Spencer, J. W., Yathiraj, P., Chung, H., Biliya, S., Djeddar, N., Burton, L. J., Mascuch, S. J., Brown, J., … Konstantinidis, K. T. (2023). Parallel deployment of passive and composite samplers for surveillance and variant profiling of SARS-CoV-2 in sewage. *Science of the Total Environment*, *866*. https://doi.org/10.1016/j.scitotenv.2022.161101

Chau, C. H., Strope, J. D., & Figg, W. D. (2020). COVID-19 Clinical Diagnostics and Testing Technology. *Pharmacotherapy*, *40*(8), 857–868. https://doi.org/10.1002/phar.2439

Chavarria-Miró, G., Anfruns-Estrada, E., Martínez-Velázquez, A., Vázquez-Portero, M., Guix, S., Paraira, M., Galofré, B., Sánchez, G., Pintó, R. M., & Bosch, A. (2021). Time Evolution of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Wastewater during the First Pandemic Wave of COVID-19 in the Metropolitan Area of Barcelona, Spain. *Applied and Environmental Microbiology*, *87*(7), 1–9. https://doi.org/10.1128/AEM.02750-20

Chen, L., Chen, A., Zhang, X. D., Robles, M. T. S., Han, H. S., Xiao, Y., Xiao, G., Pipas, J. M., & Weitz, D. A. (2024). Targeted whole-genome recovery of single viral species in a complex environmental sample. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(31). https://doi.org/10.1073/pnas.2404727121

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, *395*(10223), 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

Colosi, L. M., Barry, K. E., Kotay, S. M., Porter, M. D., Poulter, M. D., Ratliff, C., Simmons, W., Steinberg, L. I., Derek Wilson, D., Morse, R., Zmick, P., Mathers, A. J., Colosi, C. L., & Harold Drake, E. L. (2021). *Development of Wastewater Pooled Surveillance of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) from Congregate Living Settings public and environmental health microbiology* (Vol. 87). https://journals.asm.org/journal/aem

Crits-Christoph, A., Kantor, R. S., Olm, M. R., Whitney, O. N., Al-Shayeb, B., Lou, Y. C., Flamholz, A., Kennedy, L. C., Greenwald, H., Hinkle, A., Hetzel, J., Spitzer, S., Koble, J., Tan, A., Hyde, F., Schroth, G., Kuersten, S., Banfield, J. F., & Nelson, K. L. (2021). Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *MBio*, *12*(1), 1–9. https://doi.org/10.1128/mBio.02703-20

D'Aoust, P. M., Graber, T. E., Mercier, E., Montpetit, D., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Servos, M. R., Srikanthan, N., MacKenzie, M., Figeys, D., Manuel, D., Jüni, P., MacKenzie, A. E., & Delatolla, R. (2021). Catching a resurgence: Increase in SARS-CoV-2 viral RNA identified in wastewater 48 h before COVID-19 clinical tests and 96 h before hospitalizations. *Science of the Total Environment*, *770*. https://doi.org/10.1016/j.scitotenv.2021.145319

Davó, L., Seguí, R., Botija, P., Beltrán, M. J., Albert, E., Torres, I., López-Fernández, P. Á., Ortí, R., Maestre, J. F., Sánchez, G., & Navarro, D. (2021). Early detection of SARS-

CoV-2 infection cases or outbreaks at nursing homes by targeted wastewater tracking. In *Clinical Microbiology and Infection* (Vol. 27, Issue 7, pp. 1061–1063). https://doi.org/10.1016/j.cmi.2021.02.003

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

Espinosa, M. F., Verbyla, M. E., Vassalle, L., Leal, C., Leroy-Freitas, D., Machado, E., Fernandes, L., Rosa-Machado, A. T., Calábria, J., Chernicharo, C., & Mota Filho, C. R. (2022). Reduction and liquid-solid partitioning of SARS-CoV-2 and adenovirus throughout the different stages of a pilot-scale wastewater treatment plant. *Water Research*, *212*. https://doi.org/10.1016/j.watres.2022.118069

Fei Xiao, Jing Sun, Yonghao Xu, Fang Li, Xiaofang Huang, Heying Li, Jingxian Zhao, Jicheng Huang, & Jincun Zhao. (2020). Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerging Infectious Diseases*, *26*(8), 1920–1922.

Feng, S., Owens, S. M., Shrestha, A., Poretsky, R., Hartmann, E. M., & Wells, G. (2023). Intensity of sample processing methods impacts wastewater SARS-CoV-2 whole genome amplicon sequencing outcomes. *Science of the Total Environment*, *876*. https://doi.org/10.1016/j.scitotenv.2023.162572

Fisher, D., & Heymann, D. (2020). The novel coronavirus outbreak causing COVID-19. *BMC Medicine*, *18*(1), 18–20. https://doi.org/10.1186/s12916-020-01533-w

Fongaro, G., Stoco, P. H., Souza, D. S. M., Grisard, E. C., Magri, M. E., Rogovski, P., Schörner, M. A., Barazzetti, F. H., Christoff, A. P., de Oliveira, L. F. V., Bazzo, M. L., Wagner, G., Hernández, M., & Rodríguez-Lázaro, D. (2021). The presence of SARS-

CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. *Science of the Total Environment*, *778*, 1–4. https://doi.org/10.1016/j.scitotenv.2021.146198

Fontenele, R. S., Kraberger, S., Hadfield, J., Driver, E. M., Bowes, D., Holland, L. R. A., Faleye, T. O. C., Adhikari, S., Kumar, R., Inchausti, R., Holmes, W. K., Deitrick, S., Brown, P., Duty, D., Smith, T., Bhatnagar, A., Yeager, R. A., Holm, R. H., von Reitzenstein, N. H., … Varsani, A. (2021). High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *Water Research*, *205*. https://doi.org/10.1016/j.watres.2021.117710

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

González-Candelas, F., Shaw, M. A., Phan, T., Kulkarni-Kale, U., Paraskevis, D., Luciani, F., Kimura, H., & Sironi, M. (2021). One year into the pandemic: Short-term evolution of SARS-CoV-2 and emergence of new lineages. In *Infection, Genetics and Evolution* (Vol. 92). https://doi.org/10.1016/j.meegid.2021.104869

Gorbalenya, A. E., Baker, S. C., Baric, R. S., de Groot, R. J., Drosten, C., Gulyaeva, A. A., Haagmans, B. L., Lauber, C., Leontovich, A. M., Neuman, B. W., Penzar, D., Perlman, S., Poon, L. L. M., Samborskiy, D. V., Sidorov, I. A., Sola, I., & Ziebuhr, J. (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, *5*(4), 536–544. https://doi.org/10.1038/s41564-020-0695-z

Graber, T. E., Mercier, É., Bhatnagar, K., Fuzzen, M., D'Aoust, P. M., Hoang, H. D., Tian, X., Towhid, S. T., Plaza-Diaz, J., Eid, W., Alain, T., Butler, A., Goodridge, L., Servos, M., &

Delatolla, R. (2021). Near real-time determination of B.1.1.7 in proportion to total SARS-CoV-2 viral load in wastewater using an allele-specific primer extension PCR strategy. *Water Research*, *205*, 1–9. https://doi.org/10.1016/j.watres.2021.117681

Guerrero-Latorre, L., Ballesteros, I., Villacrés-Granda, I., Granda, M. G., Freire-Paspuel, B., & Ríos-Touma, B. (2020). SARS-CoV-2 in river water: Implications in low sanitation countries. *Science of the Total Environment*, *743*. https://doi.org/10.1016/j.scitotenv.2020.140832

Habtewold, J., McCarthy, D., McBean, E., Law, I., Goodridge, L., Habash, M., & Murphy, H. M. (2022). Passive sampling, a practical method for wastewater-based surveillance of SARS-CoV-2. *Environmental Research*, *206*. https://doi.org/10.1016/j.envres.2021

Haramoto, E., Malla, B., Thakali, O., & Kitajima, M. (2020). First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. *Science of the Total Environment*, *737*, 140405. https://doi.org/10.1016/j.scitotenv.2020.140405

Harris-Lovett, S., Nelson, K. L., Beamer, P., Bischel, H. N., Bivins, A., Bruder, A., Butler, C., Camenisch, T. D., De Long, S. K., Karthikeyan, S., Larsen, D. A., Meierdiercks, K., Mouser, P. J., Pagsuyoin, S., Prasek, S. M., Radniecki, T. S., Ram, J. L., Keith Roper, D., Safford, H., … Korfmacher, K. S. (2021). Wastewater surveillance for sars-cov-2 on college campuses: Initial efforts, lessons learned and research needs. *International Journal of Environmental Research and Public Health*, *18*(9). https://doi.org/10.3390/ijerph18094455

Hart, O. E., & Halden, R. U. (2020). Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Science of the Total Environment*, *730*, 138875. https://doi.org/10.1016/j.scitotenv.2020.138875

Hayes, K. E. & Gagnon, A. G. (2024). From capture to detection: A critical review of passive sampling techniques for pathogen surveillance in water and wastewater, *Water Research*, Volume 261,122024. https://doi.org/10.1016/j.watres.2024.122024.

Hu, B., Guo, H., Zhou, P., & Shi, Z. L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, *19*(3), 141–154. https://doi.org/10.1038/s41579-020-00459-7

Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Oude Munnink, B. B., Schapendonk, C. M. E., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F. M., Lycett, S., Medema, G., Koopmans, M. P. G., & De Graaf, M. (2021a). Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the netherlands and belgium. *Emerging Infectious Diseases*, *27*(5), 1405–1415. https://doi.org/10.3201/eid2705.204410

Jahn, K., Dreifuss, D., Topolsky, I., Kull, A., Ganesanandamoorthy, P., Fernandez-Cassi, X., Bänziger, C., Devaux, A. J., Stachler, E., Caduff, L., Cariti, F., Corzón, A. T., Fuhrmann, L., Chen, C., Jablonski, K. P., Nadeau, S., Feldkamp, M., Beisel, C., Aquino, C., … Beerenwinkel, N. (2022). Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. *Nature Microbiology*, *7*(8), 1151–1160. https://doi.org/10.1038/s41564-022-01185-x

Johnson, R., Muller, C. J. F., Ghoor, S., Louw, J., Archer, E., Surujlal-Naicker, S., Berkowitz, N., Volschenk, M., Bröcker, L. H. L., Wolfaardt, G., van der Walt, M., Mutshembele, A. M., Malema, S., Gelderblom, H. C., Mdhluli, M., Gray, G., Mathee, A., & Street, R. A. (2021). Qualitative and quantitative detection of SARS-CoV-2 RNA in untreated wastewater in Western Cape Province, South Africa. *South African Medical Journal*, *111*(3), 198–202. https://doi.org/10.7196/SAMJ.2021.V111I3.15154

Jones, D. L., Baluja, M. Q., Graham, D. W., Corbishley, A., McDonald, J. E., Malham, S. K., Hillary, L. S., Connor, T. R., Gaze, W. H., Moura, I. B., Wilcox, M. H., & Farkas, K. (2020). Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Science of the Total Environment, 749*(July), 141364. https://doi.org/10.1016/j.scitotenv.2020.141364

Kalendar, R., Baidyussen, A., Serikbay, D., Zotova, L., Khassanova, G., Kuzbakova, M., Jatayev, S., Hu, Y. G., Schramm, C., Anderson, P. A., Jenkins, C. L. D., Soole, K. L., & Shavrukov, Y. (2022). Modified "Allele-Specific qPCR" Method for SNP Genotyping Based on FRET. *Frontiers in Plant Science*, *12*, 1–16. https://doi.org/10.3389/fpls.2021.747886

Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, *609*(7925), 101–108. https://doi.org/10.1038/s41586-022-05049-6

Kevill, J. L., Lambert-Slosarska, K., Pellett, C., Woodhall, N., Richardson-O'Neill, I., Pântea, I., Alex-Sanders, N., Farkas, K., & Jones, D. L. (2022). Assessment of two types of passive sampler for the efficient recovery of SARS-CoV-2 and other viruses from wastewater. *Science of the Total Environment*, *838*. https://doi.org/10.1016/j.scitotenv.2022.156580

Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. *Cell*, *181*(4), 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011

Kim, S., Kennedy, L. C., Wolfe, M. K., Criddle, C. S., Duong, D. H., Topol, A., White, B. J., Kantor, R. S., Nelson, K. L., Steele, J. A., Langlois, K., Griffith, J. F., Zimmer-Faust, A. G., McLellan, S. L., Schussman, M. K., Ammerman, M., Wigginton, K. R., Bakker, K. M., & Boehm, A. B. (2022). SARS-CoV-2 RNA is enriched by orders of magnitude in primary settled solids relative to liquid wastewater at publicly owned treatment works. *Environmental Science: Water Research and Technology*, *8*(4), 757–770. https://doi.org/10.1039/d1ew00826a

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C. P., Hamilton, K. A., Haramoto, E., & Rose, J. B. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of the Total Environment*, 739. https://doi.org/10.1016/j.scitotenv.2020.139076

Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J., & Hsueh, P. R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, *55*(3), 105924. https://doi.org/10.1016/j.ijantimicag.2020.105924

Li, L., Uppal, T., Hartley, P. D., Gorzalski, A., Pandori, M., Picker, M. A., Verma, S. C., & Pagilla, K. (2022). Detecting SARS-CoV-2 variants in wastewater and their correlation with circulating variants in the communities. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20219-2

Lin, X., Glier, M., Kuchinski, K., Ross-Van Mierlo, T., McVea, D., Tyson, J. R., Prystajecky, N., & Ziels, R. M. (2021). Assessing Multiplex Tiling PCR Sequencing Approaches for Detecting Genomic Variants of SARS-CoV-2 in Municipal Wastewater. *MSystems*, *6*(5). https://doi.org/10.1128/msystems.01068-21

Lou, E. G., Sapoval, N., McCall, C., Bauhs, L., Carlson-Stadler, R., Kalvapalle, P., Lai, Y., Palmer, K., Penn, R., Rich, W., Wolken, M., Brown, P., Ensor, K. B., Hopkins, L., Treangen, T. J., & Stadler, L. B. (2022). Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 mutation monitoring in wastewater. *Science of the Total Environment, 833*. https://doi.org/10.1016/j.scitotenv.2022.155059

Mangwana, N., Archer, E., Muller, C. J. F., Preiser, W., Wolfaardt, G., Kasprzyk-Hordern, B., Carstens, A., Brocker, L., Webster, C., McCarthy, D., Street, R., Mathee, A., Louw, J., Mdhluli, M., & Johnson, R. (2022). Sewage surveillance of SARS-CoV-2 at student campus residences in the Western Cape, South Africa. *Science of the Total Environment, 851*. https://doi.org/10.1016/j.scitotenv.2022.158028

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 and Technology Letters*, *7*(7), 511–516. https://doi.org/10.1021/acs.estlett.0c00357

Mejías-Molina, C., Pico-Tomàs, A., Beltran-Rubinat, A., Martínez-Puchol, S., Corominas, L., Rusiñol, M., & Bofill-Mas, S. (2023). Effectiveness of passive sampling for the detection and genetic characterization of human viruses in wastewater. *Environmental Science: Water Research and Technology*. https://doi.org/10.1039/d2ew00867j

Mercier, E., D'Aoust, P. M., Thakali, O., Hegazy, N., Jia, J. J., Zhang, Z., Eid, W., Plaza-Diaz, J., Kabir, M. P., Fang, W., Cowan, A., Stephenson, S. E., Pisharody, L., MacKenzie, A. E., Graber, T. E., Wan, S., & Delatolla, R. (2022). Municipal and neighbourhood level wastewater surveillance and subtyping of an influenza virus outbreak. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20076-z

Mercier, E., Pisharody, L., Guy, F., Wan, S., Hegazy, N., D'Aoust, P. M., Kabir, M. P., Nguyen, T. B., Eid, W., Harvey, B., Rodenburg, E., Rutherford, C., Mackenzie, A. E., Willmore, J., Hui, C., Paes, B., Delatolla, R., & Thampi, N. (2023). Wastewater-based surveillance identifies start to the pediatric respiratory syncytial virus season in two cities in Ontario, Canada. *Frontiers in Public Health*, *11*. https://doi.org/10.3389/fpubh.2023.1261165

Naughton, C. C., Roman, F. A., Grace Alvarado, A. F., Tariqi, A. Q., Deeming, M. A., Bibby, K., Bivins, A., Rose, J. B., Medema, G., Ahmed, W., Katsivelis, P., Allan, V., Sinclair, R., Zhang, Y., Kinyua, M. N., & Author cnaughton, C. (2024). *Show us the Data: Global COVID-19 Wastewater Monitoring Efforts, Equity, and Gaps*. https://doi.org/10.1101/2021.03.14.21253564

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K. K., Wilkinson, R., & Wiedenheft, B. (2020). Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Reports Medicine*, *1*(6). https://doi.org/10.1016/j.xcrm.2020.100098

Overton, A. K., Knapp, J. J., Lawal, O. U., Gibson, R., Fedynak, A. A., Adebiyi, A. I., Maxwell, B., Cheng, L., Bee, C., Qasim, A., Atanas, K., Payne, M., Stuart, R., Fleury, M. D., Knox, N. C., Nash, D., Hungwe, Y. C., Prasla, S. R., Ho, H., … Charles, T. C. (2024). Genomic surveillance of Canadian airport wastewater samples allows early detection of emerging SARS-CoV-2 lineages. *Scientific Reports*, *14*(1), 26534. https://doi.org/10.1038/s41598-024-76925-6

Patel, M., Chaubey, A. K., Pittman, C. U., Mlsna, T., & Mohan, D. (2021). Coronavirus (SARS-CoV-2) in the environment: Occurrence, persistence, analysis in aquatic systems

and possible management. In *Science of the Total Environment* (Vol. 765). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2020.142698

Rimoldi, S. G., Stefani, F., Gigantiello, A., Polesello, S., Comandatore, F., Mileto, D., Maresca, M., Longobardi, C., Mancon, A., Romeri, F., Pagani, C., Cappelli, F., Roscioli, C., Moja, L., Gismondo, M. R., & Salerno, F. (2020). Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers. *Science of the Total Environment*, *744*. https://doi.org/10.1016/j.scitotenv.2020.140911

Safari, I., & Elahi, E. (2022). Evolution of the SARS-CoV-2 genome and emergence of variants of concern. In *Archives of Virology* (Vol. 167, Issue 2, pp. 293–305). https://doi.org/10.1007/s00705-021-05295-5

Sears, S. D., Ferreccio, C., Levine, M. M., Cordano, A. M., Monreal, J., Black, R. E., D'ottone, K., Rowe, B., & Committee, C. T. (1984). The Use of Moore Swabs for Isolation of Salmonella typhl from Irrigation Water in Santiago, Chile. In the *journal of infectious diseases*, (Vol. 149, Issue 4). https://academic.oup.com/jid/article/149/4/640/1188322

Sherchan, S. P., Shahin, S., Ward, L. M., Tandukar, S., Aw, T. G., Schmitz, B., Ahmed, W., & Kitajima, M. (2020). First detection of SARS-CoV-2 RNA in wastewater in North America: A study in Louisiana, USA. *Science of the Total Environment*, *743*, 140621. https://doi.org/10.1016/j.scitotenv.2020.140621

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28246-3

Vo, V., Tillett, R. L., Chang, C. L., Gerrity, D., Betancourt, W. Q., & Oh, E. C. (2022). SARS-CoV-2 variant detection at a university dormitory using wastewater genomic tools. *Science of the Total Environment*, *805*. https://doi.org/10.1016/j.scitotenv.2021.149930

Wang, J., Feng, H., Zhang, S., Ni, Z., Ni, L., Chen, Y., Zhuo, L., Zhong, Z., & Qu, T. (2020). SARS-CoV-2 RNA detection of hospital isolation wards hygiene monitoring during the Coronavirus Disease 2019 outbreak in a Chinese hospital. *International Journal of Infectious Diseases*, *94*, 103–106. https://doi.org/10.1016/j.ijid.2020.04.024

World Health Organization (2020a). Situation Report: World Health Organization. *Coronavirus Disease 2019. URL:* https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessd on 22 april 2022)

World Health Organization (2020b). Listings of WHO's response to COVID-19. In *World Health Organization*. URL: https://www.who.int/news/item/29-06-2020-covidtimeline. (accessd on 18 april 2022)

World Health Organization (2024). WHO COVID-19 dashboard; Number of COVID-19 cases and cumulative total death. August 4, 2024. URL: https://data.who.int/dashboards/covid19/cases (accessd on 12 january 2025)

Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C., & Zhang, Y. Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), 265–269. https://doi.org/10.1038/s41586-020-2008-3

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA* -

Journal of the American Medical Association, *323*(13), 1239–1242. https://doi.org/10.1001/jama.2020.2648

Wurtzer, S., Levert, M., Dhenain, E., Accrombessi, H., Manco, S., Fagour, N., Goulet, M., Boudaud, N., Gaillard, L., Bertrand, I., Challant, J., Masnada, S., Azimi, S., Gillon-Ritz, M., Robin, A., Mouchel, J. M., Sig, O., & Moulin, L. (2022). From Alpha to Omicron BA.2: New digital RT-PCR approach and challenges for SARS-CoV-2 VOC monitoring and normalization of variant dynamics in wastewater. *Science of the Total Environment*, *848*. https://doi.org/10.1016/j.scitotenv.2022.157740

Xiao, F., Sun, J., Xu, Y., Li, F., Huang, X., Li, H., Zhao, J., Huang, J. & Zhao, J. (2020). Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerging Infectious Diseases*, *26*(8), 1920–1922. https://doi.org/10.3201/eid2608.200681

Yaniv, K., Ozer, E., Lewis, Y., & Kushmaro, A. (2021). RT-qPCR assays for SARS-CoV-2 variants of concern in wastewater reveals compromised vaccination-induced immunity. *Water Research*, *207*(August), 117808. https://doi.org/10.1016/j.watres.2021.117808

Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. Di, Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., … Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/nejmoa2001017

2 Chapter 2: Literature Review

2.1 Genomic Structure of SARS-CoV-`2

The SARS-CoV-2 genome is roughly 29.9kb long, contains 38% GC content, 11 protein-coding genes and 12 expressed proteins (Chan et al., 2020; Lu et al., 2020b; Wu et al., 2020a). The genome contains four structural proteins namely spike (S), envelope (E), membrane (M), and nucleocapsid (N) (Cui et al., 2019). The S glycoprotein consists of 1273 amino acids and is divided into three subunits: S1, S2, and S2', each of which has a particular function during the infection of host cells (Wu et al., 2020a). The pathogenesis of SARS-CoV-2 begins with the binding of S protein to the host cell through the receptor-binding domain (RBD) in the S1 subunit, and then the fusion of the S2 subunit to the cell membrane (Zhang et al., 2020). The E protein is a relatively small structural protein with 75 amino acids, that plays an important function in viral morphogenesis and assembly (Li et al., 2020). The M protein is the second largest structural protein of SARS-CoV-2 comprising 222 amino acids that are involved in RNA packaging (Tang et al., 2020). The N protein is highly conserved and binds to viral RNA in a "bead on a string" manner through its 140 amino acid long RNA-binding domain (Rahimi et al., 2021).

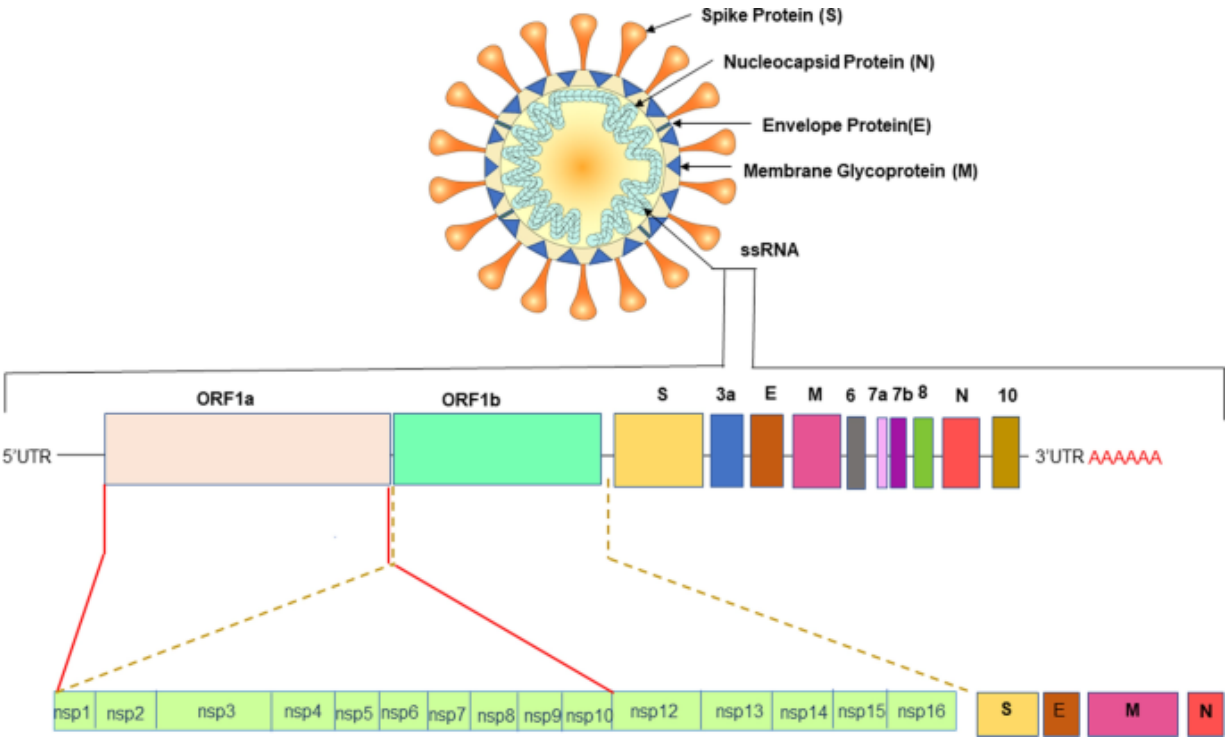

Figure 2-1: Genomic structure of SARS-CoV-2; spherical (top) and full-length genome (bottom). The ORF1ab encodes for 16 non-structural proteins (nsp1 to nsp16) and structural genes encode four structural proteins (adopted from Rastogi et al., 2020).

2.2 Genetic Diversity of SARS-CoV-2

The genetic diversity of SARS-CoV-2 is crucial for its fitness, survival, and pathogenesis, and mutations are the main sources of its genetic diversity (Rehman et al., 2020). Due to mutations, several SARS-CoV-2 variants have emerged over time, and WHO has classified them into three major groups; (a) Variant of concerns (VOCs), (b) Variant of Interests (VOIs), and (c) Variant under monitoring (VUMs) (He et al., 2021). The SARS-CoV-2 VOCs increase transmissibility, virulence, and disease severity and decrease vaccine effectiveness and other public health measures (CDC, 2021). The VOIs increase pathogenicity, decrease antigenicity, and have a detrimental effect on available treatments, or public health measures (CDC, 2021), while VUMs may alter the properties of the virus and increase risk in the future (WHO, 2021).

2.3 The Concept and Development of Wastewater based Epidemiology (WBE)

Epidemiology is the study and analysis of disease within population and regulating the disease through the identification of the risk factors (Porta, 2008). Anesthesiologist John Snow is considered as the father of epidemiology for his notable works on cholera in 1854 (Thacker, 2002). Epidemiology emerged as a discipline since World War II, while WBE was first used in scientific literature in 2001 with the idea of the presence of illicit drug residue in wastewater (Daughton, 2001). WBE is an integrated approach comprised of extracting a specific target (biomarker) in wastewater, analyzing the target, and interpreting the target to provide a holistic picture of a community (Mao et al., 2020). The biomarkers can be metabolites, endogenous substances, microorganisms, and pathogens (Choi et al., 2018). In wastewater, various disease-causing agents such as Adenovirus (Kuo et al., 2015), *E. coli* (Jiang et al., 2019), Hepatitis A, E virus (Hellmér et al., 2014; Alfonsi et al., 2018), Influenza virus (Deboosere et al., 2011; Heijnen & Medema, 2011), Norovirus (Kittigul et al., 2019), and Salmonella (Park et al., 2012) have previously been detected, which showed the prospect of WBE application for emerging pathogens. However, enteroviruses (e.g., adenovirus/norovirus) are the most common targets in WBE, whereas enveloped viruses (i.e., Influenza viruses) are less common (Corpuz et al., 2020). The enveloped viruses are mainly transmitted through droplets or aerosols and easily breakdown when they enter the wastewaters, resulting in the loss of their infectivity (Louten, 2016). Thus, it was presumed that enveloped viruses had lower persistence in wastewater, but later it was found that some enveloped viruses, such as MERS (Wigginton et al., 2015), SARS-CoV-1 (Wigginton et al., 2015) have higher persistence in wastewater, which shows the prospects of WEM for SARS-CoV-2.

2.4 The Emergence of SARS-CoV-2 WEM on a Global Scale

SARS-CoV-2 is a respiratory virus, and its RNA found in sputum, secretions, urine, and feces of symptomatic, pre-symptomatic and asymptomatic patients, which eventually enters the

wastewater network (Xiao et al., 2020; Jones et al., 2020). Thus, it was hypothesized that the SARS-CoV-2 RNA can be detected in wastewater (Foladori et al., 2020), and a group of Dutch scientists first confirmed the presence of SARS-CoV-2 RNA in wastewater in February 2020 (Medema et al., 2020). Later, the presence of SARS-CoV-2 RNA in wastewater was confirmed in Australia (Ahmed et al., 2020a), Asia (Haramoto et al., 2020), North America (Sherchan et al., 2020), South America (Fongaro et al., 2021), and Africa (Johnson et al., 2021). This WEM research extended beyond detecting SARS-CoV-2 RNA in wastewater, demonstrating a correlation between viral RNA signals and COVID-19 cases in the population. The WEM of SARS-CoV-2 also showed that it can be used as (i) an early warning tool to predict COVID-19 outbreaks days or weeks before clinical diagnosis (D'Aoust, et al., 2021), (ii) capable to identity the disease trends (Bivins et al., 2020b), (iii) assess the prevalence of infections (La Rosa et al., 2021), and (iv) determine the genetic diversity of SARS-CoV-2 (Nemudryi et al., 2020). Throughout the pandemic, WEM have provided important insight into the COVID-19 disease, and several countries including Canada, the United States (USA), Australia, and the European Union (EU) have adopted WEM as part of their regular COVID-19 surveillance (Randazzo, et al., 2020a; Gawlik et al., 2021; Hrudey & Conant, 2022). WEM was also adopted as an early warning tool on university/college campuses, particularly in North America during the pandemic (Corchis-Scott et al., 2021) Overall, throughout the COVI-19 pandemic, WEM have been expanded to over 4,648 sites across 72 countries (Naughton et al., 2024).

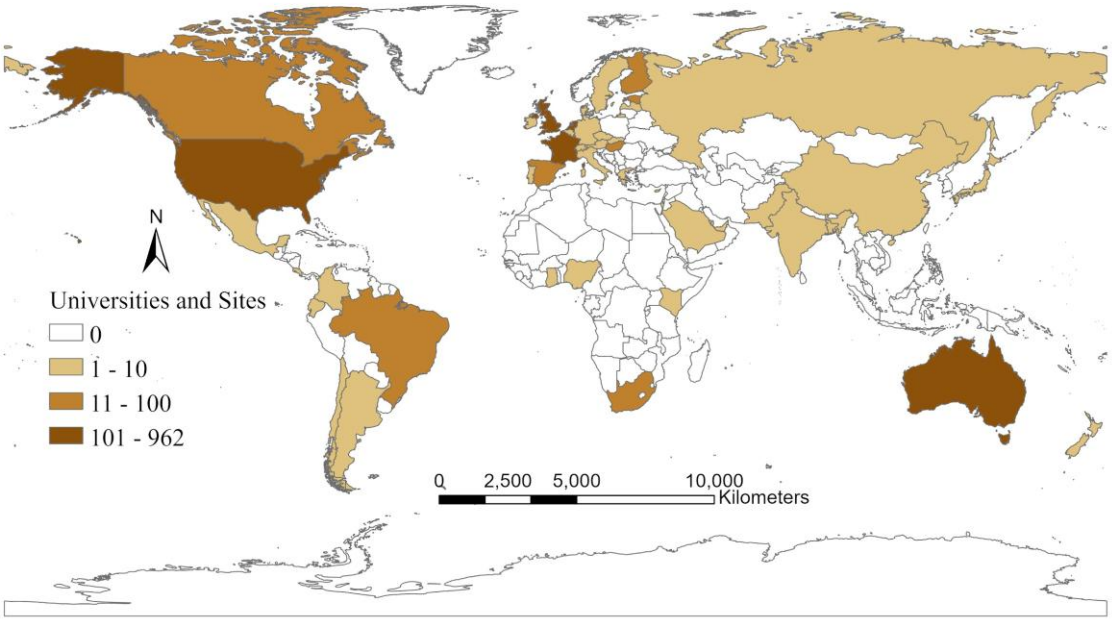

Figure 2-2: A snapshot of the global distributions of SARS-CoV-2 WEM research and programs as of December 2024 (adopted form Naughton et al., 2024).

2.5 Analytical Methodology for WEM of SARS-CoV-2

The WEM public health programs and research have contributed to the development of a standard analytical methodology for monitoring SARS-CoV-2 RNA in wastewater. The methodology incorporates a list of variables that are likely to improve the sensitivity and accuracy of the surveillance programs. The most common variables in SARS-CoV-2 WEM programs are wastewater sampling, sample storage conditions, viral concentration, total nucleic acid extraction, detection and quantification.

2.5.1 Wastewaters sampling

The genetic material of SARS-CoV-2 has been detected in wastewater and primary sludge. Wastewater samples have been collected directly from sewer networks or from wastewater treatment plants (WWTPs) prior to any treatments while primary sludge samples have been collected from the primary clarifier after the sedimentation process (Hayes et al., 2021). Grab and composite sampling are the most common methods to collect untreated wastewater as well

as primary sludge samples. Grab samples represent enriched information of a particular community at a specific time and usually collect in the morning (Ahmed et al., 2020c), whereas composite samples represent the entire information of a community for a given time period (typically 24h) (Curtis et al., 2020; Hunt et al., 2021). However, wastewater sampling entirely depends on the design and scope of the monitoring programs, with several factors influencing the decision-making process.

2.5.2 Sample types

The SARS-CoV-2 RNA is primarily shed into wastewater through human feces, respiratory secretions, and other bodily fluids which tend to be sorb onto particulate or colloidal matters. Therefore, SARS-CoV-2 RNA is preferentially detected in solid (i.e., volatile solids), although research has shown that SARS-CoV-2 RNA can also be detected in the liquid portion of wastewater (Alamin et al., 2022). However, earlier research showed that analyzing wastewater solids may yield higher viral RNA concentrations compared to liquid portion. Apart from wastewater, wastewater solids accumulated as primary clarified sludge in WWTP, and therefore, SARS-CoV-2 RNA is expected to have higher concentration in primary sludge compared to wastewater (Graham et al., 2021).

2.5.3 Sampling methods

2.5.3.1 Grab sampling

Wastewater sampling in a particular time of a day may provide enriched information on SARS-CoV-2 infection in a community. In order to catch those enhanced SARS-CoV-2 RNA in the wastewater, one or more grab samples could be taken in morning (Augusto et al., 2022), as the defection is most common in the morning. Grab sampling is a simple and convenient procedure that entails filling a container with wastewater at a specific point in time (Liu et al., 2020). Grab sampler is easy to deploy and use, requires low-technology equipment and is low-cost.

Research has demonstrated that grab samples can offer precise snapshots of COVID-19 incidence within a community (Augusto et al., 2022). However, certain people have experience diarrhea and other gastrointestinal illnesses due to SARS-CoV-2 infection, leading to irregular shedding into sewer systems (Hunt et al., 2021). Therefore, grab sampling may overlook shedding events from individuals who were not shedding during sampling period.

2.5.3.2 Composite sampling

Composite sampling is the collection of a series of grab samples multiple times a day and combined them into a single sample. This sampling method provides a representative overview of viral presence in wastewater and helps account for variations in viral load due to temporal fluctuations, thereby improving the accuracy of quantifying SARS-CoV-2 RNA in wastewater. Composite wastewater sampling can be performed using conventional autosamplers as well as recently adopted passive sampling techniques.

2.5.3.3 Automatic sampling

Autosamplers are commonly deployed for 24 hours to collect wastewater and can provide an accurate infection preview of a community. There are two types of autosampler commonly used for composite wastewater sampling: time-proportional and flow-proportional. Time-proportional auto sampling collects fixed-volume samples at defined time intervals while flow-proportional auto sampling collects variable sample volume based on the wastewater flow rate at fixed time intervals or constant volume samples in proportion to the wastewater flow rate at varying time intervals (Augusto et al., 2022). In SARS-CoV-2 WEM, time-proportional auto sampling is one of the most widely utilized methods, demonstrating success in detecting and quantifying SARS-CoV-2 RNA in wastewater (Gerrity et al., 2021; Wilson et al., 2022). However, installation of autosamplers in certain geo specific locations is complex, labor-

intensive, and requires unique specifications which lessen its user-friendliness (Curtis et al., 2020).

2.5.3.4 Passive sampling

Passive sampling involves the deployment of a device that captures analyte molecules from wastewater by using absorbent mediums. Passive sampling has been used as a promising tool to capture wide ranges of pollutants since 1987. The Moore swab is the most primitive form of passive sampler, used to trace *Salmonella Paratyphi B* from effluent sewage (Sears et al., 1984). Later, it was adopted to capture microorganisms such as fecal-borne pathogens, poliovirus, and human norovirus in water. Moore swab sampling has been successfully validated for detecting SARS-CoV-2 RNA in wastewater (Liu et al., 2020). However, various modifications to the Moore swab sampling method have been implemented throughout the pandemic to enhance its effectiveness in SARS-CoV-2 WEM. In addition to the Moore swab, various passive samplers such as the Torpedo and COSCa-ball have been designed and implemented during the COVID-19 pandemic (Li et al., 2022; Schang et al., 2021; Wilson et al., 2022). The Torpedo passive sampler consists of 3D-printed housing units, whereas the COSCa-ball passive sampler is a 10 cm diameter hollow sphere. In both the Torpedo and COSCa-ball samplers, sampling mediums are placed inside to effectively capture wastewater solids. The most common sampling mediums used in the passive sampling during the pandemic are electronegative membrane filters (Habtewold et al., 2022; Hayes et al., 2022; Liu et al., 2020; Schang et al., 2021; Wilson et al., 2022), cotton buds (Habtewold et al., 2022; Hayes et al., 2022), cheesecloth (Hayes et al., 2022), cellulose sponge (Hayes et al., 2022), tampons (Bivins et al., 2022; Kevill et al., 2022; Liu et al., 2020; Wilson et al., 2022), and medical gauzes (Habtewold et al., 2022; Rafiee et al., 2021). The effectiveness of detecting SARS-CoV-2 RNA in wastewater might vary depending on the absorbent materials used within the passive sampler. Besides, the deployment time can influence the detection of SARS-CoV-2

RNA in wastewater. One study (Vincent-Hubert et al., 2022) has evaluated various deployment durations, ranging from 4 to 96 hours, and found that a 24-hour deployment is the most effective, similar to the performance of an autosampler.

2.5.4 Sample storage conditions

Following sample collection, transportation and storage conditions are critical for the survival of SARS-CoV-2 RNA in wastewaters. Samples should be transported on ice to the laboratory and analyzed within 24 hours of arrival. However, one study reported that SARS-CoV-2 RNA can survive for up to 5 days in wastewater at 4 ^{0}C (Ahmed et al., 2020b), while others reported that SARS-CoV-2 RNA can survive for up to 27 days in wastewaters at temperatures ranging from 4 ^{0}C to 37 ^{0}C (Bivins et al., 2020a). Moreover, prolonged persistence of SARS-CoV-2 RNA has also been observed at freezing temperatures between -20 ^{0}C or -75 ^{0}C ((Hokajärvi et al., 2021).

2.5.5 Sample concentration

The SARS-CoV-2 in feces and urine has drastically diluted in sewer networks, and the recovery of SARS-CoV-2 RNA is largely dependent on the selection of suitable concentration methods. There are a variety of virus concentration methods for detecting enteric viruses in wastewater, but only a few of them have been utilized in SARS-CoV-2 WEM. The most commonly adopted concentration methods are ultrafiltration, centrifugation, polyethylene glycol, electrostatically charged membranes method, skimmed-milk flocculation, and nanotrap magnetic viral particles.

2.5.5.1 Ultrafiltration

Ultrafiltration (UF) is a filtration method in which retentate and eluate are separated using pressure or concentration gradients via a semipermeable membrane (Cheryan, 1998). The method is commonly applied for removing macromolecules from water and wastewater as well

as concentrating protein in numerous beverage industries (Clever et al., 2000; Nigam et al., 2008). Researchers adopted UF as a virus concentrating method in wastewaters and become a well-accepted approach for concentrating non-enveloped viruses such as polioviruses, noroviruses, and adenoviruses (Lu et al., 2020a). Throughout the COVID-19 pandemic, UF has been adopted as a concentration method for concentrating SARS-CoV-2 RNA in wastewater and revealed that UF can successfully concentrate SARS-CoV-2 RNA from wastewater (Ahmed et al., 2020c; Medema et al., 2020; Nemudryi et al., 2020; Wu et al., 2020b). The most common filter units used to concentrate SARS-CoV-2 RNA in wastewater are 100 kDa molecular weight cut-off (Medema et al., 2020; Nemudryi et al., 2020),10 kDa molecular weight cut-off (Ahmed et al., 2020a) and 30 kDa molecular weight cut-off (Wu et al., 2020b). The advantage of the UF method is that it does not require any preconditioning, which reduces not only the workload but also the batch effect and systematic error (Rajal et al., 2007; Lu et al., 2020a).

2.5.5.2 Centrifugation

Centrifugation is a mechanical process used to separate particles from a solution, based on the size, shape, and density (Hansen, 1999). This is a widely used method in molecular biology laboratories to isolate DNA, purify virus particles, and distinguish subtle differences in the conformation of molecules (Stephenson, 2016). The technique was previously used to concentrate bacteria, enveloped and non-enveloped viruses from wastewaters, and it was commonly employed for SARS-CoV-2 WEM over the pandemic. The key benefit of the method is that it can concentrate viruses from both the solid and liquid phases, while its main drawback is that it can only concentrate a small number of samples at once (Ahmed et al., 2020c).

2.5.5.3 Polyethylene glycol (PEG) based two-phase separation

PEG precipitation is a two-phase solution composed of dextran, methylcellulose, and water (Sim et al., 2012), and has been frequently used to concentrate enteroviruses from groundwater, river water, tap water, and wastewater (Lu et al., 2020a). Besides, the method has also been used to concentrate and purify other viruses such as bacteriophage T2, adenovirus, ECHO virus, and poliovirus from wastewater (Ikner et al., 2012). PEG was approved as a standard concentrating method for poliovirus surveillance in the environment by the WHO in 2003 (WHO, 2003). During the COVID-19 pandemic, PEG have also adopted as a concentrating method for SARS-CoV-2 WEM (Haramoto et al., 2020; Kumar et al., 2020; Wu et al., 2020b; Bertrand et al., 2021). The PEG is relatively inexpensive method and can concentrate SARS-CoV-2 RNA from both solid and liquid phases, but the method is time consuming and only a portion of concentrate is used to extract RNA (Ahmed et al., 2020c).

2.5.5.4 Electrostatically charged membranes method

Electrically charged membranes consist of a mixture of negatively and positively charged segments that trap particles/viruses based on their surface charge (Zhou & Pang, 2018). The membrane is called an anion-exchange membrane when positive ions are present, and a cation-exchange membrane when negative fixed ions are present (Zhou & Pang, 2018). Separation is achieved in an electrically charged membrane not only by pore size but also by the exclusion of co-ions (Breite et al., 2019). The most common electrostatically charged membrane for concentrating virus from water is viral adsorption elution (VIRADEL), which uses electrostatically charged microporous materials as filtration media (Ikner et al., 2012). In VIRADEL method, water is passed through a filter media to which viruses have been attached, and then eluting or removing the viruses from the filter with a solution that can be supplemented with mechanical disruption methods (Anderson-Coughlin & Kniel, 2019). The VIRADEL can be classified as electropositive membrane filtration (EPMF) and electronegative membrane

filtration (ENMF) based on the electrical surface charge (Mousazadeh et al., 2021). The EPMF is capable of extracting enterovirus and norovirus from groundwater and surface water (Fout et al., 2015). The ENMF was adopted to concentrate SARS-CoV-2 RNA from wastewater, but the presence of organic matter and high turbidity in wastewater limit the virus recovery efficiency of the method (Lu et al., 2020a). However, the recovery efficiency of ENMF has increased with the addition of $MgCl_2$ or NaCl as the addition of these salts can enhance the attachment of virus particles to a cellulose nitrate membrane filter via salt-bridging (Haramoto et al., 2020).

2.5.5.5 Skimmed-milk flocculation (SMF)

SMF method entails attaching of viruses to pre-flocculated skimmed milk proteins, allowing the flocs to settle by gravity and then dissolving the sediment in a buffer (Calgua et al., 2008). SMF is an effective method for concentrating viruses from river water, seawater, groundwater, and wastewater (Gonzales-Gustavson et al., 2017). This concentration method eluted SARS-CoV-2 RNA using glycine alkaline buffer prior to organic flocculation (Mousazadeh et al., 2021). The SMF is more effective at concentrating SARS-CoV-2 RNA from wastewater, but the method has not been widely used because it is time-consuming (Guerrero-Latorre et al., 2020; Philo et al., 2021). The main advantage of SMF over other concentration methods is that it reduces the cost as the procedure doesn't require any special equipment like filters or pressure pumps (Mousazadeh et al., 2021).

2.5.5.6 Aluminum-driven flocculation

The aluminum-driven flocculation method works on the adsorption mechanism by which freshly formed $AL(OH)_3$ adsorbs viruses from wastewater. This method is commonly applied in water and wastewater treatment plants all over the world. The approach was rarely used for concentrating SARS-CoV-2 RNA in wastewaters, but it was effective in recovering SARS-

CoV-2 RNA from wastewaters, even in low-prevalence areas (Barril et al., 2021; Randazzo, et al., 2020b).

2.5.5.7 Nanotrap magnetic virus particles

Nanotrap magnetic virus particles are hydrogel nanoparticles designed to capture and concentrate viruses from complex biological samples, such as wastewater (Huang et al., 2019). These magnetic particles are coated with polymers that exhibit a strong affinity for viral particles. When introduced into a sample or medium, the hydrogel binds selectively to the viral particles. Following this binding, an external magnetic field is applied, which causes the beads with the attached viral particles to separate from the rest of the sample. This process effectively isolates the virus from other contaminants and impurities present in the sample (Angga et al., 2022). Nanotrap magnetic virus particles method was previously utilized to extract nucleic acids from a variety of metrics such as serum, urine, sputum, whole blood, tissue, and wastewater (Parra-Guardado et al., 2022). Throughout the COVID-19 pandemic, nanotrap magnetic virus particles were extensively used to extract SARS-CoV-2 RNA from wastewater. The SARS-CoV-2 RNA extracted using nanotrap magnetic virus particles likely free from non-target materials and potential inhibitors, making it suitable for downstream analyses such as sequencing due to the improved quality and integrity of the RNA (Jiang et al., 2024). The use of nanotrap magnetic virus particles for RNA extraction offers several advantages, including rapid processing, ease of use, and does not require any specialized equipment. However, the primary drawback of this method is the higher cost compared to other methods.

2.5.6 Total nucleic acid extraction

Nucleic acid extraction is an important step in separating SARS-CoV-2 RNA from wastewater which involves cell lysis, denaturation and deactivation of RNases, removal of cellular components, and RNA recovery (Michael-Kordatou et al., 2020). The SARS-CoV-2 RNA

extraction from wastewater is usually challenging, and successful extraction depends on the selection of appropriate kits (Ahmed et al., 2020d). There are several commercial kits available for extracting SARS-CoV-2 RNA from wastewaters, but the effectiveness of RNA recovery varies for each kit (Ahmed et al., 2020d; Haramoto et al., 2020; La Rosa et al., 2020; Medema et al., 2020). Thereby, CDC (2020a) recommended a few characteristics that should be considered when choosing an extraction kit for SARS-CoV-2; (i) extraction protocol should be able to extract highly purified nucleic acid from environmental samples, (ii) extraction kit/protocol should be specifically designed to purify RNA and include RNase denaturants prior to lysis. The organic extraction, silica membrane-based spin columns, paramagnetic particle kits meet the criteria, and CDC approved those as SARS-CoV-2 RNA extraction kits (CDC, 2020a). All of the approved extraction kits are based on the either TRIzol-chloroform or lysis buffer/TRIzol, or silica membrane spin column except the magnetic virus particles (Mousazadeh et al., 2021). The silica column is simple to use, low-cost, and less prone to contamination from proteins and other chemicals, while the organic extraction approach is prone to contamination from proteins and other organic solvents such as phenol-chloroform, organic salts, and ethanol (Tavares et al., 2011).

2.5.7 Detection and quantification

The presence of SARS-CoV-2 RNA in wastewater can be determined using Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR) and Reverse Transcription Droplet Digital Polymerase Chain Reaction (RT-ddPCR). The details of each detection method are as follows.

2.5.7.1 RT-qPCR

RT-qPCR is the most commonly used method for quantifying SARS-CoV-2 RNA in wastewaters. In this method, reverse transcriptase (RT) enzyme reverses transcribed RNA

template to complementary DNA (cDNA) using short DNA sequences (primers) (Bustin & Noman, 2020). Then, a fluorescent dye or a sequence-specific, fluorescence-labeled DNA probe is employed to monitor the amplification of target DNA in real time. As amplification proceeds through multiple cycles, the accumulation of fluorescence or an electrical signal corresponds to the presence of viral cDNA, enabling quantitative determination of SARS-CoV-2 RNA concentrations based on a standard curve (Hamouda et al., 2021). The RT-qPCR method for quantifying SARS-CoV-2 RNA in wastewater is faster, requires less sample handling, decreases bench time, and reduces pipetting errors (Eftekhari et al., 2021).

2.5.7.2 RT-ddPCR

RT-ddPCR is a relatively recent technology for detecting and measuring RNA, in which a sample containing RNA is partitioned into thousands of nanoliter droplets, and each droplet undergoes its own RT-PCR reaction (Pinheiro et al., 2012). The RT-PCR reaction detects whether or not the droplets contain the RNA of interest, allowing for an absolute estimate of the amount of target molecules in the reaction based on the poisson distribution (Huggett et al., 2013). The key advantages of RT-ddPCR over RT-qPCR are absolute quantification of virus genome copy numbers without the requirement for external calibration (Barceló, 2020), increased precision (Brunetto et al., 2014), and insensitivity to PCR inhibitors (Huggett et al., 2008). The precision, specificity, and sensitivity of RT-qPCR and RT-ddPCR for the detection of SARS-CoV-2 RNA in wastewaters were compared and found that RT-ddPCR had higher sensitivity and precision (Ciesielski et al., 2021; Dumke et al., 2021), and the lower limit of detection (Ciesielski et al., 2021; Gonzalez et al., 2021). The RT-ddPCR method is currently expensive compared to RT-qPCR as it requires sophisticated instruments and expensive reagents (Alygizakis et al., 2021).

2.6 Whole genome sequencing (WGS) of SARS-CoV-2

WGS of SARS-CoV-2 entails sequencing of virus's genetic material and identifying the changes in the SARS-CoV-2 genome (Vancollie, 2021). The first genome sequence of SARS-CoV-2 (accession MN908947.3) was recovered from a clinical sample in December 2019 and has since been used as a reference sequence (Wu et al., 2020a; Zhou et al., 2020). The initial genome sequences of SARS-CoV-2 were obtained from clinical samples and later different next generation sequencing (NGS) based strategies such as shotgun metagenomics, amplicon-based, and hybrid capture-enrichment sequencing have been used to sequence SARS-CoV-2 genome in wastewater. The details of each approach are as follows.

2.6.1 Shotgun metagenomics

Shotgun metagenomics is the unbiased (shotgun) sequencing of all (meta) bacterial genomes present in one sample, enabling the analysis of complex microbial communities (Handelsman, 2005). The most common approach of shotgun metagenomics is 16S rRNA gene sequencing which interrogates the microbial complexity of a sample (Hamady & Knight, 2009). A typical workflow of the shotgun metagenomics approach consists of RNA fragmentation, cDNA synthesis, library preparation, and sequencing. The major advantages of shotgun metagenomics include it doesn't require any prior knowledge of viral sequence and eliminates the potential effects of divergent regions on capture and amplicon approaches (Chiara et al., 2021).

2.6.2 Amplicon-based sequencing

Amplicon-based sequencing involves targeting specific genomic regions, amplifying them through PCR, and then sequencing the generated amplicons. This method is widely used for detecting and analyzing specific genetic material, including mutations or gene variations (Bybee et al., 2011). The approach is highly specific and requires a good understanding of the targeted regions prior to the sequencing as the amplification exclusively depends on the

primers. Amplicon-based sequencing is the most widely used approach for sequencing SARS-CoV-2 genome which adopts an enrichment workflow consisting of first-strand cDNA synthesis followed by genome amplification with multiplex PCR (Nemudryi et al., 2020). The multiplex PCRs enable the production of pools of amplicons that cover the entire length of the viral genome (Quick et al., 2020). The amplicon sequencing is limited to primer's efficiency as well as the mutations in the primer annealing regions (Chiara et al., 2021). Thereby, amplification across the genome can be biased, with decreased coverage in specific genomic regions. For example, ARTIC v3 primer sets can't detect the presence certain mutations of Omicron VOC and the associated linages, thereby, ARTIC v4.1 primer sets have been developed to capture those mutations in the SARS-CoV-2 genome (Davis et al., 2021; Lambisia et al., 2022).

2.6.3 Hybrid capture-enrichment sequencing

Hybrid capture enrichment is a targeted NGS method that utilizes probes to hybridize the regions of interest (Drmanac et al., 2002). This sequencing strategy was originally developed for human genome sequencing and later adopted as a common sequencing approach in NGS (Albert et al., 2007). The nearly complete genome (> 90%) of SARS-CoV-2 was recovered using the hybrid capture-enrichment method which enables the detection of circulating variants of SARS-CoV-2 in wastewater (Fitzpatrick et al., 2021; Nagy-Szakal et al., 2021). The hybrid capture-enrichment method is based on the generation of a larger number of fragments using multiplex PCR and then hybridizing the targeted regions of SARS-CoV-2 (Xiao et al., 2020). The capture enrichment method is more robust to the genomic variability as the targeted regions are less dependent on the perfect complementary (Chiara et al., 2021). This method offers an unbiased representation of intra-sample variation (Maurano et al., 2020; Xiao et al., 2020). The main limitation of the target-enrichment sequencing is to develop biotinylated

probes from reference sequences that require high expertise and knowledge about the sequence of the reference genome (Gaudin & Desnues, 2018).

2.6.4 Bioinformatics workflow to analyze SARS-CoV-2 sequence

Bioinformatics is an interdisciplinary science that integrates biological data with informatics techniques (e.g., applied math, computer science, and statistics) to understand and organize the information associated with biological data (Luscombe et al., 2001). A set of bioinformatics algorithms executed in a predefined sequence to process complex biological data is collectively referred to as a bioinformatics pipeline (Leipzig, 2017). A bioinformatics pipeline encompasses a series of transformations of the massive sequence using multiple software tools, databases, and operating environments (Roy et al., 2018). Several bioinformatics pipelines have been developed to aid in the genomic surveillance of SARS-CoV-2, and the SARS-CoV-2 bioinformatics pipeline is automated, and sufficient quality control (QC) is required to assure the precise variant calling. A typical SARS-CoV-2 bioinformatics pipeline consists of base calling, demultiplexing, adapter trimming, alignment, and variant calling.

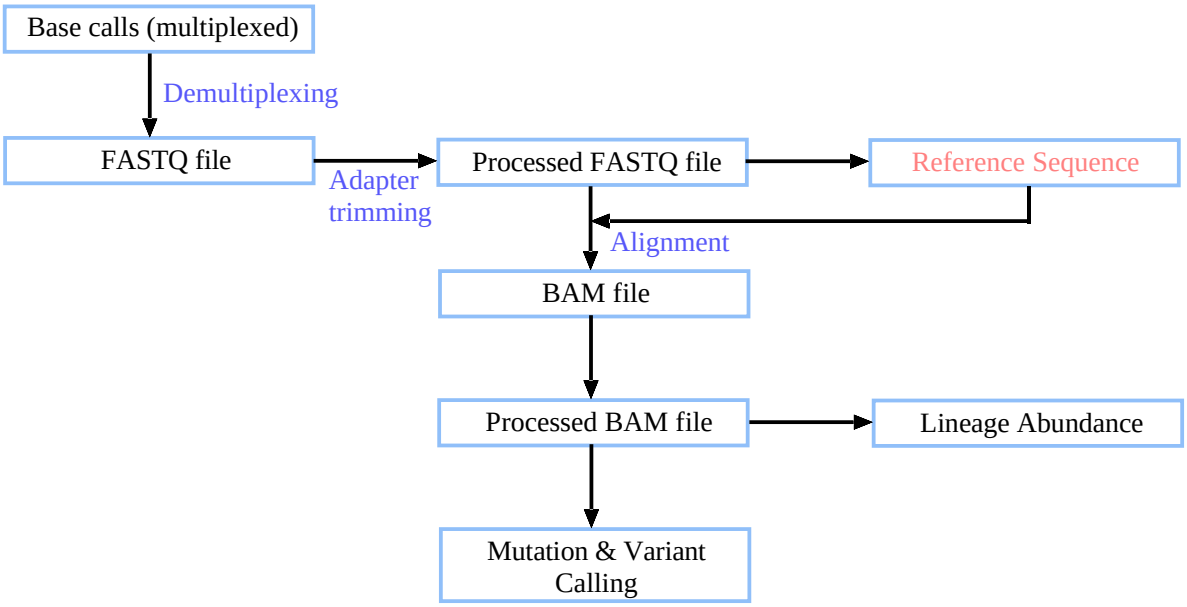

Figure 2-3: SARS-CoV-2 bioinformatics pipeline. The figure illustrate*s* the bioinformatics pipeline and its components that are usually used for the processing of whole genome sequenc*es* of SARS-CoV-2 (adopted from Roy et al., 2018).

2.6.5 Sequence generation and quality control

Sequence generation is the process that converts sensor data from the sequencing platform and identifies the sequence of nucleotides for each of the short fragments (reads) of DNA (Heather & Chain, 2016). A corresponding Phred quality score (a measure of how likely a base call in DNA sequencing is incorrect) is assigned for each nucleotide of the short fragments of DNA, and the read sequences along with Phred quality score are stored in a FASTQ file. The FASTQ files are the standard format for the representation of biological sequence information (Hosseini et al., 2016). The quality of the generated SARS-CoV-2 sequence evaluates through either FastQC (Illumina) or pycoQC (Nanopore) based on the sequencing platforms (Simon Andrews, 2020). Then the barcode or adapter trimming is performed using fastp (Illumina) or artic guppy plex (Nanopore) (Chen et al., 2018).

2.6.6 Sequence alignment

Sequence alignment is the process of determining the similarities of short DNA sequences with the reference sequence (Prjibelski et al., 2018). The generated sequences are aligned with the SARS-CoV-2 reference sequence (accession MN908947.3) using either Bowtie2 (Langmead & Salzberg, 2012) or Minimap2 (Li, 2018). A minimum Phred quality score (Q20 or Q30) is usually assigned for the alignment which increases confidence in the alignment process. The aligned sequence is stored in a standard binary alignment map (BAM) file format. Using the BAM files, the % of genome and depth (X) of coverage of the mapped sequences can be calculated using mosdepth (Pedersen & Quinlan, 2018).

2.6.7 Variant calling

Variant calling is the process of detecting the differences in nucleotides between the sample and reference sequence (Koboldt, 2020). The variant calling algorithm is based on the types of variants, such as SNVs, insertions and deletions (indels), copy number variations (CNV), and large structural alterations (insertions, inversions, and translocations). The typical input for variant calling is a set of aligned reads in the BAM file format. The variant calling of SARS-CoV-2 is performed using a number of tools such as iVAR (Grubaugh et al., 2019), BCFTools (Li et al., 2009), FreeBayes (Garrison & Marth, 2012), VarScan (Koboldt et al., 2009), LoFreq (Wilm et al., 2012), and GATK (McKenna et al., 2010). A recent study compared the sensitivity of the variant calling tools and revealed that BCFtools, FreeBayes, and VarScan called the expected mutations with higher precision, while LoFreq gave the least reliable results due to the high number of false-positive mutations detected (Bassano et al., 2023). Although the different tools indicated different levels of precision, the accuracy of variant calling is highly dependent on the quality of bases and the aligned reads (Koboldt, 2020).

2.6.8 Lineages prevalence estimation

The SARS-CoV-2 genome recovered from wastewater is derived from multiple infected individuals, making it a composite source of genetic material that contains information on various variants and lineages of the virus. To capture the dynamics of virus evolution and spread, Freyja, a bioinformatics tool is utilized to infer relative abundance of lineages in wastewater (Karthikeyan et al., 2022). A .bam file used on the Freyja workflow (v1.3.10) to determines lineage abundance through a regression approach that considers depth weighting and least absolute deviation. The Freyja packages are easily accessible and found at https://github.com/andersen-lab/Freyja. Another tool, Crykey, has been developed for the rapid identification of cryptic SARS-CoV-2 mutations in wastewater (Liu et al., 2024).

2.7 References

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

Ahmed, W., Bertsch, P. M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic, A., Riddell, S., Sherchan, S. P., Simpson, S. L., Sirikanchana, K., Symonds, E. M., Verhagen, R., Vasan, S. S., Kitajima, M., & Bivins, A. (2020b). Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*, *191*(August), 110092. https://doi.org/10.1016/j.envres.2020.110092

Ahmed, W., Bertsch, P. M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B. R., Mueller, J. F., Simpson, S. L., Smith, W. J. M., Symonds, E. M., Thomas, K. V., Verhagen, R., & Kitajima, M. (2020c). Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Science of the Total Environment*, *739*(June), 139960. https://doi.org/10.1016/j.scitotenv.2020.139960

Ahmed, W., Bivins, A., Bertsch, P. M., Bibby, K., Choi, P. M., Farkas, K., Gyawali, P., Hamilton, K. A., Haramoto, E., Kitajima, M., Simpson, S. L., Tandukar, S., Thomas, K. V., & Mueller, J. F. (2020d). Surveillance of SARS-CoV-2 RNA in wastewater: Methods optimization and quality control are crucial for generating reliable public health information. *Current Opinion in Environmental Science and Health*, *17*, 82–93.

https://doi.org/10.1016/j.coesh.2020.09.003

Alamin, M., Tsuji, S., Hata, A., Hara-Yamamura, H., & Honda, R. (2022). Selection of surrogate viruses for process control in detection of SARS-CoV-2 in wastewater. *Science of the Total Environment, 823.* https://doi.org/10.1016/j.scitotenv.2022.153737

Albert, T. J., Molla, M. N., Muzny, D. M., Nazareth, L., Wheeler, D., Song, X., Richmond, T. A., Middle, C. M., Rodesch, M. J., Packard, C. J., Weinstock, G. M., & Gibbs, R. A. (2007). Direct selection of human genomic loci by microarray hybridization. *Nature Methods*, *4*(11), 903–905. https://doi.org/10.1038/nmeth1111

Alfonsi, V., Romanò, L., Ciccaglione, A. R., La Rosa, G., Bruni, R., Zanetti, A., Libera, S. Della, Iaconelli, M., Bagnarelli, P., Capobianchi, M. R., Garbuglia, A. R., Riccardo, F., Tosti, M. E., Ferrigno, L., Crateri, S., Galli, C., Tagliacarne, C., Giordani, M. T., Morelli, E., … Megnia, A. S. (2018). Hepatitis e in Italy: 5 years of national epidemiological, virological and environmental surveillance, 2012 to 2016. *Eurosurveillance*, *23*(41). https://doi.org/10.2807/1560-7917.ES.2018.23.41.1700517

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

Anderson-Coughlin, B. L., & Kniel, K. E. (2019). Recovery and detection of enteric viruses from non-traditional irrigation water sources. *Methods and Protocols*, *2*(3), 1–8. https://doi.org/10.3390/mps2030055

Angga, M. S., Malla, B., Raya, S., Kitano, A., Xie, X., Saitoh, H., Ohnishi, N., & Haramoto, E. (2022). Development of a magnetic nanoparticle-based method for concentrating

SARS-CoV-2 in wastewater. *Science of the Total Environment, 848.* https://doi.org/10.1016/j.scitotenv.2022.157613

Augusto, M. R., Claro, I. C. M., Siqueira, A. K., Sousa, G. S., Caldereiro, C. R., Duran, A. F. A., de Miranda, T. B., Bomediano Camillo, L. de M., Cabral, A. D., & de Freitas Bueno, R. (2022). Sampling strategies for wastewater surveillance: Evaluating the variability of SARS-COV-2 RNA concentration in composite and grab samples. *Journal of Environmental Chemical Engineering,* 10(3). https://doi.org/10.1016/j.jece.2022.107478

Barceló, D. (2020). Wastewater-Based Epidemiology to monitor COVID-19 outbreak: Present and future diagnostic methods to be in your radar. *Case Studies in Chemical and Environmental Engineering*, *2*(September). https://doi.org/10.1016/j.cscee.2020.100042

Barril, P. A., Pianciola, L. A., Mazzeo, M., Ousset, M. J., Jaureguiberry, M. V., Alessandrello, M., Sánchez, G., & Oteiza, J. M. (2021). Evaluation of viral concentration methods for SARS-CoV-2 recovery from wastewaters. *Science of the Total Environment*, *756*, 144105. https://doi.org/10.1016/j.scitotenv.2020.144105

Bassano, I., Ramachandran, V.K., Khalifa, M.S., Lilley, C. J., Brown, M. R., Aerle, V., Denise, H., Rowe, W., Airey, G., Cairns, E., Pickwell, N. D., Wilson, M., Carlile, M., Holmes, N., Payne, A., Paterson, S., Loose, M., Burke, T., Wade, M. J., Grimsley, J. M. S., & Bassano, I. (2023). Evaluation of variant calling algorithms for wastewater-based epidemiology using mixed populations of SARS-CoV-2 variants in synthetic and wastewater samples. *Microbial Genomics.* DOI: 10.1099/mgen.0.000933.

Bertrand, I., Challant, J., Jeulin, H., Hartard, C., Mathieu, L., Lopez, S., Schvoerer, E., Courtois, S., & Gantzer, C. (2021). Epidemiological surveillance of SARS-CoV-2 by genome quantification in wastewater applied to a city in the northeast of France: Comparison of ultrafiltration- and protein precipitation-based methods. *International*

Journal of Hygiene and Environmental Health, *233*(August 2020). https://doi.org/10.1016/j.ijheh.2021.113692

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

Bivins, A., Greaves, J., Fischer, R., Yinda, K. C., Ahmed, W., Kitajima, M., Munster, V. J., & Bibby, K. (2020a). Persistence of SARS-CoV-2 in Water and Wastewater. *Environmental Science and Technology Letters*, *7*(12), 937–942. https://doi.org/10.1021/acs.estlett.0c00730

Bivins, A., North, D., Ahmad, A., Ahmed, W., Alm, E., Been, F., Bhattacharya, P., Bijlsma, L., Boehm, A. B., Brown, J., Buttiglieri, G., Calabro, V., Carducci, A., Castiglioni, S., Cetecioglu Gurol, Z., Chakraborty, S., Costa, F., Curcio, S., De Los Reyes, F. L., … Bibby, K. (2020b). Wastewater-Based Epidemiology: Global Collaborative to Maximize Contributions in the Fight against COVID-19. *Environmental Science and Technology*, *54*(13), 7754–7757. https://doi.org/10.1021/acs.est.0c02388

Bivins, A., Lott, M., Shaffer, M., Wu, Z., North, D., Lipp, E. K., & Bibby, K. (2022). Building-level wastewater surveillance using tampon swabs and RT-LAMP for rapid SARS-CoV-2 RNA detection. *Environmental Science: Water Research & Technology*, 8(1), 173–183. https://doi.org/10.1039/d1ew00496d

Breite, D., Went, M., Prager, A., Kühnert, M., & Schulze, A. (2019). Charge separating

microfiltration membrane with pH-dependent selectivity. *Polymers*, *11*(1). https://doi.org/10.3390/polym11010003

Brunetto, G. S., Massoud, R., Leibovitch, E. C., Caruso, B., Johnson, K., Ohayon, J., Fenton, K., Cortese, I., & Jacobson, S. (2014). Digital droplet PCR (ddPCR) for the precise quantification of human T-lymphotropic virus 1 proviral loads in peripheral blood and cerebrospinal fluid of HAM/TSP patients and identification of viral mutations. *Journal of NeuroVirology*, *20*(4), 341–351. https://doi.org/10.1007/s13365-014-0249-3

Bustin, S. A. and Nolan, T. (2020). *RT-qPCR Testing of SARS-CoV-2 : A Primer*. *Int. J. Mol. Sci.*, 21(8), 3004. https://doi.org/10.3390/ijms21083004

Bybee, S. M., Bracken-Grissom, H., Haynes, B. D., Hermansen, R. A., Byers, R. L., Clement, M. J., Udall, J. A., Wilcox, E. R., & Crandall, K. A. (2011). Targeted amplicon sequencing (TAS): A scalable next-gen approach to multilocus, multitaxa phylogenetics. *Genome Biology and Evolution*, *3*(1), 1312–1323. https://doi.org/10.1093/gbe/evr106

Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P., Hundesa, A., Wyn-Jones, A. P., López-Pila, J. M., & Girones, R. (2008). Development and application of a one-step low cost procedure to concentrate viruses from seawater samples. *Journal of Virological Methods*, *153*(2), 79–83. https://doi.org/10.1016/j.jviromet.2008.08.003

Carrillo-Reyes, J., Barragán-Trinidad, M., & Buitrón, G. (2021). Surveillance of SARS-CoV-2 in sewage and wastewater treatment plants in Mexico. *Journal of Water Process Engineering*, 40, 101815.

CDC. (2020a). *Guidance for U.S. Centers for Disease Control and Prevention Staff for the Establishment and Management of Public Health Rapid Response Teams for Disease Outbreaks*. *May*. https://stacks.cdc.gov/view/cdc/89022

CDC. (2021). SARS-CoV-2 Variant Classifications and Definitions. In *Cdc* (pp. 1–12). https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Interest

Chan, J. F. W., Yuan, S., Kok, K. H., To, K. K. W., Chu, H., Yang, J., Xing, F., Liu, J., Yip, C. C. Y., Poon, R. W. S., Tsoi, H. W., Lo, S. K. F., Chan, K. H., Poon, V. K. M., Chan, W. M., Ip, J. D., Cai, J. P., Cheng, V. C. C., Chen, H., … Yuen, K. Y. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*, *395*(10223), 514–523. https://doi.org/10.1016/S0140-6736(20)30154-9

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Cheryan, M. (1998). *Ultrafiltration and microfiltration handbook*. CRC press.

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

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., & 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

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

Clever, M., Jordt, F., Knauf, R., Räbiger, N., Ruedebusch, M., & Hilker-Scheibel, R. (2000). Process water production from river water by ultrafiltration and reverse osmosis. *Desalination*, *131*(1–3), 325–336. https://doi.org/10.1016/S0011-9164(00)90031-6

Corchis-Scott, R., Geng, Q., Seth, R., Ray, R., Beg, M., Biswas, N., Charron, L., Drouillard, K. D., D'Souza, R., Heath, D. D., Houser, C., Lawal, F., McGinlay, J., Menard, S. L., Porter, L. A., Rawlings, D., Scholl, M. L., Siu, K. W. M., Tong, Y., … McKay, R. M. L. (2021). Averting an Outbreak of SARS-CoV-2 in a University Residence Hall through Wastewater Surveillance. *Microbiology Spectrum*, *9*(2). https://doi.org/10.1128/spectrum.00792-21

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

Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, *17*(3), 181–192. https://doi.org/10.1038/s41579-018-0118-9

Curtis, K., Keeling, D., Yetka, K., Larson, A., & Gonzalez, R. (2020). Wastewater SARS-CoV-2 Concentration and Loading Variability from Grab and 24-Hour Composite Samples. *MedRxiv*,2020.07.10.20150607.

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

Daughton, C. G. (2001). AND Science of Environmental Chemistry and Mass Spectrometry : Pharmaceuticals. *J Am Soc Mass Spectrom.*, *12*(01), 1067–1076.

Davis, J. J., Long, S. W., Christensen, P. A., Olsen, R. J., Olson, R., Shukla, M., Subedi, S., Stevens, R., & Musser, J. M. (2021). Analysis of the ARTIC Version 3 and Version 4 SARS-CoV-2 Primers and Their Impact on the Detection of the G142D Amino Acid Substitution in the Spike Protein. *Microbiology Spectrum*, *9*(3), 2–6. https://doi.org/10.1128/spectrum.01803-21

Deboosere, N., Horm, S. V., Pinon, A., Gachet, J., Coldefy, C., Buchy, P., & Vialette, M. (2011). Development and validation of a concentration method for the detection of influenza a viruses from large volumes of surface water. *Applied and Environmental Microbiology*, *77*(11), 3802–3808. https://doi.org/10.1128/AEM.02484-10

Drmanac, R., Drmanac, S., Chui, G., Diaz, R., Hou, A., Jin, H., Jin, P., Kwon, S., Lacy, S., Moeur, B., Shafto, J., Swanson, D., Ukrainczyk, T., Xu, C., & Little, D. (2002). Sequencing by hybridization (SBH): advantages, achievements, and opportunities. *Advances in Biochemical Engineering/Biotechnology*, *77*, 75–101. https://doi.org/10.1007/3-540-45713-5_5

Dumke, R., Barron, M. de la C., Oertel, R., Helm, B., Kallies, R., Berendonk, T. U., & Dalpke, A. (2021). Evaluation of two methods to concentrate SARS-CoV-2 from untreatedwastewater. *Pathogens*, *10*(2), 1–7. https://doi.org/10.3390/pathogens10020195

Eftekhari, A., Alipour, M., Chodari, L., Dizaj, S. M., Ardalan, M. R., Samiei, M., Sharifi, S., Vahed, S. Z., Huseynova, I., Khalilov, R., Ahmadian, E., & Cucchiarini, M. (2021). A comprehensive review of detection methods for SARS-CoV-2. *Microorganisms*, *9*(2), 1–18. https://doi.org/10.3390/microorganisms9020232

Fei Xiao, Jing Sun, Yonghao Xu, Fang Li, Xiaofang Huang, Heying Li, Jingxian Zhao, Jicheng

Huang, & Jincun Zhao. (2020). Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerging Infectious Diseases*, *26*(8), 1920–1922.

Fitzpatrick, A. H., Rupnik, A., O'Shea, H., Crispie, F., Keaveney, S., & Cotter, P. (2021). High Throughput Sequencing for the Detection and Characterization of RNA Viruses. *Frontiers in Microbiology*, *12*(February). https://doi.org/10.3389/fmicb.2021.621719

Foladori, P., Cutrupi, F., Segata, N., Manara, S., Pinto, F., Malpei, F., Bruni, L., & La Rosa, G. (2020). SARS-CoV-2 from faeces to wastewater treatment: What do we know? A review. *Science of the Total Environment*, *743*, 140444. https://doi.org/10.1016/j.scitotenv.2020.140444

Fongaro, G., Stoco, P. H., Souza, D. S. M., Grisard, E. C., Magri, M. E., Rogovski, P., Schörner, M. A., Barazzetti, F. H., Christoff, A. P., de Oliveira, L. F. V., Bazzo, M. L., Wagner, G., Hernández, M., & Rodríguez-Lázaro, D. (2021). The presence of SARS-CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. *Science of the Total Environment*, *778*, 1–4. https://doi.org/10.1016/j.scitotenv.2021.146198

Fout, G.S., Cashdollar, J.L., Varughese, E.A., Parshionikar, S.U., Grimm, A.C. (2015). EPA Method 1615. Measurement of enterovirus and norovirus occurrence in water by culture and RT-qPCR. I. Collection of virus samples. *J. Vis. Exp*. https://doi.org/10.3791/52067.

Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing*. 1–9. http://arxiv.org/abs/1207.3907

Gaudin, M., & Desnues, C. (2018). Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Frontiers in Microbiology*, *9*(NOV), 1–9. https://doi.org/10.3389/fmicb.2018.02924

Gawlik, B., Tavazzi, S., Mariani, G., Skejo, H., Sponar, M., Higgins, T., Medema, G., &

Wintgens, T. (2021). *SARS-CoV-2 Surveillance employing Sewage - Towards a Sentinel System*. https://doi.org/10.2760/300580

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

Gibas, C., Lambirth, K., Mittal, N., Juel, M., 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., & Munir, M. (2021). Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. *The Science of the Total Environment*, 782, 146749. https://doi.org/10.1016/j.scitotenv.2021.146749

Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., da Silva, M. F. M., Hundesa, A., Amorós, I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R., & Girones, R. (2017). Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. *Journal of Microbiological Methods*, *134*, 46–53. https://doi.org/10.1016/j.mimet.2017.01.006

Gonzalez, R. A., Larson, A., Thompson, H., Carter, E., & Cassi, X. F. (2021). Redesigning SARS-CoV-2 clinical RT-qPCR assays for wastewater RT-ddPCR. *MedRxiv*, 2021.03.02.21252754. https://doi.org/10.1101/2021.03.02.21252754

Graham, K. E., Loeb, S. K., Wolfe, M. K., Catoe, D., Sinnott-Armstrong, N., Kim, S., Yamahara, K. M., Sassoubre, L. M., Men-

doza Grijalva, L. M., Roldan-Hernandez, L., Langenfeld, K., Wigginton, K. R., & Boehm, A. B. (2021). SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with COVID-19 Cases in a Large Urban Sewershed. *Environmental Science & Technology*, 55(1), 488–498. https://doi.org/10.1021/acs.est.0c06191

Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., Tan, A. L., Paul, L. M., Brackney, D. E., Grewal, S., Gurfield, N., Van Rompay, K. K. A., Isern, S., Michael, S. F., Coffey, L. L., Loman, N. J., & Andersen, K. G. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biology*, *20*(1), 1–19. https://doi.org/10.1186/s13059-018-1618-7

Guerrero-Latorre, L., Ballesteros, I., Villacrés-Granda, I., Granda, M. G., Freire-Paspuel, B., & Ríos-Touma, B. (2020). SARS-CoV-2 in river water: Implications in low sanitation countries. *Science of the Total Environment*, *743*(February), 0–4. https://doi.org/10.1016/j.scitotenv.2020.140832

Habtewold, J., McCarthy, D., McBean, E., Law, I., Goodridge, L., Habash, M., & Murphy, H. M. (2022). Passive sampling, a practical method for wastewater-based surveillance of SARS-CoV-2. *Environmental Research*, 204. https://doi.org/10.1016/j.envres.2021.1120

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research*, *19*(7), 1141–1152. https://doi.org/10.1101/gr.085464.108

Hamouda, M., Mustafa, F., Maraqa, M., Rizvi, T., & Aly Hassan, A. (2021). Wastewater surveillance for SARS-CoV-2: Lessons learnt from recent studies to define future applications. *Science of the Total Environment*, *759*(xxxx), 143493.

https://doi.org/10.1016/j.scitotenv.2020.143493

Handelsman, J. (2005). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*, *69*(1), 195–195. https://doi.org/10.1128/mmbr.69.1.195.2005

Hansen, J. C. (1999). the Modern Instrument and. *Instrumentation*, *10*(4), 163–176.

Haramoto, E., Malla, B., Thakali, O., & Kitajima, M. (2020). First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. *Science of the Total Environment*, *737*, 140405. https://doi.org/10.1016/j.scitotenv.2020.140405

Hayes, E. K., Sweeney, C. L., Anderson, L. E., Li, B., Erjavec, G. B., Gouthro, M. T., Krkosek, W. H., Stoddart, A. K., & Gagnon, G. A. (2021). A novel passive sampling approach for SARS-CoV-2 in wastewater in a Canadian province with low prevalence of COVID-19. *Environmental Science: Water Research and Technology*, *7*(9), 1576–1586. https://doi.org/10.1039/d1ew00207d

Hayes, E. K., Sweeney, C. L., Fuller, M., Erjavec, G. B., Stoddart, A. K., & Gagnon, G. A. (2022). Operational Constraints of Detecting SARS-CoV-2 on Passive Samplers using Electronegative Filters: A Kinetic and Equilibrium Analysis. *ACS ES and T Water*, 2(11), 1910–1920. https://doi.org/10.1021/acsestwater.1c00441

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, *107*(1), 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003

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. https://doi.org/10.2166/wh.2011.019

Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T., & 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. https://doi.org/10.1128/AEM.01981-14

Hokajärvi, A. M., Rytkönen, A., Tiwari, A., Kauppinen, A., Oikarinen, S., Lehto, K. M., Kankaanpää, A., Gunnar, T., Al-Hello, H., Blomqvist, S., Miettinen, I. T., Savolainen-Kopra, C., & Pitkänen, T. (2021). The detection and stability of the SARS-CoV-2 RNA biomarkers in wastewater influent in Helsinki, Finland. *Science of the Total Environment*, *770*, 145274. https://doi.org/10.1016/j.scitotenv.2021.145274

Hosseini, M., Pratas, D., & Pinho, A. J. (2016). A survey on data compression methods for biological sequences. *Information (Switzerland)*, *7*(4). https://doi.org/10.3390/info7040056

Hrudey, S. E., & Conant, B. (2022). The devil is in the details: emerging insights on the relevance of wastewater surveillance for SARS-CoV-2 to public health. *Journal of Water and Health*, *20*(1), 246–270. https://doi.org/10.2166/WH.2021.186

Huang, Z., Hu, S., Xiong, Y., Wei, H., Xu, H., Duan, H., & Lai, W. (2019). Application and development of superparamagnetic nanoparticles in sample pretreatment and immunochromatographic assay. In TrAC - *Trends in Analytical Chemistry* (Vol. 114, pp. 151–170). Elsevier B.V. https://doi.org/10.1016/j.trac.2019.03.004

Huggett, J. F., Foy, C. A., Benes, V., Emslie, K., Garson, J. A., Haynes, R., Hellemans, J., Kubista, M., Mueller, R. D., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., Wittwer, C. T., & Bustin, S. A. (2013). The digital MIQE guidelines: Minimum information for publication of quantitative digital PCR experiments. *Clinical Chemistry*, *59*(6), 892–902. https://doi.org/10.1373/clinchem.2013.206375

Huggett, J. F., Novak, T., Garson, J. A., Green, C., Morris-Jones, S. D., Miller, R. F., & Zumla, A. (2008). Differential susceptibility of PCR reactions to inhibitors: An important and

unrecognised phenomenon. *BMC Research Notes*, *1*, 1–9. https://doi.org/10.1186/1756-0500-1-70

Hunt, R. H., East, J. E., Lanas, A., Malfertheiner, P., Satsangi, J., Scarpignato, C., & Webb, G. J. (2021). COVID-19 and Gastrointestinal Disease: Implications for the Gastroenterologist. *Digestive Diseases*, *39*(2), 119–139. https://doi.org/10.1159/000512152

Ikner, L. A., Gerba, C. P., & Bright, K. R. (2012). Concentration and Recovery of Viruses from Water: A Comprehensive Review. *Food and Environmental Virology*, *4*(2), 41–67. https://doi.org/10.1007/s12560-012-9080-2

Jiang, M., Wang, A. L. W., Be, N. A., Mulakken, N., Nelson, K. L., & Kantor, R. S. (2024). Evaluation of the Impact of Concentration and Extraction Methods on the Targeted Sequencing of Human Viruses from Wastewater. *Environmental Science and Technology*, 58(19), 8239–8250. https://doi.org/10.1021/acs.est.4c00580

Johnson, R., Muller, C. J. F., Ghoor, S., Louw, J., Archer, E., Surujlal-Naicker, S., Berkowitz, N., Volschenk, M., Bröcker, L. H. L., Wolfaardt, G., van der Walt, M., Mutshembele, A. M., Malema, S., Gelderblom, H. C., Mdhluli, M., Gray, G., Mathee, A., & Street, R. A. (2021). Qualitative and quantitative detection of SARS-CoV-2 RNA in untreated wastewater in Western Cape Province, South Africa. *South African Medical Journal*, *111*(3), 198–202. https://doi.org/10.7196/SAMJ.2021.V111I3.15154

Jones, D.L., Baluja, M.Q., Graham, D.W., Corbishley, A., McDonald, J.E., Malham, S.K., ... Wilcox, M.H. (2020). Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Science of The Total Environment* 749, 141364. https://doi.org/10.1016/j.scitotenv.2020.141364

Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, 609(7925), 101–108. https://doi.org/10.1038/s41586-022-05049-6

Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. *Cell*, *181*(4), 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011

Kittigul, L., Rupprom, K., Che-arsae, M., Pombubpa, K., Thongprachum, A., Hayakawa, S., & Ushijima, H. (2019). Occurrence of noroviruses in recycled water and sewage sludge: emergence of recombinant norovirus strains. *Journal of Applied Microbiology*, *126*(4), 1290–1301. https://doi.org/10.1111/jam.14201

Koboldt, D. C. (2020). Best practices for variant calling in clinical sequencing. *Genome Medicine*, *12*(1), 1–13. https://doi.org/10.1186/s13073-020-00791-w

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., Weinstock, G. M., Wilson, R. K., & Ding, L. (2009). VarScan: Variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, *25*(17), 2283–2285. https://doi.org/10.1093/bioinformatics/btp373

Kumar, M., Patel, A. K., Shah, A. V., Raval, J., Rajpara, N., Joshi, M., & Joshi, C. G. (2020). First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. *Science of the Total Environment*, *746*, 141326. https://doi.org/10.1016/j.scitotenv.2020.141326

La Rosa, G., Mancini, P., Bonanno Ferraro, G., Veneri, C., Iaconelli, M., Bonadonna, L., Lucentini, L., & Suffredini, E. (2021). SARS-CoV-2 has been circulating in northern Italy

since December 2019: Evidence from environmental monitoring. *Science of the Total Environment*, *750* 141711. https://doi.org/10.1016/j.scitotenv.2020.141711

Lambisia, A. W., Mohammed, K. S., Makori, T. O., Ndwiga, L., Mburu, M. W., Morobe, J. M., Moraa, E. O., Musyoki, J., Murunga, N., Mwangi, J. N., Nokes, D. J., Agoti, C. N., Ochola-Oyier, L. I., & Githinji, G. (2022). Optimization of the SARS-CoV-2 ARTIC Network V4 Primers and Whole Genome Sequencing Protocol. *Frontiers in Medicine*, *9*(February), 1–8. https://doi.org/10.3389/fmed.2022.836728

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Leipzig, J. (2017). A review of bioinformatic pipeline frameworks. *Briefings in Bioinformatics*, *18*(3), 530–536. https://doi.org/10.1093/bib/bbw020

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

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

Li, S., Yuan, L., Dai, G., Chen, R. A., Liu, D. X., & Fung, T. S. (2020). Regulation of the ER Stress Response by the Ion Channel Activity of the Infectious Bronchitis Coronavirus Envelope Protein Modulates Virion Release, Apoptosis, Viral Fitness, and Pathogenesis. *Frontiers in Microbiology*, *10*(January), 1–14. https://doi.org/10.3389/fmicb.2019.03022

Li, J., Verhagen, R., Ahmed, W., Metcalfe, S., Thai, P. K., Kaserzon, S. L., Smith, W. J. M., Schang, C., Simpson, S. L., Thomas, K. V., Mueller, J. F., & Mccarthy, D. (2022). In Situ Calibration of Passive Samplers for Viruses in Wastewater. *ACS ES and T Water*, 2(11), 1881–1890. https://doi.org/10.1021/acsestwater.1c00406

Liu, P., Ibaraki, M., VanTassell, J., Geith, K., Cavallo, M., Kann, R., & Moe, C. (2020). A sensitive, simple, and low-cost method for COVID-19 wastewater surveillance at an institutional level. *Sci Total Environ.* 10;807(Pt 3):151047. doi: 10.1016/j.scitotenv.2021.151047.

Liu, Y., Sapoval, N., Gallego-García, P., Tomás, L., Posada, D., Treangen, T. J., & Stadler, L. B. (2024). Crykey: Rapid identification of SARS-CoV-2 cryptic mutations in wastewater. *Nature Communications*, 15(1). https://doi.org/10.1038/s41467-024-48334-w

Louten, J. (2016). Virus Transmission and Epidemiology. *Essential Human Virology*, 71–92. https://doi.org/10.1016/b978-0-12-800947-5.00005-3

Lu, D., Huang, Z., Luo, J., Zhang, X., & Sha, S. (2020a). Primary concentration – The critical step in implementing the wastewater based epidemiology for the COVID-19 pandemic: A mini-review. *Science of the Total Environment*, *747*. https://doi.org/10.1016/j.scitotenv.2020.141245

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., … Tan, W. (2020b). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, *395*(10224), 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8

Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine*, *40*(4), 346–358. https://doi.org/10.1055/s-0038-1634431

Mao, K., Zhang, K., Du, W., Ali, W., Feng, X., & Zhang, H. (2020). The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Current Opinion in Environmental Science and Health*, *17*, 1–7.

https://doi.org/10.1016/j.coesh.2020.04.006

Maurano, M. T., Ramaswami, S., Zappile, P., Dimartino, D., Boytard, L., Ribeiro-Dos-Santos, A. M., Vulpescu, N. A., Westby, G., Shen, G., Feng, X., Hogan, M. S., Ragonnet-Cronin, M., Geidelberg, L., Marier, C., Meyn, P., Zhang, Y., Cadley, J., Ordoñez, R., Luther, R., … Heguy, A. (2020). Sequencing identifies multiple early introductions of SARS-CoV-2 to the New York City region. *Genome Research*, *31*(12), 1781–1788. https://doi.org/10.1101/gr.266676.120

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

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 and Technology Letters*, *7*(7), 511–516. https://doi.org/10.1021/acs.estlett.0c00357

Michael-Kordatou, I., Karaolia, P., & Fatta-Kassinos, D. (2020). Sewage analysis as a tool for the COVID-19 pandemic response and management: The urgent need for optimised protocols for SARS-CoV-2 detection and quantification. *Journal of Environmental Chemical Engineering*, *8*(5), 104306. https://doi.org/10.1016/j.jece.2020.104306

Mousazadeh, M., Ashoori, R., Paital, B., Kabdaşlı, I., Frontistis, Z., Hashemi, M., Sandoval, M. A., Sherchan, S., Das, K., & Emamjomeh, M. M. (2021). Wastewater Based Epidemiology Perspective as a Faster Protocol for Detecting Coronavirus RNA in Human Populations: A Review with Specific Reference to SARS-CoV-2 Virus. *Pathogens*, *10*(8), 1008. https://doi.org/10.3390/pathogens10081008

Nagy-Szakal, D., Couto-Rodriguez, M., Wells, H. L., Barrows, J. E., Debieu, M., Butcher, K., Chen, S., Berki, A., Hager, C., Boorstein, R. J., Taylor, M. K., Jonsson, C. B., Mason, C. E., & O'Hara, N. B. (2021). Targeted Hybridization Capture of SARS-CoV-2 and Metagenomics Enables Genetic Variant Discovery and Nasal Microbiome Insights. *Microbiology Spectrum*, *9*(2). https://doi.org/10.1128/spectrum.00197-21

Naughton, C. C., Roman, F. A., Grace Alvarado, A. F., Tariqi, A. Q., Deeming, M. A., Bibby, K., Bivins, A., Rose, J. B., Medema, G., Ahmed, W., Katsivelis, P., Allan, V., Sinclair, R., Zhang, Y., Kinyua, M. N., & Author cnaughton, C. (2024). Show us the Data: Global COVID-19 Wastewater Monitoring Efforts, Equity, and Gaps. *MedRxiv*, 2021.03.14.21253564

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K. K., Wilkinson, R., & Wiedenheft, B. (2020). Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Reports Medicine*, *1*(6), 100098. https://doi.org/10.1016/j.xcrm.2020.100098

Nigam, M. O., Bansal, B., & Chen, X. D. (2008). Fouling and cleaning of whey protein concentrate fouled ultrafiltration membranes. *Desalination*, *218*(1–3), 313–322. https://doi.org/10.1016/j.desal.2007.02.027

Parra-Guardado, A. L., Sweeney, C. L., Hayes, E. K., Trueman, B. F., Huang, Y., Jamieson, R. C., Rand, J. L., Gagnon, G. A., & Stoddart, A. K. (2022). Development of a rapid pre-concentration protocol and a magnetic beads-based RNA extraction method for SARS-CoV-2 detection in raw municipal wastewater. *Environmental Science: Water Research and Technology*, 8(1), 47–61. https://doi.org/10.1039/d1ew00539a

Park, M., Lee, J. H., Shin, H., Kim, M., Choi, J., Kang, D. H., Heu, S., & Ryu, S. (2012). Characterization and comparative genomic analysis of a novel bacteriophage, SFP10,

simultaneously inhibiting both Salmonella enterica and Escherichia coli O157:H7. *Applied and Environmental Microbiology*, *78*(1), 58–69. https://doi.org/10.1128/AEM.06231-11

Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: Quick coverage calculation for genomes and exomes. *Bioinformatics*, *34*(5), 867–868. https://doi.org/10.1093/bioinformatics/btx699

Philo, S. E., Keim, E. K., Swanstrom, R., Ong, A. Q. W., Burnor, E. A., Kossik, A. L., Harrison, J. C., Demeke, B. A., Zhou, N. A., Beck, N. K., Shirai, J. H., & Meschke, J. S. (2021). A comparison of SARS-CoV-2 wastewater concentration methods for environmental surveillance. *Science of the Total Environment*, *760*, 144215. https://doi.org/10.1016/j.scitotenv.2020.144215

Pinheiro, L. B., Coleman, V. A., Hindson, C. M., Herrmann, J., Hindson, B. J., Bhat, S., & Emslie, K. R. (2012). Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical Chemistry*, *84*(2), 1003–1011. https://doi.org/10.1021/ac202578x

Porta M. (2008). *A Dictionary of Epidemiology*. 5th edn. New York, NY: Oxford University Press, 2008.

Prjibelski, A. D., Korobeynikov, A. I., & Lapidus, A. L. (2018). Sequence analysis. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (Vols. 1–3). https://doi.org/10.1016/B978-0-12-809633-8.20106-4

Quick, J., Barnes, K., & Quick, J. (2020). nCoV-2019 sequencing protocol. In *ARTIC Coronavirus Method Devolopment community* (pp. 1–24). https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w%0Ahttps://artic.network/

Rafiee, M., Isazadeh, S., Mohseni-Bandpei, A., Mohebbi, S. R., Jahangiri-rad, M., Eslami, A., Dabiri, H., Roostaei, K., Tanhaei, M., & Amereh, F. (2021). Moore swab performs equal to composite and outperforms grab sampling for SARS-CoV-2 monitoring in wastewater. *Science of the Total Environment, 790*. https://doi.org/10.1016/j.scitotenv.2021.148205

Rahimi, A., Mirzazadeh, A., & Tavakolpour, S. (2021). Genetics and genomics of SARS-CoV-2: A review of the literature with the special focus on genetic diversity and SARS-CoV-2 genome detection. *Genomics*, *113*(1P2), 1221–1232. https://doi.org/10.1016/j.ygeno.2020.09.059

Rajal, V. B., McSwain, B. S., Thompson, D. E., Leutenegger, C. M., Kildare, B. J., & Wuertz, S. (2007). Validation of hollow fiber ultrafiltration and real-time PCR using bacteriophage PP7 as surrogate for the quantification of viruses from water samples. *Water Research*, *41*(7), 1411–1422. https://doi.org/10.1016/j.watres.2006.12.034

Randazzo, W., Cuevas-Ferrando, E., Sanjuán, R., Domingo-Calap, P., & Sánchez, G. (2020). Metropolitan wastewater analysis for COVID-19 epidemiological surveillance. *International Journal of Hygiene and Environmental Health*, *230*(September). https://doi.org/10.1016/j.ijheh.2020.113621

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., & Sánchez, G. (2020b). SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Research, 181*. https://doi.org/10.1016/j.watres.2020.115942

Rastogi, M., Pandey, N., Shukla, A., & Singh, S. K. (2020). SARS coronavirus 2: from genome to infectome. *Respiratory Research*, *21*(1), 1–15. https://doi.org/10.1186/s12931-020-01581-z

Rehman SU, Shafique L, Ihsan A, L. Q. (2020). Evolutionary Trajectory for the Emergence of Novel. *Pathogens*, *2*(Figure 1), 1–12.

Roy, S., Coldren, C., Karunamurthy, A., Kip, N. S., Klee, E. W., Lincoln, S. E., Leon, A., Pullambhatla, M., Temple-Smolkin, R. L., Voelkerding, K. V., Wang, C., & Carter, A. B. (2018). Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *Journal of Molecular Diagnostics*, *20*(1), 4–27. https://doi.org/10.1016/j.jmoldx.2017.11.003

Schang, C., Crosbie, N. D., Nolan, M., Poon, R., Wang, M., Jex, A., John, N., Baker, L., Scales, P., Schmidt, J., Thorley, B. R., Hill, K., Zamyadi, A., Tseng, C. W., Henry, R., Kolotelo, P., Langeveld, J., Schilperoort, R., Shi, B., … McCarthy, D. T. (2021). Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. *Environmental Science and Technology*, 55(15), 10432–10441. https://doi.org/10.1021/acs.est.1c01530

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

Sears, S. D., Ferreccio, C., Levine, M. M., Cordano, A. M., Monreal, J., Black, R. E., D'ottone, K., Rowe, B., & Committee, C. T. (1984). The Use of Moore Swabs for Isolation of Salmonella typhl from Irrigation Water in Santiago, Chile. In the *journal of infectious diseases*, (Vol. 149, Issue 4). https://academic.oup.com/jid/article/149/4/640/1188322

Sherchan, S. P., Shahin, S., Ward, L. M., Tandukar, S., Aw, T. G., Schmitz, B., Ahmed, W., & Kitajima, M. (2020). First detection of SARS-CoV-2 RNA in wastewater in North America: A study in Louisiana, USA. *Science of the Total Environment*, *743*, 140621. https://doi.org/10.1016/j.scitotenv.2020.140621

Sim, S. L., He, T., Tscheliessnig, A., Mueller, M., Tan, R. B. H., & Jungbauer, A. (2012).

Protein precipitation by polyethylene glycol: A generalized model based on hydrodynamic radius. *Journal of Biotechnology*, *157*(2), 315–319. https://doi.org/10.1016/j.jbiotec.2011.09.028

Simon Andrews. (2020). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. In *Soil* (Vol. 5, Issue 1, pp. 47–81). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Stephenson, F. H. (2016). Centrifugation. In: Stephenson, F H. *Calculations for Molecular Biology and Biotechnology*. Academic Press, United Sates, pp. 431–438.

Tang, T., Bidon, M., Jaimes, J. A., Whittaker, G. R., & Daniel, S. (2020). Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Research*, *178*(April), 104792. https://doi.org/10.1016/j.antiviral.2020.104792

Tavares, L., Alves, P. M., Ferreira, R. B., & Santos, C. N. (2011). Comparison of different methods for DNA-free RNA isolation from SK-N-MC neuroblastoma. *BMC Research Notes*, *4*, 3. https://doi.org/10.1186/1756-0500-4-3

Thacker S. B. (2002). Historical development. In: Teutsch SM, Churchill RE, editors. *Principles and practice of public health surveillance*, 2nd ed. New York: Oxford University Press; 2002. p. 1–16.

Thompson, J. R., Nancharaiah, Y. V., Gu, X., Lee, W. L., Rajal, V. B., Haines, M. B., Girones, R., Ng, L. C., Alm, E. J., & Wuertz, S. (2020). Making waves: Wastewater surveillance of SARS-CoV-2 for population-based health management. *Water Research*, *184*. https://doi.org/10.1016/j.watres.2020.116181

Vancollie, V. (2021). *What is Genomic Surveillance? | Facts | yourgenome.org*. https://www.yourgenome.org/facts/what-is-genomic-surveillance

Vincent-Hubert, F., Wacrenier, C., Desdouits, M., Jousse, S., Schaeffer, J., Le Mehaute, P.,

Nakache-Danglot, F., Le Guyader, F. S., Bertrand, I., Boni, M., Gantzer, C., Maday, Y., Marechal, V., Mouchel, J. M., Moulin, L., & Wurtzer, S. (2022). Development of passive samplers for the detection of SARS-CoV-2 in sewage and seawater: Application for the monitoring of sewage. *Science of the Total Environment,* 833. https://doi.org/10.1016/j.scitotenv.2022.155139

Vo, V., Tillett, R. L., Chang, C. L., Gerrity, D., Betancourt, W. Q., & Oh, E. C. (2022). SARS-CoV-2 variant detection at a university dormitory using wastewater genomic tools. *Science of the Total Environment*, *805*, 149930. https://doi.org/10.1016/j.scitotenv.2021.149930

Wigginton, K. R., Ye, Y., & Ellenberg, R. M. (2015). Emerging investigators series: The source and fate of pandemic viruses in the urban water cycle. *Environmental Science: Water Research and Technology*, *1*(6), 735–746. https://doi.org/10.1039/c5ew00125k

Wilm, A., Aw, P. P. K., Bertrand, D., Yeo, G. H. T., Ong, S. H., Wong, C. H., Khor, C. C., Petric, R., Hibberd, M. L., & Nagarajan, N. (2012). LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research*, *40*(22), 11189–11201. https://doi.org/10.1093/nar/gks918

Wilson, M., Qiu, Y., Yu, J., Lee, B. E., McCarthy, D. T., & Pang, X. (2022). Comparison of Auto Sampling and Passive Sampling Methods for SARS-CoV-2 Detection in Wastewater. *Pathogens*, 11(3). https://doi.org/10.3390/pathogens11030359

World Health Organization. (2021). Tracking SARS-CoV-2 variants. URL: https://www.who.int/en/activities/tracking-SARS-Co).

World Health Organization (2003). Guidelines for environmental surveillance of poliovirus circulation Guidelines for environmental. *Vaccine and Biologicals*, *03*(03), 28. URL:

https://apps.who.int/iris/bitstream/handle/10665/67854/WHO

Wu, Fan, Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., Hu, Y., Tao, Z. W., Tian, J. H., Pei, Y. Y., Yuan, M. L., Zhang, Y. L., Dai, F. H., Liu, Y., Wang, Q. M., Zheng, J. J., Xu, L., Holmes, E. C., & Zhang, Y. Z. (2020a). A new coronavirus associated with human respiratory disease in China. *Nature*, *579*(7798), 265–269. https://doi.org/10.1038/s41586-020-2008-3

Wu, Fuqing, Zhang, J., Xiao, A., Gu, X., Lee, L., Armas, F., & Kauffman, K. (2020b). SARS-CoV-2 Titers in Wastewater Are Higher than Expected. *MSystems*, *5*(July), 1–9.

Xiao, M., Xiao, M., Liu, X., Ji, J., Ji, J., Ji, J., Li, M., Li, M., Li, M., Li, J., Li, J., Li, J., Yang, L., Sun, W., Sun, W., Sun, W., Ren, P., Ren, P., Yang, G., … Li, J. (2020). Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Medicine*, *12*(1), 1–15. https://doi.org/10.1186/s13073-020-00751-4

Zhang, H., Wang, X., Huang, Y., & Gou, X. (2020). Day-Ahead Stochastic Scheduling Model Considering Market Transactions in Multi-Energy System. *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, *2020-Septe*. https://doi.org/10.1109/APPEEC48164.2020.9220412

Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. Di, Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., … Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zhou, H. X., and Xiaodong Pang, X. (2018). Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem Rev.* 118(4): 1691–1741.

3 Chapter 3: Wastewater Solids Drive Comparability of Sampling Methods for SARS-Cov-2 Wastewater and Environmental Surveillance

3.1 Context

Chapter 3 represents the published research article entitled *Wastewater solids drive comparability of sampling methods for SARS-CoV-2 wastewater and environmental surveillance* by Kabir, M. P., Renouf, E., Pisharody, L., Mercier, E., D'Aoust, P. M., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Tomalty, E., Graber, T. E., and Delatolla, R. (Journal of Environmental Chemical Engineering, 2025). This study explores the SARS-CoV-2 RNA concentrations in wastewater solids collected by autosampler, passive samplers, and primary sludge and determines the influences of solids content on SARS-CoV-2 RNA measurements.

3.2 Abstract

Passive sampling has proven to be a reliable and cost-effective method in wastewater and environmental monitoring (WEM) during the COVID-19 pandemic. In passive sampling, wastewater solids were collected from the wastewater networks, homogenized in solutes and analyzed the supernatant to measure the SARS-CoV-2 RNA concentrations in wastewater. However, the direct impact of wastewater solids content collected via passive sampling on SARS-CoV-2 RNA measurements has not been previously evaluated. In this study, we analyzed wastewater solids collected using Auto, Torpedo, COSCa-ball samplers, and primary sludge samples from a wastewater treatment plant to measure SARS-CoV-2 RNA concentrations in wastewater. Results showed significant variation ($p < 0.05$) in wastewater solids content (i.e., TS and VS) across Auto, Torpedo, COSCa-ball samplers, and primary sludge samples. Despite differences in solids content, SARS-CoV-2 RNA concentrations in wastewater solids from passive samplers can be effectively compared ($p > 0.05$) to autosampler

and primary sludge samples. To evaluate the influences of wastewater solids content on SARS-CoV-2 RNA measurement, we used a linear mixed-effects model. The model demonstrated that wastewater solids content had no direct effect on SARS-CoV-2 RNA measurements across the sampling methods and primary sludge samples. Overall, this study established a standardized experimental approach for implementing passive samplers as a viable alternative to conventional autosampler in WEM for emerging pathogens.

3.3 Introduction

The COVID-19 pandemic has highlighted the necessity of fast, accurate, and sensitive techniques for early detection of circulating pathogens in communities (Amman et al., 2022; Carter et al., 2020). The surveillance of COVID-19 disease predominantly depends on the clinical testing of symptomatic patients, but the surveillance system can't differentiate the prevalence of asymptomatic and pre-symptomatic infections (Wang et al., 2022; Sah et al., 2021). However, the infectious agent of COVID-19 disease; SARS-CoV-2 RNA excretes through saliva, feces, and urine of both symptomatic and asymptomatic patients, which underlies the use of WEM as an alternative approach for monitoring COVID-19 disease in the communities (Ahmed et al., 2020; Medema et al., 2020). The use of WEM was previously limited to illicit drugs as well as few gastrointestinal infections, but the COVID-19 pandemic saw the widespread adoption of WEM for respiratory illnesses such as SARS-CoV-2, Influenza, RSV (Jiang et al., 2023; Kisand et al., 2023; Mercier et al., 2023; Mercier et al., 2022; Melvin et al., 2021). Over the course of the COVID-19 pandemic, WEM has been extended to 4648 sites in 72 countries to monitor the presence of numerous viruses, their occurrence, provide early warnings, and track circulating variants (Naughton et al., 2024).

The emergence of WEM has prompted the development of several wastewater sampling methods. The two most common sampling methods used in WEM are composite (time or flow-proportional) and grab sampling. Composite samples are often collected using an Autosampler

over the course of 24 hours while a grab sample is taken at a certain time of the day, usually in the morning. However, implementation of an Autosampler for composite sampling is labor intensive, expensive, and challenging at geo-specific locations, while grab sample is prone to intraday variation, making it less representative to community-level disease prevalence (Augusto et al., 2022; Harris-Lovett et al., 2021; Kitajima et al., 2020). As an alternative to conventional auto and grab sampling, passive sampling has been widely adopted as practical, affordable, and reliable sampling method for SARS-CoV-2 WEM (Mejías-Molina et al., 2023; Habtewold et al., 2022; Kevill et al., 2022). In passive sampling, sampling mediums (i.e., medical gauzes, cheesecloth, cellulose sponges, or electronegative membrane filters) are packed within styrene-like plastic devices and deployed in wastewater networks to sorb wastewater solids over time (Bivins et al., 2022). Wastewaters are comprised of a variety of suspended materials and biosolids that are likely to compete with SARS-CoV-2 viral particles for sorption in the medium. Therefore, characteristics of the captured solids on passive samplers may affect SARS-CoV-2 RNA measurements, but this association is not yet well comprehended. Most of the earlier research (Breulmann et al., 2023; Cha et al., 2023; Mejías-Molina et al., 2023; Habtewold et al., 2022) homogenized collected wastewater solids into deionized water or soluble buffer, measured SARS-CoV-2 RNA, along with other disease targets using the solute, and reported as copies/mL or copies/L. The SARS-CoV-2 RNA are partitions favorably to the solids fraction (Espinosa et al., 2022), and therefore, directly analyzing solids from passive samplers could yield higher concentrations compared to the solute. Besides, the SARS-CoV-2 RNA was found in higher concentrations in settled solids (i.e., primary sludge) compared to influent wastewater (Kim et al., 2022; D'Aoust et al., 2021a). However, it has yet to be determined whether the SARS-CoV-2 RNA concentrations in solids collected using passive samplers can be effectively compared to composite wastewater and primary sludge samples.

In this study, we investigated the wastewater solids collected using Auto, Torpedo, and COSCa-ball samplers, as well as primary sludge samples to quantify SARS-CoV-2 RNA and evaluate the influence of wastewater solids content on SARS-CoV-2 RNA measurement. The study demonstrated the capability of passive samplers for collecting wastewater solids and ensure their applicability in WEM for pandemic preparedness.

3.4 Materials and Methods

3.4.1 Wastewater sample collection, processing, and nucleic acid extraction

Twenty-one paired influent wastewater and primary sludge samples were collected from Robert O. Pickard Environmental Center (ROPEC), Ottawa, Canada from March 13 to May 4, 2023. On the same day, wastewater samples were collected using an Autosampler (n=21) (Teledyne ISCO, Lincoln, NE, USA), a Torpedo passive (n=21), and a COSCa-ball passive sampler (n=21), while primary sludge samples (n=21) were collected using only an Autosampler. The Torpedo and COSCa-ball samplers were printed with polylactic acid (PLA) filament in an Ultimaker 2+ 3-D printer (Ultimaker, New York, USA) with multiple entry points (Figure 3-1). The Torpedo sampler was packed with two medical gauzes (Swisspers, China) and the COSCa-ball sampler was packed with four medical gauzes (Swisspers) to capture solid particles from wastewater. Passive samplers were deployed for a period of 24 hours and the Autosampler was programmed to collect 24-hour's time and volume weighted wastewater and primary sludge samples. All the samples were transferred to the laboratory in ice cooler packs and analyzed within 24 hours of sampling.

The medical gauzes used in the passive sampling were immediately removed, placed in 500 mL of deionized water, and vigorously agitated with a glass rod to recover the sorbate. Thereafter, medical gauzes were squeezed and discarded, and the suspension was allowed to settle for 30 minutes. Similarly, 24-hours composite wastewater samples collected using

Autosampler were allowed to settle for 30 minutes. The supernatant from the settled samples was then decanted carefully without losing the settled solids. Next, 40 mL of the suspension containing settled solids and primary sludge were centrifuged for 45 min at 10,000 × g at 4 °C, and 250 mg of the pellet was subsequently used to extract 100 µL viral RNA using the RNeasy Power Microbiome Kit (Qiagen, Germantown, USA) on a QIACube connect Automated extraction platform as previously described (D'Aoust et al., 2021b).

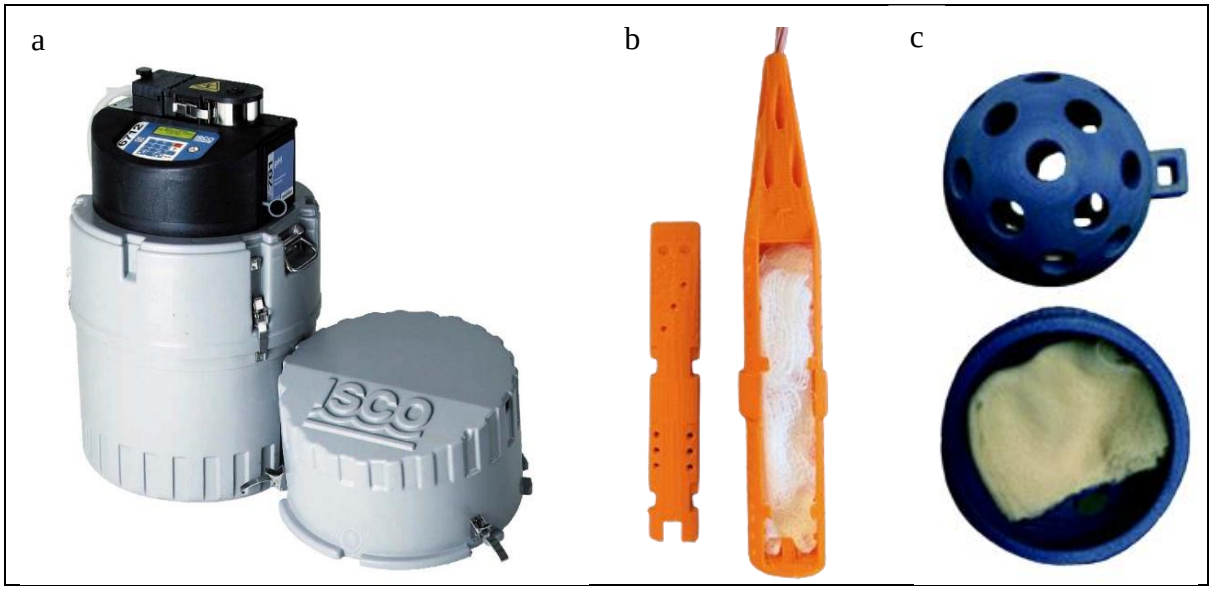

Figure 3-1: Schematics of different types of samplers used in this study to collect wastewater and primary sludge samples from ROPEC; a) conventional Autosampler used to collect 24-hours composite wastewater and primary sludge samples, b) Torpedo sampler packed with two medical gauzes, and c) COSCa-ball sampler with four medical gauzes used to collect wastewater solids.

3.4.2 Reverse transcription (RT) qPCR and SARS-CoV-2 RNA quantification

The presence of SARS-CoV-2 N1 and N2 targets and pepper mild mottle virus (PMMoV) were confirmed using singleplex one-step RT-qPCR with TaqManTM 1-Step Fast Virus Master Mix (ABI) (Thermo-Fisher, Ottawa, Canada), on a CFX Connect qPCR thermocycler (Bio-Rad, Hercules, Canada) according to the established protocol (D'Aoust et al., 2022). Briefly, 7 µL

of master mix consisting of 2.5 µL of 4X TaqMan® Fast Virus 1-step mastermix (Thermo-Fisher), 0.75 µL of 500 µM primer, and 125 µM probe mixture (IDT, Kanata, Canada), and 3.75 µL of nuclease-free water was mixed with 3 µL of RNA template. The PCR conditions were as follows: RT at 50 °C for 5 min, RT inactivation and initial denaturation at 95 °C for 20 s, 45 cycles of denaturation at 95 °C for 3 s and annealing/ extension at 60 °C for 30 s. Samples were run in triplicates with RNAse-free water as non-template control and quantified using a five-point gradient of the EDX SARS-CoV-2 standard curve (Exact Diagnostics, Texas, USA). The assay's limit of detection (ALOD) for the N1 target was approximately 2 copies/reaction, with a limit of quantification (ALOQ) of 3.2 copies/reaction. For the N2 target, the ALOD was around 2 copies/reaction, and the ALOQ was 8.1 copies/reaction. The SARS-CoV-2 RNA concentrations were normalized against quantified copies of PMMoV of each sample. All qPCR assays were performed according to the MIQE guidelines (Bustin et al., 2009).

3.4.3 Characterization of wastewater solids

The wastewater solids content such as total solids (TS) and volatile solids (VS) were analyzed using the United States Environmental Protection Agency (USEPA) standard protocol (EPA, 2001). Briefly, 250 mg pellet was transferred to a pre-weighted aluminum weighing dish (Fisherbrand, Ottawa, Canada) in triplicate. The dishes were placed into a 105 °C oven (Thermo-Fisher) for 3 hours, and afterward the incubation, they were weighed to determine the TS content. Later, the dishes were placed into a muffle furnace (Thermo-Fisher) operating at 550 °C for 30 minutes, and the solids lost to ignition were determined as VS content.

3.4.4 Linear mixed-effects model

A linear mixed-effects model (Bates et al., 2015) was employed to evaluate the effect of solids content (TS, VS) on SARS-CoV-2 RNA measurements, with log-transformed RNA

concentrations as the dependent variable and fixed effects for TS and VS. This approach enabled us to include samples with missing values at one or more time points, providing a more comprehensive analysis.

3.4.5 Statistical analysis

Wastewater solids content as well as average SARS-CoV-2 RNA concentrations were compared among Auto, Torpedo, COSCa-ball samplers, and primary sludge samples using one-way analysis of variance (ANOVA). An F-test was conducted to assess the daily variation in SARS-CoV-2 N1 and N2 measurements in solids across Auto, Torpedo, COSCa-ball sampler, and primary sludge samples. In our statistical analysis, p-values less than 0.05 were considered statistically significant. All statistical analyses and visualizations were performed using GraphPad Prism version 10.2.1 (La Jolla, California, USA), and R software packages (Version 4.3.3).

3.5 Results and Discussion

3.5.1 Wastewater solids content

The settled solids in influent wastewater collected using Auto, Torpedo, COSCa-ball sampler and primary sludge samples exhibited distinct TS and VS contents (Figure 3.2). Primary sludge samples had the highest (ANOVA, $p < 0.05$) TS content ($4.06 \pm 0.86 \times 10^4$ mg/L) followed by the settled solids in COSCa-ball ($1.52 \pm 4.38 \times 10^4$ mg/L), Autosampler ($4.61 \pm 3.15 \times 10^3$ mg/L), and Torpedo sampler ($4.65 \pm 2.13 \times 10^3$ mg/L). Similar VS measurements (ANOVA, $p < 0.05$) were observed across the sampling methods. The VS content of the solids in primary sludge, COSCa-ball, Torpedo and Autosampler was $3.34 \pm 0.52 \times 10^4$ mg/L, $1.11 \pm 0.34 \times 10^4$ mg/L, $3.61 \pm 2.17 \times 10^3$ mg/L, and $3.41 \pm 1.49 \times 10^3$ mg/L respectively. These variations of TS and VS in the settled solids between Auto and passive samplers can be attributed to a variety of factors such as the adsorption capacity of the samplers as well as design and numbers of

adsorbent material used in the passive samplers. The Autosampler typically collects a wider variety of wastewater solids with different adsorption and absorption characteristics over a specified period including suspended colloids, organic and inorganic debris (George et al., 2022). The Torpedo sampler featured a printed housing unit with numerous small entry points while the COSCa-ball sampler was a hollowed sphere with multiple large entry points, allowing more suspended colloids, solids and debris to enter the COSCa-ball samplers. Additionally, the uses of four medical gauzes in the COSCa-ball sampler resulted in the formation of thick layer of solids on the gauzes with lower water content. On the other hand, primary sludge was the same influent wastewater that underwent a period of approximately of thirty hours of physical setting at the wastewater treatment plant, leading to sample with higher TS and VS content.

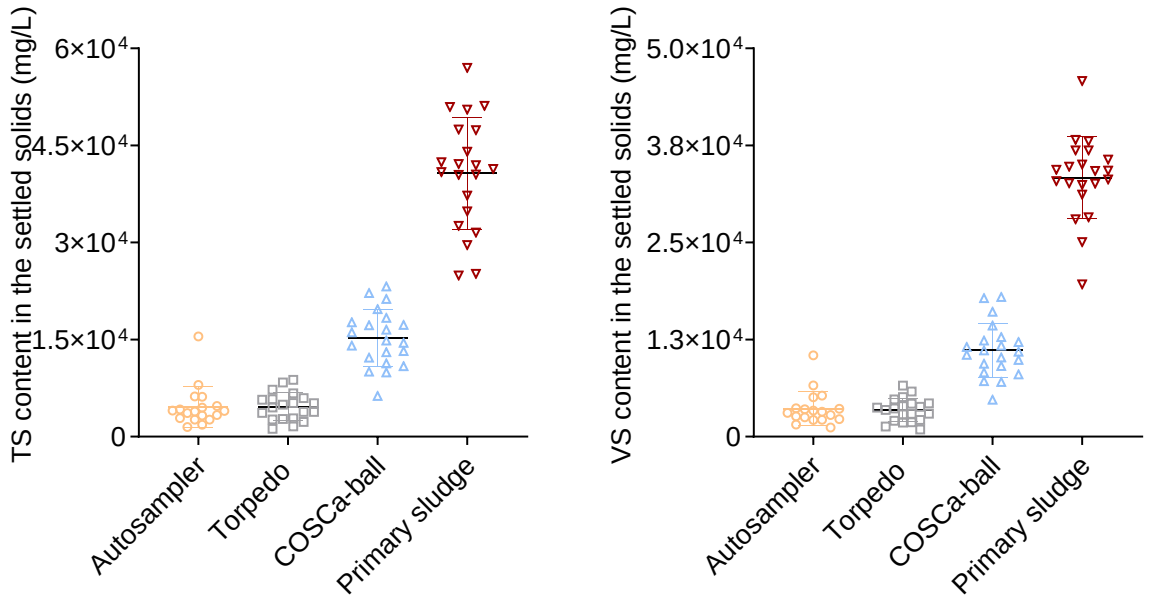

Figure 3-2: TS and VS content in the settled solids of influent wastewater collected using Auto, Torpedo, and COSCa-ball sampler and primary sludge samples. The error bars represent the standard deviation from the median, while the lower and upper data points indicate the minimum and maximum values within the data distribution, respectively.

3.5.2 Comparison of SARS-CoV-2 RNA concentrations in wastewater solids across various sampling methods and primary sludge

A total of twenty-one paired wastewater samples collected using Auto, Torpedo, and COSCa-ball samplers, along with primary sludge samples on corresponding days were analyzed between March 13 and May 4, 2023. Samples with Ct (Cycle threshold) less than 40 were considered positive, and SARS-CoV-2 RNA was quantified using the standard curve of N1 and N2 target assays. Results showed that SARS-CoV-2 RNA concentrations in solid mass fraction (copies/g) were comparable (ANOVA, $p > 0.05$) among Auto, Torpedo, COSCa-ball samplers, and primary sludge samples. The average concentrations of SARS-CoV-2 RNA in wastewater solids collected using Auto, Torpedo, and COSCa-ball sampler were $4.48 \pm 2.97\times10^3$, $6.68 \pm 4.02\times10^3$ and $5.61 \pm 3.51\times10^3$ (copies/g) respectively, while the SARS-CoV-2 RNA concentrations in primary sludge is $6.16 \pm 3.85\times10^3$ (copies/g) (Figure 3.3.A). The SARS-CoV-2 RNA concentrations in wastewater solids and primary sludge were normalized against PMMoV concentrations for each sample. Following PMMoV normalization, similar (ANOVA, $p > 0.05$) pattern in SARS-CoV-2 concentrations (copies/copies of PMMoV) was observed in wastewater across the various samplers and primary sludge samples (Figure 3.3.B). The average PMMoV normalized concentration of SARS-CoV-2 in wastewater solids collected using Auto, Torpedo, COSCa-ball and primary sludge was $6.11 \pm 5.11\times10^{-4}$, $7.55 \pm 3.6\times10^{-4}$, $6.69 \pm 5.91\times10^{-4}$ and $5.53 \pm 4.02\times10^{-4}$ copies/copies of PMMoV respectively.

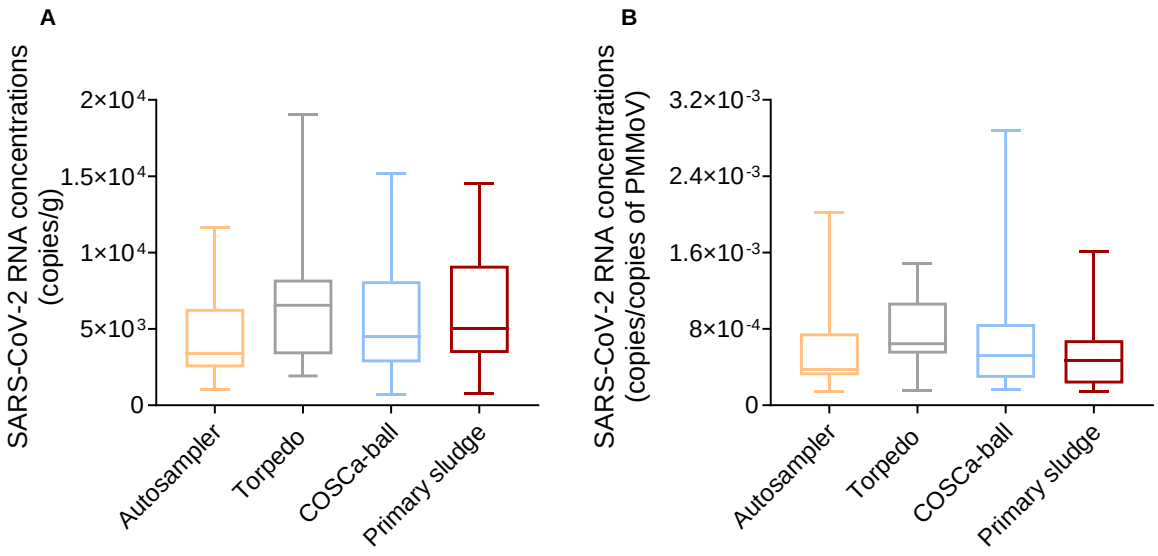

Figure 3-3: Comparison of SARS-CoV-2 RNA concentrations in wastewater solids collected using Auto, Torpedo, COSCa-ball and primary sludge samples. The box-and-whisker plots display the median (central line within the box), the 75th percentile (upper edge of the box), and the 25th percentile (lower edge of the box). The whiskers extend to the minimum and maximum values within the data distribution.

Although the average SARS-CoV-2 RNA concentrations were comparable across the sampling methods, we employed an F-test to evaluate the daily variation of the triplicate of SARS-CoV-2 N1 and N2 measurements across longitudinal time among Auto, Torpedo, COSCa-ball and primary sludge. This analysis enhances the accuracy of the sampling methods for quantifying SARS-CoV-2 RNA in wastewater solids. Results showed that approximately 67% of samples had similarity ($p > 0.05$) between Auto and Torpedo sampler as well as Autosampler and primary sludge samples, respectively, when SARS-CoV-2 was measured in copies/g (Figure 3.4.A). A similar pattern (~70%) observed between both Auto and Torpedo sampler as well as primary sludge samples in SARS-CoV-2 RNA when measured as copies/copies of PMMoV (Figure 3.4.B). Similarly, 76% exhibited similarity between Auto and COSCa-ball sampler, when SARS-CoV-2 RNA measured in (copies/g) or copies/copies of PMMoV. Overall, the above analyses suggested that quantifying SARS-CoV-2 RNA concentrations as solid mass

fraction or mass fraction PMMoV-normalized unit from passive samplers can be effectively compared to 24-hour composite wastewater and primary sludge sample.

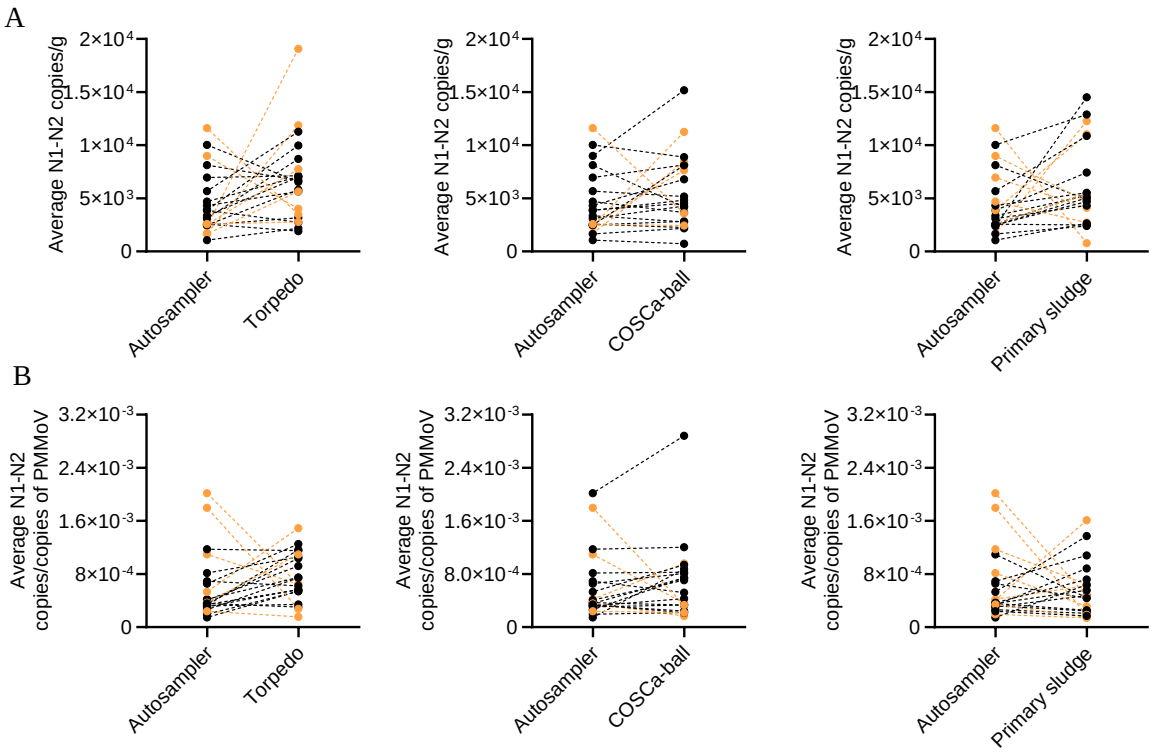

Figure 3-4: Comparison of daily SARS-CoV-2 N1 and N2 measurements in wastewater solids across a longitudinal data set from March 13, 2023, to May 4, 2023. An F-test was employed to evaluate the daily variation of SARS-CoV-2 RNA measurements among the samplers. Orange lines represent statistically significant ($p < 0.05$) variations in daily SARS-CoV-2 RNA measurements between the samples.

The study findings are consistent with prior research (Rafiee et al., 2021; Wilson et al., 2022), which demonstrated the comparable performance of passive samplers to Autosampler in SARS-CoV-2 RNA measurements in wastewater. However, aforementioned studies primarily focused on the supernatant from pre-processed passive samplers while the current study revealed that SARS-CoV-2 RNA measurements from the solids captured by passive samplers can also be effectively compared to conventional Auto sampling as well as primary sludge samples. The processing of solids from passive samplers offers significant advantages by

reducing both sample processing time and associated costs. Unlike liquid-based methods, solids-based processing involves a straightforward centrifugation step to separate solids before RNA extraction, streamlining the workflow without compromising data quality. However, the applicability of wastewater solids collected through passive samplers to other disease targets may be constrained by specific limitations, such as the presence of certain pathogens in the liquid phase or their tendency to exist in forms that do not readily adsorb onto wastewater solids. This inherent variability in pathogen adsorption limits the generalizability of passive sampling methods across diverse disease surveillance targets. Addressing these challenges will require further refinement of passive sampling techniques to enhance their versatility and expand their utility for monitoring a broader spectrum of pathogens beyond SARS-CoV-2.

3.5.3 Influence of solids content on SARS-CoV-2 RNA measurements across sampling methods

To evaluate the influence of wastewater solids content on SARS-CoV-2 RNA measurements, we employed linear mixed-effects model where solids content was treated as a fixed effect, enabling the model to assess how changes in TS and VS influence SARS-CoV-2 RNA measurements. By incorporating solids content as a fixed effect and accommodating variations across time points as random effects, the linear mixed-effects model can elucidate the impact of wastewater solids content on SARS-CoV-2 RNA measurements as well as the variations of impact across the sampling methods. Results showed that the TS and VS content of solids did not influence ($p > 0.05$) SARS-CoV-2 RNA measurements expressed as solid mass fraction or mass fraction PMMoV-normalized units across any sampling method or primary sludge (Figure 3.5). The study findings diverge from previous research (Amoah et al., 2022; Bitter et al., 2022), which reported a significant correlation between solids content (i.e., TS and VS) of wastewater collected by Autosampler and SARS-CoV-2 RNA concentrations. The earlier research processed the liquid portions of wastewater using ultrafiltration and correlated

wastewater solids content with SARS-CoV-2 RNA concentrations (copies/L). However, our research analyzed wastewater solids through direct centrifugation and therefore, we did not find any correlation with solids content and SARS-CoV-2 RNA concentrations, which suggested that association between wastewater solids (i.e., TS and VS) and SARS-CoV-2 RNA concentrations is largely dependent on the sample processing methods. Besides, the influence of solids content on SARS-CoV-2 measurements may vary across different wastewater treatment plants due to factors such as wastewater flow, wastewater constituents, and other environmental conditions (Amoah et al., 2022). In contrast, primary sludge represents a more settled, dense, and concentrated form of wastewater which provides a stable environment for viral particles. The elevated and uniform solids content facilitates effective adsorption of viral particles onto the dense solids, rendering measurements less sensitive to variations in the relative proportions of TS or VS. Consequently, the viral RNA in sludge is typically distributed uniformly across the solid matrix, effectively buffering against fluctuations in TS and VS content. This consistency and density create an optimal matrix for viral adsorption, ensuring that viral RNA is captured uniformly and minimizing the impact of solids content variability on RNA measurements.

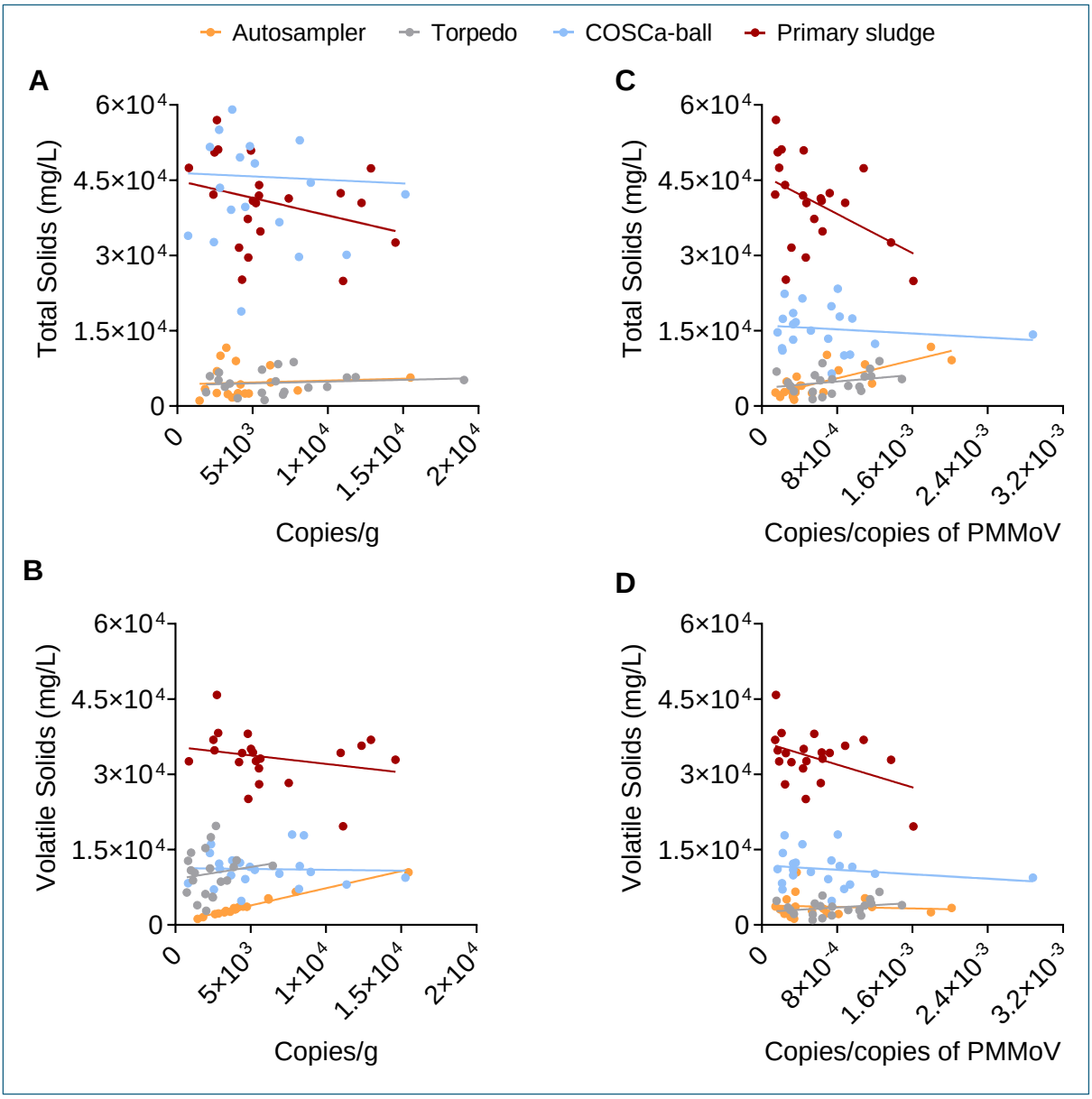

Figure 3-5: The linear mixed-effects model evaluated the influence of TS and VS content in wastewater solids collected by Auto, Torpedo, COSCa-ball samplers, and primary sludge samples on SARS-CoV-2 RNA measurements. The model reveals that TS and VS content of settled solids exert no significant influence ($p > 0.05$) on SARS-CoV-2 RNA measurement.

3.6 Conclusion

WEM is an essential tool for monitoring the prevalence and spread of COVID-19 diseases across communities. Even though WEM has been proven to be an efficient tool to alert healthcare authorities and policymakers, further optimization, and validation of sampling methods for data management are still required. In this study, we parallelly deployed autosampler and two types of passive samplers to collect wastewater solids from the inlet of a wastewater treatment plant, while primary sludge samples were collected using the autosampler for quantifying SARS-CoV-2 RNA. Results showed wastewater solids content varied among auto, Torpedo, COSCa-ball samplers and primary sludge samples. The results also suggested that SARS-CoV-2 RNA concentrations in wastewater solids obtained from passive samplers can be effectively compared to 24-hour composite wastewater and primary sludge samples. Finally, results demonstrated that wastewater solids content had no influence on SARS-CoV-2 RNA measurements. Overall, the study showed the capability of passive samplers to capture wastewater solids and demonstrated their adaptability for addressing emerging public health emergencies.

3.7 References

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

Amman, F., Markt, R., Endler, L., Hupfauf, S., Agerer, B., Schedl, A., Richter, L., Zechmeister, M., Bicher, M., Heiler, G., Triska, P., Thornton, M., Penz, T., Senekowitsch, M., Laine,

J., Keszei, Z., Klimek, P., Nägele, F., Mayr, M., … Bergthaler, A. (2022). Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nature Biotechnology*, *40*(12), 1814–1822. https://doi.org/10.1038/s41587-022-01387-y

Amoah, I. D., Abunama, T., Awolusi, O. O., Pillay, L., Pillay, K., Kumari, S., & Bux, F. (2022). Effect of selected wastewater characteristics on estimation of SARS-CoV-2 viral load in wastewater. *Environmental Research*, *203*. https://doi.org/10.1016/j.envres.2021.111877

Augusto, M. R., Claro, I. C. M., Siqueira, A. K., Sousa, G. S., Caldereiro, C. R., Duran, A. F. A., de Miranda, T. B., Bomediano Camillo, L. de M., Cabral, A. D., & de Freitas Bueno, R. (2022). Sampling strategies for wastewater surveillance: Evaluating the variability of SARS-COV-2 RNA concentration in composite and grab samples. *Journal of Environmental Chemical Engineering*, *10*(3). https://doi.org/10.1016/j.jece.2022.107478

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bitter, L. C., Kibbee, R., Jiménez, G. C., & Örmeci, B. (2022). Wastewater Surveillance of SARS-CoV-2 at a Canadian University Campus and the Impact of Wastewater Characteristics on Viral RNA Detection. *ACS ES and T Water*, *2*(11), 2034–2046. https://doi.org/10.1021/acsestwater.2c00060

Bivins, A., Kaya, D., Ahmed, W., Brown, J., Butler, C., Greaves, J., Leal, R., Maas, K., Rao, G., Sherchan, S., Sills, D., Sinclair, R., Wheeler, R. T., & Mansfeldt, C. (2022). Passive sampling to scale wastewater surveillance of infectious disease: Lessons learned from COVID-19. In *Science of the Total Environment* (Vol. 835). https://doi.org/10.1016/j.scitotenv.2022.155347

Breulmann, M., Kallies, R., Bernhard, K., Gasch, A., Müller, R. A., Harms, H., Chatzinotas, A., & van Afferden, M. (2023). A long-term passive sampling approach for wastewater-based monitoring of SARS-CoV-2 in Leipzig, Germany. *Science of the Total Environment*, *887*. https://doi.org/10.1016/j.scitotenv.2023.164143

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Carter, L. J., Garner, L. V., Smoot, J. W., Li, Y., Zhou, Q., Saveson, C. J., Sasso, J. M., Gregg, A. C., Soares, D. J., Beskid, T. R., Jervey, S. R., & Liu, C. (2020). Assay Techniques and Test Development for COVID-19 Diagnosis. *ACS Central Science*, *6*(5), 591–605. https://doi.org/10.1021/acscentsci.0c00501

Cha, G., Graham, K. E., Zhu, K. J., Rao, G., Lindner, B. G., Kocaman, K., Woo, S., D'amico, I., Bingham, L. R., Fischer, J. M., Flores, C. I., Spencer, J. W., Yathiraj, P., Chung, H., Biliya, S., Djeddar, N., Burton, L. J., Mascuch, S. J., Brown, J., … Konstantinidis, K. T. (2023). Parallel deployment of passive and composite samplers for surveillance and variant profiling of SARS-CoV-2 in sewage. *Science of the Total Environment*, *866*. https://doi.org/10.1016/j.scitotenv.2022.161101

D'Aoust, P. M., Tian, X., Towhid, S. T., Xiao, A., Mercier, E., Hegazy, N., Jia, J. J., Wan, S., Kabir, M. P., Fang, W., Fuzzen, M., Hasing, M., Yang, M. I., Sun, J., Plaza-Diaz, J., Zhang, Z., Cowan, A., Eid, W., Stephenson, S., … Delatolla, R. (2022). Wastewater to clinical case (WC) ratio of COVID-19 identifies insufficient clinical testing, onset of new

variants of concern and population immunity in urban communities. *Science of the Total Environment, 853*. https://doi.org/10.1016/j.scitotenv.2022.158547

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

D'Aoust, P. M., Towhid, T. S., Mercier, E., Hegazy, N., Tian, X., Bhatnagar, K., Zhang, Z., Naughton, C. C., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021b). COVID-19 wastewater surveillance in rural communities. *Science of the Total Environment, 801*. https://doi.org/10.1016/j.scitotenv.2021.149618

EPA. (2001). Method 1684: Total, Fixed, and Volatile Solids in Water, Solids, and Biosolids. *U. S. Environmental Protection Agency, EPA-821-R-*(January), 13.

Espinosa, M. F., Verbyla, M. E., Vassalle, L., Leal, C., Leroy-Freitas, D., Machado, E., Fernandes, L., Rosa-Machado, A. T., Calábria, J., Chernicharo, C., & Mota Filho, C. R. (2022). Reduction and liquid-solid partitioning of SARS-CoV-2 and adenovirus throughout the different stages of a pilot-scale wastewater treatment plant. *Water Research, 212*. https://doi.org/10.1016/j.watres.2022.118069

George, A. D., Kaya, D., Layton, B. A., Bailey, K., Mansell, S., Kelly, C., Williamson, K. J., & Radniecki, T. S. (2022). Impact of Sampling Type, Frequency, and Scale of the Collection System on SARS-CoV-2 Quantification Fidelity. *Environmental Science and Technology Letters*, *9*(2), 160–165. https://doi.org/10.1021/acs.estlett.1c00882

Habtewold, J., McCarthy, D., McBean, E., Law, I., Goodridge, L., Habash, M., & Murphy, H. M. (2022). Passive sampling, a practical method for wastewater-based surveillance of SARS-CoV-2. *Environmental Research*, 204. https://doi.org/10.1016/j.envres.2021.1120

Harris-Lovett, S., Nelson, K. L., Beamer, P., Bischel, H. N., Bivins, A., Bruder, A., Butler, C., Camenisch, T. D., De Long, S. K., Karthikeyan, S., Larsen, D. A., Meierdiercks, K., Mouser, P. J., Pagsuyoin, S., Prasek, S. M., Radniecki, T. S., Ram, J. L., Keith Roper, D., Safford, H., … Korfmacher, K. S. (2021). Wastewater surveillance for sars-cov-2 on college campuses: Initial efforts, lessons learned and research needs. *International Journal of Environmental Research and Public Health*, *18*(9). https://doi.org/10.3390/ijerph18094455

Jiang, G., Liu, Y., Tang, S., Kitajima, M., Haramoto, E., Arora, S., Choi, P., Jackson, G., D'Aoust, P. M., Delatolla, R., Zhang, S., Guo, Y., Wu, J., Chen, Y., Sharma, E., Prosun, T. A., Zhao, J., Kumar, M., Honda, R., … Meiman, J. (2023). Moving forward with COVID-19: Future research prospects of wastewater-based epidemiology methodologies and applications. *Current Opinion in Environmental Science & Health*. https://doi.org/10.1016/j.coesh.2023.100458

Kevill, J. L., Lambert-Slosarska, K., Pellett, C., Woodhall, N., Richardson-O'Neill, I., Pântea, I., Alex-Sanders, N., Farkas, K., & Jones, D. L. (2022). Assessment of two types of passive sampler for the efficient recovery of SARS-CoV-2 and other viruses from wastewater. *Science of the Total Environment*, *838*. https://doi.org/10.1016/j.scitotenv.2022.156580

Kim, S., Kennedy, L. C., Wolfe, M. K., Criddle, C. S., Duong, D. H., Topol, A., White, B. J., Kantor, R. S., Nelson, K. L., Steele, J. A., Langlois, K., Griffith, J. F., Zimmer-Faust, A. G., McLellan, S. L., Schussman, M. K., Ammerman, M., Wigginton, K. R., Bakker, K.

M., & Boehm, A. B. (2022). SARS-CoV-2 RNA is enriched by orders of magnitude in primary settled solids relative to liquid wastewater at publicly owned treatment works. *Environmental Science: Water Research and Technology*, *8*(4), 757–770. https://doi.org/10.1039/d1ew00826a

Kisand, V., Laas, P., Palmik-Das, K., Panksep, K., Tammert, H., Albreht, L., Allemann, H., Liepkalns, L., Vooro, K., Ritz, C., Hauryliuk, V., & Tenson, T. (2023). Prediction of COVID-19 positive cases, a nation-wide SARS-CoV-2 wastewater-based epidemiology study. *Water Research*, *231*, 119617. https://doi.org/10.1016/j.watres.2023.119617

Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C. P., Hamilton, K. A., Haramoto, E., & Rose, J. B. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Science of the Total Environment* 739, https://doi.org/10.1016/j.scitotenv.2020.139

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. Environ. Sci. Technol. Lett. 7,511–516. https://doi.org/10.1021/acs.estlett.0c00357

Mejías-Molina, C., Pico-Tomàs, A., Beltran-Rubinat, A., Martínez-Puchol, S., Corominas, L., Rusiñol, M., & Bofill-Mas, S. (2023). Effectiveness of passive sampling for the detection and genetic characterization of human viruses in wastewater. *Environmental Science: Water Research and Technology*. https://doi.org/10.1039/d2ew00867j

Melvin, R. G., Hendrickson, E. N., Chaudhry, N., Georgewill, O., Freese, R., Schacker, T. W., & Simmons, G. E. (2021). A novel wastewater-based epidemiology indexing method predicts SARS-CoV-2 disease prevalence across treatment facilities in metropolitan and regional populations. *Scientific Reports*, *11*(1), 1–9. https://doi.org/10.1038/s41598-021-00853-y

Mercier, E., D'Aoust, P. M., Thakali, O., Hegazy, N., Jia, J. J., Zhang, Z., Eid, W., Plaza-Diaz, J., Kabir, M. P., Fang, W., Cowan, A., Stephenson, S. E., Pisharody, L., MacKenzie, A. E., Graber, T. E., Wan, S., & Delatolla, R. (2022). Municipal and neighbourhood level wastewater surveillance and subtyping of an influenza virus outbreak. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20076-z

Mercier, E., Pisharody, L., Guy, F., Wan, S., Hegazy, N., D'Aoust, P. M., Kabir, M. P., Nguyen, T. B., Eid, W., Harvey, B., Rodenburg, E., Rutherford, C., Mackenzie, A. E., Willmore, J., Hui, C., Paes, B., Delatolla, R., & Thampi, N. (2023). Wastewater-based surveillance identifies start to the pediatric respiratory syncytial virus season in two cities in Ontario, Canada. *Frontiers in Public Health*, *11*. https://doi.org/10.3389/fpubh.2023.1261165

Naughton, C. C., Roman, F. A., Grace Alvarado, A. F., Tariqi, A. Q., Deeming, M. A., Bibby, K., Bivins, A., Rose, J. B., Medema, G., Ahmed, W., Katsivelis, P., Allan, V., Sinclair, R., Zhang, Y., Kinyua, M. N., & Author cnaughton, C. (2024). *Show us the Data: Global COVID-19 Wastewater Monitoring Efforts, Equity, and Gaps*. https://doi.org/10.1101/2021.03.14.21253564

Rafiee, M., Isazadeh, S., Mohseni-Bandpei, A., Mohebbi, S. R., Jahangiri-rad, M., Eslami, A., Dabiri, H., Roostaei, K., Tanhaei, M., & Amereh, F. (2021). Moore swab performs equal to composite and outperforms grab sampling for SARS-CoV-2 monitoring in wastewater. *Science of the Total Environment*, *790*. https://doi.org/10.1016/j.scitotenv.2021.148205

Sah, P., Fitzpatrick, M. C., Zimmer, C. F., Abdollahi, E., Juden-Kelly, L., Moghadas, S. M., Singer, B. H., & Galvani, A. P. (2021). *Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis*. https://doi.org/10.1073/pnas.2109229118/-/DCSupplemental

Wang, Y., Zheng, K., Gao, W., Lv, J., Yu, C., Wang, L., Wang, Z., Wang, B., Liao, C., & Li, L. (2022). Asymptomatic and pre-symptomatic infection in Coronavirus Disease 2019 pandemic. In *Medical Review* (Vol. 2, Issue 1, pp. 66–88). https://doi.org/10.1515/mr-2021-0034

Wilson, M., Qiu, Y., Yu, J., Lee, B. E., McCarthy, D. T., & Pang, X. (2022). Comparison of Auto Sampling and Passive Sampling Methods for SARS-CoV-2 Detection in Wastewater. *Pathogens*, *11*(3). https://doi.org/10.3390/pathogens11030359

4 Chapter 4: Passive Sampling for Genomic Surveillance of SARS-CoV-2 in Wastewater Resource Recovery Facility: Insights for Pandemic Preparedness

4.1 Context:

Chapter 4 represents the research article entitled "*Passive sampling for genomic surveillance of SARS-CoV-2 in wastewater resource recovery facility: Insights for pandemic preparedness* by Kabir, M. P., Plaza-Diaz, J., D'Aoust, P.M., Mercier, E., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Renouf, E., Goodridge, L., Lawal, O. U., Graber, T. E., and Delatolla, R. (2025) This article has been submitted to the journal of Water Research in 2025. This study evaluates two types of passive samplers in comparison with an autosampler at a wastewater treatment plant to assess SNV profiles and lineage prevalence in wastewater.

4.2 Abstract

Passive sampling provides a cost-effective alternative to conventional auto-sampling and serves as a valuable approach for wastewater surveillance in resource-limited settings. However, the feasibility of passive sampling for SARS-CoV-2 wastewater genomic surveillance remains underexplored in wastewater resource recovery facility (WWRF). In this study, we collected influent wastewater samples using an autosampler, COSCa-ball and Torpedo passive samplers from inlet of WWRF serving the city of Ottawa, Canada. We enriched, extracted, quantified, and sequenced samples targeting SARS-CoV-2 genome using a widely used ARTIC tiled amplicon approach. Our findings show that SARS-CoV-2 RNA concentrations were comparable ($p > 0.05$) among sampling methods. Although similar viral RNA concentrations were captured from passive samples, we found genomic recovery of SARS-CoV-2 from passive samplers was influenced by the targeted sequencing read length, with shorter (300 bp) reads resulting in poorer recovery than longer (600 bp) reads. Our study confirmed near-complete recovery of SARS-CoV-2 genomes (≥ 90%) from the autosampler,

COSCa-ball, and Torpedo samplers using longer reads. Genome sequencing parameters such as the number of raw reads, trimmed reads, mapped reads, depth of coverage and % of genome coverage was identical ($p > 0.05$) among the three sampling methods. Genomic analyses showed similar ($p > 0.05$) single nucleotide variant (SNV) profiles and lineage prevalence across sampling methods, and concordance with the available clinical surveillance. Overall, the findings suggest that passive sampling of wastewater is a viable, cost-effective alternative for population-scale genomic surveillance of SARS-CoV-2 and may allow for surveillance of other pathogens, supporting future pandemic preparedness efforts.

4.3 Introduction

The COVID-19 pandemic has transitioned into endemic phase, yet it remains a formidable threat as SARS-CoV-2 continually evolves, giving rise to diverse variants that enhance transmissibility, disease severity, and evade immune defenses (Gangavarapu et al., 2023; Harvey et al., 2021). Identifying and tracking them has become critical to early and effective public health responses. However, the sensitivity of existing strategies for variant detection through clinical genomic surveillance has suffered due to widespread vaccination efforts which has lessened the severity of infections and the need for clinical diagnosis (Ling-Hu et al., 2022; Robishaw et al., 2021). Besides, clinical genomic surveillance is costly, confined to symptomatic patients, and susceptible to sampling bias due to healthcare disparities that disproportionately affect poorer and underserved communities (Lieberman-Cribbin et al., 2020; Majid et al., 2020; Reitsma et al., 2021). In contrast, wastewater genomic surveillance (WWGS) has emerged as a robust tool for monitoring the emergence and spread of SARS-CoV-2 variants, effectively overcoming sampling bias and economic constraints as well as offering a non-invasive approach to track the circulation patterns of different SARS-CoV-2 variants at the population levels (Fontenele et al., 2021; Karthikeyan et al., 2021; Pérez-Cataluña et al., 2022). WWGS entails the collection of wastewater samples, followed by target

enrichment, RNA extraction and sequencing to identify and characterize the distinct strains or lineages of SARS-CoV-2 circulating within the population (Ai et al., 2021; Li, Uppal, et al., 2022; Vo et al., 2022). This approach allows for the early detection of prevalent SARS-CoV-2 lineages on both regional and national scales (Amman et al., 2022; Crits-Christoph et al., 2021; Izquierdo-Lara et al., 2021), as well as demonstrated the detection of circulating and cryptic (i.e., not encountered in clinical surveillance) lineages in the population level (Jahn et al., 2022; Karthikeyan et al., 2022; Smyth et al., 2022). The WWGS has also been extended to monitor other disease targets, including enteroviruses, influenza A and B, poliovirus, and respiratory syncytial viruses (Allen et al., 2024; Child et al., 2023; Klapsa et al., 2022; Lee et al., 2024; Williams et al., 2024).

WWGS has predominantly relied on composite samples, often collected using an autosampler, from WWRF or grab sampling of sewer shed access points (Nemudryi et al., 2020; Spurbeck et al., 2021; Vo et al., 2022). However, auto-sampling is expensive, and requires electricity and maintenance, while grab samples are prone to intra-day variation, making it less representative to population-level disease prevalence (Augusto et al., 2022; Harris-Lovett et al., 2021). As an alternative to traditional auto and grab sampling, passive sampling has garnered attention for its simplicity and cost-effectiveness in wastewater surveillance (WWS) (Habtewold et al., 2022; Kevill et al., 2022; Vincent-Hubert et al., 2022). In passive sampling, sampling mediums (i.e., medical gauzes, cheesecloth, cellulose sponges, or electronegative membrane filters) are packed within styrene-like plastic devices and deployed in sewer networks to sorb viral particles over time (Bivins et al., 2022). The sampling mediums are then homogenized in deionized water or soluble buffer to extract total RNA for downstream analysis. This technique is simple, requires little maintenance and no mechanical components, and is particularly useful in locations with limited infrastructure and limited access to electricity, where long-term sampling is desired. Although passive sampling has demonstrated success in WWS, it remains

underutilized in WWGS of SARS-CoV-2. To date, only five studies have focused on SARS-CoV-2 genomic analysis using passive sampling. Among these, Corchis-Scott et al. (2021) and Mangwana et al. (2022) employed allele-specific RT-qPCR for the detection of lineage-specific mutations, while Cha et al. (2023), Alamin et al. (2024) and Overton et al. (2024) sequenced the SARS-CoV-2 genome for comprehensive variant identification. However, Cha et al. (2023) was unable to recover near-complete genome (≥ 90%) from any of the wastewater samples collected by Moore swabs, whereas Alamin et al. (2024) reported successful recovery of SARS-CoV-2 genome from the majority of wastewater samples collected through Tampon swabs. Despite the potential of passive sampling in WWGS, their applications were primarily limited to tracking SARS-CoV-2 lineages in congregate settings with low sewer flow rates, such as student dorms at university campuses and wastewater from airport. However, the implementation of passive sampling in larger, high-flow environments such as the inlet of WWRF for genomic analysis remains unexplored. In these settings, passive sampling may face challenges, as the high-flow, turbulent conditions could reduce the efficiency of capturing and retaining viral particles on the sampling medium required for sequencing. Besides, there are concerns about the integrity of RNA species in such conditions, as high-flow may degrade RNA more rapidly, potentially impacting the quality of genomic data retrieved through passive sampling. Moreover, earlier studies (Fontenele et al., 2021; Nemudryi et al., 2020) suggest that recovering SARS-CoV-2 genomes requires at least 2.8×10^5 viral copies per liter in wastewater. However, later findings indicated that genome recovery depends heavily on sample processing and concentration methods (Chen et al., 2024; Feng et al., 2023). Thus, it remains uncertain whether passive sampling in WWRF and current processing methods can be effectively utilized for genomic analysis of SARS-CoV-2. Additionally, it is unclear whether SARS-CoV-2 lineage profiles obtained from passive sampling would be comparable to those from conventional auto sampling and clinical surveillance. This underscores the need for

further research to evaluate the efficacy of passive sampling at the inlet of WWRF in order to support the continued and broader application of WWGS for monitoring respiratory pathogens and emerging disease targets.

The aim of this research is to evaluate the feasibility and accuracy of passive sampling in detecting and quantifying the distribution of SARS-CoV-2 lineages from the inlet of WWRF. To validate the application of passive samplers for genomic analysis in WWRF, we sequenced and analyzed wastewater samples collected by an autosampler, a COSCa-ball, and a Torpedo passive sampler from the inlet of the Robert O. Pickard Environmental Center (ROPEC) in Ottawa, Canada. The WWRF serves a sewered population of approximately 1 million in the nation's capital, with an average flow of 4.35×10^5 m^3/day and a hydraulic residence time ranging from 2-35 hours, averaging around 12 hours.

4.4 Materials and Methods

4.4.1 Wastewater sample collection and processing

A total of 75 wastewater samples composited over a 24-hour period were collected in parallel during a period from March 14 to May 4, 2023, using an autosampler (n=25) (Teledyne ISCO, Lincoln, NE, USA), COSCa-ball (n=25) and Torpedo sampler (n=25) from the inlet of WWRF, Ottawa, Canada. The COSCa-ball and Torpedo sampler housing unit were printed in polylactic acid filament using an Ultimaker 2+ 3-D printer (Ultimaker, New York, USA). The COSCa-ball and Torpedo samplers are designed with multiple entry points to allow wastewater to pass through, effectively trapping wastewater solids in the absorbing mediums secured within the cavity. In the COSCa-ball sampler, four swatches of medical gauze (Swisspers, China) were used as the absorbing medium whereas in the Torpedo sampler, two swatches of medical gauze were used, due to the smaller diameter. The autosampler was used to collect 24 hours composite wastewater sample while COSCa-ball and Torpedo samplers were deployed for a period of 24

hours. Following each sampling day, 500 mL composite wastewater sample was collected using an autosampler, and the medical gauzes from passive samplers were retrieved in 50 mL conical tubes. All samples were transported to the laboratory in ice-cooled packs. Wastewater samples collected by autosampler were allowed to settle for 30 minutes. The gauzes were placed in 500 mL of deionized water and vigorously agitated with a glass rod to recover the sorbate (Figure 4.1). The gauzes were subsequently squeezed with a glass rod to remove sorbate, discarded the gauzes, and larger particles were also allowed to settle for 30 minutes.

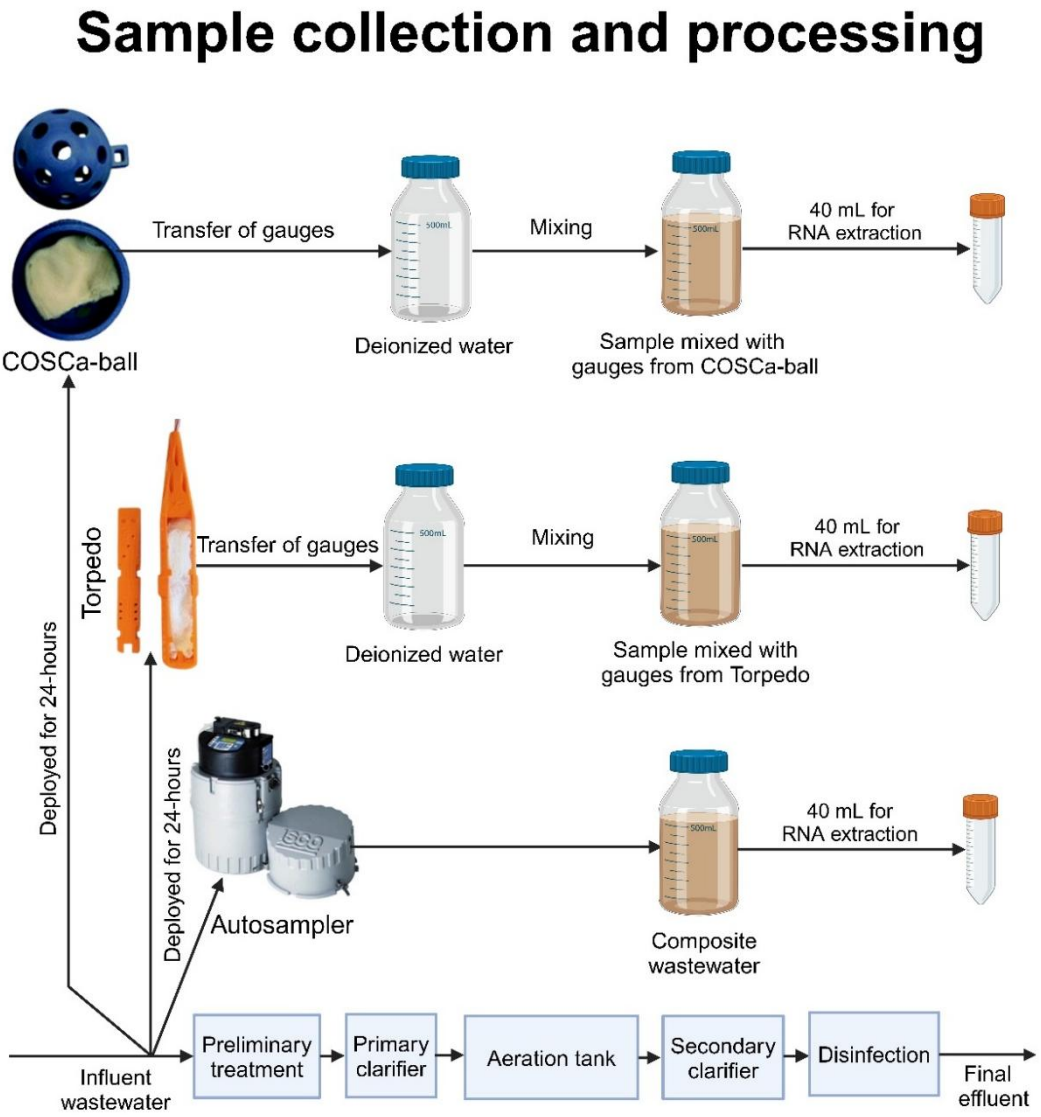

Figure 4-1: Schematic displaying collection of wastewater samples using an autosampler along with COSCa-ball, and Torpedo passive samplers, followed by processing for total RNA extraction.

4.4.2 Wastewater suspended solids

To determine the suspended solids in wastewater samples used for RNA extraction, ten paired wastewater samples were collected from the same WWRF using an autosampler, COSCa-ball, and Torpedo sampler between April 14 and May 6, 2024. The Total Suspended Solids (TSS) and Volatile Suspended Solids (VSS) of the wastewaters samples were determined using standard methods (EPA, 2001).

4.4.3 SARS-CoV-2 RNA extraction

Nanotrap® Magnetic (Ceres Nanosciences, Manassas, USA) hydrogel particles were used to enrich SARS-CoV-2 genetic material as per Kabir et al. (2025a). Briefly, 40 mL of the sample supernatant was transferred to a centrifuge tube, 600 µL of magnetic particles were added, and the tube rotated at 100 rpm for 20 minutes at 20 ºC. Following centrifugation for 10 minutes at 4 ºC at 8000 x g, the tube was placed on a magnetic rack (DynaMag-50, Invitrogen, Massachusetts, USA) to retain the particles, and the supernatant discarded. The bound material was then lysed with 140 µL of PBS buffer (pH 7.4) and 560 µL of Viral Lysis Buffer, then transferred to a 2.0 mL microcentrifuge tube and placed on a magnetic rack (DynaMag™-2 magnet, Invitrogen) for 10 minutes to complete the lysis and remove the magnetic particles. The lysate was collected without disturbing the magnetic particles and total RNA was extracted using QIAmp Viral RNA Mini Kit (Qiagen, Germantown, USA) and eluted in 70 µL of nuclease-free water.

4.4.4 Reverse Transcription (RT)-qPCR and SARS-CoV-2 RNA quantification

The presence of SARS-CoV-2 N1 and N2 targets were confirmed using singleplex one-step RT-qPCR with TaqMan® 1-Step Fast Virus Master Mix (ABI) (Thermo-Fisher, Massachusetts, USA) on a CFX Connect qPCR thermocycler (Bio-Rad, Hercules, Canada) according to the established protocol (Kabir et al., 2025b). Briefly, 7 µL of master mix

consisting of 2.5 µL of 4X TaqMan® Fast Virus 1-step mastermix (Thermo-Fisher), 0.75 µL of 500 µM primer, and 125 µM probe mixture (IDT, Kanata, Canada), and 3.75 µL of nuclease-free water was mixed with 3 µL of RNA template. The qPCR conditions are as follows: RT at 50 °C for 5 min, RT inactivation and initial denaturation at 95 °C for 20 s, 45 cycles of denaturation at 95 °C for 3 s and annealing/ extension at 60 °C for 30 s. Samples were run in triplicate with RNase-free water as non-template control and quantified using a five-point EDX SARS-CoV-2 standard curve (Exact Diagnostics, Texas, USA). The qPCR assays were performed according to the MIQE guidelines (Bustin et al., 2009), with data analysis being limited to samples that met the following quality control criteria: (a) linear standard curve (R2 $\geq$ 0.95), (b) copies/well in the linear range of the standard curve, and (c) primer efficiency (90% to 120%). Moreover, samples with a value ± 0.5 standard deviations from the mean were excluded from the study.

4.4.5 Genome sequencing of SARS-CoV-2

The RNA extracts were first treated using ezDNase™ Enzyme kit (ThermoFisher), followed by reverse transcription with SuperScript IV First-Strand Synthesis System (ThermoFisher). Complementary DNA was used to amplify viral sequences through multiplexed PCR using ARTIC primer set v4.1 and Q5 Hot Start Master Mix (BioLabs New England, Massachusetts, United States). The thermocycling conditions are as follows: initial activation at 98 °C for 30 seconds, followed by 35 cycles of denaturation at 95 °C for 15 seconds, and annealing and extension at 63 °C for 5 minutes. The expected 400 bp size of the tiled amplicons was verified using the 2100 Bioanalyzer system (Agilent Technologies, Mississauga, Canada), cleaned with 0.8x AMPure XP beads (Beckman Coulter, California, USA) and quantified in a Qubit 4 fluorometer (ThermoFisher) using the dsDNA High Sensitivity Kit (Invitrogen). Amplicon concentrations were normalized, and sequencing libraries prepared using Nextera XT DNA Library Preparation kit (Illumina, California, USA). Finally, the libraries were pooled,

normalized and sequenced on a MiSeq platform (Illumina) with a read length of 300 bp in paired-end mode.

4.4.6 Bioinformatics

A customized bioinformatics pipeline (https://github.com/nf-core/viralrecon/tree/2.6.0) was used to analyze the FASTQ files for quality control and mutation detection (Ewels et al., 2020). Briefly, the quality of raw reads was assessed using FastQC (Simon Andrews, 2020), filtered using fastp (Chen et al., 2018) in order to remove adaptor sequences, ambiguous base (N), low quality reads (Phred score < 30), and small fragments (< 50 nt). Filtered reads were then aligned to the SARS-CoV-2 reference genome (accession MN908947.3) using Bowtie2 (Langmead & Salzberg, 2012) with default parameters, and Mosdepth (Pedersen & Quinlan, 2018) was used to generate coverage information along the genome. The aligned reads were sorted using SAMtools (Li et al., 2009) and consensus sequences were generated using iVar consensus algorithm (Grubaugh et al., 2019). The consensus sequences (frequency >50%) were constructed using map reads with a coverage of > 10× and Phred score of > 30. Finally, mutation calling was performed using iVar with minimum frequency threshold (0.01), minimum Phred score (30), and minimum read depth (10×).

4.4.7 Wastewater SARS-CoV-2 lineage estimation

To capture the dynamics of virus evolution and spread, we utilized Freyja bioinformatics tool (Karthikeyan et al., 2022) to infer relative abundance of Pango lineages in wastewater samples. We used BAM files on the Freyja workflow (v1.3.10) and packages found at https://github.com/andersen-lab/Freyja to determine lineage abundance through a regression approach that considers depth weighting and least absolute deviation. All samples were analyzed using Freyja, and the results were plotted against the sampling dates to observe the

patterns of lineage prevalence in wastewater collected using autosampler, COSCa-ball, and Torpedo samplers.

4.4.8 Clinical surveillance data

Clinical genomic surveillance data was obtained from the weekly epidemiological summary of SARS-CoV-2 genomic surveillance in Ontario for the period between March 14 and May 4, 2023 (PHO, 2023). The prevalence (%) of each lineage was compiled to facilitate a comparative analysis with wastewater genomic surveillance data from the corresponding database, and lineages with a prevalence of less than 1% were classified as "Other".

4.4.9 Statistical analysis

A one-way analysis of variance (ANOVA) test was used to compare metrics (i.e., TSS and VSS concentrations and genome sequencing parameters) derived from autosampler along with the COSCa-ball and Torpedo sampled wastewater. The SARS-CoV-2 RNA concentrations were compared across the sampling methods using a Kruskal–Wallis test. The relationship between SARS-CoV-2 RNA concentrations and genome sequencing parameters were evaluated using Pearson correlation analyses. Additionally, the number of SNVs was compared across the autosampler, COSCa-ball sampler, Torpedo sampler, and clinical samples for the same time period using a linear fixed-effects model (Bates et al., 2015), while lineage prevalence was analyzed using a zero-inflated model (Hall, 2000). The selection of the statistical test was determined based on the assessment of the normality of the data. In our statistical analysis, we considered p values less than 0.05 to be significant. All statistical analyses and plotting were performed in GraphPad Prism 10.2.1 (La Jolla, California, USA), and R software packages (Version 4.3.3).

4.5 Results and Discussion

4.5.1 Wastewater suspended solids concentrations

The average TSS and VSS concentrations varied significantly (ANOVA, $p < 0.05$) among the sampling methods (Figure 4.2). Post-hoc analysis indicated that the COSCa-ball samples (361 ± 181.5 mg/L) contained significantly higher ($p < 0.05$) TSS concentrations followed by Torpedo (191.6 ± 98.29 mg/L) and autosampler (165.8 ± 78.77 mg/L) collected samples. Similarly, and as expected, VSS concentrations were also higher ($p < 0.05$) in the COSCa-ball (357.9 ± 174.7 mg/L) compared to autosampler (163.4 ± 63.93 mg/L) and Torpedo (185.2 ± 69.99 mg/L) collected samples. In contrast, TSS and VSS concentrations were similar ($p > 0.05$) between the autosampler and Torpedo samples. These variations of TSS and VSS in the COSCa-ball samples were mainly attributed to its design and the amount of sorbent material used within the COSCa-ball casing. The COSCa-ball sampler is a hollow sphere with multiple large entry points, allowing more suspended colloids, solids, and debris to enter the samplers. Additionally, the use of four swatches of medical gauze in the COSCa-ball sampler resulted in the formation of thick layer of solids on the gauzes. When the gauzes were resuspended in deionized water, the solution became less opaque compared to that obtained from autosampler and Torpedo samples, even after 10 minutes of settling. Ultimately, the supernatant used for RNA extraction from the COSCa-ball sampler is herein shown to contain higher TSS and VSS concentrations compared to autosampler, and Torpedo passive sampler collected wastewaters.

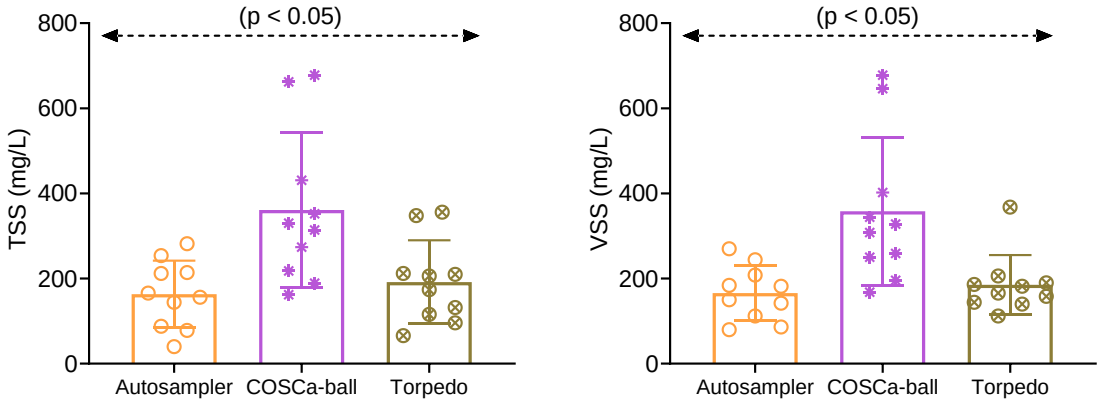

Figure 4-2: TSS and VSS of influent wastewater concentrations across various sampling methods. The error bars represent the standard deviation of TSS and VSS concentrations throughout the sampling period.

4.5.2 SARS-CoV-2 RNA concentrations

The SARS-CoV-2 RNA was detected in all 75 samples (100%) collected by autosampler, COSCa-ball, and Torpedo passive samplers, with average concentrations of $1.32 \pm 1.21 \times 10^4$, $1.98 \pm 1.97 \times 10^4$ and $2.21 \pm 2.28 \times 10^4$ copies/L respectively (Figure 4.3). Statistical analysis indicated that SARS-CoV-2 RNA concentrations were similar (Kruskal-Wallis test, $p > 0.05$) among the autosampler, COSCa-ball and Torpedo sampler collected wastewaters, although the average SARS-CoV-2 RNA concentration obtained from the Torpedo samplers was approximately twice as high as the average RNA concentration of the autosampler collected sample. This difference is associated with the higher variability in the SARS-CoV-2 RNA concentrations of the Torpedo sampler data set. Our findings align with earlier research (Kevill et al., 2022; Vincent-Hubert et al., 2022) that demonstrates that SARS-CoV-2 RNA concentrations in wastewater collected by passive samplers are often comparable to conventional composite samples collected by autosampler in WWRFs. In recognition that the comparability of autosampler and passive sampler collection methods depends on several factors, including effective deployment time of the passive sampler, selection of appropriate

sampling mediums contained within the passive sampler, and sample processing methods (Habtewold et al., 2022; Schang et al., 2021; LaTurner et al., 2021; Li et al., 2021). In this study, we deployed the passive samplers over 24 hours to align with the autosampler collection time and processed both the autosampler and passive sampler samples in parallel using Nanotrap® Magnetic Virus Particles, a suitable method for viral RNA isolation (Antkiewicz et al., 2024; Dehghan Banadaki et al., 2024). Thus, by minimizing potential confounding effects introduced by differences in sampler deployment locations and time along with downstream processing methods, we show that comparable SARS-CoV-2 RNA concentrations are obtained between the autosampler and two different passive sampler collection methods.

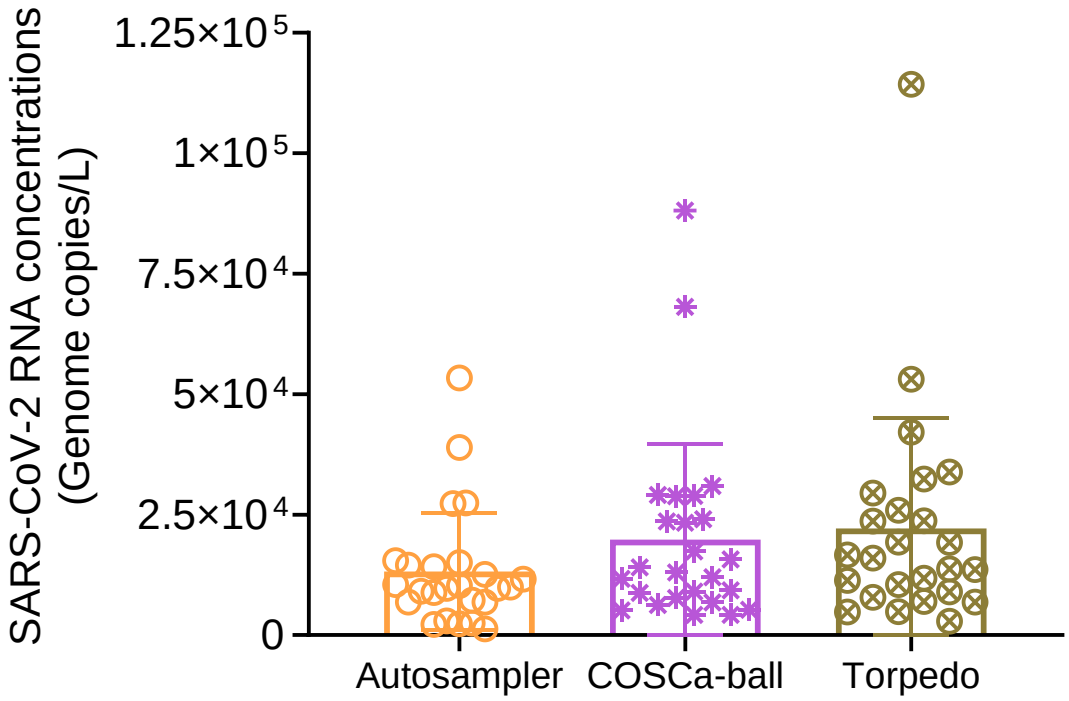

Figure 4-3: SARS-CoV-2 RNA concentrations measured by RT-qPCR in wastewater collected using an autosampler along with a COSCa-ball, and Torpedo passive samplers. The error bar shows the standard deviation of SARS-CoV-2 RNA concentrations over the monitoring period.

4.5.3 Illumina platform longer read sequencing improve genomic coverage of SARS-CoV-2 from passive samplers

To obtain the genome sequence of SARS-CoV-2, wastewater samples collected using autosampler, COSCa-ball, and Torpedo samplers were simultaneously sequenced on a MiSeq platform. Initially, we sequenced one set of samples from autosampler, COSCa-ball, and Torpedo sampler using a 2 × 150 bp run mode. The results showed lower genomic coverage for the COSCa-ball (48% at 10× depth) and Torpedo (2% at 10× depth) samples, despite generating approximately 2 million (M) raw sequence reads and a high proportion of SARS-CoV-2-mapped reads (Figure 4.4.A and Table A1). We hence subsequently re-sequenced the same set of samples on a MiSeq platform in 300 bp paired-end run mode and successfully recovered near-complete consensus genomes (≥ 90%) with nearly the same number of raw reads as the 2 x 150 bp run from the autosampler, COSCa-ball, and Torpedo samples (Figure 4.B and Table S1). The difference in genomic coverage between the 150 bp and 300 bp paired-end run mode suggests that sequencing length may have a significant role in SARS-CoV-2 genome recovery from passive samplers. This is possibly due to the fact that longer sequencing read lengths (i.e., 2 × 300 bp) can enhance genome assembly by providing greater overlap between reads, improving the resolution of complex or variable regions, and increasing overall coverage across the genome.

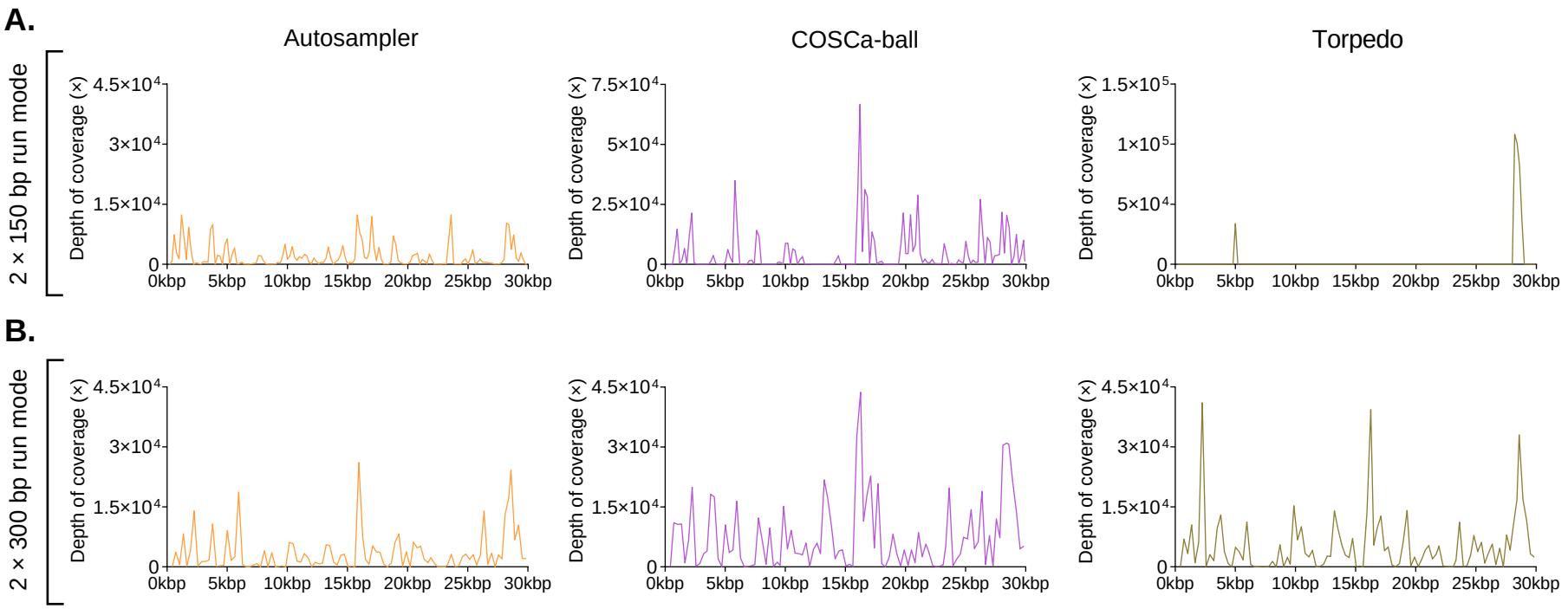

Figure 4-4: Genomic coverage of 150 bp paired-end run mode (A) and 300 bp paired-end run mode (B) across autosampler, COSCa-ball, and Torpedo samples. The 150 bp paired-end run mode yielded lower genomic coverage, with COSCa-ball achieving 48% at 10× depth and Torpedo only 2% at 10× depth, while the near-complete genome (≥ 90%) was recovered from autosampler, COSCa-ball, and Torpedo samplers using the 300 bp paired-end mode.

4.5.4 Overview of SARS-CoV-2 genome sequencing parameters across sampling methods

We sequenced all 75 samples of this study on a MiSeq platform using a 2 × 300 bp run mode, and results showed that the average number of raw reads obtained from autosampler, COSCa-ball, and Torpedo samplers were 2.61M ± 0.43M, 2.56M ± 0.38M, and 2.41M ± 0.38M, respectively (Figure 4.5.A). On average, 38.5% of the reads were filtered out for each sample, yielding high-quality trimmed reads: 1.62M ± 0.38M for auto, 1.58M ± 0.31M for COSCa-ball, and 1.46M ± 0.26M for Torpedo sampler. An average of 1.53M (94.5%) trimmed reads from the autosampler, 1.51M (95.6%) from the COSCa-ball, and 1.34M (91.9%) from the Torpedo sampler were mapped with the SARS-CoV-2 reference genome sequence (accession MN908947.3). The number of mapped reads varied from 0.55M to 2.12M (mapping rate: 46.49% to 99.82%) for the autosampler, 0.85M to 2.07M (mapping rate: 80% to 99.81%) for the COSCa-ball, and 0.82M to 1.86M (mapping rate: 54.29% to 99.87%) for the Torpedo sampler (Table A2). After mapping to the reference genome, the average genome coverage of SARS-CoV-2 at 10× depth was 93.68% ± 6.52% (range: 69% to 99%) for the autosampler, 95.28% ± 3.06% (range: 88% to 99%) for the COSCa-ball, and 96.12% ± 1.98% (range: 90% to 98%) for the Torpedo sampler (Figure 4.5.B). Among the 25 samples collected using each type of sampler, near-complete genome (≥ 90% coverage) at 10× depth was recovered from 84% of the autosampler samples, 92% of the COSCa-ball samples, and 100% of the Torpedo samples. The calculated median depth of coverage varied between 1168× to 6392× for the autosampler samples, 1391× to 6396× for the COSCa-ball and 1810× to 5682 for the Torpedo sampler (Figure 4.5.C). However, no significant differences (ANOVA, $p > 0.05$) were observed in the total number of raw reads, trimmed reads, mapped reads, percentage of genome coverage, and depth of coverage among autosampler, COSCa-ball, and Torpedo samplers. Overall, the results of this study suggested that SARS-CoV-2 genome sequencing parameters

were consistent across different sampling methods, indicating that genomic data were not limited by the sampling technique.

We further investigated the potential relationship between SARS-CoV-2 RNA concentrations and sequencing parameters using Pearson correlation analyses and no significant correlation ($p > 0.05$) was observed across the different sampling methods (Table A3). This demonstrated that genome sequencing parameters, particularly genome coverage, were not dependent on SARS-CoV-2 N1 or N2 RNA concentrations, as near-complete SARS-CoV-2 genomes were successfully recovered from the autosampler at SARS-CoV-2 RNA concentrations as low as 1.29×10^3 copies/L (corresponding to a Ct value of 38). Similarly, near-complete genomes were recovered from COSCa-ball and Torpedo samplers at the lowest SARS-CoV-2 RNA concentrations of 4.17×10^3 and 2.91×10^3 copies/L (corresponding to Ct values ~36), respectively. Our findings contrast with earlier studies (Fontenele et al., 2021; Nemudryi et al., 2020) showing that recovering SARS-CoV-2 genomes from wastewater required a minimum SARS-CoV-2 RNA concentrations of 2.8×10^5 copies/L. This suggests that low SARS-CoV-2 RNA concentrations (specifically N1 or N2 genomic regions) do not adversely affect sequencing parameters; rather, the sample processing method may play a more dominant role. Our hypothesis is in accordance with a recent finding by Feng et al. (2023), which evaluated different sample processing methods and suggested that sample processing is one of the predominant factors influencing success/failure of amplicon sequencing. Moreover, advanced sequencing technologies can enable the recovery of high-quality data, even from samples with very low SARS-CoV-2 RNA concentrations. The assumptions aligned with earlier research, where robust sequencing data were derived from low concentration samples using Illumina platforms via amplification process (Fontenele et al., 2021; Izquierdo-Lara et al., 2021; Pérez-Cataluña et al., 2021), and capture enrichment method (Crits-Christoph et al., 2021).

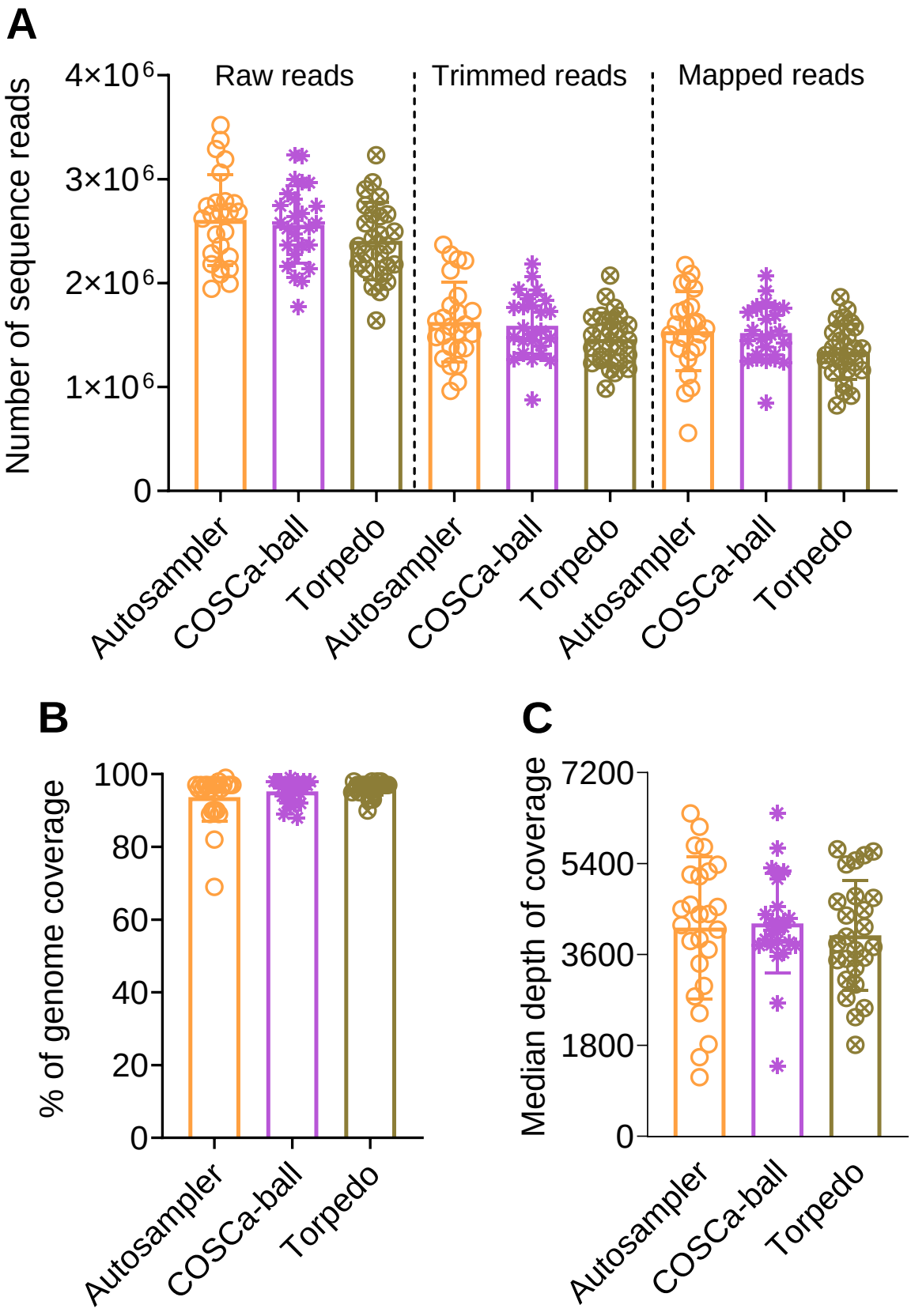
A
Number of sequence reads
4×10^6
3×10^6
2×10^6
1×10^6
0
Raw reads
Trimmed reads
Mapped reads
Autosampler
COSCa-ball
Torpedo
B
% of genome coverage
100
80
60
40
20
0
C
Median depth of coverage
7200
5400
3600
1800
0

Figure 4-5: SARS-CoV-2 genome sequencing parameters in wastewater across three different sampling methods. (A) Total number of reads, trimmed reads, and SARS-CoV-2 mapped reads, (B) % of genome coverage at minimum 10× read depth, (C) median depth (×) of coverage.

4.5.5 Evaluation of SARS-CoV-2 SNVs and INDELs across sampling methods

Genomic analysis showed on an average 105 ± 10, 102 ± 9, and 99 ± 7 individual mutations (SNVs and Indels) detected in the autosampler, COSCa-ball, and Torpedo samplers respectively in comparison with a SARS-CoV-2 reference sequence (Figure 4.6.A). Temporal trends exhibited similar (linear effect model, $p > 0.05$) SNVs profile across the sampling methods, with the exception of the autosampler on March 29, 2023, which was possibly due to lower genomic coverage (69%) on that date (Table A2). This finding aligns with previous research that demonstrated correlations between SARS-CoV-2 mutations and genomic coverage. A study conducted by Crits-Christoph et al. (2021) found that higher sequencing genome coverage improved the resolution of mutations detection in wastewater samples. Similarly, Jahn et al. (2022) showed that genomic coverage influences the ability to detect and quantify mutations of SARS-CoV-2 in wastewater, where higher coverage facilitates more reliable identification of lineages and their defining mutations.

The majority of detected SNVs and Indels in SARS-CoV-2 genome across the autosampler, COSCa-ball, and Torpedo samplers were found in the ORF1ab (40%) and S (40%) regions, with the remaining identified in the N region (~7%), followed by ORF3a (3.5%), M (3%), ORF8 (2.25%), E (2%), ORF6 (2%), and ORF10 (0.1%) of the genome respectively (Figure 4.6.B). Of the detected SNVs, ~70% were missense (resulting in amino acid changes in protein sequences) across all sampling methods, while 21% were synonymous (not altering the amino acid sequence) in autosampler and Torpedo samples, and 16% in COSCa-ball samples. Similarly, 5% of Indels

were identified in autosampler and Torpedo samples, with 7% detected in COSCa-ball samples (Figure 4.6.C). It was worth noting that deletions were more common than insertions across all genes across the sampling methods. We categorized four types of SNVs: Nonsense (premature stop codon), nonstop (point mutation in a stop codon leading to continued RNA translation), upstream (towards 5' UTR), and downstream (towards 3' UTR) as "others". Figure 4.6.C shows the distribution of these other SNVs across the autosampler, COSCa-ball, and Torpedo samples. Notably, nonsense SNVs (mostly detected in ORF8) were found only in RNA extracted from autosampler and COSCa-ball samplers, none in Torpedo samples. Furthermore, two upstream and one downstream SNVs were detected across the samples. The findings of this study align with clinical sequence analyses regarding the locations and types of SNVs and Indels detected. An early study (Gálvez et al., 2022) conducted in Cambodia during the initial phase of the pandemic revealed that the majority of SNVs and Indels were located in the ORF1ab and S genes, predominantly as missense or synonymous mutations. Subsequently, evaluations of clinical sequences available in the GISAID database corroborated these observations, indicating similar patterns in SNVs and Indels distribution and types throughout the world during five major VOCs (Saldivar-Espinoza et al., 2023). Overall, the number, locations, and types of SNVs and Indels detected in SARS-CoV-2 genome were identical (ANOVA, $p > 0.05$) across the autosampler, COSCa-ball, and Torpedo samplers.

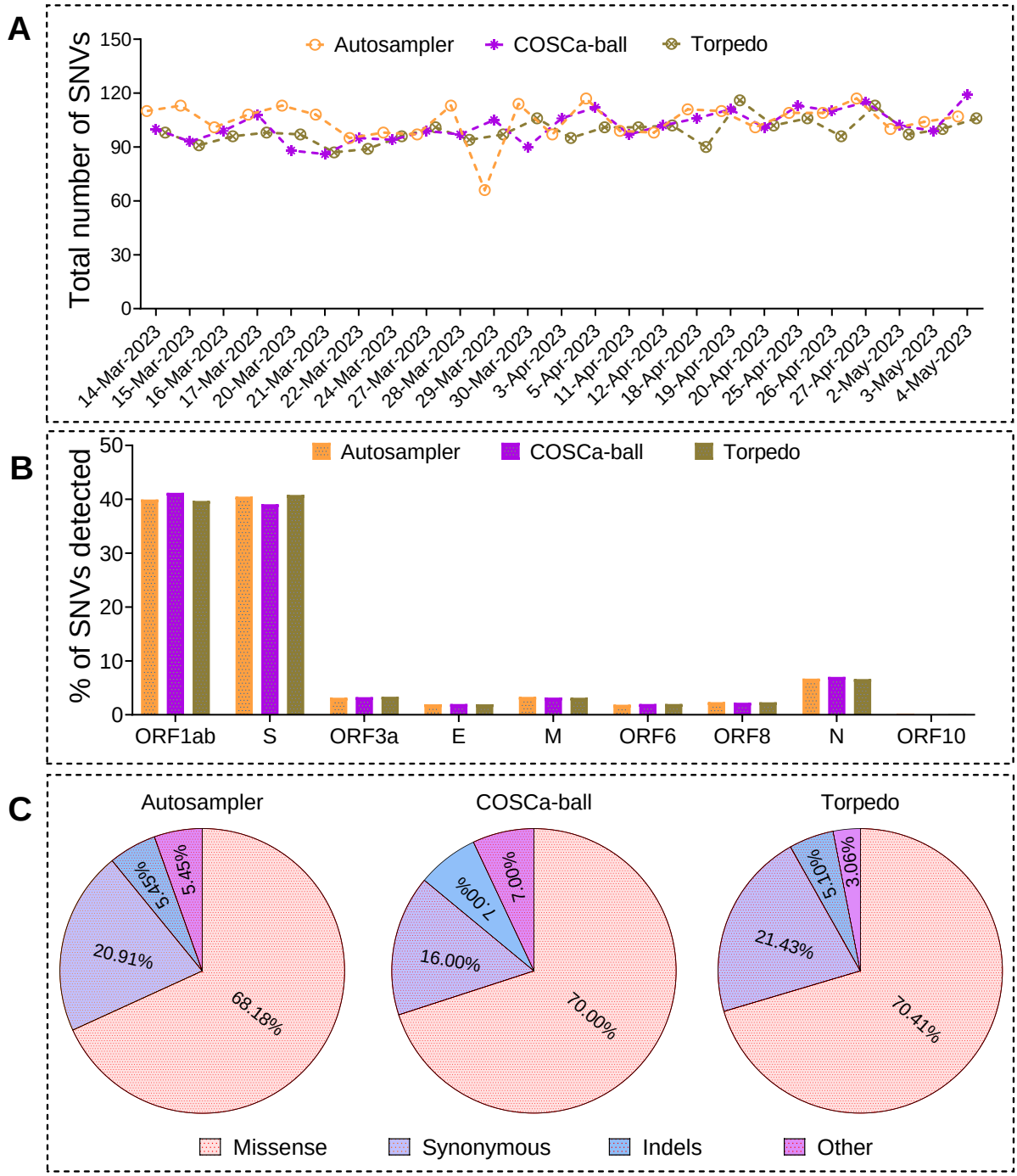

Figure 4-6: SARS-CoV-2 SNVs and Indels in wastewater samples collected using autosampler, COSCa-ball, and torpedo sampler. (A) Longitudinal assessment of the total number of detected SNVs in wastewater per sample, (B) their location and (C) type of the detected SNVs and Indels in the genome.

4.5.6 Comparable SARS-CoV-2 lineages prevalence across sampling methods

As wastewater contains a complex mixture of SARS-CoV-2 genetic material contributed from many different infected hosts and different viral lineages were in circulation during the sampling period, we evaluated the effectiveness of passive samplers in capturing the Pango lineages distribution compared to autosampler. To estimate the prevalence of lineages in RNA extracted from autosampler, COSCa-ball, and Torpedo samplers, we employed the Freyja bioinformatics tool (Karthikeyan et al., 2022). The analysis showed the detection of two parental Omicron lineages B.1.1.529* and CH.1.1 (alias B.1.1.529.2.75.3.4.1.1.1.1) along with five recombinant lineages from BJ.1 and BM.1.1.1 namely XBB, XBB.1.5, XBB.1.9, XBB.1.16, and XBB.2.3 in wastewater samples collected using various sampling methods throughout the study period (Figure A1).

Among the detected lineages, XBB.1.5 was identified as the dominant lineage across all sampling methods, closely aligning with available clinical surveillance throughout the sampling period (Figure 4.7). Notably, while clinical sequences did not report XBB during the sampling period, the wastewater samples identified its presence at approximately 30% prevalence throughout the study period in nearly all wastewater samples across the three sampling methods. XBB.2.3 was detected in wastewater but not clinical sequences, although these detections were sporadic regardless of sampling method, and were at low prevalence among sampling methods. The prevalence of the Omicron (B.1.1.529*) lineage in clinical surveillance was approximately 20% at the beginning of the sampling period and gradually declined. In contrast, the prevalence of Omicron in autosampler collected samples were consistently detected at a higher prevalence throughout March 2023, while a lower prevalence was observed for the COSCa-ball and Torpedo samples, although prevalences reached ~75% in early-to-mid April samples. The prevalence of XBB.1.9 in wastewater collected

using autosampler, COSCa-ball, and Torpedo samplers closely aligned with that observed in clinical sequences. On the other hand, XBB.1.16 was first detected in mid-March in wastewater at lower prevalence in samples collected using COSCa-ball and Torpedo samples, while it first detected in both wastewater samples collected by autosampler and clinical samples in early April 2023. Thereafter, its prevalence gradually increased in both wastewater and clinical sequences. The prevalence of "other" lineages in wastewater is shown to gradually increase with time across the sampling methods as well as in clinical sequences. However, the early detection of XBB.1.16 using COSCa-ball and Torpedo samplers highlights the potential of passive sampling for early identification of emerging SARS-CoV-2 lineage. Earlier studies (Alamin et al., 2024; Overton et al., 2024) have demonstrated the utility of passive sampling as an early warning tool for the detection of SARS-CoV-2 lineages prevalence across various scales.

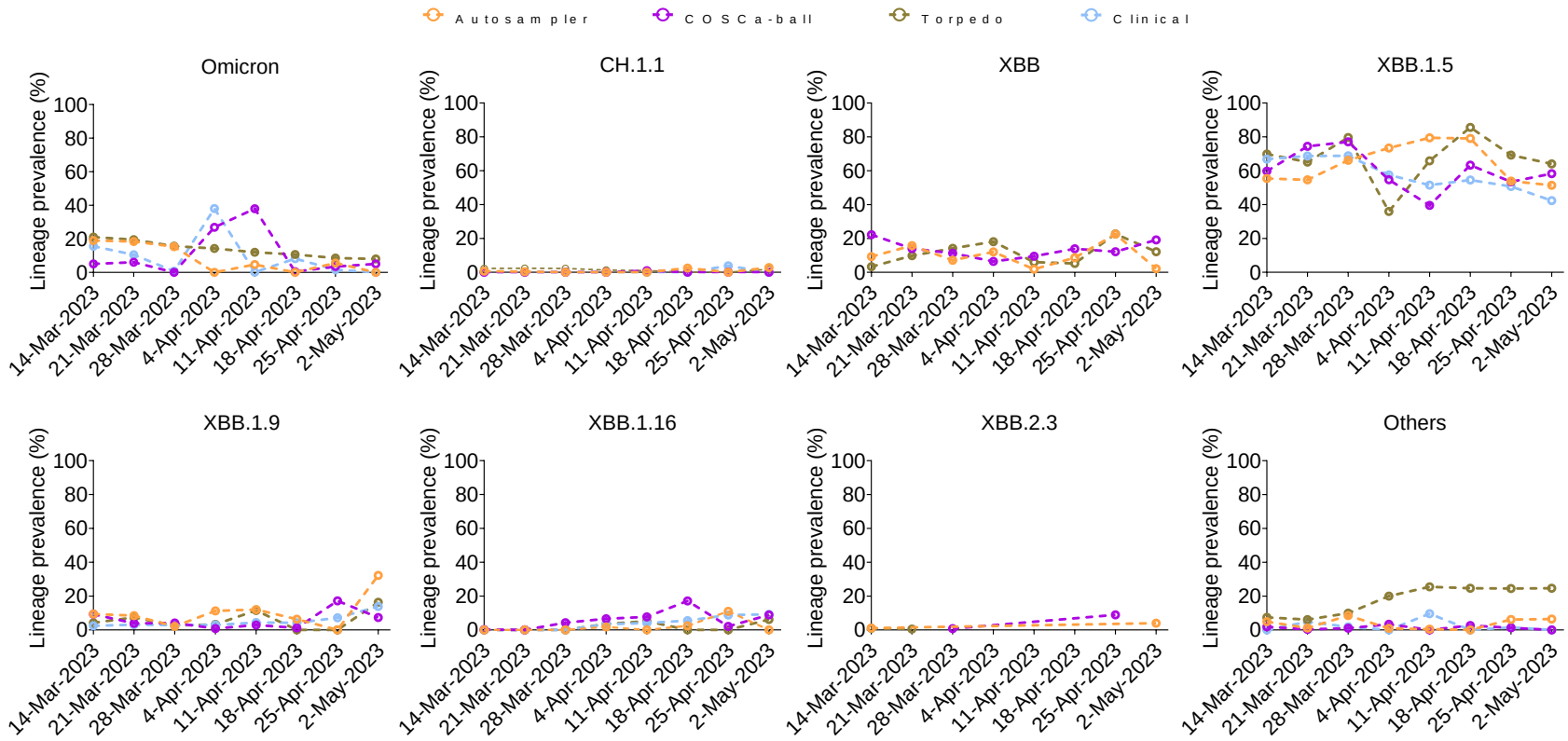

Figure 4-7: Weekly SARS-CoV-2 lineage prevalence across autosampler, COSCa-ball, Torpedo, and clinical surveillance data. The lineage prevalence in wastewater across the sampling methods was presented weekly, as the Ontario clinical surveillance data (PHO, 2023) were available on a weekly basis.

To evaluate the differences in lineages prevalence across autosampler, COSCa-ball, Torpedo and clinical surveillance data, a zero-inflated model was employed. The Zero-inflated model is suitable for datasets with an excess of zeros, as they can accurately fit the mean and variance while distinguishing between true zeros and counts modeled by a negative binomial distribution (Hall, 2000). The analysis showed that lineage prevalence across the sampling methods was comparable ($p > 0.05$) and concordant with the clinical surveillance data. However, although not statistically significant, the prevalence of each lineage in wastewater captured by the autosampler, COSCa-ball, and Torpedo samplers was consistently higher than that observed in clinical sequences throughout the sampling period. This may be attributed to the fact that WGS reflects lineages prevalence across an entire community, capturing SARS-CoV-2 RNA from both symptomatic and asymptomatic cases, whereas clinical surveillance generally reports only symptomatic individuals (Gupta et al., 2023; Li, et al., 2022). As the lineages prevalence observed in passive sampling was comparable to that obtained through autosampler, this study highlights the potential of passive sampling as a cost-effective alternative to high cost autosampler in large-scale WWRF settings for genomic surveillance of future pandemics beyond COVID-19.

4.6 WWGS using passive sampling for pandemic preparedness

WWGS has emerged as a valuable tool for monitoring circulating SARS-CoV-2 lineages within population and is based on the recovery of viral RNA fragments found in wastewaters (Khan et al., 2023; Wurtz et al., 2021). Due to its non-invasive nature in tracking the spread and evolution of circulating lineages, many countries have started using WWGS as a complementary surveillance method for population-level screening (Tiwari et al., 2023; Wardi et al., 2024). Despite the promising potential of WWGS for monitoring SARS-CoV-2 lineages, its implementation faces significant challenges arising from the mixed origin (i.e., contributions from multiple individuals), degraded state, and low concentration of genetic

material present in these sample types. Additionally, matrix and context-specific factors, such as the presence of PCR inhibitors, can critically affect its success. These complexities highlight the need for improved sample processing and more downstream haplotype phasing methods. Besides, It is also important to recognize the fast-paced progress of this research area throughout the pandemic with significant progress being made in both sample processing and data analysis. However, meaningful challenges remain, particularly for implementing WWGS in countries with limited funding resources. Besides, as the COVID-19 pandemic transitions toward endemicity, funding for WWGS has been gradually declining worldwide, limiting the ability of many regions to maintain robust surveillance efforts. To sustain wastewater surveillance through genome sequencing in a cost-effective manner, it is crucial to address key methodological issues, especially in the fundamental step of wastewater sampling.

WWGS primarily relies on composite wastewater though conventional auto sampling, which is costly, due to high capital, and ongoing operations and maintenance costs, and is limited to locations amenable to these devices (e.g., accessible to a maintenance crew, secure, temperature-controlled, and with adequate flow). As a cost-effective alternative, researchers have piloted passive sampling in small congregate settings, demonstrating its potential as a low-cost method for variant tracking (Corchis-Scott et al., 2021; Mangwana et al., 2022; Corchis-Scott et al., 2023). However, current funding for WWGS remains directed towards tracking SARS-CoV-2 lineages at the national and city level, often using WWRFs to provide community snapshots for millions of individuals with a single sample. While passive sampling could further reduce costs in these settings, its implementation in high-flow inlets of large-scale WWRFs present significant challenges. High flow rates, fluctuating wastewater composition, and diverse catchment areas can contribute to the dilution of SARS-CoV-2 viral particles and introduce variability. Additionally, SARS-CoV-2 RNA is often fragmented and degraded in wastewater (Anand et al., 2021; Zhakparov et al., 2023), raising further questions about

whether passive sampling in such environments can capture sufficient genomic material for sequencing and variant detection. Lastly, methodological standardization remains a challenge, as variations in sampling, processing, and sequencing protocols can lead to inconsistent results, further complicating large-scale deployment of passive sampling.

Our study demonstrated that SARS-CoV-2 RNA concentrations measured by N1, and N2 RT-qPCR were comparable among autosampler, COSCa-ball, and Torpedo samples. With similar N1 and N2 concentrations, our findings suggested that genomic recovery of SARS-CoV-2 from passive samplers may depend on sequencing read length, as longer read lengths enhanced genome recovery. When sequenced with longer reads (i.e., 2 × 300bp) on an Illumina platform, the genome recovery and sequencing parameters for SARS-CoV-2 were comparable across sampling methods. Genomic analyses further revealed that mutation-level (i.e., SNVs and Indels) profiles and lineages prevalence were consistent across sampling approaches and concordant with available clinical genomic surveillance, underscoring the capability of passive samplers to effectively capture viral RNA in WWRFs. This indicated that passive sampling is able to sufficiently replicate and preserve SARS-CoV-2 RNA signatures. Overall, similar mutations and lineages profiles emerged across sampling methods highlighting the potential of passive samplers for whole genome sequencing of emerging pathogens in high-flow wastewater contexts.

The 21st century has witnessed the vulnerability of global public health to an array of epidemics and pandemics, including SARS, MERS, 2009 H1N1 influenza, Ebola in West Africa, yellow fever in Angola, Cholera, Zika, Tuberculosis, Mpox, and the continuing HIV/AIDS and COVID-19 pandemics (GPMB, 2024). Additionally, climate change has accelerated the spread of vector-borne diseases such as malaria, chikungunya, Japanese encephalitis, dengue, kala-azar, and lymphatic filariasis, as warmer temperatures and shifting ecosystems create favorable conditions for novel or non-endemic pathogens to emerge or re-emerge as public health threats

(WHO, 2024). Meanwhile, antimicrobial resistance (AMR) is increasingly recognized as a silent pandemic due to the global proliferation of antibiotic-resistant genes (ARGs), which threaten to undermine current medical treatments and complicate the control of bacterial infections (Prestinaci et al., 2015). These health emergencies underscore the rising risk of future pandemics, which is evident by recent developments like the spillover of avian influenza H5N1 to cattle and humans which increases the risk of a zoonotic events (Venkatesan, 2023) and the emergence of a novel strain of a more pathogenic Mpox virus in East Africa with potential for global dissemination (Olufadewa et al., 2024). Besides, rising human interconnectedness and mobility facilitate the rapid transmission of endemic pathogens and the emergence of novel variants worldwide. In light of this evolving threat landscape, a proactive and adaptable approach to preparedness is essential. WWGS can therefore serve as a viable option for monitoring the spread and evolution of future pandemics at the population level, if the causative agents (i.e., RNA or DNA) of the diseases are excreted in significant quantities through feces and urine or other means and are able to adsorb be persistent in wastewaters. Implementing WWGS through passive sampling can enhance pandemic preparedness in a cost-effective manner, and findings of this study provide crucial evidence of its applicability of passive sampling in population levels. This reinforces the role of passive sampling may not only as a cost-efficient preparedness strategy but also as a means to advance the One Health approach, which integrates human, animal, and environmental health monitoring to address global health risks comprehensively.

4.7 Conclusion

This study demonstrated the viability of passive sampling for SARS-CoV-2 genomic surveillance at a high-flow WWRF. Our findings show that SARS-CoV-2 RNA concentrations were comparable between conventional autosampler and passive sampling techniques. Although similar viral RNA concentrations were captured from passive samples, the recovery

of genomes was influenced by sequencing read length, with longer reads enabling near-complete genome recovery. Genomic analyses revealed that SARS-CoV-2 mutation and lineage profiles were consistent across sampling methods and aligned closely with clinical surveillance data. The simplicity, affordability, and versatility of passive samplers make them a practical alternative to conventional composite sampling for SARS-CoV-2 WWGS and may also be applicable for other targets of public health concern. Overall, these findings support the use of passive sampling as a feasible and effective tool for WWGS, positioning it as a valuable method in future pandemic preparedness.

4.8 References

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

Alamin, M., Oladipo, P., Hartrick, J., Islam, N., Bahmani, A., Turner, C. L., Shuster, W., & Ram, J. L. (2024). Improved passive sampling methods for wastewater to enable more sensitive detection of SARS-CoV-2 and its variants. *Science of the Total Environment*, *950*. https://doi.org/10.1016/j.scitotenv.2024.175044

Allen, D. M., Reyne, M. I., Allingham, P., Levickas, A., Bell, S. H., Lock, J., Coey, J. D., Carson, S., Lee, A. J., McSparron, C., Nejad, B. F., McKenna, J., Shannon, M., Li, K., Curran, T., Broadbent, L. J., Downey, D. G., Power, U. F., Groves, H. E., … Gilpin, D. F. (2024). Genomic Analysis and Surveillance of Respiratory Syncytial Virus Using Wastewater-Based Epidemiology. *The Journal of Infectious Diseases*. https://doi.org/10.1093/infdis/jiae205

Amman, F., Markt, R., Endler, L., Hupfauf, S., Agerer, B., Schedl, A., Richter, L., Zechmeister, M., Bicher, M., Heiler, G., Triska, P., Thornton, M., Penz, T., Senekowitsch, M., Laine, J., Keszei, Z., Klimek, P., Nägele, F., Mayr, M., … Bergthaler, A. (2022). Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nature Biotechnology*, *40*(12), 1814–1822. https://doi.org/10.1038/s41587-022-01387-y

Anand, U., Adelodun, B., Pivato, A., Suresh, S., Indari, O., Jakhmola, S., Jha, H. C., Jha, P. K., Tripathi, V., & Di Maria, F. (2021). A review of the presence of SARS-CoV-2 RNA in wastewater and airborne particulates and its use for virus spreading surveillance. *Environmental Research*, *196*. https://doi.org/10.1016/j.envres.2021.110929

Antkiewicz, D. S., Janssen, K. H., Roguet, A., Pilch, H. E., Fahney, R. B., Mullen, P. A., Knuth, G. N., Everett, D. G., Doolittle, E. M., King, K., Wood, C., Stanley, A., Hemming, J. D. C., & Shafer, M. M. (2024). Wastewater-based protocols for SARS-CoV-2: insights into virus concentration, extraction, and quantitation methods from two years of public health surveillance. *Environmental Science: Water Research and Technology*. https://doi.org/10.1039/d3ew00958k

Augusto, M. R., Claro, I. C. M., Siqueira, A. K., Sousa, G. S., Caldereiro, C. R., Duran, A. F. A., de Miranda, T. B., Bomediano Camillo, L. de M., Cabral, A. D., & de Freitas Bueno, R. (2022). Sampling strategies for wastewater surveillance: Evaluating the variability of SARS-COV-2 RNA concentration in composite and grab samples. *Journal of Environmental Chemical Engineering*, *10*(3). https://doi.org/10.1016/j.jece.2022.107478

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bivins, A., Kaya, D., Ahmed, W., Brown, J., Butler, C., Greaves, J., Leal, R., Maas, K., Rao, G., Sherchan, S., Sills, D., Sinclair, R., Wheeler, R. T., & Mansfeldt, C. (2022). Passive sampling to scale wastewater surveillance of infectious disease: Lessons learned from COVID-19. *Science of the Total Environment* 835. https://doi.org/10.1016/j.scitotenv

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Cha, G., Graham, K. E., Zhu, K. J., Rao, G., Lindner, B. G., Kocaman, K., Woo, S., D'amico, I., Bingham, L. R., Fischer, J. M., Flores, C. I., Spencer, J. W., Yathiraj, P., Chung, H., Biliya, S., Djeddar, N., Burton, L. J., Mascuch, S. J., Brown, J., … Konstantinidis, K. T. (2023). Parallel deployment of passive and composite samplers for surveillance and variant profiling of SARS-CoV-2 in sewage. *Science of the Total Environment, 866*. https://doi.org/10.1016/j.scitotenv.2022.161101

Chen, L., Chen, A., Zhang, X. D., Robles, M. T. S., Han, H. S., Xiao, Y., Xiao, G., Pipas, J. M., & Weitz, D. A. (2024). Targeted whole-genome recovery of single viral species in a complex environmental sample. *Proceedings of the National Academy of Sciences of the United States of America, 121*(31). https://doi.org/10.1073/pnas.2404727121

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Child, H. T., Airey, G., Maloney, D. M., Parker, A., Wild, J., McGinley, S., Evens, N., Porter, J., Templeton, K., Paterson, S., Van Aerle, R., Wade, M. J., Jeffries, A. R., & Bassano, I.

(2023). Comparison of metagenomic and targeted methods for sequencing human pathogenic viruses from wastewater. *MBio*, *14*(6). https://doi.org/10.1128/mbio.01468-23

Corchis-Scott, R., Geng, Q., Al Riahi, A. M., Labak, A., Podadera, A., Ng, K. K. S., Porter, L. A., Tong, Y., Dixon, J. C., Menard, S. L., Seth, R., & McKay, R. M. (2023). Actionable wastewater surveillance: application to a university residence hall during the transition between Delta and Omicron resurgences of COVID-19. *Frontiers in Public Health*, *11*. https://doi.org/10.3389/fpubh.2023.1139423

Corchis-Scott, R., Geng, Q., Seth, R., Ray, R., Beg, M., Biswas, N., Charron, L., Drouillard, K. D., D'Souza, R., Heath, D. D., Houser, C., Lawal, F., McGinlay, J., Menard, S. L., Porter, L. A., Rawlings, D., Scholl, M. L., Siu, K. W. M., Tong, Y., … McKay, R. M. L. (2021). Averting an Outbreak of SARS-CoV-2 in a University Residence Hall through Wastewater Surveillance. *Microbiology Spectrum*, *9*(2). https://doi.org/10.1128/spectrum.00792-21

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

Dehghan Banadaki, M., Torabi, S., Rockward, A., Strike, W. D., Noble, A., Keck, J. W., & Berry, S. M. (2024). Simple SARS-CoV-2 concentration methods for wastewater surveillance in low resource settings. *Science of the Total Environment*, *912*. https://doi.org/10.1016/j.scitotenv.2023.168782

EPA. (2001). Method 1684: Total, Fixed, and Volatile Solids in Water, Solids, and Biosolids. *U. S. Environmental Protection Agency*, *EPA-821-R-*(January), 13.

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. In *Nature Biotechnology* (Vol. 38, Issue 3, pp. 276–278). Nature Research. https://doi.org/10.1038/s41587-020-0439-x

Feng, S., Owens, S. M., Shrestha, A., Poretsky, R., Hartmann, E. M., & Wells, G. (2023). Intensity of sample processing methods impacts wastewater SARS-CoV-2 whole genome amplicon sequencing outcomes. *Science of the Total Environment*, *876*. https://doi.org/10.1016/j.scitotenv.2023.162572

Fontenele, R. S., Kraberger, S., Hadfield, J., Driver, E. M., Bowes, D., Holland, L. R. A., Faleye, T. O. C., Adhikari, S., Kumar, R., Inchausti, R., Holmes, W. K., Deitrick, S., Brown, P., Duty, D., Smith, T., Bhatnagar, A., Yeager, R. A., Holm, R. H., von Reitzenstein, N. H., … Varsani, A. (2021). High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *Water Research*, *205*. https://doi.org/10.1016/j.watres.2021.117710

Gálvez, J. M., Chaparro-Solano, H. M., Pinzón-Rondón, Á. M., Albornoz, L. L., Pardo-Oviedo, J. M., Zapata-Gómez, F. A., Patiño-Aldana, A. F., Hernández-Rodríguez, A. del P., Díaz-Quiroz, M., & Ruiz-Sternberg, Á. M. (2022). Mutation profile of SARS-CoV-2 genome in a sample from the first year of the pandemic in Colombia. *Infection, Genetics and Evolution*, *97*. https://doi.org/10.1016/j.meegid.2021.105192

Gangavarapu, K., Latif, A. A., Mullen, J. L., Alkuzweny, M., Hufbauer, E., Tsueng, G., Haag, E., Zeller, M., Aceves, C. M., Zaiets, K., Cano, M., Zhou, X., Qian, Z., Sattler, R., Matteson, N. L., Levy, J. I., Lee, R. T. C., Freitas, L., Maurer-Stroh, S., … Hughes, L. D. (2023). Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-

2 variants and mutations. *Nature Methods*, *20*(4), 512–522. https://doi.org/10.1038/s41592-023-01769-3

GPMB Pandemic Risk Report (2024). *Global Preparedness Monitoring Board*, Geneva: World Health Organization; 2024. Licence: CC BY-NC-SA 3.0 IGO

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

Gupta, P., Liao, S., Ezekiel, M., Novak, N., Rossi, A., LaCross, N., Oakeson, K., & Rohrwasser, A. (2023). Wastewater Genomic Surveillance Captures Early Detection of Omicron in Utah. *Microbiology Spectrum*, *11*(3). https://doi.org/10.1128/spectrum.00391-23

Habtewold, J., McCarthy, D., McBean, E., Law, I., Goodridge, L., Habash, M., & Murphy, H. M. (2022). Passive sampling, a practical method for wastewater-based surveillance of SARS-CoV-2. *Environmental Research* 204. https://doi.org/10.1016/j.envres.2021.11205

Hall, D. B. (2000). *Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study*.

Harris-Lovett, S., Nelson, K. L., Beamer, P., Bischel, H. N., Bivins, A., Bruder, A., Butler, C., Camenisch, T. D., De Long, S. K., Karthikeyan, S., Larsen, D. A., Meierdiercks, K., Mouser, P. J., Pagsuyoin, S., Prasek, S. M., Radniecki, T. S., Ram, J. L., Keith Roper, D., Safford, H., … Korfmacher, K. S. (2021). Wastewater surveillance for sars-cov-2 on college campuses: Initial efforts, lessons learned and research needs. *International*

Journal of Environmental Research and Public Health, *18*(9). https://doi.org/10.3390/ijerph18094455

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

Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Oude Munnink, B. B., Schapendonk, C. M. E., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F. M., Lycett, S., Medema, G., Koopmans, M. P. G., & De Graaf, M. (2021). Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the netherlands and belgium. *Emerging Infectious Diseases*, *27*(5), 1405–1415. https://doi.org/10.3201/eid2705.204410

Jahn, K., Dreifuss, D., Topolsky, I., Kull, A., Ganesanandamoorthy, P., Fernandez-Cassi, X., Bänziger, C., Devaux, A. J., Stachler, E., Caduff, L., Cariti, F., Corzón, A. T., Fuhrmann, L., Chen, C., Jablonski, K. P., Nadeau, S., Feldkamp, M., Beisel, C., Aquino, C., … Beerenwinkel, N. (2022). Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. *Nature Microbiology*, *7*(8), 1151–1160. https://doi.org/10.1038/s41564-022-01185-x

Kabir, M. P., Mercier, É., Eid, W., Plaza-Diaz, J., D'Aoust, P. M., Landgraff, C., Goodridge, L., Lawal, O. U., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Thakali, O., Pisharody, L., Stephenson, S., Graber, T. E., & Delatolla, R. (2025a). Diagnostic performance of allele-specific RT-qPCR and genomic sequencing in wastewater-based surveillance of SARS-CoV-2. *Eco-Environment and Health*, *4*(1). https://doi.org/10.1016/j.eehl.2025.100135

Kabir, M. P., Renouf, E., Pisharody, L., Mercier, É., D'Aoust, P. M., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Tomalty, E., Graber, T. E., & Delatolla, R. (2025b). Wastewater solids drive comparability of sampling methods for SARS-CoV-2 wastewater and environmental surveillance. *Journal of Environmental Chemical Engineering*, *13*(1). https://doi.org/10.1016/j.jece.2025.115374

Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, *609*(7925), 101–108. https://doi.org/10.1038/s41586-022-05049-6

Karthikeyan, S., Nguyen, A., McDonald, D., Zong, Y., Ronquillo, N., Ren, J., Zou, J., Farmer, S., Humphrey, G., Henderson, D., Javidi, T., Messer, K., Anderson, C., Schooley, R., Martin, N. K., Knight, R., & Karthikeyan, C. S. (2021). *msystems*. *6*, 793–814. https://doi.org/10.1128/mSystems

Kevill, J. L., Lambert-Slosarska, K., Pellett, C., Woodhall, N., Richardson-O'Neill, I., Pântea, I., Alex-Sanders, N., Farkas, K., & Jones, D. L. (2022). Assessment of two types of passive sampler for the efficient recovery of SARS-CoV-2 and other viruses from wastewater. *Science of the Total Environment*, *838*. https://doi.org/10.1016/j.scitotenv.2022.156580

Khan, M., Li, L., Haak, L., Payen, S. H., Carine, M., Adhikari, K., Uppal, T., Hartley, P. D., Vasquez-Gross, H., Petereit, J., Verma, S. C., & Pagilla, K. (2023). Significance of wastewater surveillance in detecting the prevalence of SARS-CoV-2 variants and other respiratory viruses in the community – A multi-site evaluation. *One Health*, *16*. https://doi.org/10.1016/j.onehlt.2023.100536

Klapsa, D., Wilton, T., Zealand, A., Bujaki, E., Saxentoff, E., Troman, C., Shaw, A. G., Tedcastle, A., Majumdar, M., Mate, R., Akello, J. O., Huseynov, S., Zeb, A., Zambon, M., Bell, A., Hagan, J., Wade, M. J., Ramsay, M., Grassly, N. C., … Martin, J. (2022). Sustained detection of type 2 poliovirus in London sewage between February and July, 2022, by enhanced environmental surveillance. *The Lancet*, *400*(10362), 1531–1538. https://doi.org/10.1016/S0140-6736(22)01804-9

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

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

Lee, A. J., Carson, S., Reyne, M. I., Marshall, A., Moody, D., Allen, D. M., Allingham, P., Levickas, A., Fitzgerald, A., Bell, S. H., Lock, J., Coey, J. D., McSparron, C., Nejad, B. F., Troendle, E. P., Simpson, D. A., Courtney, D. G., Einarsson, G. G., McKenna, J. P., … Bamford, C. G. G. (2024). Wastewater monitoring of human and avian influenza A viruses in Northern Ireland: a genomic surveillance study. *The Lancet Microbe*. https://doi.org/10.1016/S2666-5247(24)00175-7

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

Li, L., Mazurowski, L., Dewan, A., Carine, M., Haak, L., Guarin, T. C., Dastjerdi, N. G., Gerrity, D., Mentzer, C., & Pagilla, K. R. (2022). Longitudinal monitoring of SARS-CoV-

2 in wastewater using viral genetic markers and the estimation of unconfirmed COVID-19 cases. *Science of the Total Environment*, *817*. https://doi.org/10.1016/j.scitotenv.2022.152958

Li, L., Uppal, T., Hartley, P. D., Gorzalski, A., Pandori, M., Picker, M. A., Verma, S. C., & Pagilla, K. (2022). Detecting SARS-CoV-2 variants in wastewater and their correlation with circulating variants in the communities. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20219-2

Li, X., Zhang, S., Shi, J., Luby, S. P., & Jiang, G. (2021). Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. In *Chemical Engineering Journal* (Vol. 415). Elsevier B.V. https://doi.org/10.1016/j.cej.2021.129039

Lieberman-Cribbin, W., Tuminello, S., Flores, R. M., & Taioli, E. (2020). Disparities in COVID-19 Testing and Positivity in New York City. *American Journal of Preventive Medicine*, *59*(3), 326–332. https://doi.org/10.1016/j.amepre.2020.06.005

Ling-Hu, T., Rios-Guzman, E., Lorenzo-Redondo, R., Ozer, E. A., & Hultquist, J. F. (2022). Challenges and Opportunities for Global Genomic Surveillance Strategies in the COVID-19 Era. In *Viruses* (Vol. 14, Issue 11). https://doi.org/10.3390/v14112532

Majid, F., Omer, S. B., & Khwaja, A. I. (2020). Optimising SARS-CoV-2 pooled testing for low-resource settings. In *The Lancet Microbe* (Vol. 1, Issue 3, pp. e101–e102). Elsevier Ltd. https://doi.org/10.1016/S2666-5247(20)30056-2

Mangwana, N., Archer, E., Muller, C. J. F., Preiser, W., Wolfaardt, G., Kasprzyk-Hordern, B., Carstens, A., Brocker, L., Webster, C., McCarthy, D., Street, R., Mathee, A., Louw, J., Mdhluli, M., & Johnson, R. (2022). Sewage surveillance of SARS-CoV-2 at student campus residences in the Western Cape, South Africa. *Science of the Total Environment*, *851*. https://doi.org/10.1016/j.scitotenv.2022.158028

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K. K., Wilkinson, R., & Wiedenheft, B. (2020). Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Reports Medicine*, *1*(6). https://doi.org/10.1016/j.xcrm.2020.100098

Olufadewa, I. I., Oladele, R. I., Olajide, O. A., Adetunji, H. T., Okoduwa, G. E., Olufadewa, T. A., & Adesina, M. A. (2024). Emergence of a more virulent clade of Mpox in Africa: Learning from history and charting a path forward. *Journal of Medicine, Surgery, and Public Health*, *3*, 100134. https://doi.org/10.1016/j.glmedi.2024.100134

Overton, A. K., Knapp, J. J., Lawal, O. U., Gibson, R., Fedynak, A. A., Adebiyi, A. I., Maxwell, B., Cheng, L., Bee, C., Qasim, A., Atanas, K., Payne, M., Stuart, R., Fleury, M. D., Knox, N. C., Nash, D., Hungwe, Y. C., Prasla, S. R., Ho, H., … Charles, T. C. (2024). Genomic surveillance of Canadian airport wastewater samples allows early detection of emerging SARS-CoV-2 lineages. *Scientific Reports*, *14*(1), 26534. https://doi.org/10.1038/s41598-024-76925-6

Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: Quick coverage calculation for genomes and exomes. *Bioinformatics*, *34*(5), 867–868. https://doi.org/10.1093/bioinformatics/btx699

Pérez-Cataluña, A., Chiner-Oms, Á., Cuevas-Ferrando, E., Díaz-Reolid, A., Falcó, I., Randazzo, W., Girón-Guzmán, I., Allende, A., Bracho, M. A., Comas, I., & Sánchez, G. (2022). Spatial and temporal distribution of SARS-CoV-2 diversity circulating in wastewater. *Water Research*, *211*. https://doi.org/10.1016/j.watres.2021.118007

Pérez-Cataluña, A., Cuevas-Ferrando, E., Randazzo, W., Falcó, I., Allende, A., & Sánchez, G. (2021). Comparing analytical methods to detect SARS-CoV-2 in wastewater. *Science of the Total Environment*, *758*. https://doi.org/10.1016/j.scitotenv.2020.143870

Public Health Ontario (2023). Archive of SARS-CoV-2 Whole Genome Sequencing in Ontario, March 12 to May 5, 2023.

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

Reitsma, M. B., Claypool, A. L., Vargo, J., Shete, P. B., McCorvie, R., Wheeler, W. H., Rocha, D. A., Myers, J. F., Murray, E. L., Bregman, B., Dominguez, D. M., Nguyen, A. D., Porse, C., Fritz, C. L., Jain, S., Watt, J. P., Salomon, J. A., & Goldhaber-Fiebert, J. D. (2021). Racial/ethnic disparities in covid-19 exposure risk, testing, and cases at the subcounty level in California. *Health Affairs*, *40*(6), 870–878. https://doi.org/10.1377/hlthaff.2021.00098

Robishaw, J. D., Alter, S. M., Solano, J. J., Shih, R. D., DeMets, D. L., Maki, D. G., & Hennekens, C. H. (2021). Genomic surveillance to combat COVID-19: challenges and opportunities. In *The Lancet Microbe* (Vol. 2, Issue 9, pp. e481–e484). Elsevier Ltd. https://doi.org/10.1016/S2666-5247(21)00121-X

Saldivar-Espinoza, B., Garcia-Segura, P., Novau-Ferré, N., Macip, G., Martínez, R., Puigbò, P., Cereto-Massagué, A., Pujadas, G., & Garcia-Vallve, S. (2023). The Mutational Landscape of SARS-CoV-2. *International Journal of Molecular Sciences*, *24*(10). https://doi.org/10.3390/ijms24109072

Schang, C., Crosbie, N. D., Nolan, M., Poon, R., Wang, M., Jex, A., John, N., Baker, L., Scales, P., Schmidt, J., Thorley, B. R., Hill, K., Zamyadi, A., Tseng, C. W., Henry, R., Kolotelo, P., Langeveld, J., Schilperoort, R., Shi, B., … McCarthy, D. T. (2021). Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. *Environmental Science and Technology*, *55*(15), 10432–10441. https://doi.org/10.1021/acs.est.1c01530

Simon Andrews. (2020). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. In *Soil* (Vol. 5, Issue 1, pp. 47–81). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28246

Spurbeck, R. R., Minard-Smith, A., & Catlin, L. (2021). Feasibility of neighborhood and building scale wastewater-based genomic epidemiology for pathogen surveillance. *Science of the Total Environment*, *789*. https://doi.org/10.1016/j.scitotenv.2021.147829

Tiwari, A., Adhikari, S., Zhang, S., Solomon, T. B., Lipponen, A., Islam, M. A., Thakali, O., Sangkham, S., Shaheen, M. N. F., Jiang, G., Haramoto, E., Mazumder, P., Malla, B., Kumar, M., Pitkänen, T., & Sherchan, S. P. (2023). Tracing COVID-19 Trails in Wastewater: A Systematic Review of SARS-CoV-2 Surveillance with Viral Variants. In *Water (Switzerland)* (Vol. 15, Issue 6). MDPI. https://doi.org/10.3390/w15061018

Venkatesan, P. (2023). Avian influenza spillover into mammals. *The Lancet Microbe*, *4*(7), e492. https://doi.org/10.1016/s2666-5247(23)00173-8

Vincent-Hubert, F., Wacrenier, C., Desdouits, M., Jousse, S., Schaeffer, J., Le Mehaute, P., Nakache-Danglot, F., Le Guyader, F. S., Bertrand, I., Boni, M., Gantzer, C., Maday, Y., Marechal, V., Mouchel, J. M., Moulin, L., & Wurtzer, S. (2022). Development of passive samplers for the detection of SARS-CoV-2 in sewage and seawater: Application for the monitoring of sewage. *Science of the Total Environment*, *833*. https://doi.org/10.1016/j.scitotenv.2022.155139

Vo, V., Tillett, R. L., Chang, C. L., Gerrity, D., Betancourt, W. Q., & Oh, E. C. (2022). SARS-CoV-2 variant detection at a university dormitory using wastewater genomic tools. *Science of the Total Environment*, *805*. https://doi.org/10.1016/j.scitotenv.2021.149930

Wardi, M., Belmouden, A., Aghrouch, M., Lotfy, A., Idaghdour, Y., & Lemkhente, Z. (2024). Wastewater genomic surveillance to track infectious disease-causing pathogens in low-income countries: Advantages, limitations, and perspectives. *Environment International*, *192*. https://doi.org/10.1016/j.envint.2024.109029

World Health Organization (WHO, 2024). Vector-borne diseases. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases. (Accessed on 08 December 2024).

Williams, R. C., Farkas, K., Garcia-Delgado, A., Adwan, L., Kevill, J. L., Cross, G., Weightman, A. J., & Jones, D. L. (2024). Simultaneous detection and characterization of common respiratory pathogens in wastewater through genomic sequencing. *Water Research*, *256*. https://doi.org/10.1016/j.watres.2024.121612

Wurtz, N., Revol, O., Jardot, P., Giraud-Gatineau, A., Houhamdi, L., Soumagnac, C., Annessi, A., Lacoste, A., Colson, P., Aherfi, S., & La Scola, B. (2021). Monitoring the circulation of SARS-CoV-2 variants by genomic analysis ofwastewater in marseille, south-east France. *Pathogens*, *10*(8), 1–14. https://doi.org/10.3390/pathogens10081042

Zhakparov, D., Quirin, Y., Xiao, Y., Battaglia, N., Holzer, M., Bühler, M., Kistler, W., Engel, D., Zumthor, J. P., Caduff, A., & Baerenfaller, K. (2023). Sequencing of SARS-CoV-2 RNA Fragments in Wastewater Detects the Spread of New Variants during Major Events. *Microorganisms*, *11*(11). https://doi.org/10.3390/microorganisms11112660

5 Chapter 5: Optimization of Primary Sludge Processing Method for Wastewater Genomic Surveillance of SARS-CoV-2

5.1 Context

Chapter 5 represents the research article entitled *Optimization of primary sludge processing method for wastewater genomic surveillance of SARS-CoV-2* by Kabir, M. P., Plaza-Diaz, J., Mercier, E., D'Aoust, P. M., Goodridge, L., Lawal, O. U., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Addo, F., Renouf, E., Graber, T. E., and Delatolla, R. This article has been submitted for publication to the journal of ACS ES&T Water in 2025. This study focuses on developing a primary sludge processing method for SARS-CoV-2 genome sequencing that yields comparable or better results to conventional wastewater genomic sequencing.

5.2 Abstract

Wastewater genomic surveillance (WWGS) of SARS-CoV-2 is typically performed using influent wastewater, but the approach is challenging due to degradation as well as low target concentrations in wastewater. This could be alleviated by utilizing the higher concentrations of SARS-CoV-2 RNA found in primary sludge, however this matrix is prone to sequencing library failures. Our study focuses on developing a robust primary sludge-based SARS-CoV-2 genome sequencing method. The study was conducted using 30 parallel influent wastewater and primary sludge samples collected during three different time periods, under three clinically predominant SARS-CoV-2 Omicron lineages in Ottawa, Canada. Results showed that our approach consistently recovered near complete (≥ 90%) SARS-CoV-2 genomes from both influent wastewater or primary sludge samples. Prevalent lineage and single nucleotide variant (SNV) profiles were identical between wastewater and primary sludge. However, there was a higher likelihood of rare (low prevalence) and Canadian cryptic (not detected in clinical surveillance) SNVs in primary sludge compared to influent wastewater. Overall, our approach

enables the sequencing of the most concentrated sources of genetic material within the wastewater matrix, providing valuable insights for public health forecasting of infectious disease prevalence beyond the COVID-19 pandemic.

5.3 Introduction

The COVID-19 pandemic caused by SARS-CoV-2 has highlighted the importance of effective surveillance systems to monitor viral spread in population levels. Clinical surveillance is the conventional method to monitor the prevalence of COVID-19, however clinical surveillance metrics overlooks the asymptomatic carriers as well as the individuals without access to medical facilities, leading to an incomplete picture of incidence and prevalence in the populations (Lieberman-Cribbin et al., 2020; Reitsma et al., 2021). As a complement to clinical surveillance, PCR based wastewater and environmental surveillance (WES) of SARS-CoV-2 RNA has emerged as a powerful tool for monitoring COVID-19 at various levels, offering early warning of outbreaks within the population (Ahmed et al., 2021; Daughton, 2020; Medema et al., 2020). With the progression of the pandemic, SARS-CoV-2 continues to evolve, leading to the emergence of multiple variants of concern (VOCs). These variants have been linked to increased transmissibility, disease severity, reinfection, and reduced vaccine effectiveness (Davies et al., 2021; Tao et al., 2021; Wang et al., 2021). To monitor the evolving landscape of SARS-CoV-2 variants, clinical genomic surveillance through the whole genome sequencing of patient samples is the standard practice. However, clinical genomic surveillance is also limited to symptomatic individuals and is further constrained by the high cost and resource-intensive nature of sequencing samples from individual patients (Vashisht et al., 2023). Besides, as the pandemic wanes, clinical genomic surveillance becomes less capable of describing the trajectory of emerging VOCs due to widespread vaccination efforts which has lessened the severity of infections and the need for clinical diagnosis (Ling-Hu et al., 2022; Robishaw et al., 2021). Therefore, WES has become in some regions and countries the only

remaining viable option to monitor the temporal changes of these variants and their lineages within the population. The qPCR-based WES is limited to measuring SARS-CoV-2 RNA in wastewater (except allele specific RT-qPCR) but cannot track the prevalence of emerging variants in the population. To monitor the evolution of SARS-CoV-2 variants, WWGS has been adopted as a reliable and cost-effective method. WWGS often entails the collection of influent wastewater samples, followed by viral RNA extraction and sequencing to identify and characterize the distinct strains or variants of SARS-CoV-2 circulating within the population (Ai et al., 2021; Li, Uppal, et al., 2022; Vo et al., 2022). WWGS has demonstrated the detection of circulating, early and cryptic mutations and viral lineages in the population (Jahn et al., 2022; Karthikeyan et al., 2022; Smyth et al., 2022). Although results of WWGS have faithfully mirrored clinical genomic surveillance in many jurisdictions, researchers routinely encounter difficulties in recovering complete genomes of SARS-CoV-2 from influent wastewater due to the fragmented nature and limited amount of the RNA target in wastewater (Bar-Or et al., 2021; Fontenele et al., 2021; Patel et al., 2021). This makes the detection and tracking of low prevalence or rare lineages and mutations of SARS-CoV-2 in wastewater challenging.

To resolve this issue, clarified primary sludge is theoretically a more appropriate location for sample collection compared to influent wastewater for WWGS, as SARS-CoV-2 RNA is found in higher concentrations in primary sludge as compared to influent wastewater (Balboa et al., 2021; D'Aoust et al., 2021a; Kim et al., 2022). However, in light of the higher concentration of SARS-CoV-2 genomic targets in primary sludge, sequencing efforts are more complex and challenging due to the anticipated presence of enriched PCR inhibitors in sludge matrices (Peccia et al., 2020). Consequently, limited efforts have been documented in scientific literature to sequence the SARS-CoV-2 genome directly from primary sludge. Early in the pandemic, Lin et al., (2021) attempted to sequence the whole genome of SARS-CoV-2 from primary sludge but failed to produce library for recovering near-complete (≥ 90%) genomes

from those samples. They expressed the reasons to be the presence of PCR inhibitors in primary sludge which inhibits library preparation, while other impurities caused undesired amplification resulting in poor quality sequencing data along with incomplete genomic coverage. Lin et al., (2021) also assumed that SARS-CoV-2 genome is less intact in primary sludge, which led to significant RNA degradation during direct sludge processing. Although, later Baaijens et al., (2022) revealed the recovery of near complete genome of SARS-CoV-2 from selected primary sludge samples (< 30% of those sampled) where the cycling threshold (Ct) value was less than 31, and the genomic recovery decreased with the Ct value increased. However, a critical comparison between these two studies suggests that differences in sludge processing methods may be the most influential factor affecting the successful recovery of near-complete genomes from primary sludge samples. This highlighted the need for optimized processing methods to recover near-complete genomes in primary sludge, and resolve challenges associated with WWGS of SARS-CoV-2 using primary sludge samples.

In this study, we developed a robust approach by diluting primary sludge in deionized water at a 1:70 ratio and processed using Nanotrap® Magnetic Virus Particles for whole genome sequencing of SARS-CoV-2. The dilution of primary sludges likely mitigates PCR inhibitors (Association of Public Health Laboratories, 2022), while the high surface area and specific binding properties of Nanotrap® Magnetic Virus Particles are expected to remove impurities (Andersen et al., 2023), thereby enriching and facilitating the recovery of intact SARS-CoV-2 genomes from primary sludge. In parallel, we processed influent wastewater using the same method. We also compared the detected lineage (i.e., a genetically distinct SARS-CoV-2 genome) profiles and SNVs (i.e., single-nucleotide variations across sequencing reads) counts and their genomic and temporal distributions between wastewater and primary sludge. We further evaluated the identified SNVs in wastewater and primary sludge against the Canadian and global clinical sequences during the sampling periods as well as throughout the pandemic

to determine the presence of rare (low prevalence) and cryptic SNVs in wastewater and primary sludge samples. Our approach establishes the robust recovery of near-complete SARS-CoV-2 genomes from both influent wastewater or primary sludge samples. Importantly, comparable, or slightly better performance metrics were obtained in primary sludge as compared to influent wastewater samples across the study periods.

5.4 Materials and Methods

5.4.1 Wastewater sampling

A total of thirty pairs of 24 hour-composite influent wastewater and primary sludge samples were collected in three distinct periods (first: April 02 to April 24, 2022; second: September 06 to October 01, 2023; and third: November 25 to December 17, 2023) from the City of Ottawa's Robert O. Pickard Environmental Center (ROPEC) using autosamplers (Hoskin Scientific, Burlington, Canada). Influent wastewater refers to the untreated municipal wastewater entering to ROPEC for treatment while primary sludge is the same influent wastewater following several hours of gravitational settling in the primary sedimentation basins at ROPEC. On each sampling day, 500 mL of influent wastewater (collected as ~20 mL aliquots per hour) and 500 mL of primary sludge (collected as ~125 mL aliquots at 6-hour intervals) were obtained over a 24-hour composite period. The samples were transferred to the laboratory in an ice cooler packed and stored at 4 °C for a maximum of 24 hours prior to analysis.

5.4.2 Concentration, nucleic acid extraction and detection of SARS-CoV-2 RNA

The primary sludge samples were processed for RNA extraction using our optimized sludge processing methods for sequencing. Briefly, primary sludge samples were first diluted in deionized water at 1:70 ratio by volume (0.7 mL of primary sludge with 49.3 mL of deionized water), mixed thoroughly and incubated at room temperature for 10 minutes in 50 mL conical

tubes. Similarly, 50 mL wastewater was left at room temperature for 10 minutes to settle the larger particles. Afterwards, 40 mL samples from the top were transferred to centrifuge tubes, followed by the addition of 600 µL Nanotrap® Magnetic Virus Particles (Ceres Nanosciences, Manassas, USA). The mixture was rotated at 100 rpm for 20 minutes at 20º C, and centrifuged for 10 minutes at 4º C at 8000 x g. The centrifuge tubes were placed on a magnetic rack (DynaMag-50, Invitrogen, Massachusetts, USA), and supernatant was discarded by pipetting without disturbing the pellet. The pellet was resuspended in 140 µL of PBS buffer (pH 7.4) and 560 µL of Viral Lysis Buffer (Qiagen, Germantown, USA). The resulting suspension was transferred to a 2.0 mL microcentrifuge tube and placed on a magnetic rack (DynaMag™-2 magnet, Invitrogen, Massachusetts, USA) for 10 minutes to remove the magnetic particles. The lysate was carefully collected by pipette, and total nucleic acid was extracted using QIAmp Viral RNA Mini Kit (Qiagen) and eluted in 70 µL of nuclease-free water. The SARS-CoV-2 RNA was detected by singleplex RT-qPCR targeting the N1 loci using TaqMan® Fast Virus 1-Step Master Mix (ThermoFisher, Ottawa, Canada) on a CFX Connect qPCR thermocycler (Bio-Rad, Hercules, Canada) as previously described protocol (D'Aoust, et al., 2021b).

5.4.3 cDNA synthesis and SARS-CoV-2 tiled amplicon generation

First, any contaminating DNA from extract was removed using ezDNase™ Enzyme kit (ThermoFisher), and RNA was reverse transcribed using random hexamer and SuperScript IV First-Strand Synthesis System (ThermoFisher). Briefly, 10 µL of DNA-free RNA was mixed with 1 µL each of 50 ng/µL random hexamers, 10 mM dNTP mix, and nuclease-free water, and incubated at 65 ^{0}C for 5 minutes. The mixtures were chilled on ice for at least for 2 minutes, and 7 µL of mastermix comprised of 4 µL of 5x SSIV Buffer, 1 µL of each 100 mM DTT, Ribonuclease Inhibitor, SuperScript IV Reverse Transcriptase (200 U/µL) was added, and serially incubated at 23º C for 10 minutes, 55º C for 10 minutes, and 80º C for 10 minutes. Subsequently, SARS-CoV-2 cDNA was amplified through multiplexed PCR using ARTIC

primer set v4.1. Each PCR reaction contained 12.5 µL Q5 Hot Start Master Mix (BioLabs, New England, Massachusetts, USA), 3.6 µL 10 uM of primer pool (IDT, Ottawa, Canada), 2.9 µL nuclease-free water and 6 µL of the cDNA. Samples were amplified with the following conditions: initial activation at 98º C for 30 seconds, followed by 35 cycles of denaturation at 95º C for 15 seconds, and annealing and extension at 63º C for 5 minutes. The amplicon size was verified in Bioanalyzer 2100 system using High Sensitivity DNA Kit (Agilent Technologies, Mississauga, Canada), purified with 0.8x AMPure XP beads (Beckman Coulter, California, USA) and quantified using a Qubit 4 fluorometer (ThermoFisher) using the dsDNA High Sensitivity Kit (Invitrogen).

5.4.4 Library preparation and sequencing

The SARS-CoV-2 library was prepared using Nextera XT DNA Library Preparation kit (Illumina, San Diego, USA) as per the manufacturer's instructions. Briefly, amplicons were tagmented at 55º C for 10 minutes, followed by a 15-cycle PCR with the following conditions: 72º C for 3 minutes, 95º C for 30 seconds and 15 cycles of 95º C for 10 seconds, 55º C for 30 seconds and 72º C for 30 seconds before a single step at 72º C for 5 minutes. The size distribution and quality of the resulting library were verified and quantified as described above. Libraries were normalized and sequenced on a MiSeq Illumina platform with paired end reads of (2 × 150 bp) for the first period of sampling and (2 × 300 bp) for second and third period of sampling.

5.4.5 Bioinformatics

A customized bioinformatics pipeline (nf-core/viralrecon) was used to analyze the raw sequence (Ewels et al., 2020). Briefly, FASTQ files were pre-processed using fastp (Chen et al., 2018), and aligned to the SARS-CoV-2 reference sequence (accession MN908947.3) using Bowtie2 (Langmead & Salzberg, 2012) to generate coverage information along the genome.

Mapped genomes were subsequently processed using iVar for consensus generation and mutations (SNVs) calling with a minimum frequency threshold (0.01), minimum quality threshold (30), and minimum read depth (30×) (Grubaugh et al., 2019). To capture the dynamics of virus evolution and spread, we utilized Freyja (Karthikeyan et al., 2022) bioinformatics to infers estimation relative abundance of lineage-defining mutations in both influent wastewater and primary sludge. We used .bam files on the Freyja workflow (v1.3.10) and packages found at https://github.com/andersen-lab/Freyja to determine lineage abundance through a regression approach that considers depth weighting and least absolute deviation.

5.4.6 Determination of species richness and Shannon entropy

The species richness and Shannon entropy were evaluated based on SNVs, following the protocols outlined in previous studies (Grubaugh et al., 2019). Specifically, species richness was defined as the total number of distinct SNVs sites identified, while Shannon entropy was calculated using the following formula:

$$H(x) = - \Sigma\, p(x) \log_2 (p(x)$$

Where, p(x) is the allele frequency at the position, x the distribution of each mutation and degree of genetic diversity in the sample.

5.4.7 Enumeration of Canadian cryptic, low-prevalence SNVs

To enumerate Canadian cryptic SNVs, we examined SNVs that were absent in Canadian clinical sequences during the sampling periods but identified in wastewater and primary sludge samples. These SNVs were cross-referenced with the comprehensive Canadian clinical sequence dataset available on the CovSPECTRUM-Canada portal (https://cov-spectrum.org/explore/Canada). SNVs not detected in Canadian clinical sequences throughout the pandemic (January 01, 2020, to April 03, 2024) were defined as Canadian cryptic SNVs. Additionally, we assessed the overall detection proportion of these SNVs in both Canadian and

global clinical sequences, as documented on the CovSPECTRUM-World platform (utilizing GISAID data) (https://cov-spectrum.org/explore/World). SNVs with a global detection proportion < 0.01 were categorized as "low prevalence"/ "rare", whereas those with a proportion > 0.01 were categorized as "prevalent".

5.4.8 Statistical analysis

The differences in N1 Ct values, percentage of genomic coverage, depth of genome coverage, and the number of detected lineages and SNVs between influent wastewater and primary sludge were compared using a paired t-test. The relationship between N1 Ct values and the percentage of genomic coverage as well as the depth of coverage was determined using Spearman correlation coefficient. The choice of each statistical test was determined by the normal distribution of the data. The statistical analyses and graphing were performed in GraphPad's Prism 10.1 (La Jolla, California, USA).

5.5 Results

5.5.1 Optimization of primary sludge processing for sequencing

To develop an optimized sludge processing method for sequencing, we compared the conventional sludge processing technique (D'Aoust, et al., 2021b) with our modified sludge processing approaches for SARS-CoV-2 RNA extraction. This comparative analysis aimed to overcome challenges posed by PCR inhibitors and RNA degradation typically present in primary sludge, thereby enhancing the recovery of high-quality viral RNA for reliable downstream sequencing applications. In our modified processing approaches, primary sludge was diluted with deionized water (50 mL) and Phosphate-Buffered Saline (PBS, 10X pH 7.4) buffer (50 mL) at varying volumetric ratios (1:50, 1:70, 1:100, 1:200, and 1:500), corresponding to 1.0 mL, 0.7 mL, 0.5 mL, 0.25 mL, and 0.1 mL of primary sludge mixed with 49 mL, 49.3 mL, 49.5 mL, 49.75 mL, and 49.9 mL of diluent respectively. Subsequently, 40

mL samples from the top layer were transferred into centrifuge tubes and concentrated using Nanotrap® Magnetic Virus Particles (Ceres Nanosciences), followed by total RNA extraction with the QIAmp Viral RNA Mini kit (Qiagen).

Following RNA extraction using both conventional and modified sludge processing methods, cDNA was synthesized using the SuperScript IV First-Strand Synthesis System (ThermoFisher). Tiled 400-bp amplicons were generated using the SARS-CoV-2 ARTIC primer set v4.1, and the amplicon size was verified using the Bioanalyzer 2100 system (Agilent Technologies). To validate amplification during cDNA synthesis and amplicon generation, a SARS-CoV-2 true-positive clinical sample (Ct comparable to primary sludge samples) was used as a positive control, while nuclease-free water served as a negative control.

The Bioanalyzer results revealed that desired amplicons (~400 bp) were successfully generated only in the clinical sample, with no amplification detected in either primary sludge samples processed using conventional methods or in the negative control (Figure A2). However, ARTIC amplification was observed at all dilution ratios in both deionized water (Figure A3) and PBS buffer (Figure A4), as well as in the positive control, but not in the negative control. Subsequent analysis of the average Ct values for the N1 and N2 loci of the SARS-CoV-2 genome indicated that the lowest Ct value was achieved with deionized water at a 1:70 dilution ratio (Table S1). Thereby, we selected a 1:70 dilution of primary sludge in deionized water for our final experiments, optimizing the processing method for improved sequencing outcomes.

5.5.2 Comparable SARS-CoV-2 consensus genome recovered from influent wastewater and primary sludge

Thirty pairs of influent wastewater and primary sludge samples were collected during three different periods under the three dominant lineages (BA.2, EG.5, EG.5.1) of Omicron VOC from ROPEC, Ottawa, Canada. The RT-qPCR of the samples showed that the average N1 Ct

values between influent wastewater (mean: 32.75 ± 1.41, range: 30.32 to 36.37) and primary sludge (mean: 32.64 ± 1.64, range: 28.28 to 36.31) were similar (paired t-test, $p > 0.05$) (Figure 1). Following RT-qPCR, all the samples were processed for tiled amplicon sequencing on a MiSeq platform. In the first sampling period, we sequenced two technical replicates of each of the paired influent wastewater and primary sludge samples with paired-end reads of 150 bp, while in the second and third periods, we sequenced the samples using paired-end reads of 300 bp. Thereby, a total of 80 sequence were generated (comprising 40 from the first period and 20 each from the second and third periods), and results showed that 37 (93%) of the influent wastewater and 35 (88%) of the primary sludge samples had genome coverage higher than 90% at 30× depth (Figure 5.1.A). The remaining three influent wastewater and five primary sludge samples with gnomic coverage lower than 90% observed during the second sampling period which could be due to multiple cycles of freezing and thawing of those samples. These repeated freeze-thaw cycles were required by technical malfunctions in the MiSeq sequencing instrument during that time. Previous studies have also demonstrated that freeze-thaw cycles of SARS-CoV-2 RNA may impact the quality and integrity of the viral genome, potentially affecting downstream analyses such as sequencing (Thapar et al., 2023; Williams et al., 2024). However, the average genome coverage in influent wastewater and primary sludge samples were 93.5% ± 5.81% and 91.6% ± 9.02% respectively at 30× depth throughout the sampling periods. The calculated depth of genome coverage varied between 1420× to 7394× in influent wastewater and from 1811× to 7750× in primary sludge samples (Figure 5.1.B). The findings indicated that our approach for processing primary sludge successfully recovers SARS-CoV-2 genomes with comparable depth to influent wastewater, overcoming the challenges encountered by other studies (Baaijens et al., 2022; Lin et al., 2021).

We then explored the potential relationship between Ct values and % genome coverage as well as depth of genome coverage in influent wastewater and primary sludge samples using

Spearman correlation analyses. No correlation ($p > 0.05$) was observed between Ct value and genome coverage as well as depth of coverage in influent wastewater and primary sludge samples. Overall, our approach demonstrated the ability to recover the near-complete genomes of SARS-CoV-2 from either primary sludge or influent wastewater samples with Ct values below 35. However, correlation analysis of the current study is aligned with the previous studies conducted using an Illumina sequencing platform via an amplification process (Fontenele et al., 2021; Izquierdo-Lara et al., 2021; Pérez-Cataluña et al., 2022), and capture-based methods (Crits-Christoph, et al., 2021). The lack of correlation between Ct values and genome coverage or depth of coverage may be attributed to the fact that RT-qPCR assays target specific small regions of the genome, whereas the entire genome is susceptible to degradation in wastewaters. Another possible reason could be the presence of regions within the genome that are non-intact or not amplified to the desired depth (30×) (Pérez-Cataluña et al., 2022).

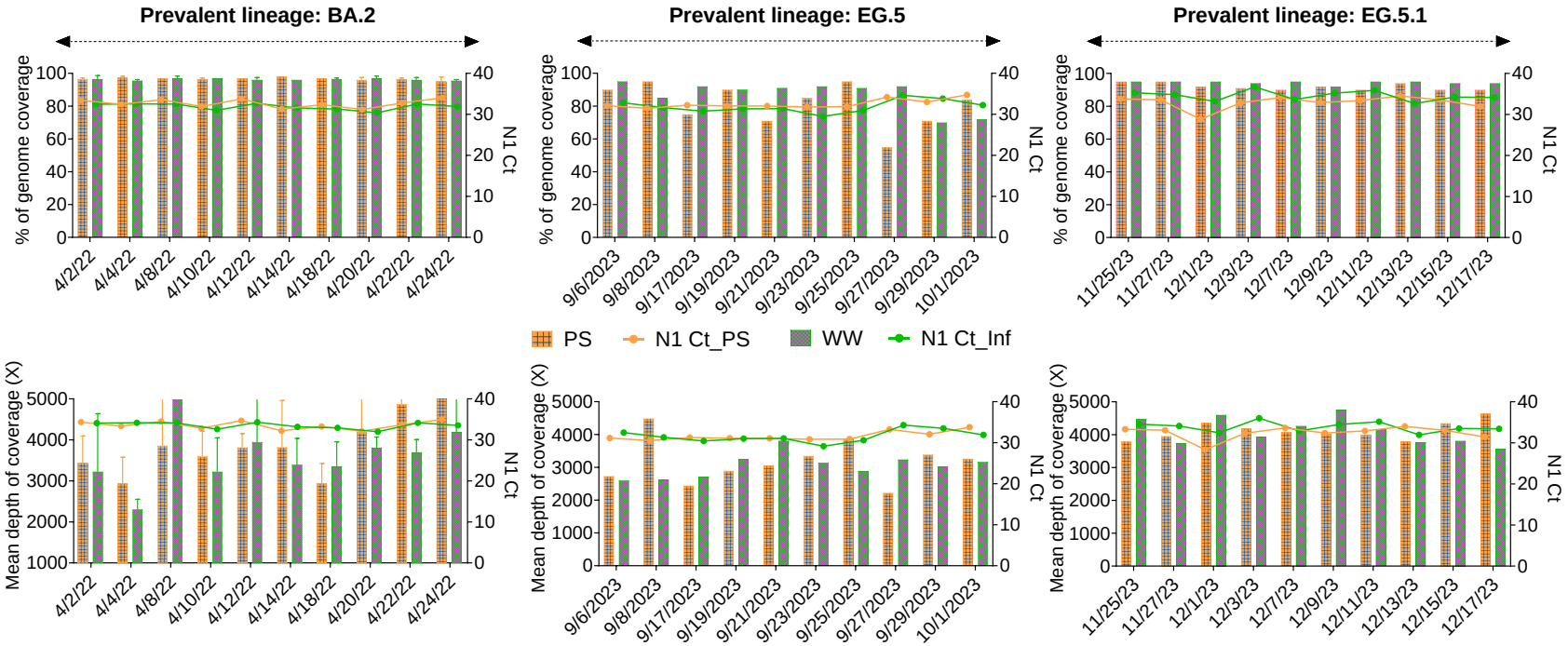

Figure 5-1: Longitudinal comparison of SARS-CoV-2 genome coverage in influent wastewater (Inf) and primary sludge (PS) across three different times periods under the three dominant lineages of Omicron VOC. A. The percentage of genome coverage (left axis) was plotted against the N1 Ct value (right axis), and B. The mean depth of coverage (×) (left axis) was plotted against the N1 Ct value. The percentage of genome coverage and mean depth of coverage on the leftmost graph were shown with error bars, as two technical replicates of each paired influent wastewater and primary sludge samples were sequenced during the first period of sampling.

5.5.3 SARS-CoV-2 lineages prevelance in influent wastewater and primary sludge

The SARS-CoV-2 RNA found in wastewaters is typically mixed from numerous infected individuals and provides sufficient information on circulating VOCs and lineages within a community (Gregory et al., 2022; Rouchka et al., 2021; Swift et al., 2021). To identify the prevalence of multiple VOCs and lineages in the mixed samples, we used Freyja bioinformatics tool to efficiently recover their relative abundance in influent wastewater and primary sludge samples. Freyja shows effective recovery of variant prevalence from mixed and/or degraded samples with genome coverage as low as 50% (Karthikeyan et al., 2022). In this study, we obtained near complete genome (≥ 90%) of SARS-CoV-2 from both influent wastewater and primary sludge (See Figure 5.1); and with the similar genome coverage, we compared the SARS-CoV-2 lineages prevalence between influent wastewater and primary sludge to determine whether these two common matrices of wastewater followed similar patterns.

In the first period of sampling, Omicron BA.2, a sub-lineage of Omicron VOC was predominant in clinical genomic surveillance in the city of Ottawa, Canada (Ottawa Public Health, 2024), and we compared the prevalence of BA.2 along with other alias of Omicron in influent wastewater and primary sludge samples. A total of 27 Omicron lineages were identified in influent wastewater and 29 in primary sludge samples, with BA.2.37 and BA.2.38 being the most prevalent lineages in both types of samples (Figure 5.2.A). Among the detected lineages, 17 were common to both influent wastewater and primary sludge, while 10 were unique to influent wastewater and 12 were exclusively detected in primary sludge. Later, several SARS-CoV-2 lineages of Omicron VOC became predominant in Ottawa, Canada. However, in July and August 2023, EG.5, another descendant of the Omicron variant, emerged, raising concerns about the potential for increased infections in the city (Ottawa Public Health, 2024). In this circumstance, we started monitoring

EG.5 in influent wastewater and primary sludge during the second period of sampling. In this period, 25 Omicron VOC lineages were recovered from influent wastewater and 33 in primary sludge, with 16 common to both sample types. In September 2023, EG.2, EG.5, EG.5.1, and XBB.1.5.28 were the prevalent lineages in both influent wastewater and primary sludge (Figure 5.2.B). In late 2023, the JN.1 subvariant of Omicron, the most mutated since parental Omicron VOC, emerged in Canada with the potential for higher infection rates and more severe symptoms (Quarleri et al., 2024). Due to its rapid spread, the WHO and the Public Health Agency of Canada classified JN.1 as a variant of interest (Dangerfield, 2023). Although the first clinical case of JN.1 in Canada was detected on October 9, 2023, its spread in Ottawa primarily began in late December 2023 (Ottawa Public Health, 2024). To track the early presence of JN.1 variant in Ottawa, Canada, we began screening in paired influent wastewater and primary sludge samples during third sampling period. JN.1 was not detected in either sample type which aligned with the clinical sequence in the city of Ottawa (Ottawa Public Health, 2024). Instead, we identified BA.2.1 and EG.5.1 as the predominant lineages in both sample types (Figure 5.2.C). Unlike the first and second periods, 26 Omicron lineages were found in influent wastewater and 21 in primary sludge, with only 10 lineages being shared between influent wastewater and primary sludge. Overall, throughout the sampling periods, the shared lineages between influent wastewater and primary sludge were detected with higher prevalence while lineages unique to either influent wastewater or primary sludge were detected with lower prevalence.

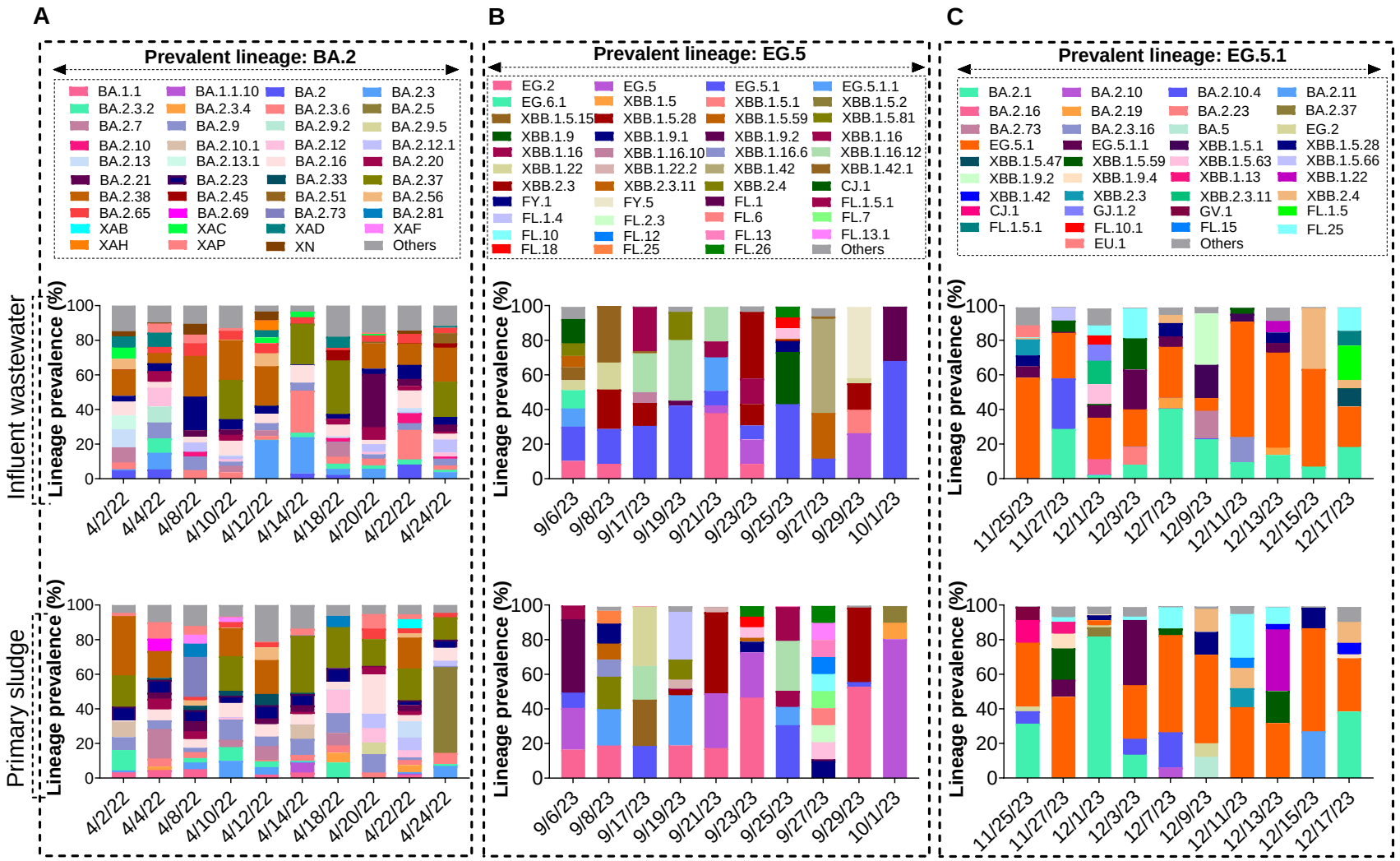

Figure 5-2: The prevalence (%) of different Omicron VOC lineages in influent wastewater (top) and primary sludge (bottom) was accessed using Freyja bioinformatics tool under the three dominant lineage BA.2 (leftmost), EG.5 (middle), and EG.5.1 (rightmost). The topmost section of the figures shows the legends for each lineage associated with each sampling period.

5.5.4 Comparative analysis of SARS-CoV-2 SNVs in influent wastewater and primary sludge

The SARS-CoV-2 SNVs were detected in influent wastewater and primary sludge samples using iVAR tool with a minimum quality threshold (30) and minimum read depth (30×) (Grubaugh et al., 2019). To ensure consistency and minimize inter-run variability of SNV calling, paired influent wastewater and primary sludge samples were sequenced within the same Illumina MiSeq run. Furthermore, sequencing was performed in distinct MiSeq runs corresponding to different predominant periods of SARS-CoV-2 Omicron lineages, each characterized by unique mutational signatures. As a result, most SNVs were inherently specific to their respective sequencing runs; however, approximately 40% of SNVs were consistently detected across all runs, reflecting a conserved set of mutations shared across multiple SARS-CoV-2 lineages. Overall, results showed that over the course of the monitoring, a total of 570 and 659 SNVs were detected in influent wastewater and primary sludge samples, respectively (Figure 5.3.A). During the first period of sampling, 100 SNVs were reported in both influent wastewater and primary sludge samples. In subsequent periods, influent wastewater and primary sludge samples showed higher counts of SNVs in both second (245 and 269) and third periods (225 and 287), although this trend was not statistically significant ($p > 0.05$). To further define SARS-CoV-2 genetic diversity between influent wastewater and primary sludge, we employed two commonly used metrics: Shannon entropy, which quantifies the uncertainty in randomly sampling an allele, and richness, which measures the number of single-nucleotide variants (Grubaugh et al., 2019). We observed statistically comparable ($p > 0.05$) genetic diversity of SARS-CoV-2 between influent wastewater and primary sludge samples throughout the monitoring periods (Figure 5.3.B). Later, we assessed each SNVs to determine whether they were common between influent wastewater and primary

sludge or exclusively present in either influent wastewater or primary sludge. The results showed that ~40% (229) SNVs were found exclusively in influent wastewater, whereas ~48% (318) unique SNVs were identified only in primary sludge. Similar to total SNV counts, comparable counts of unique SNVs were detected between the two sample types during the first period of sampling, whereas more ($p < 0.05$) unique SNVs were observed in primary sludge during the second and third periods of sampling. To access the reproducibility of the sequencing data, we monitored variation in mutation profiles between technical replicates and found consistent mutation calling between replicates across the paired samples (data not shown).

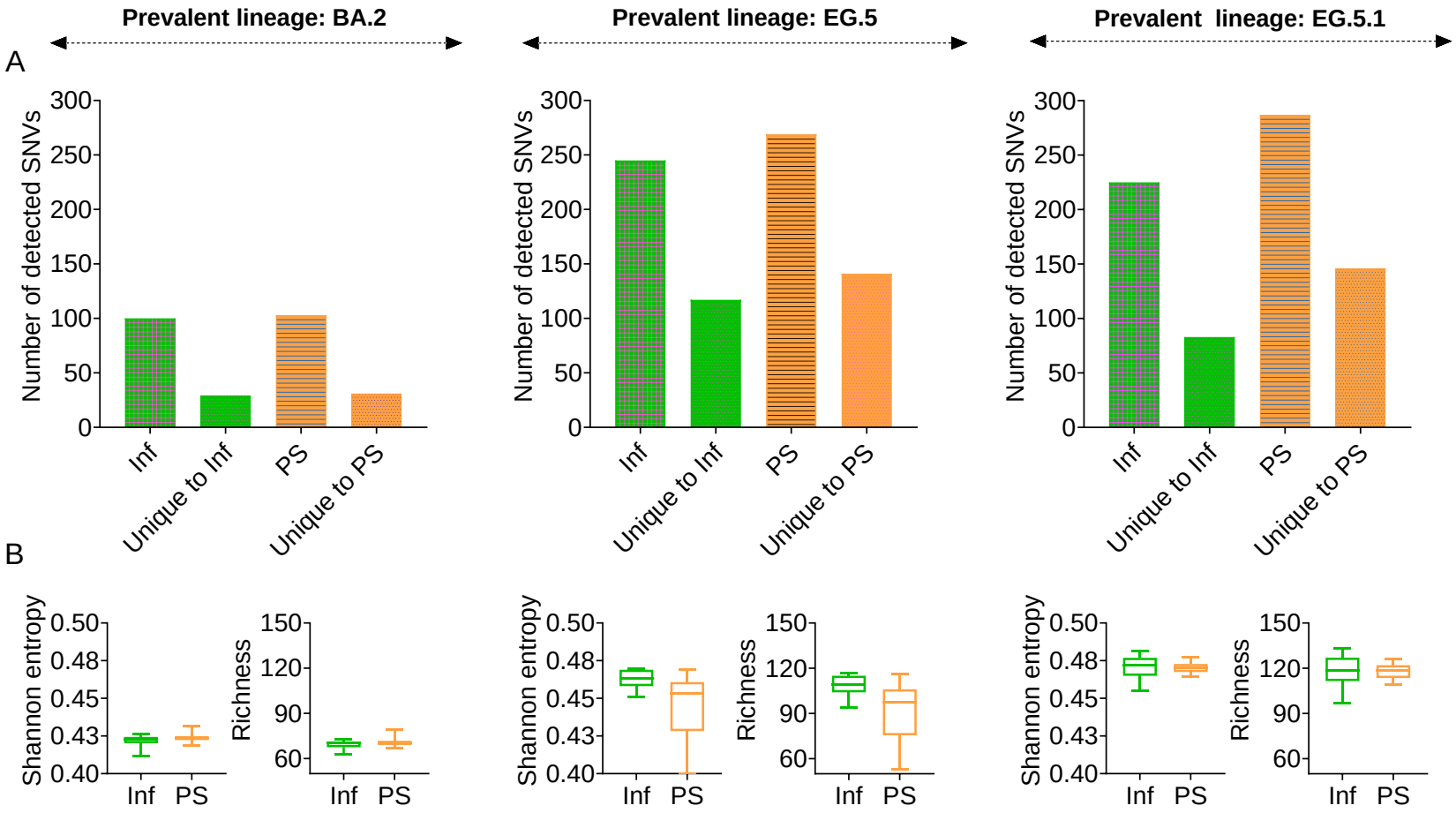

Figure 5-3: SARS-CoV-2 genetic diversity in influent wastewater (Inf) and primary sludge (PS) samples. A. Total and unique (detected exclusively in either sample type) SNVs counts in influent wastewater and primary sludge samples across the three sampling periods. B. Boxplots of Shannon entropy (leftmost of each variant type) and richness (rightmost of each variant type) during the periods dominated by lineage BA.2, EG.5, and EG.5.1. The edges of the box represent the first (bottom) and third (top) quartiles, the solid line indicates the median, and the whiskers show the maximum and minimum values.

5.5.5 Detection of more SARS-CoV-2 rare, cryptic, and low frequency SNVs in primary sludge compared to influent wastewater

We evaluated the detected SNVs in influent wastewater and primary sludge against the Canadian (Canada - CovSPECTRUM, 2024) and global (GISAID, 2024) clinical sequences during the sampling periods. Results indicated that common SNVs between influent wastewater and primary sludge were detected in both global and Canadian clinical sequences during the sampling periods. However, a total of 61 unique SNVs (six, 36 and 19 in each of the three periods, respectively) in influent wastewater and 133 unique SNVs (8, 60, and 65 in each of the three periods, respectively) in primary sludge were not detected in Canadian clinical sequences during the sampling periods (Figure 5.4). We then analyzed these unique SNVs, undetected in Canadian clinical sequences during the sampling periods, against the Global and Canadian clinical sequences throughout the pandemic (From Jan 01, 2020, to April 03, 2024). The analysis showed that 48% (29 out of 61) of the SNVs unique to influent wastewater and 46% (61 out of 133) in primary sludge were rarely (prevalence < 0.01) found in Canadian clinical sequences either before or after the sampling periods. The analysis also showed that 25% (15 out of 61) of the unique SNVs identified in wastewater and 30% (40 out of 133) in primary sludge corresponded to Canadian cryptic SNVs throughout the pandemic. The remaining unique SNVs in both influent wastewater (27%) and primary sludge (24%) were detected in Canadian clinical sequences with higher prevalence (> 0.01). All of the unique SNVs (i.e., rare, cryptic, and prevalent) found in influent wastewater and primary sludge samples were detected in global clinical sequences. However, the majority 61% (37 out of 61) of the unique SNVs in influent wastewater and about (72%; 96 out of 133) in primary sludge were rarely detected in global clinical sequences.

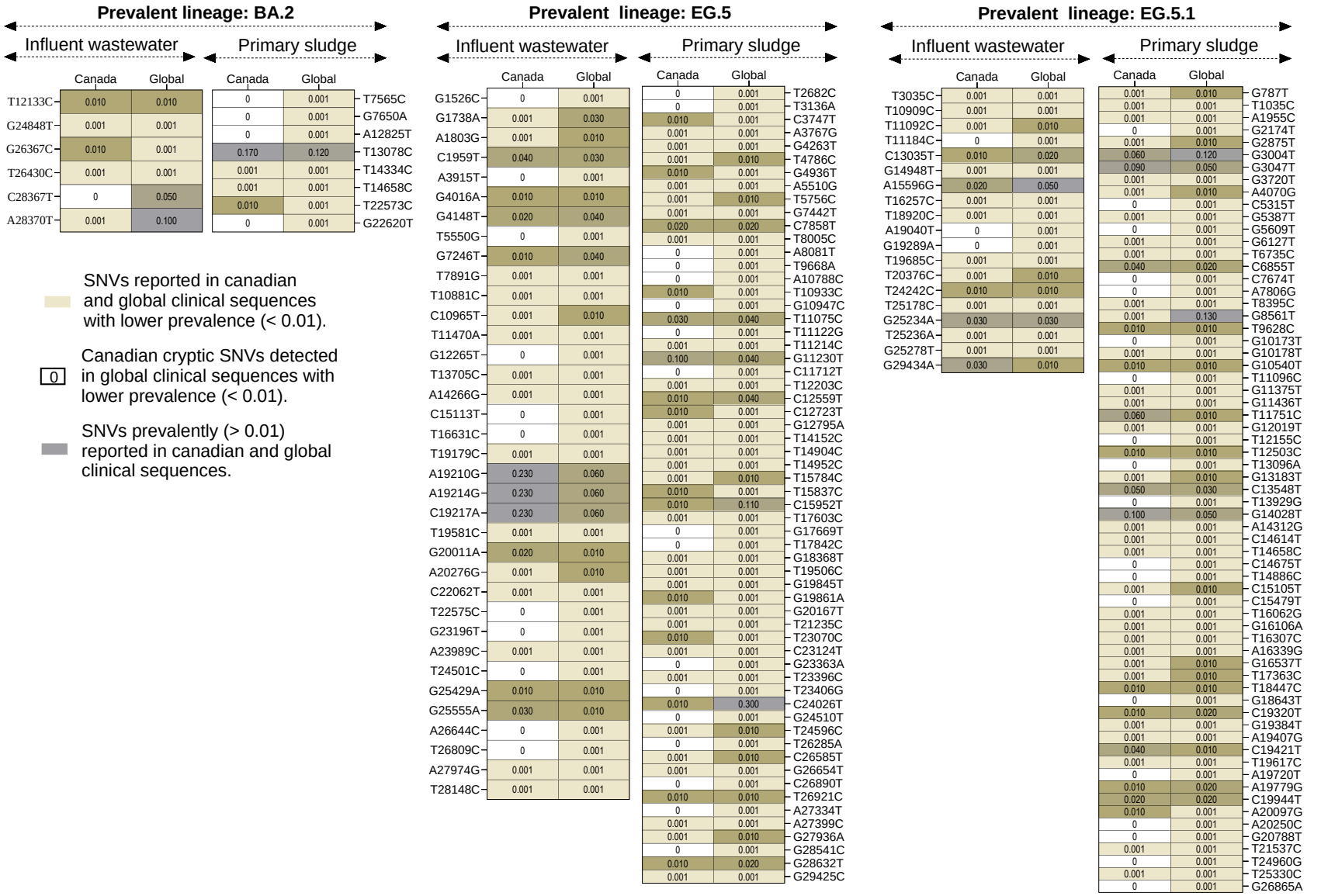
Prevalent lineage: BA.2
Influent wastewater
Primary sludge
Canada
Global
T12133C 0.010 0.010
G24848T 0.001 0.001
G26367C 0.010 0.001
T26430C 0.001 0.001
C28367T 0 0.050
A28370T 0.001 0.100
0 0.001 T7565C
0 0.001 G7650A
0 0.001 A12825T
0.170 0.120 T13078C
0.001 0.001 T14334C
0.001 0.001 T14658C
0.010 0.001 T22573C
0 0.001 G22620T
SNVs reported in canadian and global clinical sequences with lower prevalence (< 0.01).
0
Canadian cryptic SNVs detected in global clinical sequences with lower prevalence (< 0.01).
SNVs prevalently (> 0.01) reported in canadian and global clinical sequences.
Prevalent lineage: EG.5
Influent wastewater
Primary sludge
Canada
Global
G1526C 0 0.001
G1738A 0.001 0.030
A1803G 0.001 0.010
C1959T 0.040 0.030
A3915T 0 0.001
G4016A 0.010 0.010
G4148T 0.020 0.040
T5550G 0 0.001
G7246T 0.010 0.040
T7891G 0.001 0.001
T10881C 0.001 0.001
C10965T 0.001 0.010
T11470A 0.001 0.001
G12265T 0 0.001
T13705C 0.001 0.001
A14266G 0.001 0.001
C15113T 0 0.001
T16631C 0 0.001
T19179C 0.001 0.001
A19210G 0.230 0.060
A19214G 0.230 0.060
C19217A 0.230 0.060
T19581C 0.001 0.001
G20011A 0.020 0.010
A20276G 0.001 0.010
C22062T 0.001 0.001
T22575C 0 0.001
G23196T 0 0.001
A23989C 0.001 0.001
T24501C 0 0.001
G25429A 0.010 0.010
G25555A 0.030 0.010
A26644C 0 0.001
T26809C 0 0.001
A27974G 0.001 0.001
T28148C 0.001 0.001
0 0.001 T2682C
0 0.001 T3136A
0.010 0.001 C3747T
0.001 0.001 A3767G
0.001 0.001 G4263T
0.001 0.010 T4786C
0.010 0.001 G4936T
0.001 0.001 A5510G
0.001 0.010 T5756C
0.001 0.001 G7442T
0.020 0.020 C7858T
0.001 0.001 T8005C
0 0.001 A8081T
0 0.001 T9668A
0 0.001 A10788C
0.010 0.001 T10933C
0 0.001 G10947C
0.030 0.040 T11075C
0 0.001 T11122G
0.001 0.001 T11214C
0.100 0.040 G11230T
0 0.001 C11712T
0.001 0.001 T12203C
0.010 0.040 C12559T
0.010 0.001 C12723T
0.001 0.001 G12795A
0.001 0.001 T14152C
0.001 0.001 T14904C
0.001 0.001 T14952C
0.001 0.010 T15784C
0.010 0.001 T15837C
0.010 0.110 C15952T
0.001 0.001 T17603C
0 0.001 G17669T
0 0.001 T17842C
0.001 0.001 G18368T
0.001 0.001 T19506C
0.001 0.001 G19845T
0.010 0.001 G19861A
0.001 0.001 G20167T
0.001 0.001 T21235C
0.010 0.001 T23070C
0.001 0.001 C23124T
0 0.001 G23363A
0.001 0.001 T23396C
0 0.001 T23406G
0.010 0.300 C24026T
0 0.001 G24510T
0.001 0.010 T24596C
0 0.001 T26285A
0.001 0.010 C26585T
0.001 0.001 G26654T
0 0.001 C26890T
0.010 0.010 T26921C
0 0.001 A27334T
0.001 0.001 A27399C
0.001 0.010 G27936A
0 0.001 G28541C
0.010 0.020 G28632T
0.001 0.001 G29425C
Prevalent lineage: EG.5.1
Influent wastewater
Primary sludge
Canada
Global
T3035C 0.001 0.001
T10909C 0.001 0.001
T11092C 0.001 0.010
T11184C 0 0.001
C13035T 0.010 0.020
G14948T 0.001 0.001
A15596G 0.020 0.050
T16257C 0.001 0.001
T18920C 0.001 0.001
A19040T 0 0.001
G19289A 0 0.001
T19685C 0.001 0.001
T20376C 0.001 0.010
T24242C 0.010 0.010
T25178C 0.001 0.001
G25234A 0.030 0.030
T25236A 0.001 0.001
G25278T 0.001 0.001
G29434A 0.030 0.010
0.001 0.010 G787T
0.001 0.001 T1035C
0.001 0.001 A1955C
0 0.001 G2174T
0.001 0.010 G2875T
0.060 0.120 G3004T
0.090 0.050 G3047T
0.001 0.001 G3720T
0.001 0.010 A4070G
0 0.001 C5315T
0.001 0.001 G5387T
0 0.001 G5609T
0.001 0.001 G6127T
0.001 0.001 T6735C
0.040 0.020 C6855T
0 0.001 C7674T
0 0.001 A7806G
0.001 0.001 T8395C
0.001 0.130 G8561T
0.010 0.010 T9628C
0 0.001 G10173T
0.001 0.001 G10178T
0.010 0.010 G10540T
0 0.001 T11096C
0.001 0.001 G11375T
0.001 0.001 G11436T
0.060 0.010 T11751C
0.001 0.001 G12019T
0 0.001 T12155C
0.010 0.010 T12503C
0 0.001 T13096A
0.001 0.010 G13183T
0.050 0.030 C13548T
0 0.001 T13929G
0.100 0.050 G14028T
0.001 0.001 A14312G
0.001 0.001 C14614T
0.001 0.001 T14658C
0 0.001 C14675T
0 0.001 T14886C
0.001 0.010 C15105T
0 0.001 C15479T
0.001 0.001 T16062G
0.001 0.001 G16106A
0.001 0.001 T16307C
0.001 0.001 A16339G
0.001 0.010 G16537T
0.001 0.010 T17363C
0.010 0.010 T18447C
0 0.001 G18643T
0.010 0.020 C19320T
0.001 0.001 G19384T
0.001 0.001 A19407G
0.040 0.010 C19421T
0.001 0.001 T19617C
0 0.001 A19720T
0.010 0.020 A19779G
0.020 0.020 C19944T
0.010 0.001 A20097G
0 0.001 A20250C
0 0.001 G20788T
0.001 0.001 T21537C
0 0.001 T24960G
0.001 0.001 T25330C
0 0.001 G26865A

Figure 5-4: Evaluation of non-detected SNVs in influent wastewater and primary sludge during the sampling periods against the Canadian and global clinical sequences to determine their prevalence. The figure denotes SNVs with lower prevalence using light, brown-colored cells, while Canadian cryptic SNVs are shown in white with 0 colored cells and commonly reported SNVs in Canadian and global clinical sequences are represented by ash-colored cells.

5.6 Discussion

In this study, we have demonstrated that our approach of processing primary sludge for sequencing allows the recovery of near complete SARS-CoV-2 genomes, comparable to influent wastewater. Primary sludge is usually processed using direct sludge processing methods (i.e., centrifugation), resulting in higher viral RNA concentrations compared to influent wastewater (Balboa et al., 2021; D' Aoust et al., 2021a; Kim et al., 2022). Despite the higher concentrations of SARS-CoV-2 RNA extracted from primary sludge, the RNA may contain significant amounts of non-target genetic material leading to background noise that complicates the identification and assembly of SARS-CoV-2 genomes. To resolve the difficulties in primary sludge sequencing, we diluted primary sludge in deionized water at a 1:70 ratio, which might remove the impurities as well as potential inhibitors from the primary sludge. We then concentrated on pairing diluted primary sludge and influent wastewater samples using Nanotrap® Magnetic Virus Particles and extracted viral RNA using Qiagen Viral RNA Mini Kit. More exclusively, the beads and dilution may negatively select for PCR inhibitors (polymeric substances, possibly low polarity) or the beads may positively select for a narrow range of size/charge, thus excluding non-targeted particles. The high surface area and specific binding properties of Nanotrap® Magnetic Virus Particles likely facilitated the recovery of SARS-CoV-2 genomes from both influent wastewater and primary sludge. In the downstream analysis, we processed all paired samples in parallel, including cDNA synthesis, amplicon generation, library preparation, and sequenced them in the same MiSeq run. Finally,

the percentage of SARS-CoV-2 genome coverage and the depth of coverage (×) were calculated using Mosdepth with default parameters and found the recovery of similar SARS-CoV-2 genome from influent wastewater and primary sludge samples. Considering the overall similarity of the sequencing methodologies and bioinformatics pipelines employed for influent wastewater and primary sludge, our approach demonstrates effective mitigation of background noise and impurities from the extracted RNA of primary sludge, leading to the successful recovery of comparable SARS-CoV-2 genome from influent wastewater and primary sludge.

With comparable genome coverage of SARS-CoV-2 between influent wastewater and primary sludge, we utilized Freyja (Karthikeyan et al., 2022) to analyze the prevalence of SARS-CoV-2 lineages and iVAR (Grubaugh et al., 2019) to determine the presence of SNVs in the genome. Genomic analysis indicated that a similar ($p > 0.05$) number of SARS-CoV-2 lineages and SNVs were identified between influent wastewater and primary sludge. The prevalent lineages and SNVs were common between influent wastewater and primary sludge; however, rare and Canadian cryptic SNVs were exclusively found in either influent wastewater or primary sludge samples. The detection of these SNVs demonstrates the effectiveness of WWGS using influent wastewater and now primary sludge, consistent with findings from earlier research (Karthikeyan et al., 2022; Li et al., 2022; Smyth et al., 2022). Our study also showed there was a higher likelihood of detecting rare, Canadian cryptic SNVs in primary sludge compared to influent wastewater. The detection of rare and Canadian cryptic SNVs in both influent wastewater and primary sludge samples may be attributed to the evolving nature of the pandemic. As the pandemic wanes and vaccination rates increase, the severity of the disease diminishes, often presenting as mild flu-like symptoms, leading to fewer individuals seeking clinical diagnosis (Shafer et al., 2024; Smyth et al., 2022). Thereby, the detection of these SNVs in wastewater and primary sludge was confirmed the presence of a lower number of infected individuals in the community who do not undergo through clinical diagnosis. Besides, as

Ottawa serves as the capital city of Canada and hosts numerous historical landmarks, attracting visitors year-round, these individuals may contribute to the introduction of circulating, low-prevalence SNPs into the wastewater networks. However, the majority of these rare, Canadian cryptic SNVs were detected in only one sample throughout the monitoring periods. This raises questions about whether they represent true lineages and SNVs or are artifacts of sequencing errors. The sequencing errors can be due to a number of factors such as the quality of the RNA sample, sample processing, variations in cluster density as well as data processing and analysis errors (Endrullat et al., 2016; Sheng et al., 2017). We conducted simultaneous processing of paired influent wastewater and primary sludge for sequencing on the MiSeq platform and analyzed the data using identical pipelines, thereby eliminating potential errors specific to particular sample type (i.e., primary sludge). Moreover, these rare SNVs in primary sludge were also rarely detected (prevalence < 0.01) in available Canadian (Canada - CovSPECTRUM, 2024) and global (GISAID, 2024) clinical sequences. Similarly, the Canadian cryptic SNVs found in primary sludge were also rarely observed in global clinical sequences. This demonstrates that the identified rare, Canadian-cryptic SNVs in primary sludge were genuine and underscoring the power of primary sludge sequencing.

Viral particles in influent wastewater are generally present at low concentrations due to dilution, but they accumulate in primary clarified sludge within the settling tanks of wastewater treatment plants through gravitational sedimentation. The higher likelihood of rare, Canadian cryptic SNVs detection in primary sludge compared to influent wastewater demonstrated that the sequencing of SARS-CoV-2 genome in diluted wastewater may not be able to uncover all of the circulating variants in the population, which can be effectively identified through sludge sequencing. Therefore, SARS-CoV-2 genomic surveillance should prioritize primary sludge over influent wastewater due to the inherent advantages of sludge as a sample matrix. However, more impurities and enzyme inhibitors may also concentrate in primary sludge, making its

sequencing more difficult and challenging. Our approach for processing primary sludge in SARS-CoV-2 genome sequencing overcomes these barriers, allowing for the detection of rare SNVs in primary sludge. Furthermore, sludge provides a more aggregated and representative snapshot of population level viral loads, reducing variability introduced by flow rate changes and rainfall dilution in influent wastewater. These advantages make primary sludge an optimal matrix for robust SARS-CoV-2 genomic surveillance, improving sensitivity, consistency, and alignment with clinical epidemiological data to better inform public health decision-making. Overall, our approach demonstrates the potential applicability of primary sludge sequencing for recovering whole genomes to detect different variants of emerging pathogens. These insights provide critical information for robust public health strategies during the COVID-19 pandemic and in future public health emergencies.

5.7 Conclusion

In this study, we developed a robust approach for processing primary sludge samples for the whole genome sequencing of SARS-CoV-2. Using our approach, we successfully recovered near-complete genomes from approximately 90% of both influent wastewater and primary sludge samples. This demonstrates that our approach can provide valuable genomic data for WWGS, despite the challenges posed by RT and/or PCR inhibitors, impurities, lower target concentrations, and viral RNA degradation in wastewater. In addition, we identified various lineages and SNVs in SARS-CoV-2 sequences from each influent wastewater and primary sludge sample. Some of these, found at higher prevalence, were common to both influent wastewater and primary sludge, while others, with lower prevalence, were unique to either influent wastewater or primary sludge. The lower prevalence SNVs provided insights into Canadian cryptic mutations in both influent wastewater and primary sludge samples. However, our study reveals that primary sludge has the potential to detect rare and Canadian cryptic SNVs compared to influent wastewater. As a collective, the approaches we have outlined in this study

can be utilized within public health settings to inform infectious disease mitigation measures for numerous pathogens and their associated variants, particularly in scenarios where obtaining clinically derived sequences is rare or challenging.

5.8 References

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., & Mueller, J. F. (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*. https://doi.org/10.1016/j.scitotenv.2020.144216

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

Andersen, P., Barksdale, S., Barclay, R. A., Smith, N., Fernandes, J., Besse, K., Goldfarb, D., Barbero, R., Dunlap, R., Jones-Roe, T., Kelly, R., Miao, S., Ruhunusiri, C., Munns, A., Mosavi, S., Sanson, L., Munns, D., Sahoo, S., Swahn, O., … Lepene, B. (2023). Magnetic hydrogel particles improve nanopore sequencing of SARS-CoV-2 and other respiratory viruses. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-29206-7

Association of Public Health Laboratories (2022). *SARS-CoV-2 Wastewater Surveillance Testing Guide for Public Health Laboratories ® APHL SARS-CoV-2 Wastewater Testing Guide | 2*

Baaijens, J. A., Zulli, A., Ott, I. M., Nika, I., van der Lugt, M. J., Petrone, M. E., Alpert, T., Fauver, J. R., Kalinich, C. C., Vogels, C. B. F., Breban, M. I., Duvallet, C., McElroy, K. A., Ghaeli, N., Imakaev, M., Mckenzie-Bennett, M. F., Robison, K., Plocik, A., Schilling, R., … Baym, M. (2022). Lineage abundance estimation for SARS-CoV-2 in wastewater using transcriptome quantification techniques. *Genome Biology*, *23*(1). https://doi.org/10.1186/s13059-022-02805-9

Balboa, S., Mauricio-Iglesias, M., Rodriguez, S., Martínez-Lamas, L., Vasallo, F. J., Regueiro, B., & Lema, J. M. (2021). The fate of SARS-COV-2 in WWTPS points out the sludge line as a suitable spot for detection of COVID-19. *Science of the Total Environment*, *772*. https://doi.org/10.1016/j.scitotenv.2021.145268

Bar-Or, I., Weil, M., Indenbaum, V., Bucris, E., Bar-Ilan, D., Elul, M., Levi, N., Aguvaev, I., Cohen, Z., Shirazi, R., Erster, O., Sela-Brown, A., Sofer, D., Mor, O., Mendelson, E., & Zuckerman, N. S. (2021). Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel. *Science of the Total Environment*, *789*, 1–14. https://doi.org/10.1016/j.scitotenv.2021.148002

Canada - covSPECTRUM (2024). Detect and analyze variants of SARS-CoV-2. URL: https://open.cov-spectrum.org (Accessed 03 April 2024)

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Crits-Christoph, A., Kantor, R. S., Olm, M. R., Whitney, O. N., Al-Shayeb, B., Lou, Y. C., Flamholz, A., Kennedy, L. C., Greenwald, H., Hinkle, A., Hetzel, J., Spitzer, S., Koble, J., Tan, A., Hyde, F., Schroth, G., Kuersten, S., Banfield, J. F., & Nelson, K. L. (2021).

Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *MBio*, *12*(1), 1–9. https://doi.org/10.1128/mBio.02703-20

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

D'Aoust, P. M., Towhid, S. T., Mercier, É., Hegazy, N., Tian, X., Bhatnagar, K., Zhang, Z., Naughton, C. C., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021b). COVID-19 wastewater surveillance in rural communities: Comparison of lagoon and pumping station samples. *Science of the Total Environment*, *801*. https://doi.org/10.1016/j.scitotenv.2021.149618

Dangerfield, K. (2023). JN.1 is now a COVID-19 variant of interest, WHO says - National _ Globalnews.ca. published on 12.19.2023. URL: (https://globalnews.ca/news/10179450/jn-1-covid-variant-of-interest-who/). (Accessed 15 July 2024)

Daughton, C. G. (2020). Wastewater surveillance for population-wide Covid-19: The present and future. In *Science of the Total Environment* (Vol. 736). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2020.139631

Davies, N. G., Jarvis, C. I., van Zandvoort, K., Clifford, S., Sun, F. Y., Funk, S., Medley, G., Jafari, Y., Meakin, S. R., Lowe, R., Quaife, M., Waterlow, N. R., Eggo, R. M., Lei, J., Koltai, M., Krauer, F., Tully, D. C., Munday, J. D., Showering, A., … Keogh, R. H.

(2021). Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*, *593*(7858), 270–274. https://doi.org/10.1038/s41586-021-03426-1

Endrullat, C., Glökler, J., Franke, P., & Frohme, M. (2016). Standardization and quality management in next-generation sequencing. In *Applied and Translational Genomics* (Vol. 10, pp. 2–9). Elsevier B.V. https://doi.org/10.1016/j.atg.2016.06.001

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, *38*(3), 276–278. https://doi.org/10.1038/s41587-020-0439-x

Fontenele, R. S., Kraberger, S., Hadfield, J., Driver, E. M., Bowes, D., Holland, L. R. A., Faleye, T. O. C., Adhikari, S., Kumar, R., Inchausti, R., Holmes, W. K., Deitrick, S., Brown, P., Duty, D., Smith, T., Bhatnagar, A., Yeager, R. A., Holm, R. H., von Reitzenstein, N. H., … Varsani, A. (2021). High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *Water Research*, *205*. https://doi.org/10.1016/j.watres.2021.117710

GISAID (2024). GISAID - hCov19 Variants. URL: (https://gisaid.org/) (Accessed 03 April 2024)

Gregory, D. A., Trujillo, M., Rushford, C., Flury, A., Kannoly, S., San, K. M., Lyfoung, D. T., Wiseman, R. W., Bromert, K., Zhou, M. Y., Kesler, E., Bivens, N. J., Hoskins, J., Lin, C. H., O'Connor, D. H., Wieberg, C., Wenzel, J., Kantor, R. S., Dennehy, J. J., & Johnson, M. C. (2022). Genetic diversity and evolutionary convergence of cryptic SARS- CoV-2 lineages detected via wastewater sequencing. *PLoS Pathogens*, *18*(10). https://doi.org/10.1371/journal.ppat.1010636

Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., Tan, A. L., Paul, L. M., Brackney, D. E., Grewal, S., Gurfield, N., Van Rompay, K. K. A., Isern, S., Michael, S. F., Coffey, L. L., Loman, N. J., & Andersen, K. G. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biology*, *20*(1), 1–19. https://doi.org/10.1186/s13059-018-1618-7

Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Oude Munnink, B. B., Schapendonk, C. M. E., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F. M., Lycett, S., Medema, G., Koopmans, M. P. G., & De Graaf, M. (2021). Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the netherlands and belgium. *Emerging Infectious Diseases*, *27*(5), 1405–1415. https://doi.org/10.3201/eid2705.204410

Jahn, K., Dreifuss, D., Topolsky, I., Kull, A., Ganesanandamoorthy, P., Fernandez-Cassi, X., Bänziger, C., Devaux, A. J., Stachler, E., Caduff, L., Cariti, F., Corzón, A. T., Fuhrmann, L., Chen, C., Jablonski, K. P., Nadeau, S., Feldkamp, M., Beisel, C., Aquino, C., … Beerenwinkel, N. (2022). Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. *Nature Microbiology*, *7*(8), 1151–1160. https://doi.org/10.1038/s41564-022-01185-x

Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Aigner, S., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., … Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, *609*(7925), 101–108. https://doi.org/10.1038/s41586-022-05049-6

Kim, S., Kennedy, L. C., Wolfe, M. K., Criddle, C. S., Duong, D. H., Topol, A., White, B. J., Kantor, R. S., Nelson, K. L., Steele, J. A., Langlois, K., Griffith, J. F., Zimmer-Faust, A. G., McLellan, S. L., Schussman, M. K., Ammerman, M., Wigginton, K. R., Bakker, K. M., & Boehm, A. B. (2022). SARS-CoV-2 RNA is enriched by orders of magnitude in primary settled solids relative to liquid wastewater at publicly owned treatment works. *Environmental Science: Water Research and Technology*, *8*(4), 757–770. https://doi.org/10.1039/d1ew00826a

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Li, L., Uppal, T., Hartley, P. D., Gorzalski, A., Pandori, M., Picker, M. A., Verma, S. C., & Pagilla, K. (2022). Detecting SARS-CoV-2 variants in wastewater and their correlation with circulating variants in the communities. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-20219-2

Lieberman-Cribbin, W., Tuminello, S., Flores, R. M., & Taioli, E. (2020). Disparities in COVID-19 Testing and Positivity in New York City. *American Journal of Preventive Medicine*, *59*(3), 326–332. https://doi.org/10.1016/j.amepre.2020.06.005

Lin, X., Glier, M., Kuchinski, K., Ross-Van Mierlo, T., McVea, D., Tyson, J. R., Prystajecky, N., & Ziels, R. M. (2021). Assessing Multiplex Tiling PCR Sequencing Approaches for Detecting Genomic Variants of SARS-CoV-2 in Municipal Wastewater. *MSystems*, *6*(5). https://doi.org/10.1128/msystems.01068-21

Ling-Hu, T., Rios-Guzman, E., Lorenzo-Redondo, R., Ozer, E. A., & Hultquist, J. F. (2022). Challenges and Opportunities for Global Genomic Surveillance Strategies in the COVID-19 Era. In *Viruses* (Vol. 14, Issue 11). NLM (Medline). https://doi.org/10.3390/v14112532

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 and Technology Letters*, *7*(7), 511–516. https://doi.org/10.1021/acs.estlett.0c00357

Ottawa Public Health (2024). Statistics on Coronavirus 19 (COVID-19) in Ottawa. URL: (https://www.ottawapublichealth.ca/en/reports-research-and-statistics/statistics-on-coronavirus-19-covid-19-in-ottawa.aspx). Accessed (10 February 2024)

Patel, M., Chaubey, A. K., Pittman, C. U., Mlsna, T., & Mohan, D. (2021). Coronavirus (SARS-CoV-2) in the environment: Occurrence, persistence, analysis in aquatic systems and possible management. In *Science of the Total Environment* (Vol. 765). Elsevier B.V. https://doi.org/10.1016/j.scitotenv.2020.142698

Peccia, J., Zulli, A., Brackney, D. E., Grubaugh, N. D., Kaplan, E. H., Casanovas-Massana, A., Ko, A. I., Malik, A. A., Wang, D., Wang, M., Warren, J. L., Weinberger, D. M., Arnold, W., & Omer, S. B. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature Biotechnology*, *38*(10), 1164–1167. https://doi.org/10.1038/s41587-020-0684-z

Pérez-Cataluña, A., Chiner-Oms, Á., Cuevas-Ferrando, E., Díaz-Reolid, A., Falcó, I., Randazzo, W., Girón-Guzmán, I., Allende, A., Bracho, M. A., Comas, I., & Sánchez, G. (2022). Spatial and temporal distribution of SARS-CoV-2 diversity circulating in wastewater. *Water Research*, *211*. https://doi.org/10.1016/j.watres.2021.118007

Quarleri, J., Delpino, M. V., & Galvan, V. (2024). Anticipating the future of the COVID-19 pandemic: insights into the emergence of SARS-CoV-2 variant JN.1 and its projected impact on older adults. In *GeroScience* (Vol. 46, Issue 3, pp. 2879–2883). Springer

Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s11357-024-01066-7

Reitsma, M. B., Claypool, A. L., Vargo, J., Shete, P. B., McCorvie, R., Wheeler, W. H., Rocha, D. A., Myers, J. F., Murray, E. L., Bregman, B., Dominguez, D. M., Nguyen, A. D., Porse, C., Fritz, C. L., Jain, S., Watt, J. P., Salomon, J. A., & Goldhaber-Fiebert, J. D. (2021). Racial/ethnic disparities in covid-19 exposure risk, testing, and cases at the subcounty level in California. *Health Affairs*, *40*(6), 870–878. https://doi.org/10.1377/hlthaff.2021.00098

Robishaw, J. D., Alter, S. M., Solano, J. J., Shih, R. D., DeMets, D. L., Maki, D. G., & Hennekens, C. H. (2021). Genomic surveillance to combat COVID-19: challenges and opportunities. In *The Lancet Microbe* (Vol. 2, Issue 9, pp. e481–e484). Elsevier Ltd. https://doi.org/10.1016/S2666-5247(21)00121-X

Rouchka, E. C., Chariker, J. H., Saurabh, K., Waigel, S., Zacharias, W., Zhang, M., Talley, D., Santisteban, I., Puccio, M., Moyer, S., Holm, R. H., Yeager, R. A., Sokoloski, K. J., Fuqua, J., Bhatnagar, A., & Smith, T. (2021). The rapid assessment of aggregated wastewater samples for genomic surveillance of sars-cov-2 on a city-wide scale. *Pathogens*, *10*(10). https://doi.org/10.3390/pathogens10101271

Shafer, M. M., Bobholz, M. J., Vuyk, W. C., Gregory, D. A., Roguet, A., Haddock Soto, L. A., Rushford, C., Janssen, K. H., Emmen, I. E., Ries, H. J., Pilch, H. E., Mullen, P. A., Fahney, R. B., Wei, W., Lambert, M., Wenzel, J., Halfmann, P., Kawaoka, Y., Wilson, N. A., … Johnson, M. C. (2024). Tracing the origin of SARS-CoV-2 omicron-like spike sequences detected in an urban sewershed: a targeted, longitudinal surveillance study of a cryptic wastewater lineage. *The Lancet Microbe*, *5*(4), e335–e344. https://doi.org/10.1016/S2666-5247(23)00372-5

Sheng, Q., Vickers, K., Zhao, S., Wang, J., Samuels, D. C., Koues, O., Shyr, Y., & Guo, Y. (2017). Multi-perspective quality control of Illumina RNA sequencing data analysis. *Briefings in Functional Genomics*, *16*(4), 194–204. https://doi.org/10.1093/bfgp/elw035

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28246-3

Swift, C. L., Isanovic, M., Correa Velez, K. E., & Norman, R. S. (2021). Community-level SARS-CoV-2 sequence diversity revealed by wastewater sampling. *Science of the Total Environment, 801*. https://doi.org/10.1016/j.scitotenv.2021.149691

Tao, K., Tzou, P. L., Nouhin, J., Gupta, R. K., de Oliveira, T., Kosakovsky Pond, S. L., Fera, D., & Shafer, R. W. (2021). The biological and clinical significance of emerging SARS-CoV-2 variants. In *Nature Reviews Genetics* (Vol. 22, Issue 12, pp. 757–773). Nature Research. https://doi.org/10.1038/s41576-021-00408-x

Thapar, I., Langan, L. M., Davis, H., Norman, R. S., Bojes, H. K., & Brooks, B. W. (2023). Influence of storage conditions and multiple freeze-thaw cycles on N1 SARS-CoV-2, PMMoV, and BCoV signal. *Science of the Total Environment*, *896*. https://doi.org/10.1016/j.scitotenv.2023.165098

Vashisht, V., Vashisht, A., Mondal, A. K., Farmaha, J., Alptekin, A., Singh, H., Ahluwalia, P., Srinivas, A., & Kolhe, R. (2023). Genomics for Emerging Pathogen Identification and Monitoring: Prospects and Obstacles. In *BioMedInformatics* (Vol. 3, Issue 4, pp. 1145–

1177). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/biomedinformatics3040069

Vo, V., Tillett, R. L., Chang, C. L., Gerrity, D., Betancourt, W. Q., & Oh, E. C. (2022). SARS-CoV-2 variant detection at a university dormitory using wastewater genomic tools. *Science of the Total Environment, 805*. https://doi.org/10.1016/j.scitotenv.2021.149930

Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P. D., Graham, B. S., Mascola, J. R., Chang, J. Y., Yin, M. T., Sobieszczyk, M., Kyratsous, C. A., Shapiro, L., Sheng, Z., Huang, Y., & Ho, D. D. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature, 593*(7857), 130–135. https://doi.org/10.1038/s41586-021-03398-2

Williams, R. C., Perry, W. B., Lambert-Slosarska, K., Futcher, B., Pellett, C., Richardson-O'Neill, I., Paterson, S., Grimsley, J. M. S., Wade, M. J., Weightman, A. J., Farkas, K., & Jones, D. L. (2024). Examining the stability of viral RNA and DNA in wastewater: Effects of storage time, temperature, and freeze-thaw cycles. *Water Research, 259*. https://doi.org/10.1016/j.watres.2024.121879

6 Chapter 6: Diagnostic Performance of Allele-Specific RT-qPCR and Genomic Sequencing in Wastewater-Based Surveillance of SARS-CoV-2

6.1 Context

Chapter 6 represents published research article entitled *Diagnostic performance of allele-specific RT-qPCR and genomic sequencing in wastewater-based surveillance of SARS-CoV-2* by Kabir, M. P., Mercier, E., Eid, W., Plaza-Diaz, J., D'Aoust, P. M., Landgraff, C., Goodridge, L., Lawal, O. U., Wan, S., Hegazy, N., Nguyen, T., Wong, C., Thakali, O., Pisharody, L., Stephenson, S., Graber, T. E., and Delatolla, R. (Journal of Eco-Environment and Health, 2025). This study compares single-allele frequency estimation using AS-RT-qPCR, to single-allele or haplotype frequency estimations derived from amplicon-based sequencing to estimate variant prevalence in municipal wastewater samples. The study also determines the diagnostic performance of single allele (RT-qPCR), single allele (sequencing), and haplotype of each variant to ensure the accuracy of each method in wastewater genomic surveillance of SARS-CoV-2.

6.2 Abstract

Clinical genomic surveillance is regarded as the gold standard for monitoring SARS-CoV-2 variants globally. However, as the pandemic wanes, reduced testing poses a risk to effectively tracking the trajectory of these variants within populations. Wastewater genomic surveillance (WWGS) that estimates variant frequency based on its defining set of alleles derived from clinical genomic surveillance has been successfully implemented. This method has its challenges, and allele-specific RT-qPCR or RT-dPCR may instead be used as a complementary method for estimating variant prevalence. Demonstrating equivalent performance of these methods is a prerequisite for their continued application in current, and future pandemics. Here, we compared single-allele frequency using AS-RT-qPCR, to single-allele or haplotype

frequency estimations derived from amplicon-based sequencing to estimate variant prevalence in wastewater during emergent and prevalent periods of Delta, Omicron, and two sub-lineages of Omicron. We found that all three methods of frequency estimation were concordant and contained sufficient information to describe the trajectory of variant prevalence. We further confirmed the accuracy of these methods by quantifying the diagnostic performance through Youden's index. The Youden's index of AS-RT-qPCR was reduced during the low prevalence period of a particular variant while the same allele in sequencing was negatively influenced due to insufficient read depth. Youden's index of haplotype-based calls was negatively influenced when alleles were common between variants. Coupling AS-RT-qPCR with sequencing can overcome the shortcomings of either platform and provide a comprehensive picture to the stakeholders for public health responses.

6.3 Introduction

The SARS-CoV-2 RNA genome is susceptible to mutations, resulting in the emergence of viral variants throughout the COVID-19 pandemic (Boehm et al., 2021; Pachetti et al., 2020). Many of the more prevalent variants have been linked to increased transmissibility, disease severity, reinfection, and reduced vaccine effectiveness (Davies et al., 2021; Tao et al., 2021; Wang et al., 2021). These have been designated as variants of concern (VOCs), variants of interest (VOIs), and variants under monitoring (VUMs) (WHO, 2022). Identifying and tracking them through clinical genomic surveillance has become critical to early and effective global public health responses (Xiao et al., 2020; Otto et al., 2021).

Clinical genomic surveillance of SARS-CoV-2 catalogues all mutations and variants with an individual case or patient granularity, but the method is expensive, time-consuming, and requires a sufficient number of clinical tests to ascertain the epidemiological metrics such as variant prevalence in a population (Furuse, 2021; Rahimi et al., 2021). Allele-specific reverse transcription quantitative PCR (AS-RT-qPCR, or its digital PCR counterpart) is another

analytical tool that has been used to monitor SARS-CoV-2 mutations and variants in clinical specimens during the pandemic (Vega-Magaña et al., 2021; Garson et al., 2022). The clinical implementation of the method reports binary outcomes, either the presence or absence of a targeted mutation (allele). Although practically limited to monitoring one or two genetic loci for mutation, careful selection of sufficiently variant-specific mutations allows the method to be easily applied in a conventional SARS-CoV-2 diagnostic PCR testing facility and can rapidly report variant prevalence (Ratcliff et al., 2022).

Despite the wide adoption of these tools during the pandemic, sequencing, or AS-RT-qPCR-based SARS-CoV-2 clinical surveillance may not accurately estimate the frequency of a viral variant circulating in asymptomatic, pre-symptomatic, and post-symptomatic cohorts (i.e., those unlikely to be tested) and its implementation at large scale (i.e., local, national and regional scale) is not sustainable (Suhail et al., 2021). The SARS-CoV-2, a respiratory virus, and its RNA shed in various bodily fluids, feces, and urine, which eventually enter sewer networks from infected individuals, including symptomatic, asymptomatic, and presymptomatic patients (Fei Xiao et al., 2020; Jones et al., 2020). This finding leads to the emergence of wastewater-based surveillance as a non-invasive method for tracking the presence and prevalence of the virus across populations. Later, WWGS and AS-RT-qPCR-based surveillance have been rapidly developed and adopted by public health departments and agencies to monitor SARS-CoV-2 variants. The resulting data provides an aggregated, anonymous, and non-invasive measurement of the prevalence of circulating SARS-CoV-2 variants in the community (Crits-Christoph et al., 2021; Fontenele et al., 2021; Graber et al., 2021; Lee et al., 2021).

The power of WWGS of SARS-CoV-2 genetic variation is in its ability to cast a much wider net than its clinical counterpart, allowing the discovery of new and emerging or cryptic (i.e., not found in clinical genome repositories) mutations that might have clinical significance

(Gregory et al., 2022; Smyth et al., 2022). However, it is fundamentally different from its clinical counterpart which derives epidemiological metrics such as variant prevalence from discrete counts (i.e., presence/absence enumerated from individual cases). Wastewater variant frequency represents continuous variables, convolved from multiple individuals with different fractional contributions and different states of genomic degradation and fragmentation. This makes it challenging to deconvolve and estimate the abundances of contributing SARS-CoV-2 haplotypes (i.e., the set of mutations that make up the viral genome) in the wastewater context. A reductionist solution to this problem is to track a single allele diagnostic for a given variant. Methods for both approaches have been successfully implemented (i.e., haplotype-based variant frequency estimation from sequencing, and single-allele variant frequency estimation from AS-RT-qPCR) in the wastewater context, providing valuable community incidence/prevalence data to public health (Graber et al., 2021; Heijnen et al., 2021; Peterson et al., 2022). However, monitoring of a single allele is prone to signal dropout due to differential fragmentation, degradation, or expression of the locus in different wastewater contexts, thus increasing the risk of false negative detection (Barbé et al., 2022; Rios et al., 2021). In addition, the re-emergence of a single allele with multiple variants could lead to inaccurate frequency estimations.

Two studies have shown that both AS-RT-qPCR and AS-RT-dPCR are more sensitive compared to sequencing-based methods for variant detection and quantification in wastewater (Ahmed, et al., 2022; Lou et al., 2022). Previous studies (Izquierdo-Lara et al., 2021; Wurtzer et al., 2022) also revealed that single-allele frequency estimations obtained by AS-RT-dPCR were strongly correlated with single-allele or haplotype frequency estimates derived from sequencing. However, the diagnostic performance (i.e., accuracy) of these wastewater-based methods with respect to the actual, real-world data of clinical surveillance remains unclear. The diagnostic performance of an assay is derived by enumerating true positive (TP), false positive

(FP), true negative (TN), and false negative (FN) detections using a gold-standard test as a benchmark. Youden's index, which reports sensitivity and specificity in a single metric ranging from 0 to 1 and reflects test accuracy, is used widely to measure the performance of clinical diagnostic tests (Mishra et al., 2016; Rota et al., 2015; Youden, 1950; Zhang et al., 2020). A value of 1 indicates a perfect test with no false positive or negative detections, while 0 represents a test that detects equal proportions of false positives and false negatives, making it ineffective (Habibzadeh et al., 2016; Unal, 2017). Assigning this index to test the presence or absence of SARS-CoV-2 variants in wastewater samples could help assay development teams in choosing the best allele or set of alleles with which to estimate variant frequencies.

In this study, four alleles associated with B.1.617.2 (Delta; N: D63G), B.1.1.529 (Omicron; N: P13L), and Omicron sub-lineages BA.1 (S: $H69^{+}/V70^{+}$) and BA.2 (S: $H69^{-}/V70^{-}$), were individually evaluated using AS-RT-qPCR and amplicon-based sequencing methods across periods of their emergence, dominance, and extinction. Single, variant diagnostic allele frequencies measured by AS-RT-qPCR were compared to both single-allele and haplotype frequencies measured by sequencing. Both AS-RT-qPCR and sequencing based on single-allele and haplotype frequencies were found to be comparable in describing the trajectory of variant prevalence. To quantitatively evaluate the accuracy of these methods in detecting variant frequency in wastewater, Youden's index was applied for each method using Ontario, Canada's clinical genomic surveillance data as the benchmark. Youden's index confirmed the accuracy of these methods for estimating the frequency of the SARS-CoV-2 variant in wastewaters.

6.4 Materials and Methods

6.4.1 Wastewaters sampling

Thirty-six parallel (post-grit influent and primary sludge) 24-h composite samples were collected from the City of Ottawa's Robert O. Pickard Environmental Center (ROPEC), Ontario, Canada using autosamplers (Hoskin Scientific, Burlington, Canada) between November 5, 2021, and April 12, 2022. On each sampling day, 500 mL influent and primary sludge samples were collected over 24 h. Wastewater samples were transferred to the laboratory in ice cooler packs and stored at 4 °C until analysis. The primary sludge samples were analyzed by AS-RT-qPCR within 24 h of sampling in Ottawa and influent samples were shipped to National Microbiology Laboratory (NML), Winnipeg, Canada using ice cooler packs for SARS-CoV-2 whole genome sequencing.

6.4.2 RNA enrichment and extraction

For AS-RT-qPCR, total RNA from primary sludge was extracted using the RNeasy Power Microbiome Kit (Qiagen, Germantown, USA) and a QIACube connect automated extraction platform as previously described (D'Aoust, et al., 2021a). Briefly, primary sludge samples were mixed thoroughly, 40 mL transferred to centrifuge tubes, and centrifuged for 45 min at 10,000 x g at 4 ºC. The supernatant was discarded, and the pellet was further centrifuged at 10,000 x g for 5 min at 4 ºC, and 250 mg ± 5 mg of the pellet was transferred to an RNase-free microfuge tube. The pellet was lysed with 650 µL PM1 buffer with β-mercaptoethanol (100 :1 ratio) followed by the addition of Trizol LS reagent (ThermoFisher, Ottawa, Canada) before vortexing and centrifugation. Then, ~1 mL of lysate was applied to columns placed in the QIACube connect to complete the process. On-column DNase treatment was performed as per the manufacturer's instructions. 100 µL of total RNA was eluted using nuclease-free water.

For amplicon sequencing, Nanotrap® Magnetic Virus Particles (Ceres Nanosciences, Manassas, USA) were used to enrich SARS-CoV-2 RNA from 50 mL of influent wastewater. Samples were left at room temperature for 10 min in 50 mL conical tubes to settle the larger particles, and 40 mL of supernatant was transferred to centrifuge tubes. Then, 600 µL of magnetic particles were added, and samples were rotated end-over-end at 100 rpm for 20 min at 20 ºC, then centrifuged for 10 min at 4 ºC at 8000 x g. The centrifuge tubes were placed on a magnetic rack (DynaMag-50, Invitrogen, Waltham, USA), and the supernatant was discarded by pipetting without disturbing the pellet. The pellet was resuspended in 140 µL of Phosphate-buffered saline (pH 7.4) and 560 µL of Viral Lysis Buffer from QIAmp Viral RNA Mini Kit (Qiagen). The resulting suspension was transferred to a 2.0 mL microcentrifuge tube and placed on a magnetic rack (Invitrogen™ DynaMag™-2 magnet) for 10 min to remove the magnetic particles. The lysate was carefully collected by pipette, and total nucleic acid was extracted using QIAmp Viral RNA Mini Kit (Qiagen) and eluted in 68 µL of nuclease-free water. We measured A260/A280 (protein and phenol) and A260/A230 (phenol and guanidine) ratios using a Nanodrop for couple of samples extracted via centrifugation and Nanotrap Magnetic Virus Particles. We found that the A260/A280 ratio ranged from 1.8 to 2.0, while the A260/A230 ratio was higher than 2.0, indicating high purity of the extracted RNA.

6.4.3 Allele specific RT-qPCR

The presence of SARS-CoV-2 RNA in the samples was confirmed by singleplex RT-qPCR targeting N1 and N2 loci using TaqMan® Fast Virus 1-Step Master Mix (ThermoFisher) on a CFX Connect qPCR thermocycler (Bio-Rad, Hercules, Canada). Samples were analyzed in triplicates, with RNAse-free water serving as the non-template control and quantified using a five-point gradient of the EDX SARS-CoV-2 standard curve (Exact Diagnostics, Texas, USA). The EDX standard is commonly used to quantify SARS-CoV-2 targets, including E, N, ORF1ab, RdRp, and S genes, with concentrations reaching up to 200 cp/µL. In contrast, AS-

RT-qPCR was performed using previously published methods (Table A5) to quantify frequencies of the following alleles diagnostic for Delta and Omicron variants across the study period: 1) N: D63G (targeting Delta B.1.617.2 and its sub-lineages);` 2) N: P13L (targeting Omicron B.1.1.529 and its sub-lineages); 3) S: $H69^{+}/V70^{+}$ (targeting Omicron B.1.1.529.1 alias BA.1) and 4) S: $H69^{-}/V70^{-}$(targeting Omicron B.1.1.529.2 alias BA.2). The AS-RT-qPCR reactions were considered positive when threshold cycle (Ct) < 40. The use of internal controls and their recovery efficiency was conducted by D'Aoust et al. (2021b) in our laboratory. Following the methodology outlined in that publication, the data presented in this study have not been corrected for recovery efficiency. Besides, the method's limit of detection as well as the qPCR data processing were performed as per the MIQE recommendations (Bustin et al., 2009). Additionally, sample runs were discarded if they did not meet the following requirements: i) standard curves are linear ($R^2 \geq 0.95$), ii) copies/well are found in the linear range of the standard curve, iii) standard deviation among triplicates <0.5, and iv) no positives in the no-template controls.

6.4.4 SARS-CoV-2 tiled amplicon sequencing

First, DNA from wastewater extracts RNA was removed using ezDNase Enzyme kit (ThermoFisher) and cDNA was synthesized using Superscript IV First-Strand Synthesis System kit (ThermoFisher) according to the previously described protocol (Landgraff et al., 2021). Amplicons (~400bp) were generated from the cDNA using Artic V3 (samples collected between November 5 to 30, 2021) or V4.1 (samples collected between December 01, 2021, and April 12, 2022) protocols, and the size of the amplicons was confirmed through the 2100 Bioanalyzer system (Agilent Technologies, Mississauga, Canada). The amplicons were purified with 0.8X AMPure XP beads (Beckman Coulter) and quantified with a Qubit 4 fluorometer (ThermoFisher) using the dsDNA High Sensitivity Kit (ThermoFisher). The libraries were prepared using Nextera XT DNA Library Preparation Kit (Illumina, San Diego,

USA), quantified, normalized, pooled, and diluted to 10 pM. The pooled 10 pM library was sequenced on a MiSeq platform (Illumina) using PE flow cell with 300 bp paired end-run mode.

6.4.5 Mapping and variant calling

A customized "nf-core/viralrecon" bioinformatics pipeline was used to analyze the raw sequences (Ewels et al., 2020). Briefly, the quality of raw reads was assessed using FastQC (Simon Andrews, 2020), filtered using fastp (Chen et al., 2018) to remove adaptor sequences, ambiguous bases, low-quality reads (Phred score < 30), and small fragments (< 50nt). The filtered reads were then aligned to the SARS-CoV-2 reference sequence (NCBI Nucleotide accession MN908947.3) using Bowtie2 with default parameters (Langmead & Salzberg, 2012). The aligned reads were sorted using SAMtools (Li et al., 2009), and consensus sequences were generated using iVar (Grubaugh et al., 2019). The consensus sequence was constructed using map reads with coverage of > 5× and a Phred score of > 30. Finally, variant calling was performed using iVar with a minimum frequency threshold (0.01), minimum Phred score (30), and minimum read depth (30).

6.4.6 Estimation of allele frequency using AS-RT-qPCR

The frequency of the variant allele in wastewaters was calculated as a fraction of the sum of variant and reference allele RNA copies determined using allele-specific forward primers (D63G or P13L assays) or probe ($H69^+/V70^+$ assay) against a reference standard curve as previously described (Graber et al., 2021). More detailed information is found in Table A6.

6.4.7 Haplotype construction of SARS-CoV-2 variants from wastewater

To construct the haplotype of each variant, mutations specific to each lineage were first selected from public repositories including the GISAID database (GISAID, 2022), CovSPECTRUM (CovSPECTRUM-Canada, 2022), and Pango Lineages (Cov-Lineages, 2022). Mutations shared between lineages were removed to reduce overestimation or false positive detections.

A mutation was defined as specific if it was found predominately in ≥ 90% of the clinical genomes assigned to the lineage and absent or found in ≤ 10% of clinical genomes assigned to all other lineages. The average frequency across all of the specific mutations that defined a lineage (haplotype) was defined as the haplotype frequency.

6.4.8 Diagnostic performance assessment – Youden's index

The diagnostic performance of AS-RT-qPCR and amplicon-based sequencing methods for single allele and haplotype were evaluated based on Youden's index as follows (Habibzadeh et al., 2016; Unal, 2017).

Youden's index= (Sensitivity + Specificity) -1

$$\text{Sensitivity} = \frac{True\ positive\ (TP)}{(True\ positive\ (TP) + False\ negative\ (FN))}$$

$$\text{Specificity} = \frac{True\ negative\ (TN)}{(True\ negative\ (TN) + False\ positive\ (FP))}$$

Where, TP was defined as the number of events with both a positive detection of the variant in wastewater samples and ≥ 5% of clinical genomes, whereas TN was the number of events with both negative detection in wastewater and < 5% of clinical genomes sequenced in Ontario in the period assigned to the variant in question (PHO, 2022). On the contrary, FP was the number of events with variant detection in wastewater but < 5% of clinical genomes assigned, and FN was the number of non-detection events in wastewater with ≥ 5% of clinical genomes assigned. Table S2 summarizes these events.

6.4.9 Statistical analysis

The Wilcoxon matched-pairs signed-rank test was used to test for statistically significant differences (set to $p < 0.05$) between AS-RT-qPCR and the amplicon-based sequencing method. Spearman rank correlation analysis was used to compute the relation of the single

allele and haplotype frequency of each variant. The statistical analysis and graphs were generated with GraphPad Prism 10.2.1 (La Jolla, California, USA).

6.5 Results and Discussion

6.5.1 Longitudinal comparison of AS-RT-qPCR and amplicon-based sequencing in wastewaters

Primary sludge and post grit, influent samples collected on the same day between November 5, 2021, and April 12, 2022, were subjected to analysis using AS-RT-qPCR and amplicon-based sequencing. This period of time saw distinct phases of the pandemic characterized by the replacement of the endemic Delta sub-lineages with those of Omicron in the sampled population. To compare AS-RT-qPCR with amplicon sequencing for estimating SARS-CoV-2 variant frequency, alleles were chosen based on the contemporaneously unique haplotype-defining mutations, derived from clinical genomic surveillance (PANGO) as described in Materials and Methods. For sequencing, we chose unique alleles associated with B.1.617.2 (parental lineage and its sub-lineages), B.1.1.529 (parental lineage and its sub-lineages), B.1.1.529.1 (alias BA.1), and B.1.1.529.2 (alias BA.2). For AS-RT-qPCR, diagnostic alleles for Delta and its sub-lineages (N: D63G), Omicron and its sub-lineages (N: P13L), BA.1 (S: $H69^{+}/V70^{+}$), and BA.2 (S: $H69^{-}/V70^{-}$) were chosen based on contemporaneous prevalence in clinical genomic sequences. For each sampling date, variant frequency was estimated through 1) a single, diagnostic allele by AS-RT-qPCR, 2) the same allele by sequencing, and 3) the haplotype by sequencing. We computed Spearman's correlation between the two sequencing-based estimates for each variant to determine if single-allele estimates reflected the computed average across the alleles making up the haplotype. This revealed statistically significant ($p < 0.05$) positive correlations ($r = 0.860$ to 0.903) for each variant (Figure A5). Later we evaluated the performance of these three methods (i.e., AS-RT-qPCR, single-allele sequencing, and haplotype sequencing) in estimating the prevalence of variants in wastewater associated with

1) the period of waning Delta (B.1.617.2) prevalence and the concomitant emergence of Omicron (B.1.1.529) in December 2021; and 2) the replacement of the dominant BA.1 Omicron variant with the BA.2 sub-lineage in April 2022.

In the first period, N: D63G frequency assessed by AS-RT-qPCR and B.1.617.2 (Delta) haplotype frequency derived by sequencing were on average 50% and 75% in November to early December 2021 samples, respectively, whereas N: D63G frequency measured by sequencing was 100% (Figure 6.1.A).The inconsistency of N: D63G allele frequency between AS-RT-qPCR and amplicon sequencing exhibited more sensitive detection of N: D63G in the sequencing-based method. In contrast, the variability of single allele (N: D63G) and B.1.617.2 haplotype was due to the numerous missing data points of haplotype-defining alleles because of insufficient sequencing depth (< 30×) (Figure 6.2). In the middle of December 2021, Delta was rapidly replaced by Omicron. The frequency estimates derived by AS-RT-qPCR or sequencing (haplotype or single-allele) for Delta dropped precipitously in the third week of December 2021. This observation was concomitant with an equally rapid increase in Omicron frequency estimates using all three methods (Figure 6.1.B). This pattern of variant prevalence was consistent with that derived from clinical genomic surveillance in Canada during this period (CovSPECTRUM-Canada, 2022). The last days of Delta were characterized by very low overall SARS-CoV-2 signal in wastewater and this was reflected in the lower frequencies of the Delta variant using any of the methods.

The Omicron-specific N: P13L allele was first detected in primary sludge using AS-RT-qPCR on December 5, 2021 (2% of total signal), while the same allele was not detectable on that date by sequencing from influent and was instead first detected in December 9, 2021, sample at a frequency of 20%, close to the estimate by AS-RT-qPCR on that day (Figure 6.1.B). The AS-RT-qPCR estimates were derived from RNA extracted from primary sludge, which we and others (D'Aoust, et al., 2021; Yanaç et al., 2022) have previously shown to have higher SARS-

CoV-2 RNA concentrations and this may be the reason for the early detection of the N: P13L allele. Importantly, the trajectory of measured N: P13L or B.1.1.529 haplotype frequencies was consistent with that derived from clinical genomic surveillance in Canada (CovSPECTRUM-Canada, 2022). The B.1.1.529 haplotype frequency took longer to reach its peak than N: P13L due to the lower frequency of certain alleles (S: N440K, S: S477N, S: E484A, S: Q493R, S: Q498R, M: Q19E, M: A63T) (Figure 6.2).

We next assessed any differences in the three methods during the late winter-early spring period of 2022 when an Omicron BA.2 epi-wave displaced the endemic BA.1 (GISAID, 2022; Cov-Lineages, 2022). Here, allele frequencies of S: $H69^{+}/V70^{+}$ were monitored by AS-RT-qPCR or by sequencing. This allele distinguishes BA.1 from B.1.1.529 or BA.2, the latter two harboring a deletion at this locus. At the beginning of the monitoring period, S: $H69^{+}/V70^{+}$ frequencies were similar ($p > 0.05$) between AS-RT-qPCR and sequencing-based methods, but two sample dates near the end of March 2022 showed a larger deviation between AS-RT-qPCR and two sequencing-based methods (Figure 6.1.C). A similar deviation was also observed between AS-RT-qPCR (S: $H69^{-}/V70^{-}$) and BA.2 haplotype frequency estimates (Figure 6.1.D). The S: $H69^{-}/V70^{-}$ allele was realistically absent in BA.2, and during the initial stage of BA.2 epi-wave, there was no locus being monitored by AS-RT-qPCR. As epi-wave BA.2 replaced BA.1, it was presumed that the prevalence of S: $H69^{-}/V70^{-}$ could be estimated through (100-S: $H69^{+}/V70^{+}$) since the deleted "ATACATG" nucleotides in BA.1 returned with BA.2. Later, we found the similar ($p > 0.05$) trajectory of S: $H69^{-}/V70^{-}$ and BA.2 haplotype frequency in wastewater in all three methods (Figure 6.1.D). Overall, these findings revealed comparable trends of SARS-CoV-2 variant frequency across the methods employed, although certain differences were observed on particular days. These findings are consistent with earlier studies which demonstrated no discernible differences between AS-RT-qPCR or (AS-RT-dPCR) and

sequencing-based methods for estimating SARS-CoV-2 mutations or variants prevalence in wastewater (Izquierdo-Lara et al., 2023; Lou et al., 2022).

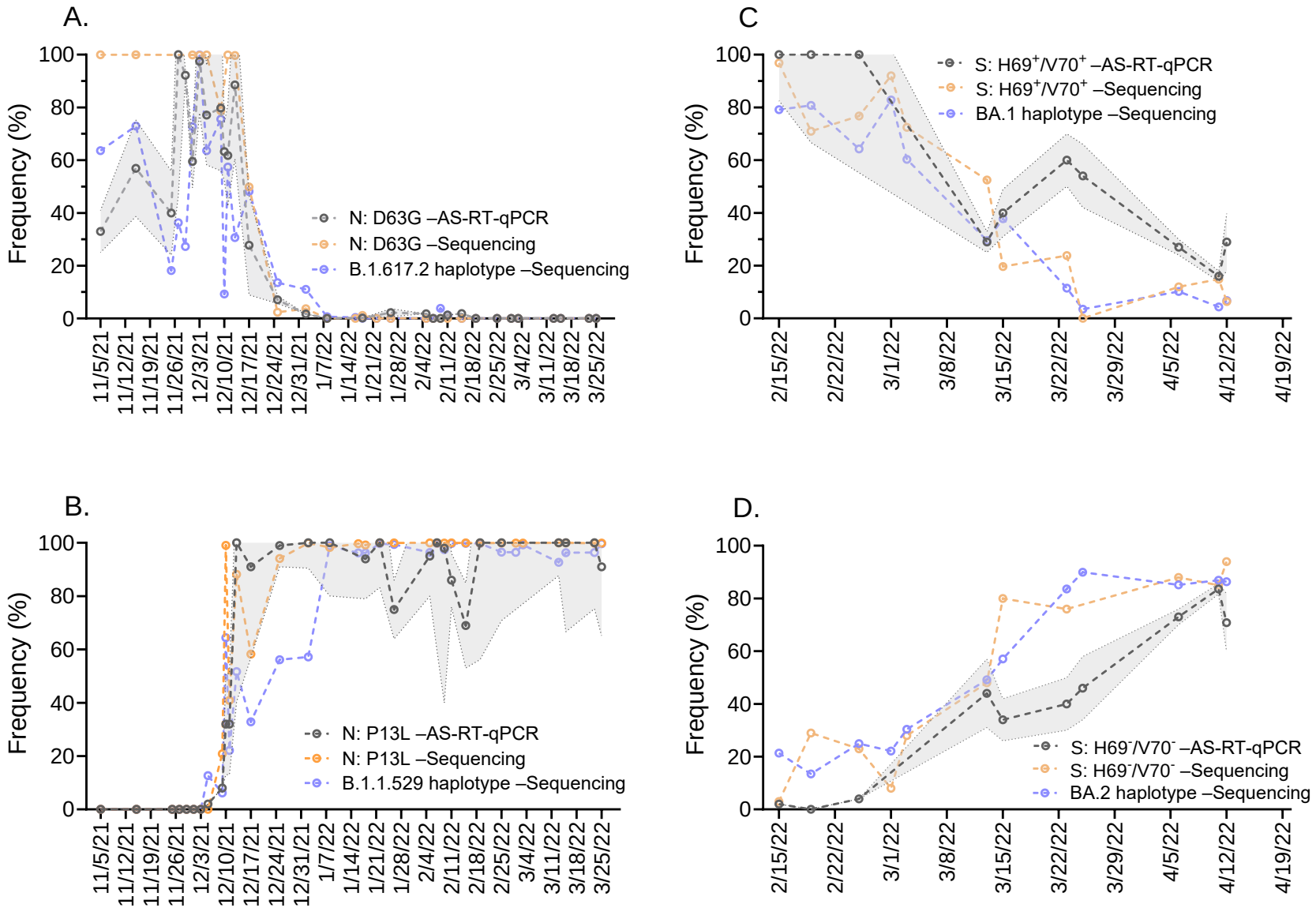
A.
Frequency (%)
N: D63G –AS-RT-qPCR
N: D63G –Sequencing
B.1.617.2 haplotype –Sequencing
B.
Frequency (%)
N: P13L –AS-RT-qPCR
N: P13L –Sequencing
B.1.1.529 haplotype –Sequencing
C
Frequency (%)
S: H69+/V70+ –AS-RT-qPCR
S: H69+/V70+ –Sequencing
BA.1 haplotype –Sequencing
D.
Frequency (%)
S: H69-/V70- –AS-RT-qPCR
S: H69-/V70- –Sequencing
BA.2 haplotype –Sequencing

Figure 6-1: Comparison of AS-RT-qPCR and amplicon-based sequencing methods for estimating SARS-CoV-2 variants frequency in wastewaters**;** A) N: D63G and B.1.617.2 haplotype, B) N: P13L and B.1.1.529 haplotype, C) S: $H69^{+}/V70^{+}$ and BA.1 haplotype, and D) S: $H69^{-}/V70^{-}$ and BA.2 haplotype. Shadow area with AS-RT-qPCR represents the standard deviation of estimation of the targeted allele of a specific variant.

6.5.2 Diagnostic performance of AS-RT-qPCR and amplicon-based sequencing in wastewaters

The accuracy of AS-RT-qPCR and amplicon-based sequencing methods was evaluated for each variant by quantifying diagnostic sensitivity and specificity across the sampling periods using the presence/absence of the variant in clinical samples as the benchmark test. In this respect, we employed Youden's index, where an index value of 1 indicates a perfect test (i.e., as sensitive, and specific as clinical genomic surveillance). Such parameterization of the method performance can help establish optimal thresholds for the inclusion/exclusion of alleles used to measure the frequency (i.e., prevalence) of SARS-CoV-2 variants in communities using wastewater-based genomic surveillance. Here, we evaluated Youden's index of targeted alleles accessed by AS-RT-qPCR, haplotype-defining mutations of each variant, and haplotype of each variant. Results showed that Youden's indices of single allele frequency by AS-RT-qPCR (N: D63G, N: P13L, S: $H69^{+}/V70^{+}$ and S: $H69^{-}/V70^{-}$) were 0.80, 0.88, 0.96 and 0.90, while those for the same targets using sequencing were 0.68, 1.00, 0.96 and 0.90, respectively (Table 6.1). Youden's indices of B.1.617.2, B.1.1.529, BA.1, and BA.2 haplotype-based detections were 0.72, 0.88, 0.88, and 0.73, respectively. Although Youden's indices across AS-RT-qPCR, single allele frequency sequencing and haplotype frequency sequencing were comparable ($p > 0.05$), the lower index of N: D63G in the single-allele sequencing method was driven by the exclusion of four data points (November

25, 27, 29, and December 10, 2021) due to insufficient sequencing depth (< 30×) that precluded mutation detection at this locus during the dominant period of Delta variant. The exclusion of data points due to insufficient sequencing depth was also observed in the other locus (Figure 6.2). Sequencing depth is a pivotal determinant of accurate SARS-CoV-2 mutation detection as low read depth reduces the likelihood of accurately identifying mutations, making it difficult to distinguish true alleles from sequencing errors (Caduff et al., 2022; Lou et al., 2022). This can result in incomplete or biased allele frequency estimates, which eventually influences Youden's index by lowering the sensitivity of the alleles. On the contrary, reduced specificity affects the Youden's index for AS-RT-qPCR which was driven by the lower proportion (~2%) of N: D63G detection in wastewater for a few days (January 26, February 5, 11 & 15, 2022) during a period when no Delta cases were detected in the Ontario clinical sequences (Figure 6.1.A). This could be due to fewer cases (so less likely to be detected and sequenced through clinical surveillance) and/or persistent infection and shedding from prior incident cases. Detection of B.1.617.2 haplotype (i.e., ORF3a: S26L, M: I82T, ORF7a: T120I, N: R203M) in this period is consistent with these possibilities (Figure 2). Similarly, the early detection of certain alleles in wastewater relative to clinical surveillance affected the specificity of B.1.1.529, BA.1, and BA.2 haplotypes, resulting in decreased Youden's indices.

Table 6-1: Youden's indices of alleles uniquely associated with B.1.617.2 (Delta), B.1.1.529 (Omicron), BA.1 (Omicron B.1.1.529.1), BA.2 (Omicron B.1.1.529.2), the haplotype of each variant, as well as N: D63G, N: P13L, S: $H69^{+}/V70^{+}$ and S: $H69^{-}/V70^{-}$ using AS-RT-qPCR in wastewaters. The light orange color cells indicate the Youden index for alleles evaluated with AS-RT-qPCR, blue represents the Youden index for the same alleles in sequencing, light blue denotes the Youden index uniquely associated with each variant evaluated by sequencing, and green represents the Youden index of the haplotype of each variant.

B.1.617.2 (Delta)				B.1.1.529 (Omicron)				Omicron BA.1				Omicron BA.2			
Allele	Sensitivity (%)	Specificity (%)	Youden's index	Allele	Sensitivity (%)	Specificity (%)	Youden's index	Allele	Sensitivity (%)	Specificity (%)	Youden's index	Allele	Sensitivity (%)	Specificity (%)	Youden's index
N: D63G	100.00	80.00	0.80	N: P13L	100.00	87.50	0.88	S: $H69^{+}/V70^{+}$	100.00	100.00	1.00	S: $H69^{-}/V70^{-}$	90.00	100.00	0.90
N: D63G	73.33	94.44	0.68	N: P13L	100.00	100.00	1.00	S: $H69^{+}/V70^{+}$	96.43	100.00	0.96	S: $H69^{-}/V70^{-}$	90.00	100.00	0.90
S: T19R	40.00	100.00	0.40	ORF1a: P3395H	88.00	100.00	0.88	ORF1a: K856R	64.29	100.00	0.64	ORF1a: S135R	85.71	100.00	0.86
S: L452R	80.00	100.00	0.80	ORF1b: I1566V	96.00	100.00	0.96	ORF1a: A1707A	92.86	100.00	0.93	ORF1a: T842I	52.38	100.00	0.52
S: P681R	80.00	100.00	0.80	S: G339D	92.00	100.00	0.92	ORF1a: SL2083I	96.43	100.00	0.96	ORF1a: G1307S	71.43	100.00	0.71
S: D950N	46.67	100.00	0.47	S: S373P	84.00	100.00	0.84	ORF1a: A2710T	85.71	100.00	0.86	ORF1a: A1352A	76.19	86.67	0.63
ORF3a: S26L	73.33	88.89	0.62	S: S375F	84.00	100.00	0.84	ORF1a: I3758V	96.43	100.00	0.96	ORF1a: L3027F	47.62	100.00	0.48
M: I82T	80.00	88.89	0.69	S: K417N	84.00	100.00	0.84	ORF1a: SLSG3673Sdel	89.29	100.00	0.89	ORF1a: V3053V	52.38	100.00	0.52
ORF7a: V82A	26.67	100.00	0.27	S: N440K	80.00	100.00	0.80	ORF1a: V4310V	96.43	100.00	0.96	ORF1a: T3090I	76.19	100.00	0.76
ORF7a: T120I	13.33	94.44	0.08	S: S477N	80.00	100.00	0.80	S: A67V	96.43	100.00	0.96	ORF1a: L3201F	61.90	100.00	0.62
N: R203M	60.00	94.44	0.54	S: E484A	80.00	100.00	0.80	S: VYY143del	96.43	100.00	0.96	ORF1A: D3311D	14.29	100.00	0.14
N: D377Y	80.00	100.00	0.80	S: Q493R	80.00	100.00	0.80	S: NL211I	71.43	100.00	0.71	ORF1a: R3394R	71.43	100.00	0.71
B.1.617.2 haplotype	100.00	72.22	0.72	S: Q498R	80.00	100.00	0.80	S: ins215EPE	67.86	100.00	0.68	ORF1a: SGF3675del	71.43	100.00	0.71
				S: N501Y	84.00	100.00	0.84	S: S371L	57.14	100.00	0.57	ORF1a: I4205I	57.14	100.00	0.57
				S: Y505H	84.00	100.00	0.84	S: G446S	60.71	100.00	0.61	ORF1b: L749L	90.48	100.00	0.90
				S: H655Y	100.00	100.00	1.00	S: G496S	53.57	100.00	0.54	ORF1b: R1315C	85.71	100.00	0.86
				S: N679K	100.00	100.00	1.00	S: T547K	96.43	87.50	0.84	ORF1b: T2163I	71.43	100.00	0.71
				S: P681H	100.00	100.00	1.00	S: N856K	100.00	87.50	0.88	ORF1b: E2196E	71.43	100.00	0.71
				S: N764K	96.00	87.50	0.84	S: L981F	96.43	87.50	0.84	S: T19I	61.90	100.00	0.62
				S: D796Y	92.00	87.50	0.80	M: D3G	82.14	100.00	0.82	S: LPPA24Sdel	57.14	100.00	0.57
				S: Q954H	100.00	87.50	0.88	BA.1 haplotype	100.00	87.50	0.88	S: V213G	76.19	100.00	0.76
				S: N969K	100.00	87.50	0.88					S: S371F	100.00	80.00	0.80
				S: D1146D	100.00	87.50	0.88					S: T376A	95.24	100.00	0.95
				ORF3a: T64T	100.00	100.00	1.00					S: D405N	95.24	100.00	0.95
				E: 9TI	100.00	100.00	1.00					S: R408S	95.24	100.00	0.95
				M: Q19E	96.00	100.00	0.96					ORF3a: T223I	66.67	100.00	0.67
				M: A63T	96.00	100.00	0.96					M: F112F	71.43	100.00	0.71
				ORF6: R20R	96.00	100.00	0.96					ORF6:D61L	71.43	100.00	0.71
				N: ERS31del	100.00	100.00	1.00					N: S413R	76.19	100.00	0.76
				N: RG203KR	92.00	100.00	0.92					BA.2 haplotype	100.00	73.33	0.73
				B.1.1.529 haplotype	100.00	87.50	0.88								

Although the Youden index threshold for a perfect diagnostic test is 1, the highest Youden index for haplotype-defining mutations of B.1.617.2, B.1.1.529, BA.1, and BA.2 variants in wastewater context were 0.80, 1.00, 0.96, and 0.95 respectively (Table 6.1). The highest Youden index for each variant signified almost an ideal target allele and the alleles, we targeted using AS-RT-qPCR or sequencing-based methods aligned perfectly with the Youden index threshold (i.e., 1). This indicated the accuracy of wastewater-based genomic surveillance using AS-RT-qPCR or sequencing-based methods against the clinical genomic surveillance for estimating SARS-CoV-2 variants prevalence in a population. However, suitable allele selection can be challenging as numerous alleles contain similar Youden index, and alleles can be influenced by sufficient sequencing depth in sequencing-based methods. Thereby, more than one allele can be separately monitored to overcome the read depth limitations of specific mutations. Youden's index of haplotype-based estimation was lower than the threshold of each variant, as monitoring haplotypes has the potential of false positive and negative detection due to the presence of certain alleles during the waning or waxing of each variant. These highlight a known issue: curating accurate haplotypes for Pango lineages is a challenging, and ongoing task in the face of rapidly emerging variants and diminishing SARS-CoV-2 clinical genomic surveillance.

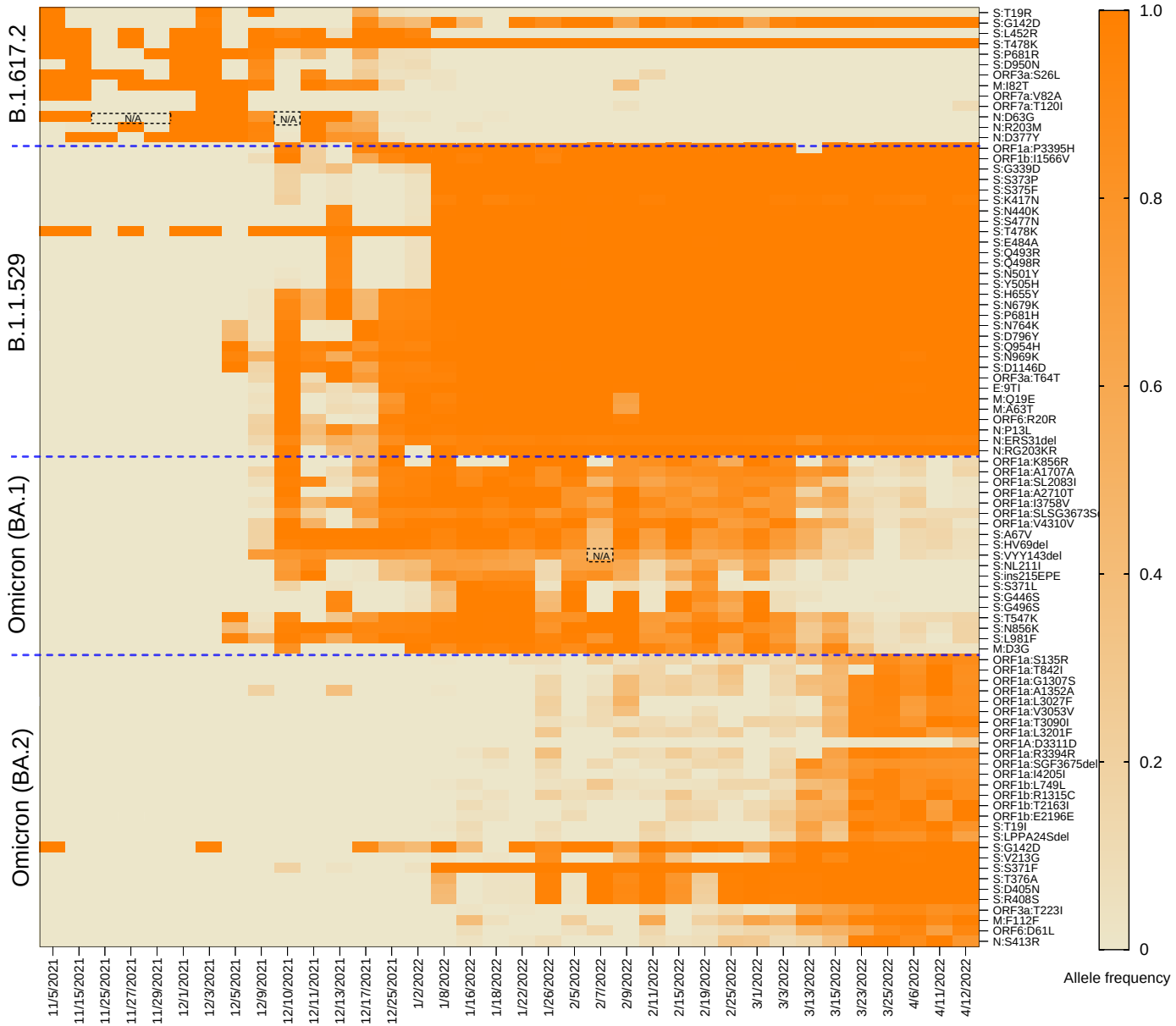
B.1.617.2
B.1.1.529
Omicron (BA.1)
Omicron (BA.2)
N/A
N/A
N/A
S:T19R
S:G142D
S:L452R
S:T478K
S:P681R
S:D950N
ORF3a:S26L
M:I82T
ORF7a:V82A
ORF7a:T120I
N:D63G
N:R203M
N:D377Y
ORF1a:P3395H
ORF1b:I1566V
S:G339D
S:S373P
S:S375F
S:K417N
S:N440K
S:S477N
S:T478K
S:E484A
S:Q493R
S:Q498R
S:N501Y
S:Y505H
S:H655Y
S:N679K
S:P681H
S:N764K
S:D796Y
S:Q954H
S:N969K
S:D1146D
ORF3a:T64T
E:9TI
M:Q19E
M:A63T
ORF6:R20R
N:P13L
N:ERS31del
N:RG203KR
ORF1a:K856R
ORF1a:A1707A
ORF1a:SL2083I
ORF1a:A2710T
ORF1a:I3758V
ORF1a:SLSG3673S
ORF1a:V4310V
S:A67V
S:HV69del
S:VYY143del
S:NI.211I
S:ins215EPE
S:S371L
S:G446S
S:G496S
S:T547K
S:N856K
S:L981F
M:D3G
ORF1a:S135R
ORF1a:T842I
ORF1a:G1307S
ORF1a:A1352A
ORF1a:L3027F
ORF1a:V3053V
ORF1a:T3090I
ORF1a:L3201F
ORF1A:D3311D
ORF1a:R3394R
ORF1a:SGF3675del
ORF1a:I4205I
ORF1b:L749L
ORF1b:R1315C
ORF1b:T2163I
ORF1b:E2196E
S:T19I
S:LPPA24Sdel
S:G142D
S:V213G
S:S371F
S:T376A
S:D405N
S:R408S
ORF3a:T223I
M:F112F
ORF6:D61L
N:S413R
11/5/2021
11/15/2021
11/25/2021
11/27/2021
11/29/2021
12/1/2021
12/3/2021
12/5/2021
12/9/2021
12/10/2021
12/11/2021
12/13/2021
12/17/2021
12/25/2021
1/2/2022
1/8/2022
1/16/2022
1/18/2022
1/22/2022
1/26/2022
2/5/2022
2/7/2022
2/9/2022
2/11/2022
2/15/2022
2/19/2022
2/25/2022
3/1/2022
3/3/2022
3/13/2022
3/15/2022
3/23/2022
3/25/2022
4/6/2022
4/11/2022
4/12/2022
1.0
0.8
0.6
0.4
0.2
0
Allele frequency

Figure 6-2: Allele frequency of variant-defining mutation of four different Pango lineages (B.1.617.2, B.1.1.529, BA.1, and BA.2) in wastewater from November 05, 2021, to April 12, 2022. The heat map represents the changes in the frequency of allele during the waning of Delta with the waxing of Omicron, as well as the waning of BA.1 with the waxing of BA.2. The missing data points of a single allele due to insufficient sequencing depth for mutation calling during the dominant period of the corresponding variant is indicated on the heatmap as N/A with black dash points.

6.5.3 Practical considerations of wastewater AS-RT-qPCR and amplicon-based sequencing implementation

AS-RT-qPCR is a rapid and inexpensive method for estimating SARS-CoV-2 variants frequency in wastewater compared to amplicon-based sequencing methods. Starting with extracted RNA, the turnaround time of AS-RT-qPCR can be as little as 2 hours including data analysis with an estimated reagent cost of ~USD$15 per sample. In contrast, amplicon-based SARS-CoV-2 genome sequencing takes about 34 hours in five consecutive steps including cDNA synthesis, amplicon generation, library preparation, sequencing, and data analysis. The reagents cost of the amplicon-based sequencing method is ~USD$100/sample (up to 24 samples with a MiSeq sequencing run). Despite the longer turnaround time and higher reagent costs, the sequencing-based method can identify the number and relative abundance of numerous mutations and variants present in wastewater without prior knowledge of the mutations. In contrast, AS-RT-qPCR requires prior knowledge of a diagnostic mutation, and this precludes its use as a method to discover emerging variants unless the new variant has undergone a reversion at a locus currently being monitored by AS-RT-qPCR (e.g., S: $H69^{+}/V70^{+}$). Significant time and expertise are needed to design and implement a new AS-RT-qPCR assay. However, the time-to-delivery of new assays has been significantly reduced

during the pandemic with a better understanding of what works well and re-using existing assays as new variants replace old ones.

The diagnostic performance evaluated using Youden's index of AS-RT-qPCR and amplicon-based sequencing methods demonstrated that all three methods can be reliably used for estimating SARS-CoV-2 variants in wastewaters, with some precautions. The Youden's index of AS-RT-qPCR was affected by reduced specificity due to low frequency detected during a period when no clinical cases were detected. On the contrary, the single-allele sequencing method was susceptible to a lower Youden's index (reduced sensitivity) when there was insufficient depth to confidently detect the mutation. Thus, monitoring more than one allele separately using both methods can help to overcome the sensitivity and specificity limitations imposed by lower frequency and read depth, and results in the increased accuracy of AS-RT-qPCR and sequencing-based methods for single allele estimation. The haplotype method, which should have high specificity, exhibited false positive detection during variant transition periods when the emerging and endemic variants shared alleles. When multiple variants share common alleles (genetic markers), it becomes more challenging to definitively assign detected alleles to a specific variant. This overlap can reduce both specificity and sensitivity, potentially resulting in an underestimation of the true positive rate or an overestimation of the false positive rate, thus complicating accurate variant identification. Thus, the accuracy of haplotype estimation requires careful haplotype curation. Accurate haplotypes can be assigned retrospectively, however prospective curation is challenging given the potential lag between the appearance of signal in wastewater and the presence of the curated genomes in clinical surveillance databases.

In general, AS-RT-qPCR offers rapid, targeted detection of specific viral variants or alleles, while sequencing provides comprehensive genomic insights, enabling the identification of novel mutations and characterization of entire viral genomes. Though both methods exhibit

similar diagnostic capabilities, their implementation faces challenges, particularly in low and lower-middle-income countries with limited resources. In rural and resource-limited settings, AS-RT-qPCR can support long-term surveillance and pandemic preparedness, where understanding the trajectory of pathogen mutations is crucial for public health response. However, AS-RT-qPCR alone may fall short when emerging variants or unknown mutations arise. In this context, sequencing facilitates the discovery of novel variants, provides transmission insights, and enables retrospective analyses of variant evolution. Integrating AS-RT-qPCR and sequencing into routine wastewater surveillance holds transformative potential for early detection and control of infectious diseases. The combined use of AS-RT-qPCR's rapid detection with sequencing's detailed genomic analysis strengthens our capacity to monitor pathogen evolution and spread, contributing to more effective public health responses and future pandemic preparedness.

6.6 Conclusion

In conclusion, our results show the similar performance of AS-RT-qPCR and sequencing-based methods to estimate SARS-CoV-2 variants frequency in wastewater using a single diagnostic allele or haplotype as a proxy of variant prevalence in a community. Later, we evaluated diagnostic performance by employing Youden's index to quantitate the accuracy of each method against the clinical genome surveillance and found that when an emerging SARS-CoV-2 variant signal is very low, the diagnostic specificity of monitoring a single-allele by AS-RT-qPCR is reduced (or clinical genomic surveillance sampling rate is insufficient). Monitoring the same allele via sequencing exhibited reduced sensitivity due to insufficient sequencing depth at the targeted locus. Youden's index derived from haplotype detections of each variant was influenced by the presence of shared alleles, highlighting the need for careful haplotype curation to accurately estimate frequencies. Therefore, the use of more than one allele using AS-RT-qPCR and amplicon-based sequencing methods is recommended during periods of low

signal and low cases to overcome the sensitivity limitations of each method. Our study recommends that any one of the methods evaluated here would provide public health stakeholders with trajectories of SARS-CoV-2 variants in their community, however, their integrations provide a comprehensive picture of the circulating variants and assist public health decision-making.

6.7 References

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

Barbé, L., Schaeffer, J., Besnard, A., Jousse, S., Wurtzer, S., Moulin, L., Le Guyader, F. S., & Desdouits, M. (2022). SARS-CoV-2 Whole-Genome Sequencing Using Oxford Nanopore Technology for Variant Monitoring in Wastewaters. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.889811

Boehm, E., Kronig, I., Neher, R. A., Eckerle, I., Vetter, P., & Kaiser, L. (2021). Novel SARS-CoV-2 variants: the pandemics within the pandemic. In *Clinical Microbiology and Infection* (Vol. 27, Issue 8, pp. 1109–1117). https://doi.org/10.1016/j.cmi.2021.05.022

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., Wittwer, C. T. (2009). The MIQE Guidelines: *M*inimum *I*nformation for Publication of *Q*uantitative Real-Time PCR *E*xperiments. *Clinical Chemistry*, (Vol 55, Issue 4, pp. 611–622). https://doi.org/10.1373/clinchem.2008.112797

Caduff, L., Dreifuss, D., Schindler, T., Devaux, A. J., Ganesanandamoorthy, P., Kull, A., Stachler, E., Fernandez-Cassi, X., Beerenwinkel, N., Kohn, T., Ort, C., & Julian, T. R.

(2022). Inferring transmission fitness advantage of SARSCoV- 2 variants of concern from wastewater samplesusing digital PCR, Switzerland, December 2020 through March 2021. *Eurosurveillance*, *27*(10). https://doi.org/10.2807/1560-7917.ES.2022.27.10.2100806

Canada - covSPECTRUM (2022). Detect and analyze variants of SARS-CoV-2. URL: https://open.cov-spectrum.org (accessed 10.12.2022)

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Cov-Lineages (2022). Pango Lineages: Latest epidemiological lineages of SARS-CoV-2. URL: https://cov-lineages.org/ (accessed 11.07.2022)

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

D'Aoust, P. M., Towhid, S. T., Mercier, É., Hegazy, N., Tian, X., Bhatnagar, K., Zhang, Z., Naughton, C. C., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021a). COVID-19 wastewater surveillance in rural communities: Comparison of lagoon and pumping station samples. *Science of the Total Environment*, 801. https://doi.org/10.1016/j.scitotenv.2021.149618

D'Aoust, P. M., Mercier, E., Montpetit, D., Jia, J. J., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Langlois, M. A., Servos, M. R., MacKenzie, M., Figeys, D., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021b). Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19

incidence and prevalence. *Water Research, 188*. https://doi.org/10.1016/j.watres.2020.116560

Davies, N. G., Jarvis, C. I., van Zandvoort, K., Clifford, S., Sun, F. Y., Funk, S., Medley, G., Jafari, Y., Meakin, S. R., Lowe, R., Quaife, M., Waterlow, N. R., Eggo, R. M., Lei, J., Koltai, M., Krauer, F., Tully, D. C., Munday, J. D., Showering, A., … Keogh, R. H. (2021). Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*, *593*(7858), 270–274. https://doi.org/10.1038/s41586-021-03426-1

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. In *Nature Biotechnology* (Vol. 38, Issue 3, pp. 276–278). Nature Research. https://doi.org/10.1038/s41587-020-0439-x

Fei Xiao, Jing Sun, Yonghao Xu, Fang Li, Xiaofang Huang, Heying Li, Jingxian Zhao, Jicheng Huang, & Jincun Zhao. (2020). Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerging Infectious Diseases*, 26(8), 1920–1922.

Fontenele, R. S., Kraberger, S., Hadfield, J., Driver, E. M., Bowes, D., Holland, L. R. A., Faleye, T. O. C., Adhikari, S., Kumar, R., Inchausti, R., Holmes, W. K., Deitrick, S., Brown, P., Duty, D., Smith, T., Bhatnagar, A., Yeager, R. A., Holm, R. H., von Reitzenstein, N. H., … Varsani, A. (2021). High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *Water Research*, *205*. https://doi.org/10.1016/j.watres.2021.117710

Furuse, Y. (2021). Genomic sequencing effort for SARS-CoV-2 by country during the pandemic. *International Journal of Infectious Diseases*, *103*, 305–307. https://doi.org/10.1016/j.ijid.2020.12.034

Garson, J. A., Badru, S., Parker, E., Tedder, R. S., & McClure, M. O. (2022). Highly sensitive and specific detection of the SARS-CoV-2 Delta variant by double-mismatch allele-specific real time reverse transcription PCR. *Journal of Clinical Virology*, *146*. https://doi.org/10.1016/j.jcv.2021.105049

GISAID (2022). GISAID - hCov19 Variants. URL: https://gisaid.org (accessed 12.06.2022)

Graber, T. E., Mercier, É., Bhatnagar, K., Fuzzen, M., D'Aoust, P. M., Hoang, H. D., Tian, X., Towhid, S. T., Plaza-Diaz, J., Eid, W., Alain, T., Butler, A., Goodridge, L., Servos, M., & Delatolla, R. (2021). Near real-time determination of B.1.1.7 in proportion to total SARS-CoV-2 viral load in wastewater using an allele-specific primer extension PCR strategy. *Water Research*, *205*. https://doi.org/10.1016/j.watres.2021.117681

Gregory, D. A., Trujillo, M., Rushford, C., Flury, A., Kannoly, S., San, K. M., Lyfoung, D. T., Wiseman, R. W., Bromert, K., Zhou, M. Y., Kesler, E., Bivens, N. J., Hoskins, J., Lin, C. H., O'Connor, D. H., Wieberg, C., Wenzel, J., Kantor, R. S., Dennehy, J. J., & Johnson, M. C. (2022). Genetic diversity and evolutionary convergence of cryptic SARS- CoV-2 lineages detected via wastewater sequencing. *PLoS Pathogens*, *18*(10). https://doi.org/10.1371/journal.ppat.1010636

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

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochemia Medica*, *26*(3), 297–307. https://doi.org/10.11613/BM.2016.034

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

Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Oude Munnink, B. B., Schapendonk, C. M. E., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F. M., Lycett, S., Medema, G., Koopmans, M. P. G., & De Graaf, M. (2021). Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the netherlands and belgium. *Emerging Infectious Diseases*, *27*(5), 1405–1415. https://doi.org/10.3201/eid2705.204410

Izquierdo-Lara, R. W., Heijnen, L., Oude Munnink, B. B., Schapendonk, C. M. E., Elsinga, G., Langeveld, J., Post, J., Prasad, D. K., Carrizosa, C., Been, F., van Beek, J., Schilperoort, R., Vriend, R., Fanoy, E., de Schepper, E. I. T., Sikkema, R. S., Molenkamp, R., Aarestrup, F. M., Medema, G., … de Graaf, M. (2023). Rise and fall of SARS-CoV-2 variants in Rotterdam: Comparison of wastewater and clinical surveillance. *Science of the Total Environment*, *873*. https://doi.org/10.1016/j.scitotenv.2023.162209

Jones, D. L., Baluja, M. Q., Graham, D. W., Corbishley, A., McDonald, J. E., Malham, S. K., Hillary, L. S., Connor, T. R., Gaze, W. H., Moura, I. B., Wilcox, M. H., & Farkas, K. (2020). Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Science of the Total Environment*, 749(July), 141364. https://doi.org/10.1016/j.scitotenv.2020.141364

Landgraff, C., Ya, L., Wang, R., Buchanan, C., Wells, M., Schonfeld, J., Bessonov, K., Ali, J., Robert, E., & Nadon, C. (2021). Metagenomic sequencing of municipal wastewater provides a near-complete SARS-CoV-2 genome sequence identified as the B.1.1.7 variant of concern from a Canadian municipality concurrent with an outbreak. *medRxiv*. https://doi.org/10.1101/2021.03.11.21253409

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lee, W. L., Imakaev, M., Armas, F., McElroy, K. A., Gu, X., Duvallet, C., Chandra, F., Chen, H., Leifels, M., Mendola, S., Floyd-O'sullivan, R., Powell, M. M., Wilson, S. T., Berge, K. L. J., Lim, C. Y. J., Wu, F., Xiao, A., Moniz, K., Ghaeli, N., … Alm, E. J. (2021). Quantitative SARS-CoV-2 Alpha Variant B.1.1.7 Tracking in Wastewater by Allele-Specific RT-qPCR. *Environmental Science and Technology Letters*, *8*(8), 675–682. https://doi.org/10.1021/acs.estlett.1c00375

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

Lou, E. G., Sapoval, N., McCall, C., Bauhs, L., Carlson-Stadler, R., Kalvapalle, P., Lai, Y., Palmer, K., Penn, R., Rich, W., Wolken, M., Brown, P., Ensor, K. B., Hopkins, L., Treangen, T. J., & Stadler, L. B. (2022). Direct comparison of RT-ddPCR and targeted amplicon sequencing for SARS-CoV-2 mutation monitoring in wastewater. *Science of the Total Environment*, *833*. https://doi.org/10.1016/j.scitotenv.2022.155059

Mishra, P., Pandey, C., Singh, U., Yadav, S., & Sharma, V. (2016). Evaluation and application of diagnostic accuracy in clinical decision-making. *International Journal of Medical*

Science and Public Health, *5*(10), 2190. https://doi.org/10.5455/ijmsph.2016.09022016401

Otto, S. P., Day, T., Arino, J., Colijn, C., Dushoff, J., Li, M., Mechai, S., Van Domselaar, G., Wu, J., Earn, D. J. D., & Ogden, N. H. (2021). The origins and potential future of SARS-CoV-2 variants of concern in the evolving COVID-19 pandemic. In *Current Biology* (Vol. 31, Issue 14, pp. R918–R929). Cell Press. https://doi.org/10.1016/j.cub.2021.06.049

Pachetti, M., Marini, B., Benedetti, F., Giudici, F., Mauro, E., Storici, P., Masciovecchio, C., Angeletti, S., Ciccozzi, M., Gallo, R. C., Zella, D., & Ippodrino, R. (2020). Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *Journal of Translational Medicine*, *18*(1). https://doi.org/10.1186/s12967-020-02344-6

Peterson, S. W., Lidder, R., Daigle, J., Wonitowy, Q., Dueck, C., Nagasawa, A., Mulvey, M. R., & Mangat, C. S. (2022). RT-qPCR detection of SARS-CoV-2 mutations S 69–70 del, S N501Y and N D3L associated with variants of concern in Canadian wastewater samples. *Science of the Total Environment*, *810*, 151283. https://doi.org/10.1016/j.scitotenv.2021.151283

Public Health Ontario (2022). SARS-CoV-2 Genomic Surveillance in Ontario. URL: https://www.publichealthontario.ca/-/media/Documents/nCoV/Archives/Genome/2022/06/SARS-CoV-2-WGS-report-2022-06-21.pdf?rev=23f767844cdf478497681f535e4d8959&sc_lang=en (accessed 20.06.2022)

Rahimi, A., Mirzazadeh, A., & Tavakolpour, S. (2021). Genetics and genomics of SARS-CoV-2: A review of the literature with the special focus on genetic diversity and SARS-CoV-2

genome detection. In *Genomics* (Vol. 113, Issue 1P2, pp. 1221–1232). Academic Press Inc. https://doi.org/10.1016/j.ygeno.2020.09.059

Ratcliff, J., Al-Beidh, F., Bibi, S., Bonsall, D., Clemens, S. A. C., Estcourt, L., Evans, A., Fish, M., Folegatti, P. M., Gordon, A. C., Jay, C., Jennings, A., Laing, E., Lambe, T., MacIntyre-Cockett, G., Menon, D., Mouncey, P. R., Nguyen, D., Pollard, A. J., … Simmonds, P. (2022). Highly Sensitive Lineage Discrimination of SARS-CoV-2 Variants through Allele-Specific Probe PCR. *Journal of Clinical Microbiology*, *60*(4). https://doi.org/10.1128/jcm.02283-21

Rios, G., Lacoux, C., Leclercq, V., Diamant, A., Lebrigand, K., Lazuka, A., Soyeux, E., Lacroix, S., Fassy, J., Couesnon, A., Thiery, R., Mari, B., Pradier, C., Waldmann, R., & Barbry, P. (2021). Monitoring SARS-CoV-2 variants alterations in Nice neighborhoods by wastewater nanopore sequencing. *The Lancet Regional Health - Europe*, *10*. https://doi.org/10.1016/j.lanepe.2021.100202

Rota, M., Antolini, L., & Valsecchi, M. G. (2015). Optimal cut-point definition in biomarkers: The case of censored failure time outcome. *BMC Medical Research Methodology*, *15*(1). https://doi.org/10.1186/s12874-015-0009-y

Simon Andrews. (2020). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. In *Soil* (Vol. 5, Issue 1, pp. 47–81).

Smyth, D. S., Trujillo, M., Gregory, D. A., Cheung, K., Gao, A., Graham, M., Guan, Y., Guldenpfennig, C., Hoxie, I., Kannoly, S., Kubota, N., Lyddon, T. D., Markman, M., Rushford, C., San, K. M., Sompanya, G., Spagnolo, F., Suarez, R., Teixeiro, E., … Dennehy, J. J. (2022). Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-282463

Suhail, Y., Afzal, J., & Kshitiz. (2021). Incorporating and addressing testing bias within estimates of epidemic dynamics for SARS-CoV-2. *BMC Medical Research Methodology*, *21*(1). https://doi.org/10.1186/s12874-020-01196-4

Tao, K., Tzou, P. L., Nouhin, J., Gupta, R. K., de Oliveira, T., Kosakovsky Pond, S. L., Fera, D., & Shafer, R. W. (2021). The biological and clinical significance of emerging SARS-CoV-2 variants. In *Nature Reviews Genetics* (Vol. 22, Issue 12, pp. 757–773). Nature Research. https://doi.org/10.1038/s41576-021-00408-x

Unal, I. (2017). Defining an optimal cut-point value in ROC analysis: An alternative approach. *Computational and Mathematical Methods in Medicine*, *2017*. https://doi.org/10.1155/2017/3762651

Vega-Magaña, N., Sánchez-Sánchez, R., Hernández-Bello, J., Venancio-Landeros, A. A., Peña-Rodríguez, M., Vega-Zepeda, R. A., Galindo-Ornelas, B., Díaz-Sánchez, M., García-Chagollán, M., Macedo-Ojeda, G., García-González, O. P., & Muñoz-Valle, J. F. (2021). RT-qPCR Assays for Rapid Detection of the N501Y, 69-70del, K417N, and E484K SARS-CoV-2 Mutations: A Screening Strategy to Identify Variants With Clinical Impact. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.672562

Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P. D., Graham, B. S., Mascola, J. R., Chang, J. Y., Yin, M. T., Sobieszczyk, M., Kyratsous, C. A., Shapiro, L., Sheng, Z., Huang, Y., & Ho, D. D. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*, *593*(7857), 130–135. https://doi.org/10.1038/s41586-021-03398-2

World Health Organization (2022). Tracking SARS-CoV-2 variants. URL: https://www.who.int/activities/tracking-SARS-CoV-2-variants (accessed 11.04.2022)

Wurtzer, S., Levert, M., Dhenain, E., Accrombessi, H., Manco, S., Fagour, N., Goulet, M., Boudaud, N., Gaillard, L., Bertrand, I., Challant, J., Masnada, S., Azimi, S., Gillon-Ritz, M., Robin, A., Mouchel, J. M., Sig, O., & Moulin, L. (2022). From Alpha to Omicron BA.2: New digital RT-PCR approach and challenges for SARS-CoV-2 VOC monitoring and normalization of variant dynamics in wastewater. *Science of the Total Environment*, *848*. https://doi.org/10.1016/j.scitotenv.2022.157740

Xiao, M., Xiao, M., Liu, X., Ji, J., Ji, J., Ji, J., Li, M., Li, M., Li, M., Li, J., Li, J., Li, J., Yang, L., Sun, W., Sun, W., Sun, W., Ren, P., Ren, P., Yang, G., … Li, J. (2020). Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Medicine*, *12*(1). https://doi.org/10.1186/s13073-020-00751-4

Yanaç, K., Adegoke, A., Wang, L., Uyaguari, M., & Yuan, Q. (2022). Detection of SARS-CoV-2 RNA throughout wastewater treatment plants and a modeling approach to understand COVID-19 infection dynamics in Winnipeg, Canada. *Science of the Total Environment*, *825*. https://doi.org/10.1016/j.scitotenv.2022.153906

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Zhang, Z. qing, Shi, B. sheng, Lu, W., Liu, D. ping, Huang, D., & Feng, Y. ling. (2020). Quantitative HBcrAg and HBcAb versus HBsAg and HBV DNA in predicting liver fibrosis levels of chronic hepatitis B patients. In *Gastroenterologia y Hepatologia* (Vol. 43, Issue 9, pp. 526–536). Ediciones Doyma, S.L. https://doi.org/10.1016/j.gastrohep.2020.03.017

7 Chapter 7: Discussion and Conclusion

This dissertation contributes to the advancement of SARS-CoV-2 genomic surveillance by systematically assessing various sampling methods, optimizing concentration and RNA extraction method for primary sludge as well as evaluating the accuracy of various variant detection methods. Specifically, this research addresses critical challenges in scaling the global application of WWGS. This dissertation demonstrates the feasibility of integrating passive sampling into WEM and highlights its potential as a future pandemic preparedness tool for WWGS. Additionally, this research optimizes a robust sludge processing method for genomic surveillance, leveraging primary sludge as an abundant and valuable resource for respiratory disease monitoring. Lastly, this research validates the accuracy of WWGS by benchmarking it against the gold standard of clinical surveillance data. Results suggest that SARS-CoV-2 RNA concentrations in wastewater solids obtained from passive samplers are comparable ($p > 0.05$) to autosampler and primary sludge, and wastewater solids content have no impact on SARS-CoV-2 RNA measurements. The study findings also suggest that recovery of genomic recovery of SARS-CoV-2 from passive samplers is influenced by sequencing read length, with longer reads enabling near-complete genome recovery. Genomic analyses reveal that SARS-CoV-2 SNVs, and lineage profiles are consistent ($p > 0.05$) across sampling methods and aligned closely with clinical surveillance data. In addition, our optimized sludge processing method consistently recovers near complete ($\geq 90\%$) SARS-CoV-2 genomes from influent wastewater or primary sludge samples. Prevalent lineage and SNV profiles are identical between wastewater and primary sludge. However, there is a higher likelihood of rare (low prevalence) and Canadian cryptic (not detected in clinical surveillance) SNVs in primary sludge compared to influent wastewater. Finally, diagnostic performance of AS-RT-qPCR, single-allele (sequencing) or haplotype frequency estimations reveals that all three methods of frequency estimation are concordant and contained sufficient information to describe the trajectory of

variant prevalence. Overall, this PhD dissertation contributes to the advance of sampling methods, sample concentration and RNA extraction for WWGS of SARS-CoV-2 by demonstrating its capability to provide real-time, scalable, cost-effective, and comprehensive data on the virus's prevalence and evolution within the populations.

7.1 Novelty and Practical Implementation

This dissertation provides novel insights into the application of passive samplers for WWGS of SARS-CoV-2 in high-flow wastewater settings. Also, this research optimizes a primary sludge concentration and RNA extraction method, achieving the first successful recovery of near-complete SARS-CoV-2 genomes from primary sludge samples. In addition, this dissertation incorporates the first comprehensive evaluation of the diagnostic performance of sequencing-based methods used to estimate variant prevalence in populations through wastewater analyses.

During the COVID-19 era, WEM and WWGS have provided crucial public health insights and contributed to the development of future pandemic preparedness plans. This is being highlighted by the detection of H5N1 and the identification of a novel, more pathogenic strain of the Mpox virus in wastewater, which has the potential for global spread. Besides, climate change has accelerated the spread of vector-borne diseases such as malaria, chikungunya, Japanese encephalitis, dengue, kala-azar, and lymphatic filariasis, as warmer temperatures and shifting ecosystems create favorable conditions for novel or non-endemic pathogens to emerge or re-emerge as public health threats. In light of this evolving threat landscape, a proactive and adaptable approach to preparedness is essential. The WWGS can therefore serve as a viable option for monitoring the spread and evolution of future pandemics at the population level. However, viral particles in wastewater are generally present at low concentrations due to sewer shed dilution but they accumulate in primary clarified sludge within WWRFs. More impurities and enzyme inhibitors may also concentrate in primary sludge, making its sequencing more

difficult and challenging. Furthermore, the implementation of WEM and WWGS to monitor multiple disease targets faces difficulties in low and lower-middle-income countries with limited resources.

This dissertation addresses these challenges by providing valuable insights that can inform targeted public health interventions. The research provides evidence of the development of robust method for genomic analysis of most abundant viral sources in wastewater matrix, the optimization of resource allocation through the use of passive samplers in high-flow wastewater settings, and the potential application of these methods in low-resource settings. The primary sludge sequencing approach enables high-resolution genomic analysis for the detection of emerging pathogens, while passive sampling based genomic analysis offers a cost-effective and scalable strategy for global WWGS implementation. Moreover, the evaluation of the accuracy of various detection and quantification methods enhances the reliability of WWGS, thereby strengthening the efficiency of pandemic preparedness and response strategies. These improvements have the potential to significantly enhance efforts to reduce transmission and control the spread of infectious diseases in future public health emergencies throughout the world.

Appendix A: Supporting Tables

Table A1: Details of the genome sequencing parameters between 150 bp paired-end run mode and 300 bp paired-end run mode across auto, COSCa-ball, and Torpedo samples.

	Sample	Input reads	Trimmed reads (fastp)	Mapped reads	Trimmed reads (iVar)	Coverage median	% Coverage > 10x	Pangolin lineage
Sequence with 150 bp paired-end run mode	Autosample_13-Apr-2023	1317002	920916	517616	516983	547	77	BA.2.73
	COSCa-ball_13-Apr-2023	1899170	1408118	1256663	1254741	5	48	Unassigned
	Torpedo_13-Apr-2023	2035898	1171360	1162411	1059008	NA	2	Unassigned
Sequence with 300 bp paired-end run mode	Autosample_13-Apr-2023	2368434	1655400	831337	830033	1674	96	BA.2.1
	COSCa-ball_13-Apr-2023	2560078	1924128	1864031	1861020	4394	98	BA.2.73
	Torpedo_13-Apr-2023	2149522	1572800	1260223	1258323	2975	94	BA.2.10

Table A2: Variations of SARS-CoV-2 genome sequencing parameters throughout the sampling period across auto, COSCa-ball and Torpedo samplers.

	Raw reads			**Trimmed reads**			**Mapped reads**			**% genome coverage at 10x**			**Depth of coverage (x)**		
	Autosampler	COSCa-ball	Torpedo	Autosampler	COSCa-ball	Torpedo	Autosampler	COSCa-ball	Torpedo	Autosampler	COSCa-ball	Torpedo	Autosampler	COSCa-ball	Torpedo
14-Mar-2023	2737972	2019212	2829512	1715204	1254482	1763600	1605708	1233091	1743253	97	92	98	3866	2636	5638
15-Mar-2023	2182310	2135164	2179422	1365746	1264832	1384638	1337778	1252057	1370598	97	98	98	3899	4146	5380
16-Mar-2023	2621116	2050638	2661770	1666418	1266544	1687424	1625661	1248594	1653060	98	97	97	4392	3793	5564
17-Mar-2023	1994402	1771856	2280744	1047346	875610	1228734	989321	848748	1139756	89	88	97	1823	1391	3005
20-Mar-2023	2679902	2363722	2748144	1634656	1442322	1626478	1616221	1426346	1619641	90	94	97	3415	4069	4721
21-Mar-2023	2787406	2856938	1912424	1731192	1780142	1162770	1721948	1775971	1160555	97	96	97	6125	5248	3817
22-Mar-2023	2136074	2805660	2969214	1372344	1763942	1868190	1368213	1760590	1865136	97	97	97	4403	5698	5682
24-Mar-2023	2467472	2967416	2496196	1602134	1939042	1597880	1599219	1923142	1574151	97	92	97	5377	3841	5459
27-Mar-2023	2768594	2669714	2661712	1783180	1723926	1672296	1772530	1687462	1667705	97	89	95	5758	3631	3535
28-Mar-2023	2254720	2274922	1961940	1478974	1505572	1259838	1469219	1502190	1258263	90	97	93	2436	3779	2738
29-Mar-2023	2285426	2581244	2430294	1272556	1447300	1309668	1267924	1443165	1307136	69	91	97	1566	4392	4370
30-Mar-2023	2774594	2368336	2902618	1528838	1320100	1542868	1506311	1312094	1523301	96	98	98	4586	4227	4754
3-Apr-23	2685394	3230442	2362980	1510334	1831462	1298438	1506620	1812242	1291560	97	93	97	5182	4222	4145
5-Apr-2023	2083702	2524360	2187870	1208560	1467994	1242504	1109990	1464371	1234852	96	98	98	2979	4317	3953
11-Apr-2023	2359888	2965506	2357568	1386682	1767694	1376810	1379176	1756937	1370965	97	96	95	4542	5096	3746
12-Apr-2023	2685544	2363384	2640664	1581614	1395866	1499080	1561837	1385516	1482610	95	96	97	4181	4253	4481
18-Apr-2023	2493310	2577204	2010772	1494104	1567386	1175114	1477254	1546608	1165161	97	96	97	4497	3873	3689
19-Apr-2023	1944180	2642898	2084386	961400	1322946	982734	940065	1312152	952516	97	97	93	2770	3898	2358
20-Apr-2023	3375880	2741266	3229510	2228920	1845434	2071160	1946210	1764430	1444854	96	99	95	5146	5312	3486
25-Apr-2023	3520804	3224192	2163654	2368728	2184686	1380518	2173216	2071441	822118	90	98	97	4091	6396	2536
26-Apr-2023	3062956	2161264	2130574	2121580	1478122	1451416	2017054	1274611	1295399	89	95	97	3696	3774	4647
27-Apr-2023	3289336	2999196	2283370	2274240	2066264	1552390	2088165	1652983	1249519	97	92	97	5724	4048	3515
2-May-2023	3190554	2476214	2572914	2216730	1716752	1686800	1999428	1534664	915782	96	98	90	6392	4539	1810
3-May-2023	2665658	2742440	1640228	1875804	1936996	1134754	1743892	1718559	1020634	99	98	96	5240	5218	3101
4-May-2023	2134172	2535284	2465458	1202204	1538512	1524038	558937	1267424	1375713	82	97	93	1168	3570	3334

Table A3: Correlation analysis was conducted between SARS-CoV-2 RNA concentrations and sequencing parameters. The results, including the p-values from Pearson correlation analysis, are presented in the table.

	Method	Number of raw reads	Number of trimmed reads	Number of mapped reads	% of genome coverage	Depth of coverage (X)
SARS-CoV-2 RNA concentrations	Autosampler	0.148	0.098	0.233	0.474	0.842
	COSCa-ball	0.573	0.754	0.407	0.058	0.187
	Torpedo	0.234	0.138	0.905	0.085	0.306

Table A4: RT-qPCR Ct values for SARS-CoV-2 N1 and N2 targets between deionized water and PBS buffer across varying ratios (1:50, 1:70, 1:100, 1:200, and 1:500). PCR technical triplicates with mean and standard deviation shown.

SARS-CoV-2 N1 target										
Dilution ratio	Deionized water					PBS				
	Ct			mean	SD	Ct			mean	SD
1:50	35.54	35.16	35.31	35.34	0.16	34.69		34.43	34.56	0.13
1:70	33.32	33.61	33.60	33.51	0.13	34.97	35.15	35.17	35.10	0.09
1:100	35.05	35.30	35.83	35.39	0.32	35.24	35.85	35.64	35.57	0.25
1:200	33.92	34.36	34.03	34.10	0.19	33.98	34.53	34.22	34.25	0.22
1:500	35.10	35.13	35.33	35.19	0.10	34.40		34.54	34.47	0.07
SARS-CoV-2 N2 target										
Dilution ratio	Deionized water					PBS				
	Ct			mean	SD	Ct			mean	SD
1:50	33.63	34.23	33.71	33.86	0.27	34.14	35.19	34.59	34.64	0.43
1:70	32.68	32.81	32.72	32.74	0.06	34.85	34.11		34.48	0.37
1:100	33.84	34.64	34.08	34.19	0.34	35.01	34.77	34.75	34.84	0.12
1:200	33.12	33.29	33.41	33.27	0.12	33.61	34.30	33.30	33.74	0.42
1:500	33.24	33.46	33.12	33.28	0.14		34.62	34.59	34.60	0.01

Table A5: Oligonucleotide sequences and AS-RT-qPCR reaction specifications of the targeted alleles associated with SARS-CoV-2 variants.

Assay	Primer/probe	Sequence	PCR reaction specifications	References
CDC N1	2019-nCoV_N1-F_Primer	GAC CCC AAA ATC AGC GAA AT	50^0C for 5 min RT, 95^0C for 20 sec, and 45 cycles of 95ºC for 3 sec, 60°C for 30 sec	(CDC, 2020)
	2019-nCoV_N1-R_Primer	TCT GGT TAC TGC CAG TTG AAT CTG		
	2019-nCoV_N1-Probe	**6-FAM**-ACC CCG CAT/**ZEN**/ TAC GTT TGG TGG ACC-**IBFQ**		
CDC N2	2019-nCoV_N2-F_Primer	TTA CAA ACA TTG GCC GCA AA		
	2019-nCoV_N2-R_Primer	GCG CGA CAT TCC GAA GAA		
	2019-nCoV_N2-Probe	**6-FAM**-ACA ATT TGC/**ZEN**/CCC CAG CGC TTC AG-**IBFQ**		
D63G (MT)	Forward primer	TCA CTC AAC ATG GCA AGA AAG G	50^0C for 5 min RT, 95^0C for 20 sec, and 44 cycles of 95ºC for 3 sec, 55°C for 45 sec	(D'Aoust et al., 2022)
	Reverse primer	GGT AGT AGC CAA TTT GGT CAT CT		
	Probe	**6-FAM**-CCT TAA ATT CCC TCG ATG ACA AGG CG-**MGB**		
N63 (WT)	Forward primer	CTC ACT CAA CAT GGC AAG AAA G		
P13L (MT)	Forward primer	CCA AAA TCA GCG AAA TGA ACC	50^0C for 5 min RT, 95^0C for 20 sec, and 45 cycles of 95ºC for 3 sec, 60°C for 30 sec	
	Reverse primer	TCT GGT TAC TGC CAG TTG AAT CTG		
	N1 probe (version 2)	**6-FAM**-CCG CAT TAC GTT TGG TGG ACC C-**MGB**		
$H69^+/V70^+$ (MT)	Forward primer	CATTCAACTCAGGACTTGTTCTTACC	50^0C for 15 min RT, 95^0C for 2 min, 40 cycles of 95^0C for 5s and 60^0C for 30s	(Peterson et al., 2022)
	Reverse primer	GGTAGGACAGGGTTATCAAACCTC		
	Probe	**FAM**-TCCATGCTATCTCTG-**MGB**		
$H69^+/V70^+$ (WT)	Forward primer	CATTCAACTCAGGACTTGTTCTTACC		
	Reverse primer	GGTAGGACAGGGTTATCAAACCTC		
	Probe	**FAM**-ATGCTATACATGTCTCTG-**MGB**		

Table A6: The details include the dates of TP, FP, TN, and FN determination in wastewater against the Ontario clinical genomic surveillance for alleles associated with each variant, haplotype of each variant, and single allele of AS-RT-qPCR.

Lineage	Alleles and haplotype	TP detection dates	Number of TP detected	FN detection dates	Number of FN detected	TN detection dates	Number of TN detected	FP detection dates	Number of FP detected
Delta (B.1.617.2)	S: T19R	05-11-2021 to 02-01-2022	6	05-11-2021 to 02-01-2022	9	03-01-2022 to 25-03-2022	18	03-01-2022 to 25-03-2022	0
	S: L452R		12		3		18		0
	S: P681R		12		3		18		0
	S: D950N		7		8		18		0
	ORF3a: S26L		11		4		16		2
	M: I82T		12		3		16		2
	ORF7a: V82A		4		11		18		0
	ORF7a: T120I		2		13		17		1
	N: D63G		11		4		17		1
	N: R203M		9		6		17		1
	N: D377Y		12		3		18		0
	B.1.617.2 haplotype		15		0		13		5
	N: D63G (AS-RT-qPCR)		15		0		14		4
Omicron (B.1.1.529)	ORF1a: P3395H	09-12-2021 to 25-03-2022	22	09-12-2021 to 25-03-2022	3	05-11-2021 to 08-12-2022	8	05-11-2021 to 08-12-2022	0
	ORF1b: I1566V		24		1		8		0
	S: G339D		23		2		8		0
	S: S373P		21		4		8		0
	S: S375F		21		4		8		0
	S: K417N		21		4		8		0
	S: N440K		20		5		8		0
	S: S477N		20		5		8		0
	S: E484A		20		5		8		0
	S: Q493R		20		5		8		0
	S: Q498R		20		5		8		0
	S: N501Y		21		4		8		0
	S: Y505H		21		4		8		0
	S: H655Y		25		0		8		0
	S: N679K		25		0		8		0
	S: P681H		25		0		8		0
	S: N764K		24		1		7		1
	S: D796Y		23		2		7		1
	S: Q954H		25		0		7		1
	S: N969K		25		0		7		1
	S: D1146D		25		0		7		1
	ORF3a: T64T		25		0		8		0
	E:9TI		25		0		8		0
	M: Q19E		24		1		8		0
	M: A63T		24		1		8		0
	ORF6:R20R		24		1		8		0
	NC		11		14		8		0
	N: P13L		25		0		8		0
	N: ERS31del		25		0		8		0
	N: RG203KR		23		2		8		0
	B.1.1.529 haplotype		25		0		7		1
	N: P13L (AS-RT-qPCR)		25		0		7		1
Omicron BA.1	ORF1a: K856R	09-12-2021 to 12-04-2022	18	09-12-2021 to 12-04-2022	10	05-11-2021 to 08-12-2022	8	05-11-2021 to 08-12-2022	0
	ORF1a: A1707A		26		2		8		0
	ORF1a:SL2083I		27		1		8		0

	ORF1a: A2710T		24		4		8		0
	ORF1a: I3758V		27		1		8		0
	ORF1a: SLSG3673Sdel		25		3		8		0
	ORF1a: V4310V		27		1		8		0
	S: A67V		27		1		8		0
	S: HV69del		27		1		8		0
	S: VYY143del		27		1		8		0
	S: NL211I		20		8		8		0
	S:ins215EPE		19		9		8		0
	S: S371L		16		12		8		0
	S: G446S		17		11		8		0
	S: G496S		15		13		8		0
	S: T547K		27		1		7		1
	S: N856K		28		0		7		1
	S: L981F		27		1		7		1
	M: D3G		23		5		8		0
	Omicron BA.1 haplotype		28		0		7		1
	S: $H69^{+}/V70^{+}$ (AS-RT-qPCR)	15-02-2022 to 12-04-2022	10	15-02-2022 to 12-04-2022	0		N/A		N/A
Omicron BA.2	ORF1a: S135R	08-01-2022 to 12-04-2022	18	08-01-2022 to 12-04-2022	3	05-11-2021 to 07-01-2022	15	05-11-2021 to 07-01-2022	0
	ORF1a: T842I		11		10		15		0
	ORF1a: G1307S		15		6		15		0
	ORF1a: A1352A		16		5		13		2
	ORF1a: L3027F		10		11		15		0
	ORF1a: V3053V		11		10		15		0
	ORF1a: T3090I		16		5		15		0
	ORF1a: L3201F		13		8		15		0
	ORF1A:D3311D		3		18		15		0
	ORF1a: R3394R		15		6		15		0
	ORF1a: SGF3675del		15		6		15		0
	ORF1a: I4205I		12		9		15		0
	ORF1b: L749L		19		2		15		0
	ORF1b: R1315C		18		3		15		0
	ORF1b: T2163I		15		6		15		0
	ORF1b: E2196E		15		6		15		0
	S: T19I		13		8		15		0
	S: LPPA24Sdel		12		9		15		0
	S: V213G		16		5		15		0
	S: S371F		21		0		12		3
	S: T376A		20		1		15		0
	S: D405N		20		1		15		0
	S: R408S		20		1		15		0
	ORF3a: T223I		14		7		15		0
	M: F112F		15		6		15		0
	ORF6:D61L		15		6		15		0
	N: S413R		16		5		15		0
	Omicron BA.2 haplotype		21		0		11		4
	S: $H69^{-}/V70^{-}$ (Sequencing_	15-02-2022 to 12-04-2022	10	15-02-2022 to 12-04-2022	1		N/A		N/A
	S: $H69^{-}/V70^{-}$ (AS-RT-qPCR)		10		1		N/A		N/A

Appendix B: Supporting Figures

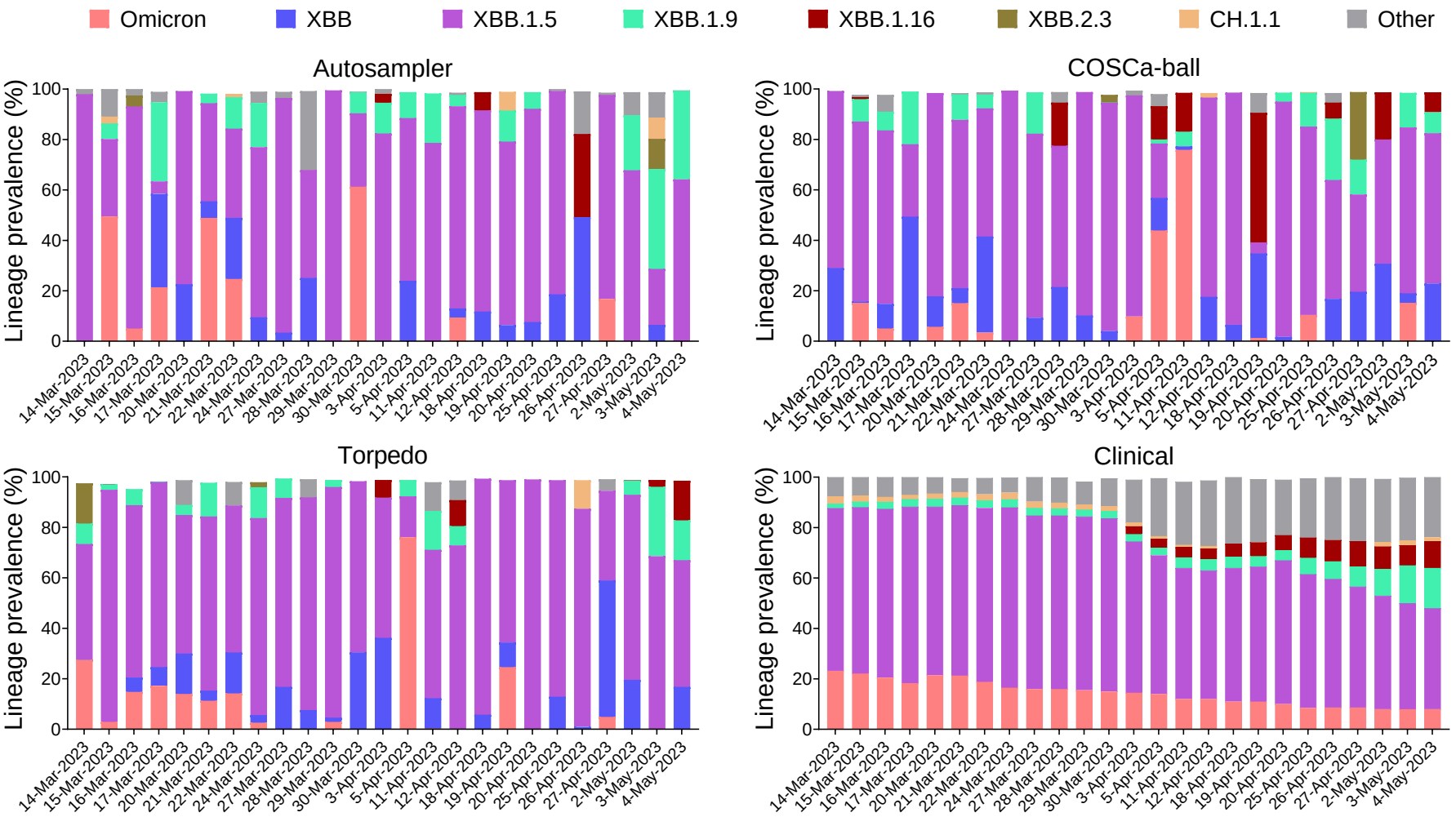

Figure A1: The SARS-CoV-2 lineages prevalence in wastewater samples collected using autosampler, COSCa-ball, and torpedo sampler as well as available clinical surveillance data from Ontario public health (PHO). Freyja analysis revealed the prevalence of two parental Omicron lineages (Omicron and CH.1.1) and five recombinant lineages of BJ.1 and BM.1.1.1 (XBB, XBB.1.5, XBB.1.9, XBB.1.16 and XBB.2.3) in wastewater samples collected using different sampling methods. Clinical cases of XBB or XBB.2.3 were not detected throughout the sampling period. The SARS-CoV-2 lineage prevalence of clinical surveillance were obtained on a weekly basis from PHO, and the daily prevalence was estimated based on the corresponding growth rate of each lineage.

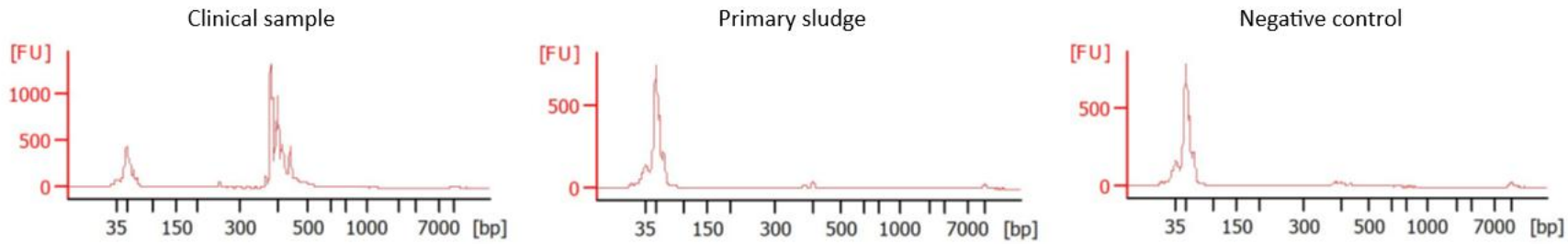

Figure A2: SARS-CoV-2 ARTIC amplification results in Bioanalyzer for three different samples: Clinical sample (Ct value similar to primary sludge sample), Primary sludge sample and negative control (nuclease-free water). The expected amplicon size is approximately 400 bp. On the graph, the x-axis represents the amplicon size (bp), while the y-axis indicates the fluorescence intensity of the amplicon.

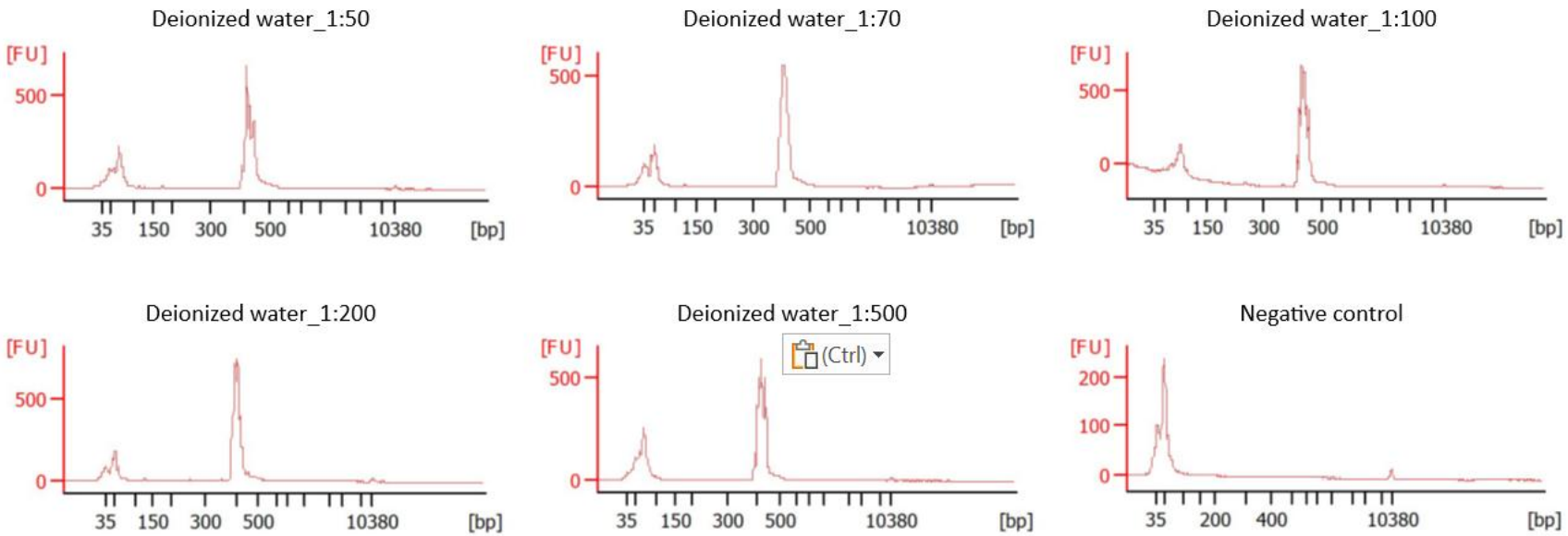

Figure A3: SARS-CoV-2 ARTIC amplification of RNA extracted from primary sludge diluted in deionized water at five different ratios (1:50, 1:70, 1:100, 1:200, and 1:500). Negative control is nuclear free water.

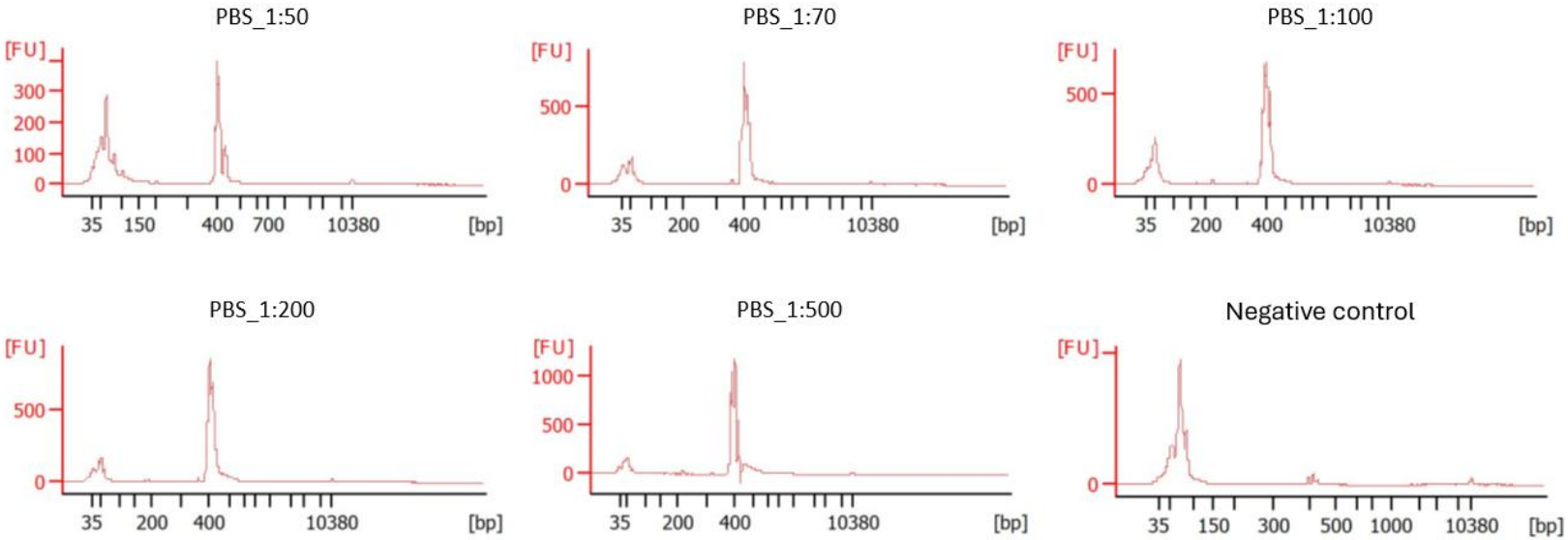

Figure A4: SARS-CoV-2 ARTIC amplification from RNA extracted from primary sludge diluted in PBS at five different ratios (1:50, 1:70, 1:100, 1:200, and 1:500). Negative control is nuclease-free water.

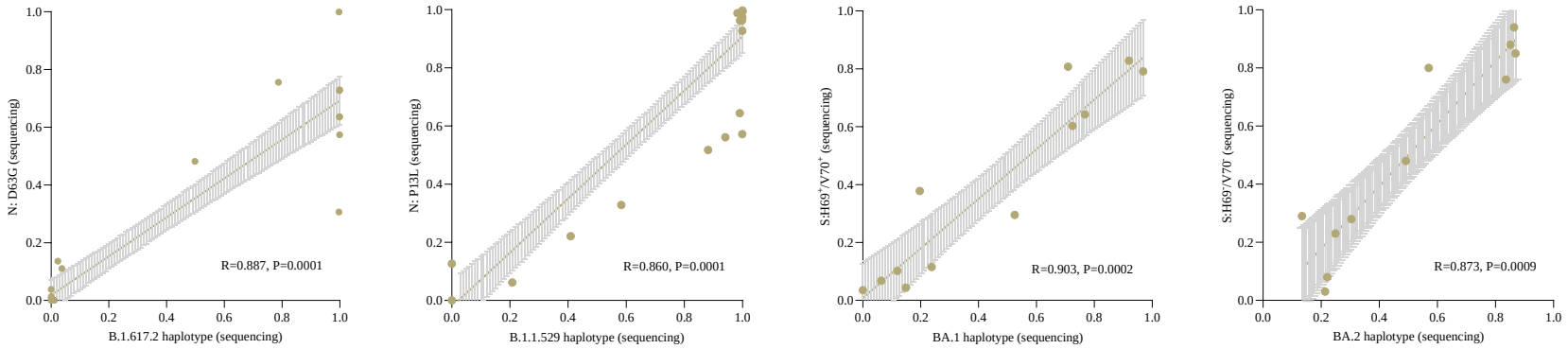

Figure A5: Spearman correlations between SARS-CoV-2 variants frequency estimations derived through single-allele and haplotype from sequenced wastewater samples. The analysis revealed significant positive correlations between N: D63G and B.1.617.2 haplotype ($r = 0.887$, $p < 0.05$), N: P13L and B.1.1.529 haplotype ($r = 0.860$, $p < 0.05$), S: $H69^{+}/V70^{+}$ and BA.1 haplotype ($r = 0.903$, $p = 0.05$), and S: $H69^{-}/V70^{-}$ and BA.2 haplotype ($r = 0.873$, $p = 0.05$). The error bar of each graph represents the uncertainty or relative variation of data points from the fitting line.