Spring 5-6-2023

DETECTION OF SARS-COV-2 MUTATIONS IN VERMILLION, SD, WASTEWATER UTILIZING PROBE-BASED RT-QPCR

Matthew J. Schmitz
University of South Dakota

DETECTION OF SARS-COV-2 MUTATIONS IN VERMILLION, SD, WASTEWATER UTILIZING PROBE-BASED RT-QPCR

By

Matthew J. Schmitz

A Thesis Submitted in Partial Fulfillment

Of the Requirements for the

University Honors Program

Division of Basic Biomedical Sciences, Sanford School of Medicine

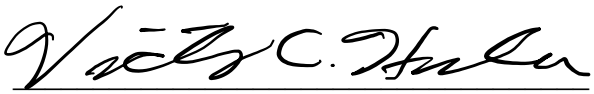
The University of South Dakota

May 2022

The members of the Honors Thesis Committee

appointed to examine the thesis of Matthew Schmitz

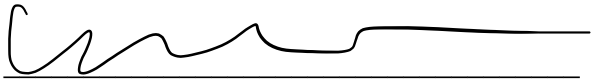
find it satisfactory and recommend that it be accepted.



Dr. Victor Huber, Ph.D.

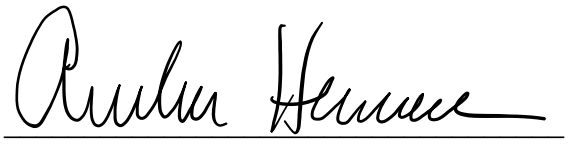
Professor of Immunology

Director of the Committee



Dr. Lisa McFadden, Ph.D.

Associate Professor of Neuroscience and Behavioral Psychology



Dr. Andrea Herrera. Ph.D.

Postdoctoral Researcher of Microbiology

ABSTRACT

Detection of SARS-CoV-2 Mutations in Vermillion, SD Wastewater
Utilizing Probe-Based RT-qPCR

By: Matthew J. Schmitz

Director: Victor Huber, Ph.D.

The COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from the subgenus *Sarbecovirus*, has presented numerous public health challenges for communities across the globe. As demonstrated by the Delta variant (B.1.617.2) and the recent Omicron variants, the virus can mutate and pose new risks for communities, such as reinfection, increased transmissibility, and variable mortality. As such, accurately monitoring the spread of specific variants has become a priority for public health agencies. However, clinical testing faces various limitations, and the increased use of at-home test kits inhibits surveillance. Wastewater-based epidemiology (WBE) can monitor infection rates of entire communities and holds the potential to track variants by detecting SARS-CoV-2 mutations in wastewater. The objective of this research endeavor was to detect SARS-CoV-2 mutations in viral RNA extracted from wastewater samples collected in Vermillion, South Dakota. Utilizing the Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02), several key mutations in the spike (S) gene were targeted utilizing probe-based RT-qPCR, allowing for the detection of specific variants. This thesis explores the evolution and application of wastewater-based epidemiology, details the protocol utilized for variant detection, assesses the efficiency and sensitivity of the Promega kit, and reports original data from samples collected in Vermillion, South Dakota, on February 1, 2022.

KEYWORDS: SARS-CoV-2, wastewater-based epidemiology (WBE), variant, spike (S) gene, reverse transcription quantitative polymerase chain reaction (RT-qPCR).

TABLE OF CONTENTS

I. ACKNOWLEDGEMENTS

I would like to start by expressing gratitude to each of my committee members, Dr. Victor Huber, Dr. Lisa McFadden, and Dr. Andrea Herrera. Dr. Huber was there at the start of my journey. More than just my Principal Investigator, he has provided necessary mentorship during my pursuit of quality research. Dr. McFadden and her research assistant Marie Severson have not only contributed critical resources to this project, but they have also endured my never-ending questions and theoretical quandaries. Dr. Andrea Herrera has been not only a kind, patient listener, but also a critical resource who has taught me many techniques involved in this project.

Thank you to the NOLOP Fellowship led by Dr. Jacob Kerby for awarding me the stipend that funded my summer research on this project. It has been my goal since my early undergraduate years to obtain this fellowship. I am grateful for the flexibility this program grants to its members, allowing us to pursue our research in an independent manner.

I would like to thank my parents for encouraging me in my endeavors. They have been my support structure, and I owe them an unpayable debt. Thank you lastly to the many professors and instructors who have contributed to my education. This project was made possible by the many people who supported and encouraged me.

II. INTRODUCTION

What Is Wastewater-based Epidemiology?

Accurately monitoring the spread of disease throughout populations allows for informed policy implementation and advanced warning for surrounding communities. As such, disease surveillance has become a priority during the COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The ability of the virus to mutate and rapidly evade established immunity has increased the demand for effective monitoring measures. The appearance of novel variants occurred rapidly across the globe during the initial stage of the pandemic and continues to occur at the time of writing this thesis. As such, low-latency detection and geographically accurate monitoring is crucial. Both of these capabilities are constituent of Wastewater-based epidemiology (WBE), which has played an increasing role throughout the pandemic in tracking the virus's spread and informing public health measures.

What is wastewater-based epidemiology? The title lends intuitive insight into its function, with "epidemiology" describing the discipline of monitoring disease spread and "wastewater" indicating the medium by which this is accomplished. Wastewater, in its broad colloquial sense, refers to the entire sewage deposit from a local community, known as a catchment. As Figure 1 illustrates, all sewage deposits from community structures focalize at the local wastewater treatment plant (WWTP) where it is further processed. The field of WBE is centered around identifying compounds, known as biomarkers, in the raw sewage that enter the wastewater treatment plant. This sewage holds an aggregate biological sample of the entire community, thus serving as an epidemiological repository for the inhabiting population.

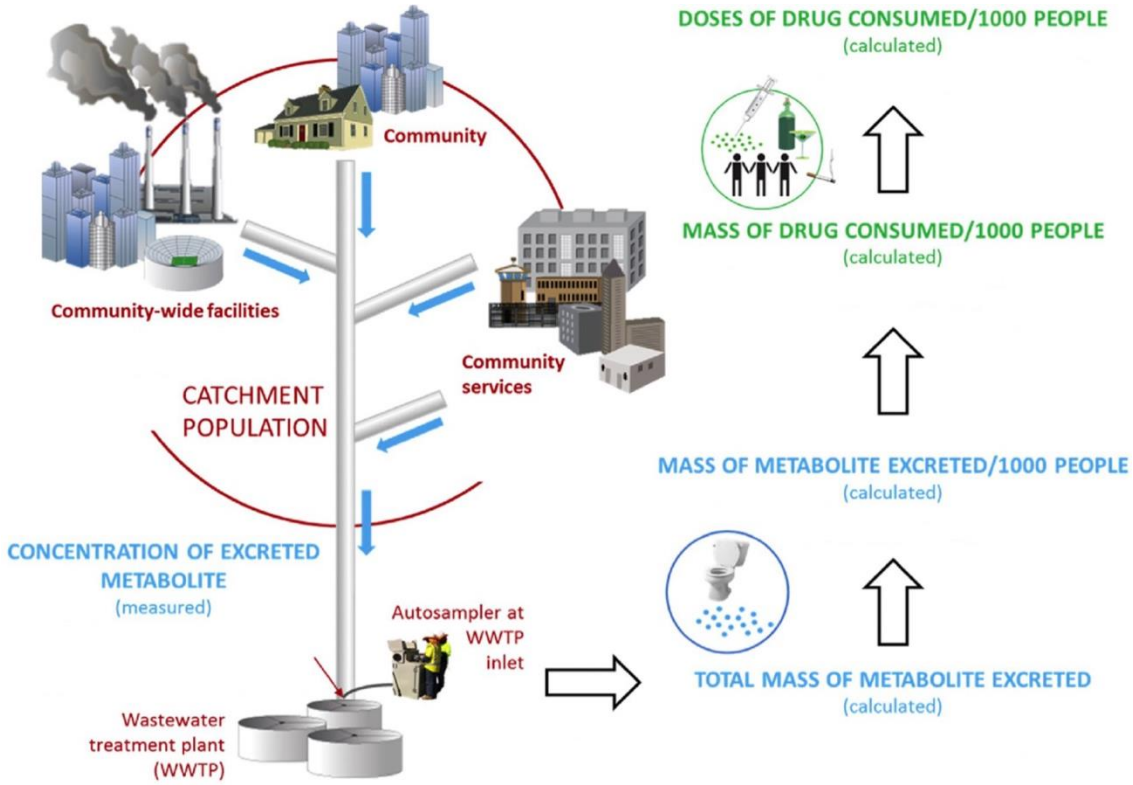


Figure 1: Process of wastewater-based epidemiology (WBE). Linear flow of wastewater from sources to treatment plant (WWTP) to laboratory testing. Specific to drug surveillance. (Choi & Tscharke, 2018, edited)

As demonstrated in Figure 1, a sampling method is implemented at the entry point to the wastewater treatment plant before the wastewater passes through various filtering and decomposition stations, which ultimately return the purified water to the environment. Next, sample treatment protocols are used to isolate a specific biomarker. Biomarkers are often metabolites of a precursor molecule of interest or genetic material of microbes. In the case of SARS-CoV-2, the relevant biomarker is the RNA genome. As depicted in Figure 1, extraction of these biomarkers from sewage samples occurs in the laboratory setting. Results are used to calculate population statistics that are both qualitative and quantitative in nature.

The concept of WBE can be traced back to English physician John Snow's epidemiological investigation of cholera outbreaks in 1854 (CDC, 2019). A specific area

of London, known as the Golden Square, experienced a significant increase in deaths caused by cholera. Death due to this disease was rather common because the groundwater was often contaminated by sewage. However, the sheer number of deaths in this location warranted an investigation. At the time, the field of epidemiology had yet to be established; thus, there were no epidemiological protocols in place to determine the source of the outbreak. Snow himself was a trained anesthesiologist, so his investigation into disease transmission was a pursuit of interest rather than training. It was his approach to the Golden Square incident that laid the groundwork for epidemiology, later earning him the title of the Founding Father of Epidemiology. His famous investigation led to the identification of a water pump contaminated by sewage runoff, which, upon disassembly, dramatically decreased the cases of cholera. The map of Golden Square, London is depicted below, with the contaminated pump centrally located.

Figure 2: Map of Golden Square, London. Dots represent cholera deaths. The pump was disassembled, thus halting the cholera outbreak. (Snow, 1936)

Snow's work established a critical concept that serves as the foundation for wastewater-based epidemiology (WBE). This idea was described more than a century

later in 2001 by environmental chemist Christian G. Daughton. Daughton stated in his article that there exists an "…intimate, immediate, and inseparable connection between humans and their environment" (Daughton, 2001, [Abstract]). In other words, humans have an unavoidable impact on their environment, a connection that can be utilized to yield valuable information concerning the health status and consumption patterns of the inhabiting population. Thus, human effects on the environment, such as wastewater drainage, hold the potential to inform public health measures. This concept serves as the cornerstone for wastewater-based epidemiology.

The first implementation of WBE in the United States occurred unintentionally, before its potential had been explicitly defined by Daughton in 2001. In the 1950s, Polio endemics were cropping up around the United States. The method of disease transmission was unknown in the early stages of the disease's spread, and asymptomatic infections were the source of substantial fear in communities across the country (CDC/NCIRD, 2015). One suspected means of transmission was the oral-fecal route, whereby wastewater contaminated the groundwater used for drinking and cooking. To test if the virus was viable in the wastewater, scientists John Paul, James Trask, and Sven Gard conducted an experiment whereby they injected monkeys with wastewater from an endemic area (Paul et al., 1940). The results of the experiment only confirmed that the polio virus was present in the wastewater drainage, rather than indicating the oral-fecal route of transmission. As such, what began as an investigation into the mode of transmission of the virus yielded results that demonstrated the capacity of wastewater to survey the presence of the virus in an infected community.

Fast-forward half a century. Daughton's 2001 article officially realized the capabilities of WBE and proposed a new potential function: quantifying illicit drug use. Daughton viewed WBE as a potential solution for the vast unknowns concerning the types and quantities of illicit drugs that were in circulation. Previously, wastewater testing had been used to determine concentrations of various pollutants, thus informing the detoxification of wastewater. There were also protocols in place for monitoring pharmaceutical drug metabolites in wastewater (Daughton, 1999). However, Daughton's 2001 article proposed utilizing wastewater surveillance as an epidemiological method to complement or possibly supplant existing epidemiological data, thus forming the field of WBE.

Daughton's proposal was put into practice several years later in Italy by Ettore Zuccato and his colleagues. Their experiment sought to compare existing cocaine-use statistics with wastewater samples (Zuccato et al., 2005). The results of their study demonstrated that established statistics significantly underestimated cocaine use in the population. This experiment served as a proof of concept that wastewater held epidemiological capabilities necessary to supplant interventional data and inform public policy. This experiment also served as a launch point for discovering many more applications of WBE. A meta-analysis written by Choi & Tscharke and colleagues detailed a wide variety of biomarkers that could be targeted to implicate the usage of compounds such as methamphetamine, heroin, tobacco, benzodiazepines, alcohol, antibiotics, and opioids (2018). The biomarkers detailed in their article were specific to non-infectious substances, whereas the COVID-19 pandemic has placed a new demand upon the capabilities of WBE: national surveillance of a viral pandemic.

Wastewater-based Epidemiology and SARS-CoV-2

The first article to propose WBE as an epidemiological tool for COVID-19 was a non-peer-reviewed scientific opinion published in the journal *Environmental Science and Technology* on February 25th of 2020 (Mao et al., 2020). The authors' rationale was based on an article published a month earlier on January 31st detailing the first case of COVID-19 in the United States (Holshue et al., 2020). This case report detailed various clinical findings, one of which stated that SARS-CoV-2 was shed in the feces of the patient. This indicated that there was a potential for detecting SARS-CoV-2 in wastewater.

The first implementation of WBE to detect the virus in the United States was described in an article by S.P. Sherchan and colleagues published on June 28th (2020). Their results demonstrated that the SARS-CoV-2 could indeed be detected in wastewater, however, their extraction methods were not efficient enough to yield consequential epidemiological data. Urgent development of standardized extraction protocols ensued, fueled by the demand for accurate epidemiological data. Recent extraction protocols, such as the Zymo Environ Water RNA Kit used for the RNA tested in this project, ensure high sensitivity. In September of 2020, the Center for Disease Control and Prevention launched the National Wastewater Surveillance System (CDC, 2023). This program reports data from WBE programs across the country and complements the CDC's COVID-19 Data Tracker, which has been utilized as a primary source for COVID-19 statistics in the US (CDC, 2020).

Utilizing WBE during the COVID-19 pandemic has yielded significant benefits, especially when used to reinforce the weaknesses of clinical testing. In most circumstances, clinical testing is used to test patients who are already feeling ill and

excludes all who elect not to get tested. Furthermore, the increased usage of at-home test kits inhibits the surveillance of clinical testing. As such, clinical testing severely underestimates infection levels in a community (Wu et al., 2020). Additionally, asymptomatic cases are known to have existed throughout the pandemic, though the relative percentage is not established (Schmitz et al, 2021; Xu et al., 2020). Regardless of the exact percentage, WBE can account for these cases given that asymptomatic individuals still shed the virus in feces (Schmitz et al., 2021).

Additional advantages of WBE over clinical testing have been demonstrated. The first concerns the advanced warning that WBE offers a community about to experience an outbreak. A meta-analysis conducted by Shimoni Shah and colleagues reported that detection of SARS-CoV-2 in the wastewater typically occurred 10 days in advance of confirmed clinical cases (2022). As such, WBE can serve as an early-warning system to inform public health measures and hospital preparation before an outbreak is underway (Sutton et al., 2020). The second advantage of WBE is its preservation of privacy. Monitoring the viral load in wastewater does not implicate individual infectivity. As such an accurate assessment of the community can be obtained without identifying individual cases. These advantages demonstrate the capabilities of WBE and serve as the foundation for the objective of this project.

Objective

The objective of this research endeavor was to identify SARS-CoV-2 variants in the wastewater collected from the Vermillion city catchment in South Dakota. This experiment differed from many studies utilizing WBE in relation to COVID-19. The

traditional approach has been to target the nucleocapsid (N) portions of the viral genome (Shah et al., 2022). The N gene is reliable for detecting the presence of the virus and estimating community infection levels, due to its low rate of mutation and stability in the wastewater (Dutta et al., 2020). However, variants of the virus, such as the Alpha, Delta, and Omicron variants, are characterized by mutations in the spike (S) gene. As such, S-gene targeting is best suited for wastewater variant monitoring. Figure 3 below illustrates the ~30 kilobase pair genome of the virus, with the S gene and N gene circled in red.

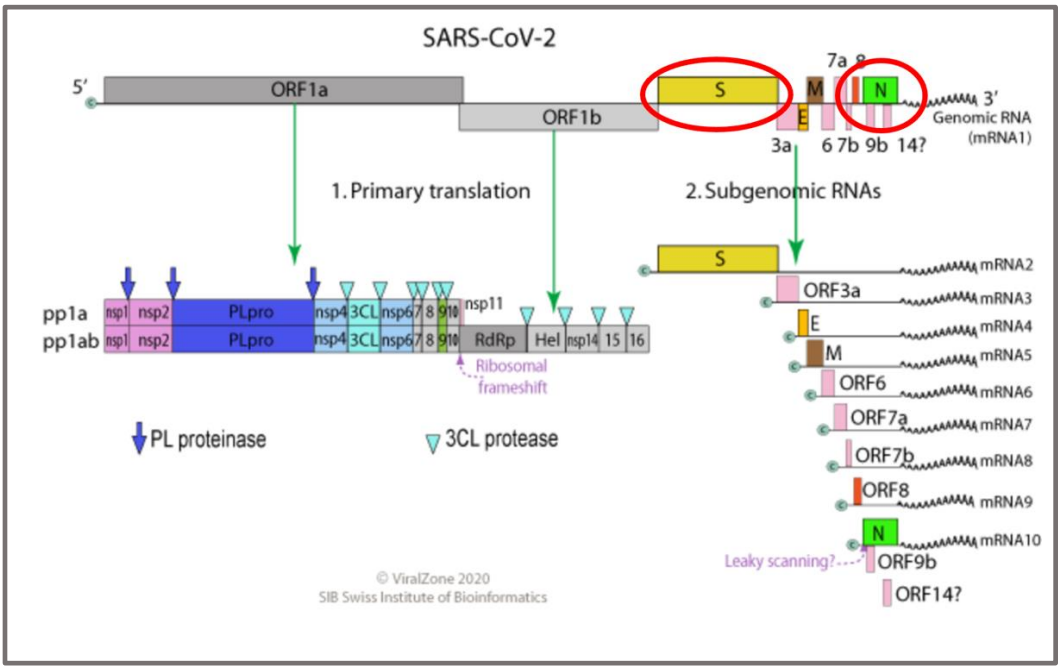


Figure 3: SARS-CoV-2 RNA genome. S-gene and N-gene circled in red (SIB, 2022)

The S gene codes for the infamous spike protein present on the surface of the viral envelope. This protein binds to angiotensin-converting enzyme II (ACE II) displayed on upper respiratory pneumocytes and is internalized, marking the beginning of infection. Mutations in the S-gene are responsible for varying levels of infectivity and mortality, and changes to this protein allow the virus to evade established immunity (Harvey et al., 2021). As such, accurate variant surveillance is imperative.

For this experiment, the S gene was targeted using the Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02), which contains eight primer-probe sets. Each set is specific to a mutation in the S gene. Samples tested with each primer-probe set provided insight into which variants were present in Vermillion at the time of sampling. This thesis will report original data concerning the sensitivity and efficiency of the Promega kit. Next, a comparison of RNA extracted from wastewater samples and sludge samples will be detailed. Lastly, a sample extracted on 2022-02-01 will be tested against all primer-probe sets, and the results will be compared to clinical data retrieved from CoVariants.org.

III. MATERIALS & METHODOLOGY

SARS-CoV-2 RNA Extraction and Storage

Weekly sampling occurred starting in June 2021 at the Vermillion Wastewater Department, located outside Vermillion, South Dakota, by the Vermillion River. Sample collection, aliquoting, and RNA extraction/storage were conducted by the McFadden laboratory, a collaborator with the Huber laboratory. Two types of sewage samples were collected: wastewater and sludge. In technical terms, “wastewater” refers to the fast-flowing component of sewage and is usually relied upon for sampling approximating a one-day average (Balboa et al., 2021). Sludge is the solid component of community sewage and is regarded as several days’ sampling average (Balboa et al., 2021). Both samples were taken from the influent of the wastewater treatment plant. Sample volumes were as follows: Wastewater: 50mL, 250 mL, and 500 mL. Sludge: 1.5 mL, 2mL, 30 mL. The raw samples did not undergo any treatment process.

RNA extraction was conducted on the 50 mL wastewater sample and 0.5 mL of the 1.5 mL sludge. The wastewater extraction assay utilized the Zymo Environ Water RNA Kit (R2042). The sludge extraction assay utilized the Zymo Quick-RNA fecal/soil Microbe MicroPrep (R2040). Extracted RNA was quantified using a Thermo Scientific Nanodrop 2000 spectrophotometer to ensure successful extraction and determine RNA concentration. Extracted RNA from WW and sludge samples were stored at -80 ˚C.

RT-qPCR Materials

Reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) was utilized to detect the presence of SARS-CoV-2 RNA in the RNA extracted from the wastewater and sludge samples. This experiment used the Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02) to target mutations in the S-gene. At the time of this work's creation, the kit is available under special order status only and is considered a prototype. The eight primer-probe sets that comprise the kit along with the variants they target are listed in Table 1.

Table 1: Primer-probe mutation targets and the variants that contain them. Sourced from Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02) instruction manual.

Part #	Volume	SARS-CoV-2 Mutation	Variants containing specific mutation
CS3174B02A	50 uL	N501Y	alpha, beta, gamma, omicron
CS3174B02B	50 uL	Del H69-/V70-	alpha, omicron (BA.1 only)
CS3174B02C	50 uL	K417N	Beta, omicron
CS3174B02D	50 uL	K417T	gamma
CS3174B02E	50 uL	Del Y144-	alpha, omicron (assay not suitable for omicron)
CS3174B02F	50 uL	E484Q	kappa
CS3174B02G	50 uL	P681R	delta, kappa
CS3174B02H	50 uL	L542R	delta, kappa, lambda

The primer-probe sets were received in separate microcentrifuge tubes containing 50 uL of each primer-probe mix. This volume contained the forward and reverse primers along with two types of hydrolysis probes specific to the amino acid target sequence in the viral genome. The two types of probes are characterized by their fluorophores. The FAM fluorophore probes detect the mutant sequence (Y in N501Y), and the HEX fluorophore probes detect the wild-type (N in N501Y). The probes used a BHQ1 quencher to mute the fluorophore. Primer and probe sequences were categorized as proprietary by Promega and are thus unavailable to be listed.

The mastermix used for this experiment, which contains the free nucleotides and polymerase enzymes, was the Promega GoTaq Enviro RT-qPCR kit (AM2011). The N1 and N2 primer-probe sets used for nucleocapsid (N-gene) detection were the IDT 2019-nCoV_N1 and N2 (10007007 & 10007008). The variant kit, mastermix, and N primer-probe sets were stored at -20 ˚C. Synthetic SARS-CoV-2 RNA positive controls were acquired from Twist Bioscience and were the following: Control 2: Original Wuhan (102024), Control 14: Alpha (103907), Control 16: Beta (104043), Control 17: Gamma (104044), Control 23: Delta (104533), and Control 50: Omicron BA.2 (105345). These positive controls were stored at -80˚C. The 96-well plates utilized for the RT-qPCR reaction were Thermo-fast (AB-1400-L).

Thawing and Plating Samples

Extracted RNA, stored at -80 ˚C, was thawed on ice, along with the positive controls. Positive controls were diluted to 1,000 copies/uL. The GoTaq Enviro Mix, GoScript Reverse Transcriptase, S-gene primer-probes, N-gene primer-probes, and

nuclease-free water were simultaneously thawed on ice. Next, the reaction mixes were prepared under a laminar flow hood disinfected with ethanol and nuclease decontaminant. A separate reaction mastermix was prepared for each primer-probe set (8 S-gene mixes and 2 N-gene mixes), each in a separate microcentrifuge tube with 20% excess volume. Vortexing and tip exchange occurred between mixture components. The volumes of each component of the mastermix per well are demonstrated in Table 2 below.

Table 2: Reaction mastermix component volumes per well. Reaction mastermix well volume is 15 uL. Add 5 uL of sample to reach total well volume of 20 uL. Sourced from Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02) instruction manual.

RT-qPCR Reaction Mastermix	Volume
GoTaq® Enviro Mastermix	10uL
GoScript™ RT	0.4uL
Primer/Probe Mix	1uL
Nuclease-free Water	3.6uL

Each mastermix was plated in the rows of a 96-well plate (i.e. N501Y mix in row A), under the laminar flow hood (15uL/well). The plate sat on a 96-well cooler to keep the plated mixtures cool. Once all the mastermixes with their probes were plated, the samples were added to the columns (i.e. Sample 1 in column 1), thus running the sample RNA against each primer-probe set (5uL/well). Positive and no-template controls were added to all 8 S-genes rows in addition to the N1 and N2 primer-probe assays. Once the plating was complete, a plate seal was applied (Applied Biosystems: 4360954). The plate was spun to remove bubbles and was subsequently placed in the machine ready to perform RT-qPCR.

Performing RT-qPCR

The machine used for this experiment was a Bio-Rad CFX Opus 96. Several initial experiments were run using an Applied Biosystems 7300 Real-Time System, but

the machine did not lend interpretable results. This issue was resolved by using the newer Bio-Rad machine. Performing RT-qPCR involves customizing the time settings for each stage of a PCR cycle and choosing the number of cycles to complete. The cycle settings used for this experiment are displayed in Table 3 below.

Table 3: Cycle Settings for RT-qPCR experiment. Sourced from Promega SARS-CoV-2 Variant Panel-8 Target kit (CS3174B02) instruction manual.

PCR cycling condition			
Step	**Temperature (˚C)**	**Time**	**Cycles**
Reverse Transcription	45	15 min	1
RT inactivation/GoTaq® activation	95	2 min	1
Denaturation	95	3 sec	40
Annealing/Ext	60	30 sec	

As demonstrated by the table, there were four steps to each RT-qPCR experiment, of which the final two are repeated numerous times until the run is complete. The first step, reverse transcription, involves the reverse transcription of the sample RNA. This adds a complementary DNA strand, known as cDNA, to the RNA strand. The second step involves a brief heating period to inactivate the reverse transcriptase enzyme before the experiment enters the cyclical and final two steps. The third and fourth steps are characteristic of a standard PCR, involving denaturation of each DNA strand and the polymerization of a new complementary strand to each separated strand. During each PCR cycle, primers and probes bind to their complementary sequences on the viral RNA and allow for duplication of both the forward and reverse strands. Once all the cycles are complete, which takes approximately two hours, the results of the experiment are ready to interpret. Up to 65 cycles were utilized for the final two stages of the run to better identify samples with irregular amplification. Samples were considered positive with a Cq < 45. The threshold was held constant at 300 RFU for all experiments.

Interpreting Results

Once the plate inside the qPCR machine completed the cycles, the data populated into CFX Maestro Software on a connected computer. For each well within the plate, an amplification curve and a quantification cycle (Cq) value were reported. The amplification curve details the fluorescence value measured at the end of each cycle. This value indicates the quantity of replicated target nucleic acid. The point at which the amplification curve crosses the threshold line is reported as the Cq. This value indirectly reports the amount of target nucleic acid present in the original sample. The lower the Cq value, the more target nucleic acid was present in the sample. Utilizing these concepts, experimental results are detailed below.

IV. RESULTS

Experiment#1: Positive Controls and RT-qPCR Performance

The first experiment completed was a verification of the intended target mutations of each S-gene primer-probe set. The protocol included with the kit detailed the same procedure and the results obtained. To independently confirm these results, the same experiment was performed. The only variation was the usage of Control 23: Delta and adding Control 50: Omicron BA.2. The protocol utilized Control 18: Kappa, mistaking it for the Delta variant. All eight of the S-gene primer-probes were run against the six Twist Synthetic RNA controls. These controls were used as positive controls in subsequent experiments. The results are displayed in Table 4 below.

Table 4: Testing Twist Synthetic SARS-CoV-2 RNA positive controls against Promega S-gene primer-probes. Displayed are the first 7 columns of a 96-well PCR plate. ncq = no-Cq. Omicron control = BA.2. NTC = no template control.

HEX: WT FAM: Mutant Ex: N501Y		1 Wuhan		2 Alpha		3 Beta		4 Gamma		5 Delta		6 Omicron		7 NTC	
		HEX	FAM	HEX	FAM	HEX	FAM	HEX	FAM	HEX	FAM	HEX	FAM	HEX	FAM
N501Y	A	32	ncq	ncq	33	ncq	32	ncq	32	32	ncq	ncq	32	ncq	ncq
DelHV 69/70	B	31	ncq	ncq	31	30	ncq	30	ncq	31	ncq	31	ncq	ncq	ncq
K417N	C	33	ncq	34	ncq	ncq	32	ncq	ncq	33	ncq	ncq	33	ncq	ncq
K417T	D	32	ncq	32	ncq	ncq	ncq	ncq	32	32	ncq	ncq	ncq	ncq	ncq
Del Y144-	E	31	ncq	ncq	31	31	ncq	31	ncq	31	ncq	36	ncq	ncq	ncq
E484Q	F	31	ncq	32	ncq	33	ncq	ncq	ncq	35	ncq	ncq	ncq	ncq	ncq
P681R	G	32	ncq	ncq	ncq	31	ncq	31	ncq	ncq	33	ncq	ncq	ncq	ncq
L452R	H	34	ncq	34	ncq	34	ncq	33	ncq	ncq	34	34	ncq	ncq	ncq

The results of this experiment confirmed the specificity of each primer-probe set for its target mutation. As shown in column 1, the original Wuhan strain registered negative for all 8 mutations. This is demonstrated by the Cq values displayed under the HEX probe, indicating that the wild-type sequence for each mutation was detected. The no-Cq (ncq) letters adjacent to these values indicate that the mutant sequence targeted by the FAM probe for each amino acid substitution was not detected. The remaining five controls each registered their targeted mutations, signified by Cq values under the FAM probe shaded with their respective colors. The HEX probe for E484Q amplified in an irregular manner for the Beta control and is denoted with red shading. A linear, jagged amplification plot indicated irregular amplification and was thus excluded. Irregular amplification was identified and excluded in future experiments. Additionally, K417N and K417T mutations occur at the same amino acid. As such, for the Beta and Omicron

controls, registering the K417N FAM probe means that the K417T wild-type sequence is absent. Vice versa is true for the Gamma control and K417T.

The next part of this experiment approximated the efficiency and sensitivity of the kit. Standard curves were performed using two of the primer-probe sets: K417T & Del Y144. Wuhan controls serially diluted in a 1:10 fashion were run in duplicates against the two primer-probe sets. The results of the standard curve assays are illustrated using CFX Maestro Software in Figures 4 & 5.

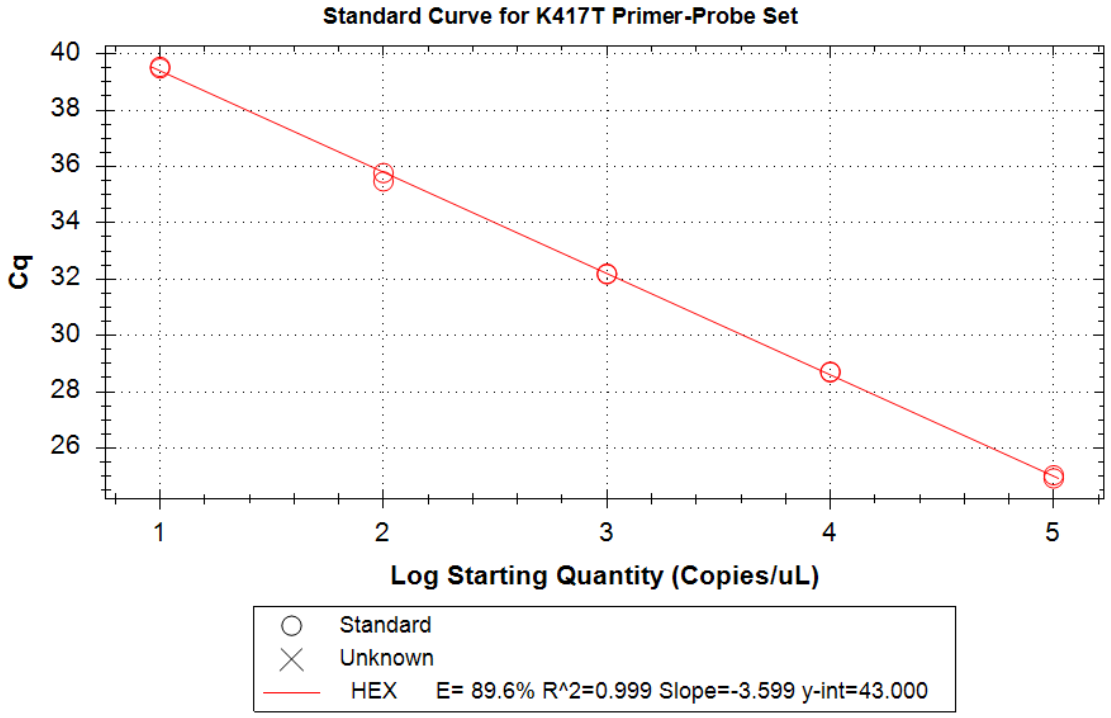


Figure 4: Standard Curve for K417T S-gene primer-probe set. X-axis depicts 10-fold dilutions with 10^5 copies/uL as the highest and 10^1 copies/uL as the lowest concentrations. Duplicates were run for each concentration on the same plate. Created using CFX Maestro software by Bio-Rad.

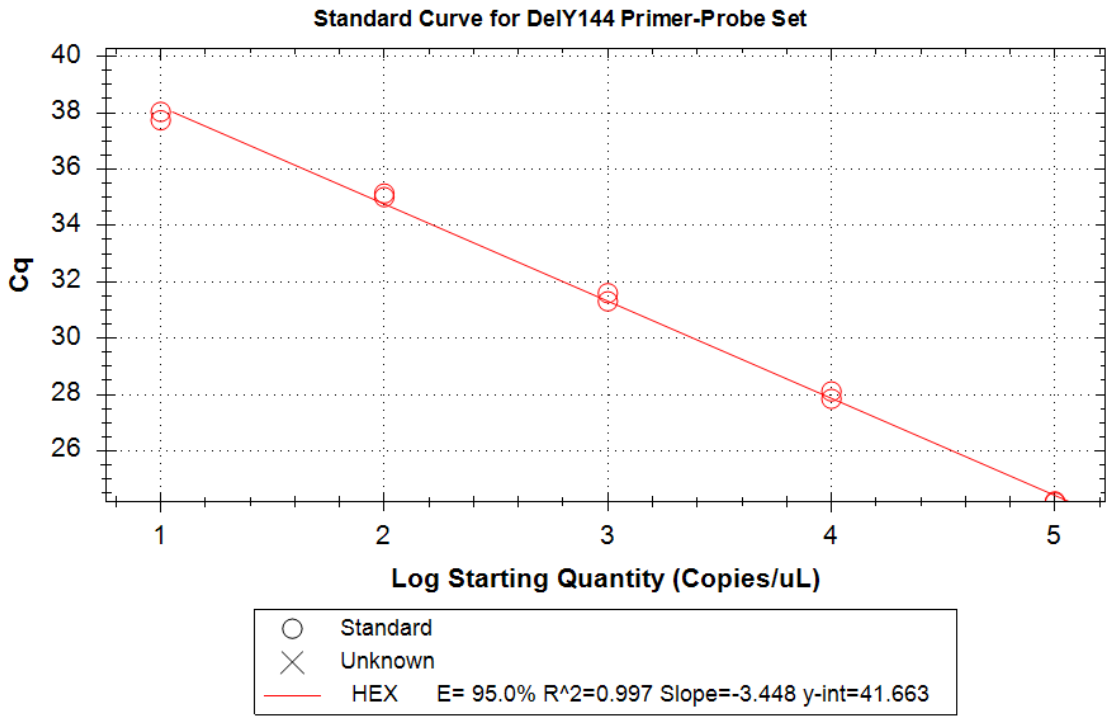


Figure 5: Standard Curve for Del Y144 S-gene primer-probe set. X-axis depicts 10-fold dilutions with 10^5 copies/uL as the highest and 10^1 copies/uL as the lowest concentrations. Duplicates were run for each concentration on the same plate. Created using CFX Maestro software by Bio-Rad.

The results from this assay indicate standard-quality efficiency and high-quality sensitivity. The efficiencies of the K417T and Del Y144 assays were 89.6% and 95% respectively. This is close to and above the target of 90% for RT-qPCR assays (Bustin et al., 2009). Concentrations of 10^0 copies/uL were also run on the plate but were excluded from the figure above due to variability. At a concentration of only 1 copy of the synthetic genome per microliter, a 5-microliter pipette draw is not guaranteed to include 5 copies of the genome. However, Cqs for these wells did register within 45 cycles. As such, the assay is likely sensitive to a single copy of synthetic RNA positive control. However, this quality of detection was only valuable insofar as it performed in wastewater and sludge samples, detailed next.

Experiment #2: Wastewater vs. Sludge

The following experiment was performed to determine the ability of the S-gene primer-probe sets to detect mutations in RNA extracted from wastewater and sludge samples. Wastewater and sludge samples extracted on 2022-2-1 were run against the K417N primer-probe set. K417N was present in all Omicron subvariants, which dominated the clinical caseload in South Dakota and the United States at that time (CoVariants). Extracted RNA was run in four concentrations, starting with the original extraction concentration and diluted in a 10-fold manner. Sludge extraction consistently yielded a greater concentration of RNA, hence the higher stock and diluted concentrations. Results from the RT-qPCR experiment are depicted in Table 5 below.

Table 5: Promega S-gene primer-probe K417N run against sludge and wastewater RNA. Samples diluted in a 10-fold manner with a no-template control included. Detected = Cq < 45. ncq = Cq > 45.

	Stock	1:10	1:100	1:1,000	NTC
[Sludge]	500.9 ng/mL	50 ng/mL	5 ng/mL	0.5 ng/mL	0 ng/mL
K417N	Detected	Detected	ncq	ncq	ncq
[Wastewater]	227.8 ng/mL	22.8 ng/mL	2.28 ng/mL	0.23 ng/mL	0 ng/mL
K417N	Detected	Detected	ncq	ncq	ncq

These results show that the S-gene primer-probe sets from the Promega kit are able to detect mutations in RNA extracted from both wastewater and sludge samples. Cqs less than 45 cycles were obtained for stock and 1:10 dilutions for both sludge and wastewater samples, indicating the presence of the K417N mutation. Furthermore, these data show that sludge and wastewater are both acceptable mediums to detect mutations. A 1:10 dilution from stock was the limit of detection (LOD) for this experiment. IDT N1 and N2-gene wastewater and sludge assays were run alongside the Promega S-gene

probes to confirm N-gene presence. The resulting Cqs were between 36 and 39 for the samples diluted 1:10 at the concentrations depicted in Table 5. Additionally, two positive controls consisting of Wuhan and Omicron synthetic RNA were run and detected the wild-type and mutant sequences respectively. Given that the Promega kit demonstrated the capability to detect one mutation, the next step was to test all 8 primer-probe sets against a sample.

Experiment #3: Detecting SARS-CoV-2 Variant

For the final experiment, RNA extracted from sludge was selected as the medium for variant detection. The rationale for this decision was two-fold. First, sludge is considered a several-day average community sample, whereas wastewater is considered a single-day sample (Balboa et al., 2021). As such, sludge is less prone to fluctuation based on day-to-day community activity. Secondly, the amplification curves resulting from sludge samples exhibited cleaner exponential amplification in Experiment #2, indicating greater RNA integrity.

This experiment tested the same sludge sample used in the previous experiment extracted on 2022-2-1 with a dilution profile of 1:20 at a concentration of 25 ng/mL. This dilution was performed due to a limited supply of sample RNA. The sample was run against all eight Promega S-gene primer-probe sets. The results of this experiment are detailed in Table 6. Presented in a columnar fashion, beginning with Omicron (BA.1) and ending with delta, are the mutation profiles of the primary variants of concern. These variants are color-coded in coordination with the mutations that are present in each.

Shown in the second column are the experimental results from running the sludge RNA against all eight primer-probe sets.

Table 6: Testing 2022-2-1 Sludge 1:20, 25 ng/mL RNA against all eight Promega S-gene primer-probe sets. Detected = Cq < 45. ncq = No Cq.

Variant Defining Mutation	Sludge RNA Sample	Variant Mutation Profile					
		Omicron (BA.1)	Omicron (BA.2)	Alpha	Beta	Gamma	Delta
N501Y	Detected	Present	Present	Present	Present	Present	Absent
DelHV69/70	Detected	Present	Absent	Present	Absent	Absent	Absent
K417N	Detected	Present	Present	Absent	Present	Absent	Absent
K417T	ncq	Absent	Absent	Absent	Absent	Present	Absent
DelY144	Not Suitable	Present	Absent	Present	Absent	Absent	Absent
E484Q	ncq	Absent	Absent	Absent	Absent	Absent	Present
P681R	ncq	Absent	Absent	Absent	Absent	Absent	Present
L452R	ncq	Absent	Absent	Absent	Absent	Absent	Present

As demonstrated by Table 6 above, the mutation profile of the sludge RNA sample aligns most closely to that of Omicron BA.1. The Del Y144 primer-probe set is not suitable to detect the tyrosine sequence deletion in Omicron variants, as stated by the Promega kit protocol. This is likely a result of the design of the hydrolysis probe. However, for the purposes of this experiment, the remaining primer-probe sets were able to identify BA.1 determinant mutations in the sludge RNA. According to clinical sequencing data from CoVariants.org, this was the dominant variant present in South Dakota from 2022-2-1 to 2022-3-28. As such, the sludge mutation profile is consistent with sequenced clinical samples. Discussed next are the implications and applications of the experimental protocols and data achieved in this thesis.

V. DISCUSSION

As stated in the protocol included with the Promega kit, the purpose of the S-gene assay is to lend detection of mutations in extracted RNA samples rather than quantification metrics. As such, results are intended only to be interpreted as mutations being “present” or “not present” in a tested sample. The kit demonstrated excellent capability in detecting target mutations in the positive Twist controls in Experiment #1. However, during this initial testing and in subsequent experiments with extracted samples, irregular amplification was observed for select probes run against select variants, specifically: E484Q HEX – Beta, E484Q HEX – Beta, K417N HEX – Omicron, and K417T HEX – Omicron. Occurrences of irregular amplification were consistent and would usually occur with aged positive controls and extracted samples, pointing to RNA integrity as a factor for irregular amplification. Irregular amplification was identified and excluded from analysis.

Given the qualitative nature of the Promega kit, the standard curves reported for the Del Y144 and K417T primer-probe sets were provided to approximate the efficiency and sensitivity of the kit. For assays attempting to quantify nucleic acid in samples, these data are essential. However, as discussed previously, this capability is outside the parameters of the Promega kit. Regardless, the Promega kit demonstrated standard-quality efficiency and high-quality sensitivity.

In Experiment #2, K417N was detected in concentrations as low as 50 ng/mL for sludge and 22.8 ng/mL for wastewater. A sludge stock dilution of 1:20 at 25 ng/mL was used for the final experiment. This dilution was utilized due to limited resources of historically extracted RNA and due to this dilution’s use as the standard for N-gene

quantification assays conducted by the McFadden Laboratory. The results of the experiment were not affected by the two-fold dilution.

The most impactful results were obtained in Experiment #3, where the mutation profile of the sample mirrored that of South Dakota clinical data sourced from CoVariants.org. This indicates that SARS-CoV-2 variant detection is possible using WBE as a detection medium. Furthermore, these results confirm that S-gene nucleic acid can hold its integrity during the extraction process to then be detected via RT-qPCR. Unfortunately, the Del Y144 primer-probe set was not capable of detecting Omicron nucleic acid. Due to the proprietary nature of the Promega primers and probes, sequences could not be analyzed or customized. However, despite the Del Y144 primer-probe's lack of Omicron detection, the detection of N501Y, Del HV 69-70, and K417N mutations was sufficient to confirm the presence of Omicron BA.1 in the wastewater.

The successful profiling of BA.1 in wastewater points to a future application of S-gene monitoring. As SARS-CoV-2 mutations accumulate and create more variants, S-gene wastewater monitoring holds the potential to serve as an early warning system for the spread of variants of concern. So far, this concept has been applied by monitoring N-gene levels quantitatively in wastewater (Shah, 2022). Utilizing S-gene surveillance, viral N-gene levels would not have to increase for the new variant to be classified as present in the community. However, given the nature of RT-qPCR, primers and probes targeting unique mutations need to be created before detection could occur in the wastewater. Developing these tools takes time and gives the new variant a period of anonymity. However, once developed, primer-probe sets could be distributed to areas susceptible to outbreak, allowing early detection of variants.

VI. CONCLUSION

Monitoring the spread of SARS-CoV-2 variants continues to be a public health priority for communities across the globe. The ability of the virus to mutate and pose new threats, such as reinfection and varied morality, calls for capable surveillance systems to inform clinical preparation and response. As detailed in this thesis, the ability of wastewater-based epidemiology to track infection levels and consumption patterns of entire communities has served an increasing role throughout history. The COVID-19 pandemic served as a catalyst to implement WBE on the front lines of disease surveillance. The objective of this research endeavor was to determine if specific S-gene mutations could be detected in wastewater utilizing the Promega kit. As demonstrated in three experiments, the Promega kit targeted the correct mutations with each of the eight primer-probe sets, exhibited quality efficiency and sensitivity, and correctly profiled Omicron BA.1 in a 2022-2-1 sludge RNA sample. These results fulfill the objective of this thesis.

VII. REFERENCES

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

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

CDC. (2020, March 28). *COVID Data Tracker*. Centers for Disease Control and Prevention. https://covid.cdc.gov/covid-data-tracker/#wastewater-surveillance

CDC. (2023). *National Wastewater Surveillance System*. Centers for Disease Control and Prevention. https://www.cdc.gov/nwss/wastewater-surveillance/index.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fhealthywater%2Fsurveillance%2Fwastewater-surveillance%2Fwastewater-surveillance.html

CDC. (2019). *Principles of Epidemiology*. Centers for Disease Control and Prevention. https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section2.html

CDC/NCIRD. (2015). *Immunology and Vaccine-Preventable Diseases – Pink Book - Polio*. https://www.cdc.gov/vaccines/pubs/pinkbook/downloads/polio.pdf

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

CoVariants: Per Country. (n.d.). Covariants.org. https://covariants.org/per-country?region=United+States

Daughton, C. G. (2001). Illicit Drugs in Municipal Sewage. *ACS Symposium Series*, 348–364. https://doi.org/10.1021/bk-2001-0791.ch020

Daughton, C. G., & Ternes, T. A. (1999). Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environmental Health Perspectives, 107*(suppl 6), 907–938. https://doi.org/10.1289/ehp.99107s6907

Dutta, N. K., Mazumdar, K., & Gordy, J. T. (2020). The Nucleocapsid Protein of SARS–CoV-2: a Target for Vaccine Development. *Journal of Virology*, *94*(13). https://doi.org/10.1128/jvi.00647-20

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

Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., Diaz, G., Cohn, A., Fox, L., Patel, A., Gerber, S. I., Kim, L., Tong, S., Lu, X., Lindstrom, S., & Pallansch, M. A. (2020). First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine*, *382*(10), 929–936. https://doi.org/10.1056/nejmoa2001191

Mao, K., Zhang, H., & Yang, Z. (2020). Can a Paper-Based Device Trace COVID-19 Sources with Wastewater-Based Epidemiology? *Environmental Science & Technology*, *54*(7), 3733–3735. https://doi.org/10.1021/acs.est.0c01174

Paul, J. R., Trask, J. D., & Gard, S. (1940). II. POLIOMYELITIC VIRUS IN URBAN SEWAGE. *Journal of Experimental Medicine*, *71*(6), 765–777. https://doi.org/10.1084/jem.71.6.765

Schmitz, B. W., Innes, G. K., Prasek, S. M., Betancourt, W. Q., Stark, E. R., Foster, A. R., Abraham, A. G., Gerba, C. P., & Pepper, I. L. (2021). Enumerating asymptomatic COVID-19 cases and estimating SARS-CoV-2 fecal shedding rates via wastewater-based epidemiology. *Science of the Total Environment*, *801*, 149794. https://doi.org/10.1016/j.scitotenv.2021.149794

Shah, S., Gwee, S. X. W., Ng, J. Q. X., Lau, N., Koh, J., & Pang, J. (2022). Wastewater surveillance to infer COVID-19 transmission: A systematic review. *Science of the Total Environment*, *804*, 150060. https://doi.org/10.1016/j.scitotenv.2021.150060

Shang, W., Kang, L., Cao, G., Wang, Y., Gao, P., Liu, J., & Liu, M. (2022). Percentage of Asymptomatic Infections among SARS-CoV-2 Omicron Variant-Positive Individuals: A Systematic Review and Meta-Analysis. *Vaccines*, *10*(7), 1049. https://doi.org/10.3390/vaccines10071049

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

(SIB) Swiss Institute of Bioinformatics. (2022). *Expasy.org.* https://viralzone.expasy.org/resources/nCoV_genome_bis.png

Snow, John. Snow on cholera: being a reprint of two papers. New York: The Commonwealth Fund; London: H. Milford, Oxford University Press, 1936. https://nrs.lib.harvard.edu/urn-3:hms.count:1168833?n=109

Wu, S. L., Mertens, A. N., Crider, Y. S., Nguyen, A., Pokpongkiat, N. N., Djajadi, S., Seth, A., Hsiang, M. S., Colford, J. M., Reingold, A., Arnold, B. F., Hubbard, A., & Benjamin-Chung, J. (2020). Substantial underestimation of SARS-CoV-2 infection in the United States. *Nature Communications*, *11*(1), 4507. https://doi.org/10.1038/s41467-020-18272-4

Xu, T., Huang, R., Zhu, L., Wang, J., Cheng, J., Zhang, B., Zhao, H., Chen, K., Shao, H., Zhu, C., Wu, C., & Liu, L. (2020). Epidemiological and clinical features of asymptomatic patients with SARS-CoV-2 infection. *Journal of Medical Virology*, *92*(10), 1884–1889. https://doi.org/10.1002/jmv.25944

Zuccato, E., Chiabrando, C., Castiglioni, S., Calamari, D., Bagnati, R., Schiarea, S., Fanelli, R. (2005, August 5). Cocaine in surface waters: A new evidence-based tool to monitor community drug abuse. *Environmental Health,* 4(14). doi: https://doi.org/10.1186/1476-069X-4-14